Curvilinear masks: an overview

Choi, Yohan, Fujimura, Aki, Shendre, Abhishek

Yohan Choi, Aki Fujimura, Abhishek Shendre, "Curvilinear masks: an overview," Proc. SPIE 11855, Photomask Technology 2021, 118550U (27 October 2021); doi: 10.1117/12.2601916

SPIE.

Event: SPIE Photomask Technology + EUV Lithography, 2021, Online Only

Curvilinear Masks: An Overview

Yohan Choi[*a], Aki Fujimura[a], Abhishek Shendre[a]
[a]D2S, Inc. (US)
[a]4040 Moorpark Avenue, Suite 250, San Jose, CA 95117 USA

ABSTRACT

Curvilinear masks are coming. With multi-beam mask writers in production, leading edge mask shops now are able to write curvilinear masks in the same mask write times as any Manhattan masks [1, 2]. As Samsung and Luminescent showed long ago [3], curvilinear mask shapes produce the best wafer process windows. For 193i masks, curvilinear SRAFs have been demonstrated in reasonable time even with the traditional variable-shaped beam (VSB) writers with good wafer results. [4]

It is widely anticipated [5] by the luminaries of the industry that curvilinear ILT shapes either are already or will be used at least for hotspots in some leading-edge layers before 2023 for both 193i and EUV masks. ILT solutions have previously focused on Manhattanizing the output to make it suitable for VSB writing, instead of using curvilinear or at least piecewise linear polygons to be specified for the desired mask shapes. With multi-beam mask writing being widely available for the leading-edge nodes, manufacturing curvilinear ILT shapes is now possible. But what about the rest of the mask making infrastructure?

This paper introduces the session on curvilinear masks by surveying the constraints and considerations around introducing curvilinear masks to mask manufacturing. The other papers of the session will address the data volume and computational complexity issues with contour geometry in polygon-based processing of data. We take note of pixel-based manipulation of data being constant in run-time performance regardless of shape, whereas manipulation of contour geometry scales in run time based on vertex or edge count. This paper also reiterates that curvilinear MRC can be significantly less complex [6]. Other aspects of the mask infrastructure including metrology, inspection and repair will be discussed.

The paper concludes with a brief discussion about curvilinear wafer targets, or curvilinear designs [7]. It suggests that allowing certain curvilinear targets can make designs more manufacturable and more resilient to manufacturing variation on the wafer, while decreasing power consumption, increasing clock speeds, and making designs smaller.

Keywords: Photomask, GPU, Inverse Lithography Technology, ILT, Curvilinear ILT, Mask Wafer Co-Optimization (MWCO), Multi-beam Mask Writer, VSB Mask Writer, Wafer process windows, MRC, Mask rule check

1. INTRODUCTION: WHY BOTHER WITH CUVILINEAR MASKS?

The mask manufacturing world is about to go through a big change. Shapes on masks used to be mostly axis-parallel rectangles or 45-degree triangles. Enabled by multi-beam mask writing, our world is now quickly transitioning to one that includes curvilinear shapes. Some masks of some chips are already largely curvilinear. Other masks of some chips are partially curvilinear, perhaps to address lithographic hotspots with curvilinear ILT. The trend to curvilinear shapes on masks seems to be moving forward for both 193i masks as well as for EUV masks.

The mask industry is already very busy with ever-increasing demand for precision of the leading-edge nodes and the additional issues surrounding EUV masks, and with the introduction of multi-beam mask writers. So, adding yet another major change by introducing curvilinear mask shapes may seem overwhelming. Yet as Figure 1 shows, industry luminaries widely expect the adoption of curvilinear masks by 2023 in this Luminaries Survey by the eBeam Initiative from 2020 [5]. In fact, 94% of the luminaries responding to this question expected at least some parts of leading-edge 193i masks to be curvilinear by 2023. Even for EUV masks, 85% of the luminaries responding to this question believed at least a part will be curvilinear. In an industry that generally changes slowly, this is a rapid change.

Photomask Technology 2021, edited by Stephen P. Renwick, Proc. of SPIE Vol. 11855,
118550U · © 2021 SPIE · CCC code: 0277-786X/21/$21 · doi: 10.1117/12.2601916

So, does this mean that the luminaries believe that the transition to curvilinear masks will be easy? Or do they believe there might there be some difficulties, but that it's worth the trouble? Or perhaps even that it is required in order to achieve the requisite process windows on wafer for the leading-edge nodes?

We believe that a combination of the latter two situations is likely to be true. It isn't easy, but as Chris Progler, CTO of Photronics, a merchant mask maker, summed up very well in an eBeam Initiative panel [8], "[A] limited number is manageable even today. But pattern inspection, repair and dispositioning are the areas that would need to be addressed for wider-spread adoption of curvilinear. Even though the writing time has been improved a lot, there are still other things that are going to prevent wide scale adoption."

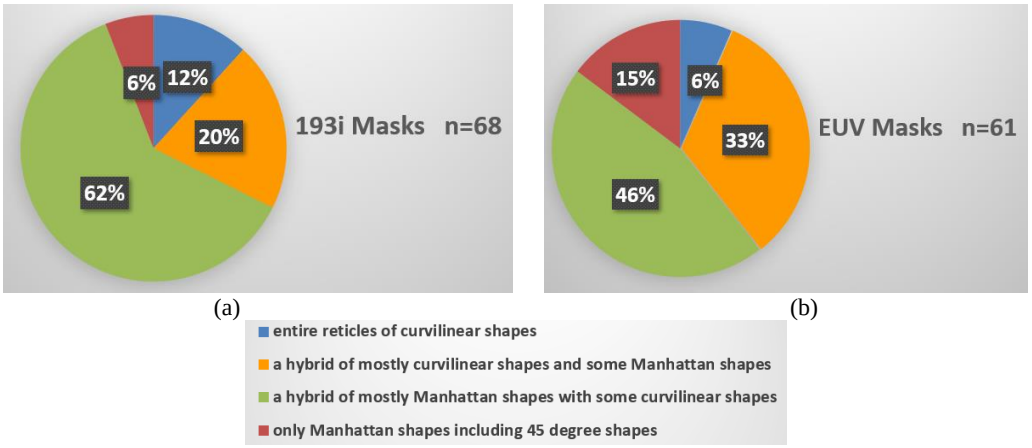


Figure 1. 2020 Luminaries Survey result for the question: "How extensively will curvilinear shapes be used for leading-edge (EUV, 193i) masks intended for high volume manufacturing (HVM) by 2023?," (a) 94% believe curvlinear shapes will be used for 193i for HVM by 2023, (b) 85% expect that EUV also needs curvlinear shapes for HVM.

Another survey question from the 2021 Luminaries Survey by eBeam Initiative [9], shown in Figure 2 reveals that the luminaries' biggest concerns in producing curvilinear masks are inspection and data infrastructures. Other than mask shop software infrastructures for curvilinear, mask back-end processes could be big challenges in mask manufacturing if all masks become 100% curvilinear.

The benefits to process window improvements on the wafer are well understood. And as we present in this paper, mask manufacturing is also simpler when target shapes are manufacturable on mask. Further, manufacturable shapes are more reliably manufacturable, which is good for both mask and wafer quality.

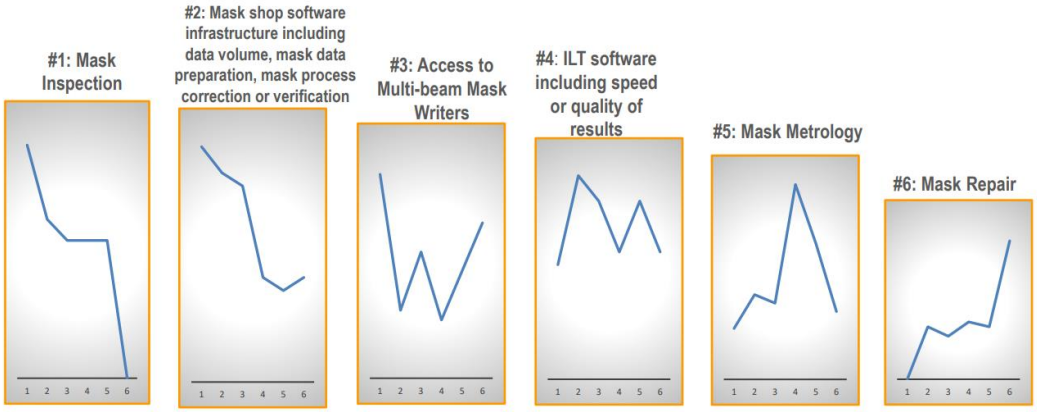


Figure 2. 2021 Luminaries Survey result to the question "Rank your biggest concerns in producing curvilinear masks"

Why do the luminaries believe curvilinear masks will be heavily deployed in EUV masks as well? In Figure 3, the luminaries indicate clearly that the primary reason to purchase multi-beam mask writers is for EUV masks. They also indicate that writing curvilinear masks is also a strong reason to purchase multi-beam mask writers. It is interesting that in the 2021 survey, the third most popular response was "Curvilinear ILT for EUV Masks" while in the 2020 survey [5], the third most popular was "Curvilinear ILT for 193i Masks." The authors believe that these responses are consistent with EUV masks needing to be written with multi-beam mask writers regardless of shape. Given that EUV masks are being written with multi-beam mask writers already, there is no penalty in the mask write time to write curvilinear shapes. It is difficult to leave the extra wafer process window on the table. [10]

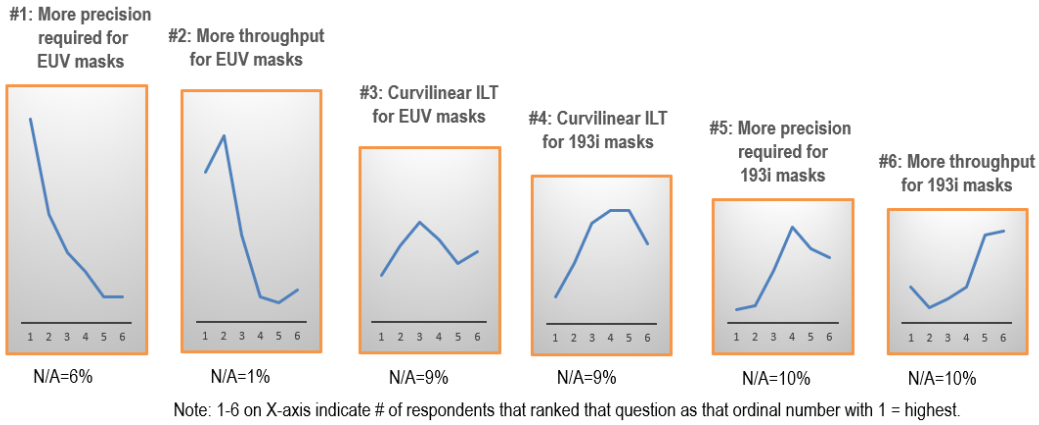


Figure 3. 2021 Luminaries Survey result to the question "Please rank the primary reasons for purchasing multi-beam mask writers." The grey bars with orange borders indicate the average of (7 – rank) that were provided, while the inset blue graphs indicate the distribution of the responses.

Multi-beam mask writers do not scale their mask write times by the complexity of the shapes being written. Given a resist sensitivity and a writing method, the write time is constant regardless of the shape. The traditional VSB mask writers write one axis-parallel rectangle shot or a 45-degree triangle shot at a time. Complex shapes are fractured into a series of shots. Diagonal or curvilinear shapes are approximated by stair-stepping a series of shots at some determined shot (stepping) size or approximated using overlapping shots as discussed further below. Mask write times of VSB machines scale by shot count.

Target shapes consisting of axis-parallel edges are sometimes referred to as Manhattan geometries. In the mask writing context, however, the term Manhattan geometries may be used to include 45-degree edges because VSB mask writers are able to write 45 degree triangles with the two short edges of equal lengths being axis-parallel. The length of these short edges cannot be longer than the maximum edge length of a rectangular shot on that machine. It is notable that an axis-parallel trapezoid's horizontal edges with 45-degree edges on left and right sides would be shot as three shots consisting of a rectangle and two 45-degree triangles.

Target shapes that do not need to be Manhattan geometries are considered curvilinear in the context of the present discussion. So specifically, piecewise linear polygons with at least one edge that is neither Manhattan nor 45 degrees to the axes, as well as infinitely curvilinear descriptions such as circles, ovals, arcs, or spline-like formats are all considered curvilinear in the context of the present discussion.

1.1 Curvilinear masks are better for the wafer

Curvilinear masks have been discussed in the semiconductor industry for years as it is well-known that curvilinear masks produce larger process windows as well as better manufacturing resilience compared to rectilinear masks (a.k.a., Manhattan masks). In 2009, Samsung and Luminescent [3] demonstrated the standard OPC mask and the curvilinear mask to evaluate wafer depth-of-focus (DOF) performance as depicted in Figure 4.

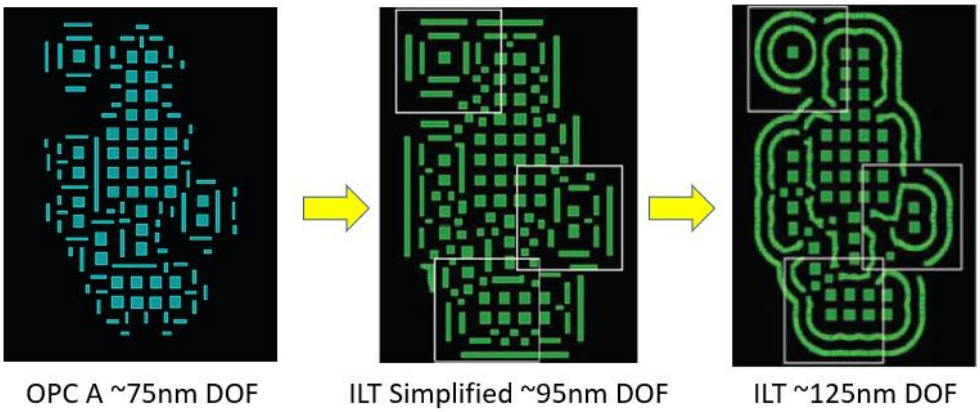


Figure 4. Samsung and Luminescent demonstrated that curvilinear ILT produces ~70% better DOF than typical OPC mask [3].

And, recently, Micron demonstrated that process window and process variation (PV) band improved substantially with curvilinear ILT as shown in Figure 5. Depth-of-focus (DoF) improvement of about 85% was demonstrated [5].

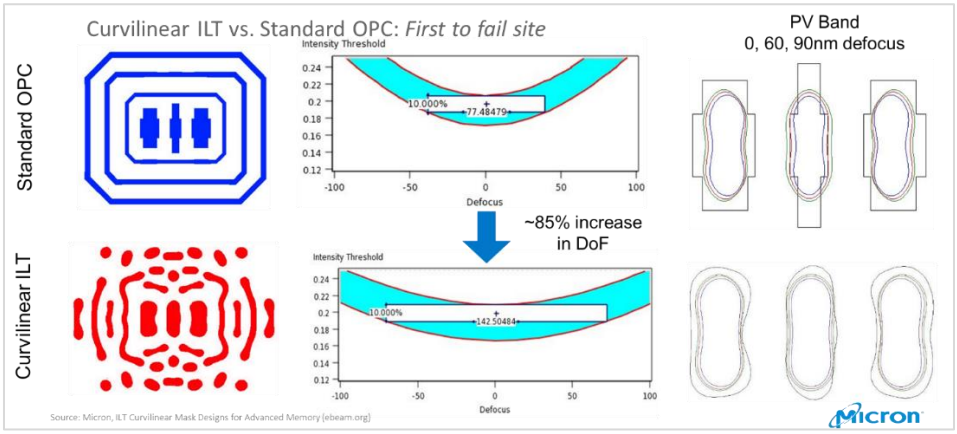


Figure 5. A standard OPC mask and curvilinear mask were evaluated on the wafer by Micron to compare the process window as well as PV band. TrueMask® ILT produced almost 2X larger process window with better PV band than standard OPC.

Wafer quality is improved by using curvilinear shapes on masks as derived by curvilinear ILT. Yet until multi-beam mask writing became available, this improvement could not be achieved practically with only VSB-based mask writing. One could theoretically use VSB shot sizes that are 10-20nm squares with offset passes like multi-beam writers use, but since these shots are made one at a time, a mask written this way would take too long to write and wouldn't be as accurate either. Conventional fracturing of the shapes that approximate the curvilinear targets with Manhattan approximations of some given tolerance can come close to the desired shape, especially with mask-simulation-based mask process correction (MPC) by taking advantage of the naturally rounding nature of mask processing. To target the desired mask shapes sufficiently accurately under nominal conditions, however, the Manhattan jogs near the curvilinear contour edges would need to be sufficiently small (such as 10-20nm jog lengths) which again increases the VSB shot count prohibitively.

1.2 Mask-Wafer Co-Optimization enables VSB writing of curvy mask shapes for 193i

The invention of mask-wafer co-optimization (MWCO) solves this problem for 193i masks [11]. For EUV masks, MWCO is not appropriate for various reasons. In any case, EUV masks are written by multi-beam writers even for Manhattan mask targets.

Prior to MWCO, a curvilinear shape could be approximated on the mask with reasonable shot count using overlapping shots [12]. The overlapped parts that tend to define the interior corners of the undulating shapes bloat out the concave parts and the natural corner rounding of the normal 1x dose takes in the outer convex parts, naturally smoothing the jaggedness of the shape produced by the union of all shots. By doing simulation-based MPC with overlapping shots, VSB shots counts can be substantially reduced. Curvilinear sub-resolution assist features (SRAFs) are the narrowest features on the mask, so a high percentage of VSB shot count of writing a complex Manhattan or a curvilinear mask is comprised of SRAFs. Yet main features affect wafer quality much more than SRAFs do. Less-important features requiring the greater percentage of shots is unfortunate, hence a methodology to use overlapping shots for SRAFs, but to use conventional fracturing for main features was developed.

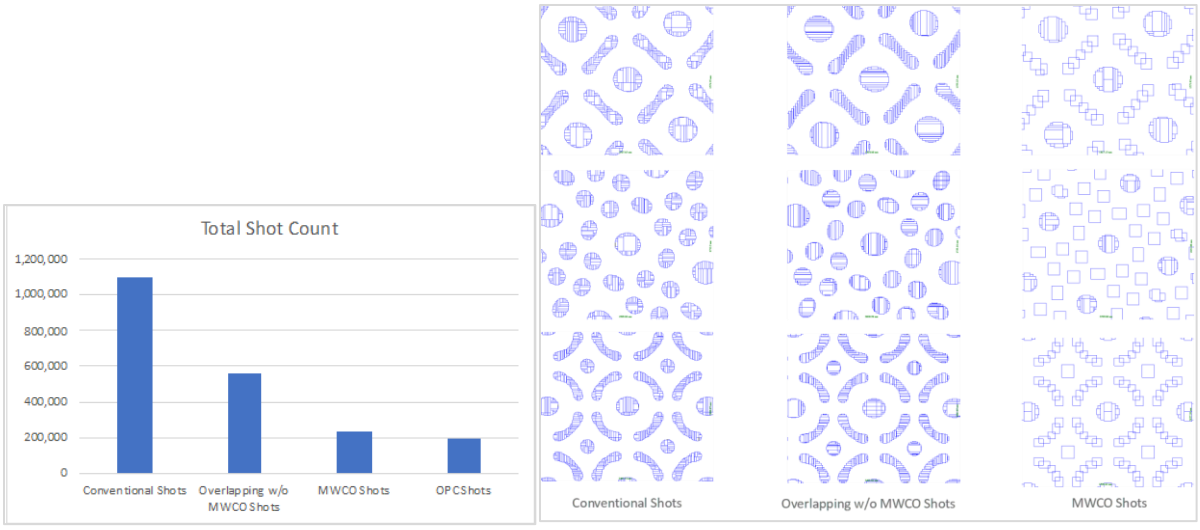
With advances in technology nodes, however, accuracy requirements became more and more stringent. It became increasingly important that SRAFs that were accurate enough "as far as 193i can see" be more precise. The initial idea with the overlapping shots was to process it purely in the mask shop, so there was discomfort that even after simulation-based MPC the undulating contour of the resulting SRAFs was approximate, with edge placement error (EPE) on mask of up to 10nm (albeit with average around the contour optimized to being within 0.1-0.2nm). There was, of course, a trade-off between shot count and contour EPE.

MWCO improved on overlapping shots by having ILT produce curvilinear mask shapes first, then doing overlapping shot generation, and finally iteratively optimizing for *wafer* quality by moving shot edges while doing mask-wafer double simulation. It is similar to how conventional ILT would Manhattanize its output shapes for VSB writing. However, it is different in that MWCO performs mask simulation prior to wafer simulation in each iterative optimization step. With the increased accuracy requirements of the leading-edge nodes, the shot corners are sufficiently close to each other as compared to the blur radius of a typical mask process using CAR resists that even with overlapping shots, rule-based corner rounding is insufficient to achieve requisite precision in computing the wafer quality. Even though using overlapping shots reduces this effect (because the shot corners are farther apart), the additional complexity of the overlapping parts using 2x dose make it necessary to do the double simulation.

In the MWCO methodology, the OPC shop performs ILT and produces a candidate VSB shot list with overlapping shots. This shot list has been simulated and iteratively optimized to produce the wafer quality deemed to be sufficient by the OPC shop. Then the shot list, along with the computed target mask shape is output to the mask shop. The mask shop is then tasked to produce the computed target mask shape, knowing that the overlapping shot list provided is able to produce that contour with a mask model that was used by the OPC shop. The actual mask model used in the mask shop may vary. There may be other differences such as shooting SRAFs with higher doses than main features to improve dose margin or linearity. This may have been anticipated in the OPC shop, but perhaps the exact amount of dose is slightly different with the batch of resist being used that day. Regardless, because the OPC shop has provided an existence proof of a shot list that was able to produce the curvilinear (undulating) target mask shapes that had been used by ILT to optimize the wafer quality, the mask shop is able to hit the target contours accurately while adjusting mask models.

It should be noted that this methodology much more faithfully produces a mask shape that is the same as what any OPC process assumed to be the mask when OPC optimized for the wafer quality, even for conventional Manhattan OPC. Simulating the mask shape during OPC/ILT is much more accurate than relying on simple corner rounding. The mask maker is given the exact target contour of the mask to produce, instead of a Manhattan approximation with inaccuracies in the corners being excused. These inaccuracies are in the 10nm range in typical processing. When masks are full of these corners and the corner-to-corner distances are within the blur radius of the mask process being used, it is extremely difficult for the mask assumed by OPC/ILT optimization to be the mask produced by the mask shop, particularly for contact or via layers where there are hardly any 1D features on the mask at the leading edge nodes, but even for line-and-space layers where lines have many small jogs after OPC/ILT. The problem lies in the handoff using a Manhattan pattern. A Manhattan pattern as the target mask is inherently an approximation. It is better for the OPC shop to specify the actual curvilinear target shape that will be physically produced on mask.

Micron produced VSB-written masks and wafers for 193i to validate the process window benefits of curvilinear ILT with MWCO as depicted in Figure 6.



(a) (b)

Figure 6. (a) VSB shot count and (b) shot configurations for three contact arrays. Note the POR OPC shot configurations are not shown in (b)

2. CURVY MASKS ARE ENABLED

So MWCO enables 193i masks written by VSB writers to benefit from a more precise handoff from the OPC shop to the mask shop by specifying the target physical mask shape to be produced on the mask, rather than a Manhattan approximation of it. Because 193i can't "see" as well, small differences in the mask don't reflect on the wafer as long as the local average areas are preserved. But even when written by VSB, curvilinear specifications are superior to Manhattan specifications when the shapes are complex. But EUV can "see" much better. Smaller mask differences reflect on the wafer with EUV. Whether for 193i or for EUV, curvilinear mask shapes produce better wafer quality. With sufficient supply of multi-beam writers, leading-edge masks are likely to be written with them in the future.

In turn, with multi-beam mask writers available in all leading-edge mask shops, the mask write times are no longer affected by the number of shapes on the mask or their complexity. The economics of mask writing is dominated by the mask writing time. The depreciating daily cost of the leading-edge mask writer necessary to write the leading-edge masks is a dominating factor by itself. But more importantly, mask yield is a function of mask write time. [13] The longer the write time, the more likely that something goes wrong during the writing of that mask. Because these are precision machines shooting negatively charged electrons at the resists, it is sensitive to environmental factors like small earthquakes or the effect on the magnetic field due to solar flare. As much as these machines and entire mask shops are designed to be resilient to these factors, it is difficult to eliminate all variation. Production schedule control is also much easier with consistent write times. The fact that multi-beam mask writers, given a resist and writing method, writes any shapes of any shape count in constant time is economically and logistically very attractive to the mask shop.

Once a mask shop has a multi-beam mask writer, curvilinear masks take no more time to write than any other. This is principally because multi-beam mask writers write with pixels, similarly to how TVs, monitors, and digital projection machines work. The device that you are reading this document with, even if it was printed on paper, was processed using pixels of a certain resolution. Pixel doses are manipulated to create grey scale per pixel. In mask writing, the naturally rounding nature of resist exposure turns these grey-scale pixels into smooth contours that divide the resist surface into exposed regions above the resist threshold and unexposed regions below. With sufficiently small pixel sizes, the generated contours become smooth, just in the same way that your eyes perceive the characters on this page as being perfectly curved characters, unless you inspect the pixels closely.

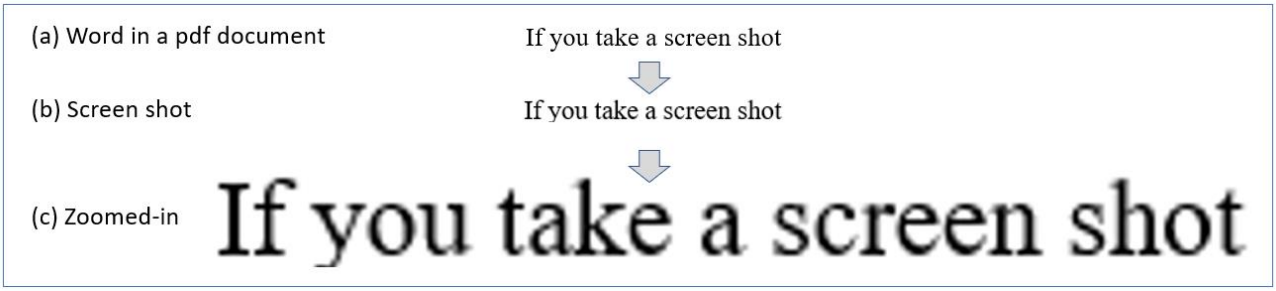


Figure 7. An example of pixelated image; the pixelated image with finite resolution is good enough for reading a PDF document. The image resolution is the trade-off in the cost of data volume and process time.

It is notable that if you zoom way in on this document with a PDF reader, for example, the characters still remain perfectly smooth because this document is prepared with fonts that have infinite resolution. If you take a screen shot of the page as was done in Figure 7, however, and then zoom way in with the same PDF reader, you will see pixels. This difference is the same difference being proposed by proponents of spline-like formats for curvilinear masks. While infinite resolution is clearly nicer, it is not necessary for mask making where all consumers of the data have a finite resolution. One could argue that parts of the data processed in software may remain infinite resolution, but 1) it isn't necessary because the destinations all have limited resolution and 2) even a simple operation like constant biasing of curvilinear contours requires an assumption of the required accuracy which amounts to a resolution limit given a certain runtime constraint. The question with infinite resolution formats like splines then becomes that of the trade-off in the cost of runtime to produce the data and the loss of accuracy in producing that data. Nevertheless, infinite resolution may be preferred for certain specific applications.

2.1 Pixel manipulation enables curvy masks

Multi-beam mask writers, which write with eBeam pixels, have practically enabled curvilinear mask writing. ILT that produces the curvilinear target mask shapes also computes in pixels. Lithography simulation is done with fast Fourier-transforms (FFTs) in pixel domain and is fastest when run with GPU-acceleration because FFT is a great application for a single-instruction multiple-data (SIMD) GPU computing machines. There are many tasks that are better with a CPU, but there are some tasks that are at least one and typically two orders of magnitude faster with GPUs. Gaussian convolutions such as those central to eBeam exposure simulations and FFTs as those central to lithography simulation (whether 193i or EUV) are both examples where SIMD computing is ideal, performed in the pixel domain. NVIDIA A40, a modern GPU as of this writing, contains 10,752 CUDA cores on one chip and is rated at 37.4 TFLOPS. [14] TFLOPS is a million times a million floating point operations per second on one chip. This can only be accomplished by keeping each processor small by effectively limiting the instructions being executed on all cores to be the same at any given time. It processes the same instructions in the same sequence on all cores (or some subset), but each core is operating on different data.

Since natural phenomena are SIMD, SIMD computing is generally the method of choice for computing large scale simulations of natural effects. The same physics, chemistry and math is acting on everything, but each element has different data like temperature or wind direction or magnetic field or blur or acid generation or gradient, or whatever else that causes complex and widely varying behavior. It is not the procedure that is different. It is the data that is different. This is true for mask simulations or wafer simulations, too.

Of course, graphical processing unit is what GPU stands for, and image processing is a SIMD computation. GPUs became enormously popular and therefore accessible in pricing to our community because of gaming. Photorealistic rendering with shading, shadowing, reflections, and smoothing in real time for interactive gaming requires this complex processing to be done fast enough to keep up with frame rates. The SIMD engine built for that purpose is ideal for pixel-based manipulation of images of all kinds. While edge-based computing of curvilinear data that is represented as polygons or spline-like formats can also be accelerated with GPUs, the degree of acceleration is greater when data is processed as images with pixel doses as the base representation. While CAD applications mostly work with edge-based contour geometry data rather than pixel doses, for mask making, many of the operations such as metrology, inspection and repair inherently work on images of the mask. GPUs are already in prominent use for processing the data in our industry because of that.

NuFlare's pixel-level dose correction (PLDC) capability, available with the MBM-2000, [15] is another example of the unique capability of GPU-acceleration doing pixel manipulation rather than edge-based contour geometry manipulation. In exactly the same way as multi-beam machines write masks in the same time regardless of mask shapes or shape count, pixel-based MPC runtime scales based on the number of pixels, not on the number of edges or vertices as is the case with edge-based contour geometry manipulation. While contour-geometry input files need to be rasterized to convert to the pixel domain, and that conversion time is a function of edge count, the total compute time is dominated by MPC itself, which is performed in pixel domain for PLDC as indicated by its name. It is performing an MPC function in addition to the edge-dose enhancement function for dose margin improvement, but it is doing so by manipulating pixel doses, not by manipulating geometries by moving edges or vertices. Unlike offline MPC functions that need them to convert the data back into edge-based contour geometries, because the PLDC capability is inline inside of the mask writer, the output is simply the pixel doses, without conversion.

Because PLDC manipulates the mask writer pixels on GPUs, the processing is extremely fast as compared to conventional MPC methods, and it is constant time regardless of the shape count or shape complexity. Thus, it is able to run inline with the machine as the machine writes, because both it and the machine work with the same pixel domain.

With the NuFlare MBM-2000, the mask data is generated using pixels, then MPC is performed inline within the mask writer as the mask is being written with pixels. In between ILT and the writer, however, there is a conversion to edge-based contour geometries in the file format. Saving pixel data in image data would result in too large a volume of data, even if compression is applied, especially since a lossless compression is needed. In addition, the rest of the mask infrastructure requires edge-based contour geometries. This is why discussion of possible curvilinear data formats are receiving much attention in the industry today. [16-19]

An additional benefit of pixel-domain computation on GPUs is deep learning. Whether applied to image processing or other applications, deep learning is highly accelerated in SIMD computing architectures because the neurons that inspired deep learning are the creation of nature. All deep learning neurons work the same way. The complex behavior deep learning achieves are the results of different data, not different procedures. As previously noted, conversion from edge-based contour geometry to pixel domain and back is expensive. Being in the pixel-domain already for other computing makes deploying deep learning inferencing as a part of the process much more efficient.

Our industry gets to benefit from all of the investment that goes back into each successive generations of GPUs from the profits generated by gaming, bitcoin mining, autonomous driving, and deep learning. For processing curvilinear masks, GPU is the right computing vehicle.

2.2 The mask ecosystem is ready for curvy ILT

What about the rest of the mask making ecosystem for curvilinear ILT? The predominant opinion of industry luminaries [9] is that mask shops are ready to handle curvilinear masks today as required. But implied in that sentiment is that the demand for curvilinear masks will be scaling over time. A limited amount is okay today, even for high-volume manufacturing (HVM) purposes (not just for test chips or experiments), but if all masks became curvilinear overnight, the luminaries anticipate some bottlenecks.

There are two key driving factors for this readiness. One, manufactured masks and pictures taken of manufactured masks are already curvilinear. Two, curvilinear masks are more manufacturable.

Even eBeam, as accurate an instrument as it is, cannot produce 90-degree corners or perfectly straight edges, particularly when practically reasonable resists are used. So, metrology, inspection, review, and repair all already deal with curvilinear shapes. In particular, the need for wafer plane mask review is anticipated to increase with curvilinear mask shapes. Metrology methodologies today rely on measuring critical dimensions (CDs), usually by measuring horizontal or vertical series of lines with various spacings and widths. While that will continue to be effective for model calibration and process tuning, it is not sufficient for metrology of masks intended for HVM at the leading edge if the critical parts of the masks use curvilinear ILT. This problem already exists today because complex OPC'ed masks input as Manhattan shapes manufacture as curvilinear shapes on mask. But becoming increasingly curvilinear both increases the need to address the issue, and also to improve the prospects of addressing it accurately.

As previously noted, asking for Manhattan shapes on mask, knowing that it will not manufacture that way creates issues. Especially for contacts or vias where even the main features contain Manhattan corners near each other (with "near"

defined by the blur radius of the mask process), understanding what mask shape is expected on the physical mask by the input Manhattan shapes requires simulation or, worse, very complex rule-based processing. By having the target mask shapes expressed as curvilinear shapes that are manufacturable on the mask everywhere, we can start to expect the physical mask to match the target mask everywhere, not just in 1D (straight) edges on both sides of the shape.

Since mask defect inspections today are performed with sophisticated image processing of large field of view pictures that capture the mask shapes approximately, deciphering defects may in part depend on expecting straight edges in some modes. Although very complex Manhattan geometries effectively turn into curvilinear shapes on the physical mask even today, specifying curvilinear shapes everywhere may require some enhancements to handle defect inspection of these masks more efficiently.

For all equipment, to the extent that today's data path may be relying on edge-based contour geometry format descriptions and internal algorithms that scale based on edge count or vertex count, addressing fully curvilinear masks for a large percentage of the masks being processed on that machine may become bound by the data path performance. Conversion of the data path into GPU-accelerated pixel-based processing will address the problem.

Figure 8. Typical photomask manufacturing flow. Mask ecosystem for curvilinear ILT is ready.

2.3 MRC (Mask Rule Checking) for curvy is simpler

Once the specification of the desired mask shape is curvilinear, we can expect the entire contour to be manufacturable in nominal contour with variation from the specified contour to be much more uniform around the contour. When specifying the mask with Manhattan shapes, a greater than 10nm variation is expected around the contour, causing most analyses that are not simulation based to check only "1D" edges that are sufficiently away from 90-degree (or 135-degree) corners. For contacts or vias, checking area and position is a sufficient substitute. ILT-specified shapes can be made to be manufacturable, so it is better to expect ILT to produce only fully manufacturable (all the way around the contour) mask shapes. In this situation, the task of mask rule checking (MRC) becomes significantly simpler. As published previously [6], conceptually, the MRC check will reduce to rolling a circle of a certain diameter (green in Figure 9) inside the mask shape, and rolling a circle of perhaps another different diameter (blue in Figure 9) outside the shape and make sure they can both roll completely around the contour such that the circle never exits the interior of the mask shape and never invades the exterior any other mask shape. In Figure 9, the red circles indicate violations.

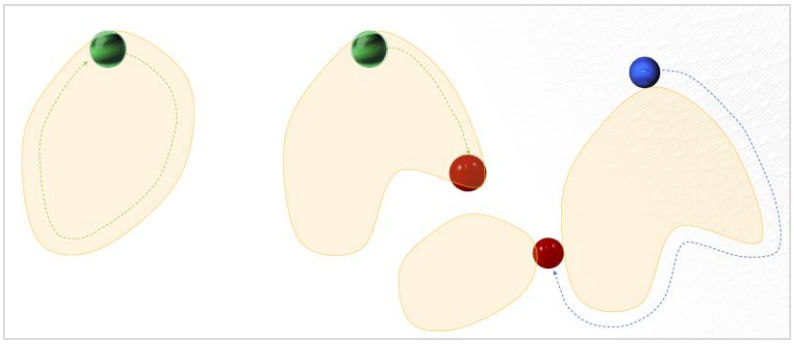


Figure 9. Conceptual examples of D2S curvilinear MRC (Mask Rule Check). As long as the minimum circle can slide entirely within and around the feature, it should meet the minimum width and space checks. Left shows that internal (width) checking passes. Middle shows internal (width) checking fails. Right shows that external (space) checking fails.

2.4 Curvy masks bias more faithfully translates to wafer

Curvilinear masks have another good feature. When mask bias is applied, the bias is more uniformly applied around the contour of a curvilinear shape.

A set of mask-wafer double-simulation experiments were performed for this study as depicted in Figure 10. By applying an x-nm bias on mask (where x is 1-nm through 10-nm) on both traditional Manhattan data and curvilinear data at a particular chosen location, the difference in effect on the resulting wafer at that location can be studied.

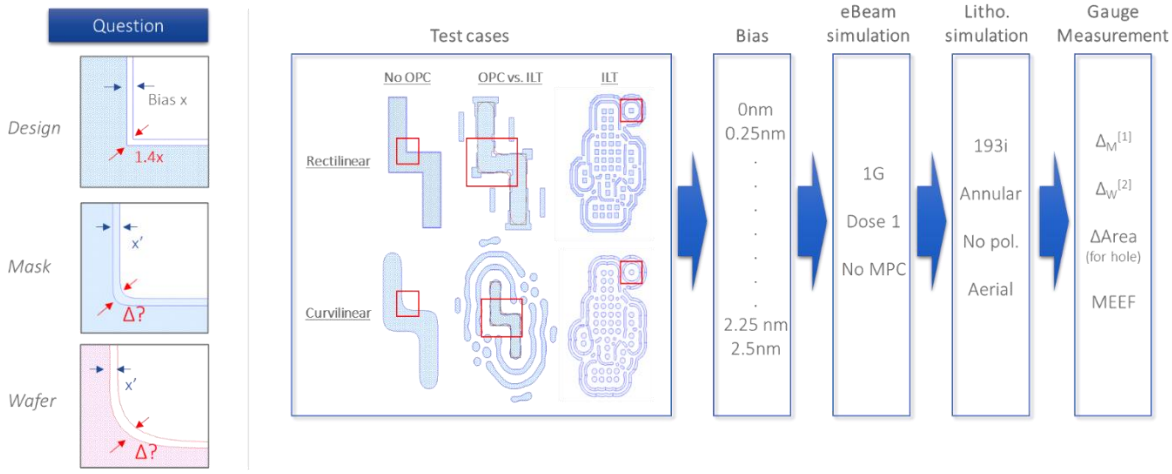


Figure 10. Left shows the fundamental question about the design bias impact on both mask and wafer at the pattern corner particularly. Right shows the experimental flow with three different inputs: (1) rectilinear and curvilinear patterns, (2) with and without OPC and (3) rectilinear ILT and curvilinear ILT.

Figure 11 depicts the first comparison study. A simple one-grid jog of a wire is depicted. Mask data specified by Manhattan shapes produce rectilinear shapes as shown in the top row. Of course, an x-nm bias of the contour would cause a 1.4x change at the 90 degree corners. Mask simulation with a simple single gaussian model suggests that this ~1.4x change translates to the expected mask shape at that corner. Wafer simulation with a simple annular light source in 193i lithography of that mask shape suggests that this results in ~1.2x change on wafer at that corner. It makes sense that particularly for 193i, the difference in mask at the corner dilutes to the surrounding area.

To contrast, an x-nm bias in the curvilinear CAD data translates to ~x-nm bias on mask and ~x-nm bias on wafer at those corners.

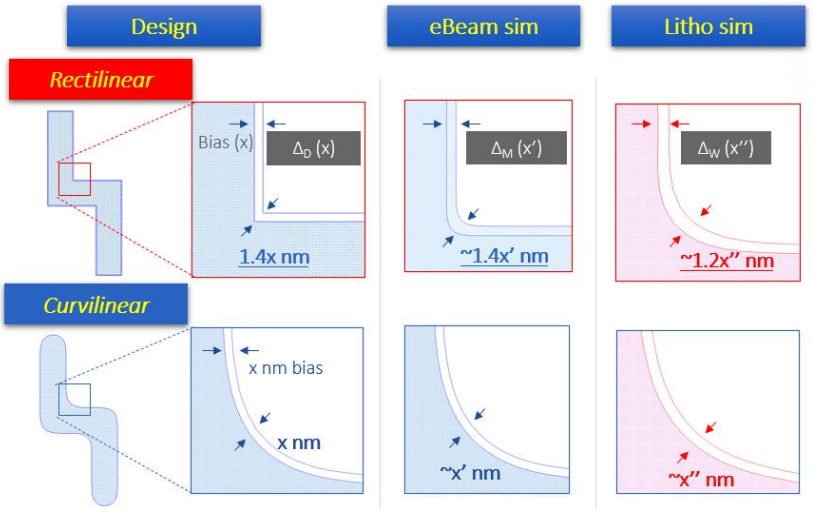
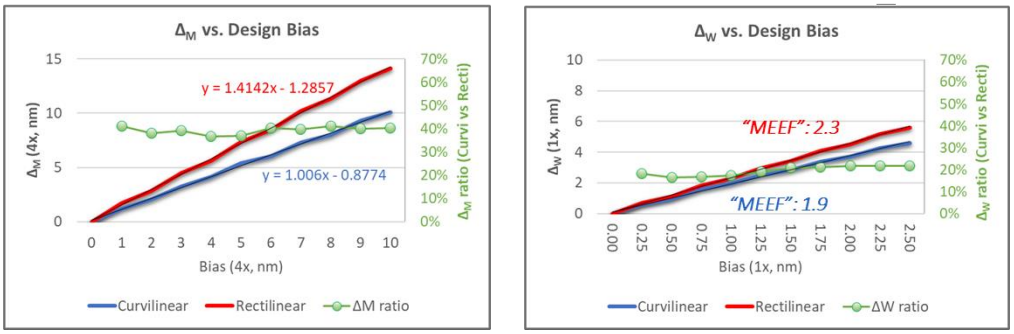


Figure 11. By design, x nm bias on pattern edge produces 1.4x Δ_D on the corner, where Δ_D is the edge displacement on the design. eBeam simulation shows the 1nm design bias also produces 1.4x Δ_M on the corner of this rectilinear test pattern, where Δ_M is the edge displacement by eBeam simulation on the mask. Even though Δ_M effect is typically diluted by the optical effects on the wafer. Lithography simulation result shows Δ_W on the pattern corner of the rectilinear pattern is still 20% bigger than that of the curvilinear pattern, where Δ_W is the edge displacement on the wafer by lithography simulation.

Figure 12 plots this across bias values from 1-nm to 10-nm. The green dots indicate the difference in the resulting bias on mask and wafer between rectilinear and curvilinear input designs for each bias amount. Across different bias amounts, we see that rectilinear shapes are biased 40% too much at that corner on mask and 20% too much at the corner on wafer. On the wafer plot, "MEEF" is also indicated, but we need to caution that this is not the standard meaning of mask-error enhancement factor (MEEF) as we are only measuring at the corner where the jog is. On a mask shape like these, the majority of the contour is not such corners, so the MEEF for the entire shape will be much better for both rectilinear and curvilinear versions.



(a) (b)

Figure 12. (a) This graph shows Δ_M on the rectilinear pattern is about 40% larger than that on the curvilinear pattern, (b) This graph shows that 1nm design bias produces ~20% larger Δ_W on the wafer for the rectilinear design compared to the curvilinear design.

But complex OPC shapes are full of 90-degree corners. Particularly at corners of the intended wafer shape, much decoration is present, such as the example depicted in Figure 13, which is a rule-based OPC version of the same pattern in Figure 11. The same 41% effect at each of the corners accumulate. In this case, Figure 14 indicates a 45-68% difference in the simulated wafer shapes at that corner for rectilinear input. As predicted, curvilinear ILT output biased by x-nm produces a nearly equivalent bias on the simulated mask, while producing a 2.4 "MEEF" (again, measuring the effect only at that corner, but for the curvilinear mask, the Figure 13 plot sees a visibly more uniform bias along the entire contour on the simulated wafer than for rule-based OPC). Please note that the y-axis of the wafer simulation results is in 1x (wafer) dimensions while for the mask simulation results, they are in 4x (mask) dimensions.

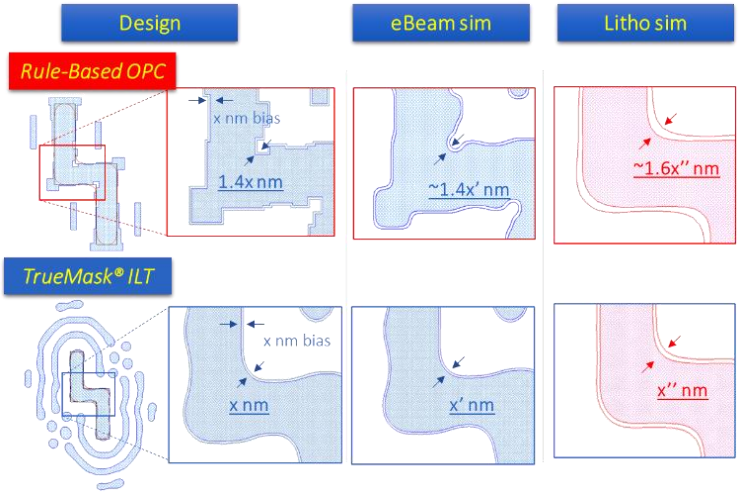
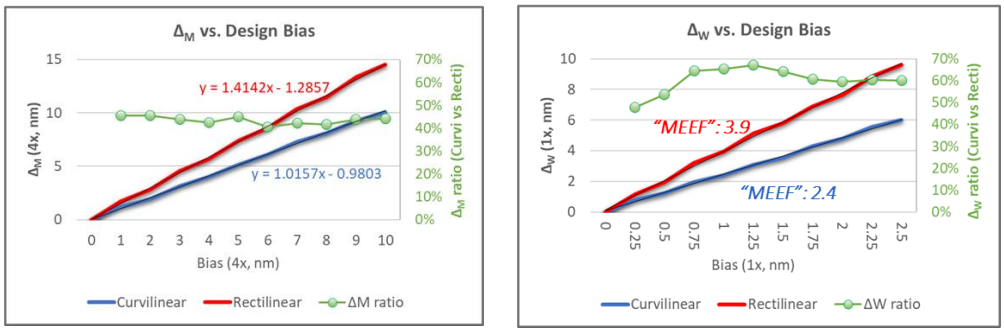


Figure 13. The x nm bias impacts are compared for the rule-based OPC pattern and the TrueMask® ILT (curvilinear) pattern in terms of Δ_M and Δ_W. Notice that the rule-based OPC has lots of jog assist patterns that produce 1.4x edge displacement on the corner.



(a) (b)

Figure 14. (a) The bias impact on the rule-based OPC pattern is still 1.4x Δ_M as described in Figure 8, where it measures the edge displacement on the corner. (b) This graph shows that rectilinear pattern has ~60% larger Δ_W than that of TrueMask ILT (Curvilinear ILT) as the excessive jogs produce 1.4x Δ_M compared to desired bias.

Figure 15, below, now compares rectilinear (Manhattanized) ILT as compared to curvilinear ILT results for a popular contact array pattern from [3]. On contacts, the four corners of a square mask pattern are sufficiently close to each other that the effect from the 41% difference in the corners distributes nearly equally to be a 10% bias on the wafer simulation result. As expected, drawing the target mask shape, as a circle, biases the mask much more uniformly and therefore the wafer result too. Note that this is measuring area of the entire circle. The 10nm bias is sufficiently small that the red and the blue plots appear to be linear, but they are actually quadratic.

Overall, such a simple operation as a mask bias produces more faithful effects across the mask shape contour when the target mask shapes are expressed as manufacturable curvilinear shapes than when expressed as rectilinear Manhattanized shapes.

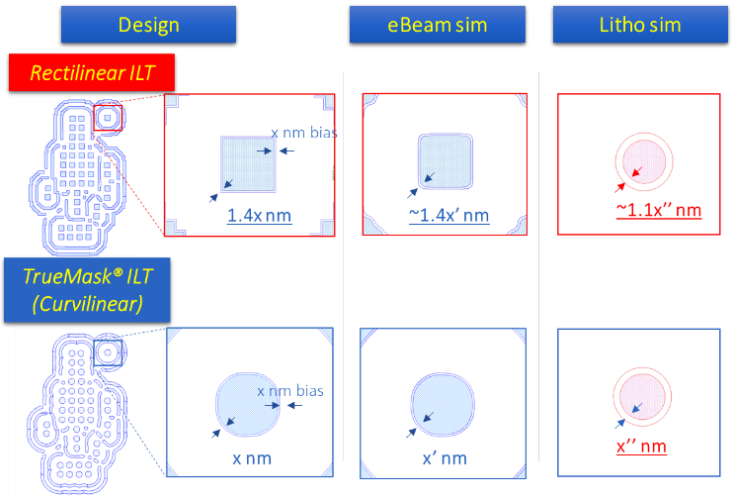
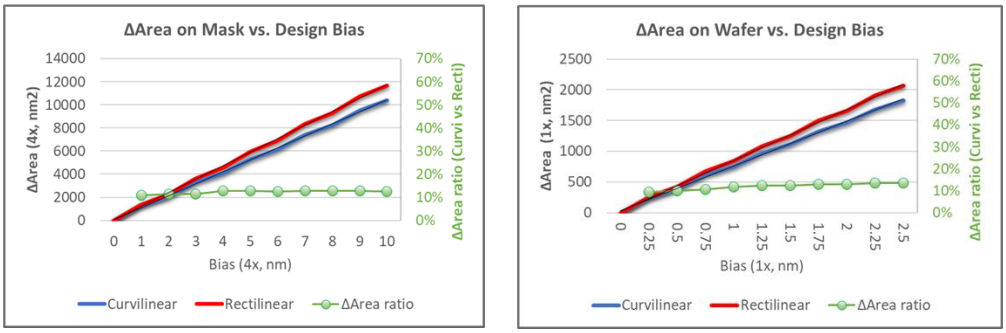


Figure 15. ILT typically provides the best process window compared to typical OPC. In this experiment, we investigated how the design bias impacts differently on rectilinear ILT and curvilinear ILT in terms of measured area on mask and wafer.

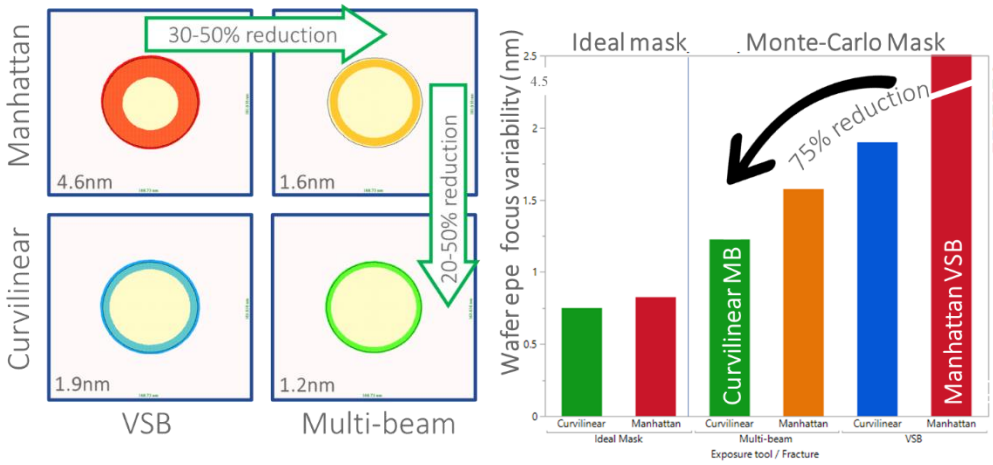


(a) (b)

Figure 16. (a) To compare the bias impact for the isolated contact hole in two different ILT designs, the area measurement was performed instead of 1D CD measurement because the bias impact on the contact hole is the energy flux on entire contact hole area. As expected, the rectilinear ILT pattern also has worse Δ_M compared to the curvilinear ILT, (b) Even though it is true that ILT in general promises the best process window, this graph shows that the curvilinear ILT still has 10% better Δ_M when the design same amount of design bias is applied on both rectilinear ILT and curvilinear ILT.

2.5 Curvy masks are more reliably manufacturable

Looking at the mask bias effect on wafer showed that curvilinear target mask shapes are better for nominal expected shapes. In addition, a previous study [20] demonstrated that uniformity on wafer is substantially better with curvilinear shapes than with Manhattanized rectilinear shapes.



(a) (b)

Figure 17. Study shows the mask variation can be reduced 75% by switching from writing Manhattan pattern on VSB mask writer to writing curvilinear mask pattern on multi-beam mask writer.

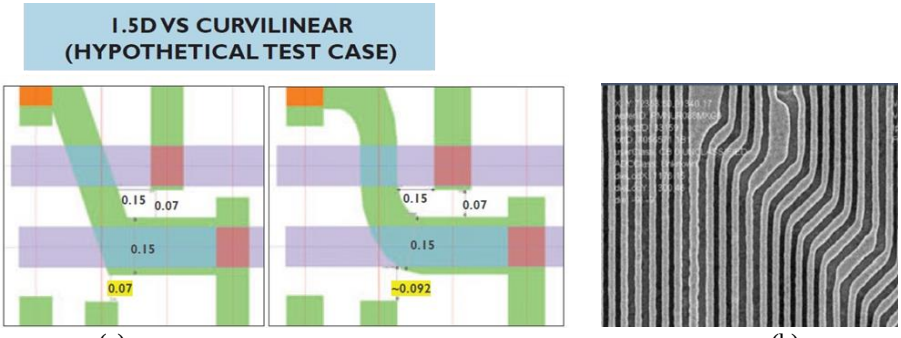
2.6 Curvy masks enable curvy designs

Semiconductor manufacturing at the leading edge, whether using 193i or EUV lithography is now anticipating a shift to curvilinear masks. Output by curvilinear ILT, 193i masks are written either with MWCO on VSB machines or multi-beam machines, and EUV masks are written with multi-beam machines. Curvilinear content may be a smaller percentage of hotspots, in which case more precision is required for curvilinear handling because it is being used exclusively for hotspots.

As we anticipate this exciting transition in mask making, an upstream effect of this change is being studied by some.

Because curvilinear ILT that outputs curvilinear mask shapes are pixel-based algorithms, there is no fundamental limit in semiconductor manufacturing then from accepting curvilinear wafer targets. Once a desired mask is manufactured, nothing in wafer production will prevent curvilinear wafer targets from being produced, so long as light sources and other wafer production effects are reasonably anticipated in producing appropriate target shapes. The target curvilinear shape that is intended to be manufacturable on wafer may not turn out to be exactly manufacturable as is, but it will likely be closer than expressing the desired wafer target than a Manhattan pattern.

Figure 18 (a) shows a picture from an Imec paper in 2019 [7]. They showed potential improvements in compacting cell designs, decreasing load, and decreasing interconnect delay. Figure 18 (b) shows a picture from a Micron presentation [21] showing that they already use manual manipulation to jog multi-bit busses using non-Manhattan shapes of varying angles. Manual manipulation is resource intensive, so it is a clear indication of the benefits being significant enough to be worth the trouble, at least for a memory maker.



(a) (b)

Figure 18. (a) An Imec paper showing curvy designs are also feasible with the reliable manufacturing of curvy masks, (b) an example wafer image from Micron with non-Manhattan design.

3. SUMMARY AND CONCLUSION

These are exciting times for the mask making community. In addition to EUV masks, and multi-beam mask writing, the prospects of curvilinear ILT and curvilinear masks anticipate a significant improvement in wafer process window that can be enjoyed by our customers. But even more than that, since curvilinear mask making is the only thing that is preventing curvilinear wafer target designs, a massive shift in semiconductor manufacturing is being enabled by the mask making community. One author came to this community after spending 20 years in the physical design community, so the fact that EDA tools today heavily embed the "Manhattan assumption" in functionality, quality of results of estimations, and in tool runtime is very well understood and appreciated. Still, the benefits in power reduction, skew reduction, delay reduction, area reduction, and yield enhancement as suggested by Imec are compelling. All of the positive effects of using manufacturable target shapes for masks cited in this paper are also true in using manufacturable target shapes for wafers. Manufacturable target shapes are more reliably manufacturable.

REFERENCES

[1] Klein, Christof, and Elmar Platzgummer. "MBMW-101: World's 1st high-throughput multi-beam mask writer." In Photomask Technology 2016, vol. 9985, p. 998505. International Society for Optics and Photonics, 2016.

[2] Hiroshi Matsumoto et al., "Multi-beam mask writer MBM-1000 and its application field." In Photomask Japan 2016: XXIII Symposium on Photomask and Next-Generation Lithography Mask Technology, vol. 9984, p. 998405. International Society for Optics and Photonics, 2016.

[3] Kim, Byung-Gook 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.

[4] Pang, Linyong Leo, Ezequiel Vidal Russell et al., "Enabling faster VSB writing of 193i curvilinear ILT masks that improve wafer process windows for advanced memory applications." In Photomask Technology 2020, vol. 11518, p. 115180W. International Society for Optics and Photonics, 2020.

[5] Aki Fujimura, Jan Willis, "eBeam Initiative Luminaries (formerly Perceptions) Survey Results." September 22, 2020.

[6] Ryan Pearman, Jeff Ungar et al., "Adopting curvilinear shapes for production ILT: challenges and opportunities.", In Photomask Technology 2019. Vol. 11148. International Society for Optics and Photonics, 2019.

[7] Amit Dounde et al., "A study of curvilinear routing in IN5 standard cells: challenges and opportunities," In Photomask Technology 2019, vol. 11148, p. 111481C. International Society for Optics and Photonics, 2019.

[8] Aki Fujimura, Jan Willis, "eBeam Initiative Panel," https://youtu.be/TGNC84Pc3FU, September 28, 2021.

[9] Aki Fujimura, Jan Willis, "eBeam Initiative Luminaries Survey Results." September 28, 2021.

[10] Aki Fujimura, Jan Willis, "eBeam Initiative Panel," https://youtu.be/k0-CIljMEig, February 23, 2021.

[11] L. Pang et al., "TrueMask ILT MWCO: full-chip curvilinear ILT in a day and full mask multi-beam and VSB writing in 12 hrs for 193i," Proc. SPIE 11327, 2020.

[12] Aki Fujimura et al., "Best depth of focus on 22-nm logic wafers with less shot count," Proc. SPIE 7748, 2010.

[13] Naoya Hayashi, Private conversation on December 2009, Tokyo, Japan

[14] NVIDIA A40 GPU datasheet, https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/a40/proviz-print-nvidia-a40-datasheet-us-nvidia-1469711-r8-web.pdf

[15] Hiroshi Matsumoto et al., "Multi-beam mask writer MBM-2000," Proc. SPIE 11855, Photomask Technology 2021.

[16] Jin Choi et al., "Curvilinear data format working group for MBMW era," Proc. SPIE 11610, Novel Patterning Technologies, 2021.

[17] Jin Choi et al., "Requirements of data technology for EUV photomask," Proc. SPIE 11148, Photomask Technology, 2019.

[18] Noriaki Nakayamada et al., "Writing curvilinear masks with MBM-2000," eBeam Initiative Presentation, https://www.youtube.com/watch?v=uQ3eOJ6daGo, 2021.

[19] Sayalee Gharat et al., "A study on various curvilinear data representations and their impact on mask manufacturing flow," Proc. SPIE 11613, Optical Microlithography XXXIV, 116130B, 2021.

[20] Ryan Pearman, et al, "How curvilinear mask patterning will enhance the EUV process window: a study using rigorous wafer+mask dual simulation", Proc. SPIE 11178, Photomask Japan 2019: XXVI Symposium on Photomask and Next-Generation Lithography Mask Technology, 1117809, 2019.

[21] Ezequiel Russell, “ILT and curvilinear designs for advanced memory nodes,” at eBeam Initiative, https://www.ebeam.org/docs/ilt-curvilinear-mask-designs-for-advanced-memory.pdf, 2020.