Study of Mask and Wafer Co-design That Utilizes a New Extreme SIMD Approach to Computing in Memory Manufacturing --Full-Chip Curvilinear ILT in a Day

Linyong (Leo) Pang[1], Ezequiel Vidal Russell[2], Bill Baggenstoss[2], Michael Lee[2], Jennefir Digaum[2], Ming-Chuan Yang[2], P. Jeffrey Ungar[1], Ali Bouaricha[1], Kechang Wang[1], Bo Su[1], Ryan Pearman[1],Aki Fujimura[1]
[1] D2S, Inc. (US)
[2] Micron Technology, Inc. (US)
[1]4040 Moorpark Avenue, Suite 250, San Jose, CA 95117 USA
[2]8000 South Federal Way, Post Office Box 6, Boise, Idaho 83707-0006

ABSTRACT

In advanced semiconductor memory manufacturing, mask and lithography are critical for patterning. In this paper we jointly study the benefits of a mask and wafer co-design that utilizes a new extreme single instruction multiple data (SIMD) approach to computing. Wafer results will be shown demonstrating the benefits of the approach. Unlike traditional EDA software that runs on customers' computer farm, this new approach leverages and maximizes GPU acceleration. In this study, software speed and quality, mask writing strategy, wafer pattern fidelity and process window are examined and analyzed.

Inverse lithography technology (ILT) has been seen as a promising solution to many of the challenges of advanced-node lithography, whether optical or EUV. However, the runtimes associated with this computational technique have limited its practical application. Until now, it has been used for critical "hotspots" on chips, but has not been used for entire chips. The solution to the runtime problem for ILT has been particularly vexing, as the traditional approach to runtime improvement – partitioning and stitching – has failed to produce satisfactory results, either in terms of runtime or in terms of quality. D2S has adopted an entirely new, stitchless approach, creating a holistically conceived, purpose-built system for ILT, This system includes a unique GPU-accelerated approach that emulates a single, giant GPU/CPU pair that can compute an entire full-chip ILT solution at once. This novel approach, systematically designed for ILT and GPU acceleration, makes full-chip ILT a practical reality in production for the first time.

For the most advanced DRAM manufacturing using 193nm immersion lithography, every aspect of design, mask, and lithography, including quality of the process, accuracy, and turn-around-time, need to be optimized. Any new technique that significantly improve one or more elements of such complete process are welcome. Recently a number of new technologies, such as multi-beam mask writer, GPU accelerated computing for mask and wafer, are emerged and are reshaping the mask and lithography. This new stitchless full chip curvilinear ILT is applied to memory chip making. We will show mask making and wafer print results, including pattern fidelity and process window, to show the actual benefits of such technologies for semiconductor manufacturing.

Keywords: Photomask, GPU, SIMD, ILT, Curvilinear ILT, Mask wafer co-optimization

1. INTRODUCTION

1.1 Curvilinear ILT Started a Decade Ago

In the last two decade for semiconductor manufacturers moving to advanced nodes, from 90nm, all the way to 5nm now – the greatest challenge has been always lithography. This is because lithography is fundamentally constrained by basic principles of optical physics. 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 mathematical approach. Such approach has been explored for many years, starting with the pioneering work of B. E. A. Saleh and his students in the 1980s [1-8]. Then in 2005 and 2006, Luminescent

Technologies (acquired by Synopsys and KLA) introduced the first industry's commercial product and the author coined the such technology as Inverse Lithography Technology (ILT)[9-17]. ILT is a rigorous computational approach to determine the mask shapes that will produce the desired on-wafer results. Given a target wafer shape and models of the lithographic optics, an inverse calculation is made to arrive at the mask pattern that will supply the desired wafer result and the best process window. Since lithography optics is a band-limited system, the ILT solution tends to be curvilinear[9] (Figure 1).

Figure 1: the original ILT mask patterns shown in the Luminescent ILT paper are curvilinear[9]

1.2 Curvilinear ILT Produces the Best Process Window

Since the late 1990s, the semiconductor industry has faced technical challenges posed by shrinking wafer geometries and the physical limitations of optical lithography to faithfully reproduce those geometries. ILT has shown great promise as a means of meeting these challenges. Numerous studies and wafer results have shown that ILT – in particular, unconstrained curvilinear ILT – can produce the best results in term of wafer-pattern fidelity and process window[18]. In this study [18] (Figure 2), authors looked the contact arrays with different pitches, for each pitch 5 variations of ILT mask patterns were generated, mask with such ILT patterns was made, wafers was printed at different process conditions, and wafer images and CDs were captured and measured. It shows that the un-constrained curvilinear ILT mask patterns produces the largest process window for all pitches.

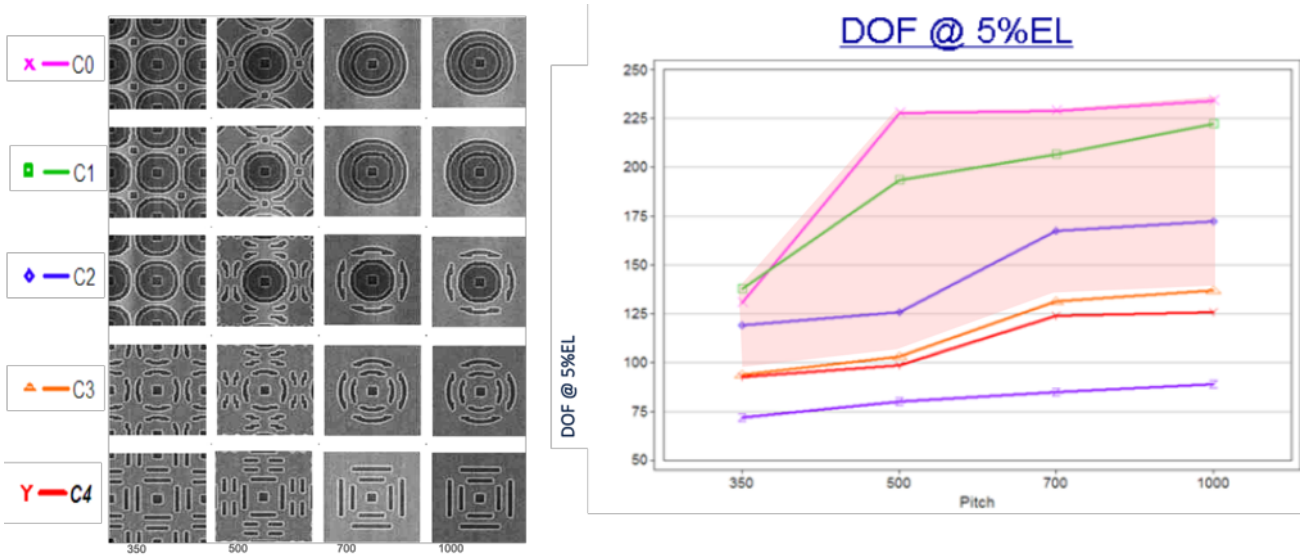


Figure 2: Study of different complexities of ILT mask patterns and their corresponding process window shows un-constrained curvilinear ILT mask patterns produce the largest process window [18]

Moving forward, ILT will be required by more and more masks, whether 193i or EUV. Optical lithography will rely more and more heavily on ILT for further progression in the roadmap to handle smaller nodes, more layers in the smaller nodes, and more aggressive design rules. With each new, smaller geometry, more areas of masks become "critical" and need ILT to ensure resolution and preserve process windows. Specific EUV effects (the non-normal, 6-degree incidence of the optical axis for the reflective optical system, as well as mask 3D effects such as mask shadowing), combined with tight

lithography error budgets require curvilinear corrections for EUV, making curvilinear ILT the best solution for EUV masks.

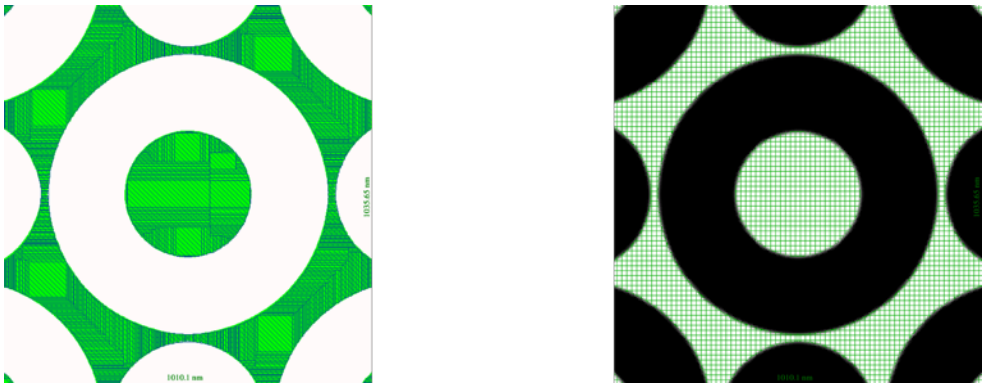
According to eBeam Initiative Survey in 2018[19], ILT has been used in critical layers in the leading technology nodes. However, it is mainly used in hot-spot fixer mode, and it is not used for all critical layers.

Two major obstacles have kept ILT from being widely applied. One of these barriers – the ability to write curvilinear mask patterns – has been removed recently by introduction of multi-beam mask writers, which can write any shape without time penalty. The other major barrier – ILT run time – was still left to be overcome.

1.3 Multi-Beam Mask Writer Enables Curvilinear ILT

The main stream mask writer, called Variable Shaped Beam (VSB) mask writer, for the leading semiconductor manufacturing was invented to write Manhattan patterns. It writes the entire mask with a single beam that can produce a rectangle shot with variable dimensions. The total write time is proportional to number of shots. This is an advantage to write Manhattan patterns, but to write a curvilinear pattern, it has to break the curvy patterns into many small shots, ending up write time too long and not practical (Figure 3a).

The mask industry recognized such challenge, and it became the major motivation to develop the new multi-beam mask writer. Multi-beam mask writer, instead of having a single variable shaped beam, has an array of 256K beams that write in a single shot, and each individual beam can be controlled to turn on and off or at partial (grey scale). Since multi-beam mask writer writes in pixel domain, write time is not affected by the patterns it write, it can write a mask with any shaped mask patterns in a constant write time, around 10 to 15 hours, including curvilinear ILT mask patterns (Figure 3b).



(a) (b)

Figure 3: (a). Conventional VSB mask writer: generates too many shots, takes too long to write; (b). Multi-beam mask writer: designed for curvilinear ILT, writes any shape in constant time[20]

1.4 Curvilinear Mask Shapes are Much More Resilient to Manufacturing Variation than Manhattan Shape

Curvilinear ILT does not only produces the best process window, but curvilinear mask shapes are also much more resilient to manufacturing variation than Manhattan shape[20]. According to Pearman's study on EUV contact using Manhattan and curvilinear shape on both VSB and multi-beam mask writer, multi-beam mask writer can reduce mask Process Variation (PV) band by 30~50% from VSB mask writer. Then switching from Manhattan pattern to curvilinear mask pattern will reduce another 20~50% mask PV bands. So, if we change from writing Manhattan mask pattern on VSB mask writer to writing curvilinear mask pattern on multi-beam mask writer, the mask variation (PV bands) can reduce 75%, very significant.

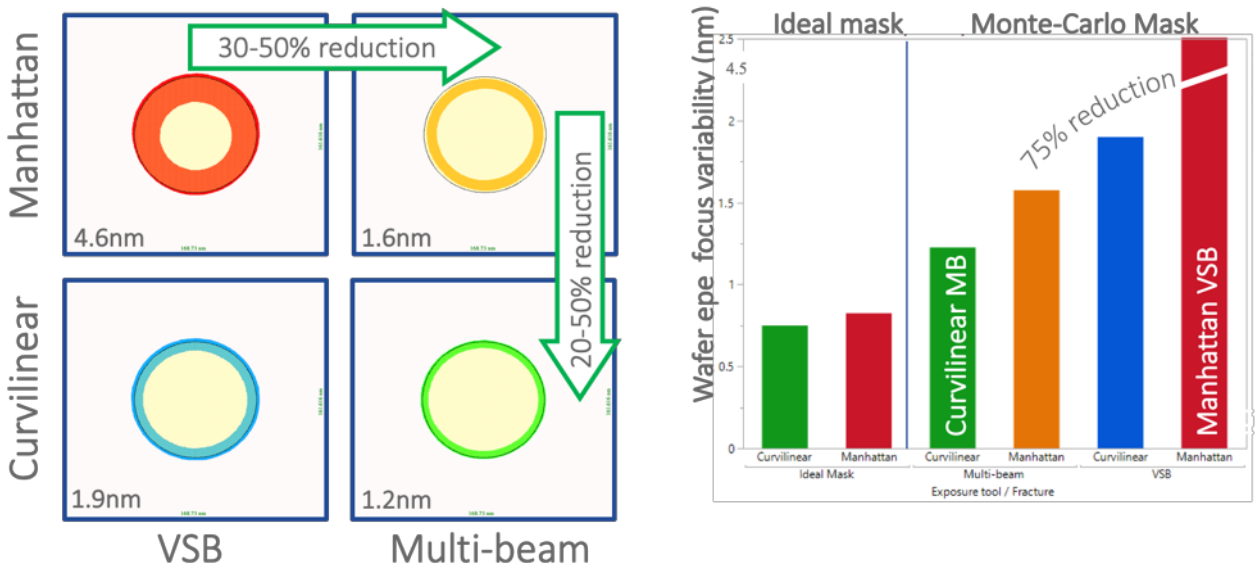


Figure 4: Study[20] shows the mask variation PV bands can be reduced 75% by switching from writing Manhattan pattern on VSB mask writer to writing curvilinear mask pattern on multi-beam mask writer

1.5 Run Time: The Challenge for Full-Chip ILT, Conventional ILT Takes Weeks to Compute for Full Chip

The biggest barrier to full-chip ILT has been runtime. The computations and models required for accurate ILT have been established and refined over the last decade since the introduction of the concept. The problem has been the sheer volume of the computations required to perform full-chip ILT and the runtimes that result.

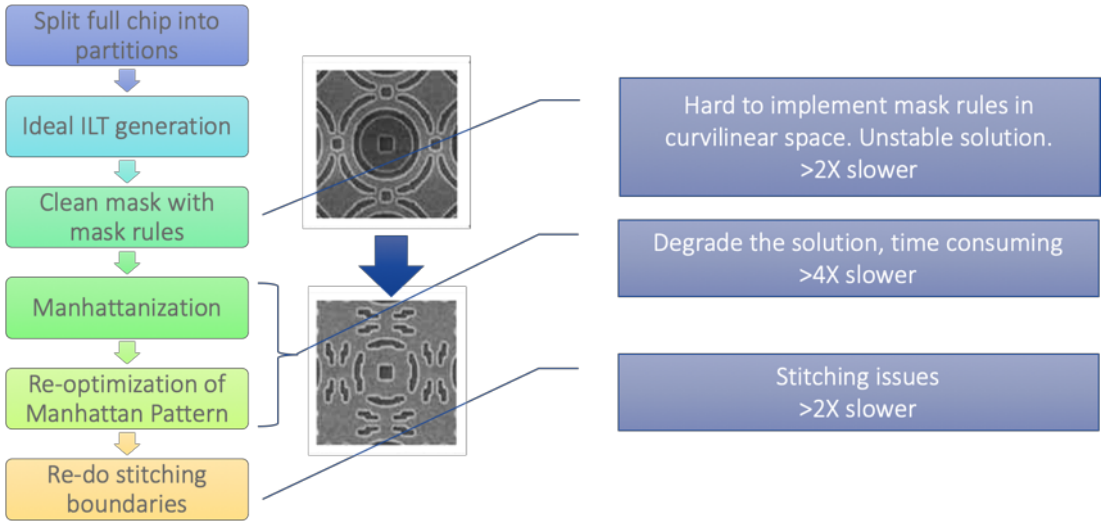


Figure 5: In the conventional ILT flow, in order to generate a solution for full chip and for VSB mask writer to write, the run time has increased order of magnitude due to all these extra steps

ILT's computation is already an order of magnitude higher than traditional OPC due to much bigger freedom of the solution space variable to compute. On top of this the standard approach for computations for full chip are too lengthy to be practical (Figure 5). Since each CPU can only handle ILT computation for a small area, conventional ILT divides the task (or in this case, the chip) into partitions, and have the computations for each partition run in parallel to save time. Each partition is passed to a CPU to process, it will first calculate the ideal ILT solution, then the ideal ILT mask solution is cleaned up to meet mask rules. After that it will go through a mask shape modification, called Manhattanization. This will

simplify the ideal mask patterns into Manhattan shape that VSB mask writer can handle. Since the mask shape is dramatically changed in this step, a re-optimization is required to enforce the new Manhattan mask pattern will meet the wafer pattern accuracy requirement and process window requirements. Then, the partitions are "stitched" back together. However, because the physics of any mask feature are impacted by the features adjacent to it, but each partition is processed separately without knowing the changes on the adjacent partitions even with the buffer region (called "halo"), it will produce inconsistency and discontinuities at the partition boundary, called "stitching error". Such stitching error must be corrected once the partitions are stitched together. The method to correct stitching error is to take regions around the partition boundary after merging plus some buffer region, re-calculate ILT solution, and then put them back. Such method will fix the existing stitching errors, but it may introduce new stitching errors at the new boundary. In addition, because the partition size that can be handled by CPU is relatively small, the buffer region due optical proximity is relatively large, these re-stitching areas are close to the size of the partition, therefore, doubling the run time. All of these steps on top of ideal ILT generation increase the run time significantly. At the end the total run time is an order of magnitude slower than the ideal ILT, which is already an order of magnitude slower than OPC.

As a result, commercial applications of ILT have been limited, and have focused mainly on smaller, high-risk portions of masks, mainly used in hot-spot fixer mode in order to make run time acceptable. A high-volume, full-chip ILT solution has been elusive.

2. STITCHLESS FULL CHIP CURVILINEAR ILT IN A DAY

Since multi-beam mask writer can write curvilinear mask in a constant write time, and curvilinear mask patterns are more resilence than Manhattan pattern, it is time to re-design a curvilinear ILT solution for multi-beam era.

2.1 The Solution: Get Rid of the Stitches

The rise, in the last decade, of the use of general-purpose graphics-processing unit (GP-GPU) computing for scientific applications has offered a new opportunity for bringing a practical full-chip ILT solution to market. GPU-accelerated computing excels at single-instruction, multiple data (SIMD) computation. This is in contrast to central-processing unit (CPU)-based computing, which excels at logical (if-then-else) computation. Simulations of natural phenomena (such as weather, or the physics effects inherent in semiconductor manufacturing) are SIMD computations, so GPU-accelerated computing is a natural fit for these operations, including ILT computations (Figure 6).

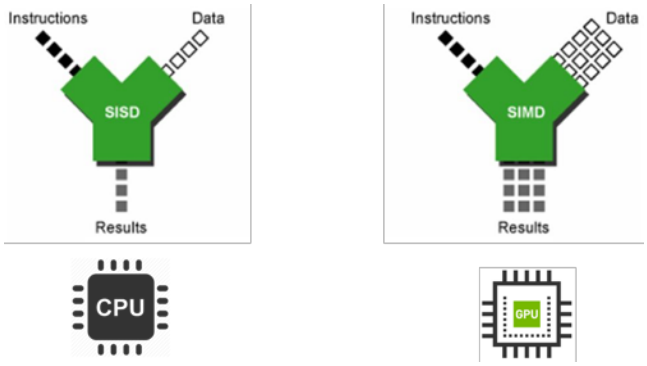


Figure 6: Illustration of CPU's SISD and GPU's SIMD

Of course, this is not a novel observation. Several attempts have been made to create commercial, full-chip ILT solutions by porting CPU-based solutions to a GPU-accelerated computing environment. However, these solutions have still fallen short in terms of acceptable turnaround time.

Partitioning/stitching has been the major culprit. Feeding chip partitions into a GPU-accelerated computing system can speed the processing of each partition. However, stitching errors and the re-computation required to address them are still show-stopping issues. D2S reasoned that what was needed was the ability to process the entire chip at once: a single, giant GPU/CPU pair that could optimize full-chip data seamlessly, without partitions.

2.2 TrueMask® ILT: Stitchless, Curvilinear Full-Chip ILT In a Day

Of course, such a giant GPU/CPU pair does not exist. However, by taking a "from the ground up," holistic approach, D2S was able to build an ILT-specific computing appliance that could *emulate* a giant GPU/CPU pair.

This approach didn't stop with the hardware, but rather included every component of a holistically conceived, purpose-built system – hardware, software, models, visualization, verification, etc. – that is designed and implemented from the ground up for GPU-acceleration and for full-chip ILT computation. Every aspect of the physics and chemistry of wafer lithography and processing, including litho simulations, mask and wafer models was examined and optimized synergistically throughout the system in order to reap the largest potential run-time benefits without compromising computational accuracy.

The product of more than ten years of development, D2S TrueMask ILT is the first commercial product to create accurate, full-chip ILT in a single day.

Stitchless

Chip partitioning and parallel computing is the most common approach to runtime reduction for full-chip computations. However, physics effects at advanced nodes are highly contextual, and partition boundaries naturally create contextual "disagreements" between items on either side of the boundaries. In addition, shifts that occur on mask can cause distortion of features that lay directly on the boundaries of a partition (think of misaligned sections of wallpaper). Handling these stitching errors – avoiding or correcting them – is one of the biggest hurdles for full-chip ILT (Figure 7a).

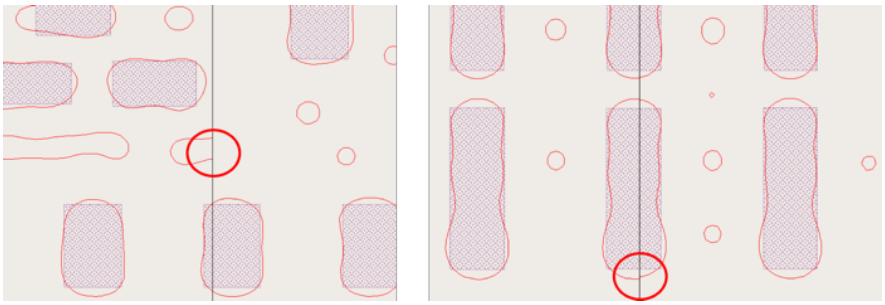


Figure 7: Stitching errors occur when a chip is partitioned for parallel computing and re-assembled

To avoid the time-consuming recursive correction passes necessary to resolve these stitching errors, D2S built the GPU-accelerated hardware platform (called the computational design platform, or CDP) and designed the software for TrueMask ILT so that the entire chip could be optimized at once. The D2S CDP has been purpose-built specifically to address simultaneous full-chip optimization. While it contains dozens of GPU-CPU pairs, TrueMask ILT, including the CDP and software, is designed to behave as though the whole system is a single, giant GPU-CPU pair that can process the mask for the entire chip simultaneously.

The system *behaves as though there are no partitions*. This means that each optimization iteration updates the entire chip as a whole, so that all proximity effects across the chip are accounted for with each update (Figure 8).

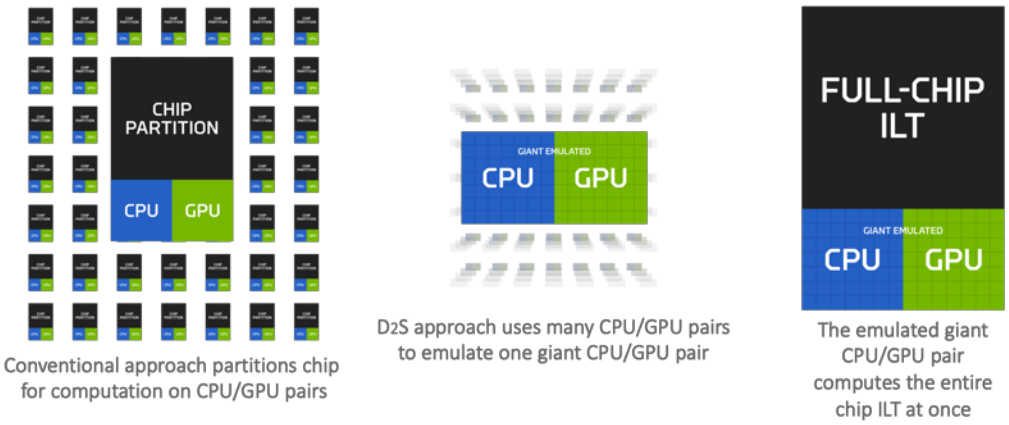


Figure 8: TrueMask ILT, although comprising many GPU/CPU pairs, has been holistically designed so that it behaves as a single GPU/CPU pair, iterating on the entire chip as a whole, and avoiding stitching errors

Because there is no partition boundaries, the solutions everywhere are continues, like what is shown in Figure 9.

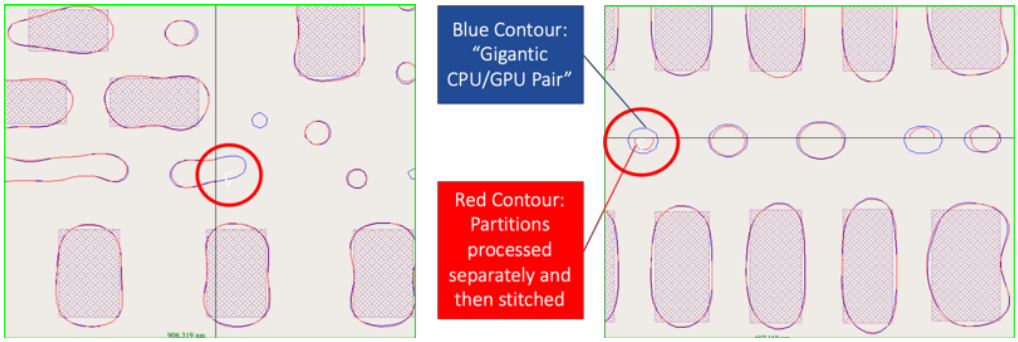
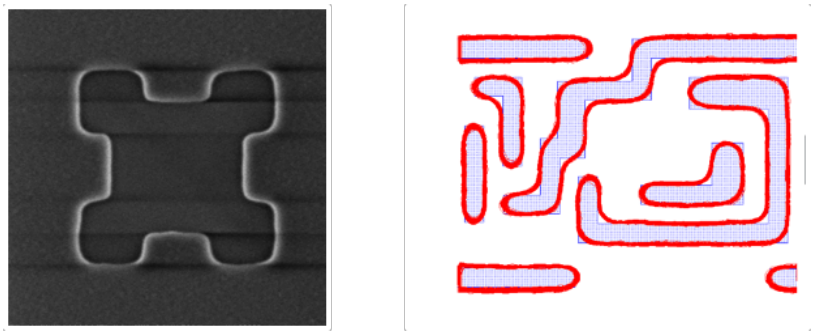


Figure 9: No stitching errors occurs in TrueMask ILT

Curvilinear

Because nothing in nature (including the physics of semiconductor manufacturing) makes 90-degree corners, manufactured masks and wafers *are all curvilinear*, even if the input geometries are rectilinear (see Figure 10). In fact, curvilinear shapes with certain minimum curvatures of shapes and spaces have been shown to be more reliably manufacturable than rectilinear shapes as we show in section 1.4[20].



(a) (b)

Figure 10: All shapes on masks and wafers are curvilinear, even if the input geometries are rectilinear: (a) Manhattan OPC mask pattern with serif shows curvilinear nature on mask; (b). Wafer pattern designed as Manhattan shows curvilinear nature from simulation

ILT is a mathematical approach that naturally produces curvilinear shapes. Traditionally, extra computation time has been needed to Manhattanize the curvilinear ILT shapes because variable-shaped beam (VSB)-based mask writing could not process curvilinear mask shapes within practical runtimes. With multi-beam mask writing now available, curvilinear shapes no longer require additional time to write. TrueMask ILT was built to leverage the power of this new world of multi-beam-based mask writing and is optimized for curvilinear mask output.

TrueMask ILT does equally well on curvilinear *input* design shapes. As multi-beam mask writers and EUV move into volume production, designers will be able to target curvilinear designs that are more manufacturable, TrueMask ILT will handle these designs with ease.

Uniquely, TrueMask ILT is able to compute curvilinear shapes efficiently because of GPU-acceleration. ILT inherently computes in the pixel domain; GPU-based computing was built for pixel-manipulation, so it is a perfect fit for this task. With its approach to emulate a giant GPU/CPU pair, TrueMask ILT computes, in essence, a rasterized image of the entire chip all at once, iterating on the full-chip ILT solution as a whole.

Full-Chip ILT

Full chip ILT has been the ultimate goal of ILT since its inception. It has been deployed only for "hotspots" and "critical areas" because the turnaround time for full-chip ILT was prohibitive. Ironically, however, stitching problems are more pronounced when "hotspot" ILT solutions need to be stitched into traditional OPC areas. There is no doubt that full-chip ILT is best, if run-time was not an issue. The unique approach of TrueMask ILT makes full-chip ILT a practical reality.

In a Day

For semiconductor companies, time is money, and time-to-market is critical for their revenue. This reality pushes semiconductor manufacturing companies, in particular, wafer fabs, to tape out and deliver wafers in its shortest time possible, which commonly constrains the budgets for OPC and ILT process time to one day. TrueMask ILT is the first commercial ILT solution that delivers full-chip ILT within this time constraint.

Integrates Curvilinear Mask Rules to Produce MRC Clean Results

Even TrueMask ILT creates curvilinear mask patterns, it still needs to meet mask rules, that is because mask process, similar to lithography process, is limited or affected by dose profile and contrast, resist resolution and etching process. Mask rules for curvilinear patterns is a whole new topic and will be discussed in a separate paper. In TrueMask ILT a set of curvilinear mask rules is integrated into the ILT optimization, therefore, TrueMask ILT produces Mask Rule Compliance (MRC) clean results. Figure 10 shows an example of TrueMask ILT without and with integrated MRC: when MRC is integrated, these features that violate minimum features are not appear in the final curvilinear ILT mask (Figure 11).

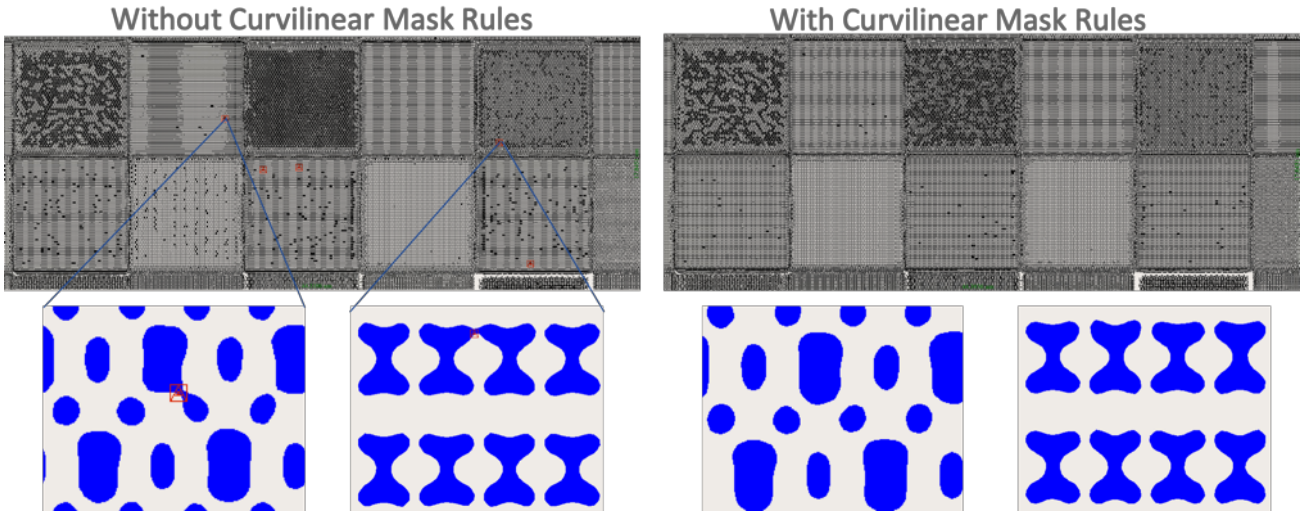


Figure 11: Comparison of TrueMask ILT without and with integrated MRC. The top row shows the layout where the red marks are MRC violations detected by D2S mask verification

Meets EPE Requirements

Although TrueMask ILT is a pixel domain implementation, its optimization can directly drive EPE to meet OPC requirements. Figure 12 shows the mask pattern, its simulated wafer contour, cost function and cost gradient at the beginning of the optimization. It is clear that the wafer contour does not hit the wafer target, cost function is not zero, and cost gradient is not flat. Figure 13 shows the situation at the end of the ILT optimization. Now the simulated wafer contour hits the wafer target, the cost function approaches zero, and gradient of the cost function is flat.

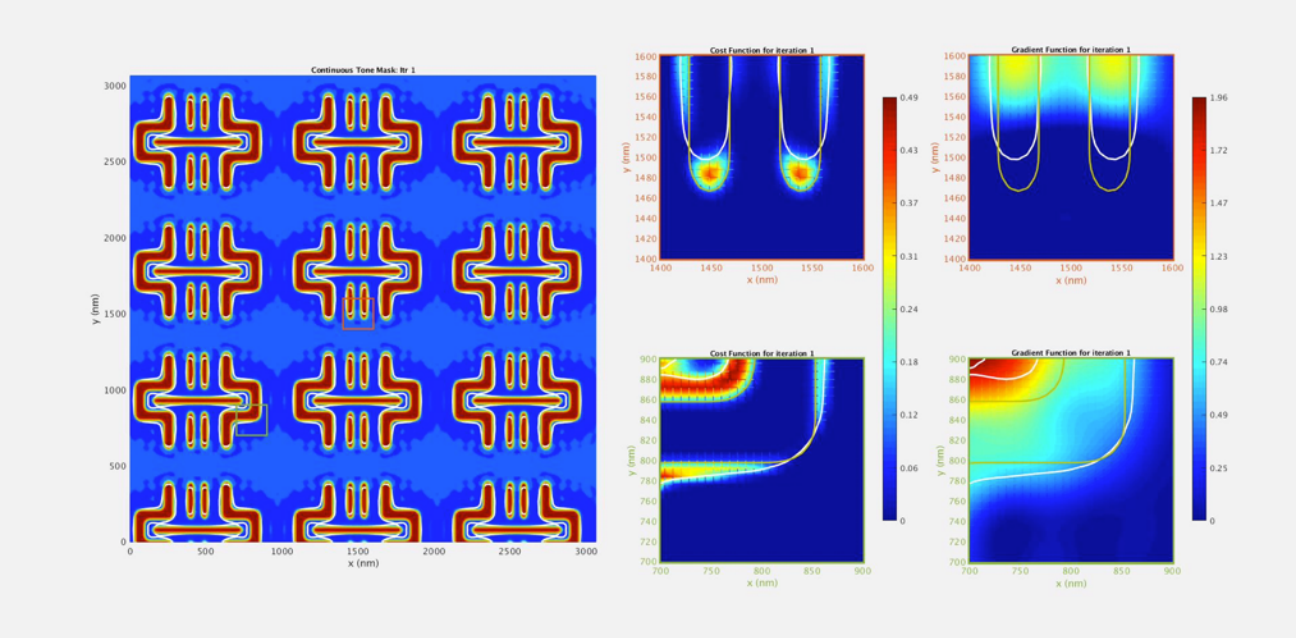


Figure 12: Mask pattern, simulated wafer contour and its target, cost function and cost gradient at the beginning of the TureMask ILT optimization

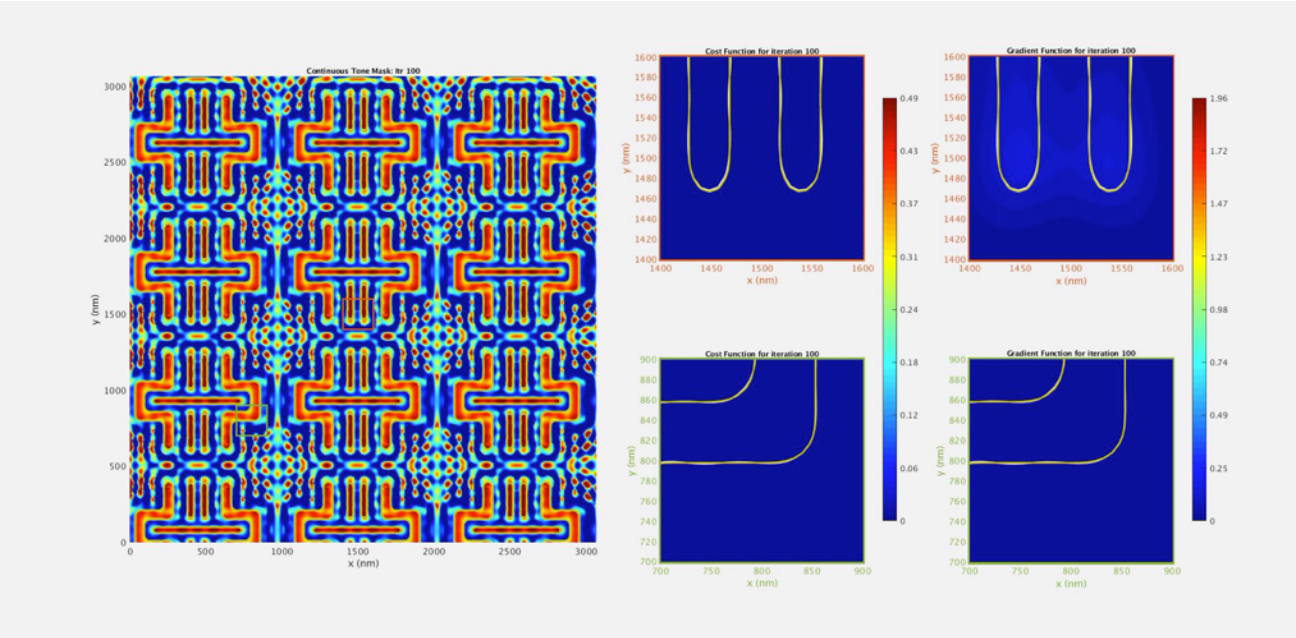


Figure 13: Mask pattern, simulated wafer contour and its target, cost function and cost gradient at the end of the TrueMask ILT optimization

Continuous and Symmetric

Solution continuity and symmetry are always the most difficult things for most ILT approaches. That is why most ILT papers only show ILT patterns for random patterns to hide their symmetry issues. TrueMask ILT has demonstrated the continuous and symmetric of its solution. Figure 14 shows a symmetric three contacts configuration. When pitches are changing from small to large, the TrueMask ILT solution gradually change while maintaining the XY symmetry.

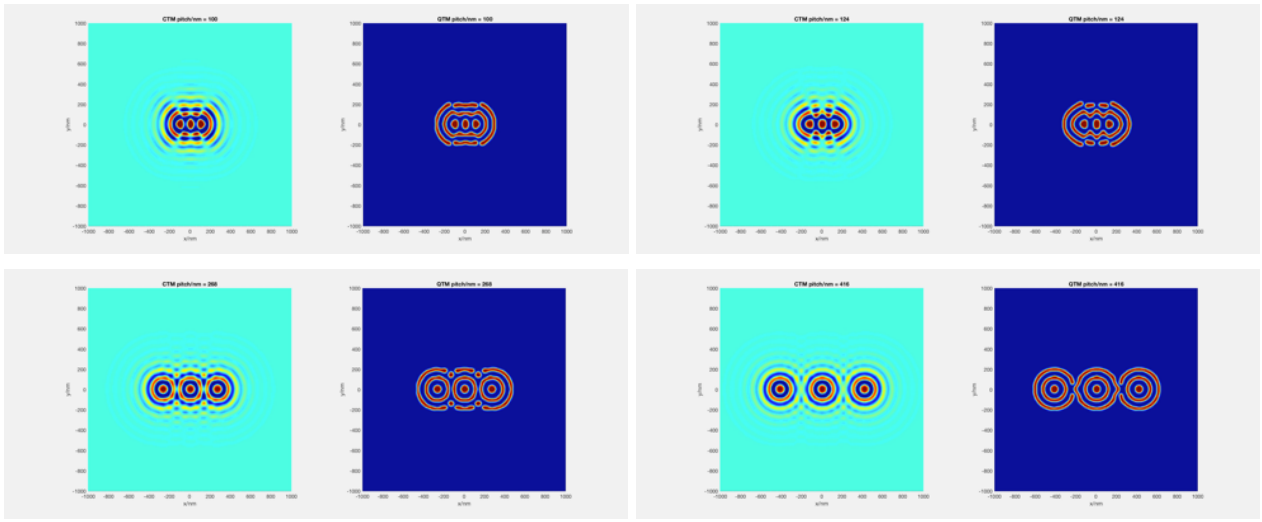


Figure 14: Continues Tone Mask (CTM) and Final ILT mask for three contacts in symmetric position at different pitches showing TrueMask ILT solutions are continuous and symmetric

On-Grid and Off-Grid invariance

Another challenge for most ILT approaches is the on-grid and off-grid invariance. TrueMask ILT has demonstrated that its solution for on-grid and off-grid are identical. Figure 15 shows an equal pitch contact array and its ILT solution. The top row is the on-grid case, while the bottom row is the off-grid case. When pitches are changing from small to large, the TrueMask ILT solution gradually change while maintaining the XY symmetry, and also the solution for off-grid case is identical to the on-grid case.

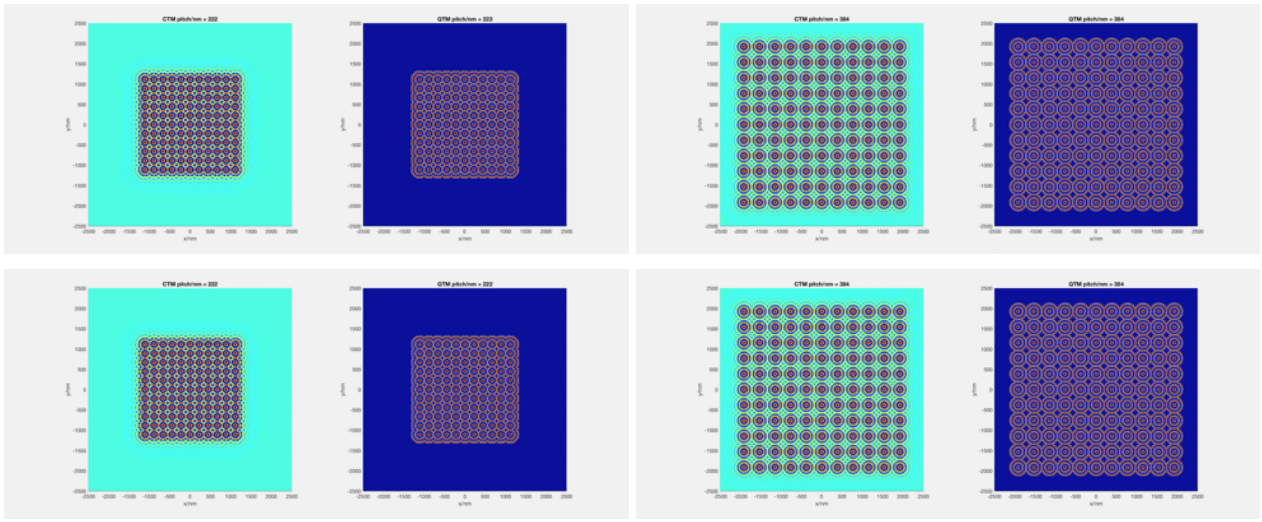


Figure 15: CTM and Final ILT mask for an equal pitch contact array for on-grid and off-grid situation. The top row are on-grid, while the bottom row are its corresponding off-grid configuration demonstrating TrueMask ILT solutions are grid invariance

Any Angle

The most challenge test for ILT is combination of multiple pitches, on-grid and off-grid, with rotation. Figure 16 shows the same equal pitch contact array and its TrueMask ILT solutions while the pitch is increasing, then also adding rotation. When pitches are changing from small to large, even with rotation, the TrueMask ILT solutions gradually change while maintaining the symmetry. Since the source is annular source, the ILT solutions are expected to be symmetric for any rotation angle, and we do see that from TrueMask ILT solutions.

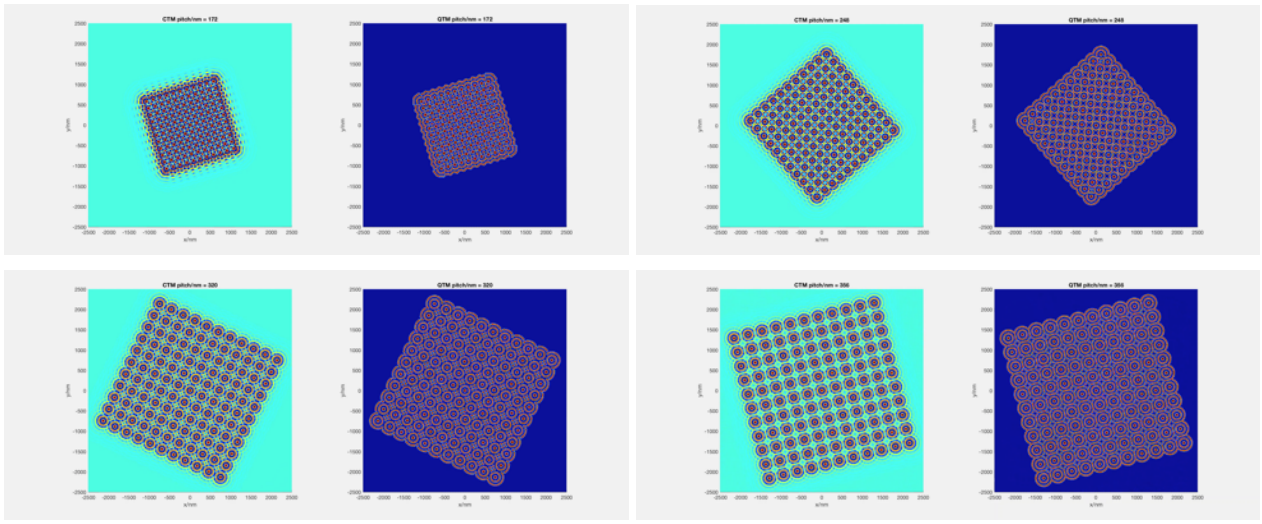


Figure 16: CTM and Final ILT mask for an equal pitch contact array at on-grid and off-grid situation, pitch change, plus rotation demonstrating TrueMask ILT solutions are symmetric and rotation invariance

3. EVALUATIN OF CURVILINEAR ILT ON MASK AND WAFER

3.1 Mask and Wafer: TrueMask ILT Results on Memory Design with Free-Form Source Demonstrated

To evaluate TrueMask ILT, masks were written and wafers were printed at Micron Technology using the process of record (POR). First TrueMask ILT model calibration was done using D2S test chip version 6. Then TrueMask ILT was run at Micron on CDP to generate curvilinear mask design.

Figure 17 shows curvilinear mask patterns generated for a free-form source used in production. The contact array has 11 by 11 contacts with equal pitch. The contact size is 40nm, and the tightest pitch is 85nm. A total of 98 curvilinear ILT mask patterns were generated by varying the pitch and rotation angle.

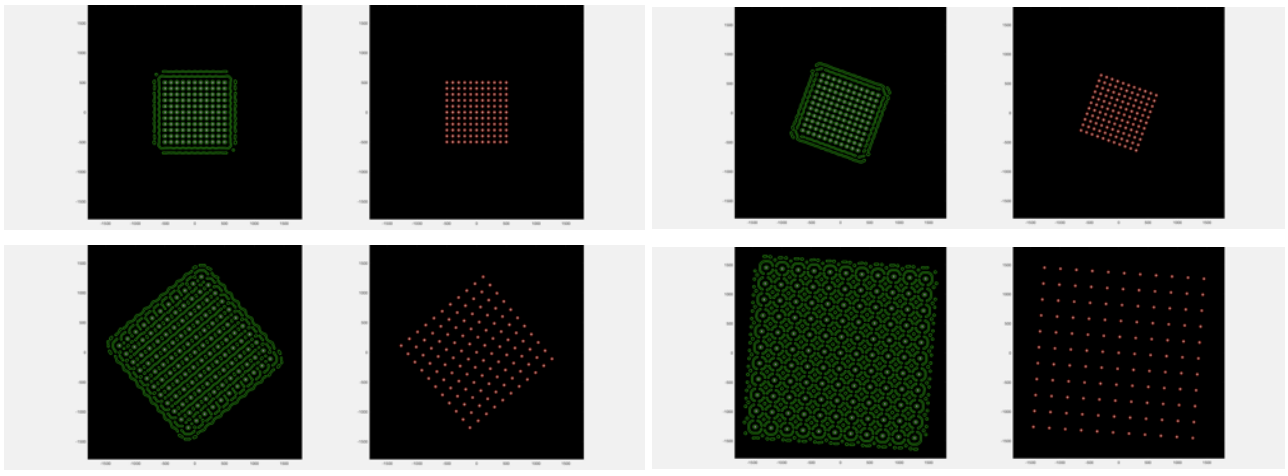
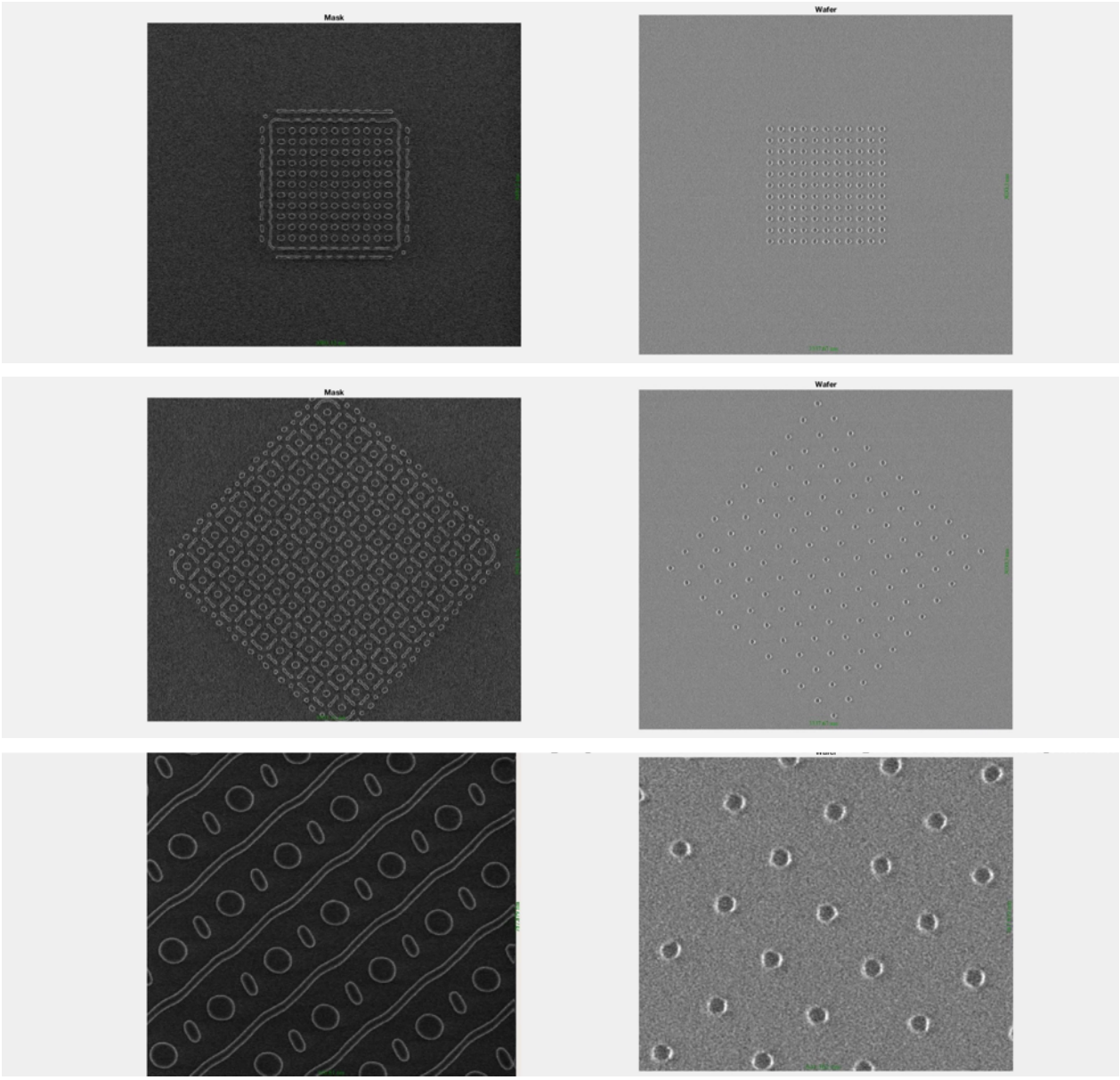


Figure 17: In each pair, the ones on the left are TrueMask ILT curvilinear mask designs for different pitches & orientations, the ones on the right are corresponding wafer target & simulated wafer contours

Figure 18 shows SEM images of some instances of the actual curvilinear mask pattern written by NuFlare's multi-beam MBM 1000, and wafer print using Micron's process of record. Mask patterns are resolved with high pattern fidelity and very smooth profile. On the wafer print, all contacts are printed evenly from the center of array to edge of the contact array.



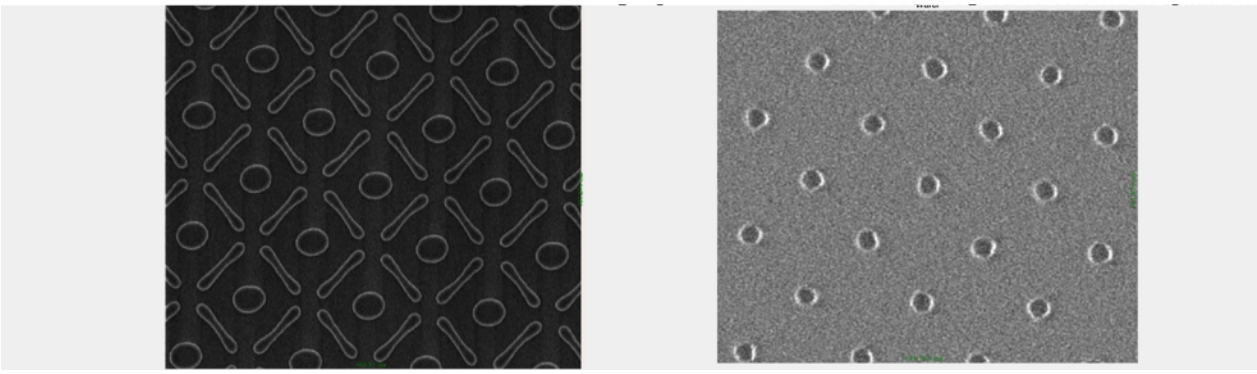


Figure 18: In each pair, the ones on the left are TrueMask ILT curvilinear mask written by NuFlare multi-beam mask writer for different pitches & orientations, the ones on the right are corresponding wafer print using Micron process

3.2 Process Window: Wafer Results Show TrueMask ILT Produces Much Larger Process Windows than OPC

The ultimate goal for curvilinear ILT is to achieve the best process window, so in this evaluation process windows were compared between OPC and TrueMask ILT using the same process.

Figure 19 shows the side-by-side wafer print comparison of OPC and TrueMask ILT at different process conditions (different focus and dose) from -60nm defocus to +60nm defocus and from 93.3% dose to 106.7% dose variation. Figure 19 randomly picked 6 process conditions from total 49 conditions, one can clearly see TrueMask ILT has much bigger process window than OPC: In many of the OPC wafer print the contacts are not printed evenly from the center of the array to edge of the array, some ones have necking problems, some ones do not even print at all. In contrast, TrueMask ILT wafer images show very consistent print for all process conditions, for all contacts no matter of their locations in the array, pitch, and angle of the rotation.

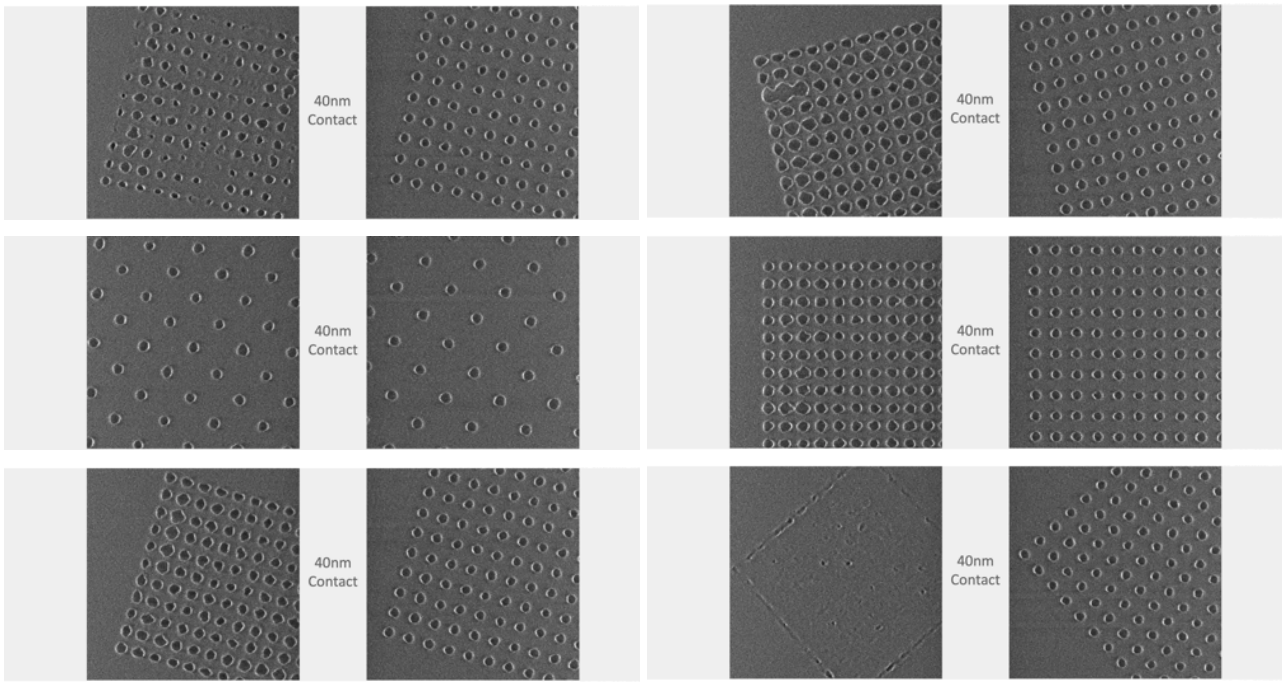


Figure 19: In each pair, the ones on the left are OPC wafer printed by Micron process at different process conditions, the ones on the right are TrueMask ILT wafer printed at the same process condition

CDs were also measured to quantify the size of the process window between OPC and TrueMask ILT. This was done on another cut layer type of design. Figure 20 shows the wafer prints for all process window matrix. The target CD is 62.8nm, all dies with CD with 10% variations are considered within process window. Figure 21 shows the CD measurements, the

conditions within process window are highlighted in green. Noticed the x axis is the focus, y axis is the dose to be consistent with process window plot. Three wafer images at process center and two process corners are also shown in zoomed-in version. Comparing to OPC, TrueMask ILT have increased the process window by over 100%.

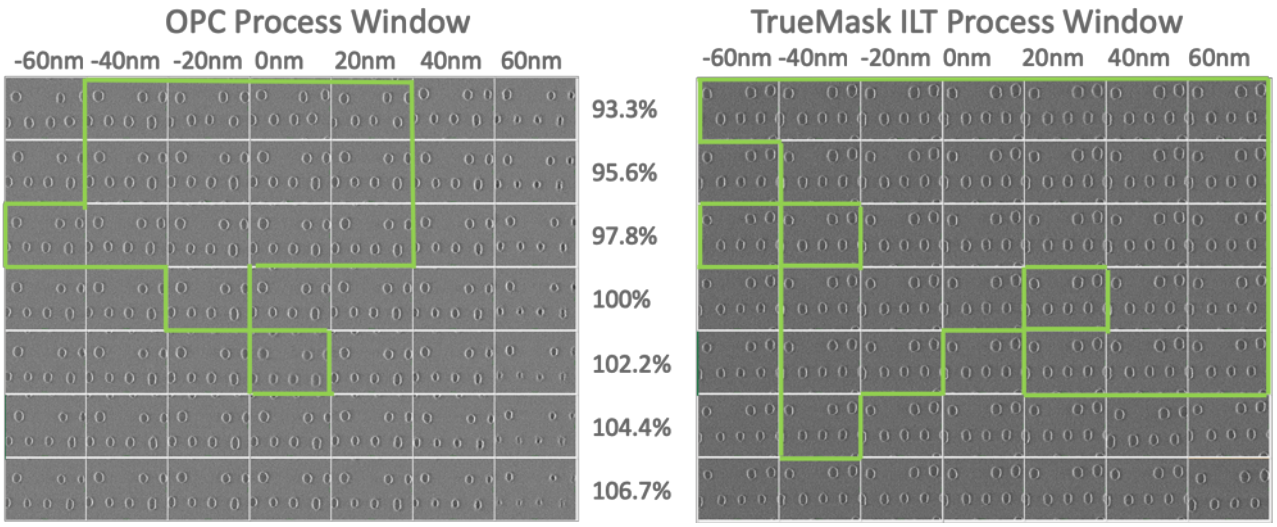


Figure 20: Wafer print matrix for a cut layer type of design. Highlighted regions are within process window

OPC Process Window

	Dose (mJ)	Defocus (nm) -60	-40	-20	0	20	40	60	Dose(%)
Latitude	21	55.4	59.6	62.5	61.2	60.5	51.6	41.8	93.3%
	21.5	56.3	57.3	63.7	61.0	58.1	52.9	40.3	95.6%
	22	57.3	60.8	61.0	60.8	57.1	48.3	41.6	97.8%
	22.5	54.0	52.8	60.1	55.8	54.7	52.1	39.3	100.0%
	23	49.4	54.4	55.3	60.3	54.5	49.2	36.1	102.2%
	23.5	49.2	50.9	54.5	55.0	51.6	46.1	33.9	104.4%
	24	54.0	56.2	54.8	49.9	50.9	44.7	28.6	106.7%

TrueMask ILT Process Window

	Dose (mJ)	Defocus (nm) -60	-40	-20	0	20	40	60	Dose(%)
Latitude	21	57.7	60.5	59.6	64.1	58.8	62.1	58.4	93.3%
	21.5	54.0	56.9	58.1	60.9	59.0	62.8	59.4	95.6%
	22	60.6	55.4	59.3	60.3	57.9	59.5	58.8	97.8%
	22.5	54.5	57.7	57.6	60.6	56.2	59.4	57.9	100.0%
	23	52.0	56.5	57.0	56.2	57.3	60.9	58.3	102.2%
	23.5	52.2	56.7	55.1	54.6	54.9	55.6	54.7	104.4%
	24	48.6	52.7	51.6	50.8	56.1	52.3	56.1	106.7%

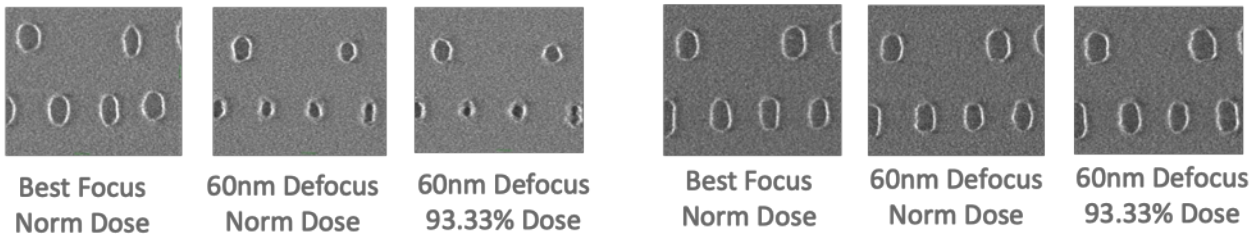


Figure 21: Process window CD measurements. The highlighted regions are within process window

4. SUMMARY AND CONCLUSIONS

Conclusion: ILT Vision Realized

For more than 10 years, the semiconductor industry has recognized the value of ILT in addressing the challenges of advanced-node lithography. Until now, runtime has been an insurmountable barrier to using ILT as a full-chip solution. Partitioning and stitching – the traditional approach to runtime improvement – have proven to be unsuccessful for full-chip ILT because of the computational time required to correct inevitable errors on partition borders.

By embracing a unique, holistically conceived, purpose-built system of GPU-accelerated hardware and software that emulates a single giant GPU/CPU pair, D2S TrueMask ILT iterates and optimizes the entire chip as a whole, making stitchless, curvilinear, full-chip ILT in a day a practical reality.

TrueMask ILT has been evaluated at Micron Technology. TrueMask ILT curvilinear mask was written by NuFlare multi-beam mask writer MBM-1000, the wafer was printed using Micron process of record. The results first show TrueMask ILT curvilinear mask pattern can be written by multi-beam mask writer with high pattern fidelity. Most importantly, the study shows TrueMask ILT has much superior wafer print quality and enlarged the process window by over 100% comparing to OPC, an astonishing result never seen or reported in the history of ILT.

5. ACKNOWLEDGEMENT

The author would like to thank NuFlare for their help to write the curvilinear masks with MBM-1000, and would like to thank ASML for their help to acquire wafer images with eP5 platform.

6. 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)

[9] Abrams, Daniel S., and Linyong Pang. "Fast inverse lithography technology." In *Optical Microlithography XIX*, vol. 6154, p. 61541J. International Society for Optics and Photonics, 2006.

[10] Pang, Linyong, Yong Liu, and Dan Abrams. "Inverse lithography technology (ILT): What is the impact to the photomask industry?." In *Photomask and Next-Generation Lithography Mask Technology XIII*, vol. 6283, p. 62830X. International Society for Optics and Photonics, 2006.

[11] Liu, Yong, Dan Abrams, Linyong Pang, and Andrew Moore. "Inverse lithography technology principles in practice: Unintuitive patterns." In *25th Annual BACUS Symposium on Photomask Technology*, vol. 5992, p. 599231. International Society for Optics and Photonics, 2005.

[12] Lin, Benjamin, Ming Feng Shieh, Jie-wei Sun, Jonathan Ho, Yan Wang, Xin Wu, Wolfgang Leitermann et al. "Inverse lithography technology at chip scale." In *Optical Microlithography XIX*, vol. 6154, p. 615414. International Society for Optics and Photonics, 2006.

[13] Hung, Chi-Yuan, Bin Zhang, Eric Guo, Linyong Pang, Yong Liu, Kechang Wang, and Grace Dai. "Pushing the lithography limit: Applying inverse lithography technology (ILT) at the 65nm generation." In *Optical Microlithography XIX*, vol. 6154, p. 61541M. International Society for Optics and Photonics, 2006.

[14] Ho, Jonathan, Yan Wang, Xin Wu, Wolfgang Leitermann, Benjamin Lin, Ming Feng Shieh, Jie-wei Sun et al. "Real-world impacts of inverse lithography technology." In *25th Annual BACUS Symposium on Photomask Technology*, vol. 5992, p. 59921Z. International Society for Optics and Photonics, 2005.

[15] Moore, Andrew, Timothy Lin, Yong Liu, Gordon Russell, Linyong Pang, and Daniel Abrams. "Inverse lithography technology at low k1: Placement and accuracy of assist features." In *Photomask Technology 2006*, vol. 6349, p. 63494T. International Society for Optics and Photonics, 2006.

[16] Hung, Chi-Yuan, Bin Zhang, Deming Tang, Eric Guo, Linyong Pang, Yong Liu, Andrew Moore, and Kechang Wang. "First 65nm tape-out using inverse lithography technology (ILT)." In *25th Annual BACUS Symposium on Photomask Technology*, vol. 5992, p. 59921U. International Society for Optics and Photonics, 2005.

[17] Chu, Chih-Wei, Becky Tsao, Karl Chiou, Snow Lee, Jerry Huang, Yong Liu, Timothy Lin, Andrew Moore, and Linyong Pang. "Enhancing DRAM printing process window by using inverse lithography technology (ILT)." In *Optical Microlithography XIX*, vol. 6154, p. 61543O. International Society for Optics and Photonics, 2006.

[18] Kim, Byung-Gook, Sung Soo Suh, Byung-Sung Kim, Sang-Gyun Woo, Han-Ku Cho, Vikram Tolani, Grace Dai et al. "Trade-off between inverse lithography mask complexity and lithographic performance." In *Photomask and Next-Generation Lithography Mask Technology XVI*, vol. 7379, p. 73791M. International Society for Optics and Photonics, 2009.

[19] 2018 eBeam Initiative Perceptions Survey Results [September 18, 2018], http://www.ebeam.org

[20] Pearman, Ryan, Jeff Ungar, Nagesh Shirali, Abishek Shendre, Mariusz Niewczas, Leo Pang, and Aki Fujimura. "How curvilinear mask patterning will enhance the EUV process window: a study using rigorous wafer+ mask dual simulation." In *Photomask Japan 2019: XXVI Symposium on Photomask and Next-Generation Lithography Mask Technology*, vol. 11178, p. 1117809. International Society for Optics and Photonics, 2019.