Laser micromachining of silicon microstructures

Investigating ablation removal rate as function of design pattern and machine parameters, to simplify the design-specific optimization process.

By:
Jens Hemmingsen s135101

Supervisors:
Jesper Hanberg (DTU Danchip)
Rune K. Christiansen (DTU Danchip)

Danmarks Tekniske Universitet
2015-08-05

Contents

1 Purpose

Laser ablation is a method used to remove material from a solid surface, by irradiating it with a laser beam. The material directly underneath the beam is heated and then evaporates. The process can be compared to old fashioned machine milling, in which a rotary tool is used to remove material from a solid surface, and indeed laser ablation is also known as laser micromachining.

There are a lot of different, sometimes interdependent, process parameters used in the milling process, which means that the ablation removal rate often depends on the specific design pattern used in a particular job.

The goal of this project is to reduce the need for design-specific process optimization or to find ways to simplify the process optimization.

2 Equipment

All processing is done on polished 100 mm silicon wafers with a thickness of 525 μm.

The laser is a microSTRUCT vario from the company 3D-Micromac AG. It is equipped with 2 different lasers, but for this project only the 50 W picosecond laser at 1064 nm with pulse repetition rate at 200 kHz is used. All processing is done at 100% laser power.

Either optical microscopes or SEM is used for inspection of the processed wafers.

3 Adjustable parameters

There are quite a few interesting parameters which will be targeted in this project. Some of these are highly interdependent, but can be adjusted separately, to reduce the influence of a given parameter on the remaining parameters during the testing. The effects of the various parameters can be divided into 2 main groups.

The effects mainly apparent in the x-y plane:

- Burn-in effects
- Shortening or lengthening of lines
- Corner sharpness
- Curved line-ends and line-starts
- Line thickness
- Downscaling burn-in

The effects mainly apparent in the z-plane:

- Redeposition
- Non-linearity of ablation depth with consecutive cutting

The parameters targeted in the project are:

Markspeed: This is the movement speed of the beam, while cutting.

Effects: Almost all affects.

Jump delay: The delay after each jump. If it is too short, a wobble is introduced. If it is too long, the scanning time is unnecessarily extended.

Effects: Curved line-starts.

Laser on delay: Controls when the beam is turned on, after reaching the begin-mark of a vector. If it is too short, the beam will be spending more time at the start point, producing a burn-in effect. If it is too long, the first part of the vector will not be drawn.

Effects: Burn-in and linelength variation.

Laser off delay: Similar to laser on delay, but in reversed order.

Effects: Burn-in and linelength variation.

Mark delay: The mark delay controls when the beam reaches the end of a vector. If the delay is too short, the jump command is executed before the entire length of the vector is reached, resulting in the end of the vector being turned towards the direction of the jump vector. If the delay is too long, the scanning time is unnecessarily extended.

Effects: Curved line-ends and linelength variation.

Polygon delay: Similar to mark delay, but "inside" vectors between polygon marks. If the delay is too short, the next polygon vector is executed, before the beam reaches the end of the current vector. The effect of this is that the polygon corners are rounded off. If the delay is too long, the beam is stopping between each vector, producing a burn-in effect.

Effects: Burn-in and corner sharpness.

4 Results

Most of the tests are made by creating a matrix with a given parameter on one axis and markspeed on the other axis. This produces an array of subtests for each parameter test, relative to the markspeed.

The markspeed values have been selected from the most commonly used values, while the parameter value range has been selected based on the default values provided from the laser manufacturer.

The resulting wafer is then inspected in an optical microscope or a SEM, and the matrix is color coded with either green, yellow or red, representing **no visual effect**, **some visual effect**, or **obvious visual effect**.

This creates a set of results that are very easy to apply to a given design, based on the sensitive or important aspects of that design.

4.1 Jump delay and mark delay results

These two delays both only negatively influence the writing, when their values are too low. A value that is set higher than needed will simply add unnecessary extra time to the process.

The optimal values for both delays are as short as possible, while not affecting the process negatively:

- For the jump delay a value of 300 μs seems to be the optimal default value.
- For the Mark delay a value of 220 μs seems to be the optimal default value.

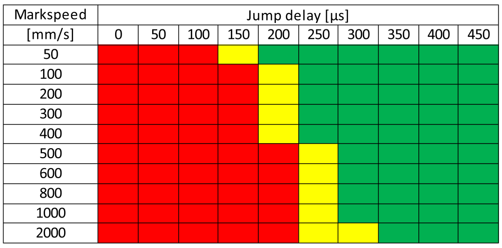

Figure 1: Jump delay results.

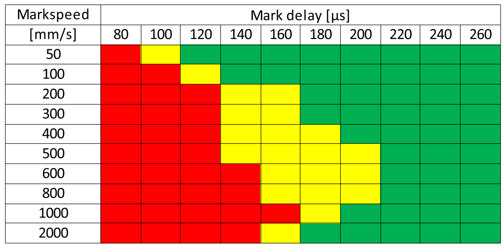

Figure 2: Mark delay results.

4.2 On delay results

The on delay has two different aspects to it; a burn-in effect when value is too low, and a linelength variation.

For measuring the linelength a vertical line of 1000 μm is used. To this length has to be added the diameter of the beam, but since the beam shape was found to be ellipsoid, we have to instead add the vertical axis, which turns out to be the minor axis of the beam.

Since the beam spotsize is affected by the markspeed, the reference spotsize for this test is chosen to be the spotsize achieved at 1000 mm/s, which is a very commonly used markspeed. At this markspeed the vertical axis of the beam is 65 μm.

This gives a total vertical length af 1065 μm. The coloring scheme used for the length matrix is based on the deviation from this value; **less than $\pm1\%$**, **between $\pm1\%$ and $\pm2.5\%$**, or **greater than $\pm2.5\%$**.

The suggested optimal delay value is selected to give the least amount of burn-in while retaining acceptable linelength accuracy over the most used range of markspeeds:

- For the on delay a value of 160 μs seems to be the optimal default value.

Regarding the funnel-shape of the yellow area in the results in figure (3): the yellow area widens as the markspeed is lowered. This simply reflects the fact that at low speeds, there is a burn-in effect along the entire track, which makes it very difficult to determine if the on delay burn-in is present. However it also makes it irrelevant to determine, because there is now a burn-in on the entire track. This effect is also present in the off delay and the polygon delay.

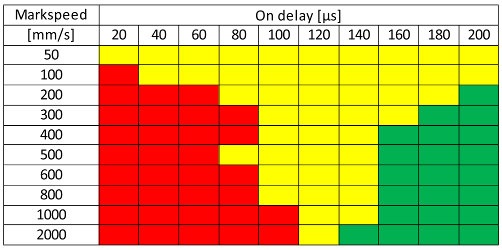

Figure 3: On delay burn-in results. Notice the funnel-shaped yellow area, due to track burn-in at low speed.

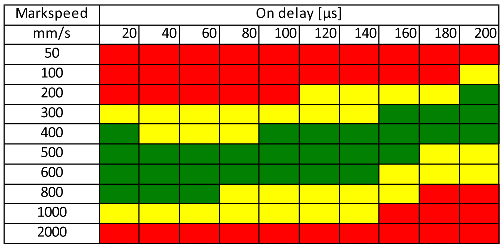

Figure 4: On delay linelength results.

4.3 Off delay results

The off delay effects are completely similar to the on delay, but in reversed order. The test procedure for the on delay test is reused for the off delay test.

Choosing an optimal value for the off delay is completely similar to the on delay. The suggested delay value is selected to give the least amount of burn-in while retaining acceptable linelength accuracy over the most used range of markspeeds:

- For the on delay a value of 240 μs seems to be the optimal default value.

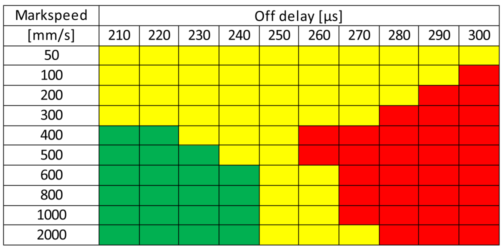

Figure 5: Off delay burn-in results. Notice the funnel-shaped yellow area, due to track burn-in at low speed.

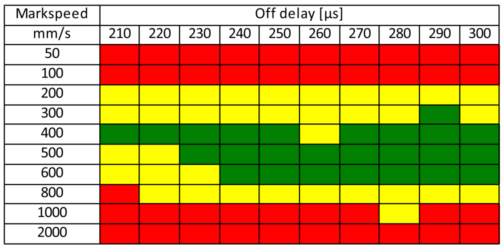

Figure 6: Off delay linelength results.

4.4 Polygon delay results

The polygon delay has two different aspects to it; a burn-in effect when the value is too low and corner rounding when the value is too high. Selecting a default value for the polygon delay is not possible, as the effects of the polygon delay are very dependant on the markspeed, and furthermore there is a tradeoff between selecting either against burn-in or selecting for corner sharpness; those two features cannot be good at the same time.

To make a decision about an optimal value of the polygon delay, it is important to first specify which of the two aspects are most important for a given design; no burn-in effects, or good corner sharpness. Next a specific markspeed has to be selected. This step involves a separate selection based on the markspeed effects.

A value can then be selected by evaluating the two polygon delay arrays, giving either:

- No burn-in, but some corner rounding.
- Some burn-in, but good corner sharpness.

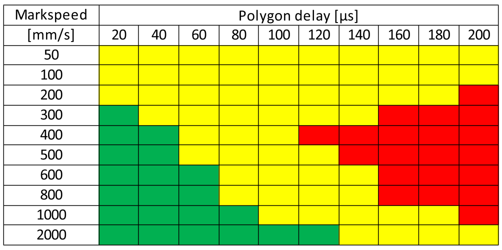

Figure 7: Polygon delay burn-in results. Notice the funnel-shaped yellow area, due to track burn-in at low speed.

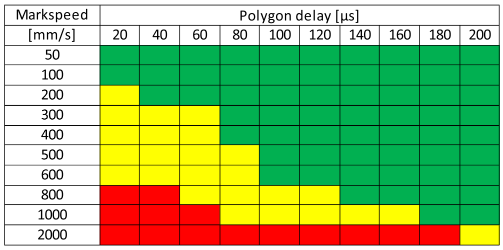

Figure 8: Polygon delay corner sharpness results.

4.5 Markspeed results

The markspeed influences all other effects, but has itself 2 independant effects; a burn-in effect along the line and linewidth variation.

Since the beam spotsize, and therefore the linewidth, is affected by the markspeed, but the beam shape is ellipsoid, a seperate test is made for horizontal and vertical lines.

The reference spotsize for this test is chosen as the spotsize achieved at 1000 mm/s, which is a very commonly used markspeed. At this markspeed the vertical width of the beam is 65 μm and the horizontal width is 88 μm.

The coloring scheme used for the width is based on the deviation from this value; **less than $\pm$10%**, **between $\pm$10% and $\pm$20%**, or **greater than $\pm$20%**.

For designs which are focused on producing a good xy-plane quality, like limiting burn-in etc., a faster speed is recommended. For cutting purposes a slower speed may sometimes be superior, although not necessarily by much, which will be discussed later in the ablation rate results:

- Good xy-plane quality: 400 mm/s or above.
- Cutting: doesn't matter.

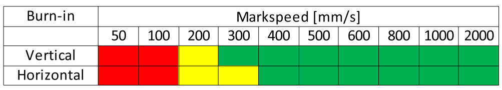

Figure 9: Markspeed burn-in results.

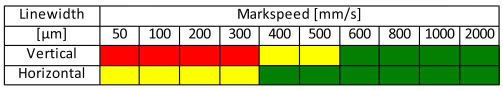

Figure 10: Markspeed linewidth results.

4.6 Polygon density results

When a given design is downscaled, it retains the number of polygons, which can create burn-in problems that would not be visible on a larger version of the structure. To quantify this in a meaningful way, a test was designed to reflect the amount of burn-in as a function of the polygon delay, the number of polygons per mm^2 and markspeed.

The results indicate, that very high density designs have to operate at 400 mm/s or higher markspeeds, and with as low a polygon delay as possible. This will of course influence the corner sharpness negatively, and a compromise between the two features has to be chosen by the user.

Furthermore tests show that a polygon density of about 100/mm^2 is the highest practical limit, structures above this limit will simply blend into each other, forming a milling area instead of individual structures.

- Markspeed: 400 mm/s or above.
- Polygon delay: 0 μs.
- Polygon density limit: $100/\text{mm}^2$.

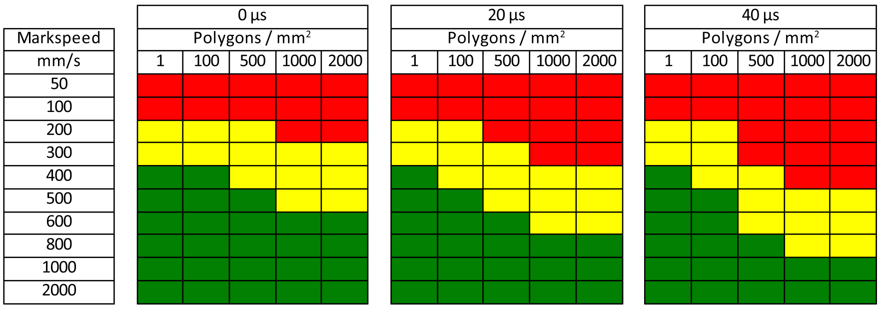

Figure 11: Polygon density results.

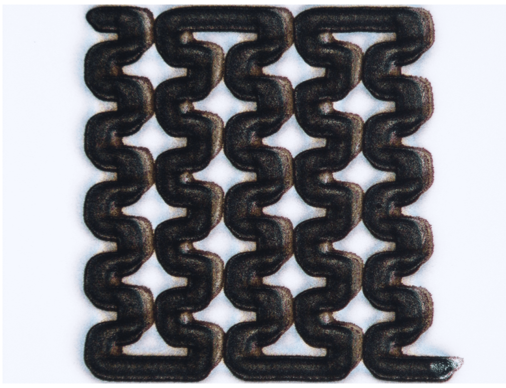

Figure 12: Polygon density of $100/\text{mm}^2$m.

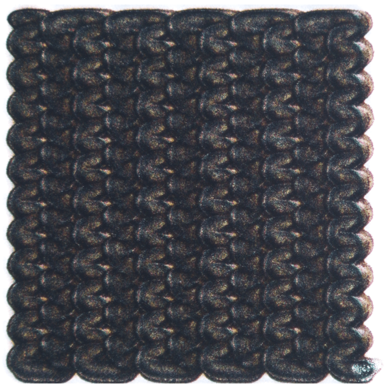

Figure 13: Polygon density of $500/\text{mm}^2$m.

4.7 Cutting - single line ablation rate

To test the ablation rate linearity relative to markspeed, a test was designed with 4 different markspeeds each with a number of iterations identical to the markspeed. Single line in this context refers to the number of repeated cuts done on one single line.

The 4 different setups are:

- Markspeed 1 x 1 iteration
- Markspeed 10 x 10 iterations
- Markspeed 100 x 100 iterations
- Markspeed 1000 x 1000 iterations

Up to about 300 μm the single line ablation rate is almost exactly linear. After about 300μ the ablation rate for single line cutting drops significantly due to redeposition; the cut is simply too narrow to effectively remove all the ablated material.

If the markspeed is matched with an identical number of iterations, the ablation rate becomes linear across speeds; ablation rate for markspeed 1 repeated only once, is the same as the ablation rate for markspeed 10 repeated 10 times, and so on.This means that there is in principle no difference in the total time it would take to do a complete wafer at any markspeed, giving that going faster simply requires more repetitions.

However many designs are made in such a way, that a lot of jumps between structures are involved, and this introduces extra processing time, which scales with both the size of the structures and the number of iterations. And since the number of iterations scales with markspeed, it may very well be that a slower markspeed will result in a shorter processing time. If no jumps are present in the design, which is possible for some designs, this dependency is removed and the total processing time becomes exactly independant of the markspeed.

The single line ablation rate, up to about 300 μm is approximately $\frac{300\,\mu\text{m}}{\text{markspeed}}$.

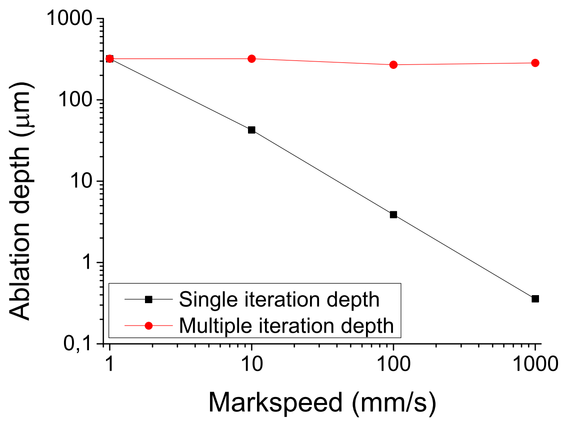

Figure 14: Ablation depth for single line cutting as function of markspeed and number of iterations. The *Single* graph shows the ablation depth of a single iteration at the given speed, while the *Multiple* graph shows the ablation rate for a markspeed with a matching number of iterations.

4.8 Cutting - parallel lines ablation rate

A superior method of cutting is to place multiple parallel lines with a slight overlap. To investigate the relationship between the markspeed and the ablation rate, a test was designed in which 10 overlapping lines are cut at various speeds and with a varying number of layers. For this test, each consecutive parallel line is shifted by 25 μm.

It turned out that markspeed 1 was completely useless for this, generating excessive amounts of foam-like redeposition taller than the cut should have been, as well as yielding no visible cut, but instead a groove of possibly amorphous, polycrystalline silicon, or redepostioned material, where the cut should have been.

The ablation rate for multi-line cutting was actually found to increase, as the ablation depth increases, meaning that the rate is lower at the first iteration and much higher on the last iteration. This is most likely due to heating, which should make the ablation process more efficient.
Unfortunately this makes it difficult to determine a specific value for the ablation rate, as it is dependant of the milling depth.

The following ablation rate is determined using a complete wafer thickness of 525 μm as the reference depth, which is similar to a real world cutting scenario.

- At markspeed 10 it takes 3 iterations to cut through an entire wafer.
- At markspeed 100 it takes 33 iterations to cut through an entire wafer.
- At markspeed 1000 it takes 500 iterations to cut through an entire wafer.

Or as a rate per iteration when cutting an entire wafer thickness:

- Markspeed 10 $\approx$ 175 μm per iteration.
- Markspeed 100 $\approx$ 16 μm per iteration.
- Markspeed 1000 $\approx$ 1 μm per iteration.

The conclusion to this is, that the most time-efficient speed is between 10-100 mm/s, with only little difference between the total processing time in this range. However the total processing time dependency of the number of jumps in the design also applies to this, meaning that designs with many jumps would benefit by selecting a lower markspeed, thus reducing the total number of jumps, and processing time.

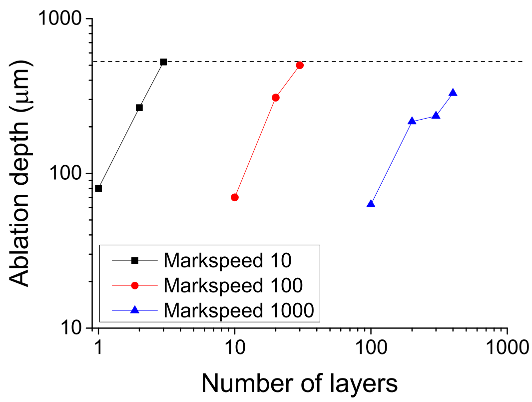

Figure 15: Ablation depth as function of markspeed and number of iterations. The dashed line represents an entire wafer width at 525 μm.

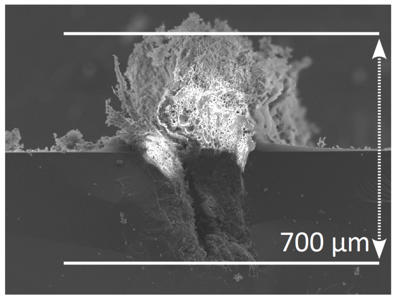

Figure 16: Markspeed 1 redeposition. Only present when running at 1 mm/s.

4.9 Cutting - minimum required number of parallel lines

Another aspect of the parallel lines is the ablation rate as a function of the number of parallel lines. This is primarily interesting in order to determine the minimum number of parallel lines needed for good multi-line cutting performance. The ablation rate with 2 parallel lines is identical to the ablation rate using single lines, and the ablation rate for cuts with 4 or more lines seem to be roughly the same. For cutting purposes this indicates that the minimum number of parallel lines are 4. At least when they are spaced 25 μ apart.

At a markspeed of 1000 mm/s the multiple line ablation rate as function of number of parallel lines is:

- 2 parallel lines $\approx$ 0.3 μm per iteration.
- 3 parallel lines $\approx$ 0.5 μm per iteration.
- 4+ parallel lines $\approx$ 1 μm per iteration.

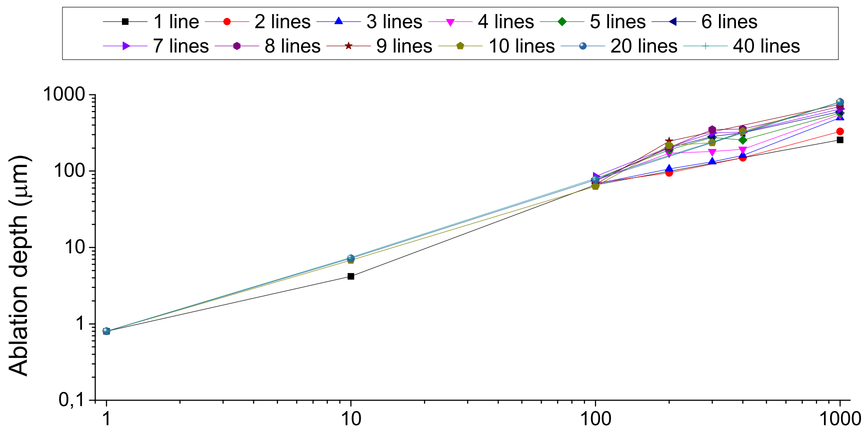

Figure 17: Ablation rate at 1000 mm/s as function of number of parallel lines.

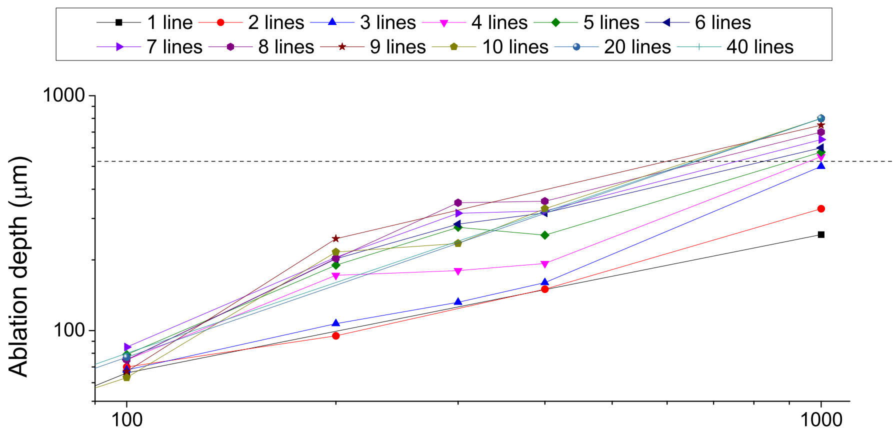

Figure 18: Ablation rate at 1000 mm/s as function of number of parallel lines. Zoom to 100 - 1000 layers. The dashed line represents an entire wafer width at 525 μm. All cuts with 4 or more parallel lines made it through an entire wafer at 1000 mm/s.

5 Summary of results

The following is a summary of the suggested optimal values found in the results section, added for ease of use.

Jump delay: The jump delay should be set to 300 μs.

Mark delay: The mark delay should be set to 220 μs.

On delay: The on delay should be set to 160 μs.

Off delay: The off delay should be set to 240 μs.

Polygon delay: The polygon delay should be set with respect to the most sensitive or important part of the design.

- No burn-in effects: Select a green combination of delay and markspeed from figure (19).
- Good corner sharpness: Select a green combination of delay and markspeed from figure (20).

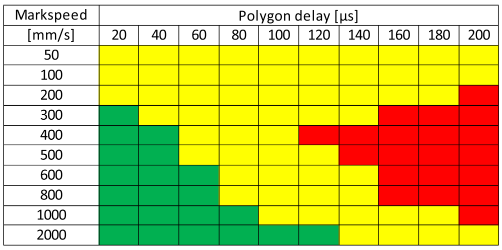

Figure 19: Select green for no burn-in.

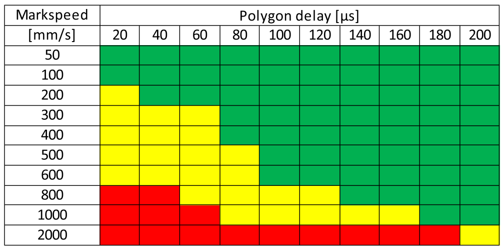

Figure 20: Select green for good corners.

Markspeed: The markspeed should be set with respect to the most sensitive or important part of the design.

- No burn-in effects: If burn-in along the lines are an issue, the markspeed should be set at 400 mm/s or higher.
- Linelength accuracy: If linelengt accuracy is an issue, the markspeed should be set at 600 mm/s or higher.

High polygon density: A density of $100/\text{mm}^2$ is the highest advised density, at these densities set the polygon delay to 0 μs and run at markspeeds of 400 mm/s or above.

Cutting with single lines: Do not use single line for cutting. Due to increased redeposition at increased depths, the ablation rate decreases after about 300 μm. Until that depth the rate is approximately $\frac{300\,\mu\text{m}}{\text{markspeed}}$.

Cutting with multiple lines: Use 4 parallel lines, spaced 25 μm apart, at 10-100 mm/s markspeed. Using this procedure the ablation rate is approximately $\frac{1600\,\mu\text{m}}{\text{markspeed}}$.

6 Further investigations

During the project some fairly interesting effects were discovered, which could be the subject of further investigations.

6.1 Cratering effect

When doing multi-line cutting with total track width of more than 250 μm, a cratering effect is observed along the edges of the track. The only reason the track needs to be this broad, is that below this limit the entire track floor is cratered, hence the effect is only differentiable, when part of the track is not cratered, which requires a trackwidth above 250 μm.

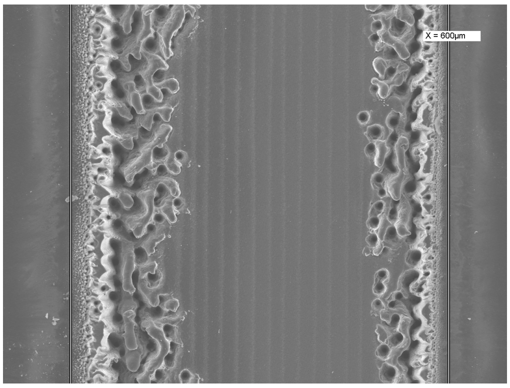

Figure 21: Cratering effect on 550 μm wide track.

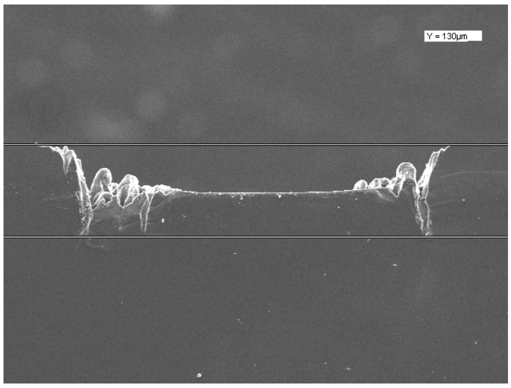

Figure 22: The crater depth is roughly twize the depth of the track floor.

During this project it was tested how inserting a pause between layers would influence this effect; it turned out to have no effect at all. Or at least no visible effect. The time between each layer was tested with no pause, 1 second, 5 seconds, and 10 seconds.

Future investigations could include:

- Varying the spacing between the parallel lines used to create the track. In this setup a spacing of 25 μm is used.
- Removing most of the outside material before creating track, leaving only a thin wall when track is completed. This could test the heat absortion, if any, from the surrounding material.

6.2 Redeposition material

When doing single line cutting, or multi-line cutting at low speeds, there is a substantial repedosition. The redeposited material on the surface of the wafer is most likely silicon oxide, or silicon nitride, but the redeposited material below the wafer surface may be different. Imaging with a SEM using backscatter electrons could determine the composition of the redeposited material.

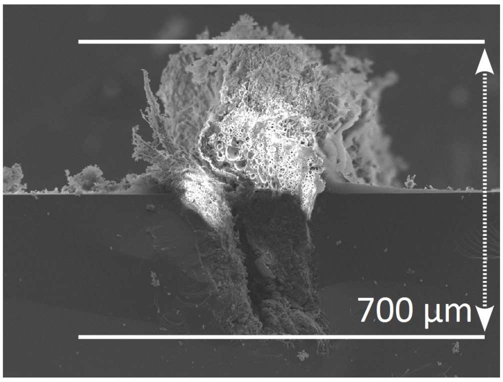

Figure 23: Redeopsition at very low speeds and multi-line track.

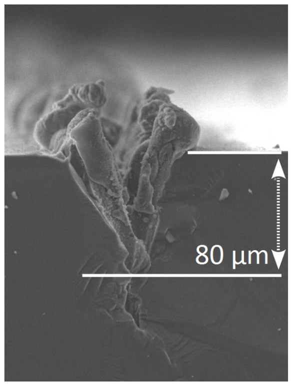

Figure 24: Redeposition of single line track.

6.3 Conchoidal fractures

When a wafer is cut manually, it is usually fractured aling the crystalline structure, resulting in very sharply defined fractures. But wafers that have been expose to laser radiation seems to have altered fracture properties; most fractures look conchoidal in nature, which could be the result of the heated monocrystalline silicon transforming into either polycrystalline silicon or amorphous silicon.

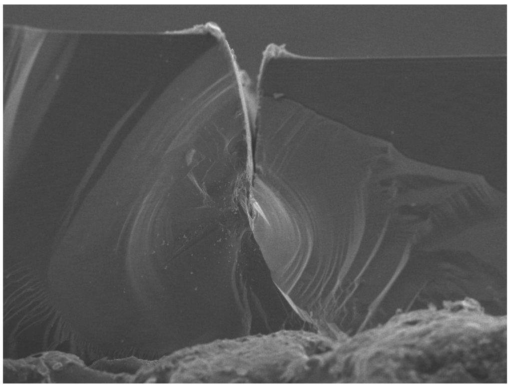

Figure 25: Conchoidal looking fracture. Note the area around the cut, where the color of the silicon is slightly lighter.

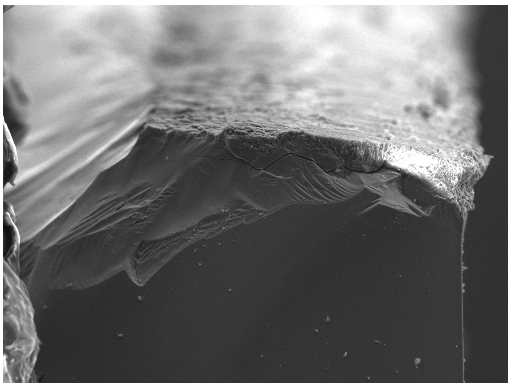

Figure 26: The fracture zones share many similarities with conchoidal fractures.

7 Conclusion

In general the design aspects of any project will have a critical influence on the process, e.g. the difference between writing a line, then *jumping* back and writing again starting from the same spot, compared to writing in one direction, and then simply writing again in the opposite direction. If writing is done at the same speed as jumping, the former obviously takes twice as long:

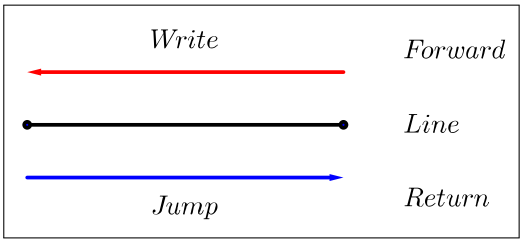

Figure 27: Writing, then jumping. Slower process.

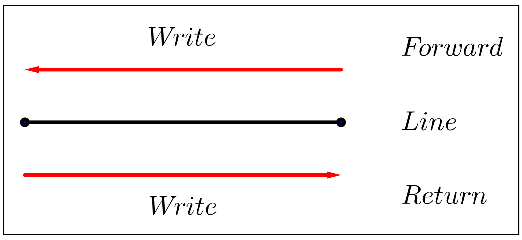

Figure 28: Writing, then reverse writing. Faster process.

This is just one of many aspects of design optimization that lies entirely on the users end; the easiest way to create ablation depth is by using the iterative method of simply repeating the pattern over and over, until the desired depth is reached, but this approach is on itself not very time efficient, as any jump related extra time will simply be multiplied with the iteration number. Problems of this kind is beyond the scope of this project, however, using the results gained from this project, many *non*-design related parameters can be easily defined.

The easiest parameters to set is definitely the jump, mark, on, and off delays. These 4 delays can simply be set at the suggested values, which should accommodate most, if not all, designs.

As such there is actually only 2 parameters for the user to optimize from; the polygon delay and the markspeed. The polygon delay and the markspeed are both parameters that the user has to select based on the design, but the selection process should be relatively easy using the guidelines described in this document. In the appendix is included an example of a user guide for selecting polygon delay and markspeed.

The only exception to this is with designs which have exceptionally many jumps (millions), where the size of the delays will begin to heavily influence the total process time. In such a case it would be necessary to focus on reducing the total process time, by first optimizing the design itself and, as a last resort, by reducing the 4 fixed delays.

A Appendix

A.1 Effects of adjustable parameters

The following images show the various effects of the adjustable parameters. All images are copied from the laser manual.

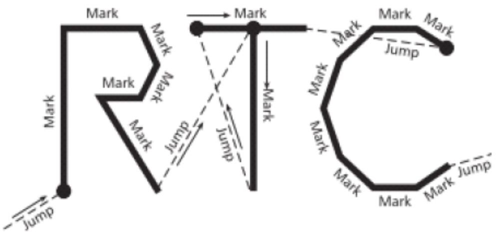

Figure 29: Laser on delay too short.

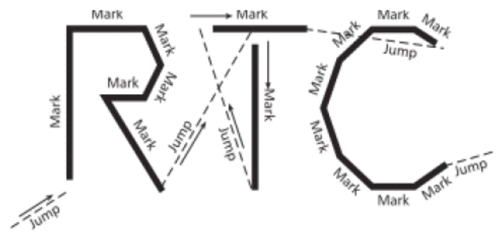

Figure 30: Laser on delay too long.

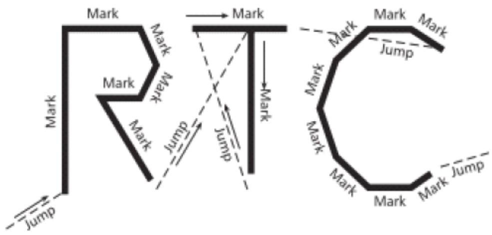

Figure 31: Laser off delay too short.

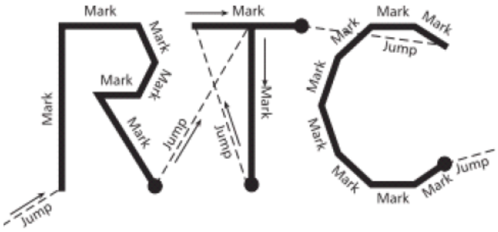

Figure 32: Laser off delay too long.

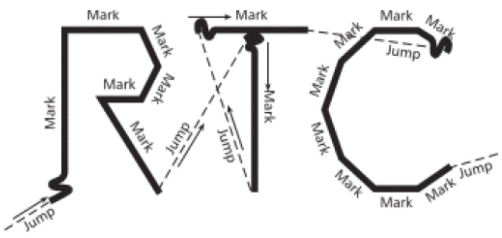

Figure 33: Jump delay too short.

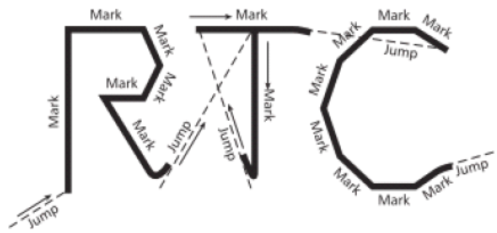

Figure 34: Mark delay too short.

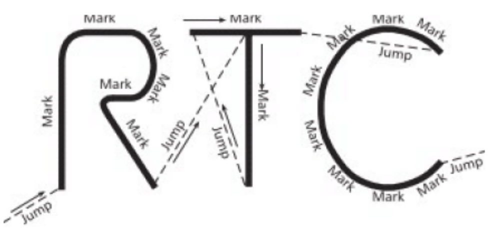

Figure 35: Polygon delay too short.

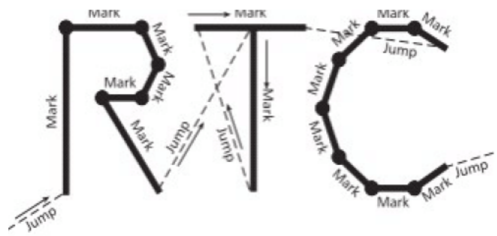

Figure 36: Polygon delay too long.

A.2 Real world images of effects

The following images are provided to give a more realistic image of the effects of the 5 delays.

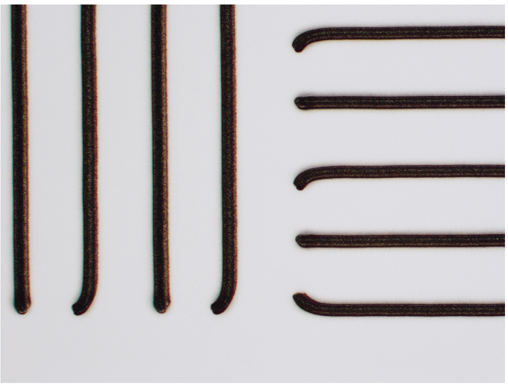

Figure 37: Jump delay too short.

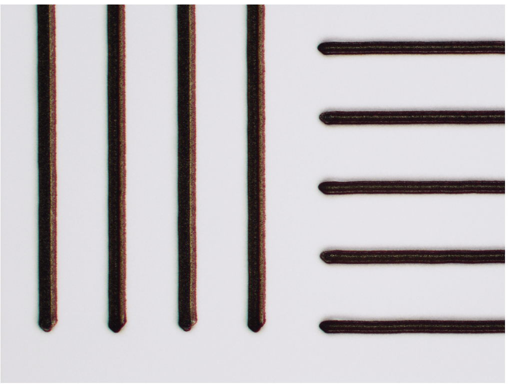

Figure 38: Jump delay OK.

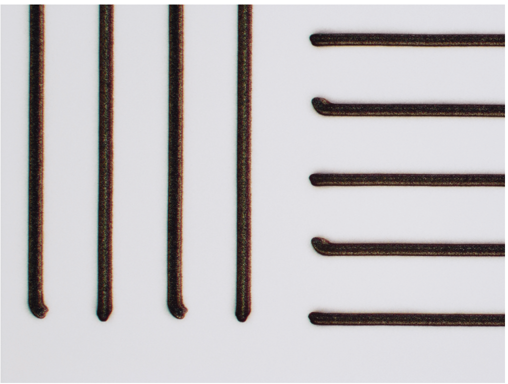

Figure 39: Mark delay too short.

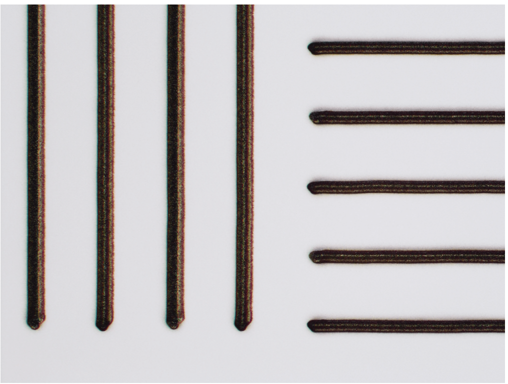

Figure 40: Mark delay OK.

Figure 41: On delay too short.

Figure 42: On delay OK.

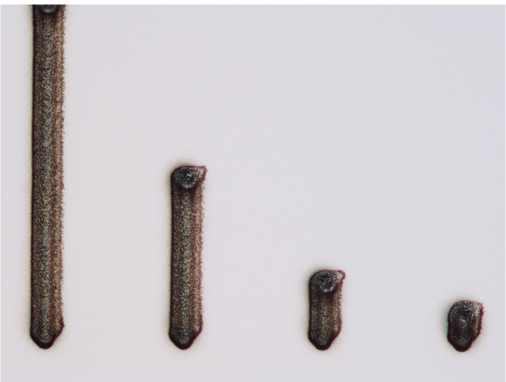

Figure 43: Off delay too short.

Figure 44: Off delay OK.

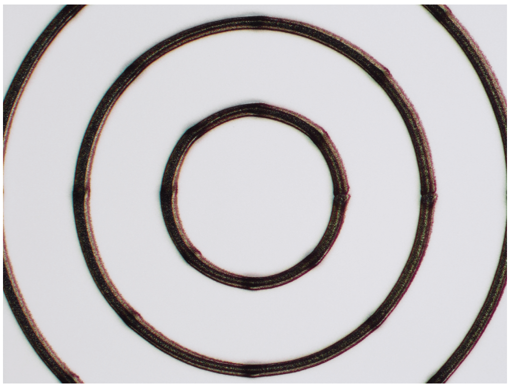

Figure 45: Polygon delay too long, resulting in burn-in.

Figure 46: Polygon delay OK.

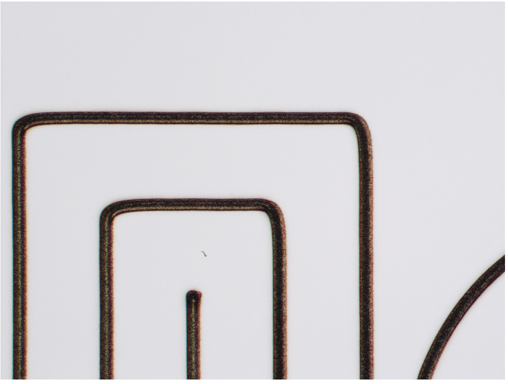

Figure 47: Polygon delay too short, resulting in rounded corners.

Figure 48: Polygon delay OK.

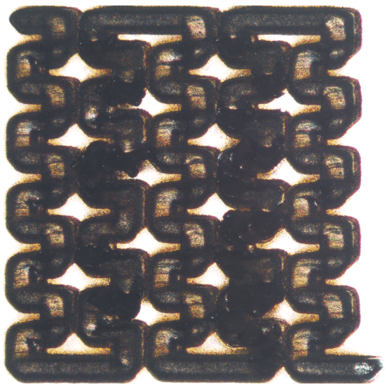

Figure 49: Polygon delay too long, resulting in burn-in at high polygon densities.

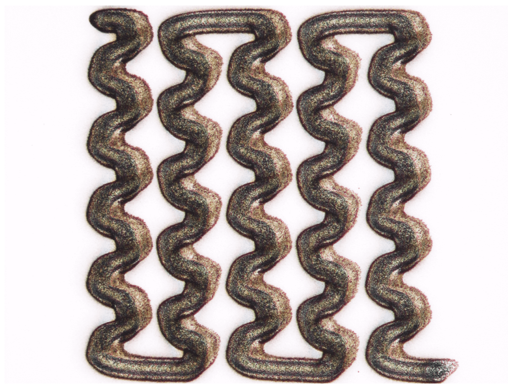

Figure 50: Polygon delay OK.

A.3 Example guide for choosing polygon delay and markspeed

The following is an attempt to illustrate how an actual guide to selecting polygon delay and markspeed could look.

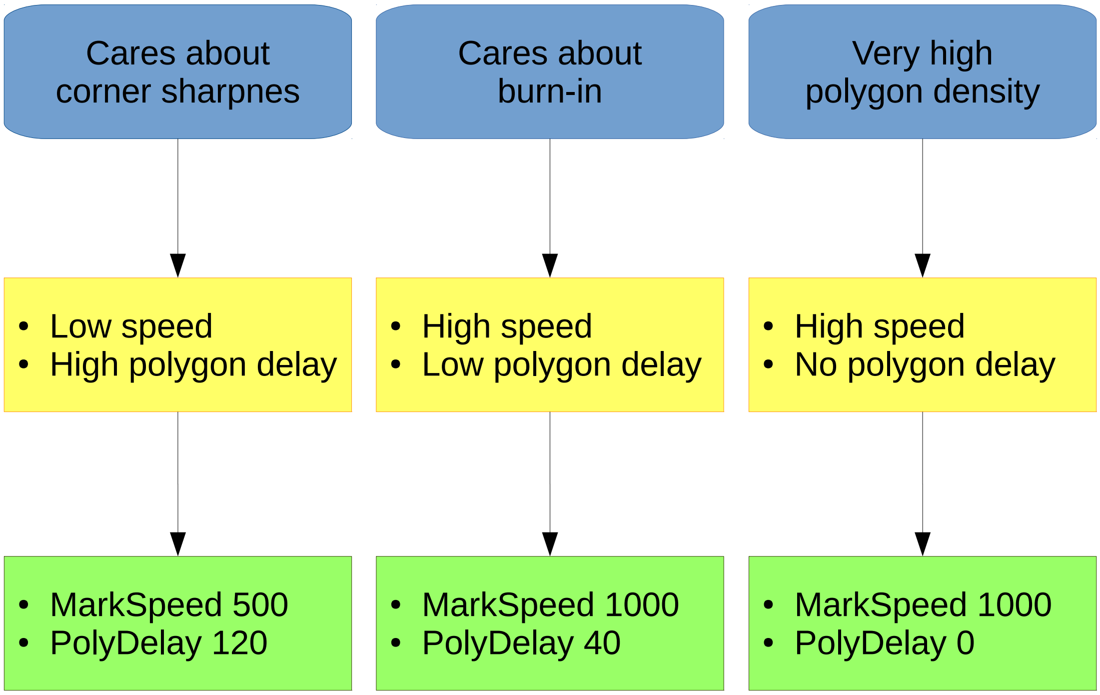

Figure 51: Crude example of how an easy-to-use guide could be created, based on the results from this project.