UCLouvain

École polytechnique de Louvain

Fabrication of a SWIR detector based on thin strained silicon

Design and fabrication

Author: **Loïc Lahaye**
Supervisor: **Jean-Pierre Raskin**
Readers: **Denis Flandre, Nicolas André, Nicolas Roisin**
Academic year 2021–2022
Master [120] : ingénieur civil électromécanicien

The writing of a thesis, even if done by one person in practice, is not possible without good support. During the year I have been surrounded by a group of very supportive people that I would like to thank here.

The first two persons that I would like to acknowledge are Nicolas André and Nicolas Roisin for their patience and dedication during this year. They assisted me during the design and fabrication of my structures by providing me with great advices, tips and knowledges. Above all, I would really like to insist on their patience. I ended up working with them in the same office almost every day and their support has never diminished.

About this office, I would also like to thank Eléonore Masarweh and Jérémy Trips who also share the office. Both of them have also been supportive throughout the year and everyone in this office made me feel accepted. It was a real pleasure to work around you all.

Having access to that office would however not have been possible without my promoter Pr. Jean-Pierre Raskin whom I'd like to thank for the trust he put in me. You gave me the opportunity to work on my own without ever being on my back but always being present if required. Also thank you for letting me "play" and experiment in such a large project. This thesis was not cheap but you trusted my work and let me follow through with it.

Furthermore, I would like to thank all the people that helped me a some point in the cleanroom. This again including the many hours spent there with Nicolas André but also with Eléonore Masarweh (help in the PECVD oxide deposition), Michaël Coulombier (wet thermal oxidation in the KOYO), Romain Tuyaerts (for his numerous advices and training with the BHF), Marie-Stéphane Colla (for the calls we had, the precious advices and for the access to SOI wafer), Alex Pip (for the many things I learned talking with him), Christian Renaux (for his patience and constant availability), Sebastien Faniel (for providing me in a rush with new wafers) and Najeh Zeidi for helping me with my cleanroom incident.

I would also like to thanks the people who spellchecked my thesis namely my father Lahaye Thierry, my brother Lahaye Nicolas and my always supportive best friend Berthus Laura.

Finally, a big shout out to the every single one of you who supported me emotionally during this time: My mother Patricia Van den Meersschaut (who added one parenthesis on page 50 to say she contributed), my other brothers and my friends (especially Permentiers Gaëtan who wrote his own thesis with me for countless hours during the night after a day at Winfab.

Contents

0 Introduction

This master thesis will tackle the topic of flexible electronics. Flexible electronics is an active modern research field that slowly finds its way in more and more applications in the world of power generation, biomedical equipment, sensing and enhancing of device performances. As we will see, many paradigms coexist under this umbrella term that is flexible electronics. Indeed, this terminology ranges from stretchable electronics while trying to avoid transferring constraints to the active material all the way to the field of strain engineering (the real topic of this thesis).

This work starts by a, rather long, state of the art. This part of the document starts by giving some background information on the history of the technology. It then proceeds to explain the different trends that exists in the flexible electronics world including strain engineering. Strain engineering could be defined as:

Modification of the electrical/optical properties of a semiconductor by application of a known level of strain either self induced or by external loading of the structure

Once this concept introduced, a theoretical analysis of the electronic properties of silicon as a semiconductor is conducted. This analysis leads to the description of the impact of strain on such properties. Once that understood, research articles and experimental results in that field are presented as a foundation for the rest of this work.

The state of the art now completed, the original part of this thesis can start. The goal was for us to design and build photodevices like PIN or NIN junctions on a flexible substrate in order to try to replicate the results of the article *"Breaking the absorption limit of silicon toward SWIR wavelength range"* by *Ajit K. Katiyar* and his team [1]. This article served as a foundation for the design of our own devices. What is presented in the second part of this work is the complete design of our membranes and photodetector. This includes, but is not limited to, the design of photolithographic masks, doping profile curves and implantation parameters, design of photodevices and mechanical simulations of the membrane deflection, stress and strain.

Once the design completed, the thesis moves to the cleanroom world for the practical part of the work. Indeed, the design was not left at the state of the drawing board but was really built at the Winfab facility of UCLouvain. In this section, the fundamentals of each technique is introduced and the designed process is presented as well as put in practice.

But without further ado let's jump into the thesis by beginning with the state of the art.

Part I
State of the art

Originally, electronic circuitry evolved as a set of components interconnected by loose metal wires. Those wires, also called jumpers or leads, where hand soldered effectively creating a *point-to-point* connection between two or more components. The components such as resistors, inductors, capacitors and so on were placed on a rigid frame, held by an insulating material hand soldered by a technician. This method meant that each circuit had to be handcrafted and that assembly time (hence amount of expensive human labor) lead to expensive final devices. In addition to that, increase in design complexity also lead to more and more complicated wiring schemes driving up the cost even further. Miniaturisation was also a challenge in that matter. [2]

Effort to start organising wiring as well as electrical components in smaller arrangements started in the early 20'th century. Famous names such as Sir Thomas Edison started experimenting with chemical deposition of conducting material on pieces of linen paper as early as 1904. Further in that century, between 1910 and 1930, multiple patents were filled across the world for various methods to deposit and etch metal insulating surfaces, hence paving the way to what we call today *traces* on printed circuit boards (PCB). Finally, in 1936, an engineer called Paul Eisler invented the modern design of the PCB as we know it today: A sandwich type design composed of an insulating rigid substrate covered with metallic copper traces and covered by a insulating resin. However, the concept only really started to gain popularity in the after war world of the 1950's. [2]

Since its invention the PCB became more and more popular, but the general concept did not change up to this day. Electrical components can be assembled on the circuit board either as through holes or as surface mounted components. Metallic traces then connect them together to form the final circuit. What has evolved however, is the amount of integration and the assembling/manufacturing cost.

Indeed, since the wiring had become traces on a hard piece of insulator, development of photolithography and etching techniques meant that smaller and smaller metallic traces could be drawn hence increasing the electrical component density on the board. The density of wiring being further increased by the development of, first, double sided PCBs and then, multiple layered PCBs.

Furthermore, parallel development in robotic allowed the introduction of automated PCB assembly using pick and place robots making the assembly cheaper and faster.

However, as said earlier in this section, the fundamental idea has not changed. Today's PCBs are most often made out of a material called FR-4 by the NEMA[1]. This material is a fire retardant composite built from a mix of fiberglass and epoxy making it both mechanically strong as well as electrically suitable for the task since it features a low water absorption level (<0.10%), good temperature resistance (up to 140°C) and high dielectric breakdown (>50 kV). ([3])

Since their invention this format has dominated electronic devices in the entire industry: rigid boards carrying components. And rigid it is since another property of FR-4 is it's mechanical stiffness with a high Young's modulus of 21 to 24 GPa.

"Young's modulus (E or Y) is the expression of a solid's stiffness or it's resistance to elastic deformation under load. It relates stress to strain along an axis" (citation: [4]). It can be computed as:

$$E = \frac{\sigma}{\epsilon} = \frac{(F/A)}{(\Delta L/L_0)} \tag{0.1}$$

With σ the mechanical stress in [Pa], ϵ the relative strain (relative deformation over the initial length), F the applied force, A the cross section and ΔL and L_0 the elongation and initial length in [m]. Young's modulus can hence be expressed in any unit of pressure and is more commonly given in [MPa] or [GPa].

Hence, as one can read from the formula, the higher the Young's modulus the stiffer and less elastic a material is. Another way to express it would be that the higher the Young's modulus, the higher the internal stresses in the material under an applied strain. For example, diamond has a modulus of 1050 to 1210 [GPA] whereas rubber has a modulus of 0.01 to 0.1 [GPa]. This explains why rubber doesn't shatter when stretched and why diamond cannot act as a bouncy ball. [4]

Note that this Young's modulus can also be computed as

$$E = 2G(1+\eta) = 3K(1-2\eta) \tag{0.2}$$

Where η is the Poisson ratio (the ratio between the expansion of the cross section and the length variation of a material under compression: $\eta = -\frac{d\epsilon_{trans}}{d\epsilon_{axial}}$), K is the bulk modulus or the resistance to compression of a volume and G is the shear modulus or the stiffness of a material relative to shear.

[1] **N**ational **E**lectrical **M**anufacturer**A**ssosiation: US trade association

To come back to the initial topic, but now with the understanding of what the Young's Modulus is, one can see what the point of this introduction is: standard PCBs, with their Young's modulus of 21 to 24 GPa are stiff and inelastic. Meaning that they are unable to bend or conform to curvature without breaking. This leads to a difficult integration of PCBs in biomedical applications such as implants or to their integration into moving or evolving structures. However, even if the idea is not new and can be traced back as early as 1947 in a publication by the US state department of commerce [5], development over the last 40 years has lead to more and more developments of *flexible circuitry*. [6]

In this first part of the thesis we are going to take a look at the different technologies grouped under the umbrella term of *Flexible electronic*. Indeed, we will see that different fields and very different applications can benefit from this new paradigm from flexible PCBs to flexible **U**ltra **T**hin **C**hips or UTC without forgetting the main topic of interest for this master thesis: strain engineering.

After this first part, the text will move to the analysis of the different applications of flexible device in optoelectronics. In this section a description of different designs of photodetectors will be given before looking at two examples of photodetectors built on top of a flexible substrate. Those two examples will be the occasion to showcase the two big categories of flexible devices: the one maximising the constraints in a material for strain engineering purposes and the one minimising them. The microfabrication process used in both cases will also be analysed.

1 Flexible electronic: a general overview

As said in the introduction to this section, flexible electronics is not a new concept at all. Indeed, flexible PCBs have been around since at least the 1950's. Flexibility being obtained by replacing the hard fiberglass of the circuit board by a polymer such as polyimide and embedding thin metallic traces into it. The *"board"* can be manufactured to feature the same capabilities as a conventional PCB plus a number of advantages:

* Copper traces
* Multiple layer and interconnect design
* Thickness of less than a millimeter
* High thermal stability
* Bendable
* Sculpted PCB

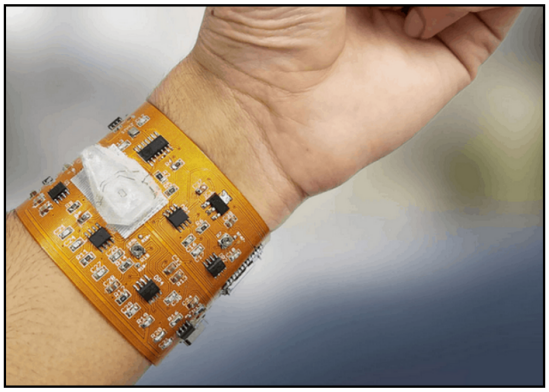


Figure 1: Flexible PCB by Swimbi [7]

Once the flexible PCB produced, conventional SMD (**S**urface **M**ounted **D**evice) components can be mounted onto it, effectively producing a circuit able to bend and conform to curvature as illustrated by Figure 1.

Those kinds of configurations have found their way in many industrial and commercial applications because of some advantages they feature over conventional rigid circuits: [8, 9]

Figure 2: Olympus stylus type camera without the cover (manufactured by Olympus). The flexible PCB can be seen wrapped around the object. This allows for a single PCB to be used instead of a combination of a mother and children set of boards. Image by Steve Jurvetson on Flickr. [10]

* Good mechanical properties: Because of their flexible nature these kinds of circuits are well suited to resist mechanical shocks, vibrations, tension or strain. Indeed, upon an applied strain or deformation they would simply bend and stress should not build up into the electrical components themselves (to a certain extent). A good and simple example would be the case of electrical ribbon/connectors often used as wires with many traces to connect a screen to a mother board in a phone or through the hinge of a laptop. Another example would be the case of an electronic bus and tram ticket used here in Belgium. These integrate a flexible antenna build from thin metallic traces embedded in a polymer, making them resistant to being dropped, bent or to be sat on.
* Another advantage is that it often reduces the BOM (**B**ill **O**f **M**aterial) required to build a complete device. Indeed, instead of manufacturing a main board and many sub-boards, a single circuit can be designed and wrapped around an object or bent

to fit into a desired space. This allows not only to drive the price down but also to reduce the risk of a failure due to a wrong assembly (less satellite boards means fewer cables, connections and soldering points) as well as making the design more compact. An example is shown in Figure 2.

* Also, polyimide (the most commonly used polymer for these kinds of applications) features a good thermal stability allowing the circuit to operate at temperatures ranging up to 400°C.

From the arguments given here above one can understand why such a technology has gained interest in many applications such as for electrical ribbon connectors, sensing, and the design of cheap or high temperature resistant electronics.

However, this technology still has one major limitation: the components mounted on it are not flexible themselves. SMD type resistors, capacitors or any other classically packaged ICs are intrinsically stiff and rigid. This limits the amount of strain that one can apply to the circuit if one wants to avoid breaking the mounted ICs/components or the solder joints.

Indeed, flexible PCBs are only the tip of the iceberg when it comes to flexible electronics. The complete discussion made up to this point only focused on the component carrier (the PCB or flex PCB) but we will now broaden the topic to flexibility at the chip and even transistor level. The next section will tackle the different trends that exist in the micro and nano-fabrication world to design flexible or bendable chips. A discussion will be made about the definition of flexibility itself. The used definition having a big influence on the device produced.

1.1 Flexible IC's origin

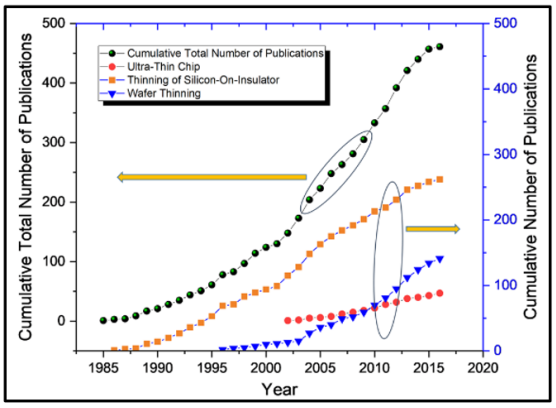


Figure 3: Number of publications mentioning or related to thin-Si. [11]

Flexible ICs, that will from now on be referred to using the term *"Flexible Electronics"*, have been a field of development for the last 40 years and is seen by many as a promising new paradigm for modern electronics. Research spanning from the development of electronic paper, fabric, flexible solar cells, flexible displays, memories and even flexible transistors have shown a growing interest from the scientific community as illustrated in Figure 3. As one can see from the plot, the number of publications have only been growing, with the first publications mainly focusing on the thinning of silicon substrate. Indeed, as silicon is thinned further and further, the materials starts becoming compliant and bendable (for thickness under 100 µm). The thinning of silicon so that the material in itself becomes flexible has given birth to the concept of UTC (**U**ltra **T**hin **C**hips) as we will see, but note for later that this is not the only way to manufacture bendable, and to an extent flexible, ICs. The use of polymer or island of rigid silicon being another route. [6, 11].

In terms of market (hence setting aside the scientific side of things), it is expected to grow from 30 billion USD in 2017 to more than 300 billion USD in 2028 according to [11]. This shows that the technology not only has an interest to researchers but also in the industry and for real life applications.

The reason for this growing interest mainly lies in the multiple possibilities offered by flexibility at "the silicon" or chip level. Here are some examples at the system level coming mainly from [6] and [11]:

* **Electronic displays:** Probably the most known and shiny example of flexible electronics. Indeed, having recently hit the market in consumer products, this technology started to be noticed even by the general public. However, OLED, the technology behind these displays, has been around since a least the 1980s and is built with a thin-film approach. Furthermore, OLED stands for **O**rganic **L**ight **E**mitting **D**iode meaning that the LED in itself is compliant. Apart from the use in mobile phones, the technology of flexible displays and even organic light emitting diodes could find its way in applications such as head-up displays integrated in the windshields of vehicles, airplanes or even glasses.

* **Sensing application:** Another example where flexible electronic could be of great interest is in the biomedical sector. Flexible sensing equipment that could be implanted in a patient's brain or any living tissue has been a hot topic in the recent year. The interest lies in the fact that not only strain, temperature, ... sensors could be implanted in a patient but also the complete accompanying electronics. This would allow for high speed data and signal processing as well as high speed wireless communication without the need of the patient carrying an external device. In addition, tight integration of electronics into living organism would allow easier energy harvesting for low power devices to operate for extended periods of time.

∗ Many other applications such as electronic skin, flexible solar array to be integrated in clothing, windows, ... application in the IoT industry and so on.

As one can see, the bottom line of all these examples is a tighter integration of electronics in everyday materials. Naturally, these examples are not representative of all possibilities the technology offers and are only examples of a system level integration of the technology. As we will see later, two great paradigms are in opposition on the topic of strain. Indeed, when designing flexible electronics one might want to either reduce the strain applied to the components/devices (transistors, diodes and so son) or, on the contrary, exploit it to do what is called *strain engineering* to change the properties of a material.

In terms of the history of the technology, as was said earlier, the hype over flexible electronics started to take off 40 years ago. However, the technology can be traced back to 1960 with the first developments of solar arrays built from thinned sub 100 µm silicon on a plastic layer, making it compliant. Further development in the field of solar arrays lead to the invention in the 1980s of what we now commonly referred to as *"Roll to Roll"* fabrication of solar panels. This technique is based on the concurrent developments of TFT made out of amorphous silicon on a polymeric substrate. The fabrication occurring on a roll of this polymer prior to it being sliced into the final desired sizes. [12]

The development of those TFTs bases on amorphous silicon, also noted a-Si, started concurrently to the thinning of silicon wafers in the early 1960s. TFTs, standing for **T**hin **F**ilm **T**ransistors, are a sub category of IGFETs (**I**nsulated **G**ate **F**ield **E**ffect **T**ransistors[2]). The definition *thin* coming from the fact that the thickness of the device is orders of magnitude smaller than its other dimensions.

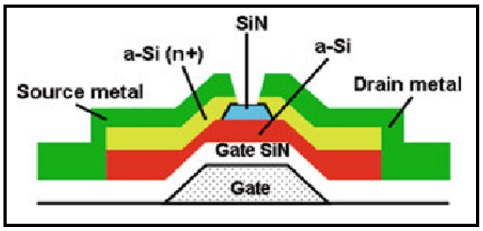


Figure 4: TFT build from PECVD deposited a-Si. Diagram from [13]

The use of a-Si is one of the key differences between a conventional MOSFET and a TFT. Indeed, where transistors where traditionally made out of monocrystalline silicon (often called single crystal silicon) that must be grown from a seed of known orientation then sliced from the ingot before polishing and thinning, the use of a-Si allows the fabrication of devices via PECVD. PECVD, standing for **P**lasma **E**nhanced **C**hemical **V**apor **D**eposition, indeed allowing for the deposition of thin silicon layer in the order of the manometer scale. Something that is hard to achieve via thinning of bulk mono-Si.

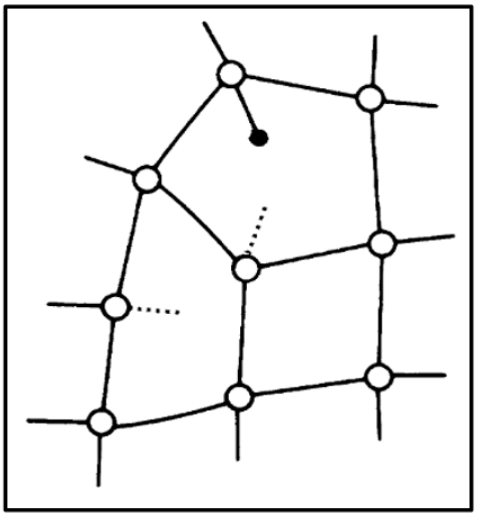


Figure 5: Structure of a-Si:H. Dangling bound in dashed lines, hydrogen atoms in dark circles and silicon atoms in white circles. Diagram from [13]

The difference between both materials lies in the fact that where mono-Si wafers are constituted of a single continuous piece of crystal, a-Si on the other hand is the non-crystalline form of silicon. This means that a-Si can be defined as short range order, but not *long range order* as in mono-Si. Silicon is known to make four covalent bounds with each adjacent silicon atom. In the case of mono-Si, a periodic and ordered repetition of atoms occurs, giving birth to a continuous lattice forming a single piece of silicon crystal, hence the name mono-Si. In a-Si the randomness of the deposition of silicon atoms (every silicon atoms has a random orientation) leads to the creation of a non-crystalline layers of unordered silicon atoms forming either covalent bounds or free dangling bounds. More information on the formation of silicon layers are given in subsubsection 1.3.1.

These dangling bounds are defined as *amphoteric*, meaning that they feature one unpaired electron able to either catch another electron and become negative or release this electron making the site positively charged. A solution to reduce the density of these dangling bounds is to use a hydrogenated silicon alloy Si:H (10% of hydrogen) where the hydrogen atom will fix to the dangling bound, effectively passivating it. The structure of a-Si:H can be seen on Figure 5. [13]

However, despite the imperfections in the crystal lattice, researchers in the 1970s have demonstrated that a passivated PECVD layer of a-Si:H had a low enough trap density to be doped and effectively be used as a semiconductor for the development of field effect transistors later giving birth to the TFT. [14]

The use of PECVD as a way to build the active silicon layer came with a series of benefits in terms of achievable device thickness as well as the emergence of new fabrication techniques such as the already mentioned *Roll to Roll* manufacturing on flexible polymers. This invention effectively increasing the production throughput as well as reducing

[2]Note that MOSFET (**M**etal **O**xide **F**ield **E**ffect **T**ransistors) are themselves a sub category of IGFET where the insulator **(I)** is the oxide

the cost. Uninterrupted PECVD deposition being achievable with, for example, a conveyor belt type CVD reactor[15]. These innovations ending up making possible the fabrication of flexible solar arrays as previously mentioned.

Naturally, all TFT aren't built on top of a flexible polymer. A common substrate being glass. The first demonstration of a flexible TFT however, traces back to 1968 with such a transistor built on paper. Since then, TFTs have been made on various materials including polyimide and the technology has reached maturity in fields such as optoelectronics by featuring field effect electron mobility up from 0.5 to 1 $cm^2V^{-1}s^{-1}$. [12, 16]

Another approach to thin silicon based bendable transistors is the use of organic materials as the semiconductor itself and not only the substrate. OTFT[3] (**O**rganic **T**hin **F**ilm **T**ransistors) were first introduced in 1964 with the demonstration of the field effect in an organic material. The use of polymer/organic material as the semiconductor comes with the benefits that the built TFT/device is by nature bendable and compliant. However, it comes with a series of challenges in terms of performances and manufacturing cost and scaling, as will be discussed in subsection 1.3. [17]

1.2 Flexible electronics: different definitions

When it comes to flexible electronics, many different paradigms coexist and share the umbrella definition of flexible. However, very different applications aiming at completely different goals exist. To highlight the different paradigms one must first consider the following question: *"What does flexible mean ?"*. This seemingly trivial question does indeed hide a key aspect of flexible electronics being that different applications might require different levels of flexibility and hence rely on different techniques. Here are some definitions/level of flexibility: [18, 11, 19]

- **Bendable:** These systems are built as released structure that are designed to flex with the level of strain >1%. Research in that field have mainly been around the development of intrinsically bendable materials such as plastic, polymer or thin chips. Hence the thinning of silicon into the sub 100 µm range, the PECVD built a-Si:H or organic TFTs, integration of chips in the neutral plane of a material and UTC all fall into that category. In these cases the device should be able to bend many times in different orientations and to conform to surfaces without breaking, but should also be able to remain functional for very small curvature radii. These device are also generally not stretchable. They are designed to accommodate the variations in strain between the outside and inside surfaces that occur upon bending. Indeed, when bending a structure the outside of the curve will stretch while the inside will compress by the bending strain $\epsilon = \frac{d}{2r}$ where d is the sheet thickness and r the curvature radius. Looking at this equation one can understand why thin chips are good candidates for bendable electronics: reducing the thickness allows for a smaller curvature radius at constant strain. The goal being to keep the maximum strain below the fracture limit of the considered material (which is a function of its thickness). Note that another approach can also be considered in the case of bendable electronics: maximising the strain applied to a material as we will see later (this being called strain engineering). Applications for bendable electronics could be for large flexible displays, roll to roll manufacturing or strain engineering, ...

- **Elastic/stretchable:** In this case the device/structure might be subject to very high and reversible levels of strain well above 1% of deformation. The electronics might be folded, bent, wrapped, stretched many many times hence requiring a completely different approach than the one used to achieve bendability. The thinning of a material or its integration into the neutral plane is here not sufficient and the generally adopted approach is to split the design into separated rigid islands interconnected using stretchable wires on a elastomer. The goal being to transfer all strain into deformation of the substrate and not transfer any stresses generated by the curvature into the island. This approach allows for tight radii of curvature. Another approach would be to go beyond the thin chip in the world of nanomembranes.

- **Permanently shaped:** This kind of flexibility is related to the two previous ones. The difference being that once conformed to a surface, the device should only experience mild changes in curvature during its lifespan hence not allowing its orientation to be changed too many times. Examples could be integration of electronics inside curved dashboards, windshields, windows, ... Another difference with the bendable approach is that if one wants to be able to conform the electronic device to any surface, the ability to handle small curvature radii might be required. In that case, as we will see later, the UTC approach is no longer an option since high levels of strain (»1%) might be encountered. A matrix/island approach with stretchable interconnects should be considered. More generally, the technique used for conformable electronics, be it bendable or stretchable, will depend on the required curvature radius.

These various requirements will call for new substrates being either thinned materials, polymers or metals. Figure 6 from [20] gives three examples of substrates for the previous definitions. The first one, thin glass, could be used in the

[3] the technology that OLED is based on

case of bendable electronics, the second one, plastic, for high strain electronics and the last one, metal, for its ability to keep its shape once formed and could hence serve for permanently shaped electronics. The table gives useful mechanical, electrical and chemical properties that will be used for future reference in this thesis.

Property	Unit	Glass (1737)	Plastics (PEN, PI)	Stainless steel (430)
Thickness	µm	100	100	100
Weight	g/m^2	250	120	800
Safe bending radius	cm	40	4	4
Roll-to-roll processable?	–	Unlikely	Likely	Yes
Visually transparent?	–	Yes	Some	No
Maximum process temperature	°C	600	180, 300	1,000
CTE	ppm/°C	4	16	10
Elastic modulus	GPa	70	5	200
Permeable to oxygen, water vapor		No	Yes	No
Coefficient of hydrolytic expansion	ppm/%RH	None	11, 11	None
Prebake required?	–	Maybe	Yes	No
Planarization required?	–	No	No	Yes
Buffer layer required? Why?	–	Maybe	yes: adhesion, chemical passivation	yes: electrical insulator, chemical passivation
Electrical conductivity	–	None	None	High
Thermal conductivity	W/m·°C	1	0.1–0.2	16
Plastic encapsulation to place electronics in neutral plane	Substrate thickness	5×	1×	8×
Deform after device fabrication	–	No	Yes	No

Figure 6: Properties of thinned glass, plastic and metal as substrates for bendable, flexible and permanently shaped electronics. Table coming from [20]

1.3 Choice of an active material for flexible electronics

This section is based on [11]. Additional sources will be explicitly mentioned otherwise

In addition to choosing a substrate that suits the specific application requirements, one must also tackle the problem of what type of semiconductor to use to obtain an active flexible device. The obvious choice would be to lean towards the use of *small-molecule* organic materials since they are by nature compliant and bendable.

The idea of using an organic type material to replace silicon into the building of active devices such as transistors is not new. The first demonstration of the field effect in such a materials indeed tracing back to 1964. The development of OFET since the late 1990s has even been able to produce pentacene-based active devices with a hole field effect mobility of 3.2 $cm^2V^{-1}s^{-1}$ giving them an effective ON current comparable to a more standard silicon based TFT.

Other materials such as graphene even feature an even greater carrier mobility of 24 000 $cm^2V^{-1}s^{-1}$, which is twice as much as mono-Si and 750 times more than a-Si as one can observe from Figure 6. Even if not considered to be an organic material, the atomic thickness of graphene makes it flexible and hence suitable for our considered application. Furthermore, looking at the mobility alone, one might think that graphene would outperform silicon in addition to featuring compliance, but this metric does not tell the entire story [17, 11].

Indeed, another thing to consider is the **T**echnology **R**eadiness **L**evel, or TRL, of a given platform. For instance, since silicon has dominated the industry for many years, it benefits of a high TRL compared to graphene which was discovered in 2004 and has not reached industrial applications yet[21]. Mass production of graphene remains a hot research topic and today's used graphene (in lab environment) comes from small scale CVD or rudimentary exfoliation [22].

Know how when it comes to micro and nanofabrication is essential to take into account and cannot be neglected. One way to asses this difference is to look at the achievable channel length when building a transistor out of a given material. This can be seen in Figure 7 under the *channel length* column. Here, one can see that the achievable channel length of transistors made out of mono-Si is currently of 14 nm compared to the 40 nm node of graphene and 2000 nm of pentacene-based device. Furthermore, still in Figure 7, one can see the difference in the achievable $\frac{I_{on}}{I_{off}}$ current ratio of silicon against graphene and pentacene.

Now that all these technological parameters have been presented one way to compare the different technological platforms available, we can look at what we call the *transition frequency* of transistors. To this end, one needs to take into consideration the parameters influencing the device performance, namely: carrier mobility, channel length, ON/OFF current ratio as well as all other parameters that are fixed. These other parameters being: the oxide capacitance, threshold voltage, channel width, ... From there one can express the transit frequency as[11]:

$$f_T \propto k\frac{\mu}{L^2} \qquad \text{that can be normalised as} \qquad f_{T_{\text{norm}}} \propto \frac{\mu}{L^2} \tag{1.1}$$

Where k is a constant depending on the device dimensions and parameters, μ is the carrier mobility and L is the channel length.

From here, and knowing that the transit frequency represents the frequency at which a transistor reaches the unity current gain, one can use it as a figure of merit to compare the performance of a given transistor with a channel made out of a different material. Again, Figure 7 has compiled examples from the literature giving us an overview of the performance of real world devices.

Here one can observe that mono-Si has a transition frequency >40 times higher than the one of graphene-based devices. Hence, from that metric one can see that mono-Si is a good candidate for the building of high performance flexible electronics.

Therefore, considering both the outstanding performance of mono-Si as well as its high TRL we can conclude that silicon is a good candidate to meet the demands of high performance flexible devices. Naturally, future development to increase the TRL of graphene or the performances of organic material (like pentacene) might reverse this trend, but this is unlikely to occur in the coming years [11]. Silicon benefiting of years of expertise indeed offering high reliability, good scalability and achievable device density as well as an extensive panel of well mastered micro and nanofabrication techniques available today.

All these considerations in mind, the next section will focus on a brief description of silicon in terms of material properties.

Material	Mobility (μ) [cm^2/V-s]	Channel length (L) [nm]	Normalized transit frequency (f_{Tnorm}) [GHz]	I_{on}/I_{off}
Monocrystalline Si	300–1200	14	4250	10^9
Amorphous Si	5–32	12,500	0.00115	10^5
III–V Semiconductors	400–12,000	75	165	10^4
MoS_2	700	300	42	10^8
WS_2	234	6000	3.8	10^8
Pentacene	1.5	2000	0.0114	10^2
CVD Graphene	24,000	40	100	10^2

Figure 7: Comparison of the performance of various devices based on different semiconductors. Table assembled by [11]

1.3.1 Silicon

This section is based on [23, 24, 25]. Additional sources will be explicitly mentioned otherwise

Since silicon is indicated as used in flexible electronics let's take a look at some of its properties before moving on. Silicon is what we call an *anisotropic* material. This means that it features different physical properties (be them electrical, mechanical, chemical or other) when observed/measured along different directions. In the case of silicon this is due to its crystallographic orientation.

Indeed, a mono-Si bulk wafer for example, is made out of silicon atoms that form four covalent bounds with adjacent atoms in a periodic long range pattern. A representation of this particular crystallographic arrangement can be seen on Figure 9. A common description of the cubic silicon crystal orientation is given as follows:

- **[...]:** Represents a given specific crystal orientation. For example [137] represents a vector made out of the 1, 3 and 7 component in an orthonormal reference frame or:$\mathbf{r} = 1\mathbf{e_x} + 3\mathbf{e_y} + 7\mathbf{e_z}$.
- **<...>:** Represents of a set of an equivalent crystal orientation. For instance, by symmetry one can define the group <100> representing the vectors [100],[010] and [001] that are equivalent from a crystallographic stand point

* **(...):** Represents a given crystal plane. Since the silicon atoms are all repeated in a periodic pattern into the crystal lattice, all atoms have the same specific orientation hence defining different crystal planes. For example, the (100) plane represents the crystal plane perpendicular to the normal vector [100].
* **{...}:** As for the <...> notation, this is an abbreviation for equivalent crystal planes. For instance {100} represents a group composed of the (100) and (-100) planes.

This particularity of silicon will have an impact on its physical properties and more notably its chemicals and mechanical ones.

Impact on the mechanical properties

Being an anisotropic material, the already given 1D definition of the Young's modulus can no longer be used because the value of the modulus will be a function of the orientation. To deal with this complication one must now consider the tensor definition of the Young's modulus given by:

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl} \tag{1.2}$$

This expression, where C is the general definition of the stiffness, is for a purely anisotropic material. However, luckily for us, the crystallographic arrangement of silicon is cubic making it symmetrical and allowing us to reduce the fourth rank C tensor made out of 81 independent terms to a third rank tensor as given here below, simplifying further calculations:

$$\begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{12} & 0 & 0 & 0\\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0\\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0\\ 0 & 0 & 0 & c_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & c_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & c_{44}\end{bmatrix} \begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \epsilon_3\\ \epsilon_4\\ \epsilon_5\\ \epsilon_6\end{bmatrix} \tag{1.3}$$

This reduction of the tensor being possible thanks to the *orthotropic* property of silicon, which is smart way to say that the cubic structure of silicon features two orthogonal symmetry planes (wood is another example).

Developing the equations (which is out of the scope of this thesis but can be found in [25]) one can use the *orthotropic* set of equations to find values for the Young's modulus of a (100) silicon wafer along the principal directions of the group <100> (values from [25]):

$$\begin{cases} E_x = E_y = 169 \text{ GPa} \\ E_z = 130 \text{ GPa} \end{cases} \quad \text{and} \quad \begin{cases} \nu_{yz} = 0.36 \\ \nu_{zx} = 0.28 \\ \nu_{xy} = 0.064 \end{cases} \quad \text{and} \quad \begin{cases} G_{yz} = G_{zx} = 79.6 \text{ GPa} \\ G_{xy} = 50.9 \text{ GPa} \end{cases} \tag{1.4}$$

where E is the Young's modulus, *nu* is the Poisson coefficient and G is the shear modulus. Additionally, the average Young's modulus, Poisson ratio and shear modulus for polycrystalline silicon (or polysilicon) are:

$$E = 160 \text{ GPa} \quad \text{and} \quad \nu = 0.22 \quad \text{and} \quad G = 65 \text{ GPa} \tag{1.5}$$

Also note that it has been empirically demonstrated that the Young's Modulus of silicon changes passed a given thickness h_b according to [26]. The modulus starting to change from thicknesses below 80 nm following:

$$E = 54.872 * h_b^{0.226} \tag{1.6}$$

which is a curve fitted onto experimental measurements.

Chemical properties and microfabrication considerations

Other critical things to consider linked to the anisotropy of silicon are the chemical properties. More precisely, as illustrated in Figure 9, the fact that mono-Si features different crystallographic planes will lead to non-negligible variations in etch rate. Indeed, when looking at the surface of a crystallographic plane, the orientation of the silicon atoms constituting it will determine the amount of *dangling bounds* exposed to the etchant in the case of wet etching. For instance, when submerged, a (100) oriented wafer will feature atoms connected to the lattice by two *back bounds* and will expose two *dangling bounds* to the etchant. However, a (111) orientated wafer will present surfacic atoms with three **back bounds** and only one *dangling bound* meaning that it is better attached to the underlying lattice and hence requiring more energy for this plane to be dissolved in a base compared to the (100) plane.

This property of silicon expresses itself by a big difference in etch rate between different crystal planes, allowing for what is called *Anisotropic wet etching*. For instance when considering the wet etching of silicon using the base KOH, the following reaction occurs:

$$\mathrm{Si} + 4\mathrm{OH}^- \longrightarrow \mathrm{Si(OH)}_4 + 4\mathrm{e}^-$$

However, due to the difference in crystallographic orientation, an etch selectivity of 1:600:400 can be observed between the (111):(110) and (100) crystal planes for an identical etchant of 20% KOH, 16% propanol and 64% of water by weight at 80°C [24].

This property allows for etching of self-limiting shapes such as the one illustrated on Figure 8. Once exposed to the etchant, the more easily etched planes will be etched first, the etching "stopping" when the etchant meets a wall with high selectivity. This gives birth to predictable shapes depending of the wafer type. For the two most commonly used wafer types in micro-processing these are an inverted pyramid of self limiting depth of:

$$d = \frac{W_{mask}}{\sqrt{2}} \tag{1.7}$$

for a (100), where W_{mask} is the dimension of a square mask. And in the case of a (110) wafer a self limiting well of depth:

$$D = \frac{(a + \sqrt{2}b)}{2\sqrt{6}} \tag{1.8}$$

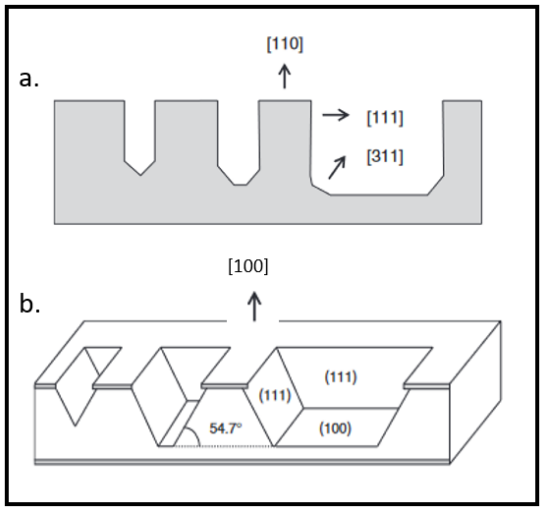


Figure 8: Self-limiting wet etching of mono-Si. **(a.)** Wet etching of a (110) type wafer and **(b.)** Wet etching of a (100) type wafer. Illustration from[23]

where a and b are the dimensions of the mask opening. These 3D shapes originating directly from the crystallographic structure of silicon and featuring, as already said, known parameters. For example, in the case of the anisotropic wet etch of a (100) silicon wafer, the walls of the pyramid are slanted by a well known 54.7° compared to the surface of the wafer. This angle being the one between the quickly etched (100) plane and the (111) etch limited plane.

This anisotropy can however be a constraint as well as a benefit since it does not allow to create any desired shape as illustrated in Figure 9. Indeed, whatever the shape drawn on the mask, the obtained geometry will always tend to align itself with a naturally self-limiting one allowed by the crystal plane. For example, drawing a circular opening in a mask for etching a (100) wafer will result in undercutting and, *in fine* the production of the classical already introduced inverted pyramid.

Isotropic etching of silicon is however not impossible at all and techniques such as dry etching can be used to obtain any desired geometry.

Wafer nomenclature

Crystal orientation being such a big deal at the mechanical and chemical level, an easy way to differentiate between the various types of wafers had to be devised. To easily discriminate between the wafers used in micro and nanofabrication labs, the wafers feature one or two flats giving information on both the crystallographic orientation and intrinsic doping level of the material.

The longer flat, or *primary flat*, is set to be aligned with the [110] direction and the presence/absence and position of the shorter or *secondary* flat allows to discriminate between p- and n-type (100) or (111) plane silicon wafer. The four combination are illustrated on Figure 9.

Different types of silicon

Finally, let's note that different sorts of silicon layers can be used in the fabrication of ICs, namely:

* **Monocrystalline silicon:** also noted mono-Si, this type of silicon is the one with the most interesting characteristics when it comes to building high performance ICs. In this case, as was already discussed, the silicon layer is one continuous piece of crystal. This uninterrupted lattice allows for a high carrier mobility compared to polysilicon or amorphous silicon as can be seen in Figure 7. Bulk wafers of mono-Si are typically built by growing an ingot from a seed of silicon featuring a known crystallographic orientation. New crystals propagating with the same structure, period and orientation of the seed crystal. The ingot must then be sliced to the desired thickness and polished to form the final wafer. Figure 10 summarises the properties of this material. [27]

* **Polycrystalline silicon:** poly-Si is at the junction between mono- and a-Si. This type of silicon layer can be obtained via an LPCVD process at 600°C resulting in a layer of silicon with grain type domain of a size of up to 100 nm. Inside each domain a piece of single continuous crystal can be found giving this material mechanical properties close to the one of mono-Si. This makes poly-Si a good and cheap alternative to using a thinned mono-Si bulk wafer. However, the size of the grain being often larger than a typical *thin film* thickness renders this material unusable for thin film deposition. [28]

* **Amorphous silicon:** a-Si is the non crystalline version of silicon. It can be obtained either by sputtering, PECVD or LPCVD at 570°C (this reducing the size of the grains/domains of poly-Si to the point where it can be considered a-Si). This material features a short range order due to its atoms being randomly ordered and positioned relatively to one another. This material can be used in thin film applications and is the most commonly used material in photovoltaic applications [28].

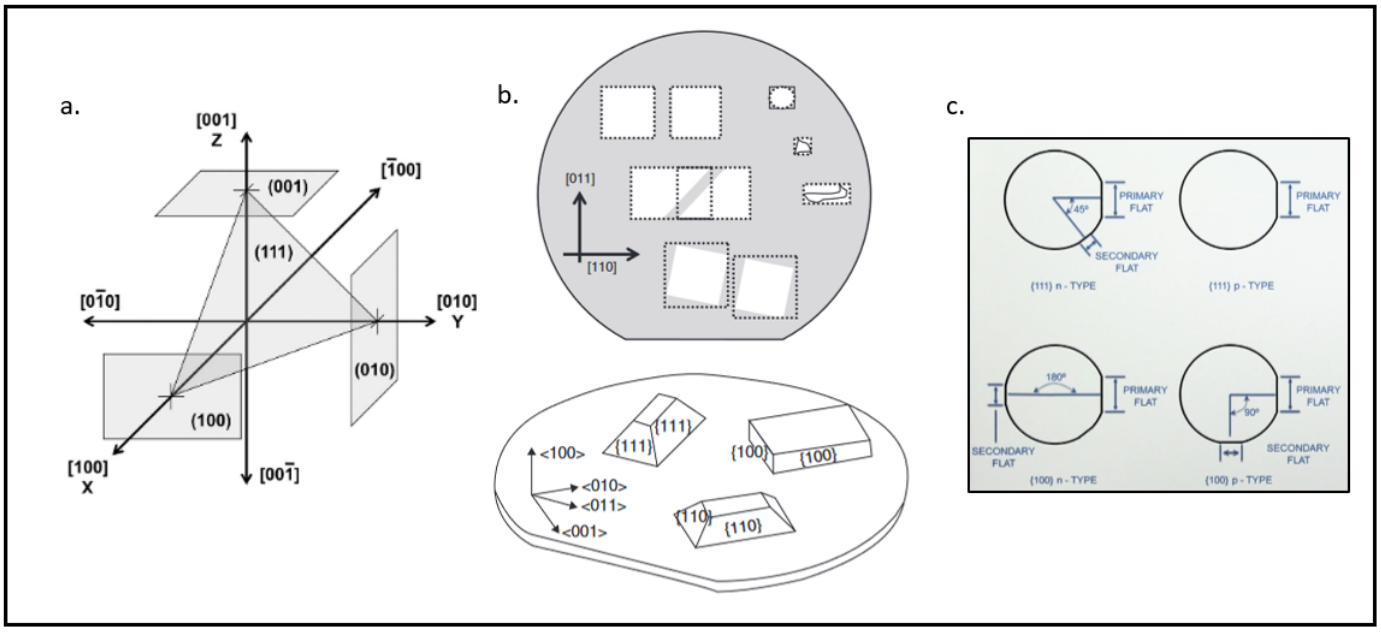


Figure 9: Representation of the crystallographic structure of silicon **(a.)** Crystallographic arrangement [25] **(b.)** Impact of the anisotropy of silicon on wet etching processing techniques [23] **(c.)** Wafer classification based on primary and secondary flats [24]

1.3.2 The challenge of using silicon

Despite being the best candidate for high performance flexible electronics, silicon comes at the cost of not being bendable as is. Indeed, its high Young's Modulus makes it a stiff and brittle material like glass. As said earlier, bendable electronics will undergo strain for a given curvature radius:

$$\epsilon = \frac{d}{2r} \tag{1.9}$$

This strain can be directly linked to a level of stress dictated by the Young's modulus, the thickness of the material and the curvature radius as dictated by the following equation and illustrated in Figure 11:[11]

$$\sigma_{st} = \frac{E * h}{2R} \tag{1.10}$$

The dashed line on the plot represents the strain level at which silicon is expected to break for >50 µm thick chips. This strain level does increase the thickness under 50 µm and a better figure of merit to predict if silicon might break under a given strain is to look at the maximal stress before fracture of silicon which is around **7 GPa**. [11] Furthermore, one must notice that silicon dies in the context of flexible electronics are often embedded in or on the surface of a flexible polymer. In this case, looking at the assembly *silicon die on polymer*, the Stoney's formula can be used to predict the stress developed on the top surface of the silicon chip following:

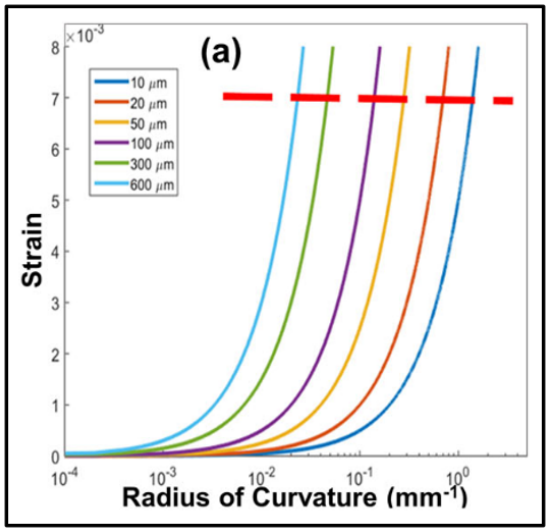


Figure 11: Strain developed in silicon for different thicknesses as a function of the curvature [11]

$$\sigma_f = \frac{E_s t_s^2}{6(1-\nu)t_f R} \tag{1.11}$$

Structural and mechanical		
Atomic weight		28.09
Atoms, total (cm^{-3})		4.995×10^{22}
Crystal structure		diamond (FCC)
Lattice constant (Å)		5.43
Density (g/cm^3)		2.33
Density of surface atoms (cm^{-2})		(100) 6.78×10^{14}
		(110) 9.59×10^{14}
		(111) 7.83×10^{14}
Young's modulus (GPa)		190 (111) crystal orientation
Yield strength (GPa)		7
Fracture strain		4%
Poisson ratio, ν		0.27
Knoop hardness (kg/mm^2)		850
Electrical		
Energy gap (eV)		1.12
Intrinsic carrier concentration (cm^{-3})		1.38×10^{10}
Intrinsic resistivity (ohm-cm)		2.3×10^5
Dielectric constant		11.8
Intrinsic Debye length (nm)		24
Mobility (drift) ($cm^2/V\text{-}s$)		1500 (electrons)
		475 (holes)
Temperature coeff. of resistivity (K^{-1})		0.0017
Thermal		
Coefficient of thermal expansion ($°C^{-1}$)		2.6×10^{-6}
Melting point (°C)		1421
Specific heat (J/kg-K)		700
Thermal conductivity (W/m-K)		150
Thermal diffusivity		0.8 cm^2/s
Optical		
Index of refraction	3.42	$\lambda = 632$ nm
	3.48	$\lambda = 1550$ nm
Energy gap wavelength	1.1 µm	(transparent at larger wavelengths)
Absorption	$>10^6\ cm^{-1}$	$\lambda = 200$–360 nm
	$10^5\ cm^{-1}$	$\lambda = 420$ nm
	$10^4\ cm^{-1}$	$\lambda = 550$ nm
	$10^3\ cm^{-1}$	$\lambda = 800$ nm
	$<0.01\ cm^{-1}$	$\lambda = 1550$ nm

Figure 10: Properties of monocrystalline silicon from [27]

where E_s is the Young's modulus of the polymer/substrate, t_s the substrate thickness, ν the Poisson's ratio, t_f the silicon's thickness and R the curvature radius.

From this point on, different practical approaches and paradigms start to emerge with the goal of achieving either bendable or flexible electronics:

* **Substrate:** From the Stoney's formula one can see that the stress is a direct function of the substrate's Young's modulus. Hence, in order to tune the stress transferred to the silicon chip upon bending one might choose a polymer with a higher or a lower modulus. For instance, if the goal is to minimise the stresses and strain, PDMS (Polydimethylsiloxane) is a better choice than polyimide ($E_{PDMS} = 360 - 870$ kPa VS $E_{PI} = 2.5$ GPa [11]).
* **Silicon thickness:** Manufacturing thinned dies, chips and ICs allows to reduce the stress built up by rendering the silicon bendable by thinning. This approach being the UTC one.
* **Two different paradigms:** Finally let's notice that two paradigm coexists when it comes to flexible electronics: avoiding stress build-up or on the contrary trying to maximise it to change the electrical and optical properties of the silicon. The latter being called *"strain engineering"*

In the upcoming sections these approaches will be described as well as illustrated by two examples linked to the world of optoelectronics.

1.4 The UTC approach for bendable electronics

This section is based on [11]. Additional sources will be explicitly mentioned otherwise

UTC, or **U**ltra **T**hin **C**hip, is a road to reach *"bendability"*, relying on the thinning beyond 50 µm of silicon. This approach allows to build complete ICs based on an intrinsically bendable material compatible with most techniques of CMOS processes.

The thinning of silicon is what renders it compliant (see Figure 12) and devices built on PDMS have been able to reach a bending radius of 6 mm without loss of functionality. The obtention of a thin layer of silicon being achievable through various techniques:

Pre-processing

Pre-processing techniques relies on reducing the thickness of the silicon wafer prior to any other processing steps required to build the ICs themselves. Existing techniques are:

* **SOI:** Probably the most obvious one. In this case the process doesn't start with a bulk silicon wafer but with a **S**ilicon **O**n **I**nsulator one. Such wafers are made out of a stack composed of a thick silicon layer (the bulk or handling layer), an insulation silicon dioxide layer (the BOX for **B**urried **Ox**yde) and the top silicon layer on which the device is built (called the functional or device layer). Despite being more expensive than classical bulk wafers ([150-1000] USD VS 25 USD), they have the advantage of featuring a very controlled top silicon layer thickness (<1% thickness variation) and can be bought with a device layer as thin as 10 nm. The IC can then be built into the device layer prior to the back etching of the SOI. The bulk-BOX-Device layer stack being used at our advantage. Indeed, using selective wet etching (for example) one can etch away the silicon under the IC with the etching stopping by itself ones it reaches the BOX. The use of SOI also allows straightforward fabrication of SiO_2 suspended membranes with a thin film silicon structure on it.

* **Anisotropic wet etching:** This approach is similar to the previous one but does not require the use of an SOI. Prior to processing the IC, a hard mask of SiO_2 or aluminum can be built on the back side of the wafer. Once the processing of the front side over, anisotropic wet etching can be used and self-limiting structures can be etched away to control the silicon layer thickness as was discussed in subsubsection 1.3.1. The major drawback of this method (also applicable to the case of SOI) is that the front side might need to be covered in a protective layer prior to the exposure to the etchant used for the back etching of the wafer.

* Other approaches such as the deposition of an thin (<µm) epitaxial layer of silicon on top of porous silicon to perform lift-off or etch stop techniques can also be used as *pre-processing* techniques.

In a nutshell, this category of fabrication methods consists in preparing the silicon thin film in advance, processing it on a carrier and then releasing the structure. Processing a thin film directly being too challenging since it increases the risk of rupturing it.

Post-processing

Opposed to the previous case, here wafer thinning is performed once the devices/ICs are completed. Methods are:

* **Back grinding:** Straightforward, cheap and easy, this technique allows to quickly remove material from the back of a wafer. The action is purely mechanical and the term **"grinding"** is not a place holder. The wafer is effectively put in contact with a rotating grinding wheel hence removing material (see Figure 12). This method is able to produce wafers as thin as 3 µm if sufficient care is taken, but it also causes a lot of damage to the crystal lattice. These imperfections in the surface finish act as stress concentration points and might be a cause of fracturing of the thinned die upon bending. This is why back grinding is often used to make an initial and quick reduction of the thickness before using another processing tool to reach the final desired thickness. These could be for instance chemical polishing or RIE (**R**eactive **I**on **E**tching). Wrapping and curling of the wafer on itself can also be another consequence of the imperfections induced by grinding. These defects can however be countered by using the *"Taiko"* process in which an outer rim of thick silicon is left to provide mechanical integrity and prevent the thin silicon film from wrapping. (see Figure 12)

* **Dry etching:** Dry etching is the case where silicon (or another material) is removed via a bombardment of charged ions. The wafer is placed in a plasma where ionised atoms (Argon ions for example) are accelerated towards the wafer. These high energy ions then impact the material with high kinetic energy hence etching it. Another type is RIE (**R**eactive **I**on **E**tching), where a reactant is added to the gas in the chamber to increase the etch rate. In this case, the phenomenon is not only physical anymore, but also chemical. An example of a silicon etchant gas would be SF_6. This techniques allows for a high level of anisotropy. However, it also comes with unwanted re-deposition of material on the wafer. Polymerisation when carbon-based gasses are used is also to be considered.

* **Dicing before grinding**: This is an extension to the common grinding process. While grinding, another risk is that the wafer might shatter. The danger is for the crack to propagate though a functional device. The solution is to pre-dice the wafer, attach it to a carrier and then grind its back. This results in thin dies on a carrier that just needs to be peeled away.

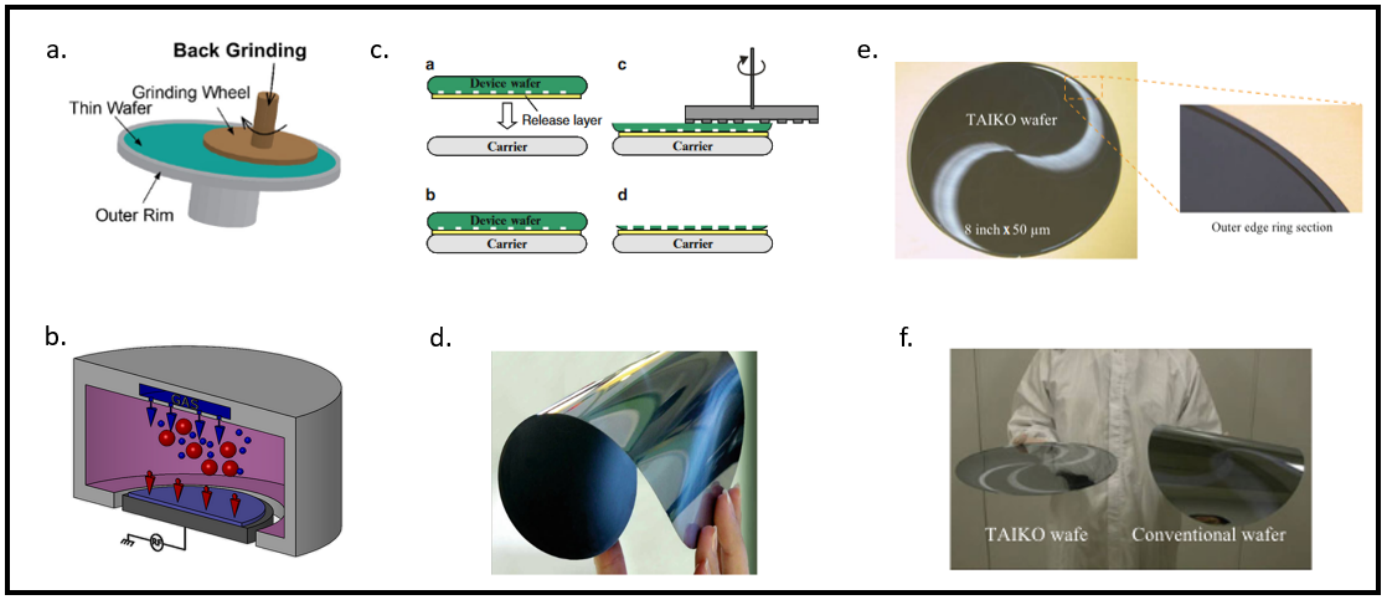


Figure 12: Example of thin film silicon processing: **(a.)** Back grinding [11], **(b.)** RIE from the *CORIAL* website, **(c.)** Dicing before grinding [29], **(d.)** 50 µm thick silicon wafer [30], **(e.)** and **(f.)** TAIKO process [31]

Other considerations

The choice of going with a pre- or post-processing path will depend on many process related parameters/challenges and is hence to be discussed on a case by case basis. Furthermore, some techniques might also not be suited for a given desired thickness and the cleanroom know-how and tooling is also to be considered when choosing a suitable approach. However, both pre- and post- allow for the fabrication of thin film devices with more or less complications for the operators. Figure 13 presents a comparison between the different thinning techniques that were previously discussed.

It is also to be noted that once built, these thin chips must be transferred onto a substrate or packaged. As we know, when bent, the interior and exterior of the curve will experience different levels of strain and hence stress will build onto the material. To increase the bendability of packaged UTCs, techniques have been developed to integrate them not on the surface of a substrate, but between two layers. This positions the UTC on the neutral plane of the substrate where stress and strain are zero in theory.

Finally, as an early introduction for a future section, UTCs are the way to go for strain engineering. Indeed, by choosing a substrate with a relatively high Young's modulus one is able to tune the strain level in the silicon hence changing its electrical and optical behavior (strain engineering). UTCs are well suited for this application since thin silicon can endure higher levels of deformation.

1.5 Stretchable electronics

This section is based on [19]. Additional sources will be explicitly mentioned otherwise

Going further than UTC, some applications might require deformation well above 1% of strain (»1%). This might be the case for biomedical devices integrated inside human tissues, electronic paper or more generally any device to be conformally wrapped around any object with a tight radius of curvature. We here also consider the case of devices that might undergo repeated deformations during their lifespan in random directions.

For such applications, the classical approach to use intrinsically bendable organic semiconductors or thinned silicon will not be sufficient. For these specific cases, research has deviated from looking into new material to focusing on the development of new mechanical configurations or on the development of nanomembranes thinner than what was studied in the UTC section (<10 nm). Let's consider two approaches.

1.5.1 Silicon islands

This approach would theoretically allow any IC, built in a specific technology node, to be made flexible and stretchable like memories or, someday, entire SOCs. The approach consists in splitting a design into small islands-like sub-module connected by stretchable interconnects. These could be made out of conventional aluminum or copper traces arranged in a horseshoe shape allowing them to be stretched like springs upon deformation of the substrate.

Process	Need for pre-processing	Material removal rate (µm/min)	Typical thickness of semiconductor layer (µm)	Challenges
Back grinding	No	0.1–10	5–10	● Deep scratches on backside ● Chipping at the edges
TAIKO	No	0.1–10	50–100	● Dicing of membrane supported on ring can lead to breakage
Dicing before grinding	No	0.1–10	10–25	● >15 µm sawlane is required ● No metal line over sawlanes
RIE/DRIE	No	0.05–10	5–30	● Non-uniform surface ● Chances of frontside contami- nation due to reactive ions
Anisotropic wet etching	Yes	0.5–2	10–100	● Sensitive to temperature and etchant concentration ● Micro-masking led hillocks formation
Epitaxial silicon over porous silicon	Yes	–	10–25	● Stacking faults due to sintering ● Warpage on thin chip during detachment from supporting pillars
Epitaxial growth and selective etching	Yes	0.17–0.2	20–50	● Low thermal budget in post-processing step due to high temperature sensitivity of etch stop layer ● Extreme control over defects in $p+$ layer
SOI box/bulk removal	No	–	12–20	● Fixing and supporting the thin chip during transfer

Figure 13: Comparison of the obtainable thickness and etch rate of different UTC related processing steps. Reduced version of a table proposed in [11]

In this case we want to prevent any strain to be transferred to the silicon in itself. The strain should be localized in the inter island space and the interconnect and flexible substrate should be able to deform elastically.

The size and thickness of the individual islands, the ability to stretch and not only bend and the spacing between each island will be a function of the polymer used. We can again confront both PDMS and polyimide. PDMS being the material of choice for stretchable electronics due to its latex-like consistency ($E_{\text{PDMS}} = 360 - 870$ kPa). On the other hand, the inability of polyimide to stretch makes it unsuitable for stretchable electronics, but it could still be used if the thickness and dimensions of the islands are limited and if they are placed on the neutral plane of the polyimide substrate. This strain limitation measure being unnecessary for PDMS substrates.

The obvious downside of this technique is that it reduces the device density by splitting a design in functional islands leaving space between them.

1.5.2 Nanomembrane

This approach is more like the one studied in the case of UTC, but thinner. In this case, sub 10 nm layers are used for either the active silicon devices or the interconnects. This allows for some exotic structures such as:

* **Wave configuration:** A thin layer of silicon or metal (to build the interconnects) is deposited on a pre-stretched layer of PDMS. Once deposited, the PDMS is relaxed and a periodic buckling appears in the deposited silicon due to the mechanical interaction and stresses that appear between the two materials upon layer deposition. This *"wavy"* configuration allows for a deformation of the nanomembrane of up to 15%. In addition, by controlling the deposition parameters and the pre-stretched conditions of the substrate, one can tune the wave frequency allowing for more or less stretchability. Indeed, a wave of high amplitude and high frequency will allow the assembly to stretch more than a low amplitude and low frequency one.

* **Pop-up:** A variation of the previous configuration is the so-called pop-up one and is used to build stretchable wires. Here the material is only bound at some locations and not on the entire plane of the substrate. The manufacturing process is however similar to the previous case: PDMS is pre-stretched whilst the metal is deposited. The layer is then bound to the substrate at the desired location (like a functional silicon island, for instance) and, upon mechanical relaxation, the unbound wires will delaminate (*"pop-up"*) from the substrate allowing them to stretch.

* **Stretchable scissor-like mesh:** Finally a scissor configuration can be used to allow for mesh deformations.

Note that in these examples the flexibility is achieved either by a combination of a nanometric membrane and a flexible substrate (or island configuration) while stretchability relies on innovative mechanical configurations (wave, scissors, ...). They are illustrated in Figure 14.

Devices based on this combination of a nanomembrane and new mechanical configuration have been demonstrated to reach bending radii of 0.05 mm while remaining functional and featuring a carrier mobility in the hundreds of $cm^2V^{-1}s^{-1}$.

Finally, even if the rigid silicon island approach has the benefit of allowing to build flexible devices at the system level, it does not allow for strain engineering. For the latter, a nanomembrane or at least thin chip approach is required. Nanomembranes having the benefit of allowing for much higher strain levels (up to 15% in the case of the wave configuration).

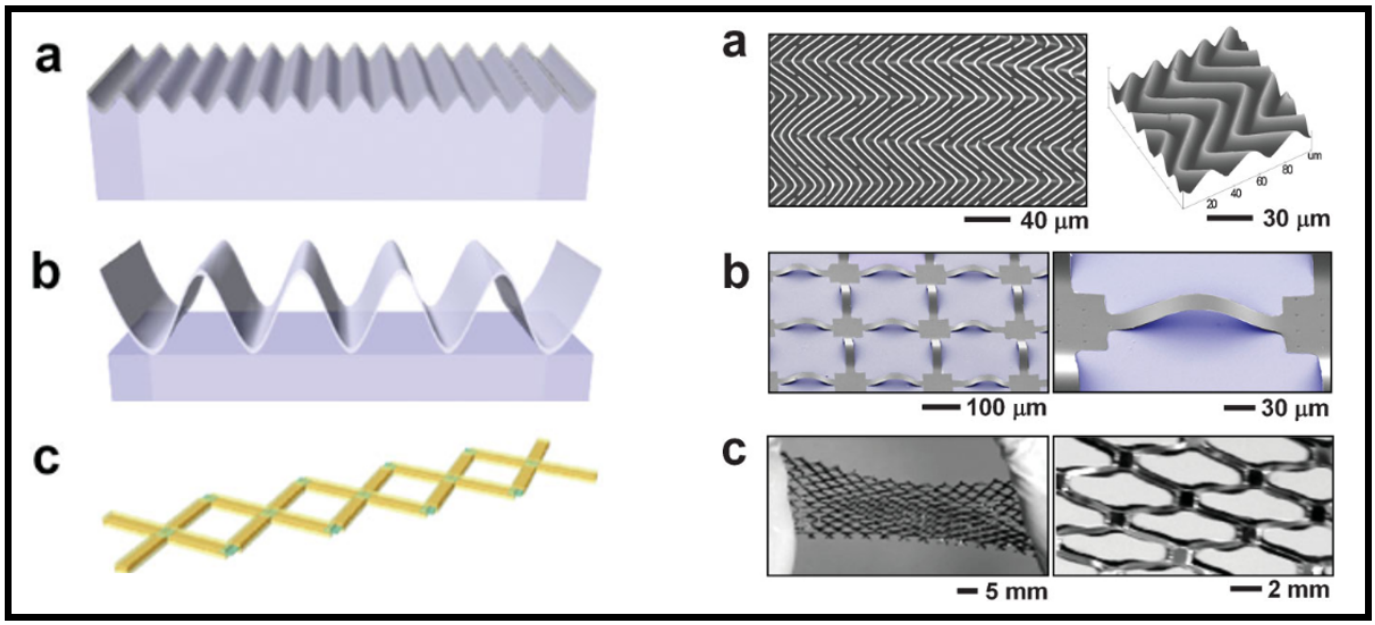


Figure 14: Example of a stretchable configuration [19]

1.5.3 Example of stretchable electronics: Hemispherical eye camera

2 Strain engineering and the application in optoelectronics

This section is based on [32, 33]. Additional sources will be explicitly mentioned otherwise

Up to here, flexibility has been introduced as a concept on its own and as a goal to achieve. Materials, structures and shapes were presented to achieve various levels of applicable strain while reducing the stress induced on the structure. These techniques ranged from trying to avoid any strain at all in the functional material (rigid silicon islands with flexible interconnects) to thinning the material down to the nanomembrane. The latter allowing to reach high levels of strain with reduced stress build-up (remember that $\epsilon = \frac{d}{2R}$ and that $E = \frac{\sigma}{\epsilon}$).

These nanomembranes (in combination with innovative shapes like the wave one) have been able to reach higher and higher levels of strain without rupture (up to 10 or 15%) in functional layers of silicon. This opens an highway to further development of a field called *Stain engineering*.

Strain engineering could be defined as:

Modification of the electrical/optical properties of a semiconductor by application of a known level of strain either self induced or by external loading of the structure

Strain engineering with small levels of deformation is already widely used in the industry since it is an easy and inexpensive way to boost the performances of transistor based devices. Indeed, the application of strain by self loading of the device allows for a reduction of the effective mass of the carriers leading to an increase (>100% increase for $\epsilon > 1\%$) in their mobility. [34]

Achieving such levels of stress in silicon is made possible by the reduction of the effective volume of each individual transistor. Self loading is induced by micro- and nanofabrication techniques taking advantage of the self induced stresses generated at the interface between two different materials.

Hence, further development in nanometric silicon membranes enabling to reach increasingly high levels of strain provides us with even more possibilities such as sufficient modification of the fundamental silicon bandgap to achieve photodetection deeper in the infrared region.

This last part is what we are going to focus on for the rest of this paper. Silicon has been used for photodetectors for many years and is now found in most digital cameras in the form of CMOS sensors. However, its fundamental bandgap of 1.12 eV does not allow it to detect wavelengths passed 1.1 µm. Hence, here strain engineering comes into action to try and shrink this bandgap to achieve detection passed this threshold.

This section will hence focus on the electrical properties of silicon and investigate how strain can help reduce its bandgap and enhance its properties.

2.1 Band structure of silicon and other electrical properties

This section is a summary of a detailed calculation made by Pr D.Flandre and V.Bayot in [35]. Complete development can be found in the source (in french). Additional sources will be explicitly mentioned otherwise

This section comes as a complement to the mechanical properties of silicon already presented in subsubsection 1.3.1. In that previous section, the focus was more on the mechanical properties of silicon whereas here the text will take a deep dive in its properties as a semiconductor.

Silicon is considered as a crystal. This crystal as an all is neutral but can be seen as a combination of positively charged ions and an equivalent number of almost free electrons. This denomination is important because we consider electrons located deep into the lattice to be fixed and hence unable to participate in conduction whereas almost free electrons on the periphery of the crystal can. The study of their energy level is of critical interest. Indeed, the level of energy they will occupy will define the conductive properties of the material.

To determine those energy levels, one must solve the Schrödinger equation to obtain the wave function $\Phi_i(\bar{r}_i)$ for each almost free electrons. The general form of the time independent Schrödinger equation:

$$\big[-\frac{\hbar - 2\Delta_i}{2m} + V(\bar{r}_i) + \sum_{j=1}^{Q} V_{ij}(\bar{r}_i)\big]\Phi_i(\bar{r}_i) = E_i\Phi_i(\bar{r}_i) \tag{2.1}$$

- $\hbar = \frac{h}{2\pi}$ the reduced Planck constant
- $\Phi_i(\bar{r}_i)$ the wave function of the almost free electron i
- E_i the energy level of the almost free electron i. The energy of all these electrons being $E = \sum_{j=1}^{Q} E_i$ where Q is the total number of almost free electrons
- $V_{ij}(\bar{r}_i)$ the potential between two almost free electrons
- $V(\bar{r}_i)$ the potential between the almost free electron i and the lattice ions

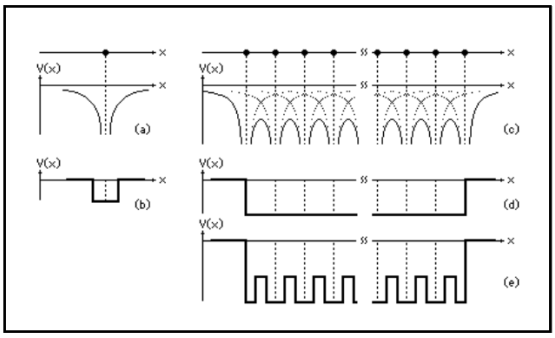


Figure 15: Model of a periodic crystal **(a.)** Sommerfeld **(b.)** Krönig-Penney. Illustration from [35]

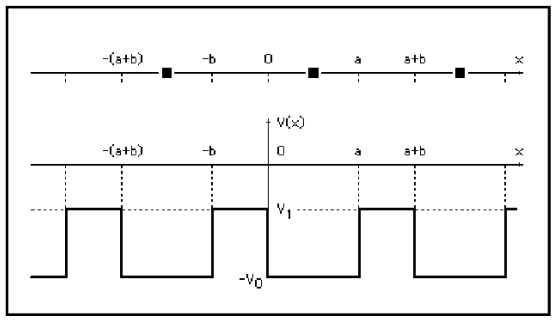


Figure 16: Krönig-Penney model. Illustration from [35]

From here, one can see that to solve for $\Phi_i(\bar{r}_i)$, one must know both $V_{ij}(\bar{r}_i)$ and $V(\bar{r}_i)$. This computation will hence be made easier by wisely choosing a model for this potential. Simplifying the expression, one could say that we need to find a model for $\sum_{j=1}^{Q} V_{ij}(\bar{r}_i) + V(\bar{r}_i) = V'(\bar{r}_i) + V(\bar{r}_i)$. Considering that monocrystalline silicon is a periodic repetition of well ordered silicon atoms, one can understand that the potential acting on each almost free electrons will be a periodic function.

The equation expressing the potential between a negative electron $-q$ and a lattice ion of charge $+qZ$ is given by:

$$V(x) = \frac{-Zq^2}{4\pi\epsilon|x|} \tag{2.2}$$

where x is the distance between the ions and the electron. Now, since a crystal is a periodic repetition of atoms, one can conclude that the potential acting on the almost free electrons will be the interaction between every individual potential

V_i, leading to representation (c) of Figure 15. To make the computation of the wave function easier, one can use the Krönig-Penney model, where the potential is simplified to a battlement like configuration with:

$$V(x) = \begin{cases} -V_0 \quad (V_0 > 0) \quad \text{for} \quad 0 \leq x \leq a \\ +V_1 \quad \text{for} \quad -b < x < 0 \end{cases} \tag{2.3}$$

This simplification is represented in Figure 15 (e) and in Figure 16. This model implies that one can simplify the Schrödinger equation, previously introduced, into two individual equations to be solved in order to obtain the energy level of the almost free electrons:

$$\begin{cases} -\dfrac{\hbar^2}{2m}\dfrac{d^2\Phi(x)}{dx^2} - V_0\Phi(x) = E\Phi(x) \text{ for } 0 \leq x \leq a \\ \\ -\dfrac{\hbar^2}{2m}\dfrac{d^2\Phi(x)}{dx^2} - V_1\Phi(x) = E\Phi(x) \text{ for } -b < x < 0 \end{cases} \tag{2.4}$$

Since the potential $V(x)$ is periodic, the solution of the Schrödinger equation should be a *Bloch* function as follows:

$$\Phi(x) = \begin{cases} e^{ikx}u_{a,k}(x) \quad \text{for} \quad 0 \leq x \leq a \\ e^{ikx}u_{b,k}(x) \quad \text{for} \quad -b < x < 0 \end{cases} \tag{2.5}$$

where u_k is a periodic function with the same period as the potential: $u_k(a+b+x) = u_k(x)$ and k is the wave number (modulus of the wave vector $|\mathbf{k}| = k = \dfrac{2\pi}{\lambda}$). Solving this equation gives us what we call the dispersion equation:

$$\cos[k(a+b)] = \cos(\alpha a)\cosh(\beta b) + \frac{b(\beta^2-\alpha^2)}{2\alpha}\sin(\alpha a)\frac{\sinh(\beta b)}{\beta b} \tag{2.6}$$

where we define α and β as follows (with m the mass of the electron):

$$\alpha = \frac{1}{\hbar}\sqrt{2m(E+V_0)} \quad \text{and} \quad \beta = \frac{1}{\hbar}\sqrt{2m(V_1-E)} \tag{2.7}$$

The dispersion equation links the energy level and the wave number k, giving the allowed energy levels inside the crystal for the almost free electrons (each combination of E and k giving a particular wave function for the considered electron). Making one more simplifications by considering that $(V_1 \longrightarrow \infty)$ and $(b \longrightarrow 0)$ but with $V_1 b$ remaining finite (to avoid undetermination), one can write:

$$\cos(ka) = \cos(\alpha a) + P\frac{\sin(\alpha a)}{\alpha a} \quad \text{with} \quad P = \frac{mV_1 ab}{\hbar^2} \tag{2.8}$$

From this simplified expression, one can see that since the cosine function is periodic and bounded between [-1,1], for each value of k, only a discrete number of αa values are possible for the equation to hold. Hence, rewriting Equation 2.7 as:

$$E = -V_0 + \frac{(\alpha\hbar)^2}{2m} = -V_0 + \frac{(\alpha a\hbar)^2}{2ma^2} \tag{2.9}$$

and using it in Equation 2.8, one can notice that we obtain a discrete number of possible values of E that we write $E_n(k)$ with $n = 1, 2, 3, ...$ This gives us the well-known result that in a periodic crystal, the electron can only occupy a discrete number of allowed energy levels. Each of these energy levels being separated by a forbidden region where electrons cannot be present (unless during the short transition period between energy bands).

$E_n(k)$ is a periodic function of period $k = \dfrac{2\pi}{a}$ where a is the wavelength of the Krönig-Penney crystal model. Note that this matches the definition of the wave number given earlier. Figure 17 here below represents the allowed energy levels of the electrons limited to the first Brillouin Zone (first period of the periodic crystal). In addition, if one takes the amplitude of each energy function, the so-called *allowed energy bands* can be defined.

This rather long development allows us to understand that almost free electrons have to exist in certain well-defined regions and avoid forbidden bands. Since lattice electrons are bound and unable to move, one must notice that the electrical properties of the crystal will then be linked to the distribution of the energy states of these valence/almost free electrons. From here, one can predict the electrical behavior of a given material by looking at these bands.

The occupation of a given allowed energy band by an electron is determined by a distribution function. For an undoped, hence intrinsic material, this statistic distribution, known as the Fermi-Dirac distribution, is given by:

$$f_{FD}(\bar{k}) = \frac{1}{1 + e^{[E_n(\bar{k}-\mu)]/kT}} \quad \text{with @ 0K} \quad \begin{cases} f_{FD}(E) = 1 & \text{if} \quad E < \mu \\ f_{FD}(E) = 0 & \text{if} \quad E > \mu \end{cases} \tag{2.10}$$

which gives the probability that an electron will occupy a given energy level $E_n(k)$. To simplify the discussion, one could consider the case where the material is at 0K. In this case, the probability to find an electron above the chemical potential μ is equal to zero. The right part of Figure 17 illustrates different possible energy band configurations. Note that the electrons will always fill the lowest energy states first. The last occupied band is called the valance band whereas the first band above it is called the conduction band. It can either be empty or partially filled with charge carriers. From here on, one can define:

* **Insulating material (a):** In this case the valence band is full of electrons ($2N$) and the conduction band is empty. The bandgap between the two is large. This forbidden gap between the conduction and valence band is defined as $E_g = E_c - E_v$ where E_c is the energy level of the conduction band and E_v is the energy level of the valence band. The bandgap for an insulator is typically of 5 eV.
* **Semiconductors (b):** Same as for an insulating material but the bandgap is smaller (1 eV) allowing carriers to jump to the conduction band from the valence band given enough energy is provided (thermal energy $E_T = k_b T = 26$ meV @300K). In the case of silicon the bandgap is 1.12 eV.
* **Metal (c):** If the bands overlap or if the valence band is partially filled, the crystal is defined as a metal.

Considering the case at room temperature (300K), one can define the Fermi level[4] which, for an intrinsic crystal, sits in the middle of the forbidden band. The doping of the crystal allows to shift this level of energy and bring it either closer to the valence or conduction band. This is the principle of doping.

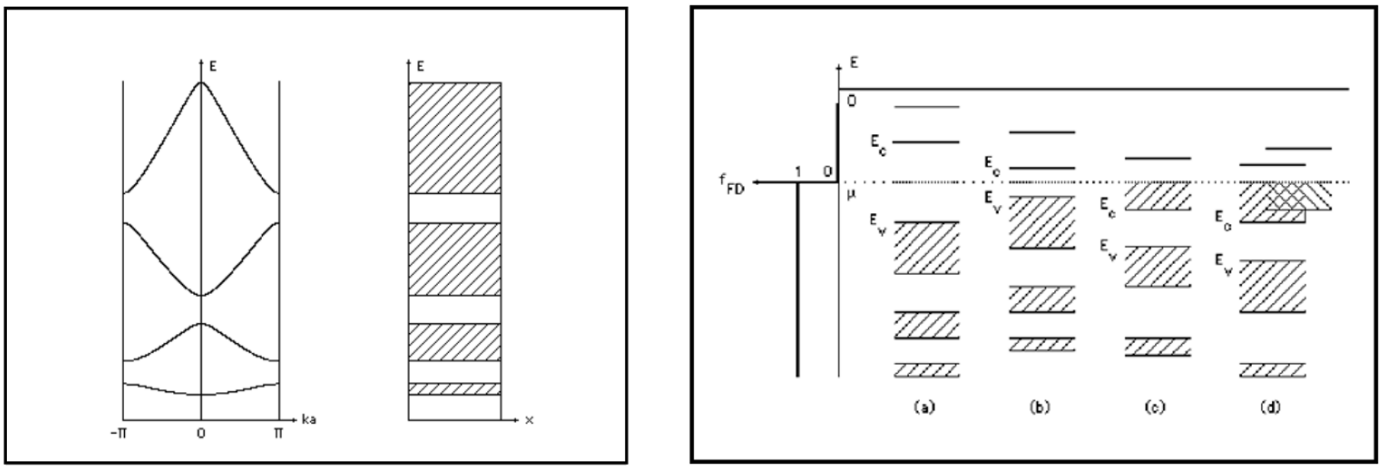


Figure 17: Allowed energy bands $E_n(k)$ limited to the first Brillouin zone **(Left)** and energy bands for an insulating, semiconductor and conductor material **(Right)**. Illustration from [35]

2.1.1 Doping of silicon

Silicon features what we call an intrinsic energy level. When undoped, the semiconductor features carriers generated only by thermal agitation. An electron can be excited and *jump* from the valence band to the conduction band. This electron now being able to participate in the electrical conduction, leaving behind what is called a hole (or a positive charge). The hole being a concept used to simplify the calculations, because in practice, a hole is just a lack of electrons. When undoped, the crystal is, as was said, in its intrinsic state and has a level of auto-generated charge carriers coming from thermal agitation given by:

$$p = n = n_i(T) = 3.78 * 10^{16} T^{3/2} \exp \frac{-7014}{T} \quad \text{@300K} \quad n_i(300) = 1,38 * 10^{10} \quad [\text{cm}]^{-3} \tag{2.11}$$

[4]Careful here: Fermi energy is the Fermi level when at 0K while the Fermi level is defined for all temperatures

here T is the temperature in Kelvin. This equation gives us the dopant concentration as a function of the temperature in an intrinsic crystal. The concentration in holes and electrons being the same since each excited electron leaves exactly one hole behind. The Fermi level in the intrinsic case is given by:

$$E_i = \frac{1}{2}\left[E_c + E_v = \frac{3}{2}kT\ln\frac{m_v}{m_c}\right] \approx \frac{1}{2}\left[E_c + E_v\right] \tag{2.12}$$

where m_c and m_v are the *effective masses* of electrons and holes (see later), respectively. This level is then, when undoped, in the middle of the forbidden band ($E_i = 0.56$ eV for undoped silicon considering a 1.12 eV bandgap and considering the valence band as the zero energy level).

However, in practice, semiconductors do not rely only on these intrinsic charge carriers but also on charges coming from impurities called doping atoms. Two kinds of doping are possible:

P doping

In this case, the silicon is *implanted* with impurities able to catch an electron, leaving a mobile hole behind. A common type of a P doping species is boron. Initially, just after the implantation, the boron atom is neutral and leads to what we call an *extrinsic energy level* to the energy band configuration of the initial crystal. An extrinsic energy level is an energy level brought by the impurities and is initially void of electrons. Note that this extrinsic energy level must be in the forbidden band of the silicon in order to change the electrical properties of the material.

Once implanted, the doping atom/impurity must be ionised by heating giving:

$$\mathrm{B} + \Delta\mathrm{E_A} \longrightarrow \mathrm{B^-} + (+\mathrm{q})$$

The energy $\Delta E_A = E_A - E_V$ in the equation represents the amount of energy required to ionise the boron atom. In other words, it is the amount of energy required to bring an electron from the valence band to the extrinsic energy level E_A. The boron atom will be ionised, become negatively charged (as it accepted one extra electron) and the electron will leave a free hole in the valence band of the silicon.

Hence, in that case, the excess carriers will be holes and the dominating currents in the material will be hole currents. Indeed $n << p$. The activation energy for boron being of 0.045 eV, meaning that all of the boron impurities are naturally ionised at 300k ($E = kT = 26$ meV @300K). If one considers that all acceptors have been ionised ($N_A = N_A^-$ with N_A the acceptor concentration and N_A^- the ionised acceptor concentration), the hole concentration will be given by:

$$p = \frac{1}{2}\left[N_A + \sqrt{N_A^2 + 4n_i^2}\right] \tag{2.13}$$

N doping

Here the ionised implanted impurities must be able to release one electron. A common N doping atom being phosphorus. Like in P doping, once implanted, the neutral impurities will bring an extrinsic level of energy E_D and upon heating will lead to the following:

$$\mathrm{P} + \Delta\mathrm{E_D} \longrightarrow \mathrm{P^+} + (-\mathrm{q})$$

The energy $\Delta E_D = E_C - E_D$ is the energy required for the excess electron of phosphorus to jump from the extrinsic energy level, become almost free and reach the conduction band, leaving behind a fixed positive charge. The excess charge carrier will, in this case, be electrons in the conduction band and their concentration, assuming complete ionisation ($N_D = N_D^+$ with N_D the donor concentration and N_D^+ the ionised donor concentration), is given by:

$$n = \frac{1}{2}\left[N_D + \sqrt{N_D^2 + 4n_i^2}\right] \tag{2.14}$$

General considerations

Figure 18 illustrates the doping of silicon. One can see that a doping atom replaces a silicon atom in the lattice upon heating. One important thing to mention is that the intrinsic Fermi level of the crystal will shift up or down as a function of the dopant species following:

$$E = \begin{cases} E_i & \text{if intrinsic crystal} \\ E_D - KT\log 2 & \text{if N doped} \\ E_A + KT\log 2 & \text{if P doped} \end{cases} \tag{2.15}$$

which is defined as the effective energy level. This quantity being used to give a common Fermi-Dirac distribution whatever the type of doping:

$$f(E_t) = \frac{1}{1 + \exp\left(E_t - \mu\right)/kT} \tag{2.16}$$

Also note that, even if heavily doped ($\geq 10^{21}$) $[\text{cm}]^{-3}$), since each mobile charge carrier is associated with a fixed opposite charge in the lattice, the crystal remains neutral:

$$\mathrm{P} + \mathrm{N_D^+} = \mathrm{n} + \mathrm{N_A^-}$$

Finally, the concentration in charge carriers of a doped crystal is in general much higher than the intrinsic carrier concentration: $\approx 10^{10}$ VS $[10^{17}; 10^{21}]$ atoms $[\text{cm}]^{-3}$.

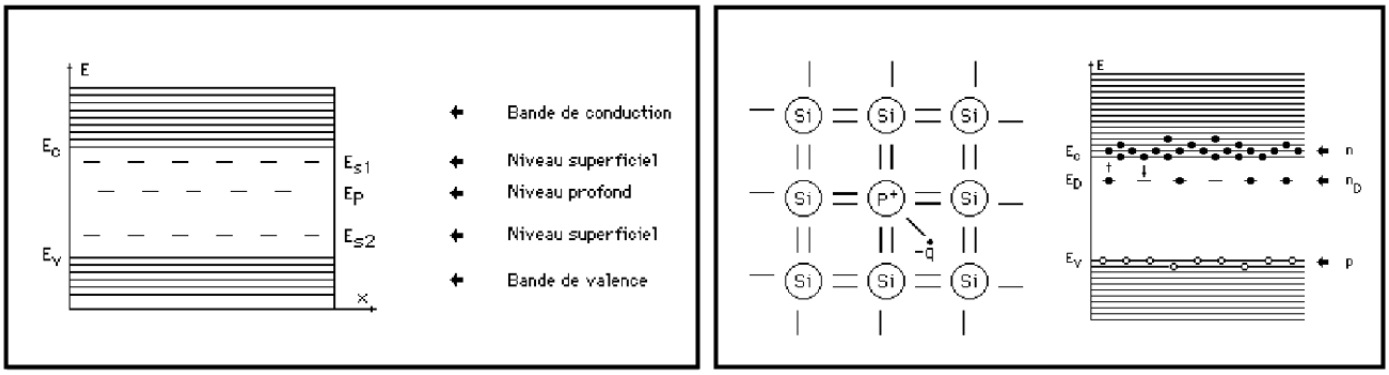


Figure 18: P doping **(Left)** and N doping **(Right)**. Illustration from [35]

2.1.2 The concept of effective mass and carrier mobility

One final consideration before looking into the impact of strain on the electrical properties of silicon is to mention the existence of a concept called the *effective mass*. The effective mass is a mathematical construct designed to easily study the impact of external forces (electric field for instance) applied on an almost free electron inside a crystal lattice. The theorem is developed around the assumption that the externally applied forces do not vary much over the length of one crystal period and are weak compared to the internal crystallographic forces. These assumptions allow us to develop the theorem without going back to the Schrödinger equation.

Let's consider a force $\bar{F}$ applied on an electron. One can link the variation of the electron energy to the applied work:

$$dE = \bar{F}.\bar{v_g}dt \tag{2.17}$$

where v_g is the group speed of the electron. Considering that the group speed of the electron is a function of its energy state:

$$\bar{v_g}(\bar{k}) = \frac{1}{\hbar}\mathrm{grad}_{\bar{k}}E_n(\bar{k}) \tag{2.18}$$

one can rewrite the previous expression as:

$$\frac{dE}{dt} = \frac{1}{\hbar}\bar{F}.\mathrm{grad}_{\bar{k}}E \tag{2.19}$$

From here, by noting $\hbar\frac{d\bar{k}}{dt} = \bar{F}$, one can derive a cousin of the classical expression $F = ma$:

$$\frac{d\bar{v_g}}{dt} = \big(\frac{1}{\mathbf{m}^*}\big)\bar{F} \Leftrightarrow \bar{F} = \mathbf{m}^*\frac{d\bar{v_g}}{dt} \tag{2.20}$$

where

$$m^* = \frac{\hbar^2}{\frac{d^2E}{dk^2}} \tag{2.21}$$

is the 1D definition of the effective mass of the electron. One can see from these expressions that the effective mass theorem allows us to manipulate the electrons in the lattice as free particles on which a force is applied.

The effective mass can hence be defined negatively or infinitely since it is a function of the second derivative of the energy state of the electron. The fact that it can be defined negatively is also the reason why the notion of positive hole has been introduced. A hole being a mobile positive charge carrier with a charge opposite to the one of the electron. It essentially is a lack of electron. The effective mass of a hole is defined in the same way as the effective mass of an electron, but with a minus sign to make the equation $F = ma$ work.

The concept of effective mass also allows to define the mobility of the charge carriers. Let's go back to the equation expressing the motion of an electron. Considering that an external electric field $\bar{E}$ is applied, we get ($-q$ the charge of the electron):

$$m_n^* = \frac{d\bar{v}}{dt} = -q\bar{E} \tag{2.22}$$

Integrating this equation and solving it for $\bar{v}$, one can get an expression for the group speed. Since the motion is induced by an external field, we will call it the *drift speed*. Without going into the details of the calculation of this particular solution and jumping directly to the result, one can write down the drift speed of the charge carriers as:

$$v_{\bar{d},n} = -\mu_n \bar{E} \qquad \text{and} \qquad v_{\bar{d},p} = \mu_p \bar{E} \tag{2.23}$$

where $\mu_p = \frac{q\tau_p}{m_p^*}$ and $\mu_n = \frac{q\tau_n}{m_n^*}$ are the hole and electron mobility, respectively, expressed in $\text{cm}^2/(\text{Vs})$. They are a function of the mean free path of the carrier that is a function of both the temperature and the doping level (more thermal agitation or more impurities leads to more collisions) and of the effective mass.

$$\begin{cases} \bar{J}_n = -qn\bar{v}_{d,n} = qn\mu_n\bar{E} \\ \bar{J}_p = qp\bar{v}_{d,p} = qp\mu_p\bar{E} \\ \sigma = q(n\mu_n + p\mu_p) \end{cases} \tag{2.24}$$

The mobility of the charge carrier has a direct impact on the current density and the resistivity of a material as expressed by Equation 2.24.

2.2 The use of silicon in optoelectronics

This section is drafted based on [35, 33]. Other sources will be mentioned explicitly

Silicon has been used for years to build photodetectors. The principle behind photodetection (that one might define as the generation of an electrical current in a crystal when exposed to light) relies on the generation of an electron that will jump from the valence band to the conduction band upon reception of energy from a photon.

This photogeneration is however possible only if the incident photon provides enough energy to the electron. The energy of the incident light being given by:

$$E = h\nu \tag{2.25}$$

Where $h = 4.135 * 10^{-15}$ [eV/Hz] is the Planck constant and ν is the frequency of the incident light. In the case of silicon, the gap between the valence and conduction band is of 1.12 eV. This implies that a photon should bring at least this amount of energy to enable the electron to cross the bandgap. Hence, silicon should be able to detect light with a wavelength of maximum (with C as the speed of light):

$$1.12[\text{eV}] \leq h\nu \Leftrightarrow \lambda \leq \frac{C}{1.12/h} = 1.1[\mu\text{m}] \tag{2.26}$$

The exact phenomenon that occurs when a photon with sufficient energy arrives in the crystal and excites an electron depends on the type of bandgap of the material. Indeed two types are possible and are defined as a function of the crystal momentum: $\mathbf{P}_{\text{crystal}} = \hbar\mathbf{k}$.

This crystal momentum represents the momentum of the electrons in a crystal lattice. This means that, as one can see in the equation, the momentum will be a function of the wave vector $\bar{k}$. As in classic mechanics, this momentum must be preserved and this explains the origin of the different types of bandgap in different materials:

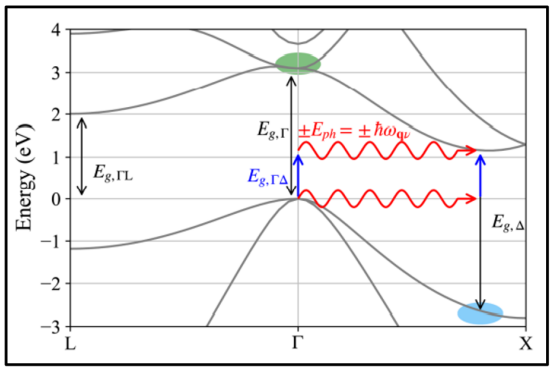


Figure 19: Indirect and direct transition in silicon. Illustration from [33]

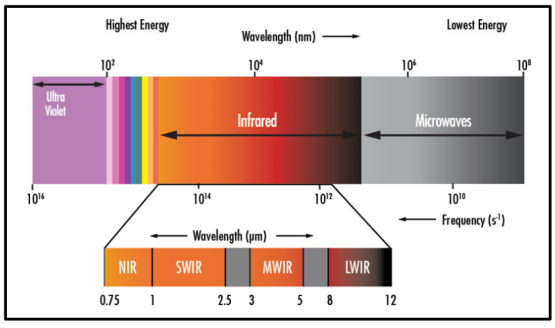


Figure 20: Light spectrum and the place of infrared. Image by Honour Optics Co.

* **Direct bandgap materials:** In this case, the k-vectors of electrons and holes are the same in the conduction and the valence band. This means that an electron can jump directly from the valence to the conduction band without any changes of the crystal momentum. The transition is direct and occurs when a photon arrives and gives enough energy to an electron to jump across the bandgap (1.12 eV for silicon). An example of a direct bandgap material would be amorphous silicon. The left part of Figure 21 shows a direct transition between the valance and conduction bands. No change in crystal momentum.

* **Indirect bandgap materials:** In this case, there is a variation of the crystal momentum between the minimum energy state of the conduction band and the maximum energy state of the valence band coming from a difference in their wave vectors. This means that a direct transition of an electron across the bandgap is not possible without violating the conservation of momentum. For this transition to occur, a *phonon* must be emitted or absorbed by the electron. This phonon represents an exchange of momenta between the crystal lattice and the electron and allows to conserve the total momentum of the crystal. The transition can be seen as carried out in two steps: first a vertical transition for the energy level of the electron to increase. It ends up at the correct energy level but with the wrong momentum. Then an exchange of momenta with the lattice resulting in an horizontal shift bringing it to the conduction band. This is shown in Figure 21 (right). This indirect photogeneration dominates for light with an incident energy above the fundamental bandgap of silicon. However, even in an indirect material, direct transitions can occur between part of the energy band that have the same k-vectors. In the case of silicon, this direct transition at a constant momentum is possible for photons of energy above 3.2 eV: the direct bandgap of monocrystalline silicon. When the incident energy is above 3.2 eV the absorption rate increases. Both direct and indirect transitions in silicon are illustrated on Figure 19.

The difference between the direct and indirect transition is a further explanation on why amorphous silicon is used for thin solar array. Indeed, since a-Si features a direct bandgap, it absorbs light better than its counterpart mono-Si and can hence be thinner. The mono-Si built solar arrays being indeed much thicker (in the hundred of microns range)[36].

In this thesis the focus will mainly be on the indirect bandgap type mono-Si. For this type of material one can define define the absorption level α as:

$$\alpha(E_\lambda) = P\Big[\frac{(E_\lambda - E_g + E_{ph})^2}{\exp\frac{E_{ph}}{k_bT} - 1} H(E_\lambda - E_g + E_{ph}) + \frac{(E_\lambda - E_g - E_{ph})^2}{1 - \exp -\frac{E_{ph}}{k_bT}} H(E_\lambda - E_g - E_{ph})\Big] \tag{2.27}$$

where we have P a constant function of the material, E_g the fundamental bandgap, E_{ph} the energy of the phonon and $H(x)$ the *Heaviside* step function. This step function allowing to write in a single equation the absorption (left part) and emission (right part) of a phonon. One can read from this equation that the absorption coefficient is small for energy values below the fundamental bandgap of a material, which is what we expected.

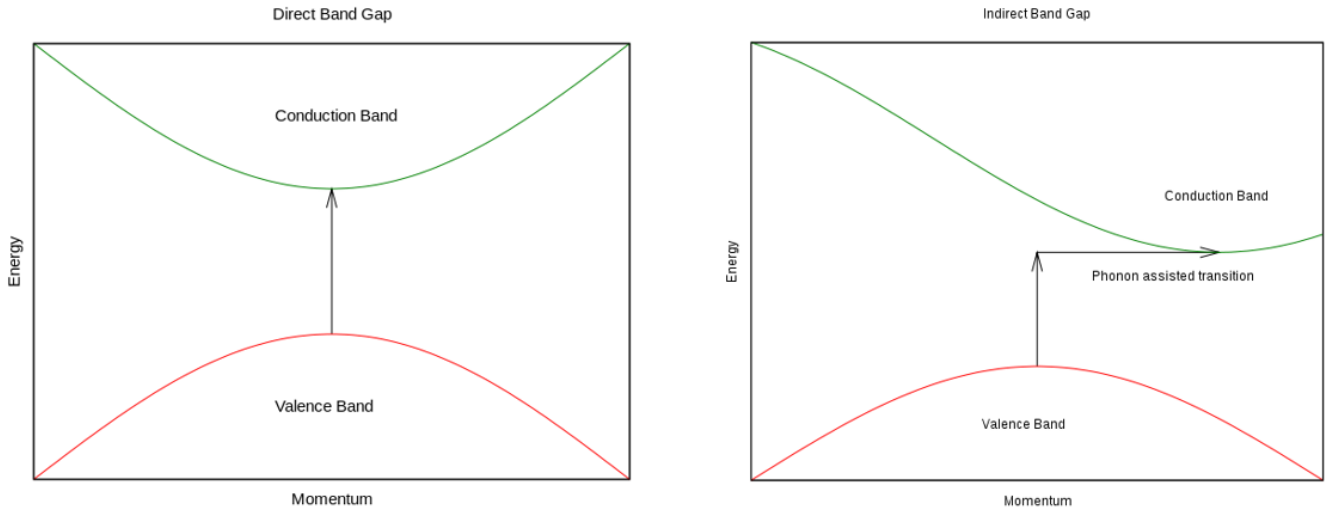


Figure 21: Direct vs indirect bandgap electron transition [36]

Going back to some more straightforward considerations, silicon is able to detect (with its 1.12 eV bandgap) most of the visible light and even a bit of the infrared spectrum stopping at the beginning of what is called the SWIR range.

Short **W**ave **In**frared is defined as the region of the infrared spectrum located in the [0.9-2.5] µm range as shown in Figure 20. Infrared sensor find many applications in high contrast photography, lidar imaging systems for autonomous vehicles or to allow for quick inspection of the interior of an object. Indeed, some plastics are transparent to SWIR enabling us to look inside a container without opening it, for example.

The issue however, is that this [0.9-2.5] µm interval means that the incident energy of photons will be comprised between [1.36-0.49] eV. So, even if small detection of the beginning of the SWIR spectrum is possible, deeper detection will not be possible with silicon. Indeed, other materials such as InGaAs are today being used to detect this specific range of

wavelength but they have the inconvenient of being expensive to produce since they are incompatible with classical CMOS processing techniques.

This is where strain engineering enters the playground. Indeed, as we will see in the following section, the application of strain to silicon allows to modify its energy band structure in a way that leads to a decrease of the bandgap and an increase of carrier mobility (Drude model). This leads to an effective improvement of the detection coefficient α as well as an increase of the cut-off wavelength of silicon (wavelength past which the absorption coefficient is below 1 $[\text{cm}]^{-1}$).

2.3 Impact of strain on the electrical properties

Let's now consider that a uniaxial strain is applied along two different orientations of the crystal: [110] and [111]. When strain is applied on a material, one can write the strain deformation tensors as follows:

$$\mathbf{E}_{[110]} = \frac{1}{2}\begin{bmatrix} (\epsilon_\perp + \epsilon_\parallel) & (\epsilon_\perp - \epsilon_\parallel) & 0 \\ (\epsilon_\perp + \epsilon_\parallel) & (\epsilon_\perp - \epsilon_\parallel) & 0 \\ 0 & 0 & 2\epsilon_\perp \end{bmatrix} \quad \text{and} \quad \mathbf{E}_{[111]} = \frac{1}{3}\begin{bmatrix} (\epsilon_\perp + 2\epsilon_\parallel) & (\epsilon_\perp - \epsilon_\parallel) & (\epsilon_\perp - \epsilon_\parallel) \\ (\epsilon_\perp - \epsilon_\parallel) & (\epsilon_\perp + 2\epsilon_\parallel) & (\epsilon_\perp - \epsilon_\parallel) \\ (\epsilon_\perp - \epsilon_\parallel) & (\epsilon_\perp - \epsilon_\parallel) & (\epsilon_\perp + 2\epsilon_\parallel) \end{bmatrix} \tag{2.28}$$

which represent the deformation that happens in the crystal upon the application of strain. As one can see from both these tensors, the impact of an external force along a single direction will generate a biaxial deformation of the crystallographic structure. This is the reason why two components are present: $\epsilon_\parallel$ which is the strain along the direction of the applied force and $\epsilon_\perp$ which is the induced deformation of the crystal lattice along the direction perpendicular to the applied force. This induced perpendicular deformation is proportional to the parallel one and can be described as: $\epsilon_\perp = -D_{[ijk]}\epsilon_\parallel$ where $D_{[ijk]}$ is function of a combination of the elastic constants C of silicon (anisotropic quantities).

From these two matrices one can already see that the crystal will not deform in the same way depending on the orientation of the applied deformation. We can hence anticipate that the electrical and optical properties will also depend on the crystallographic orientation and on the orientation of the applied strain.

The impact of this strain has been analysed in [33] which relies on numerical simulations that go way beyond the topic of this thesis. Furthermore, the paper takes a deep look into the impact of strain on different valleys of the band structure. Let's make here a summary to highlight the main effects.

2.3.1 Impact on the band structure

Upon the application of a force on the crystal, the shape of the lattice will be impacted. As we know, the band structure of a material is highly influenced by the crystal configuration and periodicity. Hence, straining the lattice will lead to a modification of the electronic bands and might hence impact the bandgap.

This is indeed what happens in simulations. As one can see in Figure 23 (a.) and (b.) the bandgap decreases for both a tensile and compressive strain and for the two studied crystallographic orientations.

The effect is however not of the same magnitude since for the [110] case the reduction of the fundamental bandgap is of -0.11 eV/% for compressive and tensile strain while in the [111] the reduction is of -0.10 eV/% for compressive and -0.05 eV/% for tensile strain (up to $\pm 2\%$ of strain).

Note that this effect of the reduction of the bandgap was also modeled by Ajit K. Katiyar and his team in [1]. In this paper a biaxial strain is applied on a matrix of 10 nm thick silicon pixels (PN junction photodetectors). They computed a reduction of the bandgap to 0.92 eV for a 1.8% level of strain and a bandgap of 0.76 eV for a strain of 3%. These bandgaps enabling, in theory, the detection of SWIR light up to a wavelength of 1.63 µm. The reduction of the bandgap is illustrated in Figure 22.

2.3.2 Impact on the effective mass

Since the electronic band structure of silicon is modified, an impact on the effective mass is expected following the discussion made in subsubsection 2.1.2. Indeed, since the effective mass depends on the second derivative of the energy state of the carrier, a modification of the shape of these bands will lead to an impact on the effective masses. This impact is shown in Figure 23 (c.) and (d.) for the valence band.

This change of effective mass comes, as was said, from the modification of the shape of the energy bands. In the case of the conduction band, the shape does not change much and this leads to a negligible modification of the effective masses. However, this is not the case for the valence band.

Upon the application of strain, the shape of the valence band changes *"a lot"*. The band transitioning from a non-parabolic and anisotropic shape to a parabolic one. This modification leads to a considerable reduction of the effective masses in the valence band for both a compressive and tensile strain applied along the [110] direction. The effect is also non-negligible for a compressive strain along the [111] direction, but is for a tensile one.

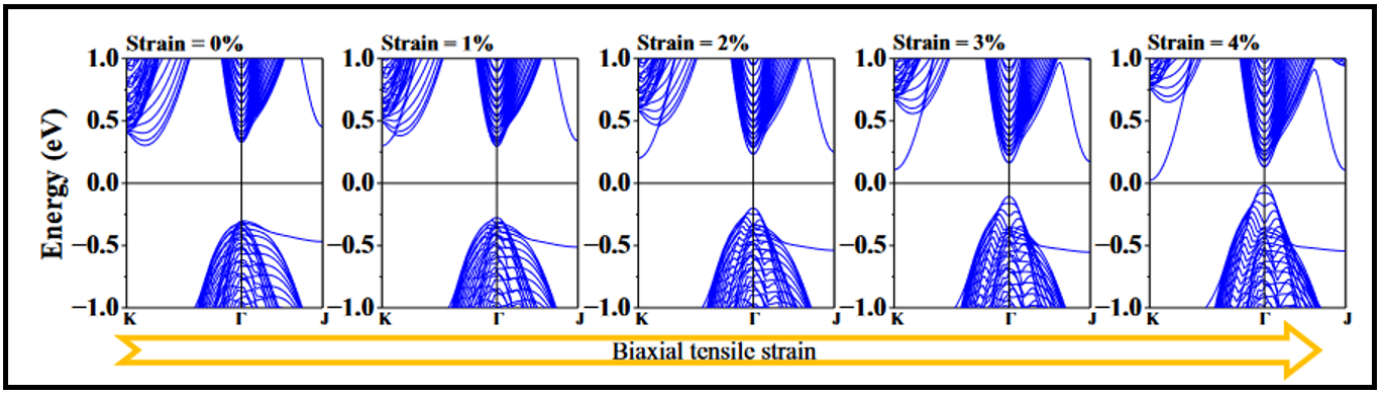


Figure 22: Impact of biaxial strain on the bandgap of 10 nm thick silicon pixels. Illustration from [1]

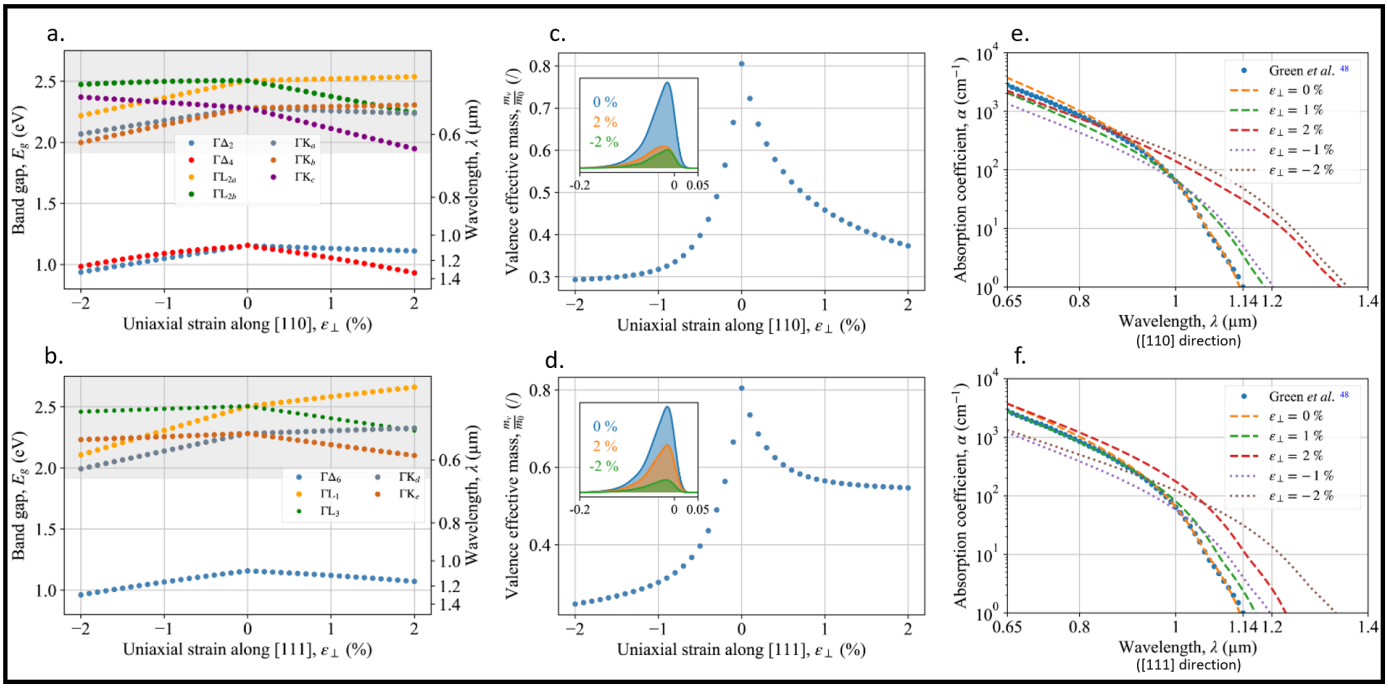


Figure 23: Impact of strain on the properties of silicon. **(a.),(b.):** Modification of the bandgap as a function of strain in the [110] and [111] directions, respectively. **(c.),(d.):** Modification of the effective mass as a function of the strain in the [110] and [111] directions, respectively. **(e.),(f.):** Modification of the cutoff wavelength as a function of the strain in the [110] and [111] direction respectively. These plots where given in an excellent article by Roisin N. [33]

These reductions will lead to an increase in carrier mobility if one relies on the math of subsubsection 2.1.2, which will in turn lead to an increase in device speed and performance.

2.3.3 Absorption coefficient

Combining the effects of the modification of the bandgap and effective masses, one can compute the variation of the absorption coefficient (the impact of the modification of the phonon frequency was found to be minor in [1] and is left here to the side). Again, Figure 23 (e.) and (f.) can be used as illustration.

On this plot, one can find not only the initial absorption limit of relaxed silicon (around 1.1 µm as one could expect), but also the absorption coefficient as a function of the strain for both crystallographic orientations.

For both orientations, a compressive strain leads to a considerable improvement of the cutoff wavelength of silicon, which is extended from 1.1 µm up to 1.35 µm for a 2% strain. Good performances are also obtained in the case of a tensile strain in the [110] direction but minimal improvements are observed for tensile strain in the [111] direction.

All in all, these results show that a considerable improvement of the cut-off wavelength is possible upon the application

of strain, allowing for detection of light in the SWIR region of the spectrum (recall: SWIR=[0.9-1.7] µm. Reached absorption in relaxed silicon = 1.1 µm. Reached absorption upon strain = 1.35 µm).

This improvement also means that thinner (and hence more flexible) silicon could be used for photodetection. Indeed, upon thinning, silicon becomes more and more transparent leading to a decrease of the absorption coefficient. This coefficient could be brought back up by applying strain. Furthermore, the thinner the silicon, the more compliant it gets hence allowing for higher strain levels. In an nutshell, one could compensate the downside of thinning by applying strain. One effect compensating the other.

2.4 Experimental validation of the impact of strain

This section is drafted based on [1]. Other sources will be mentioned explicitly

Now that the theoretical side of things has been covered, let's look at an actual test setup serving as an attempt to measure photo-currents under SWIR illumination. In an article called *"Breaking the absorption limit of silicon toward SWIR wavelength range via strain engineering"*[1] by *Ajit K. Katiyar* and his team, a 6x6 PN silicon pixel array was built on top of a flexible polyimide layer. This configuration allowing for the structure to be pressurised and illuminated with different wavelengths to measure the impact of pressure (and hence strain) on the absorption limit of silicon. Let's start by looking at the device built.

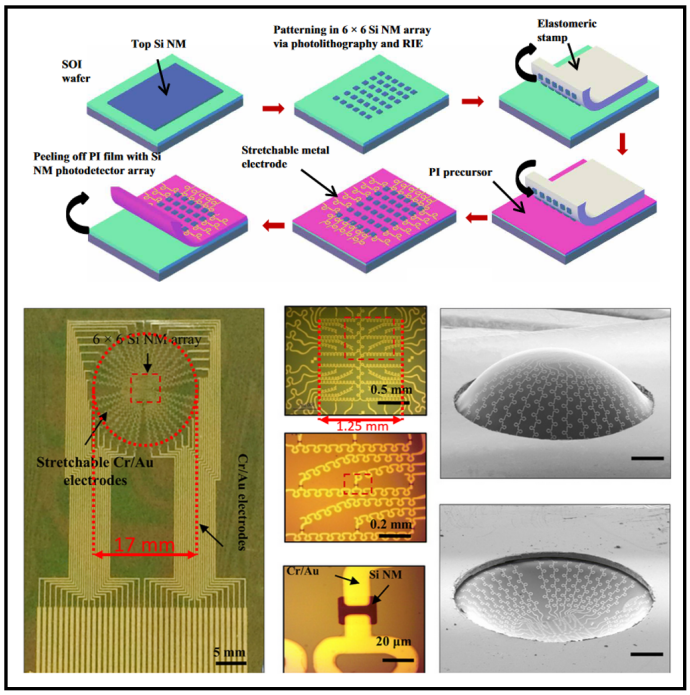


Figure 24: Fabrication process and final device (6x6 10 nm thick silicon PN pixels). Illustration by [1]

2.4.1 Fabrication process and membrane presentation

As was said in the introduction of this section, the fabricated device consists of an array of 6x6 silicon pixels built on top of a polyimide layer. The approach that was taken by the research team is the one of thin silicon allowing them to reach high levels of strain in the material without reaching a critical stress level in the silicon that would make it break (maximal stress before breaking for mono-Si is 7 GPa). Since the goal is here to induce a high level of strain in the material, the choice of polyimide is also not an accident. Indeed, as presented in subsubsection 1.3.2, polyimide features a Young's modulus of around 2.5 GPa meaning that the integrated silicon pixels will deform with it, which is the goal.

In addition, one can see in Figure 24, that the flexible interconnect approach was used. The inspiration for this comes from [37] by *R. Verplancke* and his team. Without going too much into detail, the team was able to build a set of flexible gold wires in the form of a horseshoe. This shape allowing the wiring to act like a spring and reversible strain of up to 100% were demonstrated.

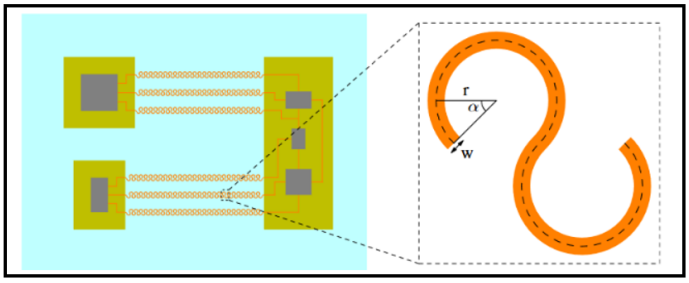


Figure 25: Horseshoe wire design by [37]. Illustration by [37]

Remembering what was said earlier in this state of the art, one might say that this approach is a mix between the island technique and thin film one. The final configuration allowing to build large membranes that should be able to deflect in a reversible way without breaking.

In terms of design, the silicon pixels are all individually addressed and the spacing between them (280 µm) was decided in order to be able to fit the wiring. The addressability as well as the number (6x6=36 PN pixels per membrane) of pixels comes from the fact that the research team wanted to use the membrane as complete photodetectors, hence requiring the ability to measure the photocurrents generated by each pixel individually. In addition, having individually connected pixels allows for a mapping of the levels of strain in different regions of the membrane.

The pixels themselves measure 20x20 µm and are packed at the center of a large membrane of 17 mm in diameter in a 1.25x1.25 mm square. It is worth noting that the 17 mm in diameter membrane is a rather unconventional size for a

membrane in the world of micro- and nanofabrication. Indeed, typical membranes are in the µm to 1 mm order. Such a large membrane is possible since it is built out of a thick polyimide layer membrane (>5µm), which is possibly doubled by commercial kapton (the article does not mention this). Note however, that a membrane of this size allows for a large deflection and hence a high induced strain in the silicon pixels.

In terms of process, due to the 10 nm thickness of the used active silicon, innovative techniques had to be used. Here is a presentation of the possibly used process based on the image provided by [1] and on the few details given in the article[5]:

1. **Thinning the active silicon from SOI wafer:** The pixels are built from lightly p-doped <100> monocrystalline silicon. The initial wafer was a standard SOI that went through a series of thinning step to get to the desired 10 nm thickness of the top silicon. The used process is a succession of thermal oxidation (see **??** followed by wet etching (section 6) of the silicon dioxide. The used oxidation technique is probably a dry process since the team reports that they can reduce the thickness of silicon using the oxidation/wet etch combination with a precision of 0.7 nm.

2. **Pixel patterning:** The pixels were then patterned using a conventional lithographic process followed by **R**eactive **I**on **E**tching (section 7).

3. **Pixel stamping:** In order to transfer the pixels to the polyimide membrane, the team of researchers decided to use an elastomeric stamping process. What they did was to cover the pixel array by a film of PDMS before etching away the bulk and BOX of the SOI wafer. This is made possible for two reasons:

   * **Selectivity:** According to [38], PDMS can be etched away using a 30% concentrated KOH solution, but must first be bombarded by protons to make it soluble to KOH. Since this is not the case here, the PDMS should not be etched by KOH. Furthermore, the etch rate of PDMS in BHF is of approximately 32 nm/min ([39]) whereas an unmasked layer of silicon dioxide can reach etching rates of up to 1 µm/min [40].
   * The second reason is the thickness of all the layers. For instance, as far as we know, the BOX layer might be thin in comparison with the thickness of the elastomeric stamp.

4. **Transfer on polyimide:** The most known form of polyimide is the cured industrial Kapton tape. However, this material exists in the form of a precursor where the polymer is dissolved in a carrier. This allows the material to be spin coated as a photoresist with a desired thickness according to the spin curve. Examples of such a product are the PI2545 or PI2611 by microchemicals. Once spin coated on a new wafer a soft bake was done to partially cure the polyimide and the pixels were laid down on it using the PDMS carrier. With the *"sandwich"* being complete, the entire assembly was baked for 5 more minutes in order to promote the adhesion of the silicon to the polyimide. Once done, the PDMS layer can be peeled and disposed of. The polyimide/Silicon assembly being baked one last time at 300°C according to the datasheet to fully cure the polyimide.

5. **Flexible wire:** Next comes the fabrication of the stretchable wires. This was done by a lift-off process. First, a layer of resist is patterned via a conventional lithographic process (Positive tone resist with a darkfield, for example). Metal (Cr/Au in this case) can then be deposited by thermal evaporation (to get a bad step coverage) before the resist is dissolved, leaving only the desired metallic wires.

6. **Final structure release:** The final device embedded in polyimide can now been peeled off. To ensure that the polyimide could be easily delaminated from the substrate several precautions must be taken such as not degassing the wafer or applying adhesion promoter prior to the polymer spin coating.

2.4.2 Experiments and results

Now that the design and the fabrication process has been introduced, let's take a look at the experimental setup. What the team did was to build a custom jig allowing to clamp the polyimide and to put the membrane under pressure using a flow of nitrogen. The control of the pressure allowing to control the level of strain applied to the silicon pixels. The pressurisation leads to the bulging of the membrane as can be seen in Figure 24 and Figure 26.

The team proposed multiple tests in the article but we will here focus our attention on two of them.

Experimental measurements of the modification of the photoresponse of silicon as a function of the wavelength and the pressure

[5]Doping steps not considered here. The article does not give details on the doping levels beside the intrinsic doping of the initial Si layer of the SOI wafer

In this first test, only one pixel of the matrix is connected to perform electrical measurements. The goal of this experiment is to measure whether or not the strained silicon would really feature an increased detection range extending in the shortwave infrared.

To this end, the membrane was placed in the pressure testing equipment and was pressurised at different levels between 760 torr (the atmosphere) and 1600 torr (2.1 atmospheres). In addition, for each pressure level, the membrane was illuminated with light sources ranging from 405 to 1550 nm (the 3.06 to 0.8 eV range).

As one could expect, when no pressure/strain is applied, the silicon does not output any photocurrent for the 1.31 and 1.55 µm of wavelength light sources. Indeed, both these sources feature photons with an incident energy smaller than the bandgap of relaxed silicon (1.12 eV). However, things start to become interesting when the pressure starts to increase and the membrane starts to be stressed.

The first thing that one can notice is that the membrane naturally takes a rounded shape [6] and hence communicates strain to the silicon pixels. The second one is represented on the plots of Figure 26.

As one can see on this plot by [1], a photocurrent starts to appear for illuminations above the cutoff wavelength of silicon when the strain starts to increase. The studied paper even reports detection of SWIR light all the way up to the 1.55 µm wavelength. This demonstration illustrates that the theory of the impact of strain is correct. Upon modification of the crystal lattice, the bandgap of silicon shrinks allowing for photodetection in the SWIR region. Furthermore, still on the same plot, one can see that the higher the strain the higher the responsivity. The definition used for the responsivity being: $R_{ph} = I_{ph}/P_{in}$ (P_{in} is the power of the incident light and I_{ph} is the generated photocurrent under a 2V biasing of the PN junction).

Strain levels up to 3.5% were reached in the silicon pixels as measured using Raman spectroscopy, which in other terms corresponds to a responsivity of 35 and 4 µA/W for an illumination under 1.31 and 1.55 µm wavelength light sources. These responsivities in the SWIR region remain low compared to the 44.21 mA/W reported under a 0.98 µm illumination, but still represent a breakthrough since this detection occurs for an energy 0.32 eV lower than what is normally required (1.55 µm corresponds to a photon energy of 0.8 eV which is 1.12-0.8=0.32 eV lower than what is normally required to cross the 1.12 eV width forbidden region of silicon).

The responsivity could however be improved by allowing a reduction of strain, hence making the silicon layer thicker and more opaque to light. Furthermore, a new design could be investigated like having a reflective layer under the silicon pixel, allowing the light to bounce back into the pixel.

Use of the membrane as a low resolution camera in the SWIR region

Now that a successful demonstration of breaking the absorption limit of silicon has been demonstrated, the team went ahead and devised a test setup to try to measure a shape using the pixel array. The test setup is illustrated in Figure 27. The idea is to shine a beam of 1.31 µm wavelength at a shadow mask that would allow light to pass through a Y shaped opening.

As one could expect, the shape does not imprint on the detector when the pixels are not strained. However, at a strain level of 1.8% (1200 torr), the image starts to become visible and only gets clearer as we reach the top pressure of 1600 torr (max strain of 3.5%). The illustration of the figure is a mapping of the generated photocurrent.

2.5 Bottom line

From all that has been presented in this (rather long) state of the art, we understand the numerous possibilities offered by the flexible electronics paradigm as an all. From here on however, the focus will stay on the topic of strain engineering and the work will follow the path of the paper that was just presented. The rest of this thesis will indeed showcase a design and microfabrication process that was undergone in order to build a matrix of strained silicon pixel inspired from the just presented article (this article might be called the *"reference article"* for the remainder of this thesis).

[6] Note that this rounded shape is a gimmick of a human eye. This non-planar detector is of interest in many optical applications since it would avoid the need to correct for the distortion induced by conventional planar CMOS photodetectors. The study of the impact of the shape of the membrane is one of the experiments made by the research group of [1] and is left to the reader's curiosity.

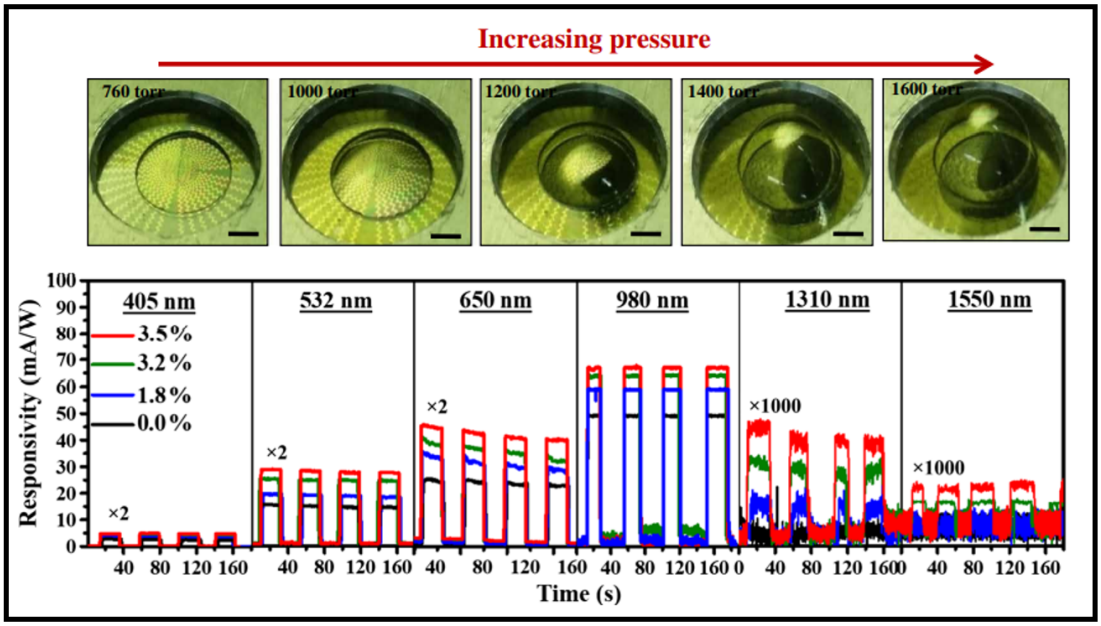


Figure 26: Modification of the photoresponse of silicon as a function of the wavelength and the pressure. Illustration from [1]

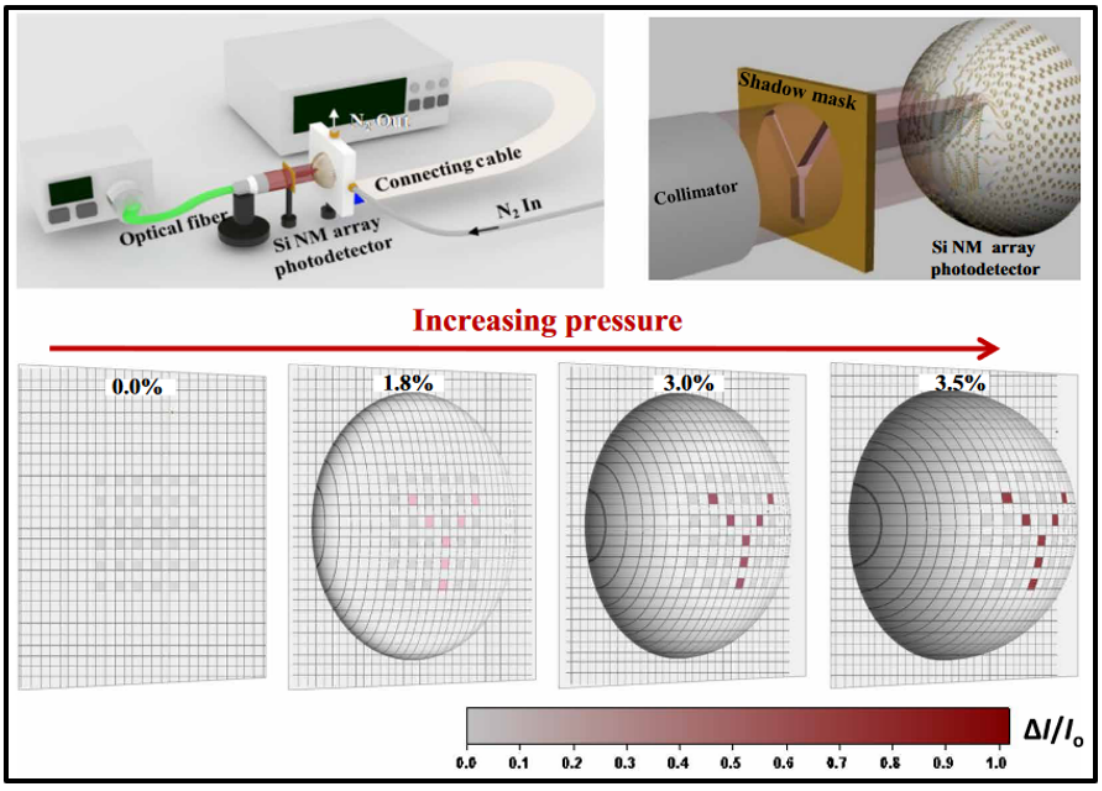


Figure 27: Use of the strained silicon pixel matrix as a camera in the SWIR region. Illustration from [1]

Part II
Design choices and presentation

Now that the situation has been assessed and that the main reference article has been studied, let's take a look at the design proposed in the thesis. This part of the document will focus on the design itself of the photodetector array and not on the process. The process aspect of the work will be discussed later on in Part IV. This section will hence describe the work that was undertaken to draw the lithography mask we had to buy to start our process. The illustrations available in this section will, for the most part, be screenshots taken from the freely available mask drawing software KLayout. The way masks are used will be discussed in the process part of this document. Obviously the preparation of a microfabrication process and the drawing of masks is done concurrently in practice since one can influence the other, but here we will separate the design from the process for the sake of clarity.

In this section of the document all the design choices that are "independent" of the process will be discussed. This includes the general dimensions and arrangement of the structures, the design of the arrays and the pixel themselves, the design of the wiring and more. These design discussions will therefore be the opportunity to present numerical simulations to back certain decisions on the thickness and sizes of the membranes (Comsol) and on the doping level of the pixels.

Since the basic design principles were strongly inspired by the main reference article used as a foundation for this work ([1]), parallels will be made in order to compare the final structures. Differences and similarities will hence be showcased here.

1 Design requirement

The first thing to do before jumping into the decisions we took was to lay down the requirements of our structures.

* Build a photodetector out of silicon on a flexible substrate to measure the impact of an applied strain on the optical performance of doped silicon
* Build different sized membranes to see the impact over the total deflection under pressure, measure the difference of stress transferred to the silicon pixels and increase the chance of having one sized membrane not breaking under the applied strain
* Build different versions of the pixels: NIN, PIN, plain N doped resistors and plain P doped resistors
* Have rectangular membranes to be bent in the four point bender and circular ones to be put under pressure.
* Have the pixel aligned along different crystallographic orientations
* Design in order to maximise the amount of tests doable with a complete processed wafer

2 General description of the final device

Before looking at each design decision in the dedicated sections, let's take a step back to look at the general design. We propose a set of membranes, either rectangular or circular, made out of a flexible polymer (polyimide in our case[7]). These membranes carry silicon pixels designed to work as strained controlled photodetectors. Each membrane, be it circular or rectangular, will form a die at the end of the process and a total of 27 dies are produced per wafer. Each of them features the following general elements:

* **A membrane:** The membrane, made out of a thin polyimide film, is what allows the final device to be flexible. The complete structures and substructure elements (wires, pixels and so on) are embedded in this polymer for later use and characterisation. Note that, initially, the structures are built on standard SOI wafer and that the polymerisation part and membrane release only comes at the end of the process. This means that the wafer can be diced once the structures are complete and that the dies can be stored for long periods of times before the membrane is released by etching the back of the wafer. This allows to easily store the samples without needing careful packaging to prevent damaging the structures/membranes. This will be further explained in Part IV of this document.

* **Silicon pixels:** Carried on the membrane are silicon pixels built out of 300 [nm] thick mono- or polycrystalline silicon. These pixels are doped to form PN or NPN junctions creating the photodetectors. Putting them on a flexible membrane will allow to perform strain engineering to tune their optical performances later on.

* **Large metallic pads:** These pads will be used to perform wire bounding or to probe the devices during the characterisation and testing phase.

[7] Also called PI

* **Stretchable electrical traces:** These traces are only featured on the circular membranes. They are designed in a "horseshoe" shape to allow them to stretch and not brake once the membrane is put under pressure.

* **Naming convention:** As one can observe in Figure 28, each die features different substructures. Hence, each of them is labelled to ease the characterisation process and to classify them. The naming convention will be discussed later.

The general dimensions of the final dies and the different features listed here above are illustrated in Figure 28. If one wanted to give a quick description of the produced structure one could describe them as: *Arrays of photodetectors on a polyimide thin flexible membrane.* However, as we will now see in the rest of this section, different variations of the structures coexist on a single wafer to allow maximising the number of different tests one could carry out using a single processed wafer. The reason for such a large variety of different designs is the prohibitive cost of microfabrication.

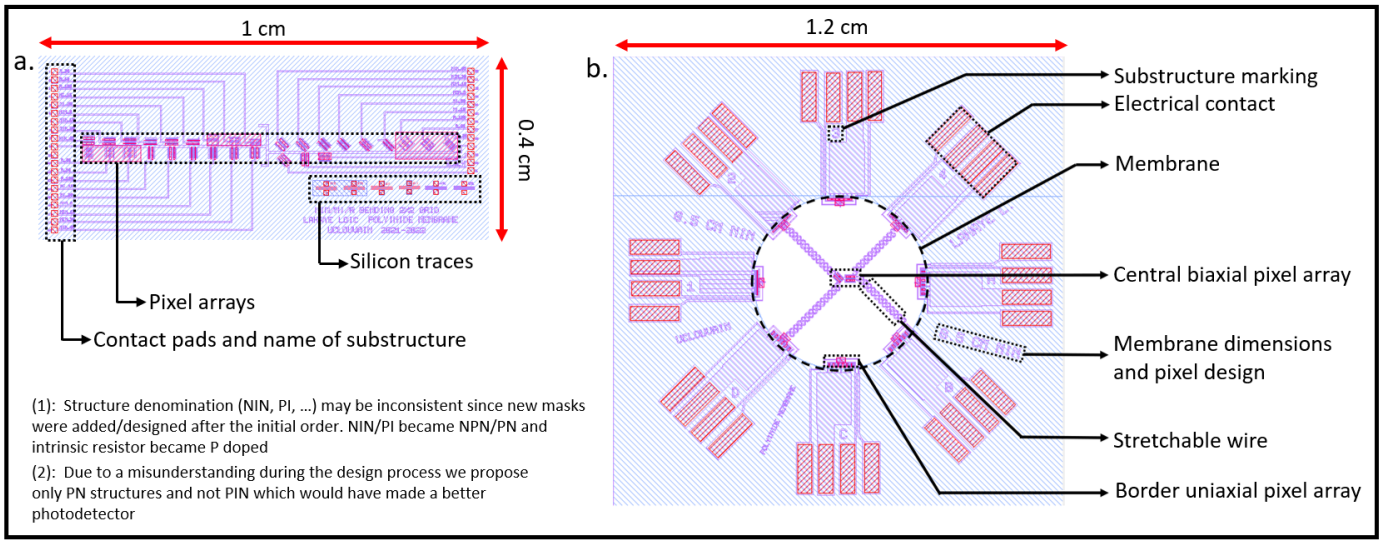


Figure 28: General presentation of the designs. **(a.)** Membrane for bending (compatible with SP1 process (see Part IV)) **(b.)** Membrane to be put under pressure (compatible with SP2 process (see Part IV as well))

3 Pixel design

Let's start the design of our structures by looking into the pixel themselves. The pixels featured in this work are of three types: P doped resistors, PN junctions and NPN junctions. They are build from SOI wafers that have a monocrystalline silicon top layer of 300 nm. This thickness was decided simply due to its availability in Winfab. In addition, it would have been risky to handle thinner layers of silicon here at the Winfab especially considering my low level of expertise at the beginning of the thesis. In addition, one might remember the oxidation and wet etching process used in the reference article to reduce the thickness of the top silicon. This technique was not used here since the procedure was not really known by the UCLouvain cleanroom staff and would have taken a considerable amount of time to implement. It was then decided, very early on in the design process, to base the work on SOI wafers featuring a top silicon layer of 300 nm and an initial doping level of 10^{15} atoms per $[\text{cm}]^{-3}$. It was believed from the beginning that a large membrane with pixels larger than the one featured in the reference work would be strained to a sufficient level to, I quote, *"At least see something as a proof of concept"*. We expected that half a percent of deformation could be achieved using the design explained below. This level of expected deformation will be challenged by the simulations done in Comsol in subsubsection 5.3.2.

3.1 The three type of pixels

As was said, three types of pixels are proposed: NPN, PN and resistors (either P or N doped). The design of the resistors is straightforward: they have the same dimensions as the PN and NPN junctions, but only have one type of doping. The level of doping being the same as the P and N levels of the NPN junctions. They will be used to measure the impact of strain on the resistivity. We expect this level to decrease due to the reduction of effective mass and hence increased mobility.

The design of the diodes and transistors however requires a bit more explanation.

3.1.1 The use of PIN and NIN structures as photodetectors

This section is drafted based on [35, 41, 42]. Other sources will be mentioned explicitly

The state of the art showed that illuminated silicon (under the correct wavelength if relaxed) can generate a photocurrent. However, using a simple piece of undoped silicon proves to be rather inefficient. Hence, devices have been devised to harness the photogeneration capabilities of silicon and improve the charge generation and charge collection. The two most common types of devices are the PIN and NIN junctions as illustrated in Figure 29. Let's start with the PIN.

The PN and PIN junctions

A PIN junction is designed by putting side by side a P-doped, an intrinsic and a N-doped region. It can operate in three different modes: photovoltaic, photoconductive or avalanche.

* **Photovoltaic mode:** In this mode, the diode is not connected to any external source and is hence unbiased. Upon illumination, carriers will be generated, as was seen in the state of the art, in the intrinsic region. Since the device is built as a PN junction, the presence of an internal electric field between the doped region will cause the charges (holes and electrons) to be collected in the P and N regions of the photodiode (holes collected in the P regions and the electrons in the N region since the field goes from N to P). This will generate a voltage at the terminals of the diode. This effect is the one used in photovoltaic cells.

* **Photoconductive mode:** In this case the diode must be reverse biased (application of a positive voltage on the cathode and a negative voltage on the anode). This reverse bias will increase the naturally present electric field between the N and P regions. This will lead to the increase of the depletion region between the P and N regions. The generated charge will hence have more chance to travel to the cathode and anode to be collected before being recombined in the lattice. One could say that this reverse bias increases the diffusion length of the charge carriers. They can travel further without recombining.

* **Avalanche mode:** If the reverse bias is increased compared to the the previous mode, the diode enters in the avalanche operating mode. This will lead to an increase in the generated photocurrent due to the breakdown of the diode. However, the exact phenomenon behind the avalanche mode is out of the scope of this thesis.

Note that here (and in Figure 29) the operation of a PIN junction is demonstrated. However, a classical PN junction can also be used as a photodetector but with worse performances. Indeed, as was seen in the state of the art, the generation of charge carrier upon illumination of silicon is an intrinsic process. This means that it does not depend on the dopant concentration. The doping of the P and N region offers the ability for the device to collect the charges more effectively. The issue with a PN junction, compared to the PIN, is that the generated carriers will have to travel a given distance in potentially highly doped regions before reaching the metallic contact pad. The high concentration of impurities reduces the diffusion length of the carrier by increasing the probability of them recombining before collection.

NIN junctions

In this case, the photodetector is built as a BJT transistor. As in the case of PIN, these kinds of photodetectors can exist in the form of an NPN with the same drawback as for the PN *VS* PIN case. Let's hence focus on the NIN case.

During operation, the collector is connected to a supply voltage and the output of the device will be on the emitter. The difference with conventional BJTs is that the current coming from the base is the result of photogenerated charges created upon its illumination. The advantage of phototransistors is that we benefit from the intrinsic gain β of the transistor as in a conventional BJT with: $I_c = \beta I_b$. Since this intrinsic gain is often in the region of $\beta = 50$ one can see that an NIN/NPN photodetector will provide the output current with natural amplification.

Note that in both the NIN and PIN cases the intrinsic region is designed to be large so that the amount of generated charges is maximised. This increase in dimension is the width of the photodevice (the 100 µm dimension in our case as depicted in section 3) and not the length of the region between the two N (in the NIN case) or between the N and P regions (in the PIN case). Indeed, the length of this intrinsic region must be chosen to be smaller than the diffusion length of the carriers in order for them to be able to be collected before recombining.

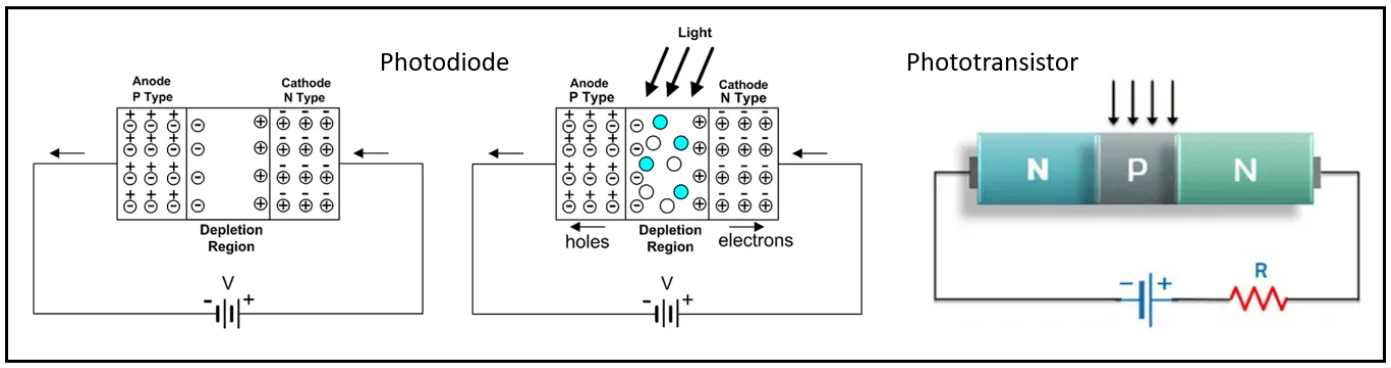


Figure 29: Schematic and operation of a PIN photodiode (illustration from [41]) and representation of a phototransistor (illustration from [42])

3.1.2 Type of photodevice for this master thesis

From what was just explained one can understand that the most efficient structures for photodetection would either be an NIN or a PIN kind of device. This is indeed true since they showcase better performances than conventional PN and NPN devices. This kind of configuration was the one that should have been chosen for this work, but a mistake was made.

As we will see later, six photolithographic masks come into play in this work. However, initially, only five were ordered. The initial plan was to build NIN transistors and NI junctions. But this design was problematic since that would mean that some of the metallic contacts would have been put directly on undoped silicon which would have resulted in the creation of a Schottky barrier and hence a bad electrical contact.

Schottky and ohmic contact

This section is drafted based on [43]. Other sources will be mentioned explicitly

As was seen in the state of the art, the doping of a semiconductor changes the Fermi level of a semiconductor. This level could either be shifted up in the case of an N-type semiconductor (hence closer to the conduction band) or down for a P-type semiconductor (hence closer to the valence band). This means that what we call the *work function* will not be the same for N- and P-type materials.

The work function is defined as the amount of energy required to extract an electron from a given material. From this definition, two types of contact arise: the Schottky or Ohmic contacts:

* **Schottky contact:** This type of contact appears when the work function of the metal is higher than the work function of the semiconductor ($\Phi_m > \Phi_{semi}$). To illustrate this let's look at Figure 30 (c.) where metal is used to contact a lightly doped N-type semiconductor. When the two materials are put in contact, the Fermi levels must align. In this case, the electrons in the conduction band of the semiconductor will flow to the metal to occupy the empty energy state in the metal above the Fermi level. This flow of electrons is possible because the energy level of the conduction band of the N-type material is higher than the highest energy level in the metal. These electrons leaving the semiconductor will leave behind fixed positive charges creating a depletion region or in other words: a diode. An electric field is created between the metal and the semiconductor and the energy band of the semiconductor curls up in the direction of the electric field. This creates a potential Φ_B that the electron must overcome to go back into the semiconductor, which is not possible without the application of an external field. The flow of electrons without biasing the junction is hence impossible. A solution to avoid getting this Schottky diode is to heavily N-dope the material. The metal will be able to suck some mobile carriers but not all. This high doping level will reduce the length of the depletion region but not its height. The carrier being able to pass by tunneling effect through the barrier (but not above).

* **Ohmic contact:** In this case we have $\Phi_m < \Phi_{semi}$. This type of contact, illustrated in Figure 30 (d.) occurs (for example) between a P-doped material and a metal. In this case, since the work function of the semiconductor is lower than the one of the metal, the energy level of the conduction band is lower than the energy level of the electrons in the metal. This leads to a natural migration of electrons from the metal to the lower energy level of the conduction band of the semiconductor. The region near the interface sees an accumulation of mobile carriers, which in turn reduces the resistivity of the junction and hence creates an Ohmic contact that features the same linear behavior as a resistor.

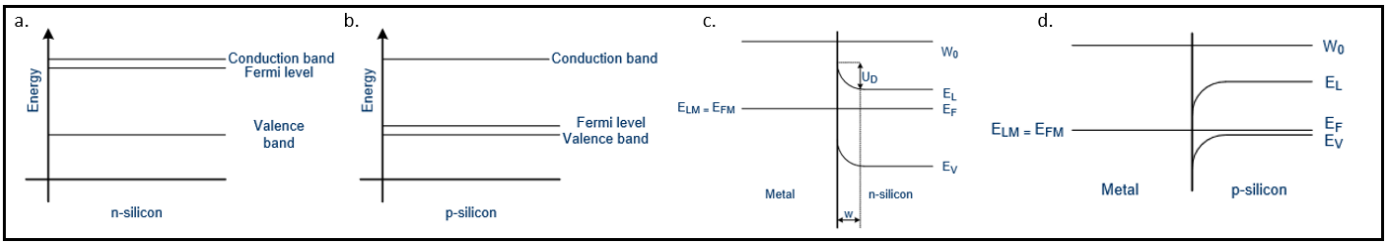


Figure 30: Impact on the energy band of putting a metal in contact with a semiconductor **(a.)** Fermi level of N-doped silicon **(b.)** Fermi level of P-doped silicon **(c.)** Schottky barrier between a metal and N-doped silicon **(d.)** Ohmic contact between a metal an P-doped silicon. Illustration from [43]

The solution, the mistake and the final photodevices

As was said, the initial plan was to go for NIN and NI junctions. From the beginning, the N-doping was set to be at saturation ($> 10^{21}$ phosphorus atoms per $[cm]^{-3}$) hence a good contact was predicted by tunneling effect. However, laying our aluminum contacts on an undoped silicon region was guaranteed to give us a Schottky diode[8].

We solved this issue by designing a new mask to dope P some regions of our pixels. And here is the mistake: Due to a misunderstanding that happened during the drawing phase of this new mask and that extended during its review phase, **ALL** the intrinsic regions of the pixels were set to become P-doped.

This means that not only the NI junction became PN instead of PIN but also that the NIN junction became NPN. This is a shame since the P regions of the NPN phototransistors are not even contacted.

The issue with this design error is that no more intrinsic region exists which, as we saw, makes for worse photodetectors compared to the ones that feature an intrinsic region.

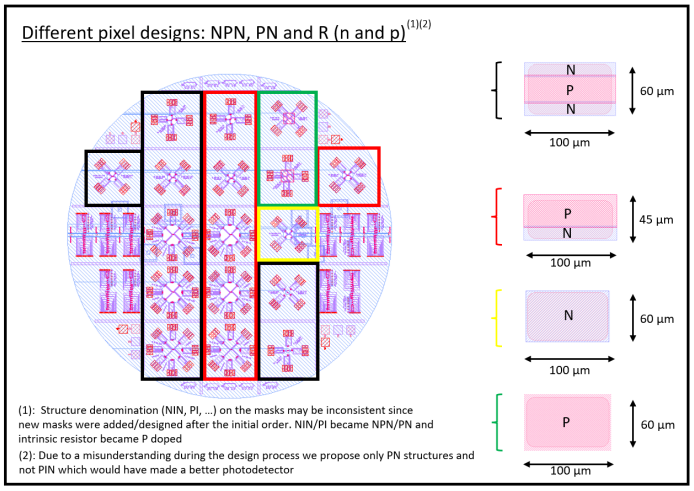


Figure 31: Different pixel designs namely NPN, PN and R_p and R_n with the location of their dedicated structures on the wafer. The rectangular dies carry all the pixel designs with alternative dimensions. For the NPN and PN devices on circular membranes (hence with standard and fixed dimensions), all N region are of 15 µm in length and the P region is of 30 µm.

The configurations, in terms of doping, of our pixels are shown in section 3 and in Figure 32. As one can see, the pixels are of the NPN and PN type (+ P and N plain doped resistors). The location of each pixel design on the wafer is shown using the color scheme. Note also that in addition to the standard dimensions of the pixels featured on the membrane, each rectangular strip has a variety of size of the P region for both the NPN and PN junctions.

The dimensions of the pixels result from a compromise between having a large enough region to generate carriers upon illumination, having enough silicon to reach a sufficient level of strain with smaller membranes (see the section about the membrane size), as well as having a large enough design to avoid needing a high (and hence hard to reach) resolution during the fabrication process.

The internal dimensions were fixed to be in coherence with the external dimensions. The decision was made early in the design phase to have the length of the charge generating region to be of 30 µm. Hence, as one can see in section 3, the dimensions (in length) are: 15/30/15 µm for the NPN stacks and 15/30 µm for the NP stack.

3.2 Doping

Now that the structures for photodetection have been determined, it is time to decide the doping level. The goal of this section is not to revisit or to repeat the physics behind the doping of a semiconductor, but to look at it in a more practical sense. Indeed, the dopant concentration will fix many parameters of our device and hence determine the final performance of our detectors. Let's take a look at the different challenges one by one before settling on a desired dopant concentration.

[8]The I region was not really intrinsic. The SOI used are lightly P-doped (around 10^{15} boron atoms per $[cm]^{-3}$), but this level of doping would result in a Schottky diode.

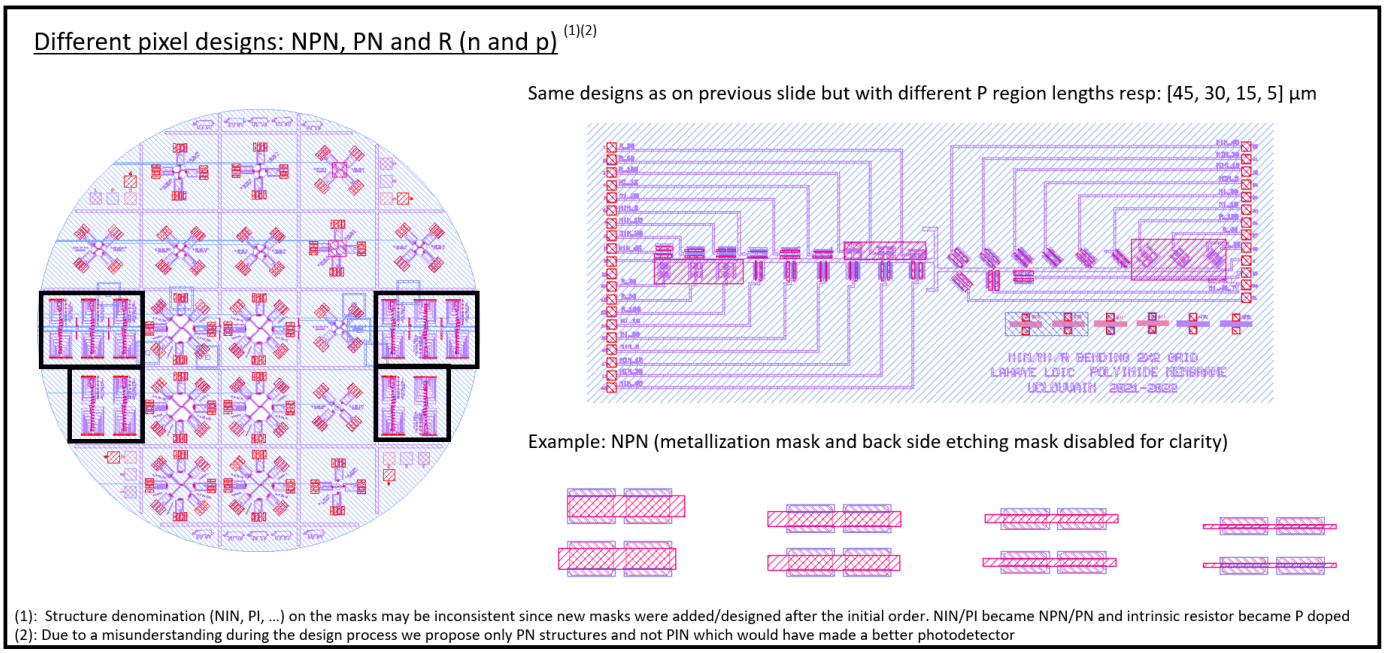


Figure 32: Variation of the length of the P regions of the pixels on the rectangular dies

Note that the numbers given in this section where obtained using an online calculator proposed by the website *pvlighthouse*. This calculator allows to compute the resistivity and diffusion length of charges inside an N- or P-doped material for any given concentration of the dopant.

Note that the decision on the doping level of the P-region will be made complicated due to the design error. Indeed, since all the charge generating regions were *in fine* drawn to be P-doped, the dopant concentration of those P-region could not simply be set at saturation level. This will be explained in details here below.

3.2.1 Optical performance

Let's remember that the ultimate goal of our devices is to act as photodetectors. This means that we want charges to be generated upon illumination and collected by our metallic contacts.

If the design had been made without any mistakes the photodetectors would have been of the NIN and PIN type. This would have made the decision on the doping level easy since both the P and N region could have been doped to saturation. Indeed, the charges would have been generated in an intrinsic region making the probability of them recombining small. This translates in a long diffusion length of the carrier and hence increases the amount of charges collected in the P and N region of the NIN or PIN pixels.

However, due to the mistake, the charge generation will occur in a P-doped region. This is problematic since this means that one cannot simply dope the P region to saturation since that would greatly reduce the diffusion length of the generated charges.

This means that the doping level of the P-region cannot be set to saturation. In addition, the P-region cannot be left at their initial[9] *"intrinsic"* doping level since this would lead to a high final resistance as well as a Schottky contact. Hence a compromise has to be made. Let's start by looking at the diffusion length and depletion length.

Diffusion length of the carrier

The most critical parameter is the diffusion length of our charge carrier. Indeed, it does not matter how many charges are generated upon illumination, if they all recombine before reaching a metallic contact this is all for nothing. Since the region tasked with the generation of charges is doomed to be P-doped, a saturation doping level cannot be considered.

Indeed, in practice, when a charge carrier is generated it must still be transported from its generation site to the metallic contact. On its path the mobile charge might encounter impurities and hence recombine and be annihilated. The length that a photogenerated charge, be it an electron or a hole, can travel depending on the dopant concentration and can be

[9]The silicon of the SOI used for the process is doped at a 1e15 atoms $[cm]^{-3}$ level from the factory

expressed as:

$$L_p = \sqrt{D_p \tau_P} \quad \text{with} \quad D_p = \mu_p \frac{kT}{q} \quad \text{and} \quad L_n = \sqrt{D_n \tau_n} \quad \text{with} \quad D_n = \mu_n \frac{kT}{q} \tag{3.1}$$

In these equations L_p and L_n are the lengths that a charge can travel before recombining. These lengths are function of the bulk lifetime of the carriers τ (in seconds) and the diffusion coefficient D expressed in cm^2/sec.

The diffusion coefficient is a function of the charge mobility and the temperature. On the other hand, the bulk lifetime τ is the parameter which is function of the dopant concentration. This indeed makes sense since the higher the concentration in impurities, the higher the risk for a charge carrier to run into a recombination site and be destroyed. For example, the diffusion length of carriers for a P-doped piece of silicon at the 1e15 and 1e21 levels are respectively of 710 and 0.0585 µm. The bulk lifetime of a carrier being much smaller in highly doped silicon.

From here, one can start to understand how we might proceed to determine the maximal doping level allowed for our P region. Let's consider the case of the PN junction. Remembering that the P and N regions measure respectively 30 and 15µm, in the worst case scenario, a generated charge should travel 30 µm before being picked up by a metallic contact. Hence, the diffusion length of the carrier should be of at least 30 µm. This illustrates perfectly why doping our P region at the 1e21 (saturation) level cannot be considered. Indeed, the 0.0585 µm diffusion length makes it virtually impossible for the charge to travel 30 µm without recombining. In a nutshell, we want our diffusion length to be larger than the 30 µm of the P region.

In addition, the recombination process follows an exponential law given by:

$$\% = \left[\exp \frac{\Delta X}{L_r}\right]^{-1} \tag{3.2}$$

Where δX is the distance traveled by the generated charge from the generation site and L_r is the recombination/diffusion length. This expression tells us that the percentage of recombination increased dramatically once the traveled distance starts to exceed the diffusion length. It also shows that recombination is less likely but not impossible for travel distances shorter than the diffusion length.

It was **arbitrarily** decided in the design phase of our devices that at least 70% of the charges should be collected. Considering the worse case scenario where the charges should travel the complete 30 µm through the P-doped region hence having $\Delta X = 30$ µm we have:

$$70\% \leq \left[\exp \frac{\Delta X}{L_r}\right]^{-1} \Leftrightarrow L_r \geq \frac{\Delta X}{\ln (70\%)^{-1}}\Big|_{\Delta X = 30\mu m} \geq 84\mu m \tag{3.3}$$

From that estimation, and using the previously introduced calculator, one can see that the maximal P-doping level allowing to achieve this 70% charge collection would be of 1e18 boron atoms per $[cm]^3$ (see Table 1). Any doping level smaller or equal than that should yield a sufficient diffusion length.

Note that the complete discussion here was only concerned with the concentration of boron, without considering the concentration of the N-type doping atoms. This is because the charges will be generated only in P-type regions and never in an N-type one. The level of doping of the N regions has no influence on the charge generation and could in theory be set to saturation which is indeed what we did as we will see later.

	1e15	**1e16**	**1e17**	**1e18**	**1e19**	**1e20**	**1e21**
Diffusion length (For boron (P type))	710	554	344	82	7.8	0.6	0.058

Table 1: Diffusion length in µm as a function of the concentration in boron atoms

Depletion length

When a P and a N region are put into contact, the mobile carriers are going to migrate. The excess electrons from the donor atoms N_D will migrate to the P side and the excess holes from the acceptor electrons will migrate to the N side. This phenomena occurs spontaneously to equalise the carrier concentration. By this process the energy level will curve and align. A region called the *depletion length* is created. This region features a space charge resulting in the creation of an electric field across the PN junction. This internal electric field is the direct consequence of the created space charge near the interface. Indeed, the N region will feature a positive charge (depletion in electrons/negative charge carriers) and the P side a negative charge (depletion in holes). The width of this depletion length is the sum of the depletion lengths in both the N and P region which are themselves function of the concentration in donors and acceptors:

$$l_p = \sqrt{\frac{2\epsilon}{q}\frac{N_D}{N_A + (N_A + N_D)}(\Phi_0 - V)} \qquad \text{and} \qquad l_n\sqrt{\frac{2\epsilon}{q}\frac{N_A}{N_D + (N_A + N_D)}(\Phi_0 - V)} \tag{3.4}$$

These equations make use of what we define as the internal potential or contact potential. This potential is the difference between the potential *"far"* in the P and region: $\Phi_0 = \Phi_{no} - \Phi_{po}$.It can be computed as:

$$\Phi_0 = \frac{kT}{q}\ln\frac{N_A N_D}{n_i^2} \tag{3.5}$$

The study of this depletion length is of interest when creating a photodetector for two reasons. The first one is that this depletion length must be smaller than the diffusion length of our photogenerated charges to avoid total recombination. This can be checked by computing the lengths of the depletion length. In our case, considering $q = 1.6 * 10^{-19}$ [C], $\epsilon = \epsilon_0\epsilon_r = 8.85 * 10^{-12} * 11.68 = 1.03 * 10^{-10}$ [F/m], one can get values for the depletion length given in Table 2 (@300 K and V=0).

One can see that the depletion length is in all cases smaller than the >1µm order of magnitude diffusion length found for similar levels of dopant. One can see from these values that the depletion length decreases as the concentration in dopant increases. Indeed, since the excess in charge carriers increases with the level of doping, contacting a P and N region will still result in a diffusion of carriers, but the charges will not need to come from as far in the material to equalise the concentrations and energy levels.

Finally, note that increasing the dopant concentration will result in a stronger internal electric field between the N and P regions. This stronger field facilitating the acceleration and pick up of photogenerated charges and hence their collections.

	1e17/1e21 PN	**1e18/1e21 PN**	**1e19/1e21 PN**
Total depletion length in µm	0.11	0.037	0.012

Table 2: Total depletion length as a function of the concentration of dopant (300k and V=0).

3.2.2 Resistivity

The second thing to consider is the impact of doping on the resistivity. Indeed, adding a dopant will generate, upon its activation, an excess of mobile charge carriers which in turn will lower the resistivity of the material compared to its intrinsic state (this is detailed in subsubsection 2.1.2).

In our case the goal was to lower the resistance of our pixels to be in the [1-100]$k\Omega$ range. This is in order to not have the testing equipment max out during the characterisation phase, especially if we want to measure the impact of strain on carrier mobility by measuring the resistors featured in our design. The final resistance can be computed as:

$$R = \frac{\rho l}{A} \tag{3.6}$$

where ρ is the resistivity of the material. This quantity depends on the concentration in charge carriers as was demonstrated in subsubsection 2.1.2. The other quantities in the equation, A and l are respectively the length (60 µm) and cross section (100x0.3 µm) of the considered resistor.

Table 3 here below summarises the total resistance across our pixels (P- and N-doped resistors). As one can see, as the doping increases the resistance decreases. Since the target is to have a resistance in the $k\Omega$ range one can see that the minimal concentration of dopant should be of 1e17 (atoms [cm]−3) for both phosphorus and boron.

This consideration on the resistance of our pixels narrows down the possible range of dopant concentration. Considering the depletion length alone we saw that we should have a P-doping smaller or equal to 1e18. Considering the need for the total resistance to be between [1-100]$k\Omega$ the possible doping value now lies in the [1e17-1e18] atoms of boron per [cm]^3 range.

As for the N-region, still considering a doping to saturation, one gets a total resistance of 0.62$k\Omega$ for the pixel dimensions given in section 3.

	1e15	1e16	1e17	1e18	1e19	1e20	1e21
Boron (P type)	27M	2.92M	412k	96k	19k	2.5k	0.51k
Phosphorus (N type)	9.1M	1.01M	177k	51k	11.7k	1.7k	0.62k

Table 3: Resistance of a 60x100x0.3 µm block of doped silicon as a function of the doping level and the type of impurities (60 is the length of the resistor and 100x0.3 is the cross section)

3.2.3 Metallic contact (Aluminum on silicon)

As was previously explained, the metallic contacts should be placed on doped material to avoid the Schottky barrier. As was seen in the section explaining the problematic of contacting silicon, one way to avoid the Schottky barrier is to dope the material either N++ or P+. Once again, considering the fact that the wanted doping level of the N-region is 1e21 (hence N++), a good metallic contact should be obtained thanks to carrier going through a very high but thin potential barrier via the tunneling effect. The problematic is hence, once again, on the level of the P-doping.

	1e17 P	1e18 P	1e19 P	1e21 N
Final resistance of a 60x(100x0.3) [µm] doped silicon pixel	27M	2.92M	412k	0.6k
Expected Schottky barrier (Al on doped Si)	Yes	No	No	No
Total depletion length in µm (Combination of the given P level and 1e21 N level)	0.11	0.037	0.012	NA
Diffusion length in doped silicon in µm	344	82	7.8	0.012

Figure 33: Summary of the expected parameters of our doped silicon

Contrarily to the previous two parameters (diffusion length and resistance), predicting if contacting a P-type semiconductor (by looking at the doping level) will result in an ohmic contact is not as straightforward as using a small equation. The math behind this phenomenon being more complicated, the problem was solved by relying on the experience of Nicolas André.

In the past, he had already been confronted with the task of P-doping silicon to make an ohmic contact. Figure 34 shows the doping curves obtained by a Silvaco simulation for two different implantation parameters. The figure on the left shows a maximal dopant concentration of a bit less than 1e18 atoms $[cm]^{-3}$ and a concentration of 1e17 atoms $[cm]^{-3}$ at the surface of the material[10]. This level of impurities was insufficient to reach the ohmicity. However, the 1e19 and 1e18 atoms $[cm]^{-3}$ as illustrated by the figure on the right for the maximal and surface concentration, respectively, did yield an ohmic contact.

This final consideration allows us to settle on a doping level of 1e18 atoms $[cm]^{-3}$ for our P-regions. A summary of the expected properties is given in Figure 33.

3.2.4 Implantation parameters

Now that the doping levels have been determined (see the summary in Figure 33, it is time to look at the implantation parameters. The *implantation* of a semiconductor is the process during which impurities are introduced into the material in order to dope it. An implantation process depends on four main parameters:

* **The type of impurities:** Probably the most obvious parameter of all to consider for an implantation process. As was already explained before, the impurities to introduce into a semiconductor will determine its electric behavior later on. These types, called *species*, can either dope the material P by capturing an electron and generating a mobile hole or N by releasing an extra electron. Two common species are boron and phosphorus for P and N doping, respectively.

[10] **Clarification:** Silvaco has a strange way of defining (or at least representing) the origin of the X-axis. The vertical line on the plot somewhere between 0.04 and 0.08 represents the separation between the thin layer of thermal oxide and the silicon. The concentration curve on the left of this line should not be considered for two reasons. The first being that it represents the concentration in the oxide layer on top of the device and second because the definition of the mesh is too rough for an accurate computation in this region. Indeed, the layer of thermal oxide covering the silicon for the implantation step has a thickness of around 27 nm. As one can see, only three points are computed for this region. In any case, only the curve on the right of the vertical separation should be considered. The *"Surface concentration"* being the level on the separation line at the beginning of the *"right"* curve and the maximal value also being on the right part of the plot

[11] Postdoc researcher at UCLouvain

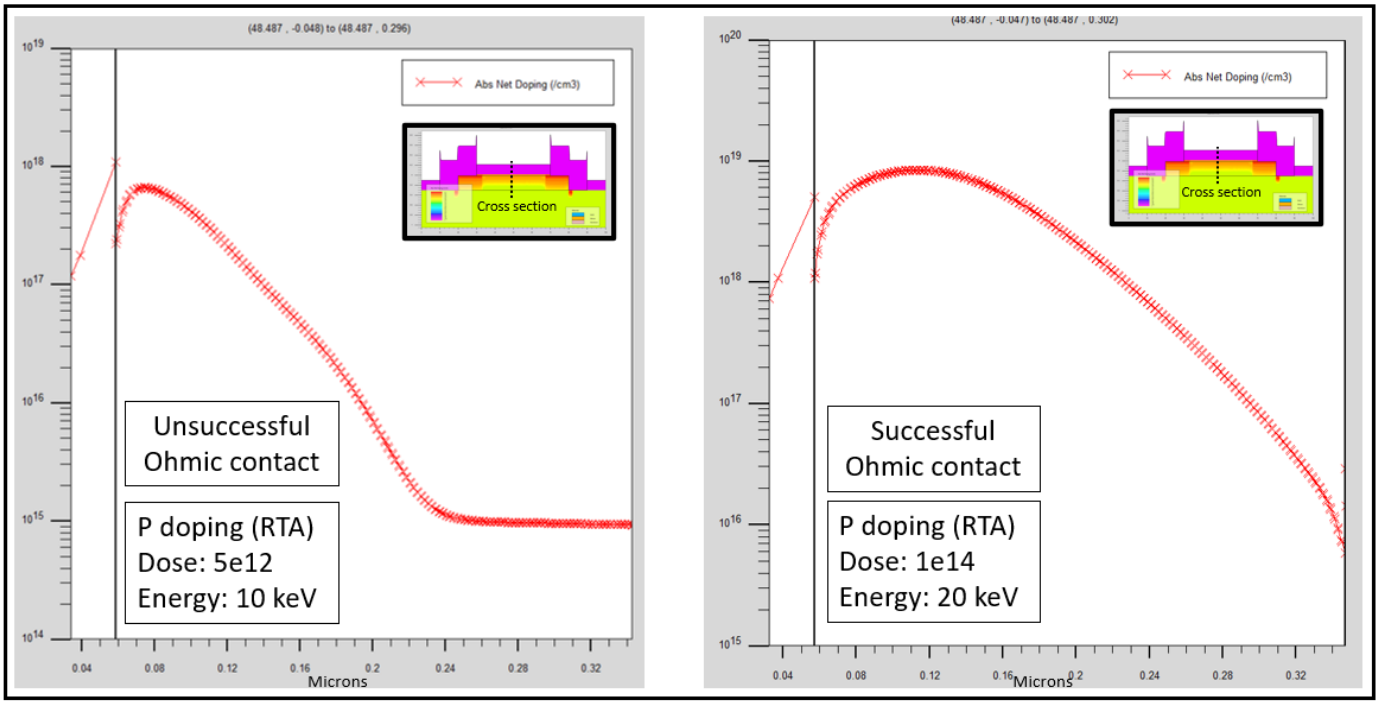


Figure 34: Boron atoms concentration as a function of the device depth proposed by Nicolas André[11] to try to achieve ohmicity of the electrical contact (silicon aluminum)

* **The dose:** During the implantation itself, the sample is placed in a chamber and bombarded with a given amount of the atoms of the desired species by $[cm]^2$. The dose determines this amount of atoms
* **The energy:** In order for the atoms to penetrate into the lattice one should not simply drop them onto the sample like a handful of sand dropped on the floor. The atoms must be shot, like bullets, toward the material to penetrate it and enter the crystal lattice. This is the implantation energy.
* **Annealing:** Adding impurities to the material is however not sufficient to consider the device doped. Indeed, energy must still be brought to the impurities to activate them and generate the charge carriers. This energy is brought to the atoms under the form of heating. Upon heating, two distinct phenomena will occur. The first one is that the impurity will replace a silicon atom inside the crystal lattice and the second one is that the thermal energy will allow to ionise the impurity as was explained in the state of the art. This is generally done by a process called **R**apid **T**hermal **A**nnealing or RTA. A typical RTA consists of heating the implanted semiconductor to a temperature up to a thousand degree centigrade for one or two minutes. The short time during which this process is done prevents the impurities from diffusing into the material and make the final electrical proprieties and dopant concentration shift to the desired value. Note that, after the process of activation of dopant atoms, steps involving high temperatures (>650°C) should be avoided.

These parameters should be decided and communicated to the person/company that will be tasked to implant our wafer. One can not simply reach out to IBS (**I**on **B**eam **S**ervice the french compagy taskey to implant our wafer) and ask them for a specific level of impurities.

Since the doping profile is hard to calculate by hand, one must rely on numerical simulation to fix the energy and dose for their implantation. For this thesis, such computations were done using a software of Silvaco. Silvaco is a company well known for producing EDA tools for the semiconductor industry. Note that the curves represented in Figure 34 were obtained through Silvaco.

In order to obtain the concentration profile, the pixels had to be designed in **DeckBuild**, where the design is done by describing the device as well as the process steps by lines of *"code"*. For example, in our case, the decision was taken to model our NPN pixels.

In practice this is done by first telling the software the material with which we start. In our case: monocrystalline silicon doped at a 1e15 level with boron atoms. Then, once the mesh and initial dimensions are given to the tool, one can describe the process. Since the process written in Silvaco is the same as the one done in practice we won't go over it here but note the following consideration:

* The dimensions of the modeled device are the one of the NPN pixel proposed in this thesis. The dimensions of the N and P region being the same as what was drawn on the mask (15/30/15 µm).
* The stack is the same as in reality. We have an oxide with a 300 nm thick layer of mono-Si on top.
* All the process steps are given to Silvaco including oxidation, metal depositions, resist deposition, etching, annealing, ...
* This was implicitly implied but, if one wants to implant a specific region, a resist mask should be build on top of the structure to prevent the impurities from penetrating where we don't want them to.
* The implantation process always takes place across a layer of thermal oxide. In practice, this oxide has a thickness of 27 nm (it is the target at least). This oxide was grown in Winfab prior to implantation and was modeled in Silvaco. A second role for this oxide, and probably the most important one, is that it reduces the incident energy of the atoms arriving on the sample to be doped. This is done to avoid having the impurities go through the complete thickness of the silicon.
* The RTA furnace of Winfab was broken at the time at which the annealing should have been done. Hence a longer annealing process was carried out in the Tempress[12] at 850°C for 30 minutes. This should provide enough energy for the impurities to ionise and replace silicon atoms in the lattice. The downside of this technique is that it makes the impurities diffuse more in the material, but this was accounted for in the simulation.

Once the device is described and the process (up to the moment of implantation) is completed, it is time to simulate the implantation. Like in practice this is done in three phases. First implanting the boron atoms, then the phosphorus atoms (or in the reverse order) and finally the annealing.

By playing with the dose and energy parameters of the implantation we shape the dopant profile inside of our semiconductor until we get the desired curve.

The final profile on which we settled, as well as the implantation parameters, are given in Figure 36. In addition, a view of the doping is shown on each plot.

The implantation parameters for the N type region were straightforward to obtain since the goal was to reach saturation. A high dose and energy level could hence be used. The result is shown on the right side of the figure.

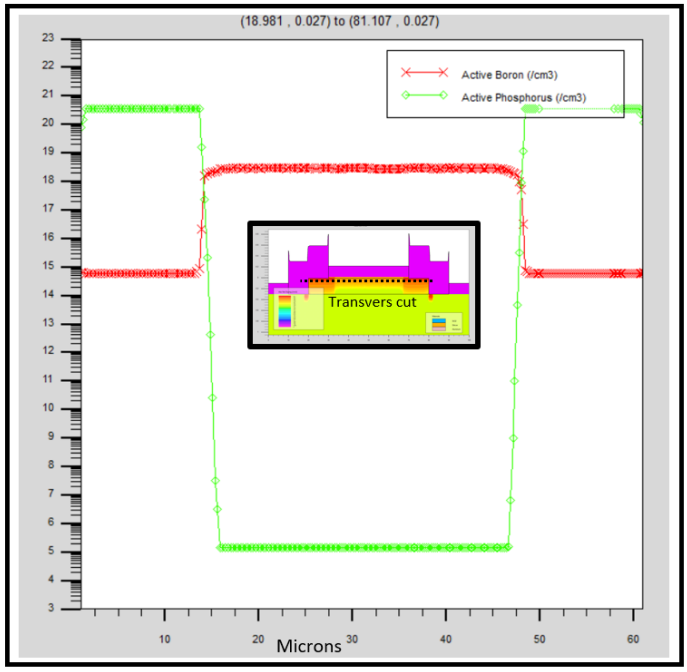


Figure 35: Doping profile inside of an NPN type pixel. Transversal cut/view

However, fine tuning of both the implantation dose and energy were required to reach this profile for the P region. It is worth noticing that this curve takes inspiration from the work of Nicolas André, but by still considering the analysis made previously. Indeed, the curve from Nicolas André (and hence the associated implantation parameters) could not have been used in our case, because the 1e19 maximal dopant concentration would have killed the photoelectric effect by reducing the diffusion length of the carrier too much. The final curve is hence a trade-off between ensuring an ohmic contact and not ruining the photogeneration capabilities of our pixels.

The curve hence features the same doping level of 1e18 boron atoms per $[cm]^3$ at the surface, but a lower maximal concentration inside the semiconductor of around 4e18 boron atoms per $[cm]^3$. This curve is believed to be a good compromise between good contact and good electrical and optical performances.

Finally, as one can see in Figure 35, the longer annealing process at 850°C instead of a conventional RTA did not lead to any noticeable diffusion. The boron and phosphorus concentration curves crossing at the 15 and 45 µm mark.

[12] A furnace that was used to grow oxide before being replaced by the Koyo. Its main purpose is now to degas wafers

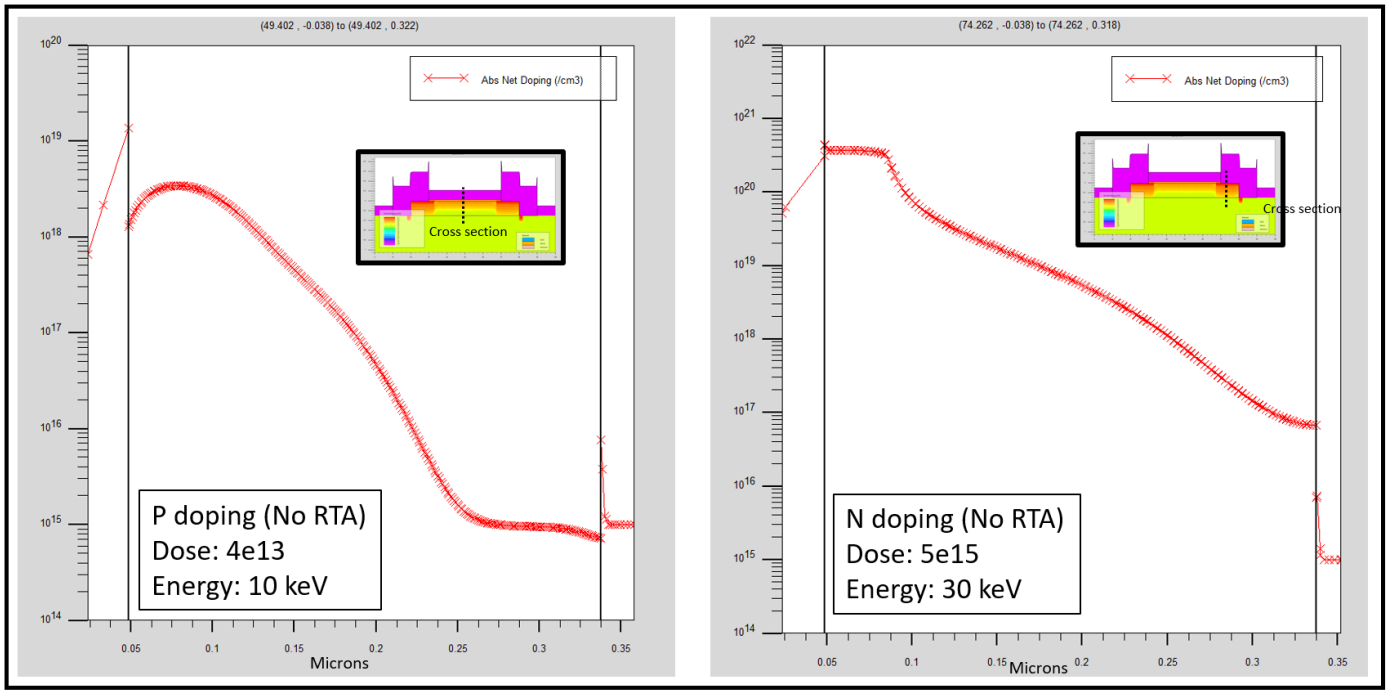


Figure 36: Final doping parameters and profile inside the silicon. P doping on the left and N doping on the right.

3.3 Pixel configuration

Another thing to consider is the way the pixels are positioned on the membrane. Their positioning and configuration can be analysed in two different ways.

Strain point of view

The pixels on the membranes are either positioned on the edge or at the center. The reason behind it is that we want to be able to look at the impact of both uni- and biaxial strain on the optoelectronic performances of our NPN and PN junction. Furthermore, the pixels on the edge of the membranes, and hence under an uniaxial load, are positioned to either endure transverse or longitudinal deformations. Each substructure on the edge of the membrane also feature two P doped resistors for strain measurement. Both the uni- and biaxial configuration can be seen in Figure 37. Note that the pixel under transverse uniaxial load are labeled from 1 to 4 while the ones for longitudinal uniaxial load are from A to D.

The impact of each of these configurations will be analysed in the section about the Comsol simulations.

Crystal orientation point of view

As was seen in the state of the art, the crystallographic orientation has a non negligible influence on the performance of strained silicon. Hence, our pixels are placed on the wafer to be aligned either along the <110> or <100> sets of crystal orientation (the wafer having a top silicon layer in the {111} crystal plane). This is shown in Figure 38.

Electrical connections point of view

Contrarily to what was done in [1], our arrays of pixels are not designed to be individually addressed. Indeed, they are arranged in clusters (either 2x2 matrices at the center of the membrane or lines of pixels at the rim of the membranes) and all share the same electrical contact.

In practice, each cluster of pixel can be contacted individually, but the pixel themselves cannot. This means that, for example, upon SWIR illumination the measured photogenerated current will be made out of a combination of the generated current of all the pixel inside the cluster. Our structures will hence not be able to act as low resolution *"cameras"*.

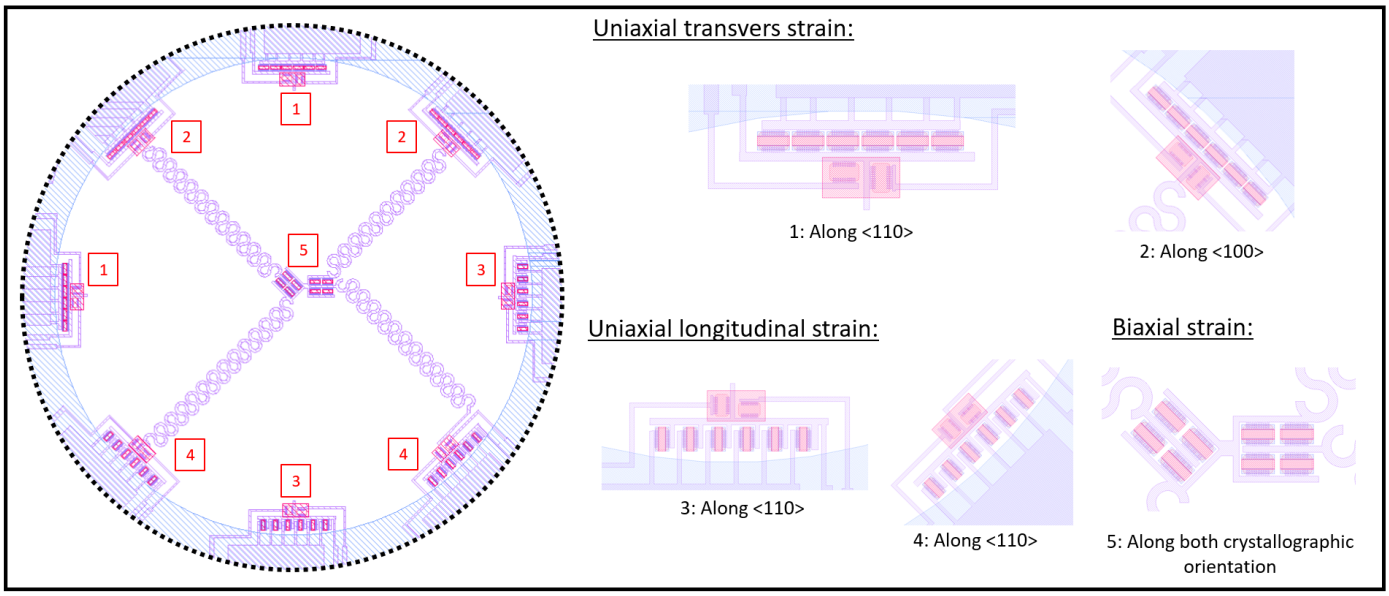


Figure 37: Configuration of the pixels on the membrane. Each membrane features clusters of pixel for both uni- or biaxial loading and along the different orientations of the crystal. In the case of the uniaxial configuration the loading can be either transverse or longitudinal.

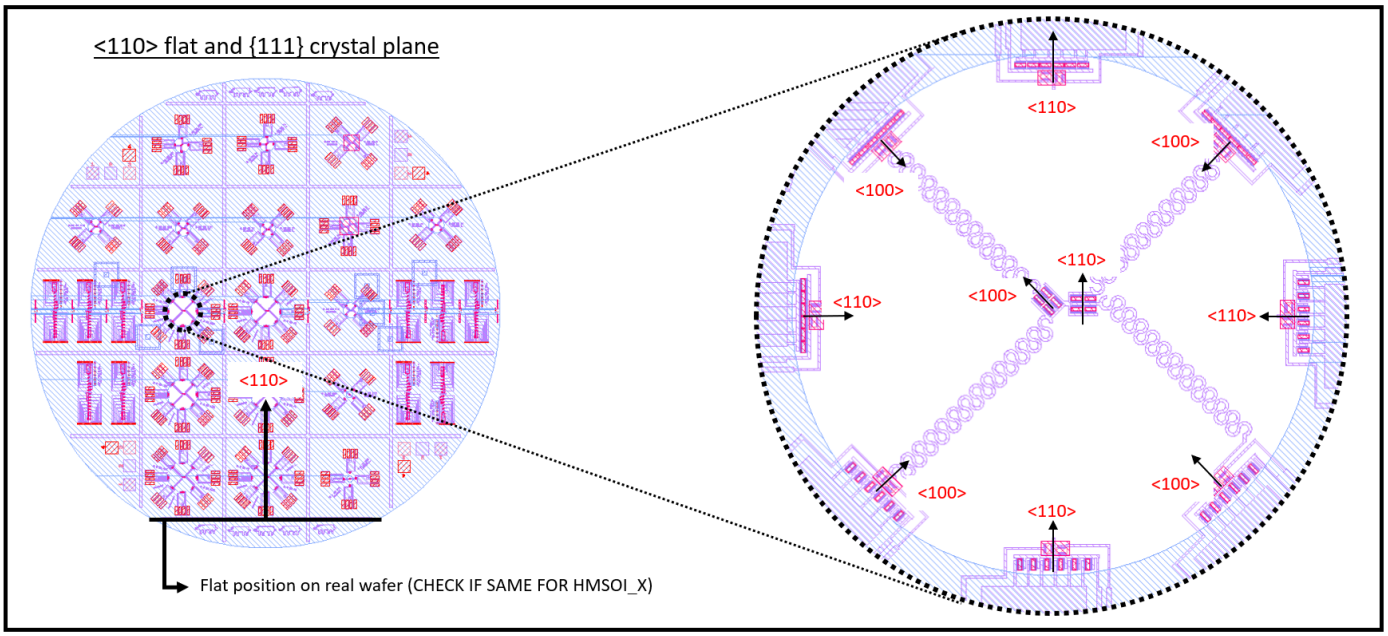


Figure 38: Pixel orientation along different crystallographic directions

3.4 Pixel electrical connections

In order to get a measurement out of our pixels they need to be electrically connected. The connections at the scale of the membrane are handled by wires shaped like horseshoes as was already introduced in the state of the art. The design of these wires is the same as in the reference article and hence similar to those in Figure 25 designed by [37]. This wiring should be able to cope with deformation up to 100%.

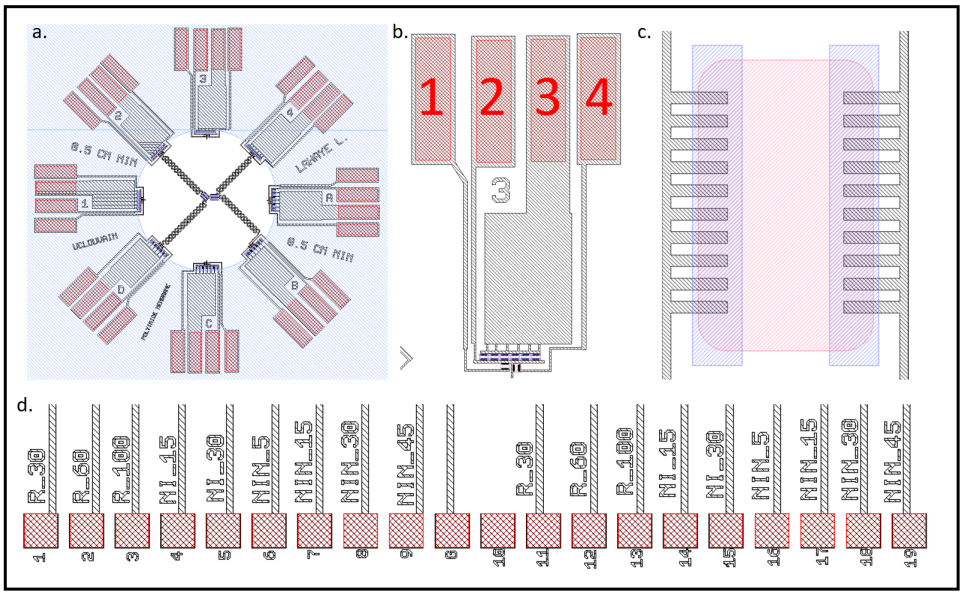


Figure 39: Design of the electrical contacts on the pixels and contact pads **(a.)** Complete die **(b.)** Single substructure. Center pixel for biaxial load contacted by probing both pads number two of the 2/B and 4/D substructures. Resistors for uniaxial transverse (longitudinal) loading uses pads 1+2 (2+4) and NPN or PN for uniaxial loading uses pads 2+3. NPN/PN under transverse (longitudinal) uniaxial loading labeled from 1 to 4 (A to D)**(c.)** Contacting fingers **(d.)** Contact pads for the pixels on the rectangular dies. Each pad is labeled with a description of the structure it contacts

In terms of contacting the pixels themselves, a finger type design was used. This was however found to be a bad idea from a processing stand point. Indeed, at some point during the process, an oxide is grown on the pixels. This means that this oxide must be open prior to the metallization to allow the metal to contact the silicon.

The issue with this finger design is that it does not allow for misalignment during the photolithography. In our case, since enough margin was left to deal with misalignment, the fingers were always contacting the silicon. However, this requires precision and patience during the alignment phase of the mask and the wafer.

An easy improvement would be to design contact holes. Holes would be opened in the oxide to access the silicon and a plain aluminum pad (the width of the pixel) could have been laid on top making misalignment impossible.

The pixels are not individually addressed. We measure them all by averaging their effect by clusters. For example, the devices inside one of the matrices at the center of the membrane all share the same two electrical connections. Their output will hence be averaged. The same goes for the other substructures.

Note that the metal used is not pure aluminum but Al-Si (Aluminum silicon). This alloy, made of 1% of silicon by weight, allowing for a better electrical contact between the aluminum and the silicon.

In terms of the contacting scheme, here is how to measure each substructure:

* **Center pixel cluster for biaxial straining:** They can be probed using two diametrically opposed number two pads (pads number two of the 2/B or 4/D substructures).
* **Transverse resistors:** Those can be probed using pads number 1 and 2
* **Longitudinal resistors:** Those can be probed using pads number 2 and 4
* **Edge NPN/PN:** Pads 2 and 3 should be used.

4 Additional design considerations and final masks

As was already said, the fabrication relies on a set of masks to perform successive photolithographic steps. In our design, six masks come into play. A complete description of each of them would be of little interest. Hence here is a quick description of each mask (all the features can be seen in Figure 40).

* **Strain mesa lightfield:** This mask will be used to build the initial shape of the pixels into the top silicon of our SOI wafers. The dimension of each silicon island being the one given in the pixel design section.

* **Strain P dop darkfield**: Used for the implantation of boron atoms. This mask defines windows for the boron atoms to penetrate and only dope some regions of the pixels
* **Strain N dop darkfield:** Same as previous mask but for the N implantation
* **Strain open passiv darkfield:** This mask is used to open access into the polyimide once coated onto the wafer. Without this mask we would be unable to access the contact pads
* **Strain membrane darkfield:** This mask will be used to etch away material from the backside of the wafer and release our membranes
* **Strain alu line lightfield:** Used for the definition of the wiring and contact pads.

Some more small details have to be added to all of these masks to make them usable in practice. Among those considerations one can find alignment markers, windows and quality markers. Those are useful to align the different layers of the design on top of one another during the fabrication process and to assess the quality of a given step. This will be seen in more details in the section detailing the fabrication process. Large squares for ellipsometric measurements are also featured on the masks.

Finally, note that all the dies are aligned on the wafer as can be seen in section 3. This is so that dicing will be facilitated at the end of the process.

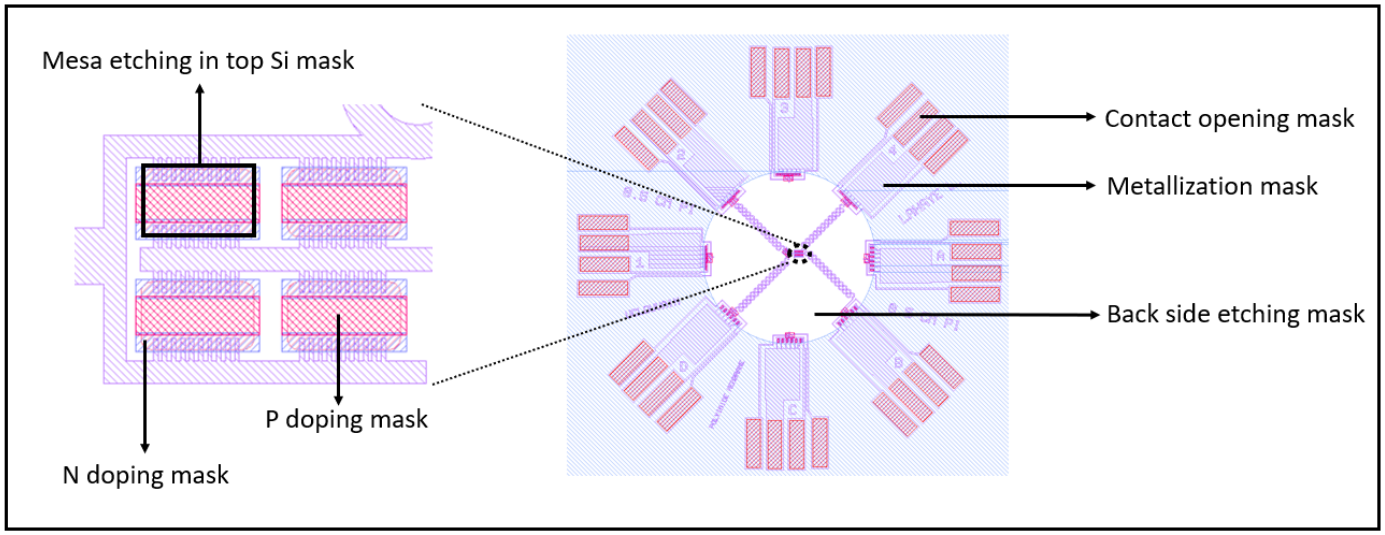


Figure 40: Overview of all the masks

5 Membrane dimensions, shapes and material

5.1 Choice of the polymer

The membranes are built out of a flexible polymer called polyimide. Initially, this material was chosen to match the selected material of [1]. However, upon closer inspection, this choice is coherent since polyimide has the advantage of featuring a high Young's Modulus as explained in subsubsection 1.3.2.

As was explained in the state of the art, many different materials are considered for fabrication of flexible devices. In our case, since the membranes will be quite large (mm range) the use of silicon dioxide or thin glass would have proven to be challenging. Indeed, a compromise should have been found between structural integrity (the membrane should not break upon release during processing) and bendability (the thicker the membrane, the less flexible it gets). Hence, these materials were set aside and different polymers were considered.

The first one, PDMS, has the advantage of being able to stretch a lot. Its Young Modulus of 360 to 870 kPa indeed allowing it to stretch and bend like latex. However, in our case this property is not really positive since, upon the application of strain, most of the deformation would occur in the PDMS and in between the pixels, reducing the amount of deformation transferred to the silicon.

Another idea, suggested by Nicolas André[13] was to use a thick photoresist as our polymer. The suggested resist was the AZ4562 by MicroChemicals[46] which is commonly used to create a thick resist mask for photolithographic processes (up to 9 µm). The advantage being that, since its intended purpose is to be used in photolithography, it is designed to be photodefinable. This would allow for easy opening of the contact pads at the end of the process. The issue however, is that the mechanical properties are not known since it is not designed to be used as a mechanical layer.

Cured Film Properties (400°C Final Cure)

Tensile strength (MPa)	260
Elongation (%)	100
Modulus (GPa)	2.3
Stress for 10µm film thickness(MPa)	18
Moisture uptake at 35% Humidity (%)	1.2
Moisture uptake at 85% Humidity (%)	3.1
Dielectric constant in X,Y plane (at 1 kHz, 50% RH)	3.5
Dielectric constant in Z plane (at 1kHz, 50% RH)	3.3
Dissipation factor	0.002
Dielectric breakdown field (volts/cm)	> 2 x 10^6
Coefficient of thermal expansion,1µm film (ppm)	13
Glass transition temperature	> 400°C
Decomposition temperature	580°C
Weight loss (% at 500°C, 60 min)	1.86
Refractive index	1.78

Figure 41: Properties of PI2545 by MicroChemicals [44]. Poisson ratio can be considered equal to 0.35 [45]

Finally we arrive at polyimide. The most known type of it is the industrial cured Kapton film. However, as was previously said, it exists in the form of a precursor allowing us to spin coat it. The cured film does feature one of the properties of polyimide that we are mainly interested in: the *"high"* Young's modulus of approximately 2.5 GPa (function of the exact product. A range between 2.5 and 8 GPa is expected). This modulus allowing us to transfer a non negligible amount of strain to the silicon pixels upon bending or pressurisation of the membranes.

Different types of polyimide precursor exist and hence a choice has to be made. In our case the decision was conditioned by the product available at the UCLouvain cleanroom (Winfab). Two products were available: PI2545 and PI2611. Initially, PI2611 was considered since it allows for thicker spin coated layers. However, for reason that will be discussed in **??** we had to select PI2545 which features the mechanical properties given in Figure 41. The obtainable thickness for a single spin coat of this polymer is between 1 and 3 µm.

5.2 Membrane shape and size

This section is drafted based on [35, 41, 42]. Other sources will be mentioned explicitly

Multiple sizes and shapes of membrane are proposed on a single wafer. Figure 42 represents the different shapes and sizes as they were drawn in CAD. Two shapes are available, rectangular and circular. Here is an overview.

5.2.1 Circular membranes vs rectangular membranes

Two different shapes of membranes are proposed on each processed wafer: Circular ones (17 per wafer) and rectangular ones (10 per wafer). The reason comes from the fact that this allows us to use the produced samples in two different characterisation setups: A four point bender and an "in house build" pressure testing equipment allowing us to pressurise circular samples up to a bar (designed by Roisin N. and Delhaye T.). Those pieces of equipment will be further described in the characterisation section.

Note that the rectangular dies are designed to be transferred on a flexible plastic strip prior to their use in the four points bender.

Finally, the rectangular and circular dies do not carry the same structures and pixel designs as we will show in later sections. Indeed, the circular membrane comes in different sizes and with different pixels design. However, each die (each membrane) only carries one type of pixel design. For example, per wafer, there are two membranes of 5 mm in diameter that bear NPN pixels. On those two membranes all the pixels are identical. Other membranes however, might come in other sizes like 1 or 2.5 mm in diameter and carry a set of PN pixels or N doped resistors and so on.

On the other hand, all the rectangular dies are the same but each carry all the different pixel designs (and more) featured on all the different circular membranes combined. For instance, each rectangular die has NPN, PN and resistors that vary in dimensions.

5.2.2 Membrane dimensions

Further looking into the design of the circular membranes, three sizes are available per wafer: six 5 mm ones, nine 2.5 mm ones and two 1 mm ones (all diameters). These three sizes are believed to be a good compromise between being able to demonstrate a transfer of a silicon based micro-device on a large piece of flexible substrate (as was done in the reference article) and being able to make an actual measurement under pressure without the membrane breaking.

[13] Postdoc at the UCLouvain

The reference article only features one membrane shape and size. The biggest difference here lies in the fact that their membranes are very large compared to ours. Indeed, where their membrane measures around 17 mm we took the decision to considerably size down ours to 5 mm in diameter for the biggest one. The reasons are:

* **Integrity during testing:** The bigger the membrane, the more prone to rupture it is. Hence, shaving the dimension down also decreases the risk of a membrane breaking once put under pressure.
* **Practical fabrication consideration:** One thing to not overlook when doing micro and nanofabrication is the "know how" of the facility we are working at. This point, even if not being a scientific or technical consideration but more a human one, is of critical importance. Indeed, the last time someone made a polyimide membrane at UCLouvain was Walewyns Thomas for his thesis: *A low-power low-cost MEMS-based gas sensing* in 2016 [47]. Him being long gone, this means that no one at Winfab[14] had experience with the spin coating of a polyimide precursor which is needed for this project. Hence making the membranes smaller means less place for defect and hence increased chance of mechanical integrity.
* **Place on the wafer:** Since we are here working on 3 inches wafers, the available space is limited. Hence, should we have build a membrane the size of the one proposed in our reference article we would have only been able to put four structures per wafer instead of the 27 of the final design.

In a nutshell the goal is to be able to first demonstrate that transferring a silicon based technology on a large piece of flexible substrate is possible: Here we want to validate the "flexible electronic" part of the thesis, and be able to produce a strong enough membrane that will hold the pressure to be able to perform some strain engineering. The next subsection will look into the simulation of the designed membranes as well as the amount of stress that one can generate inside the designed silicon pixels.

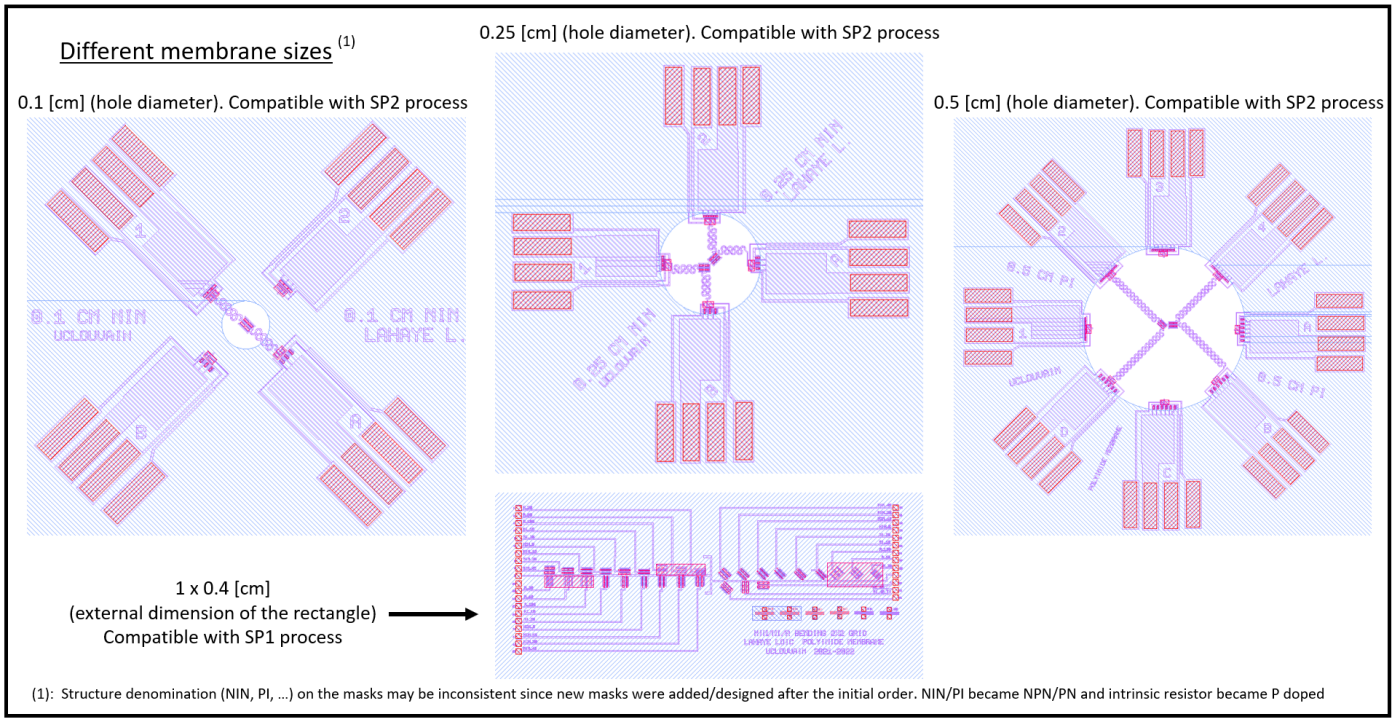


Figure 42: Membrane sizes and shapes available on a single wafer

5.3 Membrane thickness

In order to determine the thickness of the membrane, computations were required. For initial testing in the lab a single spin coated layer of 2 µm will be used to validate the process and perform initial testing. However, an analysis on membrane of various thickness has been performed in order to predict the behaviour of the final devices once pressurised. Thicker membrane could be obtained by repeating multiple times the spin coat/cure cycle: making the layer 2 µm thicker with each spin coat.

[14] The clean room facility of the UCLouvain

5.3.1 Mathematical modeling

For many cases, good approximation of a membrane deflection can be found using simple mathematical formulas. A very popular model being the one of the deflection of a circular diaphragm presented in the book *Practical Mems*[48]:

$$w(r) = \frac{pa^4}{64D}\left(1 - \frac{r^2}{a^2}\right)^2 \tag{5.1}$$

where $w(r)$ expresses the deflection of the membrane as a function of the radial position, p is the pressure, a the radius. D is defined as the *flexural rigidity* and can be defined using the Young's modulus of the membrane E, the thickness h and the Poisson ratio ν:

$$D = \frac{Eh^3}{12(1-\nu^2)} \tag{5.2}$$

This model works well for small diaphragms where the resistance to deflection is dominated by the resistance to bending of the material. For instance, this model could be used for a 1 µm of diameter thin silicon diaphragm. However, it starts to show its limits for larger membrane. The introduced formula fits in the definition of the *Small deflection theory*. For larger membranes, the resistance to deflection will be dominated by the intrinsic stress of the material giving us a slightly modified version of the previous formula (σ_i is the built-in stress of the material):

$$w(r) = \frac{pa^4}{4\sigma_i h}\left(1 - \frac{r^2}{a^2}\right)^2 \tag{5.3}$$

Additional models exist and are presented either in *practical mems*[48] or in an excellent article by *William P. Eaton* and his team [49]. In this paper, an analytical solution is presented to predict the deflection of large membranes. This analysis could be used here but my computation using their model proved to be inconclusive. Since a deep dive into the mathematics of membrane deflection is not the main topic of the thesis, the subject is left to the curiosity of the reader. In addition, even if those models allow us to *"easily"* predict the deflection of a membrane, they struggle to predict the behavior of complex structures like the one presented in this work. Indeed, in our case, silicon pixels are built into the membrane and their effect on the stress induced into the polymer should be taken into account since they might represent stress concentration points leading to potential failure of the membrane.

5.3.2 Comsol simulation setup

In order to provide us with the best possible prediction and to avoid struggling with analytical models for too long, we decided to make a simulation of our membrane. The software that we used is called *Comsol Multiphysics* and is a finite element method simulator. It is build to provide the user with a *"relatively"* intuitive graphical interface allowing to build a 1-, 2-, or 3D model of their structure and run physical simulations on it.

	$\mathbf{E_{PI}}$ **[GPa]**	$\boldsymbol{\nu_{PI}}$ **[-]**	$\boldsymbol{\rho_{PI}}$ **[kg/m^3]**
PI2545 Membrane	2.3	0.35	1300
Mono-Si Pixels	170	0.28	2329

Figure 43: Material properties of polyimide and silicon used in the Comsol simulation

In our case, the design was implemented in 3D in Comsol. Two independent 3D models were built and can be seen in Figure 44:

* **Fix radius membrane:** The first model that was build is the one of the largest membrane proposed in this thesis. The designed modeled membrane has a radius of 1250 µm and features three sets of pixels: The pixels at the center along two directions and pixels on the edge of the membrane in both the longitudinal and transverse direction. The dimensions of the pixels are the ones presented previously, namely: 60x100x0.3 µm. Analysis points were placed at various places of interest: at the center of the membrane both at the middle of the PI thickness and on the surface of the fill and on the pixels. For the pixels point of analysis, the chosen positions are: on the edge and at the center of the XY dimensions as well as at the center of the pixel thickness (0.15 µm) as shown in Figure 44. The studied parameters for this model are the pressure (sweeped between [-5;5] bar with a step of one bar) and the film thickness (evaluated at [2,5,10,20,30,40] µm). The measurement points are represented in Figure 44 (a.).
* **Variable radius membrane:** The second model is a simplified version of the pixel arrangement allowing to automatically sweep the radius of the membrane (in addition to sweeping the applied pressure and film thickness). The position of the pixel is linked to the radius so that their position remains accurate. The pixel dimensions are the same as previously as well as the measurement points.

Once the 3D model completed, one must define the material of each structure. In our case, a model (included in Comsol) was used for mono-Si and a modified version of the included model was used to fit the real parameters of the selected polyimide according to the datasheet (PI2545 [45, 44]).

What is left to do is to define the *Solid mechanics* properties of the different regions. For instance, the polyimide membrane and silicon pixels are defined using a model for linear free elastic materials. This tells the simulator to solve the set of equations 5.4 for each point of the mesh. In order of appearance these equations are respectively the *transformation gradient*, the *Green-Lagrange equation* and the *infinitesimal deformation equation*. They feature: **F** the tensor of deformation, **E** the *Green-Lagrange* deformation tensor, ϵ the small displacement under the hypothesis of small perturbation and **u** the displacement.

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} = (\nabla_{\mathbf{0}}\mathbf{x})^T \qquad \mathbf{E} = \frac{1}{2}(\mathbf{F}^T.\mathbf{F} - \mathbf{I}) \qquad \epsilon = \frac{1}{2}\big((\nabla \mathbf{u})^T + (\nabla \mathbf{u})\big) \tag{5.4}$$

In addition, one must define boundaries and initial conditions that might be expressed as:

$$\mathbf{u}|_{p=0} = \mathbf{0} \qquad \frac{\partial \mathbf{u}}{\partial t}|_{p=0} \qquad \mathbf{u}|_r = \mathbf{O} \qquad S.\mathbf{n} = \mathbf{F} \tag{5.5}$$

These conditions telling us in order of appearance that the displacement and displacement speed of the membrane are null without pressure, that the edge of the membrane are fixed and that the load is an applied pressure generating a force normal to the surface of the membrane.

Now that all of these elements have been setup, all that is left to do is leave the simulator to run and analyse the results.

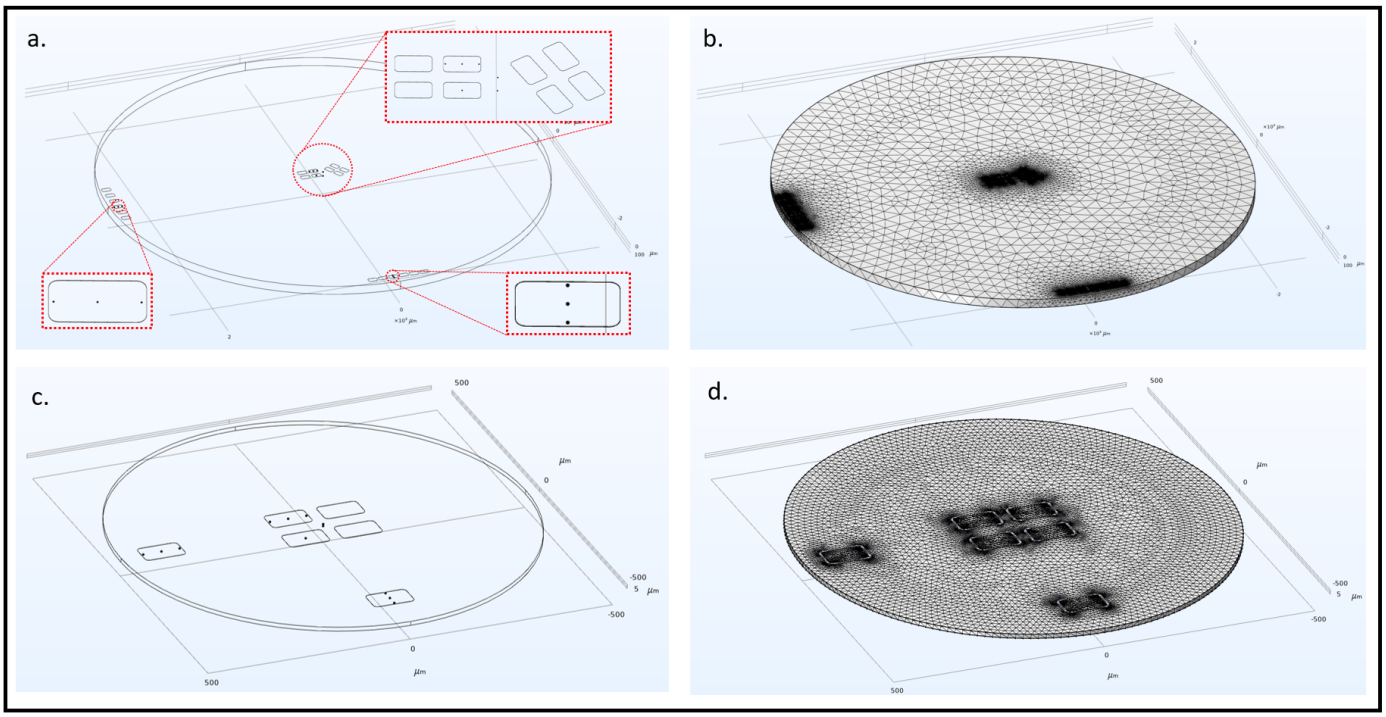


Figure 44: 3D model of the membrane in Comsol multiphysics. **(a.)** 3D model of the 1250 µm radius membrane for pressure and thickness sweep **(b.)** Produced mesh for **f**inite **e**lement **m**ethod (FEM) analysis **(c.)** Simplified membrane for pressure, thickness and radius sweep **(d.)** Produced mesh for FEM analysis

5.3.3 Simulation results

The Figures 45, 47, 48, 49 and 50 show the results of the Comsol simulation. The plots represent the deflection of the membranes for different thicknesses and radius. In addition, plots are given for the radial (uniaxial) stresses and strains of central pixels (of lateral pixels) as well as for the polyimide at the center of the membrane. Many more results can be obtained from Comsol but these are the most valuable for the case we are interested in. As a visual treat, one can see a 3D visualisation of our 5 mm wide membrane under load. The plots are, from top to bottom (Figure 45 (a.)):the strain inside the polyimide and the membrane deflection. Those 3D plots are interactive in Comsol and allow the user to manipulate them in 3D and probe spots for values with its mouse. This method is, however, not very practical and the use of conventional plots is preferred here. Beware that, even if accurate, the 3D visualisation is shown here as an illustration only. Indeed, a high pressure and thin membrane was used to get a visually pleasing result only in Figure 45. Here is a quick description of the provided plots:

* **Figure 47:** This plot features the evolution of the biaxial stress and strain for the largest membrane of 5 mm. The sweeped parameters are hence the pressure in the [-5;0] bar range and the thickness of the membrane. The negative sign of the pressure is there to make the simulation run. For some reason that is still out of my reach today the simulation did not work with positive pressures. As one can see, the deflection of the membrane increases with not only the pressure, but also as the membrane becomes thinner. This translates in a higher level of strain in the pixel when the membrane becomes thinner. From the plot, it can also be observed that the level of strain in the silicon pixels and in the polyimide are not that different. This confirms our theory that the polyimide, due to its high modulus, transfers a lot of constraint to be transferred to the silicon allowing it to bend with it. The maximal strain that has been witnessed in the silicon was one of 2% for a membrane of 10µm and a pressure of 5 bar. This, however, coming at the cost of 6% of deformation in the polyimide.

* **Figure 48:** Here the situation is a little bit different. The thickness of the polyimide has been fixed at 10 µm while the two other radius featured in our design were tested, namely the 1mm and 2.5 mm wide membranes. In this case, since the deflection is smaller (440µm VS 220µm VS 80µm of deflection for the 5, 2.5 and 1 mm membranes at 10 µm thickness) the strains are a little bit smaller than one could have anticipated. Indeed, in this case the maximal biaxial strain at the surface of a silicon pixel located at the center of the membrane is of 1.12 % for the 2.5 mm membrane at 5 bar and 10 µm. However, this decrease in strain comes with a non negligible reduction of the stress in both the polyimide and the silicon. Indeed, when looking a the 5 mm and 2.5 mm 10µm thick membranes put under a similar 5 bar load, we observe strains of 2% (6%) in the silicon (polyimide) with a stress of 5 (0.22) GPa in the silicon (polyimide). This stress decreasing to 2.8 (0.18) GPa in the silicon (polyimide) for a strain of 1.12 % (3.2%) in the silicon and polyimide. This trend in the reduction of stress and strain is also true for the 1 mm membranes compared to the two others. This means that good levels of strain could be achieved with smaller membrane while reducing the stress and hence the risk of shattering our pixels. Note that sufficient levels of strain close to the percent could already be achieved at only one bar using the largest membrane. Indeed, in practice, we do not expect our membrane to withstand 5 bars due to imperfections in the fabrication process.

* **Figure 49, Figure 50:** These plots show the evolution of the uniaxial stress and strain inside the lateral pixels. The same general remarks can be made as was done previously about the impact of the membrane thickness and size. An interesting effect that has been made visible by the simulation, however, is that the membrane does not directly take a convex shape. Indeed, starting from the edge it first undergoes a concave deformation before transitioning to the convex shape as is shown in . This is due to the high modulus of the polyimide that does not really want to deform. Plus a sharp transition from the edge to an immediately convex shape would mean that the polyimide would tear. This is not the case here, the polyimide smoothly transitions from the edge to the final convex shape that one expects. This, however, has an impact (greater on the longitudinal pixels since they cover a larger part of the curve) on the stress and strain profile of the lateral pixels. Indeed, upon inspection of the curves one can see that each pixel at the edge of the membrane goes through both compressive and tensile uniaxial loading. Since silicon does not react in the same way to compression and traction (in terms of effective mass and bandgap), the impact on its electrical performance is hard to determine.

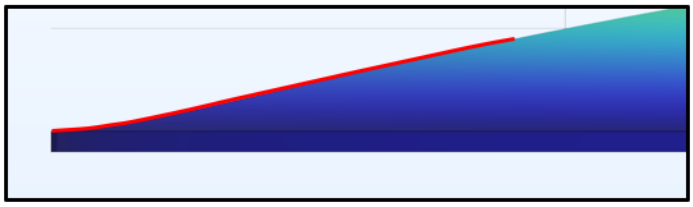


Figure 46: Transition from the concave to the convex shape at the edge of the membrane upon pressurisation

Note that in all of these simulations, the maximal yield strength of silicon of 7 GPa (as reported by [27]) was never exceeded. This does however not necessarily mean that the pixels will resist the bending conditions. Indeed, the fracture phenomena of silicon is dominated by the defect in the material. Hence, our pixel could break at a stress level way below 7 GPa. The polyimide should also never reach its maximal tensile strength of 260 MPa as one can see from the plot, but the same comment about the strength of silicon is applicable here: manufacturing defects in the polyimide layer could lead to its failure before it reaches its tensile strength of 260 MPa. In addition, concentration of stress in the polyimide around the pixel lead to high levels of strain (up to 10% for a 5µm thick 5 mm membrane under *"only"* one bar).

Since the simulation was done prior to the real fabrication of the device, it will later be used as an indicator for the polyimide thickness that we should apply. As we will see, the precursor that we wish to use allows to spin coat a 2 µm thick layer. Simulation results for this 2 µm thick membrane are summarized in the table here below. This table is build to serve as a lookup table during testing. The 2 µm value being a thickness that could be obtained without trouble in Winfab and the pressure values are achievable at the *Welcome* characterisation lab of UClouvain. No values for the 5 mm wide membrane are reported since, looking at the simulation results, it is not expected to hold even a bar. Multiple spin coats might be required.

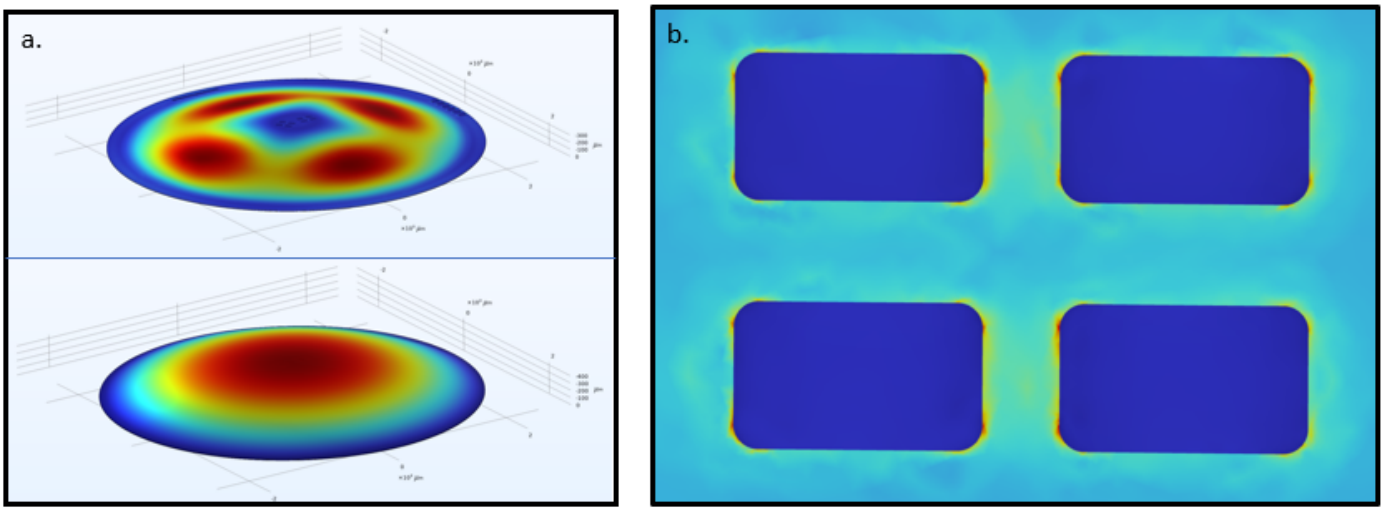


Figure 45: Visual representation of the simulation results **(a.TOP):** Strain in a 2µm thick and 5 mm wide membrane under a 5 bar load **(a.BOTTOM):** Deflection of a 2µm thick and 5 mm wide membrane under a 1 bar load **(b.)** Stress concentration around the pixels

	Membrane deflection [µm]	**Biaxial stress center membrane [GPa]**	**Biaxial stress center pixel [GPa]**	**Biaxial strain center membrane [-]**	**Biaxial strain center pixel [-]**
1 mm membrane @ 2 µm and 1 bar	60	0.086	0.64	0.015	0.0028
1 mm membrane @ 2µm and 2 bars	75	0.136	1.04	0.025	0.0045
2.5 mm membrane @ 2 µm and 1 bar	200	0.16	1.24	0.028	0.0054

Table 4: Lookup table for strain and stress value for a 2 µm membrane part one

	Uniaxial stress longitudinal [GPa]	**Uniaxial strain longitudinal [-]**	**Uniaxial stress transvers [GPa]**	**Uniaxial strain transvers [-]**
1 mm membrane @ 2 µm and 1 bar	0.5	0.0026	0.3	0.0015
1 mm membrane @ 2µm and 2 bars	0.78	0.0042	0.53	0.0025
2.5 mm membrane @ 2 µm and 1 bar	0.81	0.004	0.53	0.0028

Table 5: Lookup table for strain and stress values for a 2 µm membrane part two

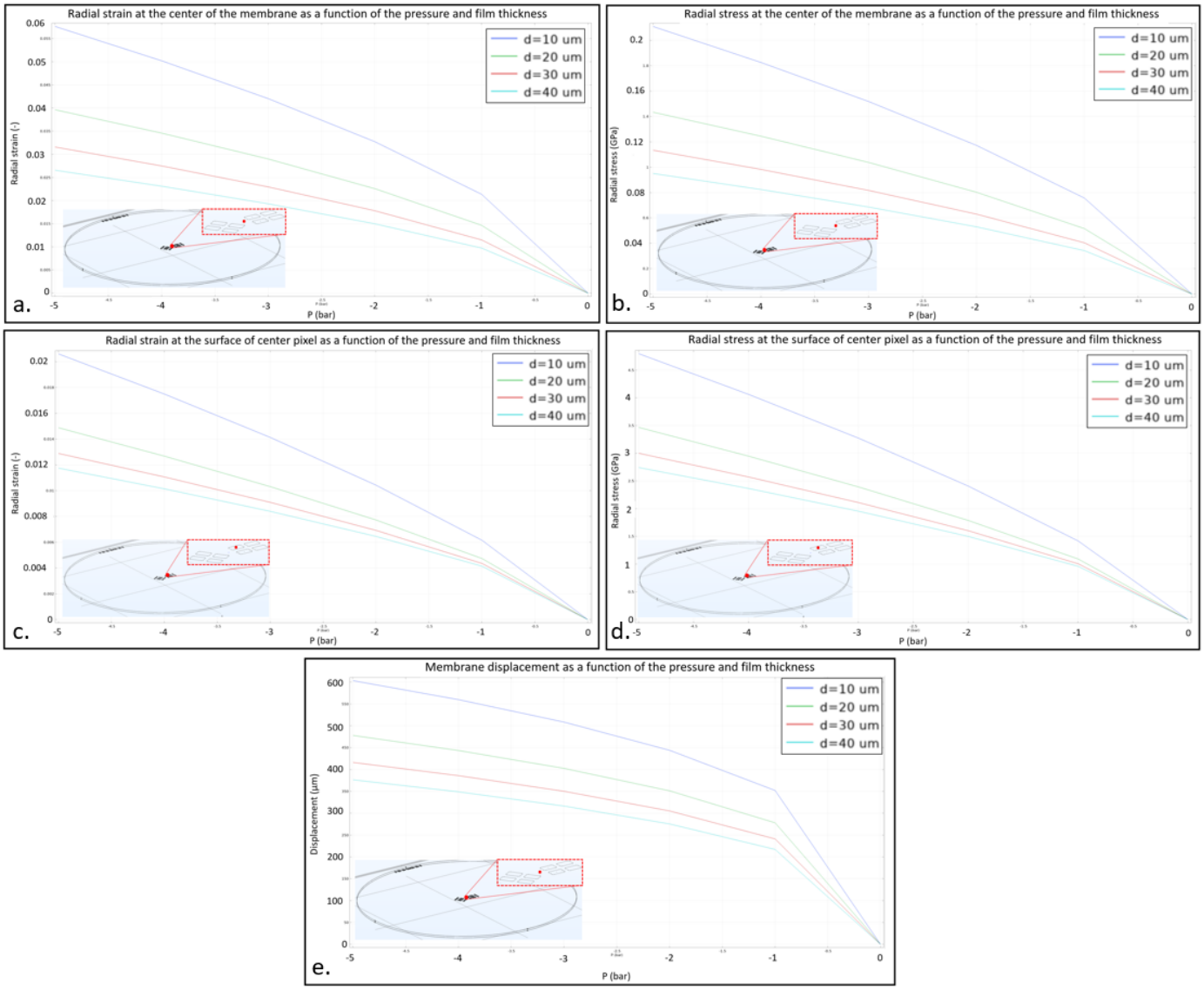


Figure 47: Biaxial stress and strain + deflection as a function of the pressure and thickness for the 5 mm membrane **(a.)** Strain at the center of the membrane in the PI **(b.)** Stress at the center of the membrane in the PI **(c.)** Strain at the surface of a pixel **(d.)** Stress at the surface of a pixel and **(d.)** Membrane deflection at the center

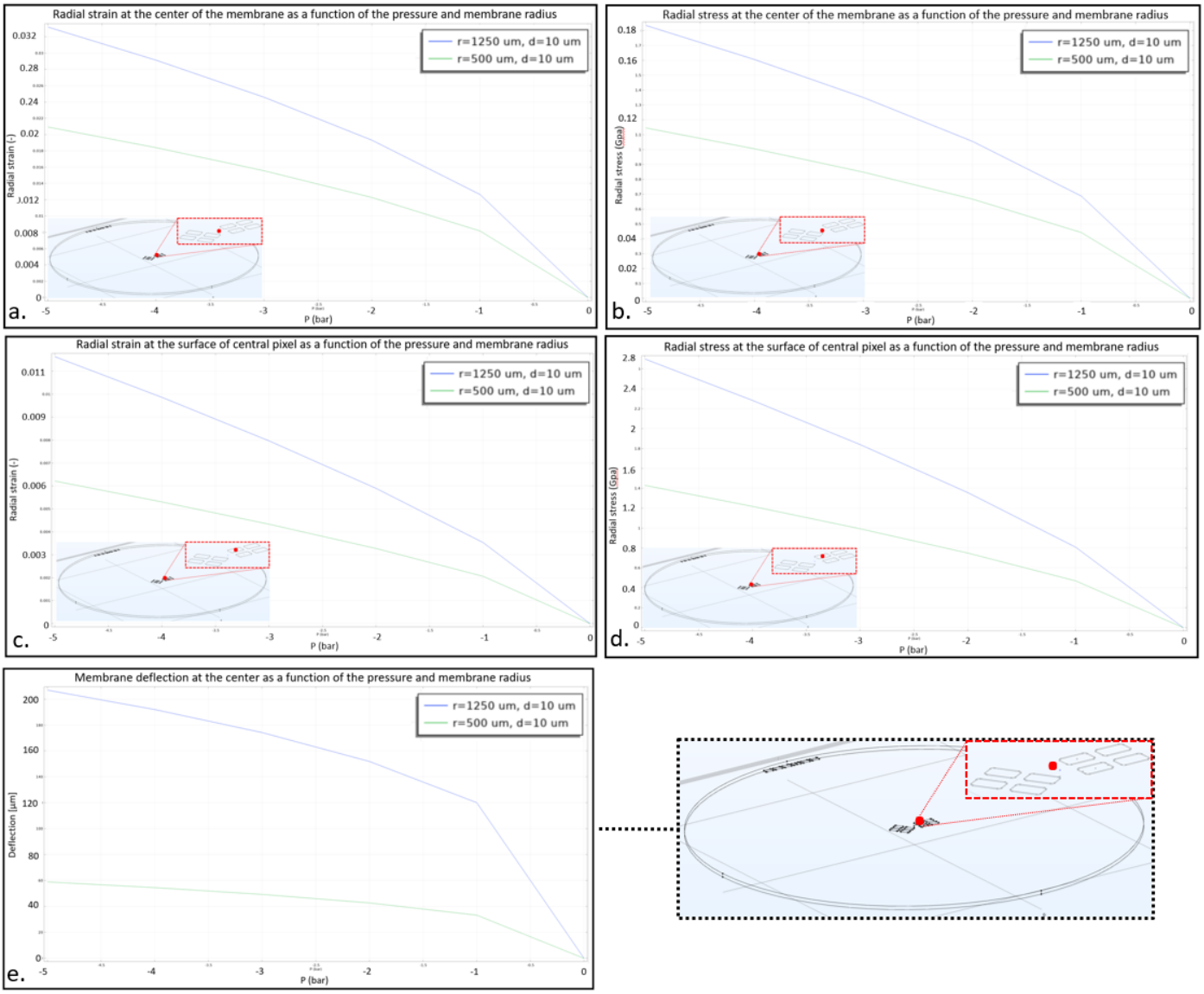


Figure 48: Biaxial stress and strain + deflection as a function of the pressure ans radius for a 5µm thick membrane. **(a.)** Strain at the center of the membrane in the PI **(b.)** Stress at the center of the membrane in the PI **(c.)** Strain at the surface of a pixel **(d.)** Stress at the surface of a pixel and **(d.)** Membrane deflection at the center

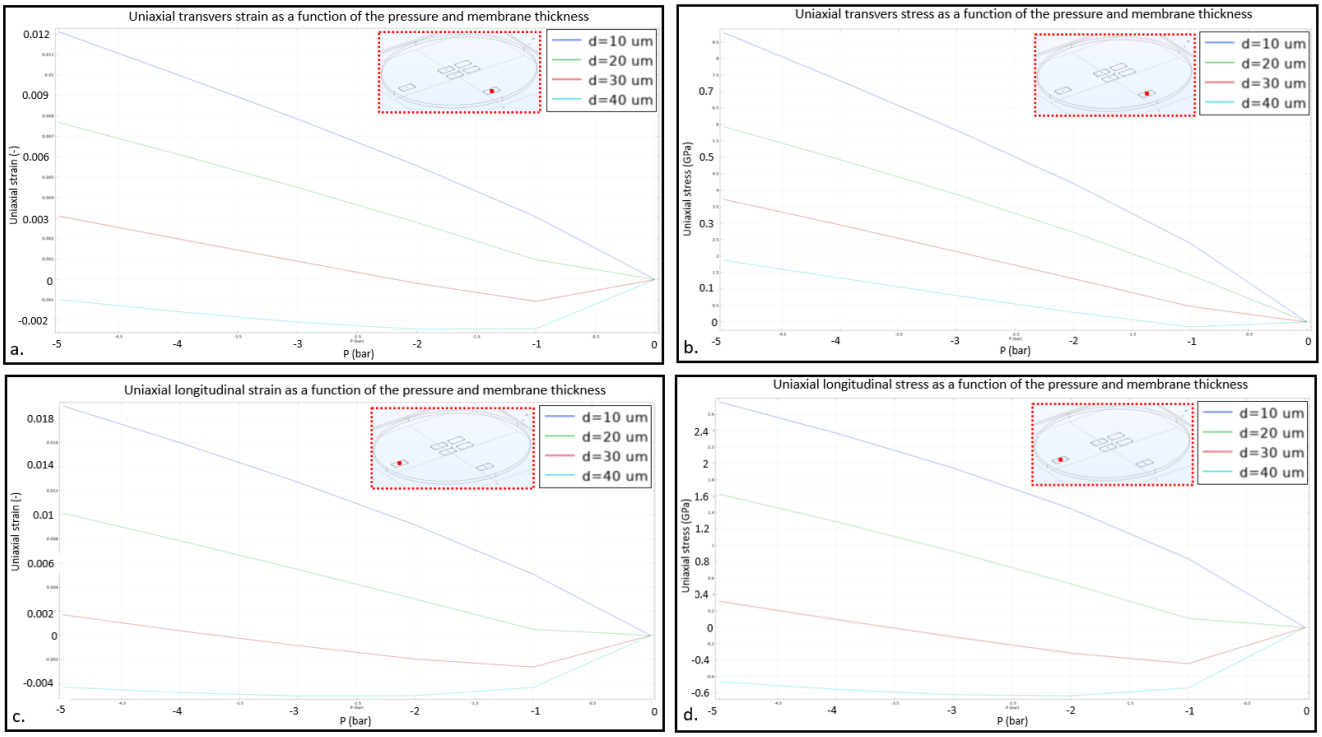


Figure 49: Uniaxial stress and strain as a function of the pressure and membrane thickness for the 5 mm membrane. **(a.)** Strain at the center of the membrane in the PI **(b.)** Stress at the center of the membrane in the PI **(c.)** Strain at the surface of a pixel **(d.)** Stress at the surface of a pixel and **(d.)** Membrane deflection at the center

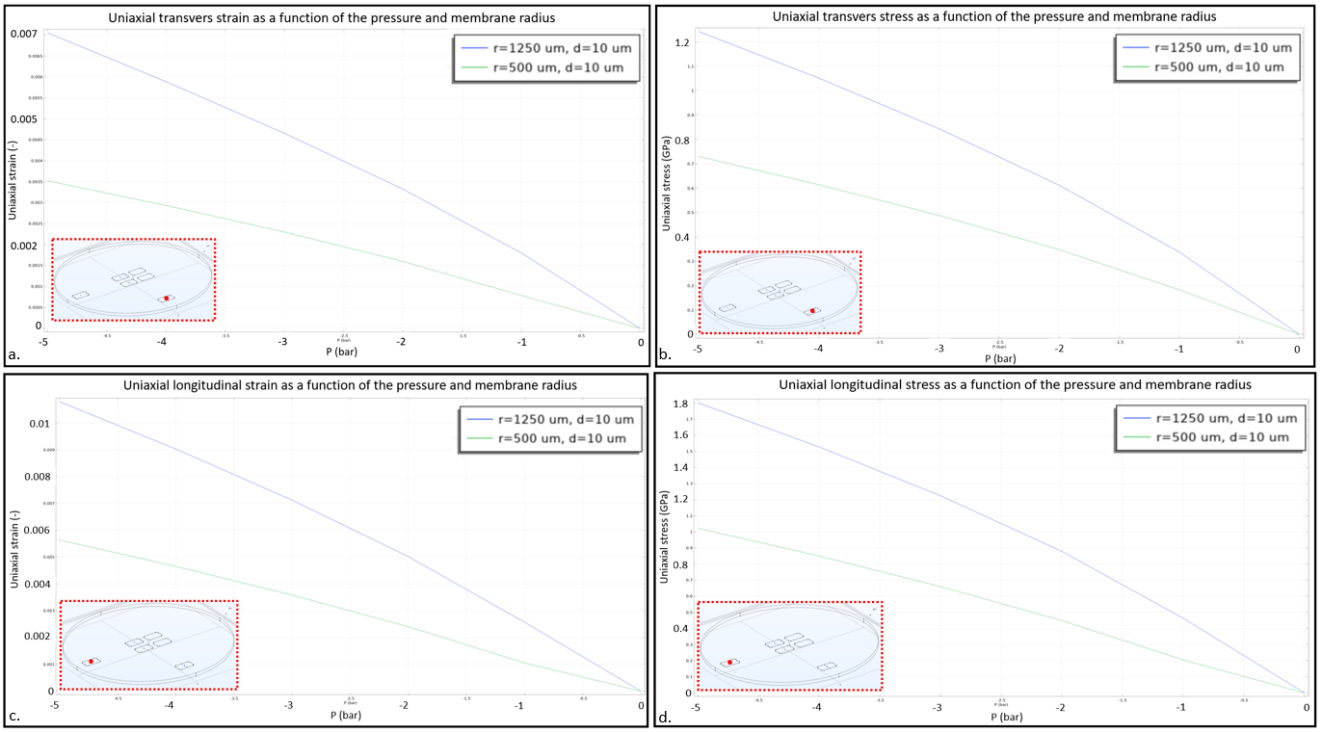


Figure 50: Uniaxial stress and strain as a function of the pressure and membrane radius for a 5µm thick membrane **(a.)** Strain at the center of the membrane in the PI **(b.)** Stress at the center of the membrane in the PI **(c.)** Strain at the surface of a pixel **(d.)** Stress at the surface of a pixel and **(d.)** Membrane deflection at the center

6 Comparison with the reference article[1]

As was said multiple times, the presented design is inspired form the design made by Ajit K. Katiyar and their team [1]. It is hence completely natural to compare the two designs.

6.1 Final design

The final version proposed for the master thesis, and the one that ended up being built in the cleanroom, is rather different from the original reference work. Here are the key differences as illustrated in Figure 51. Since this complete subsection is dedicated to comparing the two, the reader can consider that all comparisons are being made between the master thesis design and the work of Ajit K. Katiyar and their team in [1]. This article being often referred to as *"the reference article"*. This comparison will also serve as a summary of what devices are proposed in this thesis.

* **Membrane sizes:** Multiple membrane sizes and shapes are proposed. The reference work features only one size of membranes whereas our work has three to make sure that a least one size might perform well under pressure. Rectangular dies are also featured to allow four point bending and analysis in the Raman for strain measurement.
* **Connections:** Another important difference is the electrical connections. Indeed, as one can see in Figure 51, the pixels in the design of the reference article each have an individual connection. They share a common ground, but the other terminals can be individually addressed. In our case, to make the wiring easier, all the pixels in a cluster share the same connections, with one common negative (or ground) and one positive terminal. This difference is not negligible. Indeed, as we saw in the state of the art, the device produced by Ajit K. Katiyar's team can be considered as a real infrared (and in our case even SWIR if strained) low resolution *"camera"*. Since each of their pixels has an individual connection they are able to output an image as we saw in the state of the art. In addition, they can measure the variation of the strain as a function of the position on the membrane. In our case this won't be possible. Indeed, since each cluster only has one plus and one minus connection we will only be able to output an averaged measurement over the total number of pixels in the cluster. This means that we won't be able to output an "image" as was done in the reference article. The impact of strain being averaged on a given amount of pixel inside each substructure.
* **Pixel arrangement:** In the reference article only one kind of pixel configuration is proposed: a 6x6 pixels matrix. In our case we propose both a matrix type of 2x2 pixels arrangement at the center of the membrane as well as a linear type arrangement at the border of the membrane. In addition, the substructures are designed so that they align along both the <110> and the <100> crystallographic orientation of silicon. This means that we will be able, using our designed structures, to measure both the impact of a biaxial and uniaxial strain as well as the impact of the crystallographic orientation on the device performance.
* **Pixel number:** The pixel matrix of the reference article is of size 6x6. In our case, since we were aiming to measure an averaged effect of strain over multiple pixels and because we wanted to make the design of the wiring easier, the size of our matrix was set to 2x2. This is also why we were able to decrease the size of the membrane so much compared to the reference design.
* **Pixel design:** The pixel in themselves have differences. Where the reference article only has 20x20x0.01 µm PN junctions, our work features NPN, PN and N- and P-doped resistors. For microfabrication related reasons our pixels are thicker: around 0.3 µm.

6.1.1 General device variation

More generally, compared to the reference work, one processed wafer with our final design proposes more flexibility and opportunity for future work and research. Indeed, one processed wafer provides us with 27 dies featuring different membrane sizes, photodetector pixel designs, membrane shapes and so on.

6.2 Initial design

Even though making a historic timeline of the design evolution is a bit sterile, it is interesting to have a quick look at what the original proposition was for this thesis. Figure 52 provides a comparison between the reference work and the initially proposed design. Here we can see a striking similarity between the two in terms of general dimensions and features (addressable pixels and only one kind of substructure). Furthermore Figure 51 and Figure 52 can be compared to see the evolution of the design over the course of the second semester.

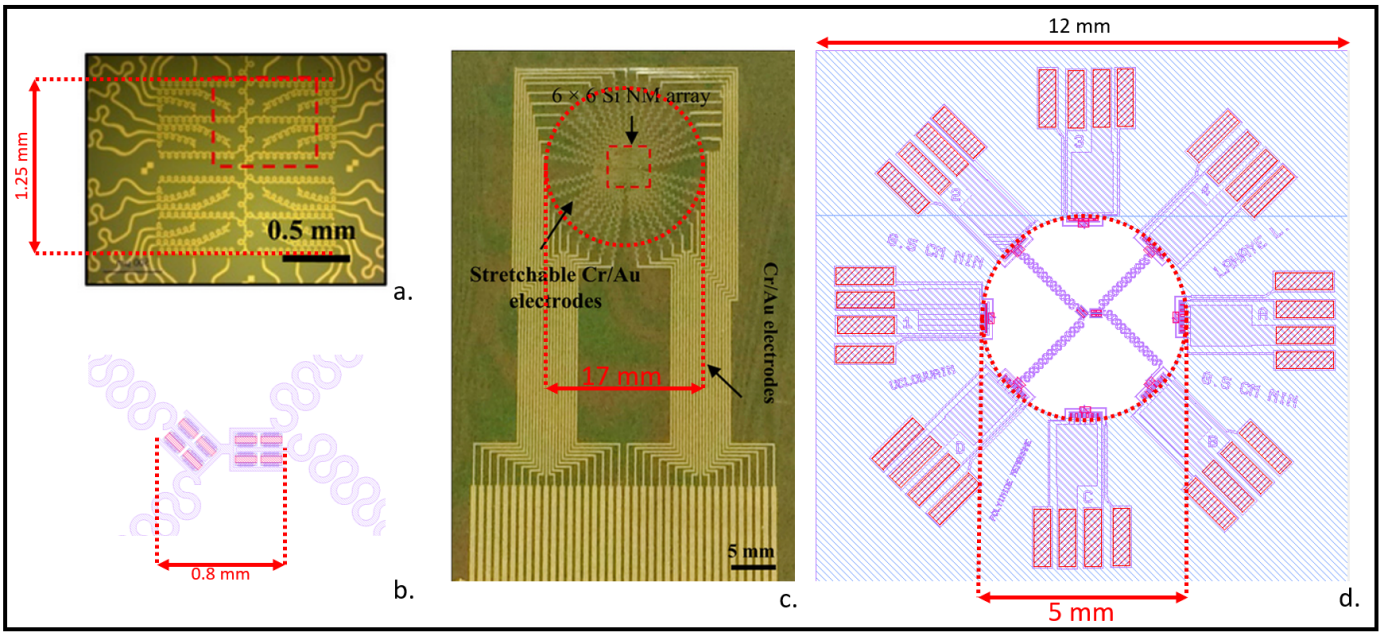


Figure 51: Final design for the master thesis compared to the reference design [1] **(a.)** 6x6 pixel matrix of reference design **(b.)** 2x2 pixel matrix along two different crystallographic orientation **(c.)** Complete design of the reference work **(d.)** Largest membrane proposed for the master's thesis. **(a.)** and **(c.)** coming from [1]

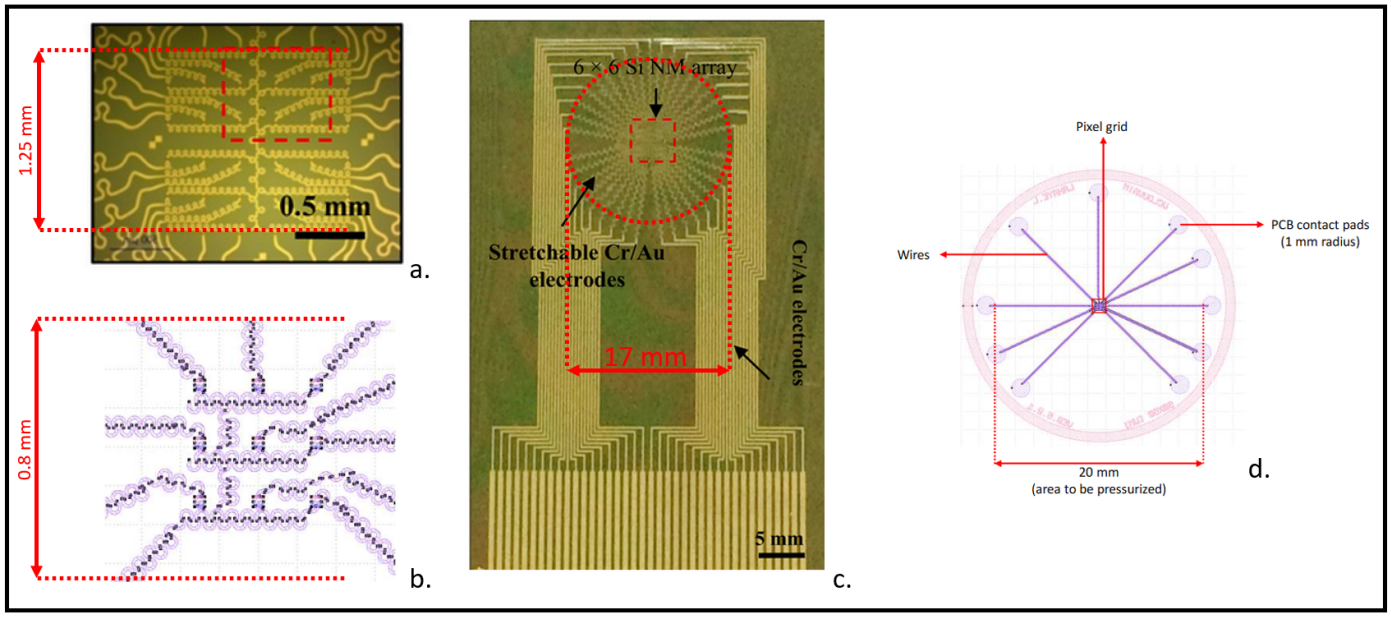


Figure 52: Initial design proposed in December for the master's thesis compared to the reference design. **(a.)** and **(c.)** coming from [1]

Part III
Processing technique presentation

This section is drafted based on [40, 50]. Other sources will be mentioned explicitly

The goal of this section is to be used as an extension to the previous one. The fundamental principles of the different microfabrication processes involved in the main *Process* section have been put here in order to make the text more readable. General indications, explanations and parameters will be given here with the ones specific to our process given in the *Process* section.

1 A word on cleanroom environment and safety

Micro- and nanofabrication of dies and chip must be built under controlled levels of cleanliness. Without going into the details of the different ISO and levels of particles in the air, a cleanroom is an environment designed to limit as much as possible the introduction and generation of contaminating particle such as dust. For that purpose, the users must wear integral cleanroom jumpsuits including a mask, hair-cap and overshoes. Naturally, since the jumpsuits themselves come from the outside of the cleanroom they will carry dust within but won't generate it. These combinations are engineered to generate as little particles as possible compared to the never ending generation of dead skin cells of fabric plumes always coming out of a person. In addition, the user should also wear gloves. These are not meant primarily to protect the person itself, but rather to prevent him from putting greases and oil on every surface.

Should contaminants still be generated in the cleanroom, a permanent laminar flow of filtered air is constantly pushed through the cleanroom from the ceiling and evacuated by the perforated floor. This laminar flow ensures that any particle will be carried down and not travel its way in the cleanroom onto some valuable samples.

All these precautions for cleanliness are taken so that the user does not have to worry too much about contaminant particles collecting on their precious samples and potentially ruining further processing steps and hence their work. Indeed, due to the considered scale (from mm to nm size devices), some grains of dust can look like an elephant compared to an ant.

Different styles of cleanrooms exists such as the ballroom or the mini-environment. In the first case, the whole room is built with the highest standards in terms of air contamination. The operators/users are often required to wear complete respiratory devices to avoid the introduction of undesired moisture and particles from their breath and eyes. This kind of configuration slowly starts to fall out of fashion because its cost is prohibitive. The other approach (the one that we have at Winfab) is the mini-environment one. In this case the cleanroom is kept at a lower ISO[15] around 6 or 5 and houses closed equipment kept under *flux*[16]. The machines are hence surrounded by air just coming back from filtration and their inside is kept at a high ISO of 1 or 2. This technique allows to drive the cost down since it does not require to control and process a large volume of air at all times. It also allows to relieve some pressure on the users by not requiring them to wear complete respirators.

1.1 A word on health and safety

Micro and nanofabrication of devices is done via what is referred to as a *Process*. During this process the wafers will go through many different steps involving many gases and dangerous chemical products. The nature of the work of a cleanroom users makes it such that the handling of chemicals becomes a routine. The installation of this routine is where the danger lurks. Indeed, if all the procedures are followed on the dot at each time and if the equipment is checked regularly, no accident should ever occur. But as with every accident, a false sensation of safety combined with time crunch inevitably leads to mistakes, lowered attention and safety standards. It is easy to forget that the cleanroom is a dangerous environment but it is. Deadly gases and chemical product are being used that could kill someone without them even noticing. Please never forget this when working in any lab environment:

- Never make a habit of seemingly routine tasks. Repetitive and quite boring tasks such as the standard cleaning often leads the user to carrying them out like a routine. If the procedure in itself is not complicated, the chemical products still remains dangerous.
- Never work alone and always check that someone else is around. Two example could be given for this. First: you could be handling certain chemicals or pieces of equipment and cause an accident and second: a gas with no odour

[15] The ISO defines the cleanliness level. An ISO 1 has less than 10 particles per m^3 whereas an ISO 5 would have less than 100000 particles per m^3.

[16] *Flux* is the term used to refer to the laminar flow

starts leaking. If no one is around at the time, the user has breathed too much of the chemical there is no one to save him.

* Never consider a funny smell as normal. Funny smells could be anything from something as simple as resist burning on a hot plate or a ruptured gas line.

* Avoid doing any manipulations in a rush

Permanent training and vigilance of the users is hence paramount to avoid accidents. The user should come in the cleanroom fully prepared and should not improvise a recipe. If all security protocols are followed and if the process steps are performed accordingly with the procedure, everything should go smoothly. A final

Avoid killing yourself in cleanroom ! Respect the procedures, don't rush and don't improvise !

2 Cleaning and preparation

A successful process often relies on proper preparation of the wafer. Indeed, one can have the best settings in the world for a given process step, if the wafer is contaminated, the chance of it ending up with a failure is high. Contamination can induce a lot of different problems such as damaging equipment in the worst case or simply having bad results such as poor adhesion of a deposited material. A good example is the spin coating of resist for a wet etch application. If one does not prepare the wafer appropriately, the resist might delaminate during further processing hence resulting in a bad etch resolution.

Contamination can take different forms:

* **Dust:** can be deposited on the wafer during characterization, handling and transportation. Indeed, even if a wafer should be handled with care and under flux and even if the cleanroom is supposed to be clean, a grain of dust could always arrive at some point

* **Organic contamination:** This type of contamination can result from either a leftover of resist of a prior step or from human contamination. The second type might come from the gloves of the user after them repositioning their facial mask, for example.

* **Metal or thin films *"contamination"*:** The use of the word contaminant depends on the case. For instance, a thin layer of oxide will always form naturally on top of a silicon wafer. This layer, while not qualifying for the term contaminant, is however often undesirable. Metal is another case and can be more complicated to handle. For instance copper is often considered as the worst form of contamination since it diffuses in silicon, ruining the wafer !

The techniques given here below are among the most used in Winfab for wafer preparation.

2.1 Standard cleaning

This technique, the most common one, is a standard procedure that should be done prior to the use of some high ISO type equipment (like the KOYO, furnaces, ...) and to enhance adhesion of later coming thin films to be deposited on the wafer. This procedure relies on a succession of baths of various chemicals and is defined a bit differently in every lab.

2.1.1 Bath one: piranha solution

The basic principle relies on the use of what one calls the *piranha solution*. This solution is a mix of H_2SO_4:H_2O_2 in a (4:1) ratio[17] (this ratio varies from lab to lab. This ratio is the one from the Franssila [40]) at a temperature of 120°C. The result is a solution being able to dissolve any organic contaminant in a two step process:

$$H_2SO_4 + H_2O \longrightarrow H_3O^+ + HSO_4^- \quad \text{then} \quad H_2SO_4 + H_2O_2 \longrightarrow H_3O^+ + HSO_4^- + O$$

The first step of the reaction is *Dehydration*. The sulfuric acid reacts with water to change it in H_3O^+ and hence remove the water from the contaminant. This removed water leaves behind the carbon that can be dissolved by oxygen. This oxygen is coming from the reaction between the sulfuric acid and the hydrogen peroxide.

This reaction will hence be able to strip any organic contaminant. One can understand that this chemical would cause problems if splashed onto the operator. This is why protection glasses are required in addition to the gloves. This process should also be performed under extraction in a fume hood. This video by the Youtuber *NileRed* shows what happens to a sausage put in piranha solution.

[17] H_2SO_4 being sulfuric acid while H_2O_2 is hydrogen peroxide

2.1.2 Second bath: Hydrofluoric acid

HF is a chemical often used to etch away silicon dioxide. The etch rate at which this occurs depends on a lot of parameters like the type of deposited oxide (thermal or PECVD for example), but can exceed the µm/min rate for PECVD. A thermal oxide on the other hand would have an etch rate closer to 100 nm/min. In most of the cases, the layer to remove will be a layer of native oxide formed at the surface of the silicon wafer by natural oxidation. This layer is typically very thin, in the order of a couple of nanometers, and close, in terms of its properties, to a thermal oxide.

The chemical reaction for the etching of SiO_2 in an HF bath is given by:

$$SiO_2 + HF \longrightarrow H_2O + SiF_4$$

Note that the F^- ion is not the etchant of silicon dioxide but H^+ is. A 1% HF is used to remove the native oxide. Note that pure HF does not etch away silicon.

2.1.3 The procedure

The procedure goes as follows:

* **Bath preparation and cleaning:** The first step is to change the bath of piranha solution according to the notebook. Indeed, depending on the time they have been sitting they might need to be emptied out and refilled. The baths are first emptied using the aspiration hose for waste chemicals then rinsed three times with DI water. Once done they must be filled again with the right ratio and amount of chemicals. The same procedure goes for the HF bath that must be changed every day. Some time should then be granted to the baths for them to cool down to 120°C (the reaction of mixing H_2SO_4 and H_2O_2 is exothermic).
* **Piranha bath 1:** The wafers are placed in a first bath of piranha solution for 10 minutes.
* **Rinsing:** The wafers are put in a bath of DI water for 10 minutes under a continuous flow of DI water and nitrogen bubbles. This procedure allows to rinse the wafers to stop the chemical reaction, remove all piranha solution and knock off particles loosely attached to the wafers. The bath should also be emptied and filled five times under continuous rinsing to further increase the odds that particles will leave the surface of the wafers.
* **Piranha bath 2 + rinsing bath 2**: Repeat the previous two steps in two other piranha and bubbling DI water baths.
* **HF bath:** This bath should be done only if the user wants to remove an oxide layer. If not, this step must be skipped. It consists in putting the wafer ten seconds in the HF bath and then putting it into a bubbling water bath for five minutes.
* **Blue fountain:** This fountain is a bubbling bath of DI water like the previous ones. The difference is that the flow of DI water is running around the clock every day of the year. This water fountain is the cleanest spot of the cleanroom. The wafer should be put in for at least ten more minutes. The name of the fountain is simply because the container in which it is at Winfab is blue.
* **Rinser dryer:** Also called the space washing machine. The wafers and their carriers are placed in a tumbler. They are then sprayed with DI water under continuous rotation and then spin dried under a flow of nitrogen. This step completes the standard cleaning process.

Note that extra care should be taken regarding the piranha solution and *particularly* the HF. Indeed, while the piranha solution has a spectacular effect on flesh, the fact that it hurts immediately should be a sufficient indicator that something is wrong. Actions like rinsing of the exposed area can then be done directly. However, HF is more insidious. If splashed or breathed in by the user, HF will not hurt directly, but rather penetrate the tissue until it reaches nerves or bones. Here, though chemical reactions way beyond my level of chemistry understanding, the HF will form complexes and deactivate the cells of the bones and nerves. This leads, in the end, to the nervous system breaking down and the bones dissolving.

This procedure **MUST** be done under a fuming hood, the gloves should be checked for holes (having two pairs of gloves is not an unseen practice) and protective glasses have to be worn.

2.2 Barrel plasma cleaning

One conventional quick way to partially clean a wafer is to put it in what is called the barrel. The barrel is a non directional plasma generator used to strip resist or polymer layers from wafers. This process does not etch away silicon (or silicon dioxide for that matter) or metal.

The chamber of the barrel is put under a low vacuum in the order of the millibar and an oxygen plasma is generated. As was said, the process is non directional and hence both the front and back side of the wafer are cleaned at once. The etching of polymers results from the mechanical bombardment of the wafer of ions generated in the plasma.

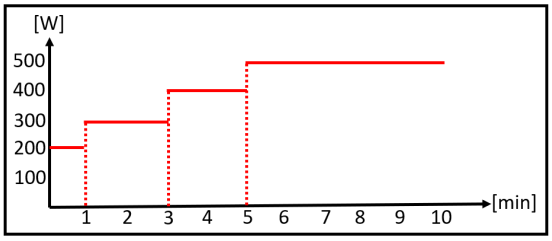


Figure 53: Barrel power profile

The procedure lasts ten minutes and the operator has to increase the power of the plasma manually following the ramp given in Figure 53. The short duration and ease of operation of the barrel tempts some cleanroom users to only do barrel cleaning instead of standard cleaning. Even if effective in most situations, this technique is however not enough to clean the wafers prior to their introduction in a high ISO type machine. But as a quick turnaround to strip resist after a fail lithography this technique is very efficient. In addition, it does not put moisture in the wafer and hence no degassing is required afterward.

2.3 Fuming nitric acid

Without going into the details of the chemistry, fuming nitric acid (which is just 100% concentrated HNO_3) can be used to remove organic traces from a wafer or sample. This type of cleaning is used to replace the standard one after metallization. The first reason is that metals cannot be put in the cleaning bath due to the risk of contaminating them and the second because the HF and piranha solution would damage the metallization.

Fuming nitric acid is very selective toward silicon, silicon dioxide and aluminum, which is why it can be used once metallic contacts have been laid down in a process. The procedure is similar to the one used for the standard cleaning but not as long. The wafer are submerged into the acid for five minutes and then rinsed under a continuous flow of DI water. The main difference is that, since fuming nitric acid is rarely used in cleanrooms (or at least at the UCLouvain Winfab) it is not available in large quantities. The product must hence be put back into the bottle and the bottle carefully rinsed afterwards.

One consideration about this acid though is that, as the name implies, it generates a lot of fumes. The quantity is a function of the amount of polymers to be dissolved. This procedure **MUST** be done under an extraction hood.

2.3.1 Adhesion promotion

Often considered to be a part of the wafer cleaning process, adhesion promotion is a parameter that must be taken seriously. Adhesion of resist in photolithography or more generally polymer adhesion is often of critical importance for the success of upcoming processing steps such as dry and wet etching.

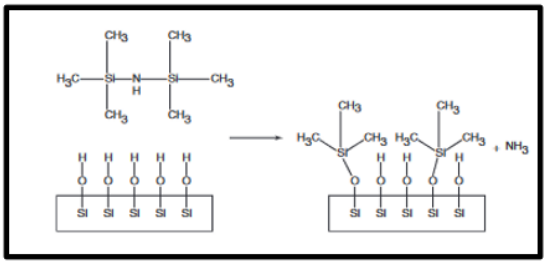


Figure 54: HMDS priming [40]

Where some process steps kind of naturally promote adhesion (like sputtering), the previously cited examples of polymer need special care to be able to adhere to the desired substrate. A common practice is to try to remove as much water as possible from the wafer. This can be done by degassing the wafer in a furnace.

Heat degassing is in most cases a quick process consisting of putting the wafer in a furnace for 15 minutes at high temperatures like 850°C or higher. At Winfab this can be done in the TEMPRESS[18]. Note, however, that degassing at such high temperatures is not possible for doped materials. Indeed, after implantation and RTA of the wafers, high temperature processes (>650°C) should be avoided for the dopant to not diffuse into the material and modify the designed doping profile. Degassing should hence be done at lower temperatures like 120-200°C for an extended period of time. A good practice after a standard cleaning is to leave the wafer to degas in a 120°C oven for at least four hours and up to a week.

Once done, an HMDS (**H**examethyl **D**isilazane) priming of the wafer can be performed. This step is done by exposing the wafer to HMDS vapor under vacuum. As can be seen in Figure 54, the silicon atom of the HMDS bounds to the oxygen at the surface of the wafer where a water molecule used to be adsorbed. This leaves a silane group at the surface

[18]The TEMPRESS is a retired furnace used in the past for thermal oxidation of silicon wafers or nitride deposition by CVD. It was replaced by the KOYO1,2 and 3 and is now used for degassing. This piece of equipment falls into the hot wall tube category. A complete thesis could be written on heat and gas uniformity of distribution, but simply note that in this case the heat comes from an electrical current running in a wire around a silica type glass. The whole system is hence heated. These devices usually stay at a standby temperature of 650°C instead of being turned on and off at the risk of breaking the glass tube (thermal expansion). Inductive heating also exists and allows to get a local temperature gradient.

of the wafer making it hydrophobic and facilitating adhesion of the upcoming resist/polymer. A typical priming process would be done in the LPIII of Winfab and takes around 30 minutes. Once completed, the wafer should be coated in the next 20 minutes before the effect of the HMDS wears off.

A common test to check if a layer is properly adhered to the underlying substrate is the *tape test*. This tape, even if rudimentary, is used very often in the cleanroom environment. The idea is to put a strip of adhesive tape on the deposited layer and then to peel it off. If the layer comes off with the tape, the adhesion is poor and actions must be taken, but if it stays on the wafer the adhesion is considered sufficient and the process can go on.

3 Thin film deposition

The deposition of thin film often refers to the deposition of material that cannot be spin coated. An overview of the existing techniques will be given here with the one used in our process investigated in more details.

3.1 CVD

Chemical **V**apor **D**eposition is a technique used to conformally deposit a layer of material ranging from the atomic thickness in the case of **A**tomic **L**ayer **D**eposition to the µm range. As was said, this process is *conformal*, meaning that the deposition of material will occur everywhere on the wafer and should fill in all cavities.

The idea behind this technique is that the material to deposit enters the reactor under the form of a gas. This gas is not directly the vapor of what one wants to deposit but rather a precursor that, upon chemical reaction at the surface of the wafer will become the final desired material hence the **C** for **C**hemical in CVD. The two chemical equations here below show the typical reaction allowing to deposit silicon and silicon dioxide:

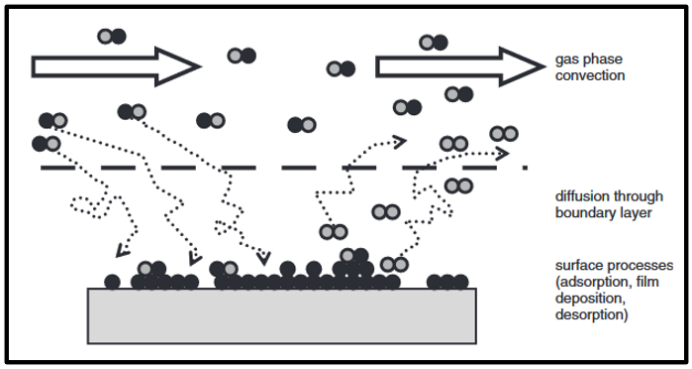


Figure 55: Principle of CVD. Illustration from [40]

$$SiH_{4(g)} \longrightarrow Si_{(s)} + 2H_{2(g)} \quad @600°C \qquad \text{and} \qquad SiCl_{4(g)} + 2H_2 \longrightarrow SiO_{2(s)} + 4HCl_{(g)} \quad @900°C$$

In these equations, one can notice the precursor on the side of the reactant and the desired species on the product side. Other materials such as silicon nitride can be deposited as well. Figure 55 shows the basic principle behind CVD with the introduction of the reactant and the extraction of the unwanted gasses at the output.

The rate at which a given material is deposited on a wafer will be a function of the temperature, the flow of input gas and the probability of nucleation. The two first parameters, temperature and mass flow, are going to determine the regime of operation of our reaction. Two cases are possible:

* **Reaction controlled:** As was hinted by the given temperatures in Figure 3.1, the CVD process and especially the deposition rate is a function of the temperature. Indeed, in order for the precursor to react, an activation energy E_A must be provided to the gas as shown in Figure 56 (a.). This activation energy usually comes from heating. We say that we are in the reaction controlled regime when enough gas is put into the reactor, but not enough energy. This lack of energy will not allow all the gas molecules to react and hence keep the reaction on leash. The deposition rate should hence increase upon an increase of the temperature in the reactor.

* **Mass transfer controlled:** In this case the reaction is limited by the lack of reactant. Enough energy is provided, but the reaction is slowed down by the lack of reactive molecules.

The equation here below[19] in combination with Figure 56 (a.) shows the impact of these parameters on the deposition rate $t_{film}(x)$. Note the presence in the equation of the parameter P_{Growth} which is the nucleation probability and the concentration of the gas C_{gas}. Note that typical deposition rates by CVD are in the order of 0.1-10 nm/s

$$t_{\text{film}}(x) = e^{-\frac{E_a}{k_b T}} P_{\text{Growth}} * C_{\text{gas}} \tag{3.1}$$

Something noticeable here is that the deposition rate is a function of X. This variable X is the distance travelled by the gas over the wafer when it is horizontal. When the gas is introduced in the reactor, it will flow above the wafer from

[19] Equation from [15]

the entrance to the exhaust port all while reacting. What happens is that the concentration in reactant will decrease along the length of the wafer resulting in a non uniform deposition rate and hence thickness of deposited layer. This X parameter influences the equation by appearing in the concentration of the gas.

One way to improve on that and try to eliminate the non uniformity and variation of the gas concentration is to increase what one calls the mean free path. The means free path, expressed here below, is the distance an atom of molecule can travel in a gas without making any collisions. Increasing the mean free path will make possible for molecules of the reactant further in the reactor to be used, hence increasing the uniformity of the deposition.

$$\lambda = \sqrt{\frac{\pi RT}{2M}}\frac{\eta}{P} \tag{3.2}$$

where R is the ideal gas constant, T the temperature, P the pressure and η the viscosity. As one could have anticipated, reducing the pressure leads to a longer mean free path. This makes sense since at lower pressure there are fewer molecules in the chamber and the number of possible collisions decreases. Other advantages include a reduction in the partial pressure of contaminant gasses as well as less likely reactions in the gas phase leading to precipitate falling down on the sample and contaminating it. The list here below gives an overview of the techniques, which are classified by their operating pressure:

* **APCVD:** This process occurs at **A p**tmospheric **P**ressure. The advantages are that the cost of the equipment is reduced due to the lack of vacuum pump and that the deposition rate is higher than the other CVD techniques at the cost of a higher gas consumption. The impact of the high pressure at which the deposition occurs is that the mean free path of the molecules is small and hence that the deposition rate is non uniform along the wafer. The process is typically mass controlled if done at high temperature (but the range is wide: [300-1200]°C). Another downside of this technique is the formation and contamination of the sample by precipitate formed in the gas phase. The throughput of this technique is also often low. Indeed, to cope with the non uniformity of the process the wafers are often processed one by one in a shower head style reactor. The gas does not flow parallel to the wafer but comes from above it.

* **SACVD:** standing for **S**ub **A**tmospheric CVD is typically done in the [1000-10] mBar range and allows to reduce the precipitates and limit the gas consumption. The uniformity is increased compared to the previous technique.

* **LPCVD: L**ow **P**ressure CVD in done in the [1-0.1] mBar range. The mean free path is in this case long enough so that the deposition is considered to be uniform. This allows to stack the wafers in a silicate boat to process them by batches and not one by one, resulting in an increased throughput even if the deposition rate is in general 10-100 times slower than in the case of APCVD. We are in the reaction limited regime with such reactors due to their operation in the [400-900]°C

* **PECVD: P**lasma **E**nhanced CVD is shown in Figure 56 (a.) and its effect in **??**. In this configuration the deposition is done with the assistance of a plasma. The reactant enters the chamber and its energy level is increased by the plasma. This means that a lower temperature can be used to maintain the reaction and hence provides the user with an higher compatibility of usable substrate. Typical temperatures are in the 300°C region (much lower than LPCVD).

* **Other:** ALD for **A**tomic **L**ayer **D**eposition of UHVCVD (**U**ltra **H**igh **V**acuum CVD) are other techniques allowing to deposit layers of one atom thickness (ALD) or further increase CVD (UHVCVD). UHVCVD operates at a typical pressure of 10^{-3} mBar and around [500-600]°C.

4 Thermal Oxidation

As we previously saw, silicon dioxide can be deposited on a wafer by PECVD. Another technique is called thermal oxidation. In this case the oxide layer is not deposited but grown. Indeed, the idea is to oxidise the silicon wafer and turn the silicon in silicon dioxide. Two reactions are possible (both at 900°C):

Wet (130 nm/hour): $Si_{(s)} + 2H_2O_{(g)} \longrightarrow SiO_{2(s)} + 2H_{2(g)}$ or Dry (30 nm/hour):$Si_{(s)} + O_{2(g)} \longrightarrow SiO_{2(s)}$

One process or the other should be chosen by considering the compatibility of a given sample/wafer to water vapor. One of the most important differences between the two being the growth rate as can be seen from the equation even though, in both cases, this rate is low compared to the 10 nm/s of CVD. The difference between the rate of etch and dry thermal oxidation comes from the fact that water molecules have it easier to penetrate the silicon and silicon dioxide layer. This results in a higher concentration of the oxidant and hence a higher growth rate.

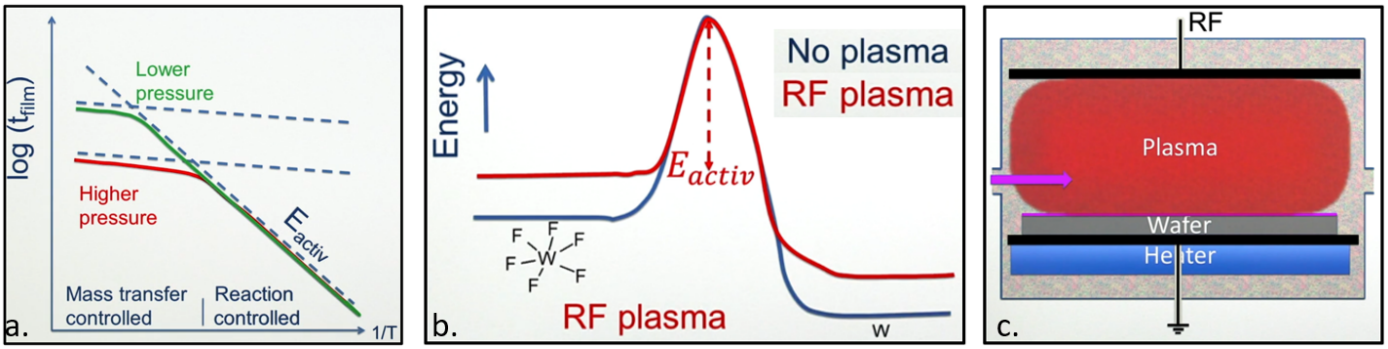


Figure 56: **(a.)** Link between the film deposition rate, the temperature and the pressure on the activation energy **(b.)** Effect of the plasma on the activation energy in PECVD **(c.)** Schematic of a PECVD reactor. All illustration coming from [15]

Note that the oxidation process will increase the volume of the layer due to the introduction of oxygen atoms and that this process consumes some silicon. This means that the growth occurs in both direction and that for a desired t_{ox} thickness, a layer of $0.45t_{ox}$ is consumed. The thickness of the oxide layer can be computed using the Deal-Groove model given by (where Z is the oxide thickness, t the time, C the oxygen molar concentration and ν the the molar volume of silicon dioxide):

$$t = \frac{Z}{KC_s\nu} + \frac{Z^2}{2DC_s\nu} \tag{4.1}$$

Finally, note that the oxidation rate is initially linearly dependant on the duration. However, as the oxide thickness increases, the oxygen of water molecules takes more time to diffuse to the underlying silicon and the oxidation rate starts to depend on the square root of the time.

4.1 CVD oxide vs thermal oxide

Two techniques were introduced to produce silicon dioxide. The choice of a particular technique will depend on the desired properties of the final oxide as well as process related questions:

* **Quality of the oxide:** Thermal oxide are of better quality than CVD oxide in terms of electrical, mechanical and chemical properties. Indeed the CVD oxide often features a rougher surface, smaller breakdown voltage and also a considerably higher etch rate than thermal oxide. About the last part, the etch rate of both thermal and PECVD silicon dioxide in BHF was measured in the context of this master's thesis and it was found that the first one had an etch rate of 86 nm/min compared to >300 nm/min (probably in the µm/min range. See the process section).
* **Duration:** As we saw, thermal oxidation is orders of magnitude slower than CVD. Hence, if the quality of the oxide is not too important, one could prefer a quick CVD deposition at a rate of up to 10 nm/s compared to one at a rate of maximum 130 nm/hour (CVD VS wet thermal oxidation maximal deposition/growing rate).
* **Temperature:** Some materials are not compatible with the high (≥900°C) temperatures of thermal oxidation. For example, thermal oxidation stops being an option once a polymer is deposited on a wafer (it would burn). Hence, in those case one might prefer PECVD and the associated lower temperature (300°C).

5 PVD

The PVD acronym, standing for **P**hysical **V**apor **D**eposition, represents another category of techniques allowing the deposition of a thin film material. The difference with the previously introduced CVD techniques is that in this case, the material to be deposited is put into the machine as is. A physical process, by the application of heat or ion bombardment, will then make that material comes out from its source to be deposited on the wafer. The science behind this kind of deposition is hence mechanical and does not rely on chemical reactions. Let's introduce the two most commonly used PVD techniques in this section. Note that both of them are done in the same piece of equipment at Winfab: the VST.

5.1 Thermal evaporation

In thermal evaporation, the material is placed in a crucible. Upon heating of the crucible this material will vaporise and the vapor molecules will travel to the wafer.

The heating of the crucible could be done either by passing a high current in a wire strapped around the crucible, but the best results are obtained using an electron beam as represented in Figure 57. In this case, an electron beam is generated and aimed at the surface of the material in the crucible. The beam can either be set to target a specific point at the surface of the material or to be swept in a periodic motion. The decision to sweep the beam or not depending on the thermal conductivity of the material to be evaporated.

This configuration allows to reach temperature all the way up to more than 3650K meaning that even tungsten can be deposited via thermal evaporation. This also means that an e-beam evaporator would be able to deposit all material featuring a lower vaporization point. This includes gold (1063K), copper (1088K), platinum (1770K) and so on. This explains why this technique is widely used to deposit thin metallic films.

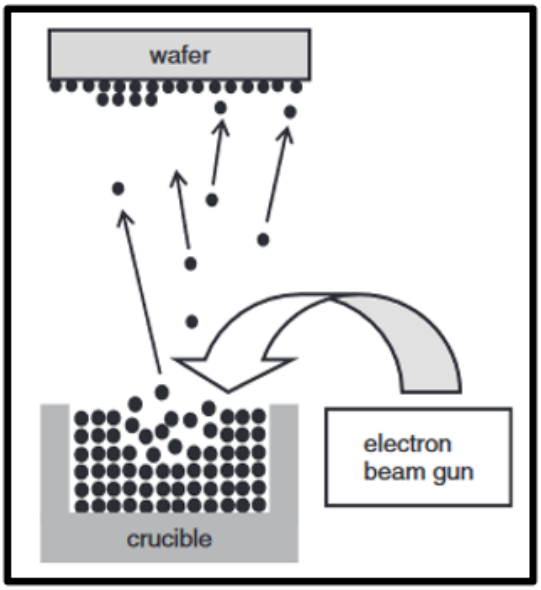


Figure 57: Principle behind thermal evaporation. The illustration is the specific case of electron beam heating. Illustration from [40]

As was the case for CVD, the pressure has a important role in thermal evaporation. Typically, a thermal evaporation process will either require an HV (10^{-7} torr) or UHV (10^{-11} torr) environment in order to work. This low pressure indeed makes the vaporization/vapor generation easier.

Each material (be it metal, water or other) has what is called a specific vapor pressure. This vapor pressure expresses the will or tendency of a molecule to come out of the liquid or solid phase to enter the gas phase. At any time, and for any material, a continuous exchange of atoms from the gas to the liquid or solid state (and vice versa) takes place. Hence, the higher the vapor pressure, the higher the vapor flux will be.

As one could however guess, the vapor pressure of a metal is relatively low compared to the one of, let's say, water. This explains why it is easier to vaporise water compared to a metal. The equation here below links the generated vapor flux as a function of this vapor pressure and the pressure inside the chamber:

$$\Phi_e = \frac{\alpha N_A (P_v - P)}{\sqrt{2\pi M R T}} \tag{5.1}$$

where we have: P_v and P the vapor pressure of the considered material and the pressure in the chamber, N_A the Avogadro number, M the molar mass and R and T the ideal gas constant and the temperature. From this equation, the impact of the chamber pressure can clearly be seen. Indeed, as the pressure lowers, it becomes easier for the atoms to enter the gas phase and the vapor flux increases. Hence, having a low pressure in the chamber is a way to increase the flux and hence the deposition rate. Note that increasing the vapor pressure is another way, but this also means changing the material which is here not the goal.

High vacuum is hence mandatory in order for thermal evaporation to work. One side effect of this is that, especially under UHV, the mean free path λ of the generated gas molecule will be of approximately the length of the reactor (meter range). This mean free path is the same as the one already introduced in the CVD section. As we saw there, reducing the pressure leads to less interaction between the gas molecules and hence a longer possible straight path without any collision for the molecule. This provides the thermal evaporation process with one particularity that one must take into account: its lack of conformality in the deposition.

As the material vaporizes, the atoms will leave the crucible following a straight line as shown in Figure 57. Since the mean free path of the gas molecule under UHV is approximately of the same length as the thermal evaporation chamber this means that the angle of incidence of the incoming atoms when they arrive at the wafer will be very high. This makes for a very directional deposition that could be at the origin of a shadowing effect and hence poor step coverage. To deal with this limitation, the wafer are installed on a planetary system that rotates them and hence makes the angle of incidence of the arriving atoms more random, improving the step coverage.

Also, note that atoms arrive on the wafer with very little energy. Their only energy comes from the heat they got from the electron beam to leave the crucible. This means that they will not impact the wafer with a lot of kinetic energy. This is both positive and negative. The good news is that the wafer will not be heated making this process compatible with

a lot of materials, but the bad news is that the arriving atoms will not be able to knock off any loosely bounded atoms, hence limiting the adhesion. A cleaned wafer is hence paramount to ensure proper adhesion of the deposited layer[20].

5.2 Sputtering

In sputtering, no evaporation is required. The process consist in the bombardment of a target (made of the material to be deposited) by highly energetic ions. Two main kinds of sputtering are possible:

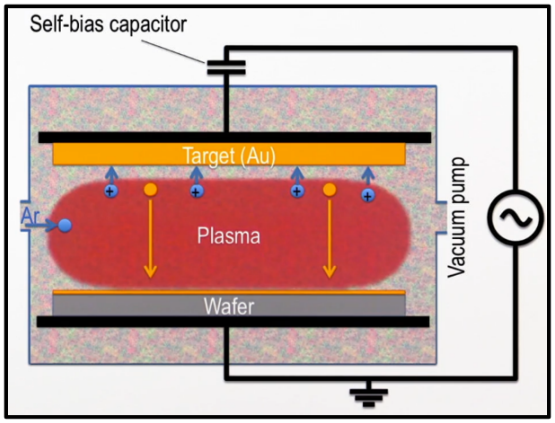


Figure 58: Concept behind sputtering. Illustration from [50]

* **DC:** In this case, the target is connected to a negative DC voltage (cathode) while the substrate is connected to a positive voltage (making it the anode). The increase of this DC voltage generates a plasma in the chamber. Upon the introduction of a gas, such as argon for example, positive and negative ions will be generated due to the gas breaking down in the plasma ($Ar \longrightarrow Ar^+ + e^-$). The positive ions will then be accelerated by the strong electric field and collide with the target. Since the electric field involved in sputtering is high, the ions hits the material with a lot of kinetic energy and hence strip atoms from the target (a bit like when one sands his car to strip rust). The ejected atoms will hence be ejected toward the substrate and deposit on it. An issue with DC sputtering is that it only works with conductive materials. Indeed, if the target is an insulator, the positive ions colliding with it will have nowhere to go and will stay there. This results in a charging of the target, a decrease of the negative DC bias and, in time, the death of the plasma[21]. This explains why DC sputtering can only be used for conductive targets (to evacuate the positive charges arriving on the negative target and maintain the high electric field between the target and wafer).

* **RF:** To overcome the charging of the target, RF (**R**adio **F**requency) sputtering was introduced. This technique, allowing the use of insulating materials as the target, generates its plasma using an RF source. This high frequency RF source is connected to the target through a self bias capacitor and is tasked with the generation of the plasma. At high frequency though, the positive ions are too heavy to follow the quickly alternating field. On the other hand, the electrons, being much lighter, can and will impact the target and the wafer alternatively. Due to the self biasing capacitor, electrons will accumulate on the target and hence create an electric field allowing for the attraction of the heavy positive ions toward it. The rest of the process is the same: the high energy collision strips atoms from the target that fly toward the wafer. Figure 58 illustrates the technique.

The highly energetic collisions between the ions and the target means that the stripped atoms will be ejected at high speed and hence also hit the wafer with a high speed and hence high energy. This comes with both a good and a bad side. The good thing is that the high velocity incoming atom will be able to strip loosely bounded atoms from the surface of the wafer. The sputtering process self-promoting its adhesion. The negative side however is that, as a function of the settings of the field and material to be sputtered, the collision between the atoms and the wafer could generate a lot of heat hence requiring cooling of the wafer, should it carry materials not able to handle high temperatures (300°C or more).

The sputtering process allows for deposition of metals as well as dielectrics and has a typical deposition rate of 1-10 nm/s. This process, by its principle, does not require the assistance of a high vacuum. This means that, if the vacuum is not pulled, the incident atoms can make a lot of collisions with the molecule in the chamber due a short mean free path. This means that the incident angle of deposition is random and hence a good step coverage is possible. The deposition is not directional and no shadowing should occur. In practice, however, both thermal evaporation and sputtering are done with the same equipment meaning that vacuum is also pulled from sputtering which can help increase directionality.

6 Wet etching

Wet etching is a process during which a wafer is submerged in a liquid etchant. The nature of this etchant as well as the properties of the etch will depend on the material to remove and the compatibility of the etchant with the rest of the materials present on the wafer. Here are different techniques and considerations.

[20] Application of HMDS is a bit useless because in practice the machine takes hour to achieve the desired level of vacuum leaving plenty of time for the effect of HMDS to wear of. One could argue that during pumping, water gets pulled out of the system and hence HMDS is useful. This is also true. I do not have a definitive answer here except that too much precaution is never a bad thing in microfabrication.

[21] No plasma no ions, no ions no collisions, no collision no deposition, and no deposition... no deposition

6.1 Silicon

As we saw in the state of the art, silicon can either be monocrystalline, polycrystalline or amorphous. The last two types of silicon, by the lack of crystal structure on the long range will both etch *esoterically* (in all direction). Monocrystalline silicon, on the other hand, features an anisotropic wet etch characteristic when exposed to KOH. This particular aspect of silicon was already seen in subsubsection 1.3.1 and will hence not be repeated here. Let's however look at another technique used to etch away silicon, but in an isotropic way. Since this technique was not used in this thesis, only a quick overlook will be given here. [50] or [40] should be consulted for more details.

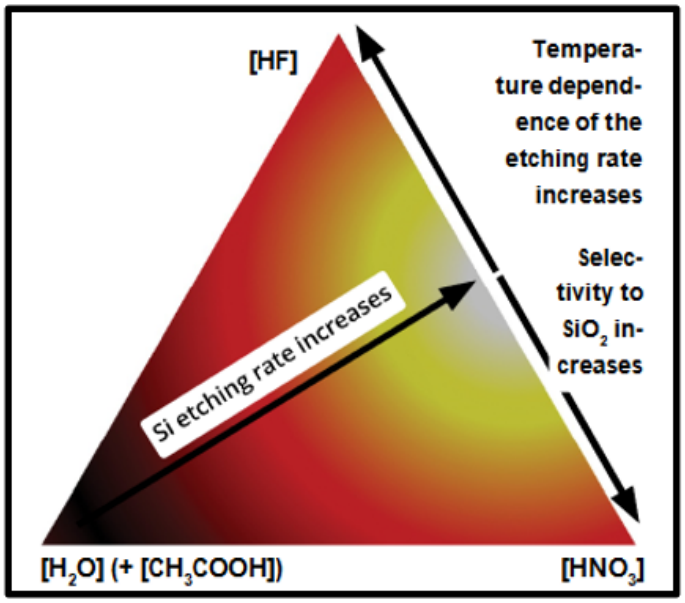


Figure 59: Etching triangle of silicon by [51]

Monocrystalline silicon can be etched away isotropically by a technique relying on oxidation. Two methods are possible, the first one relying on a succession of thermal oxidation followed by a conventional BHF etch of the silicon dioxide or by using a solution made out of nitric acid HNO_3, hydrofluoric acid HF and acetic acid.

Alone, none of them etche away silicon. However, diluted nitric acid has the potential to oxidise silicon. The oxide being later etched away by the hydrofluoric acid (the acetic acid is here to improve wetting of the wafer. Indeed, SiO_2 is hydrophobic). The amount of both nitric and hydrofluoric acid will determine the etch rate of silicon and a triangular type diagram can be drawn as was done by *MicroChemicals* in Figure 59. Note that this technique offers no selectivity toward silicon dioxide.

TMAH (**T**etramethylammonium **H**ydroxide)) can also be used to isotropically etch away silicon. Even if the process does not feature a perfect isotropy (30:1 for the (100):(111) crystal plane) as the previous technique, it comes with a selectivity of 2000:1 toward silicon dioxide and an etch rate of 0.5 µm/min. This technique is also the most common in cleanrooms.

6.2 Silicon dioxide

As we already saw, silicon dioxide can be etched away by hydrofluoric acid. The already-seen chemical reaction:

$$SiO_2 + 6HF \longrightarrow H_2SiF + 2H_2O \qquad @25°C \text{ and } 49\% \text{ HF} \tag{6.1}$$

is the reaction allowing the dissociation of silicon dioxide and hence its etching. A typical concentration like the one shown in the equation does however feature a very high etch rate making it impractical to use. In addition, the very low PH of HF makes it incompatible with conventional resist masks used for wet etching applications. Hence, in order to have more control on the etch rate and make this wet etch compatible with photoresist, a solution has been devised: BHF

Before even looking at what BHF is, let's take a look at how the etching reaction occurs. When put into a solution, the hydrofluoric acid splits into both H^+ and F^-. The species responsible for the actual etching process and hence dissolution of the silicon dioxide being the H^+. From here, one can understand that to reduce the etch rate of the process, some H^+ ions should be removed. This is done by the addition in the solution of ammonium fluoride NH_4F in a ratio of 6:1 to 10:1 (40% wt NH_4F and 49% wt HF). This particular mix results in what is called BHF or **B**uffered HF where the NH_4F acts as a buffer to reduce the etch rate.

When added to the solution, the ammonium fluoride will dissociate in both NH_4^+ and F^-. This excess in F^- will shift the reaction equilibrium and lead to a decrease in H^+ ions. This can be understood by looking at the equation of the reaction constant of the HF dissociation:

$$K = \frac{[H]^+[F]^-}{[HF]} = 6.7 * 10^{-4} \tag{6.2}$$

This reaction constant will dictate the possible concentration of the different species in solution. Hence, since adding NH_4F adds an excess in F^-, the only solution to keep the above equation true is for the concentration in H^+ to decrease. This decrease will translate into an effective reduction of the etch rate of silicon dioxide (in the order of the µm/min at 25°C for CVD silicon dioxide).

Another benefit of BHF is that its PH is also higher than the one of conventional HF, meaning that BHF can be used with classical resist masks.

6.3 Aluminum

Aluminum can be wet etched by almost all of the silicon etchants. For instance, KOH etches away aluminum at a rate of 60 nm/min for thermally evaporated aluminum and 400 nm/min for sputtered Al-Si. This implies that all the processing steps of the silicon must be done prior to the metallization of the wafer. If this is not possible in the context of a given process, care must be taken to protect the aluminum. Such care could take the form of protection of the aluminum by covering it with resist. [52]

The opposite, however, (etching aluminum but not silicon) is possible. Indeed, a common aluminum etchant is phosphoric acid H_3PO4. This acid is a very good etchant of aluminum with an etch rate of >500 nm/min for thermally evaporated aluminum and of almost 1 µm/min for sputtered Al-Si according to [52]. One major advantage of this etchant is that it features a virtually infinite selectivity to silicon (once again [52]).

Note, however, that this process is isotropic so the width of the aluminum traces should be designed with in mind the desired thickness of the traces. Indeed, because of the under etch, a 1 µm wide and 2 µm thick trace is not possible. This is well illustrated in the failed wet etch attempt of our process.

7 Dry etching

Dry etching refers to all techniques used to etch away a material by the use of a gas or a plasma. Three techniques could be highlighted:

* **Isotropic:** This technique relies on the use of a gas as an etchant, but without the assistance of a plasma. A well known example would be the use of XeF_2 to etch away silicon. This process is not driven by ion bombardment and no sidewall polymerisation occurs. The result is a high level of isotropy and under-etch should be considered while designing the traces and features of a micro device.
  * $2XeF_{2(g)} + Si_{(s)} \longrightarrow 2Xe_{(g)} + SiF_{4(g)}$
* **Anisotropic:** In this case, which is illustrated in Figure 60, the etching process is dominated by ion bombardment. The substrate is connected to an RF source generating a plasma. The plasma will break the argon molecule (for example) introduced in the reactor and the positively charged ions will be accelerated toward the wafer hence etching it. The anisotropy comes from the fact that this process is purely mechanical. This technique hence does not rely on any chemical reactions and is shown in Figure 60. Note again the presence of the self biasing capacitor allowing the wafer to charge negatively[22].
* **RIE:** The previously cited example is however not very efficient. In practice, one would prefer to use reactive species to increase the etch rate. In **R**eactive **I**on **E**tching, reactants are added at the input and, upon ionisation in the plasma, will generate reactive radicals able to etch the desired material. Down below is an example for RIE of silicon using the CF_4 gas. Note that pure RIE (and not DRIE/Bosh or ICP) does not rely on ion bombardment. The etch profile is hence highly isotropic.
  * $CF_4 + e^- \longrightarrow CF_3 + F + e^-$
  * $Si + 4F \longrightarrow SiF_4$
* **DRIE:** In this case, the bias of the wafer is not done by a self biasing capacitor, but by a separated DC source. The voltage applied on the wafer is hence the sum of an HF RF voltage and DC bias. This allows to tune both the intensity of the plasma and the intensity of the ion bombardment. Indeed, **D**eep **R**eactive **I**on **E**tching is different from RIE in the sense that both mechanical and chemical etching are featured. A typical DRIE process for silicon would involve the introduction in the reactor of two gases: SF_6 and C_4F_8. The first gas being the etchant and the second one, called a chlorocarbon gas, allowing to control the level of anisotropy by controlling the polymerisation point. In the plasma, both gases will break down. Three phenomena then take place:
  * **Chemical isotropic etch:** The generated F radical will chemically and isotropically etch away the silicon
  * **Anisotropic ions bombardment:** The ions accelerated by the field induced by the DC bias will collide with the wafer, hence mechanically etching it. This part of the process is the reason for the **D** in DRIE. This letter meaning **D**eep. Indeed, the accelerated ions have the possibility to go in deep tranches.

[22] As in sputtering. The difference being that here the wafer is the cathode.

- **Polymerisation:** The generation of carbon radicals from the C_4F_8 will induce a polymerisation process at the surface of the wafer and on the sidewalls of etched trenches. This wall polymerisation protects the wafer against lateral etching and hence increases the anisotropy. By fine tuning of what is called the *polymerisation point*, an equilibrium between the deposition of polymers on the wall and its etching by ion bombardment can be found resulting in an anisotropic etching of the material. The etching of the bottom of the trench is not made slower since the bombardment of ions perpendicularly to the wafer is highly energetic. Polymer deposition elsewhere than on the walls (where the bombardment is less intense) can be considered null or not significant. A well developed example of DRIE would be the BOSH process.

- **ICP-RIE:** Up to here, we have considered that the wafer was connected to the RF source for plasma generation and (in the case of DRIE) to a DC bias for ion bombardment. However, in practice, it is not rare to decouple the generation of the plasma from the biasing. A possible path is to go for an **I**nductively **C**oupled **P**lasma generation such as what we have in the Corial at Winfab. In this case, the plasma is generated by a coil surrounding a ceramic reactor. This allows to decrease the temperature of the wafer during etching, in addition to having a fine control of the plasma in terms of both its position and power.

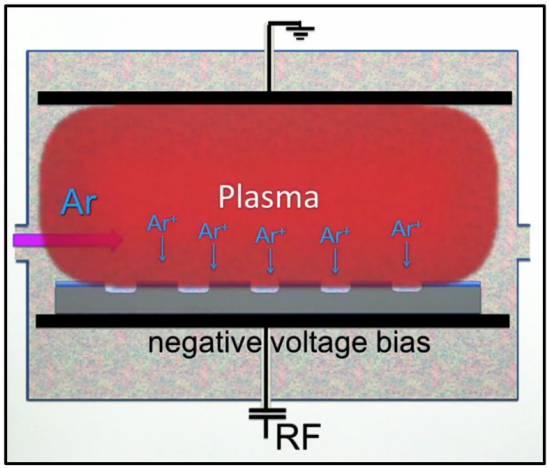


Figure 60: Principle behind anisotropic dry etching. Illustration from [50]

8 Photolithography

Lithography is a process during which a design/pattern is transferred from a reference to a desired substrate. The most well-known type of lithography uses a light source to expose a photoresist and then, upon development, reveals the drawn pattern. This specific kind is called photolithography.

The basic principle is as follows: a layer of photodefinable resist is laid on top of the wafer one wishes to transfer the pattern on. Once done, the light is shone through a photomask. The photomask carries the reference design and will block or allow light to pass through and hence define the exposed zones of the resist underneath. Let's look at the different parameters and considerations to bear in mind for a photolithographic process.

8.1 Mask design

The masks are, as was told, what carry the reference design. They allow to define a pattern at once via photolithography by blocking or allowing light through. This allowing for a high throughput of the process.

The alternative of using masks would be to etch the pattern, using for instance direct writing, which takes hours. This direct writing is hence preferably used to build a photomask. This mask starts as a piece of chromium covered glass. The chromium being the layer in which the pattern is engraved. Here are some considerations about photomasks:

- **Polarity:** A mask can either be a dark- or lightfield. In the first case, the mask is completely covered in chromium in which openings for the design are etched. A lightfield on the other hand is completely transparent, except for the pattern which is in chromium. The mask for the metallization is shown in Figure 74. This mask is an example of lightfield. The decision to use a dark- of lightfield mask should be done during the definition of the process and in accordance to the resist to be used.

* **Alignment mark:** Often many photolithography are required in order to complete a design. One step do define the silicon pattern, one to define the metal traces and so on. Hence, alignment marks like the one shown in the process part of this thesis should be included on the mask. Those are used to align a new mask on a previous pattern. In the case of a darkfield mask, windows in the hundred of µm in width should be added allowing the user to see through it. Indeed, in photolithography, the mask is placed above the wafer to expose.

8.2 Type of resist

Photoresist is the resist used for photolithography. This kind of resist reacts with the exposure light and can be of two sort: positive or negative. In both cases, the incident light will induce changes in the polymer bounds of the resist. In the case of a positive resist, upon illumination, bounds will be broken down and the resist will become more soluble where exposed. The negative tone ones do the opposite: when illuminated, new polymeric bounds will be created making the resist less soluble where exposed. This is shown in Figure 61.

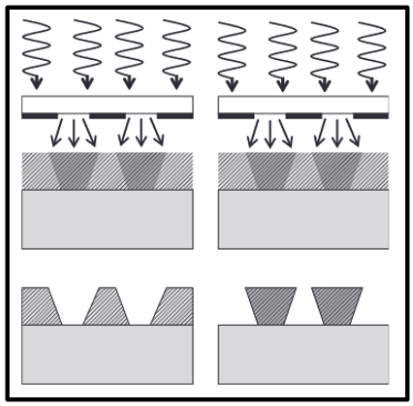


Figure 61: Positive **(Left)** and negative **(Right)** photolithography. Illustration from [40]

The amount of light and hence energy required for a sufficient modification of the chemistry of the resist is what one calls the exposition dose. This dose, expressed in mJ/cm^2 is given by the manufacturer and changes from resist to resist. For conventional positive and negative tone ones, the range is between [150-200] mJ/cm^2 (but this will fluctuate as a function of the resist and its spin coated thickness. The thicker the resist, the more energy will be required. Values out of this range are not excluded at all.)

By the nature of the reaction of polymerisation, positive and negative tone resists will feature sloped angled walls. For example, considering positive tone resist, the resist at the surface will receive a higher dose that underlying one. Hence, more polymer bound will be broken near the surface than near the bottom. This will translate in practice in inward angle walls (narrower at the bottom near the wafer and wider at the top). This is also shown in Figure 61.

Finally, note that the polymer is carried in a solvent to allow for the resist to be spin coated. A typical processing cycle for such resist would be:

* **Spin coating:** The resist is spin coated on the wafer accordingly to the datasheet. The final resist thickness and hence exposition dose is determined by the rotation speed. This thickness to rotation speed relation is given by what is called a spin curve, which must be provided by the resist manufacturer. Spin curves for the used polymers in our process are given in Figure 63.

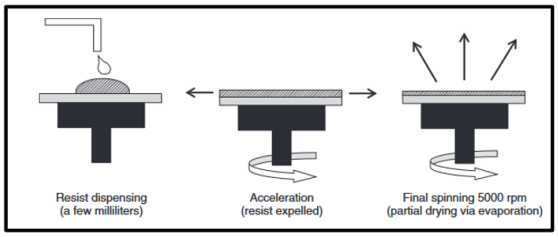


Figure 62: Illustration of spin coating. Illustration from [40]

* **Soft bake:** After the spin coating, the resist is still in a liquid form. In order to allow the wafer to be further processed, the resist must go through a soft bake. A typical baking of 90 seconds at 90°C is performed to drive out most of the solvent and make the resist *cured enough* for further processing.

* **Exposure:** The resist is exposed to light through the mask for pattern generation.

  PEB: The **P**ost **E**xposition **B**ake allows to cure the exposed/non exposed resist (in function of the tone) and make it non soluble. The exposed/non exposed part will be insensitive to this step. The post exposure bake allows the resist one wishes to keep (again depending on the tone) to become more resistant during the development. This bake typically lasts 90 seconds at 110°C.

* **Development:** A mild base is then used to strip away the soluble resist. The pattern becomes visible during this step.

* **Hard bake:** An hard bake of 120 seconds at 130°C is then done to fully cure the resist by drying out the remaining solvent. This makes the resist more resistant for further processing steps such as wet or dry etching.

* Once the resist mask has been used, it can be stripped away by acetone or in an oxygen plasma (barrel).

8.3 Resolution

The achievable resolution of a photolithographic process step is a function of both the light source, dose and distance between the photomask and wafer. The impact of the light source is directly linked to its wavelength. In conventional photolithography, the energy transferred to the photoresist for the polymeric reaction comes from the energy of incident photons. Typically, the sent wavelengths are in the UV part of the spectrum with the most common wavelength being of 365 nm.

However, UV remains light. This means that diffraction on the mask will impact the achievable resolution of the process according to:

$$\text{Line Width} \approx \sqrt{\lambda(g + d/2)} \tag{8.1}$$

where the linewidth is the smallest achievable width for a given wavelength λ, resist thickness g and air gap d between the mask and wafer. Hence, for a conventional wavelength of 365 nm, a typically achievable linewidth would be of around one micro meter depending on the air gap. This means that features drawn on the mask that are smaller than that won't be printed on the resist.

Reducing the wavelength and hence the diffraction of light on the mask traces/features leads to a bump in resolution, but at a high cost. The associated processes are **D**eep UV, x-ray and e-beam exposition.

The cheaper way to get the best possible resolution being:

* **Decreasing the air gap:** This defines the exposure position. Hard vacuum, hard contact and soft contact are common positions where the wafer is respectively further and further away from the mask. The best resolutions are obtained when the wafer is pulled against the mask by a vacuum. However, this often comes at the cost of mask contamination by resist or resist peeling from the wafer if the adhesion was not properly promoted.

* **Decrease the resist thickness:** A typical resist thickness is in the sub micrometer to one micrometer range. This gives us a resist layer thick enough to deal with ion bombardment for the duration of most dry etching processes. Reducing it could lead to an increased resolution, but at the cost of a less resilient resist mask.

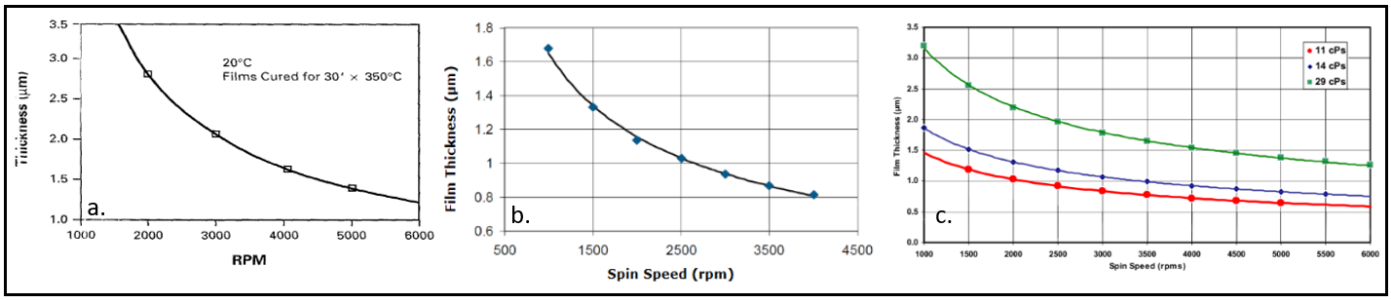


Figure 63: Spin curves for the polymers and resists used in our process **(a.):** Polyimide PI2545. Illustration and product by [44] **(b.):** AZ nLOFF 5510 negative tone photoresist. Illustration and product by [53] **(c.):** AZ MIR 701 positive tone photoresist. Illustration and product by [54]

9 Grinding

Back grinding is a method used to remove layers in the order of the hundred of micrometers from the wafer to *"quickly"* make it thinner. For example, in our process, the wafer were thinned from 750 µm to 300 µm in four hours. The advantage of this technique is that it allows for a fast reduction of the wafer thickness, but at the cost of massive crystal damage that can extend as deep as 20 µm below the grinded surface.

This process involves a rotating grinding wheel rubbing against rotating wafer and can achieve a typical etch rate in the order of a micrometer per minute.

Extra care should be given to the preparation of the wafer in order to protect the non grinded face. A layer of protective resist is commonly spin coated on the wafer before applying a thick layer of adhesive protective plastic. Protection of the non grinded face of the wafer is paramount due to the fact that the wafer is held in place by vacuum on a porous rock chuck. Note that good cleaning of the wafer after grinding is also crucial since the wafer will be highly contaminated by dust.

10 Profilometer

The profilometer is a characterisation device allowing to measure the profile and topography of a wafer. It can hence be used to measure not only steps induced by structures on the wafer but also to measure the curvature of the complete wafer. A concrete example of why this could be of interest is to easily compute the stresses generated by the deposition or removal of a layer of material from the wafer. Indeed, each material features an intrinsic compressive or tensile stress. For example, cured spin coated polyimide generates a tensile stress of [50-200] MPa [44]. The addition/removal of material from the wafer hence changes its curvature. By measuring it before and after processing, one can compute the stress in the film as:

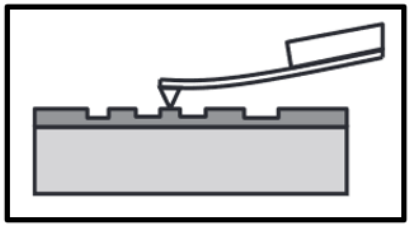


Figure 64: Schematic representation of a profilometer. Illustration from [40]

$$\sigma = \frac{E_s}{6(1-eta)} \frac{t_s^2}{t_f} \left(\frac{1}{R} - \frac{1}{R_0}\right) \tag{10.1}$$

where R_0 and R are the initial and final curvature, t_s is the substrate thickness, t_f the film thickness, E_s the Young's Modulus of the substrate and η its Poisson coefficient.

A profilometer works by dragging a magnetically suspended needle across the wafer. The variation in the needle height being sensed by a coil which generates the measurement. The principle of operation is, to an extent, similar to the way a record player works.

11 Ellipsometer

A ellipsometer is a device used to measure the thickness of transparent or translucent material. It works by shinning a laser at a precise angle at the wafer and sweeping the light source's wavelength. The light is then transmitted, reflected and diffracted by the different layers of material making up the stack of the wafer. A sensor will then measure the amplitude, phase and polarity of the transmitted light.

Prior to the measurement, the layer stack of the wafer, an approximation of the thickness of each layer, and their optical properties should be known. The computer will then make an estimation of the diffraction pattern and will try to fit the measurement to the estimated curve.

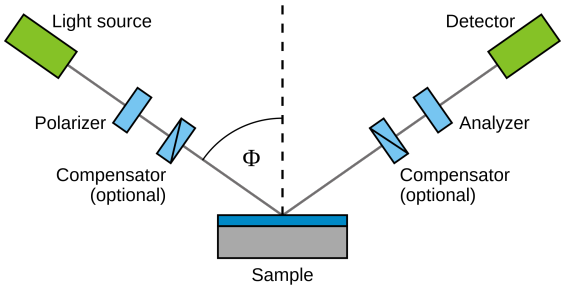


Figure 65: Schematic representation of ellipsometry. Illustration from [55]

A successful ellipsometric measurement can not be made if the optical coefficient of the material are not known. Indeed, many combinations of thickness, stack and diffraction coefficients might give the same measured curve. In most cases, the stack and optical parameters are known. This allows the computer to only having to fit the measure and estimation curve by sweeping the thickness of the stack's layers.

Part IV
Process

Now that the design of the pixels and membrane has been discussed, let's look at the micro and nanofabrication process devised for this master thesis. To keep the document readable, the process in itself has been decorrelated from the explanation of the fundamentals of each used technique. All the techniques are presented in a generic way in Part III. This is done so that this section can focus on what was effectively done in the cleanroom without getting tangled in theoretical considerations.

This section will hence start by looking at the different processes that have been considered for this thesis and will then proceed to a complete run-through of the work that was done in the cleanroom of UCLouvain: the *Winfab*.

1 The choice of the initial wafers

The first thing to consider was the type of wafer to be used. As explained in the state of the art, the reference article [1] uses thinned monocrystalline silicon (10 nm) which is the top layer of an SOI wafer. In our case, as was already explained, using thin silicon would have proven to be challenging both in terms of manipulating/processing such a thin layer, but also in the way such a layer was obtained.

Indeed, what the team did was use a succession of thermal oxidation of the silicon and then etching of the silicon dioxide. Reducing the thickness of the silicon to under 100 nm would have been possible at Winfab using the KOYO1 and then performing a BHF wet etch. The issue with that, as with many other problems in micro- and nanofabrication, is that no recipe exists at Winfab to do that step in a controlled manner. A process should have been designed and it was decided very early on that we should instead start the process with a standard non altered SOI wafer to save some time.

The choice of the silicon on insulator wafer themselves was tailored by the available stock at UCLouvain. At the beginning of the process, wafers featuring a top silicon layer of 150, 300, 500 and 700 nm were available. The final decision was taken to go for the 300 nm one because the stock of 150 nm top silicon SOI was low and that 300 nm thick was hence the thinnest top silicon available. The wafers, taken from the stock of *SOItech* hence feature a top silicon layer of 300 nm with a {111} plane crystal orientation. The different crystallographic orientation can be seen in Figure 38.

As we will see, the process being long and featuring different variations, four wafers were ordered and named SOI1, 2, 3 and 4. The idea was to have three of the wafers go through the SP1 and SP2 processes and one go through the SP3 process. However, as we will see, the SP3 process was dropped and hence, *in fine*, all wafers went through the SP1 and SP2 processes. This high number of wafers does however allow us to make mistakes. Since processing takes a lot of time it would be too risky to only process one wafer. Should it be ruined at the end, the entire process would have to be rebooted from scratch. The time available for a master thesis leaving not a lot of room for process issues, having several wafers gives the opportunity to the operator to make mistakes and not worry about losing one wafer in the combat.

As we know, the crystalline orientation of silicon will influence its response to strain. One test that we wanted to make was to see how polycrystalline silicon would react when strained. For this reason, the decision was taken to manufacture what is called in French: *"Des SOIs du pauvre"* that one could call in English: *"The poor man's SOI"*. The subsection here below details the fabrication process of what we will call HMSOI.

1.1 Homemade SOI manufacturing

To asses the effect of strain on polycrystalline silicon, what we did was to manufacture DIY (**D**o *I*t **Y**ourself) SOI wafers. The idea was to replicate the real SOI that we got from SOItec.

The first thing to do was hence to take some measurements of those real SOIs prior to the design of a process recipe. The wafers were measured using the ellipsometer (the bulk thickness was measured in a vise). The results are shown in Table 6. The expected values from the datasheets were respectively for the top silicon, BOX and bulk of 0.3, 0.1 and 725 µm. The measured values hence matching our expectations.

Figure 66 shows a schematic view of the process used to build our HMSOIs. Here are some details:

* **Wafer preparation:** The initial wafers used for this process are 3 inch monocrystalline bulk silicon ones. Their thickness was measured using a vise and found to be of 380 µm. This thickness difference with the SOI is however not critical since both the SOI and HMSOI will, in term, be grinded down to 280 µm (step 9 of section 2). Also note that (seen from the flat position of the bulk silicon wafers) the top crystal plane orientation is not the same as the one of the true SOI. This is irrelevant here due to the fact that the goal is in any case to have an active

top layer of poly-Si. The wafers were marked using a diamond needle[23] with their names for future reference. The wafers then went through standard cleaning which is mandatory prior to entering the KOY01 furnace.

* **Wet thermal oxidation:** Wet thermal oxidation was used in order to produce the BOX of our DIY SOI. Oxidation was preferred over CVD deposition since the BOX of conventional SOIs are made of thermal oxide. Since our goal was to gimmick the 1 µm BOX of our real SOIs, the wet option was chosen to save some time. The maximal achievable oxide thickness using the Winfab's KOYO1 being 600 nm, this is the final expected thickness. In practice, through ellipsometric measurement, the real thickness was found to be of approximately 540 nm, meaning a deviation of 10% from the expected thickness. This thickness was obtained after a four hour long wet thermal oxidation at 950°C hence an effective deposition rate of 135 nm/hours which is in line with what was said in section 4. Note that the consumed thickness of silicon is here negligible in comparison with its 380 µm thickness (we have in theory consumed $0.45t_{ox} = 0.45 * 540 = 240$ nm of silicon or 0.06% of silicon). Also note that both sides of the wafers were oxidised.

* **LPCVD:** What was left was to deposit a 300 nm thick layer of polycrystalline silicon on top of the wafers. This was again done in a KOYO, but this time the KOYO2. Again, the wafers were placed in the KOYO after a complete standard cleaning (but without a HF bath to avoid removing the just grown silicon dioxide). The deposition, on the top of the silicon dioxide, was done using a standard recipe stored in the KOYO1, which called for a deposition at 620°C. The exact duration of the deposition itself is hard to determine due to the pumping time and transfer time in and out of the furnace, but could be estimated to fit in the [0.5-55] min range due to a deposition rate of [0.1-10] nm/s. The final value measured with the ellipsometer is given in Table 6. The LPCVD process occurs on only one side of the wafers.

These wafers are a copy of the real SOIs featured in our process. They will hence follow the same route as their cousin true SOIs. Also, to come back to the discriminating name of *"Poor's man SOI"*, let's say that the price of each of these HMSOIs was estimated to be of around 180€. This estimation takes into account the price of the initial bulk wafer (11€) and the cost of each process step. All the pricing came from an internal web page of the Winfab website. All things considered, these wafers cost around 80€ more than store bought SOIs which are around 100€ (still not taking into account the time of the involved personnel).

Finally, the complete derailed process with every single little step is given in section 1.

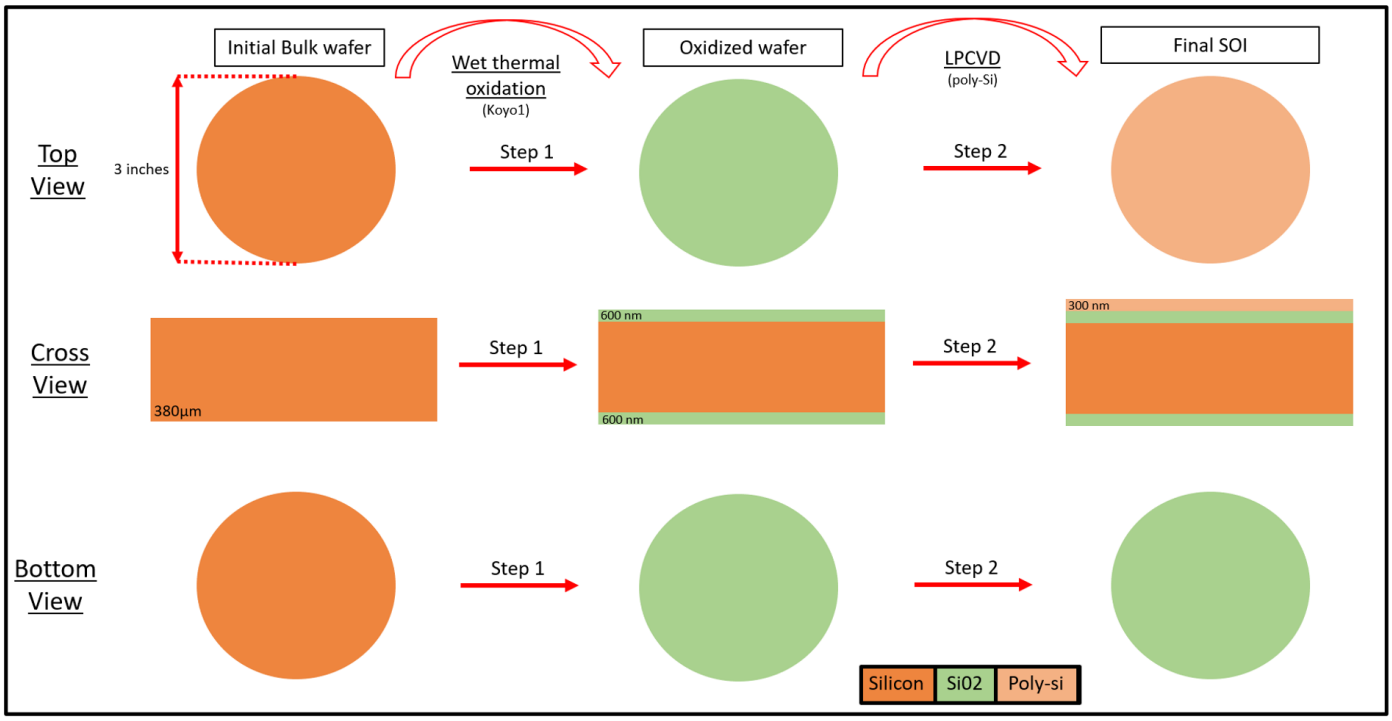


Figure 66: Fabrication process to build homemade SOI wafers

[23] These are like pencils with a pointy diamond head that is used to engrave wafers

	SOI_1	SOI_2	SOI_3	SOI_4	HM_SOI_1	HM_SOI_2	HM_SOI_3
Top silicon [nm]	303	303	303	302	329	330	331
BOX [nm]	1050	1050	1050	1050	542	543	544
Bulk [µm]	725	725	725	725	380	380	380

Table 6: Initial thicknesses of the SOI and HMSOI wafers. Used ellipsometer coefficient for the silicon dioxide: $N_0 = 1.452$ and $N_1 = 36$

2 Considered process design

Now that all the wafers are ready to go, let's start by taking a step back to look at what we have to build. When zooming out of all the design details about the thickness of the membranes, the doping of the transistors and so on, the goal of the here introduced microfabrication process is to build the following device:

"Set of doped silicon pixels with aluminum interconnects on a polyimide membrane"

Three different versions of the process have been considered for this master thesis. They can be described as follows and the final devices can be seen in Figure 67:

* **SP1**: **S**ub **P**rocess 1 has been designed in order to be compatible with the rectangular dies featured on the mask and visible in Figure 42. The idea for these dies was for them to follow the SP1 process almost all the way to the end. Indeed, once the polyimide spin coated and the wafers diced, the SP1 and SP2 processes part ways. From there on, the rectangular dies should have been bound to a layer of gelPak for structural integrity before etching away all of the remaining bulk and then all of the BOX. The final result would have been stripped of plastic carrying the polyimide membrane embedding the pixels and electrical connections. The contacting of the pixels would have been made from *"underneath"* the device. Indeed, the contact would have been covered by the polyimide and gelPak from the top while being accessible from the bottom due to the complete removal of the BOX and bulk. This process variation allowing for the fabrication of completely flexible dies.

* **SP2:** This process is, as was said, the same as the previous one all the way to the end. Indeed, at the end of this process, an under etching hard mask of aluminum is built to allow etching away the bulk and the BOX only below the membrane. This means, as one can see in Figure 67, that the dies will be made of a flexible and free standing polyimide membrane with a *"thick"* 280 µm bulk silicon frame (the remains of the SOI stack to be precise).

* **SP3:** The SP3 process was scrapped and hence not developed/validated for this thesis. The final possible configuration is shown in Figure 67, but is only here as an illustration and to start a potential discussion for another process. The goal of the SP3 was to use the 1 mm in diameter membrane designs that one can see in Figure 42 to build a silicon dioxide membrane. A layer of silicon nitride would have had to be deposited as passivation to reduce the compressive stress of the silicon dioxide.

Note that, as one can see in section 12, all three proposed processes share the same beginning at least up to step number 8.

3 Execution of the SP2 process

In this section, the execution of the SP2 process will be demonstrated. This includes, but is not limited to: the explanation and justification of the process steps, the parameters and settings as well as measurements, results and discussion. Let's start with an overview of what is ahead of us:

3.1 SP2 process overview

The process to be followed is illustrated in section 12. A total of 18 steps modifying the stack (the succession of layers) have to be performed. However, if one only looks at the main process steps, we end up with the following list:

* **Photolithography 1:** Fabrication of the pixels in the top silicon. Step 1 and 2 on the schematic
* **Photolithography 2:** N doping of the pixels. Step 3 and 5 of the schematic. [1]
* **Photolithography 3:** P doping of the pixels. Step 4 of the process schematic[1]
* **Photolithography 4:** Contact opening prior to metallization. Steps 6 and 7 of the schematic

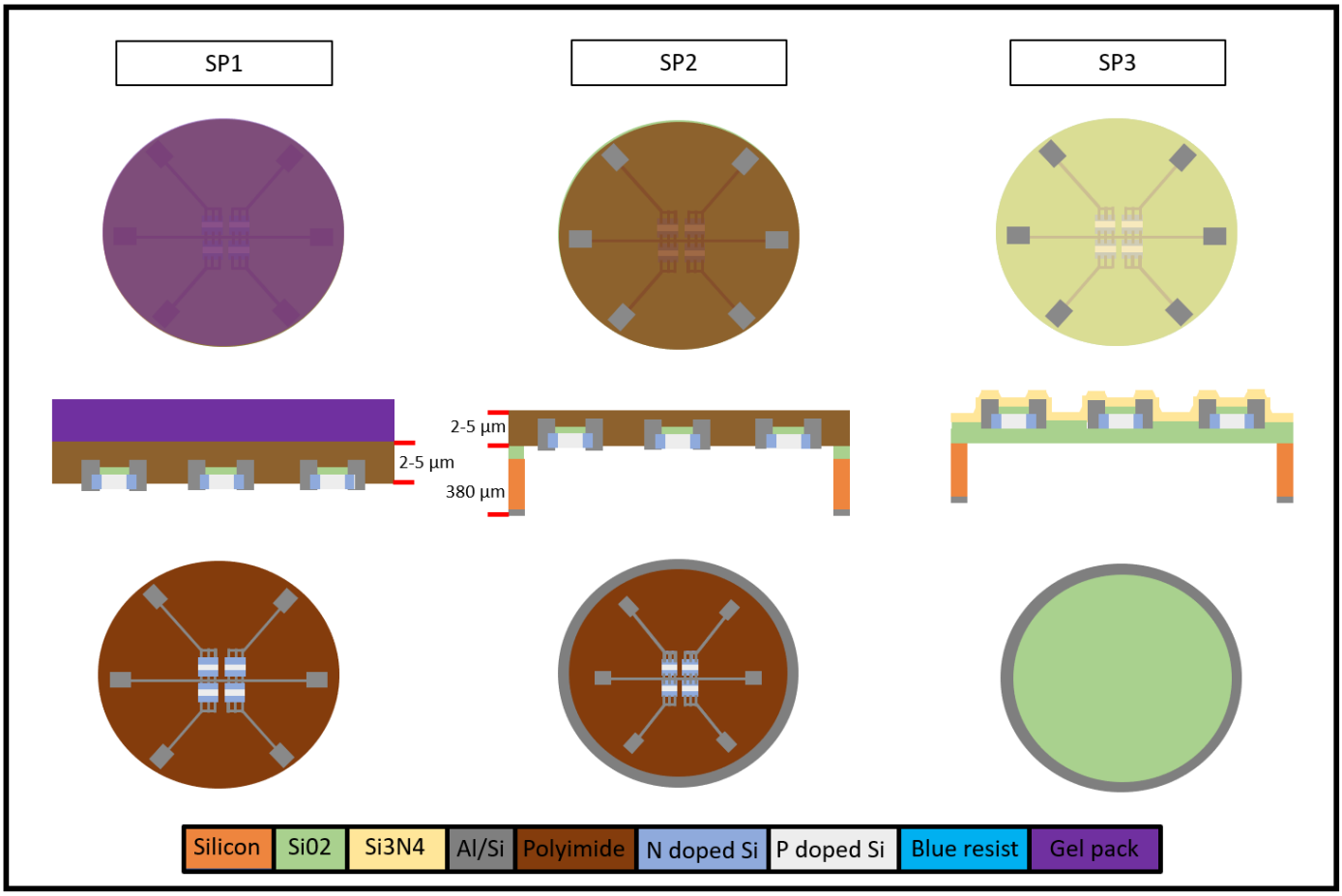


Figure 67: Comparison between the final device obtainable via the SP1, SP2, SP3 processes. **(Left)** Loose flexible dies **(Middle)** Free standing membranes on a thick silicon frame and **(Right)** Passivated silicon dioxide membrane

* **Photolithography 5:** Metallization of the wafer. Definition of the electrical contacts. Step 8 of the schematic
* **Wafer thickness reduction:** The wafers are thinned by back grinding. This is step 9 of the process
* **Photolithography 6:** Back side metallization. Step 10, 11 and 12.
* **Photolithography 7:** Polyimide spin coating and contact opening. Step 13 and 14.
* **End steps:** Dicing and membrane release.

Where Green is for the completed process steps, Yellow for the attempted process steps and Red for the process steps remaining. All these steps justify the number of masks (6) that we had to design. Indeed, each photolithography requires a custom mask to transfer the design on the wafer. In our case, we have six masks and seven photolithographies, because the fourth and fifth ones share the same one. The masks used in this process are given here below in Figure 68 in order of appearance.

One last comment before starting the description of process. The schematics and descriptions that are given here below only present the most important operations. The real life process has over 80 individual steps including the processing of test wafers, cleaning and characterisation. A typical process is described by the combination of a schematic and a step-by-step detailed list of every single small processing steps with all the parameters and comments. This description is the real life process that the wafers go through. It is given in section 2.

[23] Note that on all of the schematics of section 12, the P doping comes before the N doping. In practice the opposite was done (N then P) because we had difficulties deciding on a P doping level. The order of the doping does not alter the performance in any way

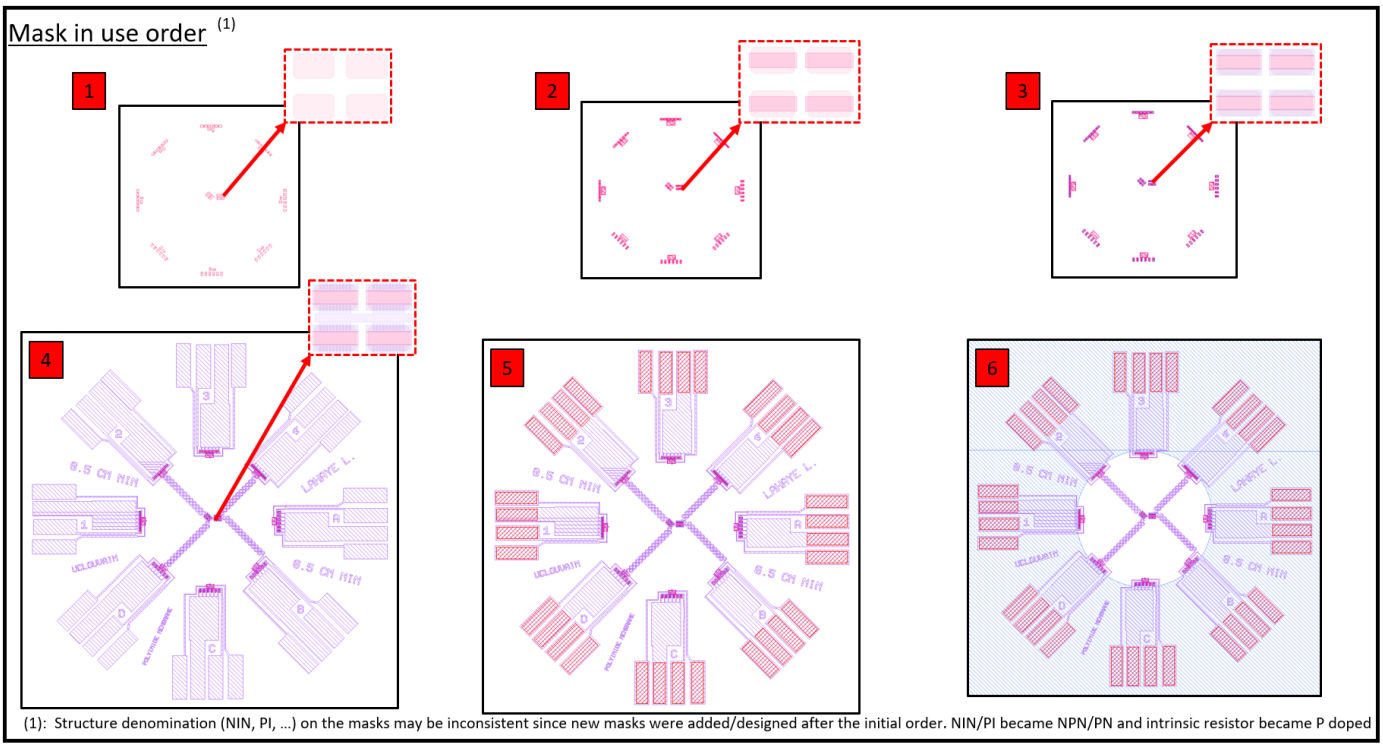


Figure 68: Masks used in the SP2 process by order of use.

4 Photolithography N°1: Silicon pixel definition

This first step of the process concerns the fabrication of the silicon pixels in the top layer of the SOI and HMSOI wafers. This part illustrates the steps 1 and 2 of the process schematics and runs from step 1 to 14 in the complete process sheet. We start with the SOI and HMSOI wafers. Let's start with the first photolithography.

4.1 Lithography 1

The goal of this step is to transfer the pattern for the silicon pixel onto the wafer. Since this is the first lithography to be done, here is the protocol followed to prepare the surface/promote the adhesion of the resist (this procedure was then done for all the other photolithographies and will hence not be repeated):

* **Standard cleaning:** Two baths of piranha solution and a bath of HF to remove organic contamination and the layer of native oxide (a native oxide is naturally created in a matter of minutes on top of air-exposed silicon).
* **Nitrogen degassing:** Quick degassing in the TEMPRESS for 15 minutes at 800°C under a 1.5 lmin flow of nitrogen.
* **HMDS application:** a 30 minutes process during which gaseous HMDM is applied on the surface of the wafer.

With the wafers ready, it was time to apply the photoresist. The mask featured in Figure 68 was bought as a lightfield. This means than what is drawn translates in chromium drawing on the real mask. The goal of this photolithography is to cover the regions where the pixels will be and hence protect them from the etchant coming just afterwards. Since the mask is a lightfield and that we want the final resist left on the wafers to be a copy of what's on the mask, the resist to be used has to be a positive tone. This way, the exposed resist will be stripped away during development.

The positive tone AZMIR701 photoresist by [54] was spin coated using the *Suss Gamma 80* which is an automatic resist processing robot. As was seen in the part of this document about the general principle of photolithography, resist must go through several steps namely: coating, soft baking, exposure, post exposure bake, development and hard bake. The *Suss Gamma 80* is a robot that performs all of those steps apart from the exposure. The robot was hence set to deposit a layer of positive tone resist at a spin coating speed of 4000 RPM which results in a resist layer of 0.7 µm according to Figure 63.

Once the soft bake (90°C for 90 seconds) was done, the wafer could be moved to the *Suss MA6* which is a device used to align a mask and wafer and perform the exposition of the photoresist through the mask with a desired dose. The dose

was selected to be of 180 mJ/cm^2 in a hard vacuum mode to get the best resolution possible. All the settings for the coming lithography will be formatted as is done in Table 7 and not detailed unless a problem occurred. Note that this first lithography did not require an alignment since it was the first one.

Figure 69 shows the result of this first step. One can see from the quality marker that the exposition settings were correct since all features are visible. The pink color is silicon.

Mode	Wavelength [nm]	Dose [mJ/cm²]	Contact mode	Mask	Resist	Spin [RPM]	SF [°C/s]	PEB [°C/s]	HB [°C/s]
TSA	365	180	H-V	Light	AZMIR701+	4000	90/90	110/90	130/120

Table 7: Settings for the first photolithography (mesa manufacturing in the top silicon)

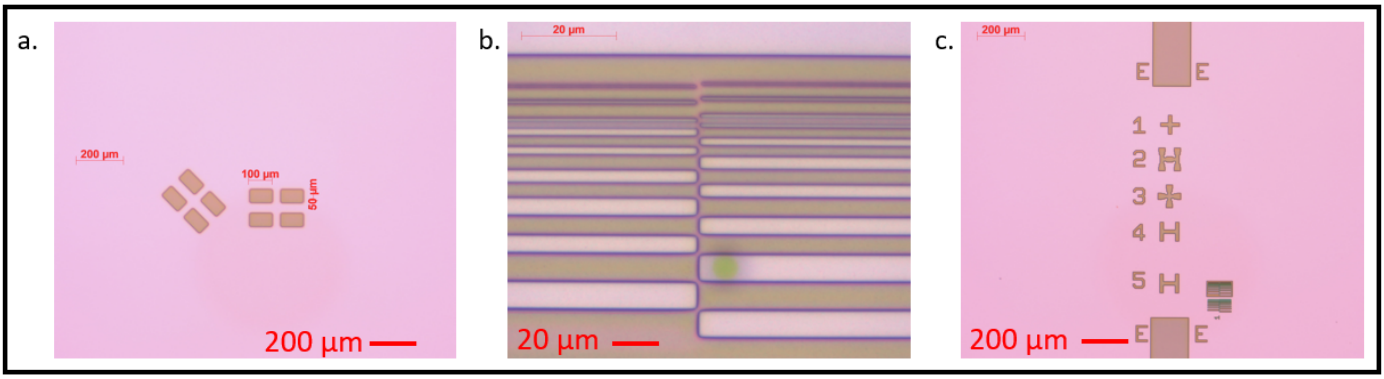


Figure 69: Results after the first photolithography for the manufacturing of the mesa in the top silicon

4.2 Top silicon etching: ICP-RIE in the Corial

Due to the fact that the mask featured pixels oriented along both the <100> and <110> crystal axes as can be seen in Figure 38, a KOH wet etch could not be used. Indeed, as was seen in subsubsection 1.3.1, the anisotropic nature of mono-Si wet etch would render the fabrication of the 45° tilted pixels impossible. Indeed, a considerable under etch would occur under the mask along the least etch resistant crystal planes. The result would be a rectangle with dimensions embedding the tilted pixel design. This effect is explained in subsubsection 1.3.1.

Hence, to deal with that issue, an ICP-RIE dry etch in the Corial was selected as our etching method. The go-to recipe at Winfab for silicon etching in the Corial involves the use of SF_6. The dissociation of this gas in the plasma generating fluorine radicals is able to etch the silicon. However, at the time this etching was set to take place, there was no more SF_6 available. In order for us not to wait, we decided to design our own dry etch recipe of silicon using CHF_3 gas.

A first test wafer was used in order to determine the etch rate of silicon. The wafer, B1, had already been coated and patterned like all of the wafers in the previous lithographic process. It was then determined through profilometric measurements with the Dektak that the etch rate of mono-Si was roughly of 2.7 nm/s or 160 nm/min under the reported conditions of Figure 70 in the reactor during etching.

CHF3 [sccm]	O2 [sccm]	RF [w]	ICP [w]	P [mTorr]
50	3	50	800	3025

Figure 70: Setting of the Corial during operation

After that initial test, the wafers were processed one by one in the reactor. The theoretically required etch time of 1.875 minutes (112 seconds) computed based on the thickness to be etched and reported etch rate is however not exactly the etch time used in practice. Indeed, the conditions of the chamber have a large influence on the etch rate. For example, the etch rate of the first wafer will not be the same as the one of the third processed wafer. This is due to the fact that the conditions (temperature, contamination) of the reactor evolve continuously over time. Hence, care was taken and, as reported in Table 8, HMSOI1 (the first processed wafer after the test wafer B1) was etched in multiple runs (90 then 30 then 30 seconds) with intermediate measurements taken each time to assess the state of the etch. The etch time of the next wafer is each time adapted from the one of the previous wafer explaining the regularity in etching time. Note that no significant variation in etch time is observed. This excess of precaution was taken because we were the first to try this recipe.

Something important to note with the Corial is that the etching has a radial profile. Indeed, the wafer will be etched more quickly at the center than on the edge. This is a second reason why etching periods longer than 112 seconds were used (to make sure that the silicon edge was completely etched). ICP-RIE having a poor selectivity, over-etching at the center (and in general) means that the BOX was etched a bit. The final thickness of the BOX being reported in Table 8. The unequal thicknesses of the center and edge of the wafers confirming the radial profile of the etching in the Corial.

Figure 71 shows the results of the etching of the top silicon. The etching went smoothly overall but, on some rare occasions, the pixels were destroyed by the dry etching process. This was only observed on the real SOI wafers and not the HMSOIs. It is hence not likely that this problem originates from a delimitation of the top silicon from the BOX. Our theory is that, upon development, resist was stripped away from the pixels due to the fact that only some small island had to stick to the wafer. They could have been damaged by the flow of rinsing DI water poured on the wafer to neutralise the developer. Nevertheless, this was only reported on a couple of structures and did not require to re-do the step with a new wafer. Unfortunately, the yield of this step has not been reported.

	SOI_1	**SOI_2**	**SOI_3**	**SOI_4**	**HM_SOI_1**	**HM_SOI_2**	**HM_SOI_3**
Etch time [s]	120	120 5	125	125 5	90 30 30	130	120 10
Border [nm]	962	952	933	945	292	402	481
Center [nm]	983	985	977	986	378	477	527
Dektak [nm]	396	395	401	398	604	503	458

Table 8: Thickness of the BOX after the dry etching of the top silicon. The Dektak measurement represents the step from the BOX to the top of the resist. The measured stack by the Dektak step is from bottom to top: BOXTop-Siresist. Resist was measured at 589 nm with the ellipsometer

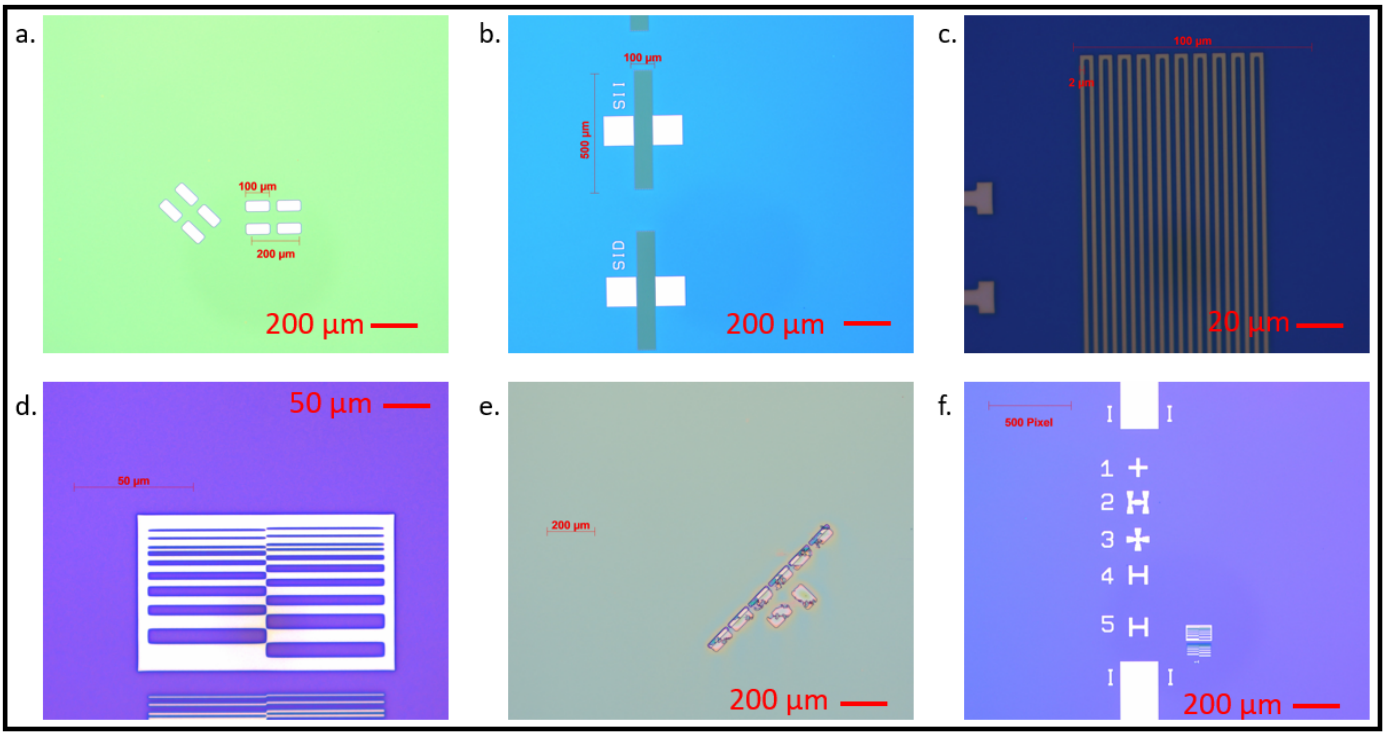


Figure 71: Result after the dry etching of the top silicon

5 Lithography N°2 and 3: P and N-doping

This step covers points three to six of the schematic and runs from step 15 to step 36 of section 2. The goal of this step is to prepare the wafer to be implanted at **I**on **B**eam **S**ervice or IBS. IBS is a french company located in Marseille France. In order for them to implant our wafers, some work has to be done on our side first. The first one, which we

already discussed, is to fix the dose, energy and species of the implantation. These parameters we send them are the one they will end up using to implant our wafers. The second thing to do is to define the regions we want to implant. This is where the lithography numbers two and three come into play. However, prior to this, an oxide layer has to be grown on top of the pixels. Let's tackle that first.

5.1 Wet thermal oxidation

The implanting process must take place across a layer of oxide. This layer of oxide, which as been taken into account in the simulation is used to slow down the incident doping atoms. This layer can however not be the native oxide for two reasons. The first one is that this oxide is too thin (in the 1-3 nm range) and the second is that it is irregular. This would lead to a less homogeneous level of doping across surface or the pixel.

Our wafers are then, after a standard cleaning including the HF bath to remove the native oxide, put into the KOY1 for a dry thermal oxidation. There, a 27 nm thick layer of oxide is grown resulting in a 12 nm decrease in thickness of our top silicon pixels (4% of the total thickness).

5.2 Lithography N°2: N-doping

This lithography is the one used to define the N region of our NPN and PN junction. The mask to be used can be seen in Figure 68 and is a darkfield mask. The goal here is to cover all the pixels and only open windows in the resist where we want the doping atoms to be implanted. This means that a combination of the darkfield mask and positive tone resist can be used. The regions to be implanted will be exposed and washed out during development.

The difference between the previous lithography and this one is that here the mask should be aligned with the patterns already on the wafer. This was done by aligning the markers from the mask and the wafer. The mask being a dark field, its alignment mark is in fact a hole in the shape of a plus sign and must correspond to the plus sign mark etched in the top silicon layer. The result of the alignment can be seen in Figure 72 (f.). The alignment in itself is just a matter of adjusting the X, Y and tilt of the mask relative to the wafer.

As in the previous lithography, the *Suss Gamma 80* was used to deposit a layer of AZMIR701 and perform the soft bake. Initially, all the same settings as previously were used for the exposition and baking. However, upon inspection under the microscope, it was shown that the resist had been stripped away from the pixels as can be seen in Figure 72 (d.) and (e.). Our hypothesis is that this comes from the hard vacuum contact. Indeed, the surface of adhesion of the resist on the pixels is small. We think that the resist delaminated upon release of the contact between the mask and wafer. This time, the damage was such that the step had to be done a second time. The resist was hence stripped in the barrel following the procedure explained in the previous part and put to the photolithographic process a second time. For the second run however, the contact mode was changed from hard vacuum to soft contact. The results of this second run can be seen in Figure 72 (a., b., c. and f.) and all the settings are given by Table 9.

Once this step completed, the wafers were packaged and sent to IBS. The turnaround time for the wafer to come back from France being one week.

Mode	**Wavelength [nm]**	**Dose [mJ/cm²]**	**Contact mode**	**Mask**	**Resist**	**Spin [RPM]**	**SF [°C/s]**	**PEB [°C/s]**	**HB [°C/s]**
TSA	365	180	S-C	Dark	AZMIR701+	4000	90/90	110/90	130/120

Table 9: Settings for the third photolithography for the N implantation

5.3 Lithography N°3: P-doping

Once the wafers have come back from France, the first thing to do is to inspect the wafers under the microscope. We here want to check whether the resist layers held up and did not burn under the particle bombardment. In our case, it did not happen and processing could go on.

What we did was first to strip the resist in the barrel and perform a standard cleaning. The degassing step had to be done at a lower temperature than normal to avoid the activation and disuse of the doping atoms. Indeed, since we will have to activate the boron P-doping atoms later, it is preferable to only bake at high temperature once to avoid diffusing the phosphorus atoms of the N doping twice. Hence, the degassing was done in the TEMPRESS at 600°C for 15 minutes instead of 850°C.

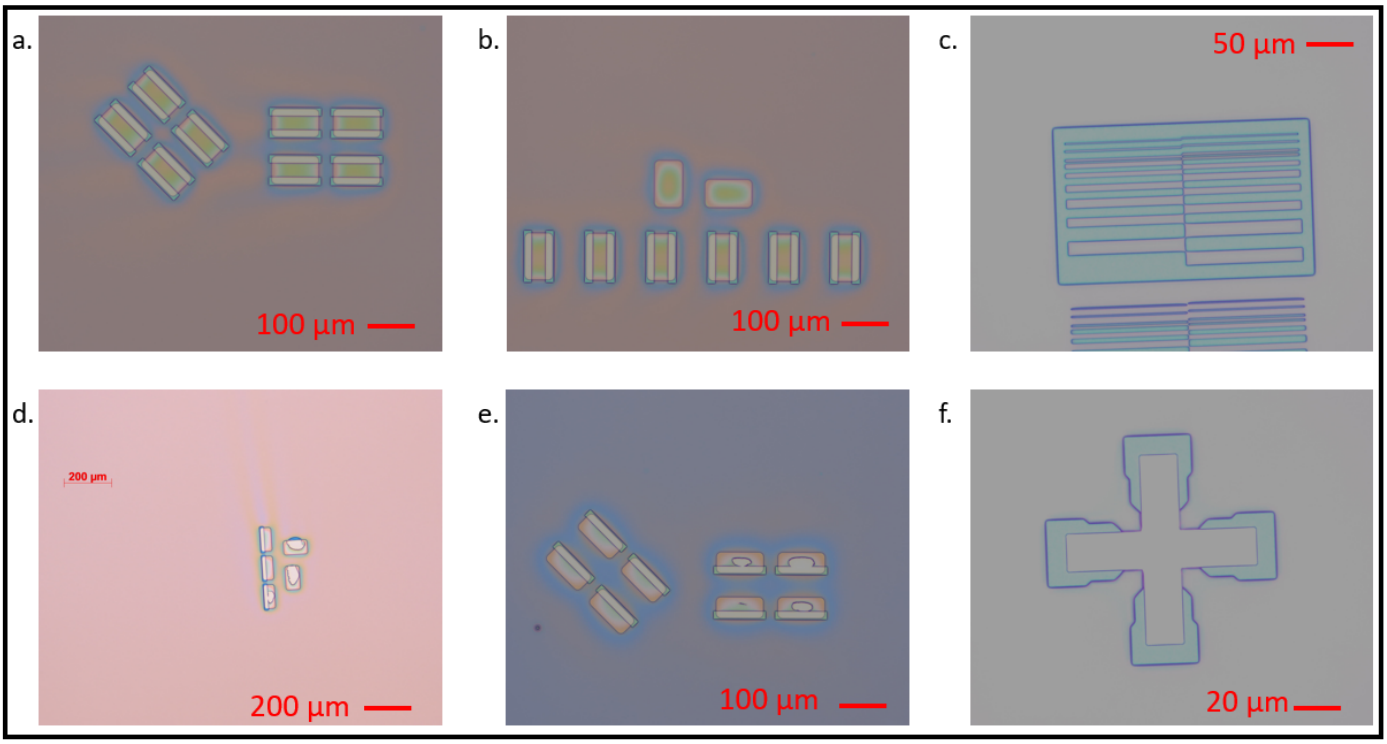


Figure 72: Result of the second photolithography for the N-doping of the pixels

Once done, the wafers went through priming in the LPIII and then a similar positive photolithography as the one for the N doping was done with the third mask of Figure 68. The settings being the same (Table 10), the procedure will not be explained a second time. This time, exposing directly in soft contact, no problem was noticed upon inspection under the microscope.

One can clearly see in Figure 73 the successful alignment (a.) as well as the opened windows in the resist. The resist in pictures (b.) and (c.) has an orange tint whereas silicon dioxide is blue and silicon is white.

Once the inspection completed, the wafer were once again sent to IBS for implantation.

Mode	Wavelength [nm]	Dose [mJ/cm²]	Contact mode	Mask	Resist	Spin [RPM]	SF [°C/s]	PEB [°C/s]	HB [°C/s]
TSA	365	180	S-C	Dark	AZMIR701+	4000	90/90	110/90	130/120

Table 10: Settings for the second photolithography for the P implantation

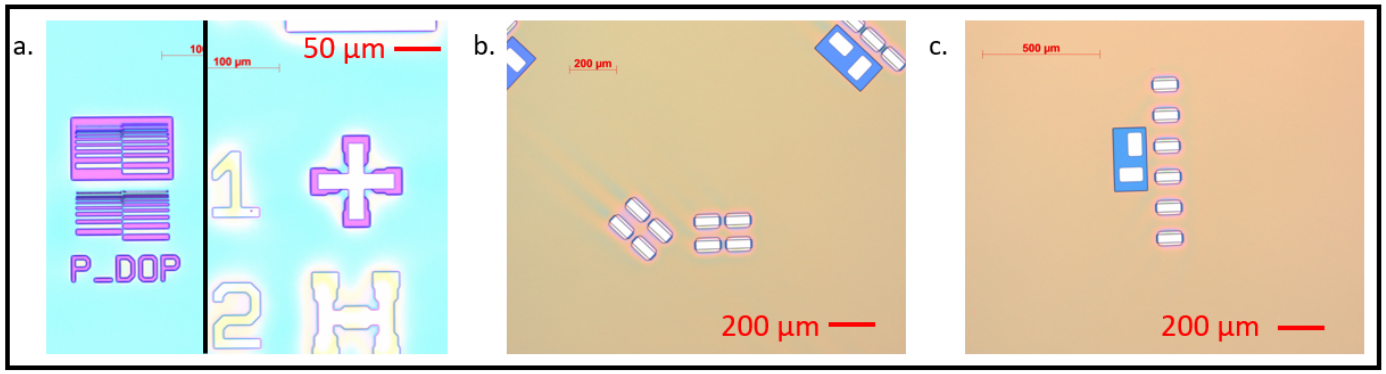


Figure 73: Result of the second photolithography for the P-doping of the pixels

5.4 Wafer RTA and passivation

Once the wafers come back from their second trip to IBS, we once again inspect them to see if the resist is intact. Since it was the case, we immediately moved on to proceed with the RTA and the passivation.

After the resist had been stripped from the wafers in the barrel, they went through standard cleaning to allow them to enter the TEMPRESS. There, a DIY RTA took place. As a quick reminder, **R**apid **T**hermal **A**nnealing is the step during which the doping impurities are activated and fixed into the crystal lattice. A normal RTA is usually done at 950°C or above for a minute. In our case, a 30 minutes long baking at 850°C replaced it because the RTA of the Winfab is currently out of order. This was however accounted for in the simulation of the doping profiles.

This fake RTA served at the same time as a degassing of the wafers. Indeed, we took the opportunity of the wafers going through cleaning and degassing to deposit a layer of passivating silicon dioxide right after.

Indeed, currently the pixels' only electrical insulation is a 27 nm thick layer of thermal oxide. This is not enough and, as a precaution, a thicker 200 nm layer of silicon was deposited onto the front of the wafers.

The depositions were made in the CVD reactor of the *Oxford Plasmalab*[24]. The used technique being PECVD. A recipe available on the computer of the Oxford was used to deposit the silicon dioxide and as can be seen in Table 11, the expected deposition rate was of 45 nm/min for the given parameters.

However, before performing the deposition on our real wafers, a test was made on a bulk silicon wafer (the wafer B5 in section 2). The deposition took 245 seconds and the final oxide thickness, as measured by the ellipsometer, was of 163 nm (both at the center and edge of the wafer). This means that the actual rate was of 40 nm/min.

The recipe was hence modified and the deposition time was increased to 275 second to get a deposition thickness of 184 nm.

The results are provided in Table 12. As one can see, the thickness is, each time, close to the targeted value.

SiH4 5% [sccm]	**N20 [sccm]**	**N2 [sccm]**	**P [torr]**	**T [°C]**	**Expected rate [nm/min]**	**Real rate [nm/min]**
100	700	1600	1	300	45	40

Table 11: Settings and rates for a deposition of a 200 nm layer of silicon dioxide (PECVD the *Oxford Plasma lab*)

SOI_1 [nm]	**SOI_2 [nm]**	**SOI_3 [nm]**	**SOI_4 [nm]**	**HMSOI1 [nm]**	**HMSOI2 [nm]**	**HMSOI3 [nm]**
182	188	179	187	181	197	187

Table 12: PECVD silicon dioxide thickness after a 276 second deposition at 40 nm/min. Measured on the ellipsometer at the center of the wafer.

6 Photolithography N°4: Contact opening in the oxide

Before moving on to the metallization of the wafer, the spot where the aluminum will contact the silicon must be opened. Indeed, at this point, the pixels are completely covered by the just deposited layer of PECVD silicon dioxide. The contact fingers must now be opened prior to the metallization for an electrical contact to occur. This is made via a photolithography and a wet etch of the silicon dioxide. This section covers the steps six and seven of the schematic and the steps 37 to 45 of section 2.

6.1 Photolithography N°4: Electrical contact opening definition

The goal of this lithography is to create a mask to be used in wet etch application. The goal of the mask is to protect the deposited silicon dioxide layer everywhere except where the electrical contact will be.

This lithography is done using the same mask as the one to be used for the metallization process. Indeed, this mask (the fourth in Figure 68), defines the pattern of all the wiring including the contacting fingers. It is thus also suited to open the mentioned fingers in the oxide layer.

[24]The oxford plasmalab is, a machine able to be operated to do either ICP-RIE etching, as with the Corial, or CVD deposition. It might be referred to as *The Oxford* later in the text

The mask was brought as a lightfield as can be seen in Figure 74. This means that, since we want to open this pattern in the resist, a negative tone should be used. Upon exposure, the illuminated resist will become less soluble and the pattern will be stripped away in the developer leaving the PECVD oxide exposed.

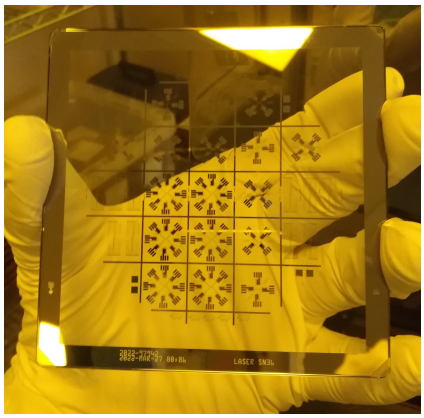


Figure 74: Metallization mask in the yellow room of Winfab

After careful preparation of the surface like always (but without the HF bath during standard cleaning to not remove the oxide), some AZnLOFF 5510 by [53] was spin coated on the wafer by the *Suss Gamma 80* automatic coater. Once done, the exposure was carried out in the *Suss MA6*. Note in Table 13 that the settings for this resist are not the same as before. This is due to the fact the resist is different from the one used before. These settings are in accordance with the datasheet provided by the manufacturer [53]. The soft contact mode between the mask and the wafer was kept in order to ensure the success of this step. Indeed, even if this comes at the cost of some resolution, the fact that we have designed our masks with a 2 µm of acceptable error of margin compensates for it.

Figure 75 shows the result of this step. As one can see from picture (d.), the alignment of the mask and the wafer is spot on. The other picture shows the resist pattern successfully defined. Note that no damage or imperfection in the pattern was seen on any of the wafers.

Mode	Wavelength [nm]	Dose [mJ/cm²]	Contact mode	Mask	Resist	Spin [RPM]	SF [°C/s]	PEB [°C/s]	HB [°C/s]
TSA	365	120	S-C	Light	AZNLOFF 5510	3000	90/90	110/60	130/120

Table 13: Settings for the lithography N°4. Negative tone resist and lightfield mask. Expected resist thickness of 0.9µm.

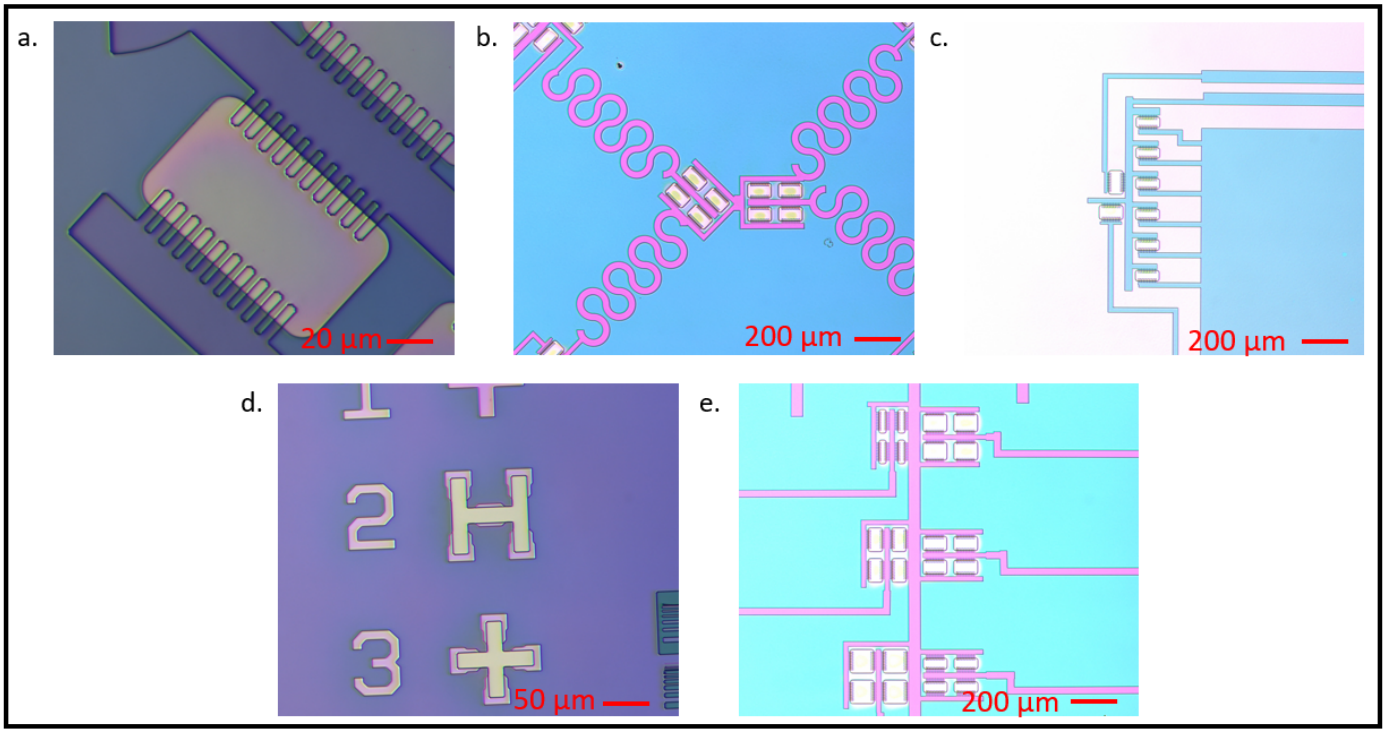


Figure 75: Results of the fourth photolithography for the opening of the contact pads in the layer of silicon dioxide.

6.2 Wet etching of the oxide

Now that the photoresist mask has been defined, the oxide can be opened. Note that in our case one should not only etch away a layer of PECVD oxide, but also the thin 27 nm layer of previously grown thermal oxide. The etch rate of

both of them had to be determined to avoid a too important under etch (etching of silicon dioxide in BHF being fully isotropic).

To this end, the B2 wafer used as a test for the CVD deposition, was repurposed for an etching test. In parallel, a bunch of bulk test wafers (B3, B4 and B5) were wet thermally oxidised in the KOYO1 following the same procedure as in subsection 1.1. This oxidation yielded the same results as previously with a silicon dioxide layer measured at 538, 539 and 539 nm, respectively.

The test wafers B2 (PECVD) and B5 (thermal) were then dipped in a conventional BHF solution of HF/ NH4F 1:7 for 30 seconds. The etch rate of both oxides were then computed based on the results of Table 14. As expected, the etch rate of the PECVD oxide was much higher than the one of the thermal oxide. Note however, that the final measured thickness of the PECVD oxide on B2 can be considered as a complete removal of the oxide layer and an imprecision of the ellipsometric measurement. This would mean, if no oxide is left, that the etch rate is potentially much higher. According to the literature, this etch rate of PECVD silicon dioxide in BHF could even reach the µm/min[52]. On the other hand, the etch rate of the thermal oxide is well into the expected range.

From here, and considering the fact that our oxide layer is made out of a stack of 27 nm thermal oxide and 180 nm PECVD oxide, a total etch time of 18 + 30 seconds was decided.

However, since we were suspicious about the computed etch rate of the PECVD oxide and were expecting it to be much higher, the real wafers were not processed all at once on the first try. Indeed, at first, the SOI1 wafer was sent *"to the battle field"* alone to allow the measurement of the under etch and, if needed, the etch time could be reduced/increased. This was also done in order to check if everything was properly etched. One thing of which we were also afraid was that the highly doped layer of thermal silicon dioxide would feature an even lower etch rate than the measured 86 nm/min of undoped thermal oxide. If that was the case, the etch time would need to be increased in order to go through that layer of doped oxide, hence leaving more time for a lateral under etch of the PECVD oxide layer.

In the end, we ended up leaving the SOI1 wafer for 55 seconds in the BHF bath. From there, the wafer was brought to the microscope and an under etch of 1.45 µm was noticed. This means, as we were suspecting, that the actual etch rate of the PECVD silicon dioxide layer is of around 1.58 µm/min which corresponds to the literature [52].

This 55 second etch duration is the time we ended up using. Figure 76 shows the results of the wet etching of the silicon dioxide. Due to the fact that our contacting fingers measure 4 µm in width, and since a 2 µm margin is allowed, the 1.4 µm under etch of the oxide is considered as acceptable for our application. Furthermore, as one can see from the pictures, the silicon dioxide has been well preserved everywhere else. This step is hence considered a success.

One final consideration: As one can see from both Figure 75 and Figure 76, the fact that we used the metallization mask leads to etching of the pattern in other areas than only the finger contact pads. Indeed, all the spots where metallic traces will end up have been exposed to the BHF. Since all these exposed regions (outside of the contacting fingers) are on the BOX, this will lead to a reduction of its thickness at those given locations. This means that the metallic traces will end up being nested in tranches dug in the BOX. The final thickness of the BOX where those tranches have been opened are given for each wafer in Table 15.

B2 Initial [nm]	**B2 Final [nm]**	**Thermal oxide Etch rate [nm/min]**	**B5 Initial [nm]**	**B5 Final [nm]**	**PECVD Etch rate [nm/min]**	**SOI1 Etch rate [µm/min]**
532	495	86	163	1.16	>324	1.58

Table 14: 30 second dip test in BHF to determine the etch rate of PECVD and thermal oxide and etch rate of PECVD from testing on the SOI1 wafer. Expected etch rate of 63 nm/min for the thermal oxide and 1 µm/min for the PECVD oxide.

SOI_1	**SOI_2**	**SOI_3**	**SOI_4**	**HM_SOI_1**	**HM_SOI_2**	**HM_SOI_3**
182	188	179	187	181	197	187

Table 15: BOX thickness of the wafers after the oxide wet etch.

7 Photolithography N°5: Metallization

This step is about the deposition of the metal traces onto the wafers to contact the silicon pixels where the oxide was just removed. This step had to be done right after the opening of the contacts to limit the thickness of the immediately forming native oxide. Since both process were done in a row, the native oxide should be in the 1-2 nm range and should

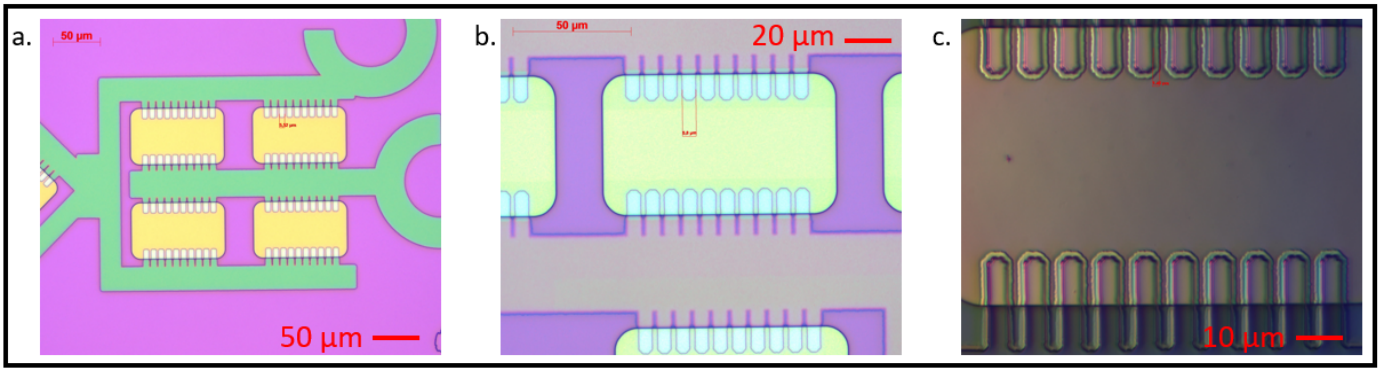


Figure 76: Result of the contact opening in the top oxide covering the pixels.

hence easily breakdown not causing an issue. This section illustrates the step 8 of the process and covers the steps 46 to 55 of section 2.

7.1 Al-Si sputtering

Right after the wafers were wet etched in BHF, the resist masks were stripped in the barrel and the wafers where brought to the VST. The VST is a chamber featuring a planetary type rotating wafer holder allowing for both thermal evaporation and sputtering.

For this first deposition, the decision was taken to go for a deposition of a 600 nm thick layer of aluminum-silicon. Al-Si is an alloy made out of 1% silicon by weight. This alloy is typically used instead of pure aluminum for contacting silicon since it helps with the ohmicity of the electrical contacts. The desired 600 nm thickness was decided with in mind the project of later wet etching the Al-Si to form the electrical contacts. Hence, since this process is highly isotropic and due to the fact that the traces are 4 µm in width, a higher thickness was not advisable. Indeed, a 600 nm layer already means that an under etch of 600 nm from both sides of the trace would occur reducing the total width of the trace by 1.2 µm from 4 to 2.8 µm. Increasing the thickness also meaning increasing the under etch. 600 nm was considered to be a good compromise between a successful step coverage of the silicon pixels while featuring a wide enough post etch aluminum contacting finger. One spoil which should be mentioned, however, is that the wet etch of aluminum proved to be ineffective in this situation. Since a dry etch in the Corial was used as a replacement, due the anisotropy of this etching method, a thicker ($\geq$ 600 nm) Al-Si layer could have been used without any problem.

Furthermore, still about the step coverage and thickness layer of the Al-Si layer, one should remember that the metallic traces will be in recessed tranches. This means that the step to cross will not only be of the thickness of the silicon pixel but of the depth of the trench plus the silicon thickness. In practice, looking at the worst case shown in Table 15 the biggest step to cross would be of (1050-857)+300 = 493 nm (where 1050 nm is the initial thickness of the BOX of the *"true"* SOI). This means that, even in this worst case, the 600 nm thick desired layer of Al-Si should be enough.

The choice of sputtering, should someone ask, as opposed to thermal evaporation was motivated by the fact that this technique allows for a better step coverage, but, more importantly, allows the deposition of alloys. Thermal evaporation of alloys is indeed not often seen in practice due to the difference of vapor pressure of the considered material, hence the difficulty to have them both evaporate in a controlled manner.

The sputtering was performed under a DC generated plasma only (hence making the process DC sputtering). The used recipe was one developed by Nicolas André for a deposition of a 1.3 µm thick Al-Si layer in 100 minutes. Hence, the decision was taken in our case to go for a 50 minutes long deposition, thus aiming for a final thickness of 650 nm. The parameters of the sputter deposition are given in Table 16 here below. Note that a test wafer was put in the sputtering chamber for later use (B6).

DC power [W]	**DC bias [V]**	**Current [A]**	**P [mTorr]**	**T [°C]**	**Heating source**	**Target type**	**Duration [min]**
306	387	0.8	5	250	Lamp	Al-Si	50

Table 16: Parameters for the DC sputter of a 600 nm layer of Al-Si.

7.2 Photolithography N°5: Definition of the electrical traces

Now that the Al-Si alloy has been deposited everywhere on the wafer, it is time to define the pattern of the electrical traces. This is done through a photolithographic step using once again the metallization mask. This time however, contrarily to the previous lithography, a positive tone resist was used. Indeed, remember that the mask was bought as a lightfield. Since we want the aluminum to be protected where we want it to stay, the exposed resist should leave and the non exposed one should stay on. This justifies the use of a positive resist.

Since the dopant had already been activated, special care had to be taken for degassing. Indeed, since we would prefer to avoid diffusing of the implanted atoms, the wafer was degassed for several hours in a 120°C oven. This temperature being way too low to allow for doping atoms to diffuse in the material. Once done, conventional HMDS priming was carried out and positive AZMIR701 was spin coated on the front side of the wafer.

Initially, the same settings as the ones used in Table 9 were used. However, a completely unexpected result occurred. Indeed, upon inspection under the microscope of the used test wafer (B6), it was found that the pattern of the contacting fingers was completely removed. This kind of missing featured in the photoresist is a clear indication that an overexposure occurred. This was weird at first since the settings of the *MA6* were giving a dose of only 190 mJ/cm^2. This dose being well in the norms. It later appeared upon testing that this phenomenon originated from the soft contact setting between the mask and the wafer. The association of this increased distance (compared to hard vacuum) and high reflectivity of the Al-Si layer led to light bouncing multiple times between the wafer and the mask hence considerably increasing the exposition dose.

The contact was then put back to hard vacuum and this effect went away giving us a successful definition of the pattern in the resist as one can see on the right picture of Figure 77. The complete settings for this lithography are given in Table 17.

Mode	**Wavelength [nm]**	**Dose [mJ/cm²]**	**Contact mode**	**Mask**	**Resist**	**Spin [RPM]**	**SF [°C/s]**	**PEB [°C/s]**	**HB [°C/s]**
TSA	365	180	H-V	Light	AZMIR 701	4000	90/90	110/90	130/120

Table 17: Settings used for the fifth photolithography.

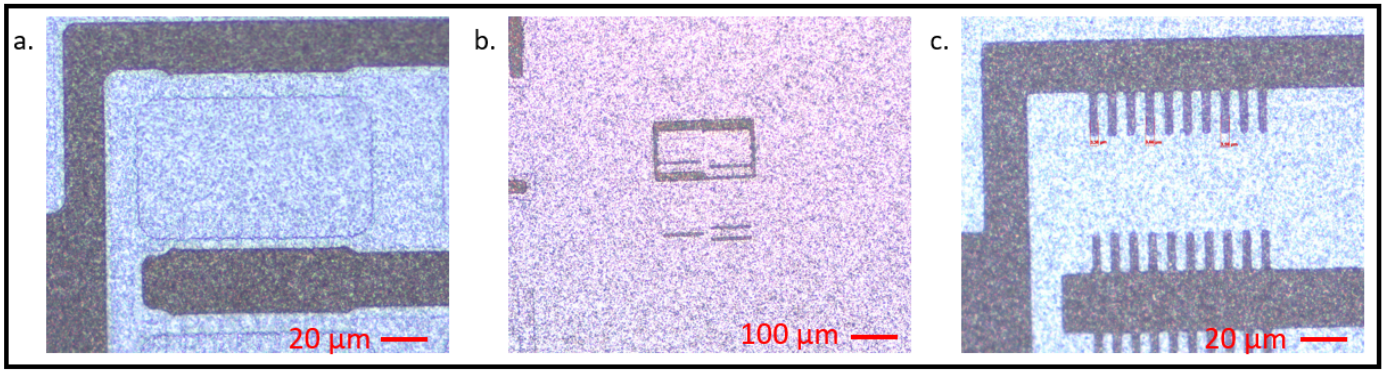


Figure 77: **(a., b.)** Failed N°5 photolithography and **(c.)** Successful N°5 photolithography

7.3 Al etching

All that is now left to do is to etch away the top layer of aluminum to obtain the metallic wires and contacts. Two techniques were used.

7.3.1 Wet etching

The initial plan was, to save time, to go for a wet etch of the aluminum in H_3PO_4. This method allows for a quick etching of aluminum thanks to a high etch rate. The main drawback of this technique, however, is that the etch rate has high dependency on the temperature of the solution.

In practice, what is done is that the phosphoric acid is poured in a beaker and brought to 60°C. Once up to temperature, the wafers are submerged into the solution and the etching is considered over when the bubbling stops. This bubbling

comes from the release of hydrogen gas during the reaction. As one sees from this description, this process is not very controlled. This is due to the fact that since the etch rate is highly dependant on the temperature, relying on a chronometer to determine the required duration of the etch is not possible. For example, our first attempt at wet etching the aluminum started with a solution at 60°C. Then, from the observation of the bubbles, the etch was considered done after 2 minutes. On the second try however, due to the previous etch, the solution had dropped by 4 degrees. This time etch took more than 5 minutes to complete, hence more than twice the time.

This high dependence makes the process highly uncontrollable. In any case, once the etch completes, the wafers are rinsed for ten minutes under running DI water to stop further under etching (this process being isotropic).

One interesting thing emerged under microscope inspection of the wafer. Since the used metal was an Al-Si alloy, some undissolved silicon clusters remained on the wafer. This is due to the fact that phosphoric acid has an almost infinite selectivity toward silicon, as said in [52]. Those clusters are not really an issue beside a potential negative impact on the optical performance of the final device. Indeed, they might interfere a little with the incident light due to the fact that they are on the surface of the silicon pixels.

One solution to remove such clusters is to use the so called *"Sauce de André"*. This product is an aggressive solution of $HNO_3 : H_2O : HF$ in a 50:20:1 ratio. To use it, some must be poured in a container and the wafer should be dipped in it for two seconds. A longer time would lead to a catastrophe since this solution is a rapid etchant of silicon. The impact on the cluster is clearly visible by comparing both the pictures **(a.)** and **(b.)** featured in Figure 78. As one can see, the density of silicon clusters is reduced by a significant amount.

The success of this step, however, is limited to a successful demonstration of the efficiency of the *Sauce de André*. Indeed, even we initially thought that the etch went well by looking at the contact fingers with the resist mask still on, this was not the case. Indeed, removal of the resist in the barrel revealed a massive under etch of more than 3 µm leaving with sad thin contact fingers of 1 µm or less in width. Figure 78 **(a., b., c., e., f.)** show the contact fingers with the resist still on while picture **(d.)** was taken after the removal of the resist.

This result can be explained by our inability to control the etch rate of the highly isotropic H_3PO_4 Al wet etch process. Indeed, since the etch is stopped based on the ability of the operator to detect the presence of bubbles, it is more than likely that the wafer might stay too long in the etchant, resulting in an extended under etch. Note that even if the wafer was pulled out of the etchant exactly at the correct moment, an under etch would still be present. In our case, since the deposited layer was 650 nm thick, the minimal amount of under etch was of 650 nm on both sides of the finger. Each extra second in the etchant leading to an increase of this under etch.

Note that leaving the wafer in longer than required is an obligation since the top of the wafer will etch more slowly than the bottom. Indeed, the beaker being on a hot plate, the liquid at the bottom will be hotter and hence feature and increased etch rate. The thickness of the liquid not helping convection.

In a nutshell, all the HMSOIs went through this process before us noticing the issue. The other wafers, the SOIs, will hence instead be processed using dry etching.

7.3.2 ICP-RIE of the top Al-Si

In order to increase the anisotropy of the etching of the Al-SI, it was decided to use dry etching for the remaining SOI wafers. The wafer were placed one by one in the Corial and a recipe based (CQ9 AL-Si NAZ) on chlorine was used. As can be seen in Figure 80 from the successive etch times, extra care was taken in order to not damage the metallization of our SOI wafers. Hence, after each pass in the reactor, the wafer was removed and inspected under the microscope. The etching being considered done only when no aluminum particles were left to be visible under the highest zoom of the microscope.

Figure 81 shows the results of this step after the removal of the resist. The fingers were measured to be 3 µm thick which is only one micrometer shy of the nominal designed 4 µm thickness. This method hence provided a much higher level of anisotropy compared to the previously done wet etch. The alignment between the metallic fingers and the contact pad can also be seen as good.

Concerning the silicon clusters, it was decided to leave them to avoid damaging our metallic lines by exposing them to the violent *Sauce de André*.

BCl3 sccm	**Cl2 sccm**	**Ar sccm**
10	25	20

Figure 79: Species in the reactor.

SOI_1	**SOI_2**	**SOI_3**	**SOI_4**
30			40
40	40	40	40
30	40	40	40
15	20	40	40
15			30
20			

Figure 80: Etch time in [s] for the SOI wafer for an ICP-RIE dry etching of 650 nm of Al-Si in the Corial. The mean etch rate can be estimated as being close to five nanometers per seconds.

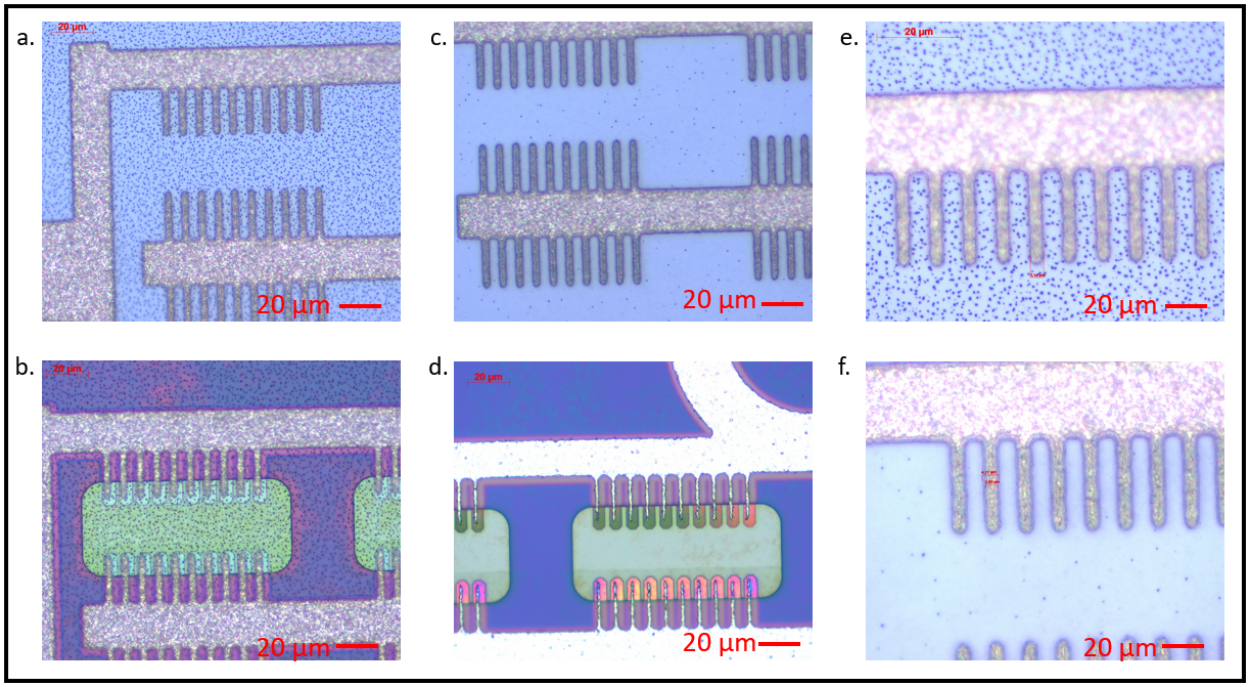


Figure 78: Results of the Al-Si wet etching in phosphoric acid.

Finally, a preliminary electrical measurement proved that the contacting of the pixels was properly made by featuring resistances in the kilo to mega Ohm range. This values of the resistance were expected since before baking the Al-Si at 432°C a Schottky type contact takes place. This explains why the resistance value varies a lot from one device to another and with the applied voltage. Upon baking, the Al-Si will diffuse in the silicon and the energy levels of both materials will align hence creating the Ohmic contact by killing the potential barrier.

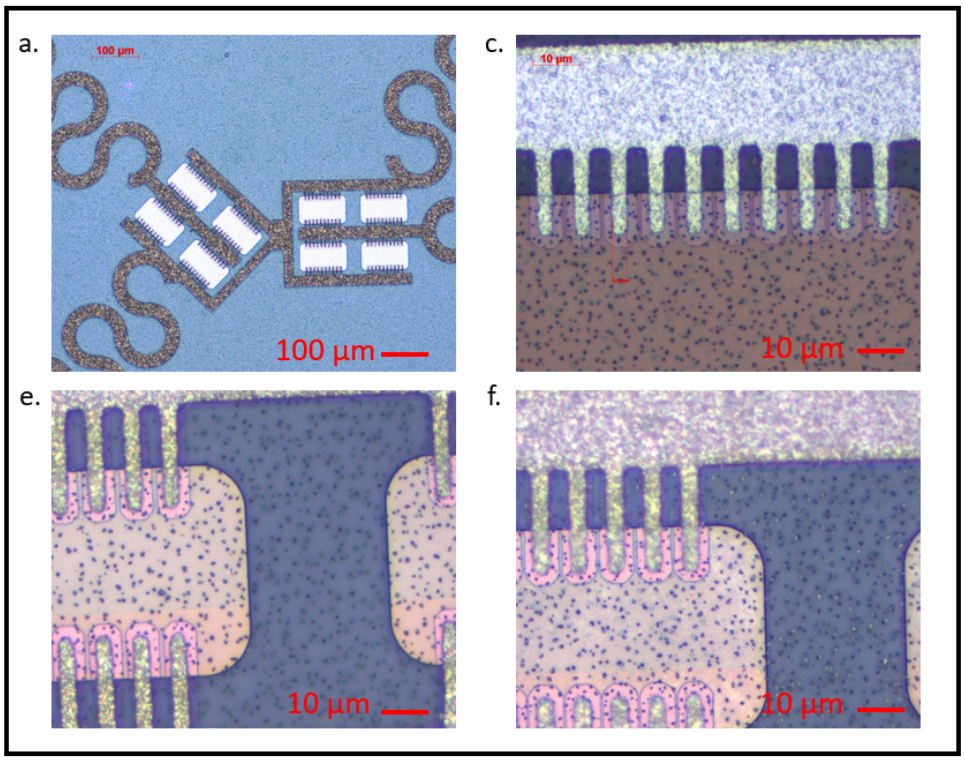


Figure 81: Results of the dry etching of As-Si in the Corial ICP-RIE.

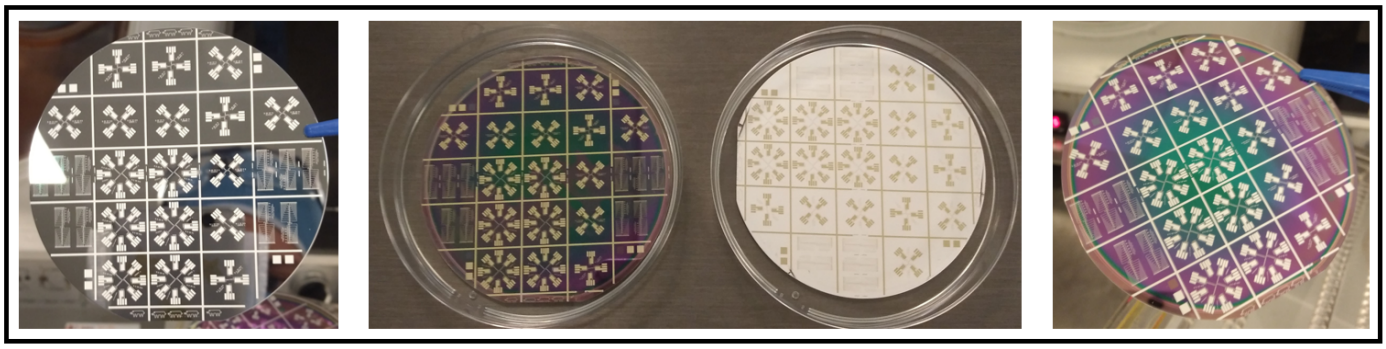


Figure 82: Test wafer B6 and SOI wafer before and after the etching of the top layer al-si. **(Left)** Test wafer B6 after the wet etch. The etching looks successful at the macroscopic scale **(Center)** SOI wafer after (left) and before (right) of the dry etching of the top Al-Si and **(Right)** SOI wafer after the dry etch of the top Al-Si

8 Back grinding

The release of the membranes will later on be done by either a XeF2 isotropic dry etch or an anisotropic Bosh process. In both cases, to make this release easier and faster, the decision was taken to first thin the wafer down by using back grinding. The plan was to use it in order to decrease the thickness of the SOI and HMSOI wafers from 750/380 µm to 280 µm. This section covers the step 9 of the schematic and the steps 56 to 62 of section 2.

After the resist of the metallization step had been removed, some protective measures were taken to protect the front sides of the wafers. The first one was to deposit a layer of blue resist. The blue resist is a non photodefinable resist used to protect wafers. It was spin coated using the automatic *Suss Gamma 80*. Once done, a piece of 125 µm thick adhesive plastic was cut to the side of the wafers and applied on them in the same way one would a apply a screen protector to a phone or a table. The role of this thick film is to absorb any irregularities in the surface of the wafer that could generate a stress point leading to the shatter of the wafer during the actual grinding phase. Care was required during this phase as to not introduce dust particles between this film and the wafer and to avoid large air pockets.

Once the wafers protected and secured by a vacuum on the chuck of the machine, the grinding could begin. The calibration of the machine had to take into account the additional thickness of the plastic film. Indeed, in practice, we tell the grinding wheel to start just above the surface of the wafers to slowly enter into contact with them. Since the zero is the surface of the chuck, we had to tell the machine to start at 750 + 125 µm in height to account for both the thickness of the film and the SOI wafer. Furthermore, we stopped the machine at 425 µm in height which, when removing the film, should leave us with a 300 µm thick set of wafers.

The final thickness of the wafers was measured in a vise to be of approximately 277 µm which is in line with what we were expecting. Considering the fact that the grinding took four hours, this gives us an effective grinding rate of 118 µm/hours. Note that extra care should now be taken when manipulating those thin wafers.

Once completed, the plastic layer could be peeled off and the blue resist stripped away with acetone. At this stage, the contamination of the wafers was high due to the silicon dust generated by the grinding. Heavy rinsing under running DI water was performed for 10 minutes to remove it.

9 Photolithography N°6: Back side hard mask

Now that the wafers have been made thinner, the back side aluminum hard mask for the back etching and release of the membrane could be built. This section covers the steps 10 to 12 of the schematic and the steps 63 to 71 of section 2.

The metallization of the back sides of the wafers had to be done once again in the VST. Hence, cleaning was required, firstly to not contaminate the VST and secondly to promote the adhesion of the aluminum layer. The problem is that, since the electrical contacts were already built at this stage, a standard cleaning could not be used anymore. Indeed, the piranha solution is known to damage the surface of aluminum. A non standard cleaning of the wafer in fuming nitric acid was thus decided in order to remove any organic contaminant without damaging the silicon, silicon dioxide or aluminum featured on the wafers. This step is done by dipping the wafers for five minutes in the acid before rinsing.

Since the wafers had just been cleaned the occasion was taken to put them in the 432°C oven for 30 minutes to render the aluminum on silicon contact effective. Indeed, upon heating, the aluminum will diffuse a bit in the silicon, their

energy levels will align and the generated Schottky diode will vanish leaving the place to a, supposedly, ohmic contact. This temperature does not generate an issue relative to the diffusion of doping atoms.

Once this intermediate step was done, the wafers were moved to the VST after having covered their top surface with protective blue resist.

9.1 Thermal evaporation

Thermal evaporation in the VST was used to deposit a layer of 300 nm of aluminum. Pure aluminum is here used as opposed to the Al-Si used for contacting the silicon. The reason for that is that pure aluminum will be more resistant to a silicon etchant than to an alloy made partially out of silicon. The second reason it that the goal here is to avoid an electrical contact. Thus having silicon in the metal to promote the ohmicity of the contact is of low interest. In addition, aluminum pellets used for thermal evaporation are cheaper than using the Al-Si target. The thickness of 300 nm was decided on an arbitrary basis. Typical aluminum hard masks are in the 200 to 300 nm thick range. We picked 300 nm as a safety net to ensure to have a thick enough layer everywhere on the wafers.

The recipe called for a deposition rate of 5 Å/s where Å is called an Ångström and is worth 0.1 nm. At this rate the deposition was set to last 600 seconds for a deposition of 3000 Å or 300 nm. The parameter for the e-Beam thermal evaporation are given in the table here below 18.

Just note that for both the sputter and the e-beam thermal evaporation in the VST, even if the deposition in itself is short, the entire procedure takes a lot of time (a matter of 5 hours at least). This is due to the fact that it takes a long time for the machines to reach the required levels of vacuum. Indeed, one must not only remove the air from the chamber but also allow for all the moisture in the machine to evaporate and be sucked out. Pumping taking in general, between four hours to a night (the longer being the better). All of this, again, for a deposition in the minutes range. This means that planning ahead and maybe combining the procedure with the process of someone else is required when using the VST.

Rate [A/s]	**Time [s]**	**Target thickness [nm]**	**Op Pressure [mTorr]**	**e-beam [KeV]**	**e-beam [mA]**	**Heating**
5	600	300	0.028	9.6	180	No

Table 18: Settings for the thermal evaporation of Al in the VST

9.2 Photolithography N°6

This lithographic step is used to define the opening on the back side aluminum hard mask. The mask was bought as a darkfield and is the last in Figure 68. Due to its polarity, a positive tone resist was used. Indeed, we want what is drawn on the mask to be opened in the resist layer to be able to expose the underlying aluminum to an etchant. After conventional adhesion promotion steps were performed, the lithography went on as in the previous cases with the settings of Table 19.

The one thing to mention is that this alignment of the mask had to be done with the back side of the wafer. We refer to this as a back side alignment or BSA. To do this a picture of the mask is first done from the bottom (we say that the image is *grabbed*) and then the wafer is aligned on that picture. The rest of the procedure is the same.

Mode	**Wavelength [nm]**	**Dose [mJ/cm²]**	**Contact mode**	**Mask**	**Resist**	**Spin [RPM]**	**SF [°C/s]**	**PEB [°C/s]**	**HB [°C/s]**
BSA	365	200	H-V	Dark	AZMIR 701	4000	90/90	110/90	130/120

Table 19: Settings for the sixth lithography for the definition of the back side aluminum hard mask.

9.3 Al wet etch

Since the opening in the masks were in the mm range, an aluminum wet etch in phosphoric acid could be attempted. Indeed, as we saw from the picture of the previous step relating the wet etch process, the macroscopic features did all came out nicely. Hence, since the openings range from 1 to 5 mm, the wet etch could be used.

The procedure is the same: dipping the wafers in the H_3PO_4 until the bubbling stops at a temperature of roughly 60°C. The results are shown in Figure 83. Please note the lack of clusters due to the use of pure aluminum instead of the Al-Si alloys. The parallel lines comes from the pattern of the back grinded underlying silicon.

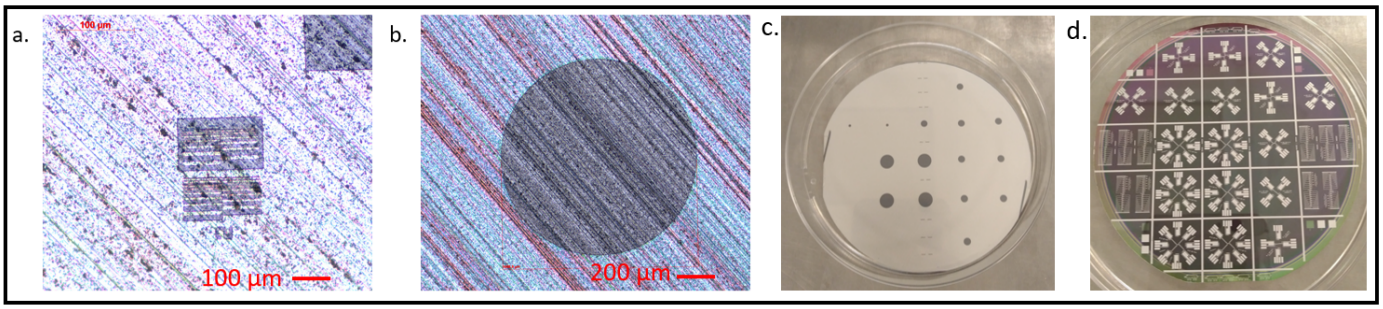


Figure 83: Opening of the Aluminum hard mask.

10 Photolithography N°6: PI processing

We now arrive at the most innovative part of this process. Before going into this part, let's spoils that this step was not completed properly and is hence left to be attempted after the hand in of the master thesis. This step was set to take place on the 12th of August 2022. However, due to cross reservation of the equipment and delay in the cleanroom that day, the process was postponed several times that day. The actual work ended starting late that day, leading to more rush and ending in an hospital room. This means that the process could not be completed in time. The concept and general idea behind it was however considered to be promising and will, as said previously, be completed between the thesis hand in and the oral presentation.

The goal of this step was to spin coat the polyimide on the wafer surface and to define and open the electrical contact pads in the polymer. The first technique that was considered was to do a dry etch as was done in [47]. However, the issue with that step would have been in the difficulty of finding a proper etch mask. Indeed, cured polyimide is hard to etch and a thick or hard mask would have been required.

It was then found, by reading the datasheet of the spin coatable PI2545 by [44], that this particular precursor could be etched before curing by a mild basis and conventional developer. This means that a simple photolithographic resist mask would be enough to define and open patterns in the polyimide prior to complete curing. This meant that the original plans to go for a spin coat of a 10 µm thick layer of PI2611 had to be dropped. Indeed, this precursor is not etched, even when uncured, by a mild basis. The issue with going for PI2545, as one can see on the spin curve, is that the maximal polyimide thickness to be spin coated in a single run is of 2-3 µm. However, it is believed that the ability to easily etch away partially cured PI2545 is more critical than the ability to spin coat a 10 µm layer at once. Indeed, should we need a thicker layer, multiple spin coated layers could be applied. The use of a wet etchable PI2545 is avoiding us the trouble to try to open the contacts using either the laser or a dry etch technique.

The technique here below, which is heavily inspired from the one presented in the PI2545 datasheet [44], gives the procedure to be followed for the spin coating of the polyimide and opening of the electrical contacts. This illustrates the steps 13 and 14 of the process schematic and runs over the steps 72 to 83 of section 2.

1. **Surface preparation:** Since the polyimide layer will be the one carrying our silicon pixels and tasked with transferring load and strain to them, good adhesion is paramount. To this end, the metallized wafer should be cleaned in fuming nitric acid following the associated protocol. Careful degassing of the wafer at 120°C for multiple hours (from 5h to a week) should then be done prior to the HMDS deposition in the LPIII.

2. **PI spin coating:** A 2 µm thick layer of polyimide should be manually spin coated according to the spin curve. The rotation speed is in our case set to be of 2000 RPM for 60 seconds. The coater should be protected by an aluminum crown to collect all spilled liquid polyimide. Polyimide precursor is a thick and sticky product, like old honey, that is not attacked by either acetone, propanol of methanol. The only method found to remove uncured polyimide is to wash it off using fuming nitric acid. Since this chemical can not be used in the coater, careful protection of the wall of the coater with the mentioned aluminum crown is required to eliminate the need to try cleaning it after the fact. The protection of the coater is shown in Figure 84. The fact that polyimide is *almost* impossible to clean is the reason why it cannot be spin coated in the *Suss Gamma 80.*

3. **PI soft bake:** A soft bake of the polyimide is done on a hot plate for three minutes at 140°C to partially cure it.

4. **Photolithography N°7:** A conventional photolithography is done. The mask is the fifth one to appear in Figure 68. It was bought as a darkfield mask and as such should be used in combination with a positive tone resist. Since the polyimide has now been partially cured, the resist can be spin coated on top of it using the manual coater. Once coated, the soft bake, exposure and PEB of the resist are done as usual using the *Suss gamma 80*

and *Suss MA6*. The settings being reported in Table 20. Note the absence of an hard bake due to the fact that the etching of the partially cured polyimide occurs during development.

5. **Puddle development:** Conventional developer is applied to the wafer. This will develop the resist and etch away the underneath partially cured polyimide. The development duration should be of a minute, but a visual verification of the contact opening might be required. Once the development completed, the wafer should be rinsed with DI water for a least two minutes. After that, the resist can be stripped by acetone

6. **PI hard bake:** The polyimide must now be baked. The baking starting at room temperature rising up to 200°C with a 4°C/s ramp and held at this temperature for 30 minutes under a nitrogen flow. Then the temperature should be increased to 350°C with a ramp of 2-5°C/s and held at this temperature for an hour under an nitrogen flow. The cool down ramp should not be over 5°C per minute. This step completes the processing of the polyimide.

Mode	**Wavelength [nm]**	**Dose [mJ/cm²]**	**Contact mode**	**Mask**	**Resist**	**Spin [RPM]**	**SF [°C/s]**	**PEB [°C/s]**	**HB [°C/s]**
TSA	365	200	H-V	Dark	AZMIR 701	4000	90/90	120/90	NA

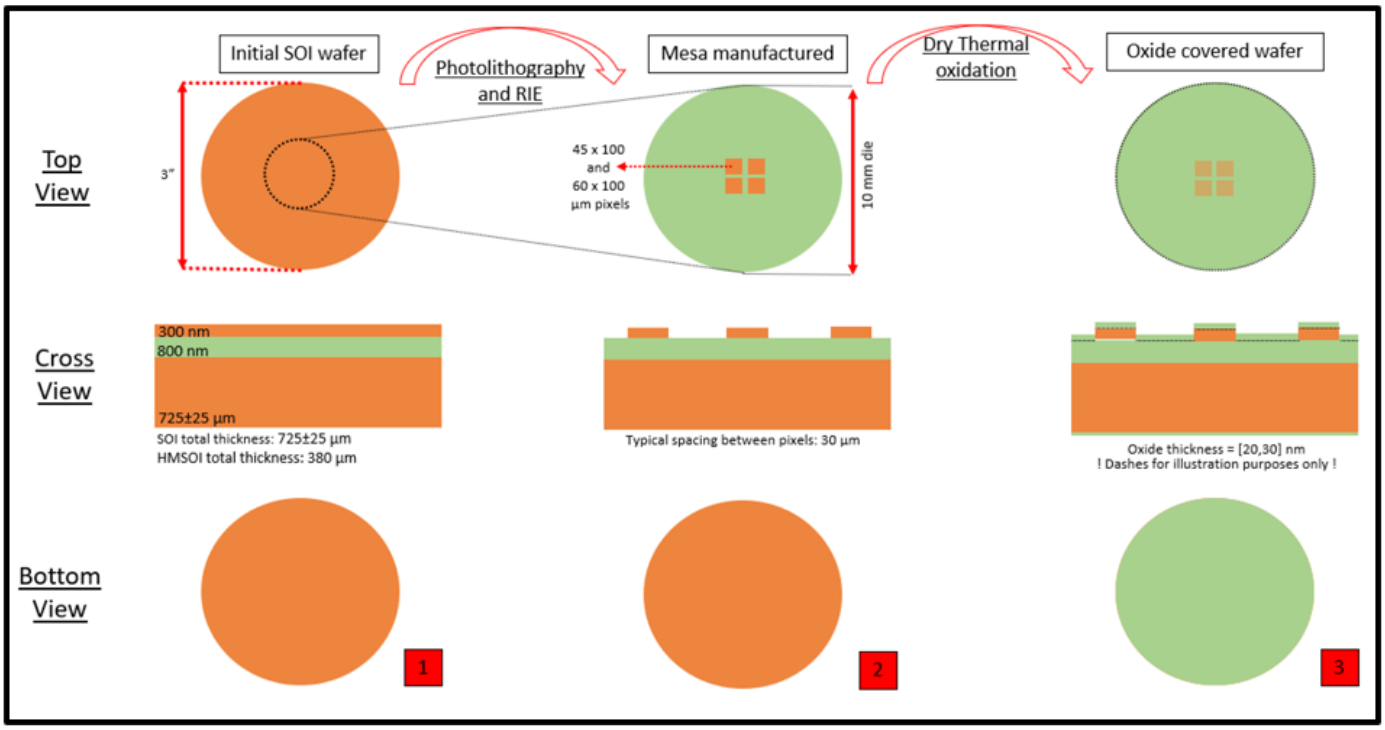
Table 20: Settings for the photolithography for the definition of the pattern in the top polyimide.

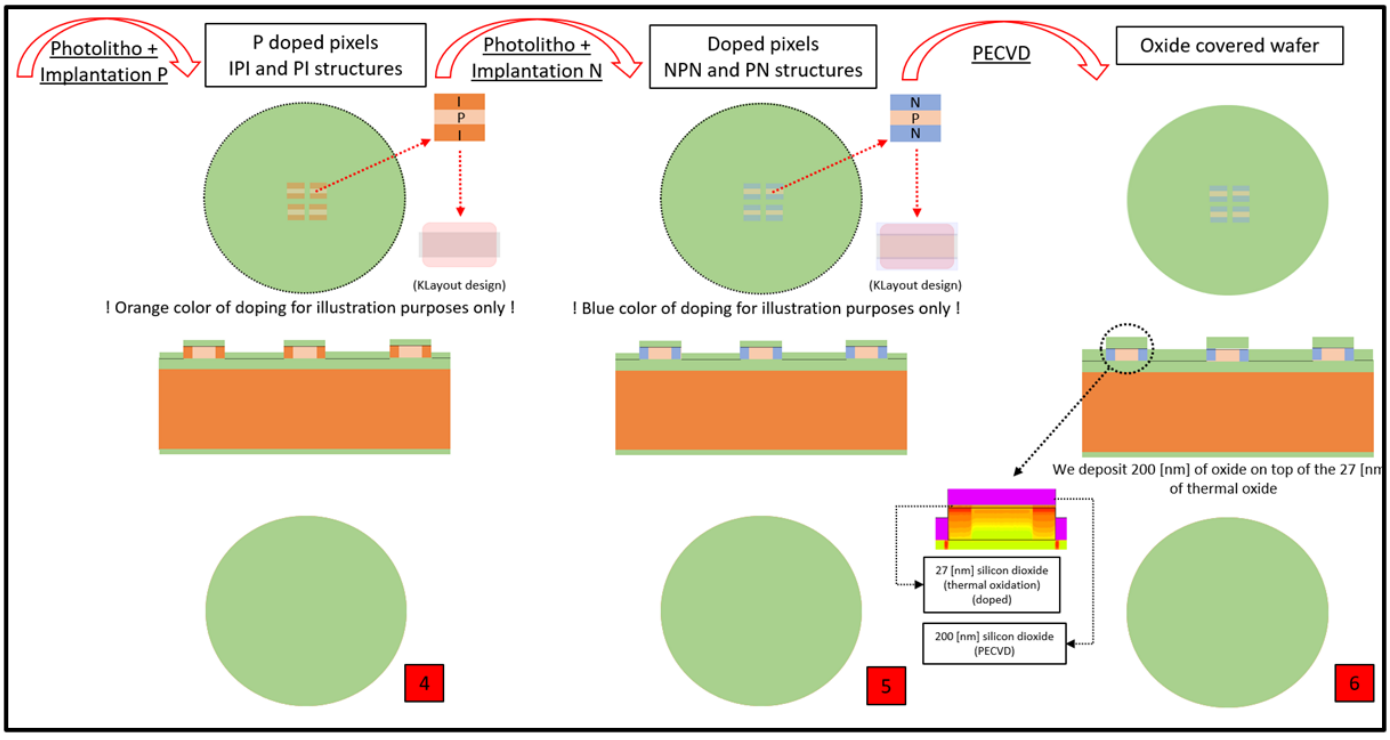
Figure 84: Protection of the spin coater with an aluminum crown.

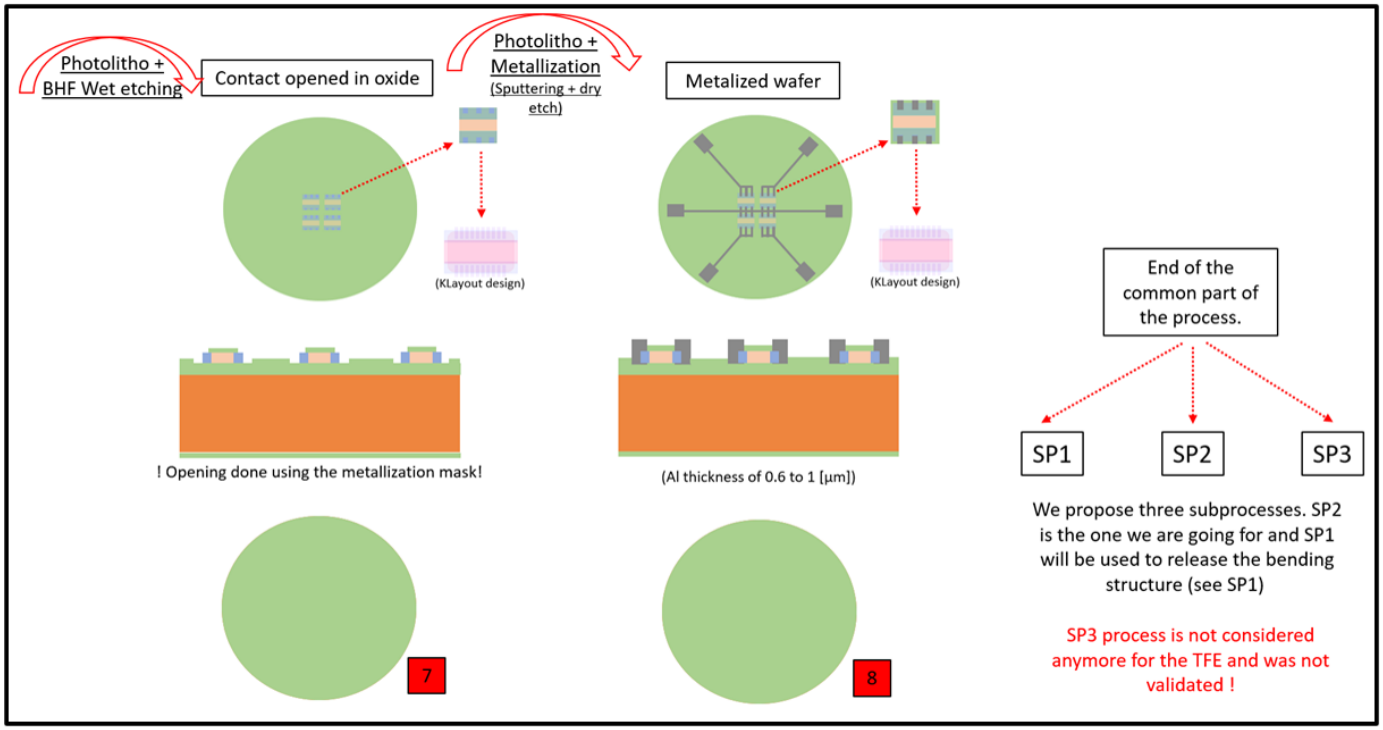
11 Dicing and membrane release

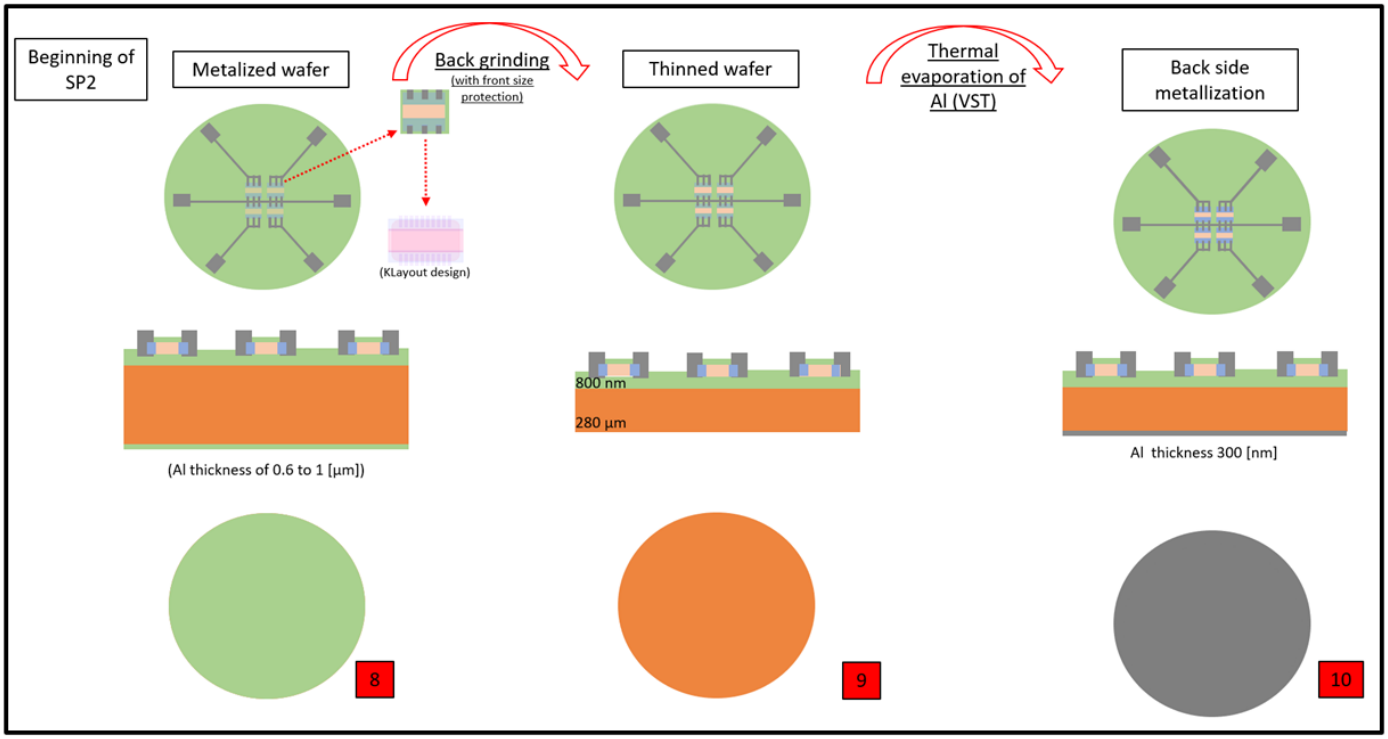
All that remains now is to dice the wafer. Once done, the membranes are either released using an XeF_2 dry etch of the silicon or using a Bosh process. The advantage of the Bosh being its anisotropy. Having an over etch in the lateral dimension is here not really bad since that would only mean that the membranes would be a bit wider. At this point, if one wanted to switch from the SP2 process to the SP1, the back side Al mask must be removed from the concerned dies and these should be bounded to a layer of gelPak prior to the etching of the underlying silicon.

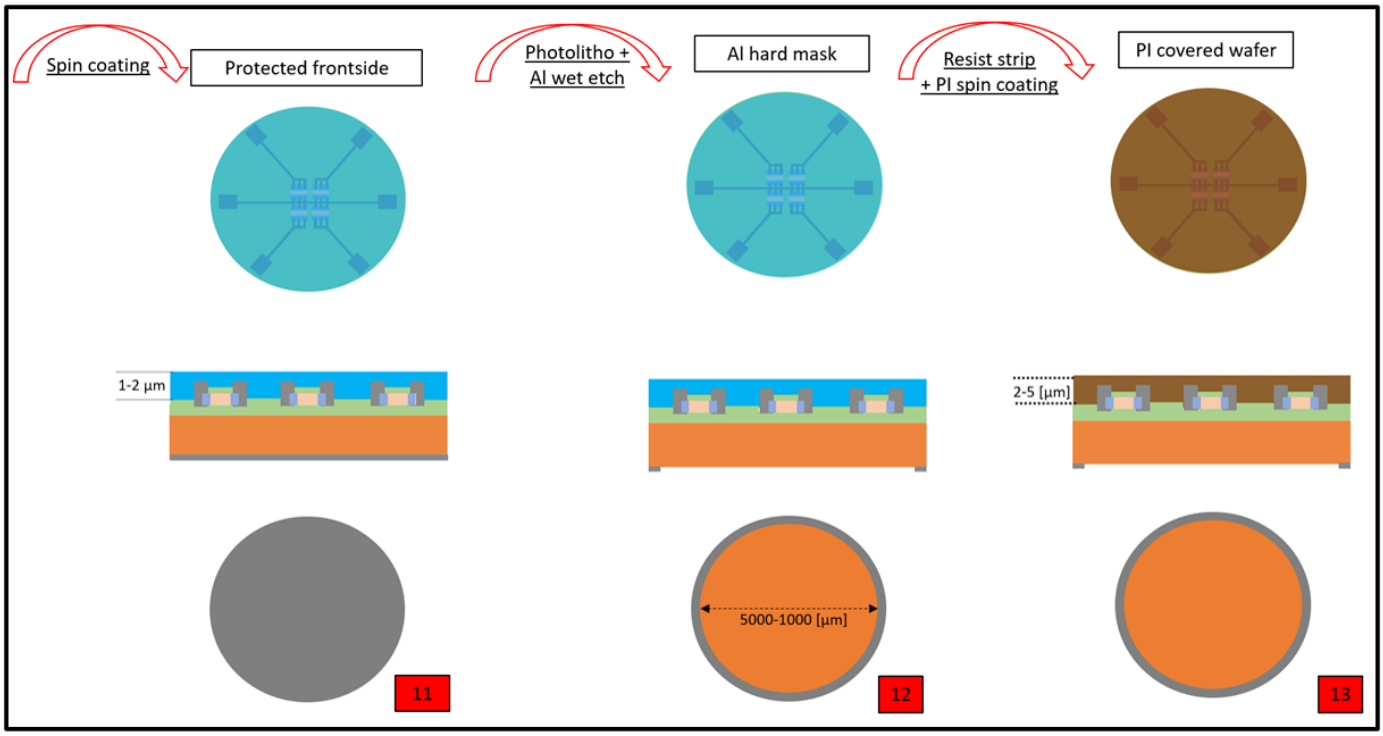
12 SP1 and SP2 process schematic

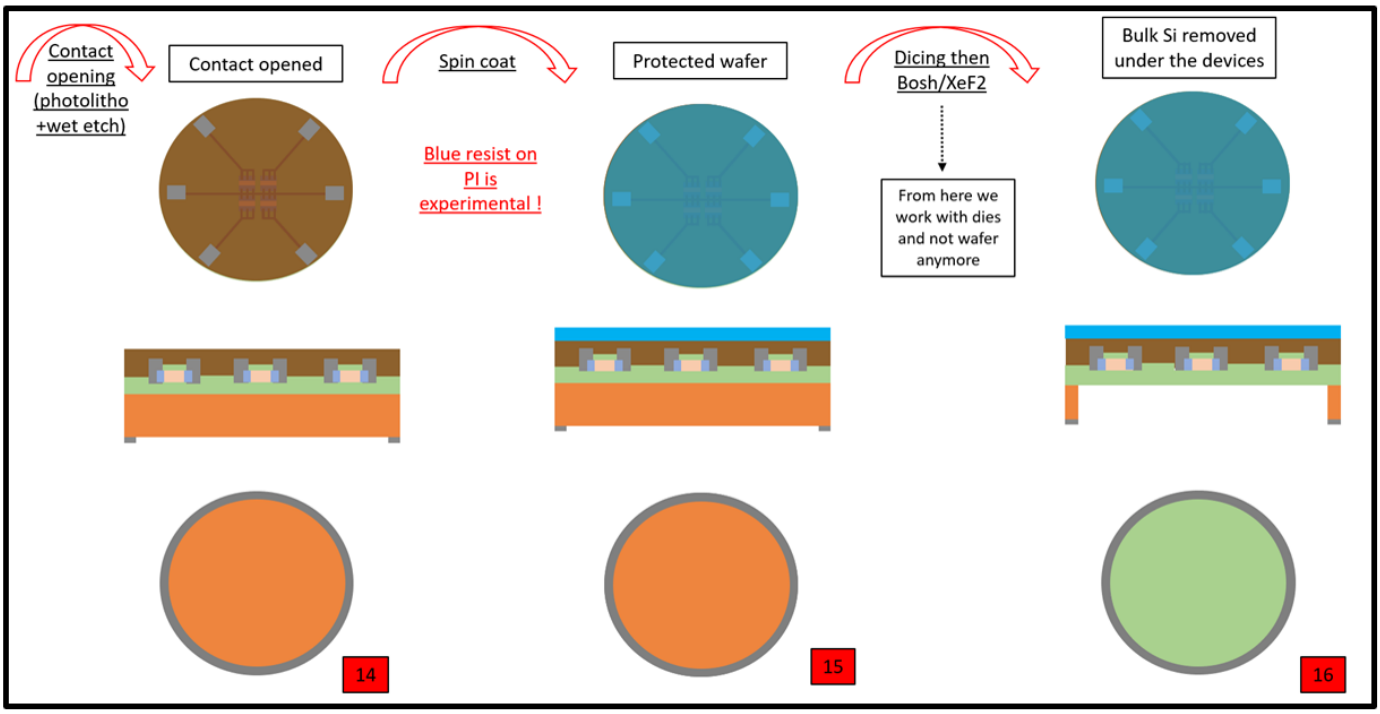


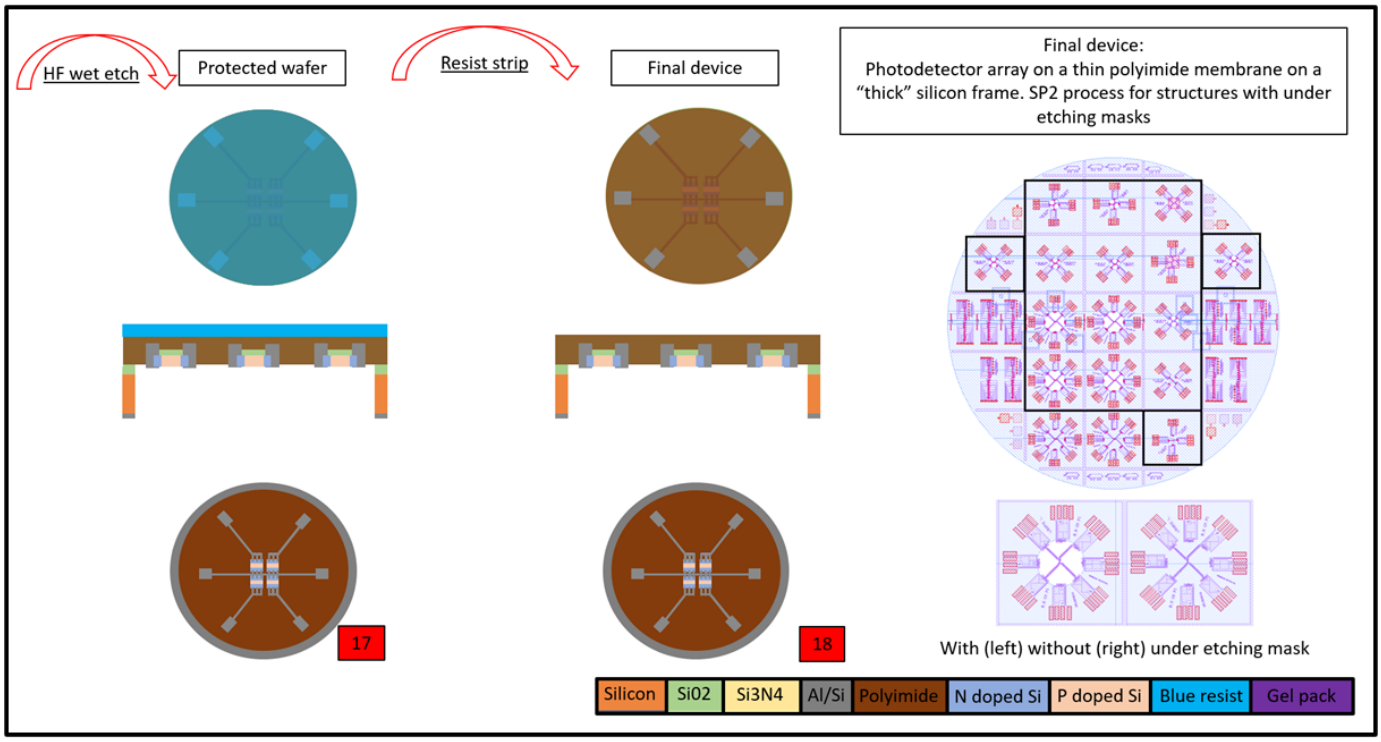



Photolitho + BHF Wet etching
Contact opened in oxide
Photolitho + Metallization (Sputtering + dry etch)
(KLayout design)
Metalized wafer
(KLayout design)
End of the common part of the process.
SP1
SP2
SP3
! Opening done using the metallization mask!
(Al thickness of 0.6 to 1 [µm])
We propose three subprocesses. SP2 is the one we are going for and SP1 will be used to release the bending structure (see SP1)
SP3 process is not considered anymore for the TFE and was not validated !
7
8


Beginning of SP2
Metalized wafer
Back grinding (with front size protection)
(KLayout design)
Thinned wafer
Thermal evaporation of Al (VST)
Back side metallization
800 nm
280 µm
(Al thickness of 0.6 to 1 [µm])
Al thickness 300 [nm]
8
9
10


Spin coating
Protected frontside
Photolitho + Al wet etch
Al hard mask
Resist strip + PI spin coating
PI covered wafer
1-2 µm
2-5 [µm]
5000-1000 [µm]
11
12
13


Contact opening (photolitho +wet etch)
Contact opened
Spin coat
Blue resist on PI is experimental !
Protected wafer
Dicing then Bosh/XeF2
From here we work with dies and not wafer anymore
Bulk Si removed under the devices
14
15
16


HF wet etch
Protected wafer
Resist strip
Final device
Final device:
Photodetector array on a thin polyimide membrane on a "thick" silicon frame. SP2 process for structures with under etching masks
17
18
With (left) without (right) under etching mask
Silicon
Si02
Si3N4
Al/Si
Polyimide
N doped Si
P doped Si
Blue resist
Gel pack

Part V
Conclusion and sequel perspective

After almost 95 pages of excruciating text, equations, diagrams and considerations, here is the end of this thesis. The work started in September 2021 and was able to produce both simulation results as well as practical cleanroom results. Here are the highlights

Without coming back on the state of the art, the first thing that was done was to design a set of photomasks for our later coming process. Six masks ended up being drawn and bought allowing for the manufacturing of released polyimide membranes carrying a set of photodetectors. The ordering of such masks required taking a lot of decisions beforehand.

Among them was the design of the photodetector pixels themselves. Those ended up being a set of PN and NPN junctions doped at a 10e21 and 10e18 atoms/cm^3 for the N and P regions respectively. The decision on the doping levels but also on the device dimensions was done while considering the diffusion length of the charge carriers, the depletion length and their impact on the optoelectrical properties of silicon. Silvaco simulations were performed in order to decide on implantation parameters. Additional care was taken in order to avoid ending up with a Schottky contact between the aluminum contacting pads and the doped silicon. The decision of using silicon being motivated in the state of the art by its high level of performance and RTL compared to other technologies such as graphene or organic semiconductors.

In parallel to the design of the pixels, a deep look on the membrane dimensions and characteristics was taken. This included the modeling of such membranes in Comsol to simulate the impact of the applied pressure and membrane thicknesses and radii. Those simulations, based on the use of polyimide as a flexible substrate due to its high Young's modulus, demonstrated that levels of strains up to half a percent could easily be obtained in 300 nm thick silicon pixels on a 5 µm membranes under a one bar loading. It was also shown that, with the proposed design, neither the membrane nor the silicon pixels should increase their maximum yield strength.

The final design proposes pixels of different types, orientations relative to the crystal and membrane shapes and sizes. 27 devices are featured per wafer. Note that those devices were designed in close collaboration with Nicolas Roisin. Indeed, should the final devices work as intended, his intention is to use them for his PhD.

The processing of the wafers themselves to build the final devices started by designing a microfabrication process from scratch. This process, made out of 85 individual steps was carried out on three homemade SOIs (SOI made by deposition of polycrystalline silicon on a BOX of thermal oxide grown from a bulk silicon wafer) and four real SOIs with a 300 nm thick top layer of monocrystalline silicon. The goal of using those two wafer types being to asses the performances of polycrystalline silicon compared to mono-Si upon the application of strain. The wafers all went through the same process made out of seven individual lithographic steps.

The result at the end of this thesis is the successful fabrication of a set of, theoretically functional devices (in terms of NPN and PN), on SOI silicon wafers. The last step of dicing the wafers and releasing the membrane could indeed not be done due to a lack of time. This is where the disappointment in this thesis resides.

Indeed, a lot of additional work could still be done. For instance, the process should be completed to demonstrate the successful fabrication of polyimide based membranes and after that, the characterisation of the optical performance and effect of strain on the silicon could be performed. The plans for characterising the final released membrane includes putting the structure in the pressuriser equipment designed by Thibault Delhaye and Nicolas Roisin. This device is just what we need since it allows to pressurise membranes up to a bar while putting the structure under illumination. Various wavelengths up to 1.3 µm could be used while playing the applied pressure. The ohmicity of the contact should also be measured after the 432°C baking of the deposited aluminum on the silicon pixels. Finally, putting our rectangular dies into the four point bender to perform Raman measurements is also something yet to be done. This additional work will be attempted during the period left between the hand out of this work and the oral presentation.

All things considered, one could say that a sequel to this thesis could easily be written by considering the potential work left to do. In addition, some considerations were voluntarily left out of the final manuscript in an attempt to not make the page count jump to 200. Hence, the publication of an *Addendum* is at this stage not excluded.

One clear win however of this master thesis, apart from the successful fabrication up to the present point of the micro devices, is the gained training of its writer. Indeed, should the reader consider this thesis a sufficient amount of work for a PhD to start, the writer of this work would begin this endeavour with a solid background and know how on micro and nanofabrication techniques acquired during the month of work in cleanroom.

Part VI
Appendix

1 Fabrication process for the homemade SOI

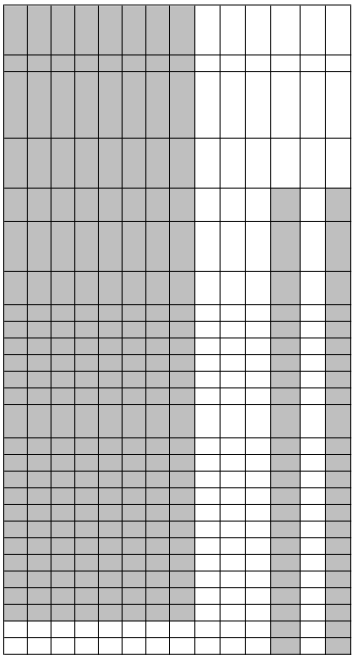
	DIY SOI fabrication from Bulk silicon wafer										
Op.nr	Sequence	Parameter 1	Parameter 2	Parameter 3	Thickness	HM_SOI_1	HM_SOI_2	HM_SOI_3	Price per unit or per hour	Quantity/time	Total
1	Custom	Marquage pointe diamant							0	0	0
2	Mesure de courbure	Dektak							2,2	3	6,5
3	Cleaning standard								33	1	33
4	Wet thermal oxydation	Koyo 1	SiO2		600 n				108	1	108
5	Mesure de courbure	Dektak							2,2	3	6,6
6	Mesure ellipsometrique	Sentech							2,2	3	6,6
7	Cleaning standard								33	1	33
8	LPCVD poly-silicon **frontside**	Koyo 2	Poly-Si	620°C	300 n				108	3	324
9	Mesure de courbure	Dektak							2,2	3	6,6
10	Mesure ellipsometrique	Sentech							2,2	3	6,6
											530

Consommable	Quantity	Price per unit	Total
Bulk Si wafer	3	11	33
Process total cost:	563,38		
Price per wafer:	187,7933333		
Price per device:	NA		

2 Fabrication process for the SWIR detector according to the SP2

Op.nr	Sequence	Parameter 1	Parameter 2	Parameter 3	Thickness	SOI_PI_1	SOI_PI_2	SOI_3	SOI_4	HM_SOI_1	HM_SOI_2	HM_SOI_3	HM_SOI_4	B1	B2	B3	B4	B5	B6	Price per u or h	Quantity/time	Total	
1	Custom	Marquage pointe diamant																		0	0	0	
2	Custom	Pollissage des bords 3" SOI																		0	0	0	
3	Custom	Retirage resine de protection																		0	0	0	
4	Mesure de courbure	Dektak																		2,16	7	15,12	
5	Mesure ellipsometrique	Sentech																		2,16	7	15,12	
6	Standard cleaning	Wet bench																		32,5	1	32,5	
7	Nitrogen Degazing	TEMPRESS	800°C	0h15'																27,1	1	27,06	
8	HMDS adhesion promotion	LPIII																		0	0	0	
9	**Lithography positive frontside (1)**	Strain-Mesa, layer 1, lighfield	AZ MIR701 (+)		0,7 µm															5	8	40	
10	Inspection au microscope	Zeiss Axio microscope																		2,2	8	17,6	
11	Top silicon RIE	Oxford Plasmalab ICP-RIE: Bosh etch	SF6	See notes																32,5	8	259,84	
12	Resist removal	Acetone + methanol+ DI water rince	Wet bench																	0	0	0	
13	Inspection au microscope	Zeiss Axio microscope																		2,2	8	17,6	
14	Mesure de courbure	Dektak																		2,2	8	17,6	
15	Standard cleaning	Wet bench	HF bath																	33	1	33	
16	SiO2 Wet thermal oxydation	Koyo1			27 nm															5,14	8	41,12	Note that we lost some silicon in the process (approx 25 nm)
17	Nitrogen Degazing	TEMPRESS N2																		27	1	27	We skip this steps because we coat directly after the oxidation
18	**Lithography positive frontside (2)**	Strain_DOPING, layer 4, darkfield	AZ MIR701 (+)		0,7 µm															5	8	40	
19	Inspection au microscope	Zeiss Axio microscope																		2,2	8	17,6	
20	Implantation @IBS	Saturation Phosphorus N++	5,00E+15	30 KeV																992	1	992,19	
21	Inspection au microscope	Zeiss Axio microscope																		2,2	7	15,4	Check to see if burnt
22	Resist removal	Barrel	500 W	10 min																5,4	1	5,4	
23	Standard cleaning	Wet bench	NO HF																	33	1	33	
24	Nitrogen Degazing	TEMPRESS	600	0h15'																27	1	27	600°c parce que on ne veut pas trop diffusé nos porteurs
25	HMDS adhesion promotion	LPIII																		0	0	0	
26	**Lithography positive frontside (3)**	Strain-Si-P-dop-M, layer 7, darkfield	AZ MIR701 (+)		0,7 µm															5	8	40	
27	Inspection au microscope	Zeiss Axio microscope																		2,2	8	17,6	
28	Implantation @IBS	Boron P+																		1025	1	1025	
29	Inspection au microscope	Zeiss Axio microscope																		2,2	7	15,4	Check to see if burnt
30	Mesure de courbure	Dektak																		2,2	8	17,6	
31	Resist removal	Barrel	500 w	0h10'																5,4	1	5,4	
32	Standard cleaning	Wet bench	NO HF except B5																	33	1	33	
33	Nitrogen Degazing/HM_RTA	TEMPRESS	850		30															27	1	27	La machine a rta est cassée donc on fait comme ça. Pris en compte dans la simu
34	Depot oxyde PECVD	Oxford Plasmalab PECVD			200 nm															32,5	4	129,92	Test to be done on B5 first to check recipe. It is 32,48€ per pass (2 wafers per pass)
35	Mesure de courbure	Dektak																		2,2	8	17,6	
36	Mesure ellipsometrique	Sentech																		2,16	1	2,16	
37	Standard cleaning	Wet bench																		33	1	33	
38	Nitrogen degasing	TEMPRESS N2	850		15															27	1	27	We do not degas the other wafer to not difuse the dopant
39	SiO2 Wet thermal oxydation	Koyo1		250°C	600 nm															108	1	108	
40	HMDS adhesion promotion	LPIII																				0	Try without nitrogen degasing first. If it does not work add degasing but at 600°C
41	**Lithography Negative frontside (4)**	Strain-ALU LINES, layer 2, lightfield	Negative tone																	5	8	40	For contact opening. Negative tone or image reversal
42	Inspection au microscope	Zeiss Axio microscope																		2,2	8	17,6	Can we inspect all wafers ? The BHF bath is for uncontaminated wafer only. Is using the microscope a to high contamination source ?
43	SiO2 wet etching	HF/ NH4F 1 :7	63 nm/min	4 min																7,58	1	7,58	Contact opening. Etch rate to be measured first on B2 (PECVD) and B3 (thermal oxyde)
44	Mesure ellipsometrique	Sentech																		2,16	2	4,32	We check the etch rate on B2 and B3 on ellipso before moving on to the real wafers
45	Inspection au microscope	Zeiss Axio microscope																				0	We control the under etch
46	Resist removal	barrel	500 W	10 min																5,4	1	5,4	
47	Standard cleaning	Wet bench	NO HF																	33	1	33	
48	Al-Si deposition (sputtering)	VST			600n																	0	Test the deposition on B6
49	HMDS adhesion promotion	LPIII																				0	
50	**Lithography Positive frontside (5)**	Strain-ALU LINES, layer 2, lightfield	AZ MIR701 (+)		0,7 µm															5	8	40	
51	Inspection au microscope	Zeiss Axio microscope																		2,2	8	17,6	
52	Al dry etch	Cl2 corial																		32	7	224	Controle etch rate on B6 Wet etch did not work we went with dry etching in corial (pricing of oxford)
53	Resist removal	Barrel	500W		10															5,4	1	5,4	
54	Inspection au microscope	Zeiss Axio microscope																		2,2	8	17,6	
55	Electric contact measurment	In welcom with prober																					Check electric contact before Al 432 baking
	Beginning SP2 process																						
56	Blue resist spin coating	Suss gamma 80 coater	Frontside																				Protection of the wafer
57	Application of protective plastic film	Front side	125 µm																				Protection of the wafer
58	Back grinding	Back side			300µm (final)																		

59	Rincing	Rince under DI water for 3 min			To remove particle. Frotter avec essui tout de chambre propre
60	Protective plastic peel				Careful fragile wafer
61	Blue resist strip	Rince in acetone, propanol, DI			Propanol after acetone to avoid leaving condensed water traces
62	Thickness measurment				Measure final thickness of grinded wafer
63	Cleaning (non standard)	Fuming nitric acid	10 min clean	10 min rince	Fuming nitric acid does not attack Al
64	Al 432°C baking	Al furnace	432°C 30 mins	H2+N2 flow	For Al to diffuse into silicon and suppress the schottky barriers
65	Blue resist spin coating	Suss gamma 80 coater	Frontside		Protection of the wafer
66	Al deposition	VST	5A/s	300 nm	
67	HMDS application	LPIII	23 min		
68	**Lithgraphy positive backside (6)**	Strain-MEMBRANES, layer 5, darkfield	AZ MIR701 (+)	0.7 µm	
69	Al wet etch	H3PO4	60°C	300 nm/min	
70	Resist strip	barrel	500 W	10 min	
71	Inspection au microscope	Zeiss Axio microscope			
72	Cleaning (non standard)	Fuming nitric acid	10 min clean	10 min rince	Fuming nitric acid does not attack Al
73	Soft degasing	HOT plate 300°C	15 mins		
74	HMDS application	LPIII	23 min		
75	PI2545 spin coating	500 rpm for 5''/2000 rpm for 45''	Manual coater		
76	Soft bake	140°C	3 mins		
77	Resist spin coating	Suss gamma 80 coater	AZ MIR701 (+)	0.7 µm	
78	Soft bake	Suss gamma 80 coater	90°C	90 sec	
79	**Lithgraphy positive frontside (7) (exp only)**	strain-OPEN PASSIV, Layer 6, darkfield	200 mJ/cm^2		
80	PEB	Suss gamma 80 coater	120°C	90 sec	
81	Development in puddle	Manual until PI is removed on pads			
82	Resist strip	Acetone			
83	Hard Bake of PI		350°C		
84	Dicing				
85	Bosh process or XeF2				



References

[1] A. Katiyar, K. Thai, W. Yun, J. Lee, and J. Ahn, "Breaking the absorption limit of si toward swir wavelength range via strain engineering," *Sci Adv*, 2020. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7439440/.

[2] Wikipedia contributors, "Fr-4." https://en.wikipedia.org/wiki/Printed_circuit_board, 2022. [Online; accessed 27-July-2022].

[3] "Nema fr4." https://advancedmaterialscience.com/electrical-insulation/nema-fr4/. [Online; accessed 27-July-2022].

[4] H. Anne Marie, "What is young's modulus?." https://www.thoughtco.com/youngs-modulus-4176297. [Online; accessed 27-July-2022].

[5] B. Cledo and C. Roger W., "Printed circuit techniques." https://nvlpubs.nist.gov/nistpubs/Legacy/circ/nbscircular468.pdf, 1947.

[6] W. Wiiliam S and S. Alberto, *Flexible electronics: materials and application*. Springer, 2009.

[7] Swimbi, "10 features that that the best flex pcb manufacturer have in china." https://swimbi.com/flex-pcb-manufacturer/. [Online; accessed 29-July-2022].

[8] PCBway, "What is flexible pcb." https://www.pcbway.com/pcb_prototype/What_is_Flexible_PCB.html. [Online; accessed 29-July-2022].

[9] G. Peter, "What is flexible pcb: Design, application and manufacturing details." https://sfxpcb.com/what-is-flexible-pcb/. [Online; accessed 29-July-2022].

[10] J. Steve, "Olympus stylus: part of a study in nude photography." https://www.flickr.com/photos/44124348109@N01/2265519. [Online; accessed 29-July-2022].

[11] G. Shoubhik, N. William Taube, L. Leandro, and D. Ravinder, "Ultra-thin chips for high-performance flexible electronics," *Nature*, 2017.

[12] W. Wiiliam S and S. Alberto, *Flexible electronics: materials and application*, pp. 1–2. Springer, 2009.

[13] B. S D, *Introduction to Thin Film Transistors: physics and technology of TFT's*, ch. 5. Springer, 2013.

[14] B. S D, *Introduction to Thin Film Transistors: physics and technology of TFT's*, ch. 1. Springer, 2013.

[15] B. Juergen and G. Martin, "Micro and nanofabrication (mems). week 2: Chemical vapor deposition (cvd)." https://learning.edx.org/course/course-v1:EPFLx+memsX+3T2017/home. Online EDX course on micro and nanofabrication techniques by the EPFL [Online; accessed 30-July-2022].

[16] W. Wiiliam S and S. Alberto, *Flexible electronics: materials and application*, p. 9. Springer, 2009.

[17] W. Wiiliam S and S. Alberto, *Flexible electronics: materials and application*, p. 10. Springer, 2009.

[18] W. Wiiliam S and S. Alberto, *Flexible electronics: materials and application*, pp. 4–8. Springer, 2009.

[19] K. Dae-Hyeong and R. John A, "Stretchable electronics: Materials strategies and devices," *Advanced Material*, 2008.

[20] W. Wiiliam S and S. Alberto, *Flexible electronics: materials and application*, pp. 5–8. Springer, 2009.

[21] A. K. Geim, "Graphene: status and prospects," *science*, vol. 324, no. 5934, pp. 1530–1534, 2009.

[22] B. Deng, Z. Liu, and H. Peng, "Toward mass production of cvd graphene films," *Advanced Materials*, vol. 31, no. 9, p. 1800996, 2019.

[23] S. Franssila, *Introduction to microfabrication*, ch. 20. John Wiley & Sons, 2010.

[24] B. Juergen and G. Martin, "Micro and nanofabrication (mems). week 6: Anisotropic wet etching of silicon in alkaline baths." https://learning.edx.org/course/course-v1:EPFLx+memsX+3T2017/home. Online EDX course on micro and nanofabrication techniques by the EPFL [Online; accessed 31-July-2022].

[25] H. Matthew A, N. William D, and K. Thomas W, "What is the young's modulus of silicon?," *JOURNAL OF MICROELECTROMECHANICAL SYSTEMS*, 2010.

[26] F. Alexander I, W. An-Bang, and C. Henry H, "Thickness dependence of nanofilm elastic modulus," *Applied Physics Letters*, 2009.

[27] S. Franssila, *Introduction to microfabrication*, ch. 4. John Wiley & Sons, 2010.

[28] S. Franssila, *Introduction to microfabrication*, ch. 5. John Wiley & Sons, 2010.

[29] B. achim N, *Ultra-Thin chip technology and applications*, ch. 4. Springer, 2011.

[30] B. achim N, *Ultra-Thin chip technology and applications*, ch. 3. Springer, 2011.

[31] B. achim N, *Ultra-Thin chip technology and applications*, ch. 5. Springer, 2011.

[32] M. JL, C. C, M. P, G. L, M. V.H, F. F, T. S, and R. F, "operando high strain measurment in silicon ultra-thin film on polymer." Poster produced for the Institut nanosciences et cryogénie and the ESRF.

[33] N. Roisin, G. Brunin, G.-M. Rignanese, D. Flandre, and J.-P. Raskin, "Indirect light absorption model for highly strained silicon infrared sensors," *Journal of Applied Physics*, vol. 130, no. 5, p. 055105, 2021.

[34] L. Ju, S. Zhiwei, and M. Evan, "Elastic strain engineering for unprecedented materials properties," *MRS bulletin*, 2014.

[35] B. Vincent and F. Denis, "Elec1330 note de cours." Syllabus on semiconductors provided for the *LELEC1330 Dispositifs electroniques* course given at the UCLouvain.

[36] wikipedia, "Direct and indirect band gaps." https://en.wikipedia.org/wiki/Direct_and_indirect_band_gaps. Images in open access Source given for further general reading.

[37] V. R, B. F, C. D, and V. J, "Thin-film stretchable electronics technology based on meandering interconnections: fabrication and mechanical performance," *IOP science*, 2011. https://iopscience.iop.org/article/10.1088/0960-1317/22/1/015002.

[38] S. S.Z. and C. C., "Selective etching of pdms: Etching technique for application as a positive tone resist," *Elsevier*, 2018.

[39] V. M.-E., T. K., K. K., V. N., T. A., and G. E., "Plasma etching technology for fabrication and surface modification of plastic microfluidic devices," *ispc-conference*.

[40] S. Franssila, *Introduction to microfabrication*. John Wiley & Sons, 2010.

[41] P. Art, "The basics of photodiodes and phototransistors and how to apply them." https://www.digikey.com/en/articles/the-basics-of-photodiodes-and-phototransistors-and-how-to-apply-them, 2022. [Online; accessed 14-August-2022].

[42] "Phototransistors: What are they and how do they work?." https://www.electrical4u.com/phototransistor/, 2020. [Online; accessed 14-August-2022].

[43] "Metallization: Metal semiconductor junction." https://www.halbleiter.org/en/metallization/metal-semiconductor-contact/. Course about semiconductor on the Halbleiter website.

[44] MicroChemicals, "Product bulletin pi-2545 wet etch applications." Obtained on direct demand via email. Contact microchemicals: https://www.hdmicrosystems.com/products/pi-2545-wet-etch.html.

[45] "Material: Polyimide film." https://www.memsnet.org/material/polyimidefilm/. [Online; accessed 14-August-2022].

[46] MicroChemicals, "Technical datasheet az 4500 series thick film photoresist." https://www.microchemicals.com/micro/tds_az_4500_series.pdf.

[47] T. Walewyns, "A low-power low-cost mems-based gas sensing," *Dial*, 2016. PhD thesis at UCLouvain.

[48] K. Ville, *Practical MEMS*. Small Gear Publishing, 2009.

[49] E. William P, B. Fernando, S. James, and P. David, "A new analytical solution for diaphragm deflection and its application to a surface-micromachined pressure sensor," *Sandia Nationnal laboratories*.

[50] B. Juergen and G. Martin, "Micro and nanofabrication (mems)." https://learning.edx.org/course/course-v1:EPFLx+memsX+3T2017/home. Online EDX course on micro and nanofabrication techniques by the EPFL [Online; accessed 30-July-2022].

[51] "Wet-chemical etching of silicon and sio2." https://www.microchemicals.eu/technical_information/silicon_etching.pdf.

[52] W. Kirt R and M. Richard, "Etch rate for micromachining processing." Found on the moodle of the LELEC2560 course given at the UCLouvain.

[53] "Technical datasheet az nlof negative tone photoresist for single layer lift-off." Published by MERCK and found internally at the UCLouvain on the Winfab website.

[54] "Technical datasheet azmir 701 positive tone photoresist." Published by MERCK and found internally at the UCLouvain on the Winfab website.

[55] Buntgarn, "Ellipsometry." https://commons.wikimedia.org/w/index.php?curid=3332636. Published on wikipedia.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**