nanusens
CREATIVE NANODEVICES

The challenges and solutions of building MEMS devices using the BEOL metal layers of a solid-state CMOS semiconductor process

This paper covers the problems faced by current production methods for MEMS sensors, in particular the difficulty of ramping up production to meet the ever-increasing demand for sensors. It then discusses the challenges of the solution of building MEMS using only standard CMOS processes in a fab and how they have been solved by Nanusens using techniques that are now patent pending. The focus of this paper is on linear inertial sensors as a worked example of how these techniques can be used. Further papers will explore their use for other types of sensors that can also be built using CMOS and, crucially, simultaneously on the same die at the same time.

By

Dr Josep Montanyà i Silvestre, CEO Nanusens

Nanusens Ltd
EPIC - White Rock Business Park
Waddeton Close
Paignton TQ4 7RZ - England

Nanusens SLU
Av. del Parc Tecnològic 3 - CENT
Parc Tecnològic del Vallès
08290 Cerdanyola del Vallès · Spain

1. State of the art – Problems to solve

The main problem with MEMS devices nowadays is the need for a bespoke manufacturing process. This has not happened with solid state electronics, which has converged towards a manufacturing standard called CMOS that has many variants and is mainly classified in terms of its node. This is the minimum feature size that the process can resolve at the FEOL (Front End Of Line) which is the first portion of IC Fabrication where the individual devices are patterned.

In practical terms, most of the MEMS ICs found in the market today consist on a package that has two dice inside. One of these dice comes from a CMOS wafer and the other from the MEMS wafer being produced at a bespoke process. The dice inside the package are typically wire-bonded and packaged using a plastic package. For combo ICs requiring several MEMS devices, the package may need to have more than two dice inside -- one CMOS die with the control electronics, and several MEMS dice, each built in a different manufacturing process.

This requirement of a proprietary manufacturing process for each MEMS manufacturer and for each type of MEMS device, has several problems, i.e., cost, size, time-to-market, performance and volume production capability.

Since 90% of the semiconductor industry consists in solid state ICs and most of them are built with CMOS processing, most semiconductor companies use the so-called fabless model and outsource all the production to large CMOS foundries whose business is only producing CMOS wafers. This generates an economy of scale world-wide, typically x100 or more compared to the largest MEMS foundries. That is why the cost of MEMS processing is more expensive than the cost of CMOS wafers.

However, the cost of a MEMS wafer could be less than the cost of a CMOS wafer, especially if we consider lower nodes, but that is because of the much-increased complexity of a CMOS process compared to a MEMS one. But, for the same level of complexity, CMOS would have a much lower cost than any MEMS process. If the MEMS could be built using the same CMOS process, the cost of the whole IC would be drastically reduced. This is because, to start with we would not need two dice, but just one inside the package. By eliminating the MEMS die, we would also simplify the packaging.

There is an interest in reducing the size of the ICs, especially for applications like smartphones and even more with wearables and, in particular, earbuds, where the space is very constrained. The best packaging technique used today to minimize the whole size of the IC is WLCSP (Wafer Level Chip Scale Packaging). This is essentially a deposition of a sealing layer on top of the wafer to protect it, the bumping of the pads, and then dicing the wafer into individual dice, with a previous optional back-grinding. There can be additional steps in the process like RDL (Redistribution Layers), but they are not essential or needed in all the implementations. Importantly, WLCSP cannot be used if there is a need to package more than one die. If the MEMS could be built within the same CMOS die, we would be able to use WLCSP, thus largely reducing the overall package die. And, if it was possible to implement different types of MEMS devices into the same CMOS process, then we could be able to build combo chips packaged in one WLCSP. This would mean even greater space saving compared to the multi-die plastic packages being used today.

When a new type of MEMS device is developed, there is a need to develop a new manufacturing process to build that new MEMS device. Due to the volumes that this process will have to handle, which being for consumer applications will be very large volumes, and the high yield expected in order to minimize cost, this is a complex project that will typically will need several years and involve high costs to set up. If the MEMS could be built using CMOS that already exists, is ready to produce large volumes and at a low cost, then time-to-market would be minimal, because only the device would have to be developed. There would be no need to spend time (and cost) to develop each bespoke manufacturing process.

Due to the different economy of scale existing for CMOS and MEMS processes, the equipment used in CMOS are state-of-the-art, while for MEMS they are usually legacy equipment to reduce the cost of set up. This means that the smallest feature size, also called critical dimension, it is usually smaller with CMOS processes than with MEMS processes. Therefore, if we could build the MEMS structures using the CMOS manufacturing process, we would be able to manufacture MEMS devices with small feature sizes. This would help to improve the device performance, because it would be possible to build softer springs/membranes and also smaller gaps. In addition to this, if we could build the MEMS using CMOS, we would minimize the parasitic capacitance that appears when connecting the MEMS to the electronic interface (typically sensing/driving) circuitry in the CMOS die. This is usually done via wire-bonding inside the plastic package, thus adding capacitances typically in the order of 1pF to 10pF. With the MEMS built beside the other circuitry on the one die just using same CMOS process, this parasitic capacitance coming from the connection of the MEMS to the electronics it would be reduced down to typically 1fF to 10fF. This is a x100 to x1000 reduction. Since the parasitic capacitance reduces the performance of the MEMS device as it masks the capacitance of the MEMS device, by reducing it we can improve the performance of the final MEMS IC. Improving performance in

this case means increasing the sensitivity of the sensor, reducing its power consumption, or a combination of the two.

Finally, as it has been mentioned, mainstream CMOS foundries have volume production capabilities which are more than x100 times larger than the main MEMS foundries. So, if we could build the MEMS devices using CMOS processing, we would benefit from this larger volume production capability. This would allow us to tackle new markets, like IoT (Internet of Things), which otherwise it would be impossible. Nowadays, it is difficult to serve the existing MEMS market by the MEMS providers, because their volume production capability is limited. IoT is expected to increase the current MEMS market volume by a factor of x100 or more but this can only be achieved if the MEMS devices are built using a mainstream CMOS foundry.

Monolithic

Some companies use a monolithic solution to build the MEMS together with the CMOS. This leaves a single die, at the end of the manufacturing process, having both the CMOS and the MEMS. The two options to build this are either bonding the MEMS and the CMOS wafer, after being manufactured separately, or building the MEMS wafer above the finished CMOS wafer, instead of starting with a blanket silicon wafer. In both cases, there is a need for a bespoke MEMS manufacturing process. These monolithic solutions reduce the size of the IC, because there is no need for wirebonding and it is possible to use WLCSP with them. They also improve the performance slightly due to the lower parasitic interconnection capacitance between the MEMS and the CMOS parts, which typically reduces down to between 100fF and 1pF.

However, this approach still has the problems of cost, time-to-market and volume production capability, as they still need a full, bespoke MEMS process. While the size is reduced and performance is increased, it would still be considerably better if we were capable of building the MEMS using the CMOS process because, in terms of size, the profile will always be larger as we will have two dice one on top of the other. But with both components made simultaneously in CMOS, it is one die and this can be back-grinded to be thinner. With respect to performance, this is a x10 increment compared to the conventional two dice package solution, but, when building the MEMS device using the CMOS process, we would get a x100 increment.

And ultimately these monolithic solutions only work if we have a single MEMS device or sensor. If we need a combo chip combining different types of sensors, this cannot be applied anymore. But, building all these MEMS devices using the CMOS process, we keep having a single die solution, that can be backgrinded. So, when we move to combo chips, the cost and size advantages of building the MEMS using the CMOS process increase significantly.

The cost and size reduction, even more when we move to combo chips, comes in part from the reduction of many bonding pads, which are no longer needed.

Building with CMOS

Given the advantages that are known, a number of solutions have been proposed to build the MEMS devices using the CMOS manufacturing process. Initial solutions proposed a modification of the CMOS processing and to add some steps in order to build the MEMS devices. Depending on whether these were performed at the beginning, in the middle or at the end of the CMOS process, the solution was called pre-processing, intra-processing or post-processing.

The modification of the CMOS process was required because MEMS devices need to perform a mechanical movement so they need some empty space inside the IC to perform this movement. And these empty spaces are something that CMOS cannot produce. Another reason for the modification was to add layers of different material, or with different mechanical properties, that were not found in the CMOS process.

Given the very large cost of implementing modern CMOS processes in mainstream foundries and the cost of keeping them stable to keep producing very large volumes while keeping a very high yield, the pre- and intra-process modifications have been abandoned. The only remaining option is CMOS post-processing to implement the MEMS.

CMOS post-processing means that, after completing the manufacturing of the CMOS wafers, they go through a few additional manufacturing steps, where the MEMS is implemented. However, unlike the previously explained monolithic approach, consisting of either wafer bonding or building the MEMS above the CMOS wafer, in this case we simply create the empty spaces needed to allow the mechanical movement of the MEMS. And the MEMS is then built using the materials that exist inside the CMOS wafer.

Although one possibility would be to implement the MEMS using polysilicon, this requires a deep etch to get there, either from the top of the wafer, therefore etching through all the BEOL first, or from the back, requiring a deep etch through the silicon substrate. This requires a complex process that would not be cost effective.

The only solution that remains is to use the materials existing in the BEOL of the CMOS to implement the MEMS. Since the BEOL is the topmost part of the CMOS die, this will require the minimal post-processing and hence the minimal cost.

Different solutions have been proposed for this [3, 6] using combinations of plasma and/or wet etching with HF and other chemicals. These processes are difficult to take into volume production with a high yield, especially when involving wet etching.

The simplest post-processing approach that has been ever proposed consists of using a single vapour HF (vHF) mask-less, post-processing step. The vHF etches away the silicon oxide existing in between the metal layers of the BEOL and it leaves all the metals. This was proposed by Baolab [1,7], and there are papers with research projects on this approach [2]. Also, the designs disclosed in other papers like [4,5] could be released with this approach. Because of its simplicity, it is the lowest cost, CMOS post-processing approach. And it can be implemented in the same CMOS foundry or in the packaging or assembly house.

In this approach, the MEMS device is built using the metal layers, typically Al or AlCu and W, but there could be also others like Cu. And it is possible to trap oxide inside metal casing with the right designs. Other materials may be used, but they must exist in the CMOS BEOL.

Most of the previous approaches use a special package, like a laminate one such as LGA, in order to protect the MEMS. This increases the cost and the size, thus minimizing or eliminating the size and cost advantages, that we would otherwise get building the MEMS using the CMOS process.

Baolab proposed using the top most metal layer to protect the MEMS, while having small holes that would allow the vHF to go inside the MEMS cavity. Later a second set of post-processing steps consisting in Al sputtering and patterning would be applied to properly seal the MEMS device. This is typically a 10% cost adder to the CMOS process.

This simplifies the packaging requirements, and it is no longer needed to use a laminate or other special packaging. Instead, any standard packaging technique, such as QFN, could be used. This reduces cost and size of the final IC.

In addition to the top metal layer, the bottom metal layer was used to complete a metal cavity where the MEMS device is located. This was done to limit the etching of vHF towards the bottom, given that most CMOS processes do have a doped silicon oxide below M1, that is the bottommost metal layer. Doped silicon oxide reacts very aggressively to vHF, quickly increasing the etching speed and leaving very nasty residues that would be difficult to eliminate. This makes the designs portable to most CMOS processes, as otherwise it would only be applicable to special ones not having doped silicon oxide under the bottom most metal layer of the BEOL.

Baolab's solution, like the other solutions using the materials in the BEOL to implement the MEMS devices, surrounds the MEMS device with metallic walls that define the MEMS cavity within the ASIC die. This way the electronics are placed around it. The implementation of these metallic walls is made with a stack-up of metal layers (usually made with aluminium) and vias (usually tungsten). However, the materials could be different, mainly copper, if we go to lower CMOS nodes, below 0.18um processes.

In principle, this is not a straight vertical wall as DRC rules require metal layers to extend beyond the edges of the vias. However, it is possible to make some exceptions to this if we are interested in increasing the lateral area exposed on the wall, for instance, in case of an in-plane capacitive sensor. This would then be a DRV that the foundry would accept.

Using Baolab's solution, vertical metallic walls in principle connect the top and bottom metal planes, thus shorting electrically all the MEMS cavity. Usually, we will not be interested in this, or, at least, not happening everywhere in all the cavity. In order to solve this issue, Baolab used vertical, interleaved, anchor structures. These structures force the vHF to go up and down across the silicon oxide layers until it is exhausted, thus leaving some unetched silicon oxide. This way we get a mechanical consistent wall without electrically shorting the top and bottom metal plates.

One reason why this is particularly effective is that usually the silicon oxide layers deposited between the metal layers of the BEOL of a CMOS process consist on two different sublayers, each having different oxide densities. Therefore, one of these layers etches away with vHF slower than the other. This way, etching the silicon oxide with vHF in the vertical direction is more difficult (it takes more time) than etching horizontally, as then the etching propagates faster along one of the silicon oxide sublayers. With these anchor structures we force the vHF to etch across all the sublayers at a slow etching rate, without being able to propagate quickly through the fast ones.

These interleaved anchors can be used also to add columns or pillars at different locations of the MEMS in order to provide more consistency to the top metal plane. This is particularly important to support the later sealing with Al

sputtering so that the top metal plane is not bent, which would end up destroying or leaving unusable the MEMS device.

The main problem with these anchor walls is that, although they provide mechanical robustness while leaving the top and bottom metal planes electrically disconnected, the electrical capacitance between them is very large. This is because of the large surfaces that are placed close to each other inside the interleaved anchor structure, one connected to the top and the other to the bottom plates, and, even worse, having an important fraction of this filled with silicon oxide.

Another problem related to the previous one is that there is a critical trade-off between this parasitic capacitance between top and bottom metal plates versus production yield and reliability. In order to minimize this parasitic capacitance, we can both minimize the length of the anchor structure, reducing the number of fingers and/or their heights, and/or we can increase the etching time. This way, if we want to minimize this parasitic capacitance, we will have small anchor structures with minimal quantity of silicon oxide remaining inside them after the vHF etching. However, this will be a very weak structure, prone to easy mechanical failure due to mechanical shock, vibration, or simply when sealing or packaging the device. Also, it will result to low yield. Because a slight over-etch will remove completely the silicon oxide inside the anchor structure, producing a collapse of the top and bottom parts and leaving the device completely unusable. In production, we need to avoid this requirement for a critical vHF etching, because it will always lead to a low yield. The reason is that the etching speed and the silicon oxide etched inside the MEMS cavity depends not only on the vHF machine and recipe applied, but also on the CMOS processing. While we can have a tight control on the vHF machine and its recipe, we cannot control the CMOS process, which typically has tolerances in the order of 30%.

In addition to potentially requiring all the metal layers to implement the MEMS devices, thus needing special packaging processes, requiring specific CMOS processes without doped silicon oxide under the bottommost metal plate and the large parasitic capacitances, the two major problems with all the solutions using the materials in the CMOS BEOL to implement the MEMS are yield and reliability. These problems are more critical when we use the Baolab approach, with top and bottom metal planes. However, if we don't use them, then the process becomes more complex and expensive, thus losing the cost advantage and also the volume production, time to market and even the performance advantages.

One major problem found when using the BEOL metals of a CMOS process to implement MEMS devices, is the vertical stress gradient. This is something that is minimized in bespoke MEMS manufacturing processes. However, in a CMOS, since these metal lines where not intended to implement mechanical structures but just electrical connections surrounded by silicon oxide in a solid-state IC, residual stresses are less of a concern, resulting typically in large values. In addition to large residual stress, we usually find a large vertical stress gradient. This results in the metals bending or curling, typically upwards, but, depending on the layers, it can be downwards, especially the top one.

This bending is a big concern when we use top and bottom metal planes. Because then, the vertical gap spacing available above the device is minimal and it can easily touch it. When the MEMS device touches the top or bottom metal planes, it becomes unusable. This leads to very poor yield and reliability.

One possibility to reduce this problem would be to increase this vertical gap distance, reducing the number of metal layers used for the MEMS device itself. However, this would reduce performance in the out-of-plane direction, as the gap would be larger and so the relative capacitance variation for a given sensor for the same displacement would be reduced. Also, in the case of an inertial sensor, we would be forced to a smaller proof mass and not being able to use all the available metal layers, reducing even more the performance.

There is a need therefore to minimize the curvature height of the MEMS device. This is defined as the maximum vertical displacement, in the out-of-plane direction, of any metal layer of the MEMS device.

One known solution to this problem is to stack up two or more metal layers. In this way, we increase the radius of curvature of the resulting metal structure, thus reducing the total curvature height. However, while this is a good solution for some parts of MEMS devices such as the proof mass of a inertial sensor that we want to make as large as possible to improve the sensitivity of the sensor, for other parts like the springs, it leads to very large stiffness, thus significantly reducing the sensitivity. Stiffness is inversely proportional to the third power of the length versus thickness, so increasing the thickness leads very quickly to very stiff springs. This means very low sensitivity for sensors and large driving voltages for actuators. Furthermore, the stack up of many layers is limited by the number of metal layers in the process, and, if there is a need to modify or use a CMOS process with a larger number of metal layers in the BEOL, this will increase quickly its cost.

As a summary, there is a need to find the right designs to implement MEMS devices using the BEOL materials already existing in a standard CMOS process, using vHF to etch away part of the silicon oxide within the MEMS cavity, packaging with WLCSP, and for these devices to have a very high yield, reliability and performance.

A remaining challenge

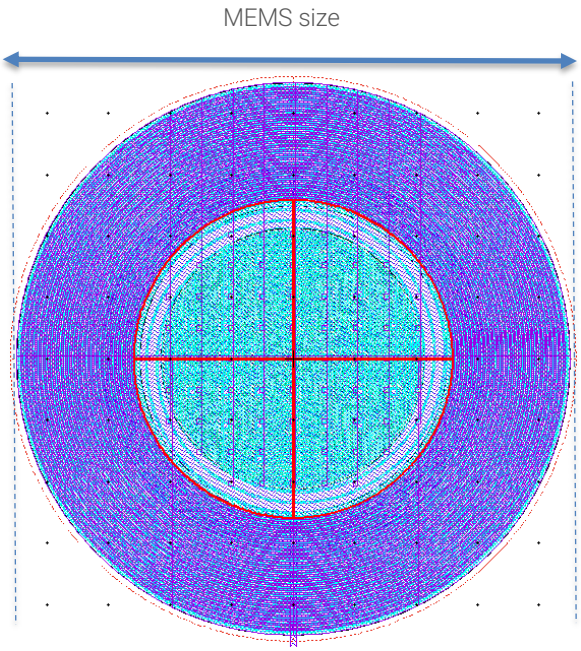
Yet another problem with using a vHF etching, post process step after the CMOS is that the SiN passivation layer deposited and patterned on the top of the CMOS wafer is partially etched away by vHF. This means that, in practice, unless a very short vHF etching step is performed, the SiN passivation layer will be largely or totally etched away. This will leave nasty residues in the wafer and it will expose all the wafer with the ASIC area that should not have its silicon oxide etched away.

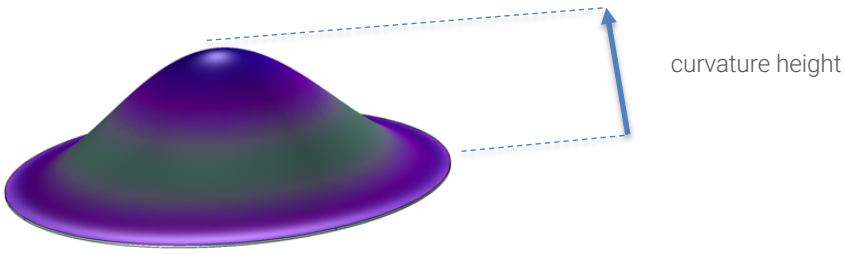
A known solution to prevent this is to increase the silicon content of the passivation, typically measured by the Refractive Index or RI of the layer. Although technically it is not something complex to do, this requires a process tweak and, for large mainstream foundries, it is very challenging for them accept it. Ultimately, this requirement means that we will no longer be able to use a completely standard CMOS process, and hence we will be losing to some extent the advantages of low cost, short time to market and high-volume production capability.

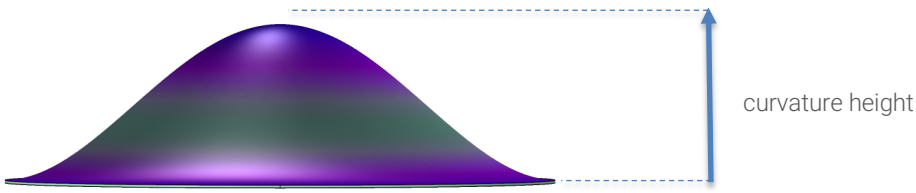
2. Nanusens solution

The Nanusens solution uses vHF etching to etch away part of the silicon oxide in the BEOL of a CMOS process, thus releasing material existing in the BEOL that will constitute the MEMS device. It uses a bottom metal plane and a top metal plane with an array of small holes to allow the vHF to go inside the MEMS cavity.

A key concept of the invention is to limit the overall MEMS size on the layout to between 50µm and 150µm, and preferably to less than 100µm. For a given curvature radius of a MEMS device or element, the total curvature height depends on the horizontal size. Therefore, if the device is small enough, the curvature height will be limited, despite having a large vertical stress gradient.







A second key concept is the design of the springs. In order to have good performance with such small devices, we need small and soft springs, while also keeping a small curvature height. These seem to be contradictory requirements. Short springs means that they will be very stiff, so sensitivity (performance) of the MEMS sensors will be reduced. To have soft springs we will need to minimize their thickness, which means minimizing metal stacking or not using it at all. But this will increase the vertical stress gradient, thus quickly increasing the total curvature height.

The preferred solution to the spring design is to use not one but a set of at least three springs, distributed evenly around the device and rotated around the device's central axis, so that, by symmetry, the MEMS device cannot be tilted after being released with the vHF etching. In the case of an inertial sensor, the MEMS device is typically a central proof mass, made with several metal layers stacked up, so that it is quite flat compared to the springs around it. This proof mass will also be typically larger than the springs in order to have enough sensitivity. If the spring or springs that hold it are curved, then if the proof mass is tilted, it will end up having a large curvature height, despite of the proof mass itself being relatively flat. However, if the springs are located evenly around it, and there are at least 3 of them, then the proof mass will experience a small vertical displacement due to the curvature of the springs, but it will be flat, hence not contributing to the total vertical height with its large size.

In a preferred embodiment, the central mass will have a circular shape, and the springs around it will have a spiral shape, and be made with only one metal layer or a stack-up of just two metal layers. Using circular and rounded shapes for the proof mass, the springs, and, in general, for all or as many parts of the MEMS device as possible, avoids the high mechanical stress that would otherwise be accumulated on the straight angles of the device geometry. These rounded shapes then facilitate the balance of the stress of the springs and leads to the horizontal tilt that we need for the central mass when we use an array of at least 3 evenly distributed springs around it.

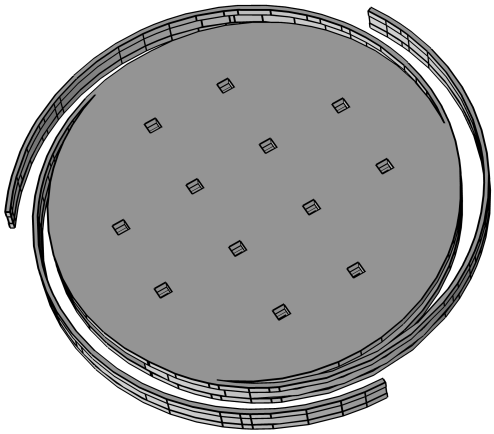


Figure 1. Proof mass with 3 spiral springs evenly distributed around it.

Although the number of springs can be increased, so that we can use 4 or more (and this would further help in achieving the horizontal tilt of the central proof mass), we will usually have only 3 springs, because otherwise the overall stiffness of the device would increase proportionally to the number of springs, thus reducing its sensitivity.

Within this paper, we use the term "inertial sensor" to refer to a variety of devices that sense acceleration. This includes at least an accelerometer, a motion detector and a bone conduction sensor. Their physical principle is the same, but their difference is in the frequencies that they detect. Also, their bandwidth, whether they need to sense DC or not, and also their resolution/sensitivity requirements.

Another embodiment would be to use straight pairs of springs, located at opposite sides of the proof mass, so that each pair of springs lies on the same straight line. This solution works because the residual stress of the metal lines is found to be tensile in most CMOS processes, sometimes with the exception of the top most metal layer. This way the curvature height is minimized. However, this solution leads to a rather high stiffness, which is also highly dependent on the device temperature. Therefore, it is a solution that may be applicable to some MEMS devices, especially when a high mechanical resonant frequency is needed, and when temperature dependence of the spring stiffness is not critical.

If the proof mass (in the case of an inertial sensor) or the central part of the MEMS device attached to the springs, is thicker than the spring (being made of a larger stack-up of metal layers), then a preferred embodiment is for the springs to be made using the upper metal layers, if we connect the central part of the MEMS or proof mass to the top

metal plane. This will minimize the parasitic capacitance of the springs and the supporting external rings towards the lower metal plate.

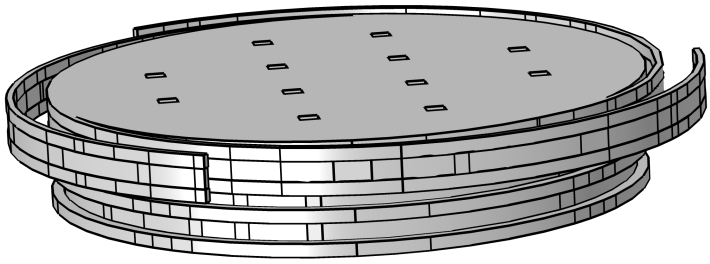


Figure 2. 4 metal layer proof mass with 2 metal layer springs built with the upper 2 metal layers.

A preferred embodiment for an accelerometer, using a 6-metal layer process, where M1 is the bottom layer and M6 is the top one, would be as follows. M1 would be used to implement the bottom plane and M6 the top metal plane, which would be shorted to the proof mass and springs. The proof mass would be implemented using as stack-up of M2 to M5. And the springs would be implemented using either only M5 or a stack-up of M4 and M5.

If we set the total diameter of the proof mass plus the springs, then there is an optimum for the length or maximum angle that the springs rotate. Longer springs (longer total angle) means that they will be softer, but the proof mass will be smaller. Shorter springs will allow for a larger proof mass but the springs will be stiffer. Therefore, there is always an optimum, which will depend on the specific design and process properties. In general, however it is difficult to build reliable devices and with good yield with springs having angles of more than 360°. That is, each individual spring will not complete a turn around the circular proof mass.

Another parameter that affects design yield, performance and reliability is the spacing around the springs. That is, the horizontal spacing or gap existing at each point of the springs without any other spring, proof mass, anchor or any part of the device (except at the edges of the spring, as they start and end up being fused with the outer ring or anchor/wall and the proof mass or other MEMS part). This obviously needs to be at least equal to the minimum spacing set by the process DRC. However, in practice, we will set it to a higher value, between x2 and x10 of this minimum spacing set by the DRC of the process. A preferred value is x5. For instance, with a 180nm process having a minimum gap DRC between metals of 300nm, we will preferably set this horizontal spacing for the springs as 1.5µm. The trade-off here is that, for very large spacing, we will reduce sensitivity/performance, because we will reduce more the proof mass area and hence mass for the same spring length (hence to achieve a given softness/stiffness for the springs). But reducing this horizontal spacing between the springs, will lead to poor yield and reliability.

Another key aspect of the springs, which is especially important when they are soft such as with the case when we use a set of 3 or more spiral springs evenly distributed around a central proof mass, is to make them short enough compared to the displacement (both vertical, so out-of-plane, and horizontal, so in-plane). Since we have small proof mass and we want very soft springs in order to achieve enough sensitivity, this would, in principle, be prone to stiction issues, which would lead to very poor reliability for these MEMS devices.

MEMS devices are designed to operate in the linear region. This is because they have large dimensions with large springs to make them soft enough despite their thickness and the maximum displacement they can have, covering all the gap they have above, below, in front of or on a side, is very small. This way the MEMS spring obeys Hook's law, having a constant stiffness, that is generating a mechanical restoring force that is proportional with the displacement.

In our case, given that the displacement is large compared to the spring length, the mechanical restoring force starts being proportional to the displacement, but, after a given initial displacement, it is no longer lineal and it increases faster. This way, although the springs are soft for small displacements around the equilibrium point, which is where the sensor will operate thus having very good sensitivity, in the case that the proof mass experiences a larger displacement (for instance, if submitted to a shock or strong vibration, touching the surrounding walls, roof or floor), the mechanical restoring force would be much larger at that point. This way the MEMS devices go back to the equilibrium position and detach from the stiction forces thanks to this increased mechanical restoring force at the point of contact.

In a more detailed view of this phenomenon, all springs are non-linear. However, while other MEMS devices experience only small displacements, ours experience larger displacements so that they enter into the non-linear region of the mechanical restoring force versus displacement before touching the surrounding walls, roof or floor.

We can count the length of the spring in two ways. One is the straight distance from one end to the other. The second measure is the whole distance along all the spring, following its meanders and curves. We will consider the longest of these measurements to be the "length" of the spring. In our designs, the shortest ratio between the minimal displacement that can lead to the spring touching a surrounding wall, roof or floor, and the length of any spring, is at least 1%, while in some designs it is 5% or even 10%. This principle can also work for lower ratios, but then the robustness is not so good. However, for 0.5% or even 0.1% it might still give good results, depending on the particular process and overall design.

Such a short ratio is key to our devices and not found in other MEMS designs. It allows us to have soft springs and short gaps for sensitive, high performance devices and able to be packaged with all packaging technologies, including WLCSP, while at the same time having high yield and reliability.

A third key concept is the design of the vertical walls, or more precisely, the definition of the MEMS area or the limit of the lateral etching of the silicon oxide, and the mechanical anchors or supports of the MEMS. As explained before, other designs use vertical metal walls or anchors.

On the first case, using vertical metallic walls, we cannot use a top metal layer and/or a bottom metal layer to seal the device from the top and/or bottom. This means that we will need a special CMOS process without doped silicon below the lowest metal layer of the BEOL, and/or a special, more expensive packaging technique like laminate substrate to properly protect the MEMS cavity from the top, and usually also a more expensive, post processing etching sequence. This may be avoided totally or partially if the MEMS device can have the top and bottom metal planes electrically shorted with the surrounding walls, which usually would not be possible.

The second option, using anchors, although it disconnects electrically the top and bottom metal layers, creates a large parasitic electrical capacitance between them that degrades the device performance. Furthermore, any attempt to improve the performance by reducing this parasitic capacitance with a reduction of the anchor structure or increasing the vHF etching time, reduces the yield and reliability of the device.

Our devices do not have the vertical metal walls connecting the top and bottom metal planes, nor the capacitive anchors. Instead, there are two different solutions that we use.

One is to extend the parts of the MEMS device located in between the top and bottom metal planes or electrodes, typically the springs, but it can also be other electrodes so that they end up buried into the silicon oxide, so far away in the horizontal direction that the vHF does not reach there. In practice, we have seen that there is no need for a long distance. For a typical 180nm CMOS node, it is enough with 20µm of metal around the MEMS device. So, in other words, there is metal placed around the MEMS device that we release, which holds it because the outer edge of it has silicon oxide around that it is not etched way. Preferably, this surrounding metal has a circular shape on its outer edge, but many other shapes would be possible.

In the previous solution, we had at least 3 electrical disconnected parts, where the top metal plane, the bottom metal plane and the part of the MEMS device in between them, are electrically disconnected, and more parts could be made electrically disconnected. Another possibility can be applied when, if instead of having 3 or more electrically disconnected parts, we had only 2. In this case, we can attach the part of the MEMS device in between the top and bottom metal planes to one of them by means of a vertical metal wall, but not to the other. Preferably we will connect it to the top one. This is because the top metal layer is usually less flat and more curved than the bottom metal layer. This is because the bottom metal layer is not detached from the underlying silicon oxide.

In order to increase the mechanical consistency of these external rings, we can use an array of vias to join them. To make them even more robust, we can use, instead of the regular square via array, an array of concentric via rings like those used inside the proof mass. However, in this case, there will be no holes going through the proof mass, and hence the rings will not have to be disrupted and they can be continuous.

Another important design aspect that can be applied for these last two options, is that, since we do not use neither vertical metal walls shorting the top and bottom metal planes nor we join them with capacitive anchors, we can then reduce the size of the bottom metal plane compared to the size of the top metal plane that will be larger. The reason for this is that the lateral over-etching will be larger for the top metal than for the bottom metal layer. The lateral over-etching is the distance that goes from the outermost edge of the MEMS devices that we need to release, to the outermost location where there is silicon oxide etched after the vHF post processing step. That is, since we do not use metal vertical walls shorting from top to bottom nor capacitive anchors, because, instead, we surround the MEMS device by metal area that extends into the silicon oxide around so part of this metal area will have its surrounding silicon oxide etched away during the vHF step while beyond some point it will no longer be etched.

With this approach, we reduce the parasitic capacitance between the top and bottom metal planes, which is also the parasitic capacitance between the top or bottom metal planes and the moveable part of the MEMS device in the case of using the second approach discussed above when we short the part between the top and bottom metal planes with one of these planes. This reduction of the parasitic capacitance results in improving the device sensitivity or performance.

The size for this reduction of the width of the external ring at the lower metal plate compared to the top one, it will depend on the CMOS process and overall design, but it will be typically between 10% and 50% with a preferred value of 30%. A preferred width for the outer ring at the top metal plate is 20µm, which means that the external ring width at the bottom metal plate will have a preferred size of 6µm. If the central disc (proof mass plus springs) has a diameter of 50µm, then the total size of the top plate would have a diameter of 90µm and the diameter for the whole bottom plate would have a diameter of 62µm.

If we anchor the part in between the top and bottom metal planes with metal areas extending around this part that end up buried into the unetched silicon oxide, then we can also make them with a size that is smaller than the top metal plane but larger than the bottom metal plane. Because, in this case, its over-etching it will be in between the top and bottom metal planes.

The external ring for this middle plate will be typically between 30% and 70% the size of the ring for the top metal plane, being the preferred value of 50%. However, it will ultimately depend on the specific CMOS process and the overall design.

If the top metal plane has a circular shape, it consists on an internal disc adjacent to an external ring surrounding it. The internal disc has an array of holes in it to allow the vHF to go inside the MEMS cavity, while the external ring is solid (with the possible exception explained below for building a trench to isolate most of this external ring electrically).

The extension of the internal disc is, in principle, the extension of the MEMS that we want to release with the vHF. However, another innovation is to reduce the extension of the inner disc, thus not placing release holes around the outer part of the MEMS that needs to be released. Since vHF can travel a relatively long distance, all the MEMS will be released and we will minimize the over-etching on the external ring, thus being able to reduce the size of this external ring. This will reduce the parasitic capacitance between the top and bottom metal plates, thus improving the performance of the MEMS. In case of using this reduced extension for the release holes, the passivation opening can also be reduced, since we only need to open it above the area having the array of holes.

The reduction of the internal disc that we can do will depend on the CMOS process, but typically it will be between 2µm and 20µm on each side, with a preferred value of 6µm. That is the disc diameter will be reduced between 4µm and 40µm, with a preferred diameter reduction of 12µm.

Doing this reduction of the internal disc, we can reduce the external ring with the same value at all the metal layers that have such external rings.

For clarification, although we talk about internal disc and external ring, in practice, the layout of the top metal layer will be a single disc. Then the etching holes array will be located in the centre, covering an area that is defined as the internal disc, and the surrounding solid area without etching holes, we call the external ring.

The above descriptions are valid also in the case of having several electrically disconnected parts in between the top and bottom metal layers. In this case, each one will have its own metal extensions buried into the silicon oxide, and they will all be electrically disconnected between them, although there will always be some electrical parasitic capacitances.

Although the preferred approach is to have this exterior metal area to support the MEMS parts in between the top and bottom metal planes, surrounding all the released MEMS inside to provide better mechanical consistency, it is not strictly necessary. This can be useful especially in the case discussed above where there are two or more electrically disconnected MEMS parts in between the top and bottom metal layers. One example of this is for an in-plane inertial sensor where we could also place several lateral electrodes to sense acceleration in different directions.

Another variant that reduces the parasitic electrical capacitance between the top metal plane (and also the middle part of the MEMS device if electrically shorted to it by means of, for instance, a vertical metal connection to it) and the bottom metal plane, is to add a very short trench around all the top metal plane at a certain distance from the MEMS that needs to be released. In principle, this trench should be located at half of the over-etching distance. This means at about 10µm, as the overall length of this metal area around the MEMS is about 20µm. But the distance can be made shorter, down to 5µm or even less, down to almost zero. Preferably it will be located at a distance between 5µm and 15µm. And the extension of the top metal plate will be a distance of about 20µm, but it could be between

5μm and 30μm, depending on the specific CMOS process properties, and the overall MEMS design and the required vHF properties and recipe.

The width of this trench needs to be minimal, as allowed by the process. This will be typically 0.8μm, but, in general, it will fall between 0.5μm and 2μm, depending on the process and, in particular, the thickness of the top metal layer. The need for a metal ring beyond this trench is the reason why we talk about a circular narrow trench around the MEMS, rather than just saying that we shorten the outer diameter of the outer ring, to sustain the passivation. With this trench, we will divide the outer ring into two parts, one inside the other, that will be electrically disconnected, and mechanical disconnected as well. Although there will be some parasitic electrical capacitance between them and also they will ultimately be connected to the silicon oxide and hence they will not move relative one to the other. Therefore, it could be questioned why we need to keep the outside part of this divided outer ring. The reason is that there will be over-etching during the vHF post processing step, so that the silicon oxide located between the passivation and this top metal layer will be etched away, thus leaving the passivation very fragile. For this reason, it is better to keep the outermost metal ring, in case the passivation breaks so that it can be supported. However, depending on the process properties and overall design, it might be possible just to remove this external part of the outer ring, and, instead of building a trench for the outer ring, just reduce its diameter. This would reduce the parasitic capacitances even more.

The preferred embodiment will have a short vertical metal wall surrounding the MEMS device and connected to the top metal plane. This vertical metal wall can be connected to a moving part of the MEMS located in between the top and the bottom metal planes, like the spring anchors. The purpose of this short (i.e., not going down to the bottom metal plane) wall is to block the vHF to etch horizontally under the top metal plane towards the outer edges of it, forcing the vHF to go down the vertical wall first and then back upwards to be able to etch under the top metal plane towards the outer edges. Depending on the design, this short vertical metal wall can also provide mechanical consistency and/or electrical connection to other parts of the MEMS device, like, for instance, the anchors of the springs.

Instead of using the capacitive anchors, another possibility to achieve a mechanical connection, without electrically shorting two parts of the MEMS, is to use a MIM layer in the MEMS process. This is usually not etched away with vHF, or at least etched slowly, although it depends on the specific CMOS process. This provides a more compact solution than the capacitive anchors. However, the electrical capacitance tends to be larger and the mechanical robustness is usually not very good. However, it might still be useful in some designs, depending on the MEMS device, process and overall design.

For some designs, it may also be useful to use horizontal capacitive anchors instead of vertical capacitive anchors. It is also possible to use a mixed design and to implement feedthroughs to pass connections across MEMS metal walls or planes using the same design principle.

The array of holes at the top metal plane will be as small as possible. Typically, they will be smaller than what is allowed by the DRC of the process but big enough to make sure that they are open through all the top metal thickness. This minimum size will depend on the specific CMOS process and particularly on the top metal thickness. A typical value that works very well is 0.8μm. Below this, it is usually difficult for them to open completely, which would lead to a low yield in production. Larger values may not be properly filled when we apply the sealing layer as explained later. So there is a trade-off and we cannot have neither too small holes that would not open when patterning the top metal layer during the CMOS process, nor too big that would not be properly sealed later during the packaging.

That is why the typical hole size will be between 0.5μm and 1.5μm, with a preferred value of 0.8μm. However, depending on the CMOS process, the top metal thickness, and the sealing material, thickness and process being used, the etching hole size may vary.

Given that the holes are so small, they will be drawn as square holes, because, in reality, any other shape would not make any difference, as we will be forcing the resolution of the process, and they will be partially rounded during the manufacturing of the device.

The separation between the holes on the top metal plane will need to be ideally similar to the vertical length of the vertical distance from the top to the bottom metal layers. In reality, it could be argued that we could space the etching holes horizontally on the top metal layer up to a distance of at least twice this vertical distance. And it could be even more, given that vHF etches slowly in the vertical direction compared to the horizontal direction due to the multiple oxide sublayers with different densities and etching speeds. Because the goal is to make sure that we etch properly all the volume inside the MEMS cavity, so for this we need to place the holes close enough, but at the same time we want to separate them as much as possible otherwise we would have a weak top metal plane with so many holes and little metal remaining that, later, would not be able to withstand the sealing on top of it when we package the device as explained further on.

We have found experimentally that a value that works well is to have the etching holes spaced a distance between 50% and 200% of the height of the metal stack. This height is counted from the lowest point of the bottom metal layer up to the highest point of the top metal layer. A preferred value is to separate the holes a distance equal to this height (so 100%). The hole distance is measured from the centre of one hole to the centre of another hole in both the horizontal (X) and vertical (Y) directions.

In order to allow the vHF to go down to the lowest level of silicon oxide so that all the silicon oxide that needs to be removed is properly etched in all the cavity, we will need to place the same array of holes going through all the MEMS device inside the cavity. This might be shifted laterally with respect to the holes on the top metal plane, although the preferred embodiment will be just to place them at the same locations. If these holes go through structures that have silicon oxide trapped inside, like the proof mass, we will need to surround these holes with via walls to avoid the vHF going inside through these holes and etching away the silicon oxide that we want to keep unetched. Given the small size of these holes, that will be made preferably with square shapes and we will implement these via barriers as square rings.

A fourth key concept is the usage of the sealing layer existing in the WLCSP process, also referred to as repassivation, typically made with Polyimide (PI) but it could also be Benzocyclobuten (BCB) or others, to seal the MEMS cavity. This avoids the need for a specific aluminium sputtering and patterning process, reducing the complexity and cost of the post-processing, which is further reduced to the vHF etching and post backing. In addition to this, using PI or BCB has been seen to offer a better seal as it provides a better covering of the array of holes on the top metal layer. In contrast, aluminium sputtering requires a very thick deposition and, even then, due to the conformality of the deposition, some holes may not be properly sealed. This does not happen with PI, which seals all the holes very well.

Should we want to use another type of package that is not WLCSP, we could still apply a PI or BCB or other coating and patterning (even aluminium sputtering, although it would not be ideal, but it could be done with enough thickness and the proper set of parameters), and then proceed with whichever packaging process.

Another key concept is not using metal filing structures within the MEMS cavity. In order to compensate for the metal residual stress, CMOS designs need to have a constant metal density across all the area of the ASIC. In order to achieve this, once the ASIC design is finalized, an automatic process called “metal filling” is performed, which fills all the empty areas with random small metal shapes in order to achieve the required target metal density. This metal filing must not be performed within the MEMS cavity. Otherwise, after applying the vHF, all these tiny metal filing structures would be released and they would attach by stiction to the MEMS devices preventing it from working properly, or not letting it to work at all.

All the explanations given in this paper can be applied to different CMOS nodes, to different metal stacks and even different solid-state semiconductor processes. Also, when we talk about the top metal layer and the bottom metal layer, this are usually the top most and the bottom most metal layers in the process layer stack. However, it may be applicable to other metal layers. In the case of building an inertial sensor in a 6-metal layer process, we will usually need all the metal layers available in order to maximize the thickness and hence the mass of the sensor proof mass. However, if the process has more metal layers available or if we build another type of MEMS device, or even for a inertial sensor if we can manage to get the required specifications, we may not need to use all the metal layers available in the metal stack. In this case, we will preferably use those located on the top, thus leaving the metal layers located at the bottom to be used for the ASIC to make electrical connections into the ASIC. In this case, there would not be any dedicated area to implement the MEMS, but instead the MEMS would be implemented above the ASIC.

Actually, in all cases the active area (FEOL) below the MEMS can be used to implement the ASIC. However, if there are no metal layers available to be used for the connections, because they are all used to implement the MEMS, it will be difficult to implement a useful part of the ASIC below the MEMS. Still depending on the process and the specific ASIC design, it may be useful to implement large transistors or other circuitry requiring little wiring, and/or polysilicon lines if available may be used for this wiring.

When not all the metal layers of the process stack are used to implement the MEMS, then all the explanations of this paper must be understood in the following way. The “top” and the “bottom” metal layers are not then the top most and bottom most of the metal stack, but they are then the top most and the bottom most of the metal layers used to implement the MEMS device.

Although the preferred embodiment is using circular and round shapes, the disclosed structures can be applied to other types of shapes.

Still another key concept is the sensing electronics to interface the MEMS, when they are capacitive MEMS sensors, like, for instance, accelerometers, bone conduction sensors, motion detectors, ultra-sound sensors or any other capacitive sensor. Our MEMS capacitive sensors have a uniquely small capacitance. This is because of the small size and also the minimal parasitic capacitance, thanks to the several inventions explained previously. And also, because

of the proximity of the ASIC attached to the MEMS edge, there is no need to wire the MEMS to another die where the ASIC would be located, nor connect to the top of the wafer where the ASIC would be placed in a wafer bonding scheme nor connect in the case of building the MEMS above the ASIC CMOS wafer.

The capacitance of our MEMS sensors is typically in the order of 10fF to 100fF, so about 50fF. This is about 100 times smaller than commercial MEMS devices for consumer electronics. This enables us to implement a completely different sensing scheme that would not be feasible with other MEMS devices because it would require too much power consumption.

The sensing of the MEMS capacitance is done by means of building a ring oscillator where at least one of the capacitances of the loop is our MEMS device. This ring oscillator will feed a counter that will be read and reset every sample period. The output of the counter will already be digital and it will give us the value of the capacitance.

This approach has many advantages. First of all, it simplifies the analog design as it is all digital with the only exception being the ring oscillator. This means that there are many analog blocks that we would otherwise need that we avoid here, like transconductance amplifier, programable gain amplifier, A/D converter, analog filters, chopper and capacitance mismatch compensation, among others. This simplification has many advantages: smaller ASIC area so lower cost in production, reduced design time so faster time to market and lower development cost, easy porting to other CMOS nodes and processes, and lower power consumption.

The lower power consumption comes from the fact that we avoid so many analog blocks that would be power hungry. In exchange of these blocks, however, we will need to be charging and discharging continuously our MEMS sensor capacitance at a very high frequency, typically between 10MHz to 100MHz, but depending on the MEMS design, CMOS process, and target specs for the sensor. This would consume too much power for the usual capacitances in the order of several pF. But as our MEMS sensors have capacitances x100 times lower, this will not imply more but actually less power consumption, thus making our sensors very power efficient, in addition to the other advantages mentioned above for this sensing scheme.

This ring oscillator will vary its frequency depending on many factors like supply voltage and its noise, temperature, and also process variations. To compensate for this, we have a second ring oscillator that uses another MEMS device built very close to the first one, so that it will see almost the same process, voltage and temperature variations. This second MEMS device (or devices if we use more than one MEMS into the ring oscillator loop) will be slightly different with a stiffer spring. Preferably this will be made using a wider and/or thicker (using more metal stack-up) spring. This way, the capacitance of this second sensor will move very little due to the magnitude that the sensor measures (like acceleration in the case of an accelerometer), but it will change in the same way as the first one due to all the other factors, like supply voltage, process and temperature variations.

We let the two ring oscillators actuate a different counter each until the second one reaches a predetermined value. In that case, we will read the first counter, which will give us the value of the magnitude being sensed, and then we will reset the two counters and start counting again. This predetermined value will usually be programmable so that we can define different sampling frequencies. When the sampling frequency is low, the ring oscillators and/or the counters will be disabled, thus minimizing power consumption.

A completely different solution would be connecting electrically, in parallel, an array of MEMS devices. Each device would need to be completely isolated from the others, each one with its own passivation window. And they would only be connected by the corresponding tracks or lines that would make the electrical connections. In this way we can increase the total or absolute capacitance of the MEMS, while keeping the same relative capacitance variation for a given external stimulus to be sensed (like acceleration in the case of the accelerometer). This would allow the usage of conventional sensing electronics used with capacitive MEMS sensors.

In order to increase the proof mass but without increasing its size, we build metal walls around all the perimeter of the MEMS device. This way silicon oxide is trapped inside the proof mass, and it is not etched away by the vHF. Furthermore, we fill the proof mass with plenty of vias, because these are made with tungsten, which has a higher density than silicon oxide and aluminium, which is the material of the metal layers. In order to further increase the effective density and hence the total mass of the proof mass, we can use larger and closer via arrays than what is allowed by the DRC of the process. In the case of a circular shaped proof mass, we can use also concentric via rings, spaced with the same distance than the thickness of the rings, preferably making this distance and rings width equal to the via size and via spacing defined by the CMOS process DRCs. It happens that, although the vias of a CMOS process typically need to be squares of a fixed size, in practice, we can extend these vias in one dimension, provided we keep the specified via size in the other dimension. Otherwise, the wafer would not be properly manufactured. Since we will need to make holes across the proof mass, these circular rings may have to be interrupted around the holes. Other shapes are possible for both the proof mass and the rings or via filing structures inside it.

Another key innovation is the modification of the pads. This is because there will be passivation openings not only above our MEMS devices, but also above each of the pads. That is actually the reason why there is a passivation opening in the CMOS process. This means that when we apply the vHF post processing, the oxide under the passivation (that is between the passivation and the top metal layer) will be etched away. If the top metal layer at the pad is not large enough, the silicon etching will go beyond it and etch under the passivation without metal underneath. If this happens, a lot of silicon oxide around the pad will be etched away and the top passivation will end up having no oxide below. As a consequence, the passivation can break and also a lot of silicon oxide can be etched away, ruining part of the ASIC electronic circuitry. A solution to solve this is to extend at least the top metal layer (we can extend more or all the other metal layers to provide better consistency) more than with a conventional pad design. This extension will depend on the details of the specific process and vHF etching that we will apply. Usually, the metal will have an extension of between 15µm and 25µm beyond the passivation opening in all directions and preferably this extension will be 20µm. There is no need to use rounded shapes, so usually the pad will keep having a square design for the passivation opening and so also the metals that define it. However other shapes are also possible.

Most of the innovations described in this paper can be applied to many different devices, including but not limited to inertial sensors, gyroscopes, pressure sensors, ultra-sound sensors and transducers like CMUTs, loudspeakers, magnetometers and compasses, microphones, RF switches, tuneable capacitors, RF inductors, temperature sensors, and many more.

In order to avoid requiring the CMOS foundry to increase the silicon content of the passivation, we can use a special recipe and equipment developed by memsstar Ltd (Scotland) that is patent protected. After the vHF etching step, we bake the wafer, in order to sublimate the fluorine residues.

Thanks to the smaller size and cost, and higher performance of the MEMS devices disclosed in this paper, plus their high-volume production capability and short time to market, it is possible to build smaller and higher performance smartphones, wearables and earbuds, having more functionality and longer life and autonomy thanks to having more space for a larger battery. These nanosensors are also an enabler for many Internet of Things (IoT) applications where there is a requirement for ultra-low cost, very low power, high performance, very small sensors, produced in very large volumes.

3. Physical designs

This is the preferred embodiment for an out-of-plane inertial sensor. It is built using a 6-metal BEOL CMOS process. Metal layers are numbered M1 (bottom) to M6 (top). There are 4 via layers, numbered V1 (between M1 and M2) up to V5 (between M5 and M6).

The proof mass is circular, has 3 spiral springs evenly placed around it, and the diameter of the proof mass plus the springs is 50µm. The proof mass is made with a stack-up of metal layers M2 to M5, and the springs stacking up metal layers M4 and M5. The proof mass has concentric rings, stopped at the etching holes that go vertically across it, made with via layers V2 to V4.

The top metal plane has the same diameter plus an external ring of 20µm width around it, which is equivalent of a circle having 90µm of diameter. There is an array of 0.8µm holes spaced approximately 5µm between the centre of each two holes in both X and Y directions.

The springs have a surrounding ring of 20µm width around them, built on the same layers M4 and M5. There is an array of concentric vias joining these rings and the top metal plane on these rings and they are implemented at layers V4 and V5.

There is a circular bottom metal plane made with M1 with a diameter of also 50µm plus an exterior ring of 6µm width, so smaller than the external rings of the top layers for the top metal plane and the support of the springs. Together the internal circle and the external ring are equivalent to a circle of 62µm diameter. The difference between the internal circle and the external ring comes from the fact that, after the vHF etching, the external ring is partially etched and the edges keep buried (and subjected) to the silicon oxide.

The top passivation is opened with a circular shape of 50µm that is above the MEMS device.

The following figure shows the complete layout, with all the BEOL layers shown in different colours.

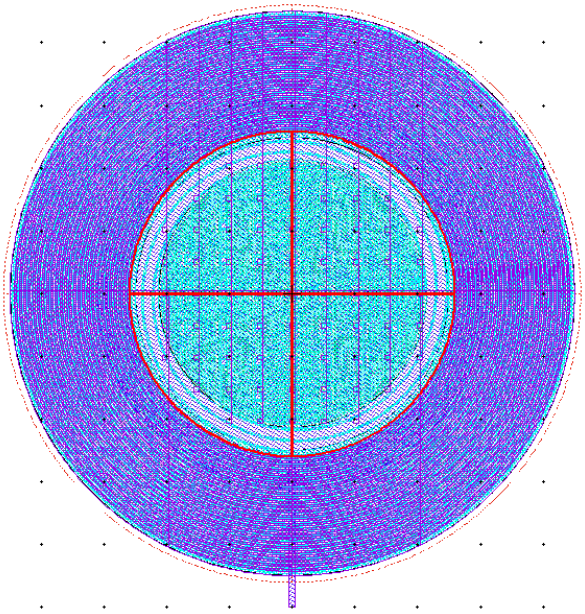


Figure 3. Complete layout for out-of-plane inertial sensor.

The following figures show the layout for each of the individual layers. Layer V1 is not shown because it is empty.

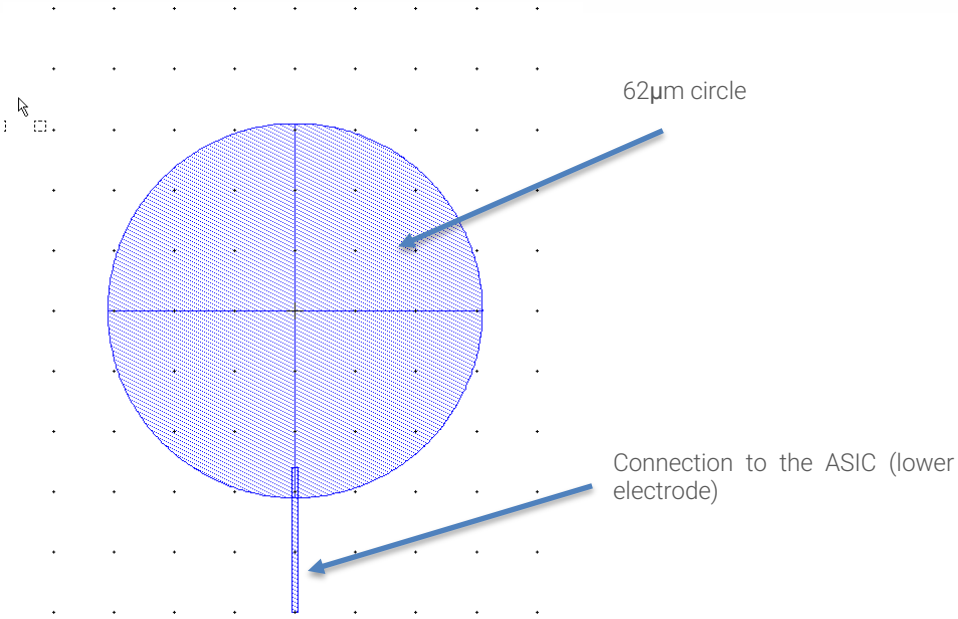


Figure 4. M1 layer.

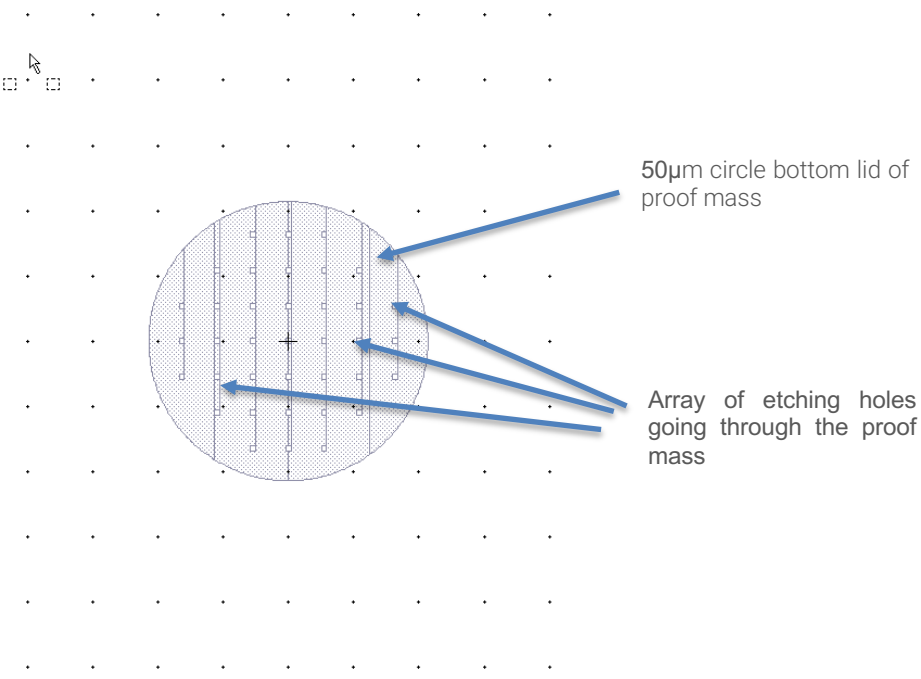


Figure 5. M2 layer.

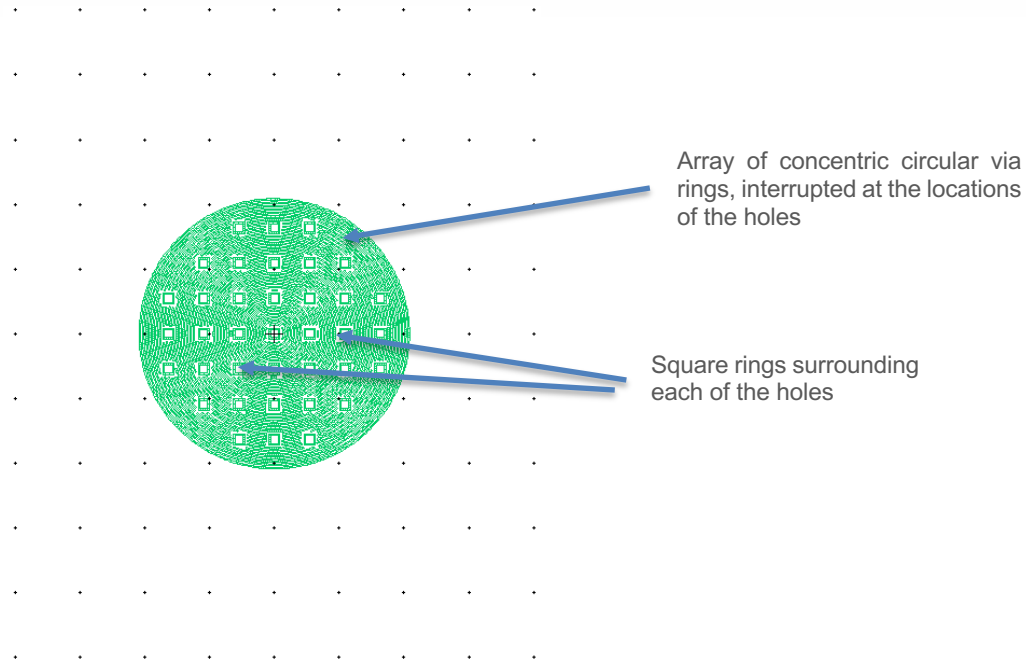


Figure 6. V2 layer.

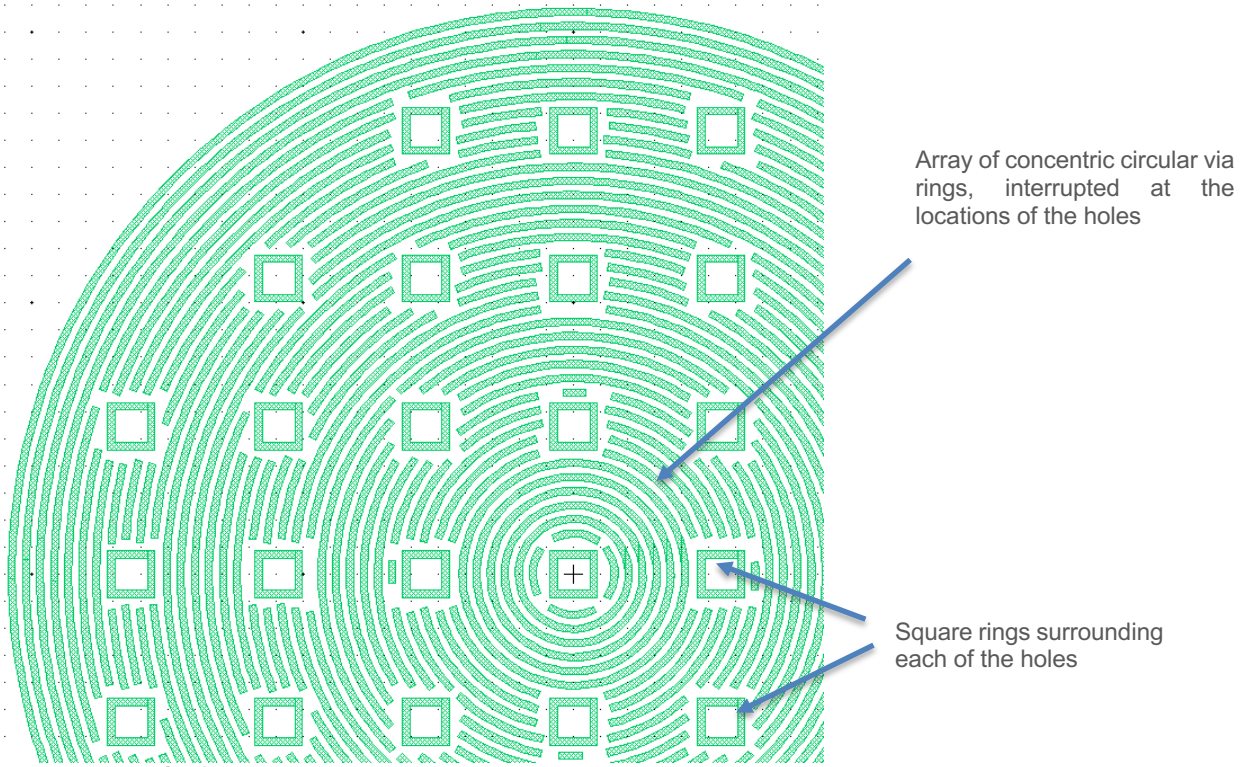


Figure 7. Zoom in of V2 layer.

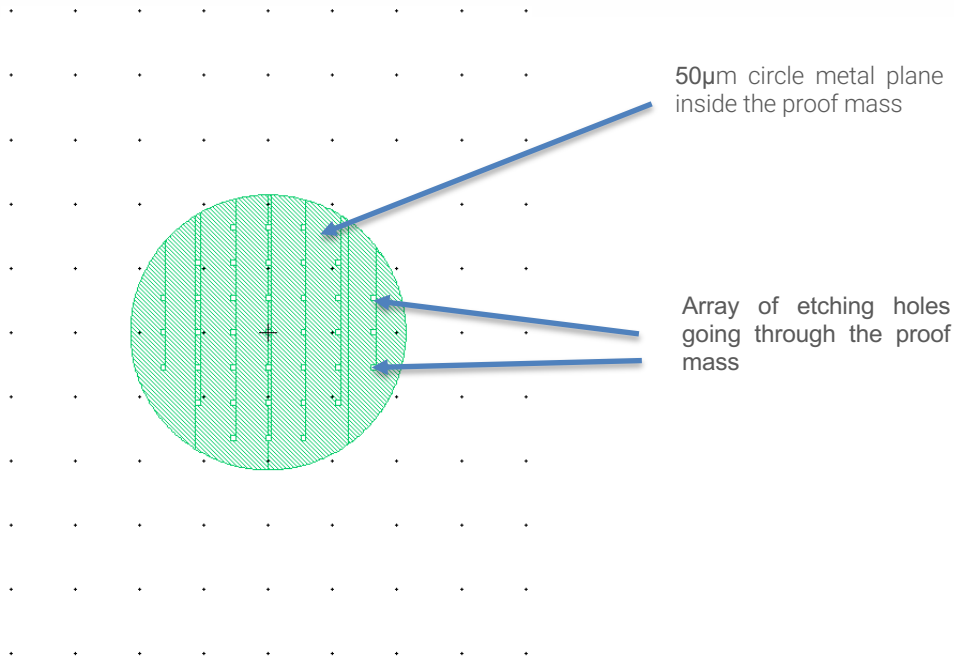


Figure 8. M3 layer.

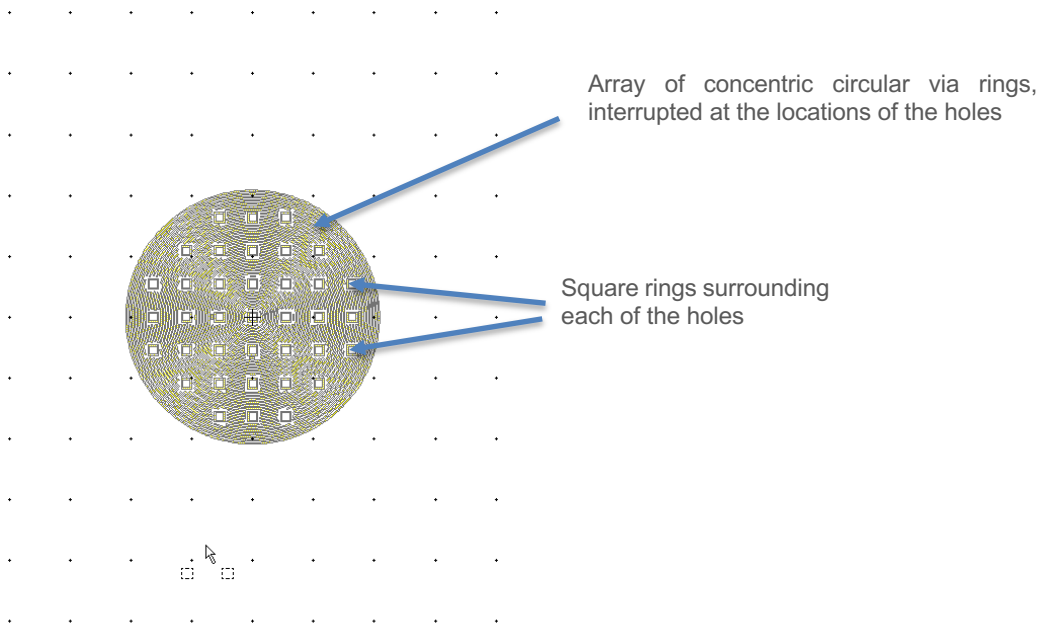


Figure 9. V3 layer. Exactly the same as V2 layer.

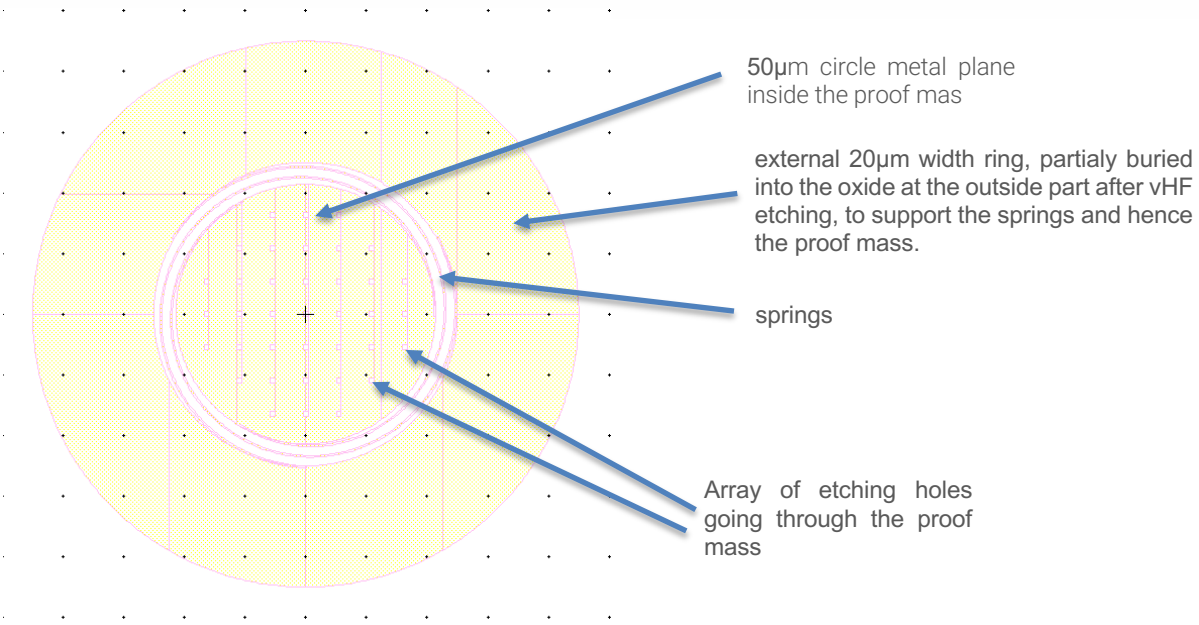


Figure 10. M4 layer.

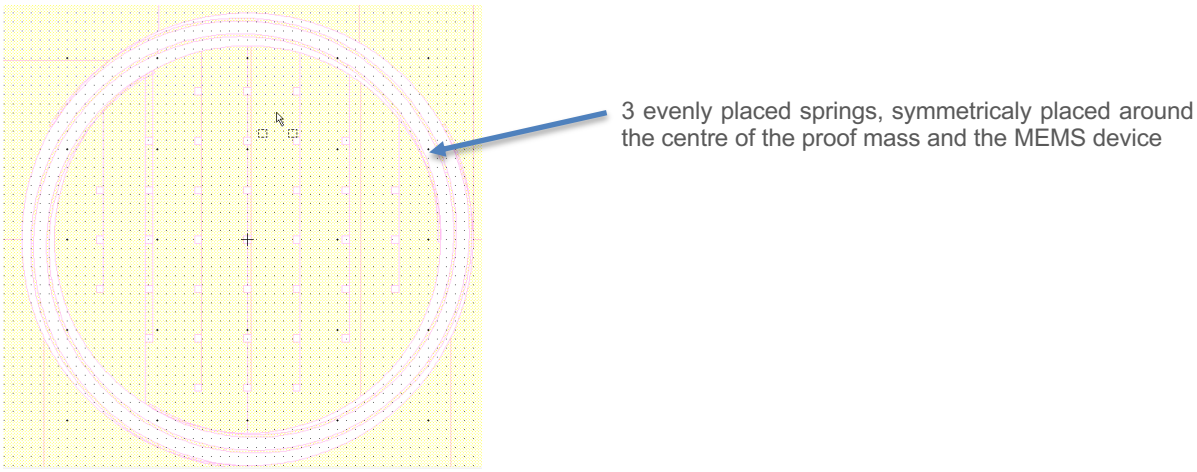


Figure 11. Zoom of proof mass and springs in M4 layer.

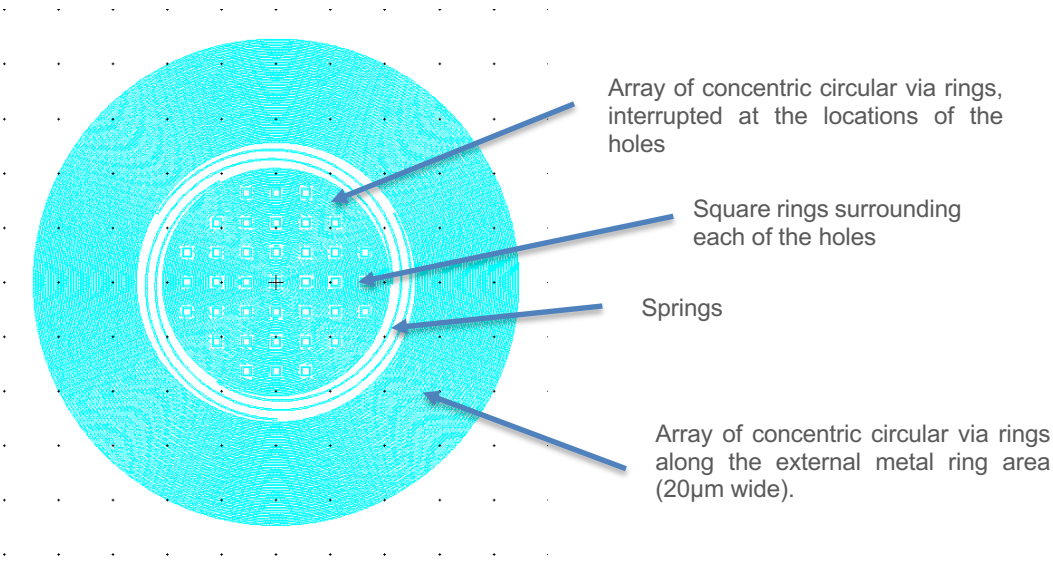


Figure 12. V4 layer.

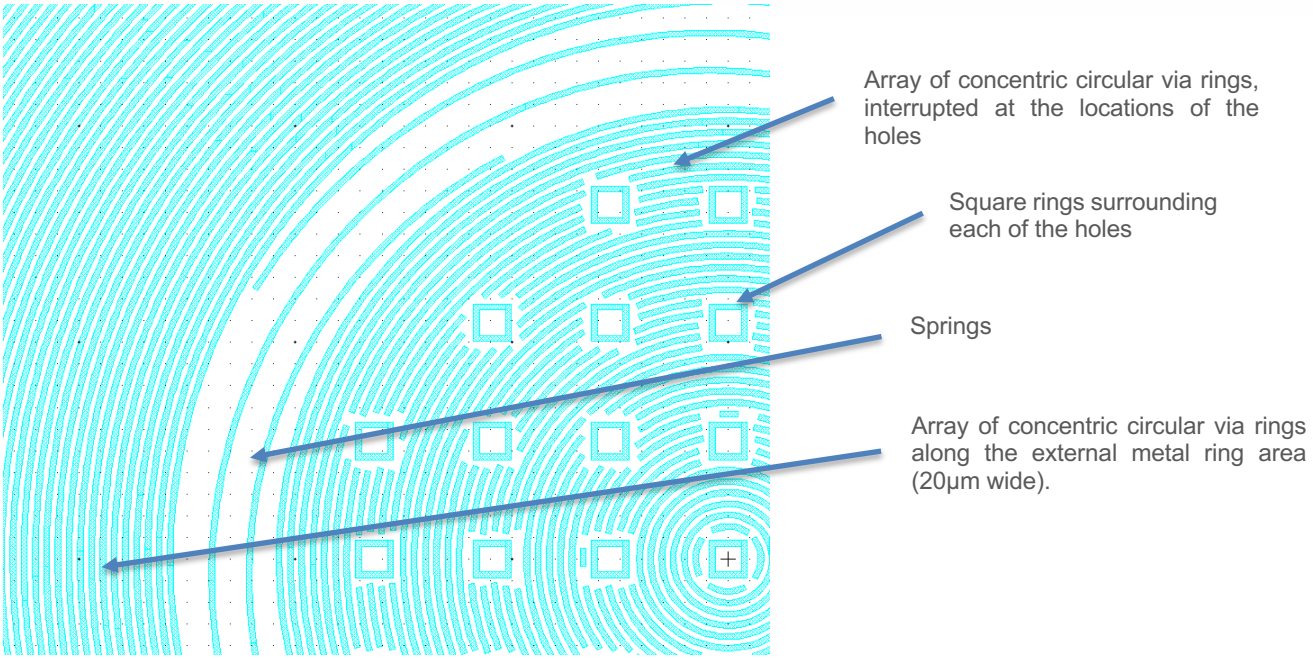


Figure 13. V4 zoom in.

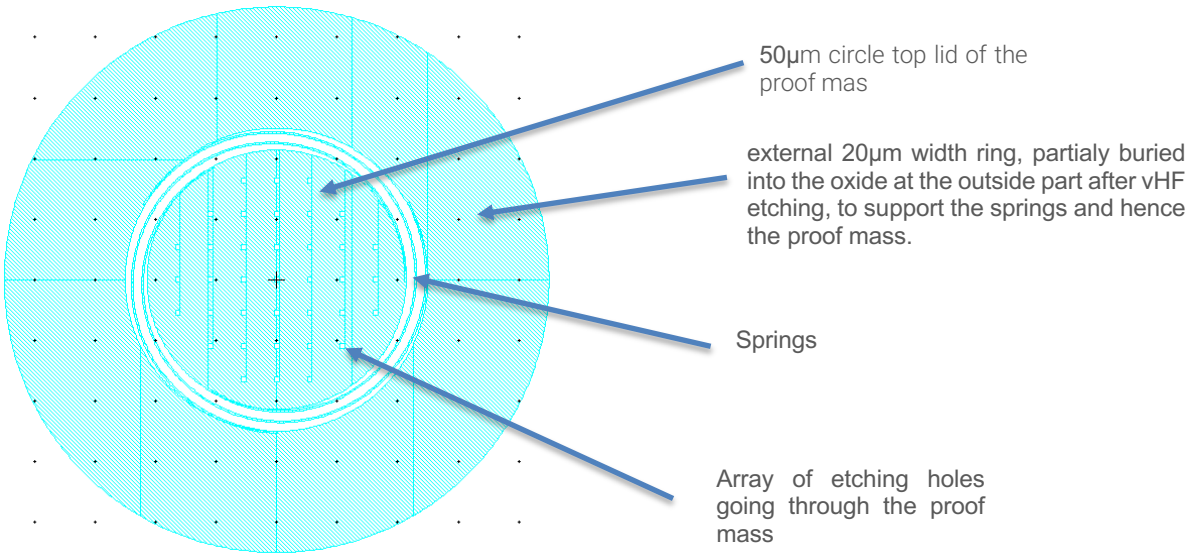
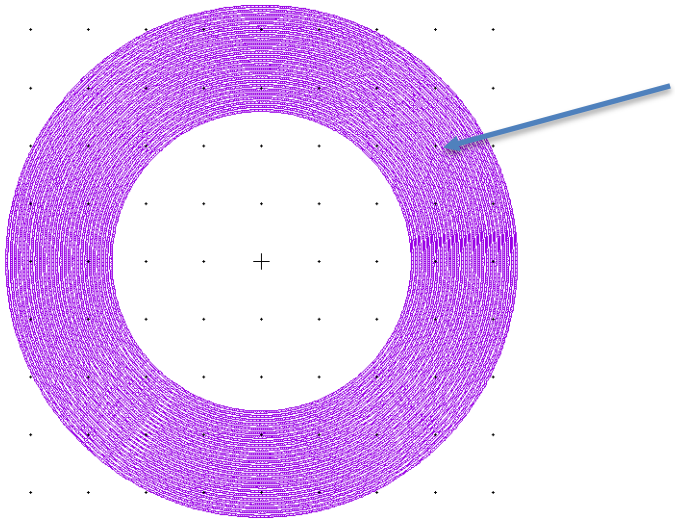
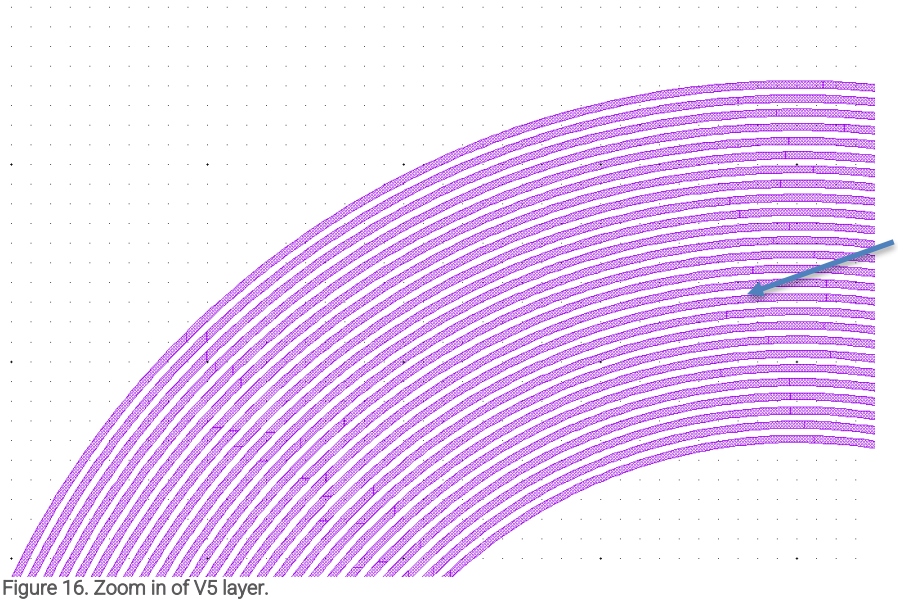


Figure 14. M5 layer. Identical as M4 layer.



Array of concentric circular via rings along the external metal ring area (20µm wide).

Figure 15. V5 layer.



Array of concentric circular via rings along the external metal ring area (20µm wide).

Figure 16. Zoom in of V5 layer.

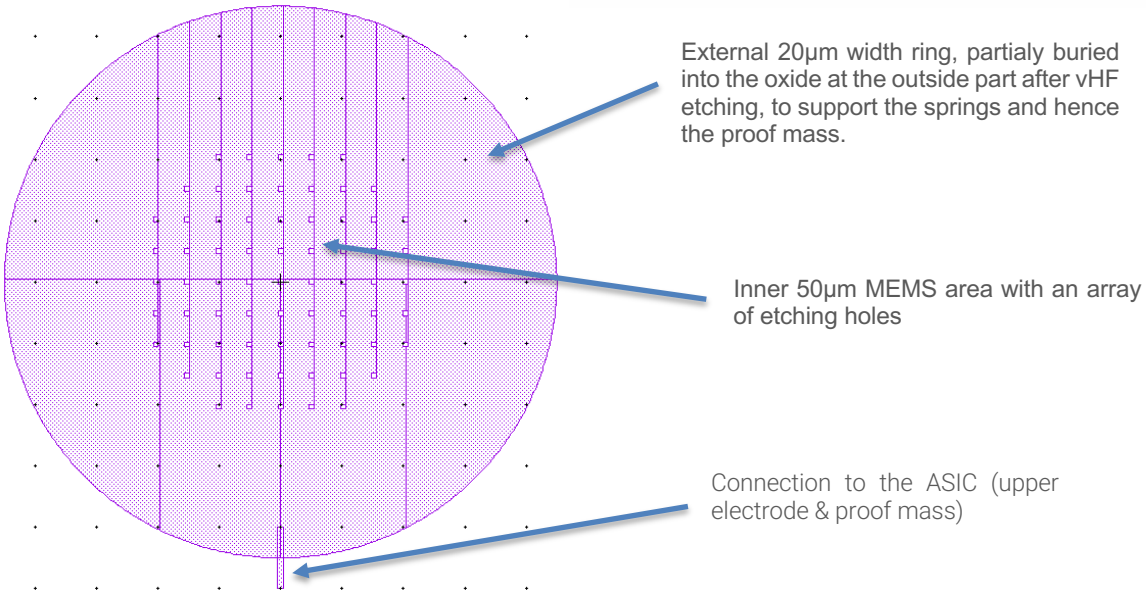


Figure 17. M6 layer.

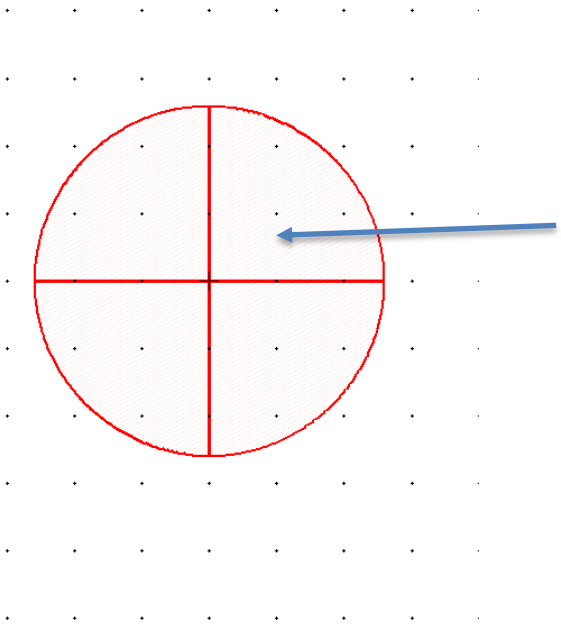
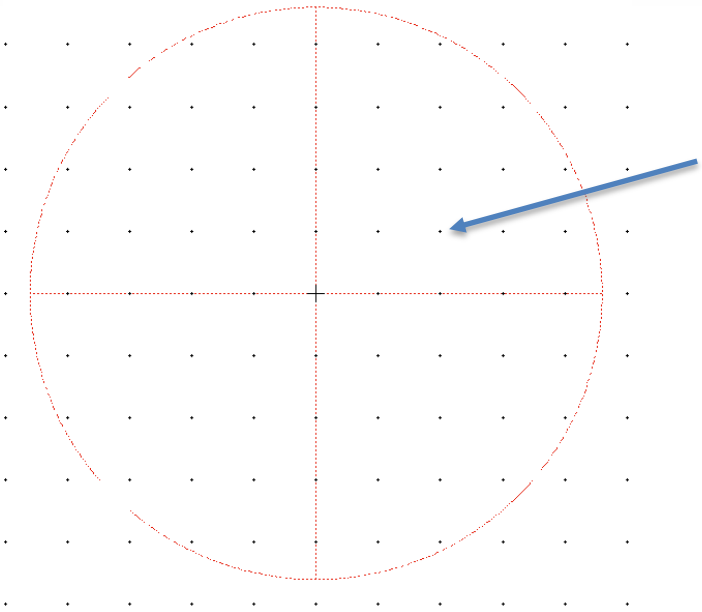


Figure 18. Passivation opening.



92µm circle covering all the extension of the MEMS, including 1µm beyond the outer metal ring. We need to make sure that there is no metal filing inside any part of the MEMS device.

Figure 19. Layer to block the metal filing structures.

The following figures shows a 3D CAD drawing of the metal layers of the device, without showing the silicon oxide, with and without transparency, seen from the top.

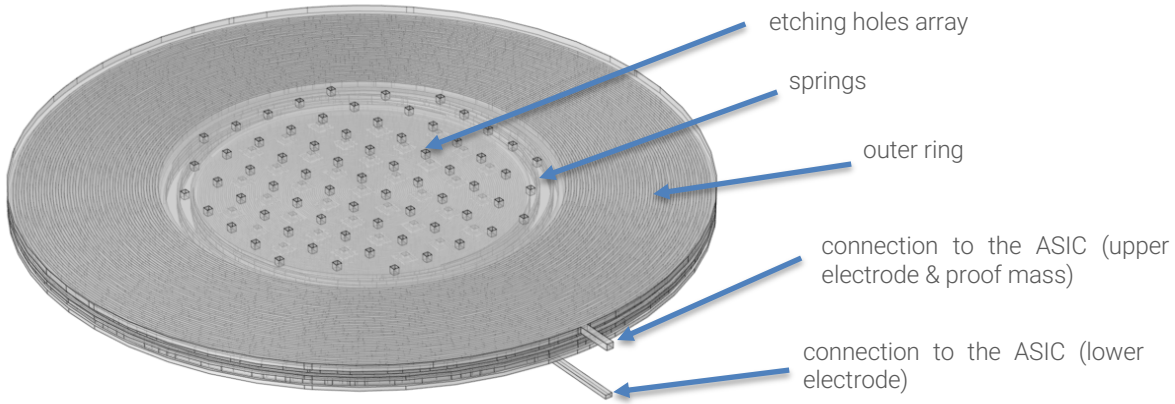


Figure 20. Full 3D CAD model of the metal layers with transparency. Top view.

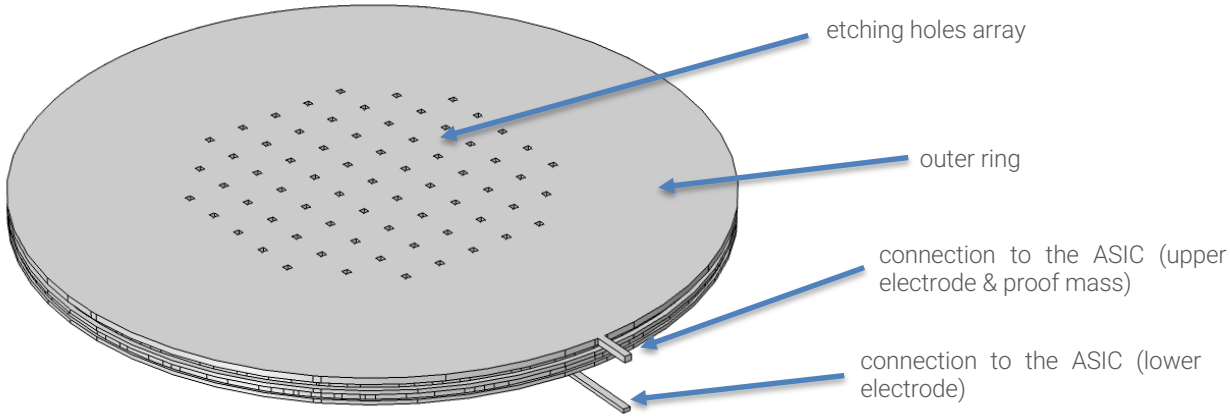


Figure 21. Full 3D CAD model of the metal layers without transparency. Top view.

The following images show the same but seen from the bottom.

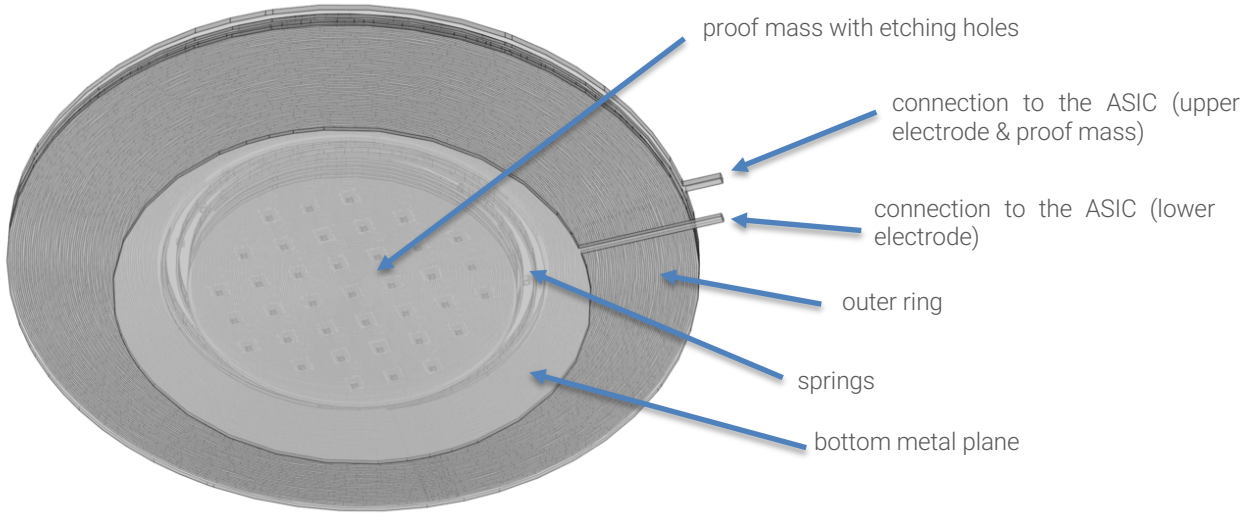


Figure 22. Full 3D CAD model of the metal layers with transparency. Bottom view.

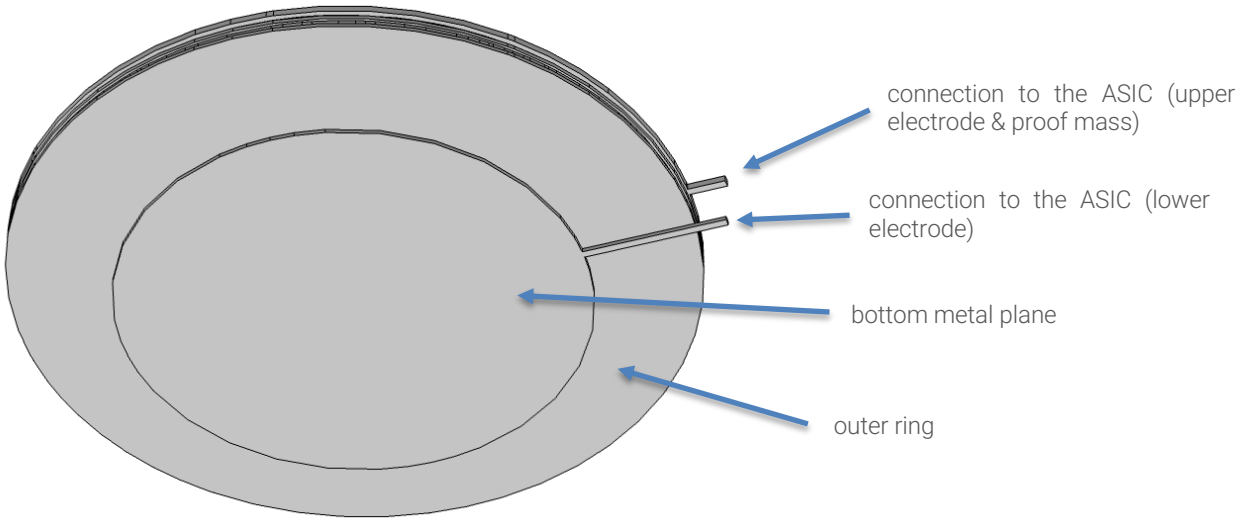


Figure 23. Full 3D CAD model of the metal layers without transparency. Bottom view.

The following figure shows the same seen from one side and without transparency.

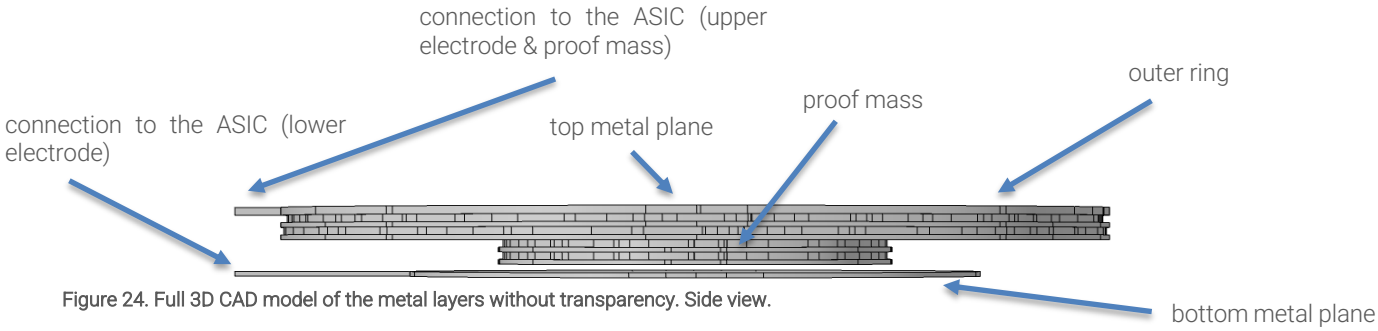


Figure 24. Full 3D CAD model of the metal layers without transparency. Side view.

The following figures show in a sequence all the metal layers, starting with M1 and then adding one by one each metal layer (except V1 which is empty) up to M6.

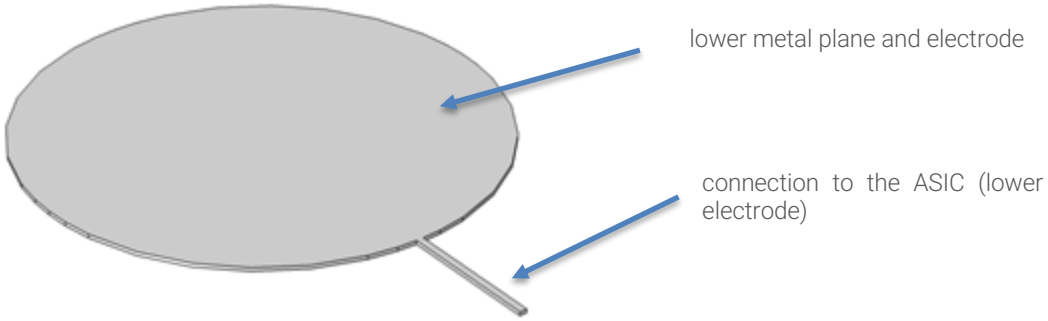


Figure 25. 3D CAD drawing of the metal layers without transparency. Only M1.

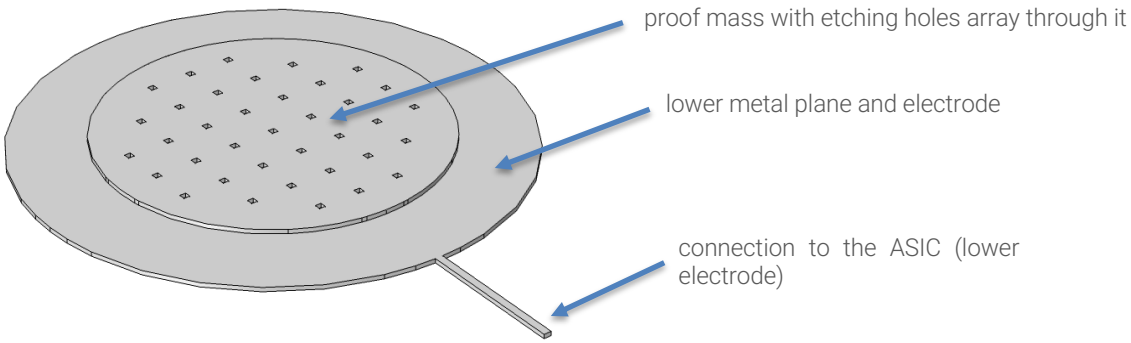


Figure 26. 3D CAD drawing of the metal layers without transparency. M1 and M2.

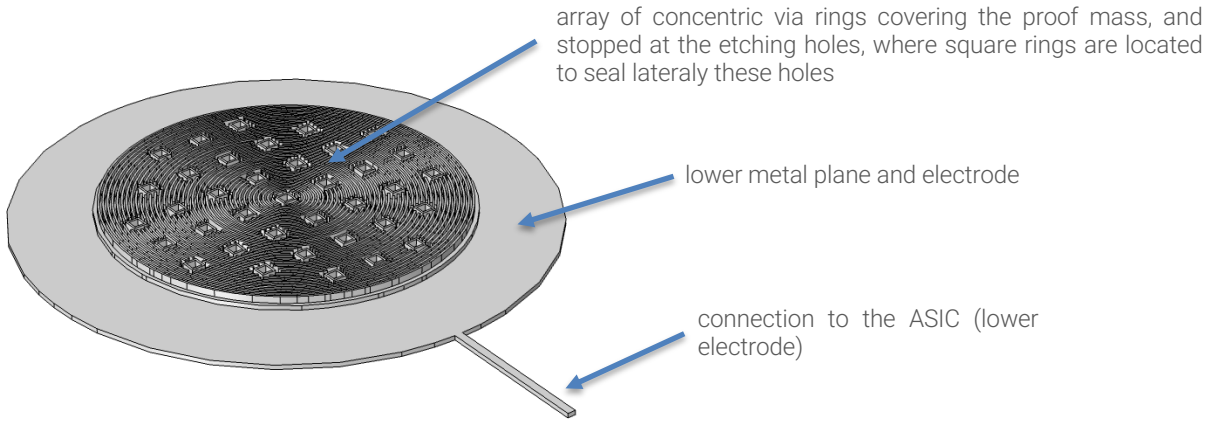


Figure 27. 3D CAD drawing of the metal layers without transparency. M1, M2 and V2.

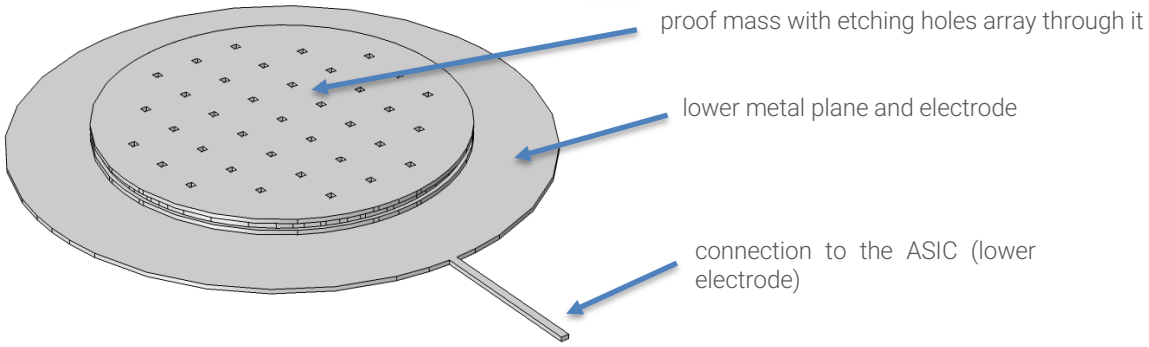


Figure 28. 3D CAD drawing of the metal layers without transparency. M1, M2, V2 and M3.

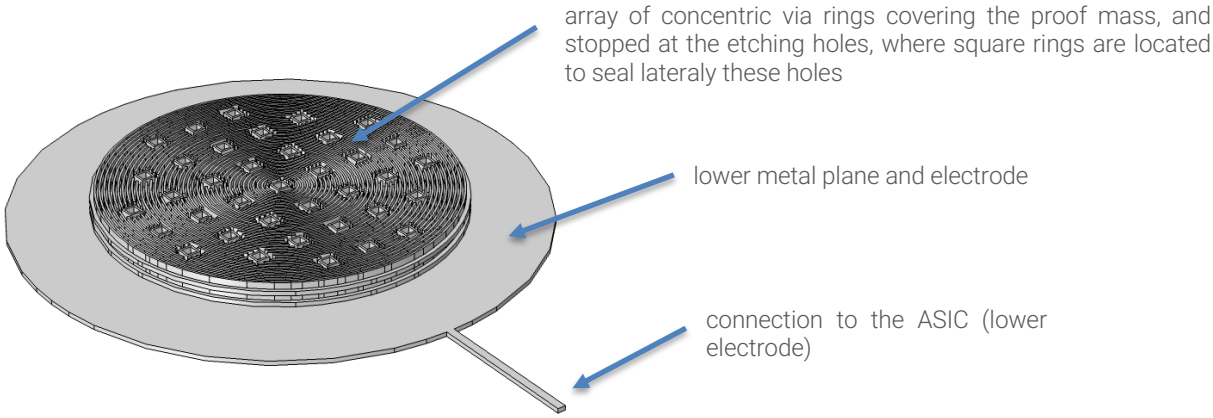


Figure 29. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3 and V3.

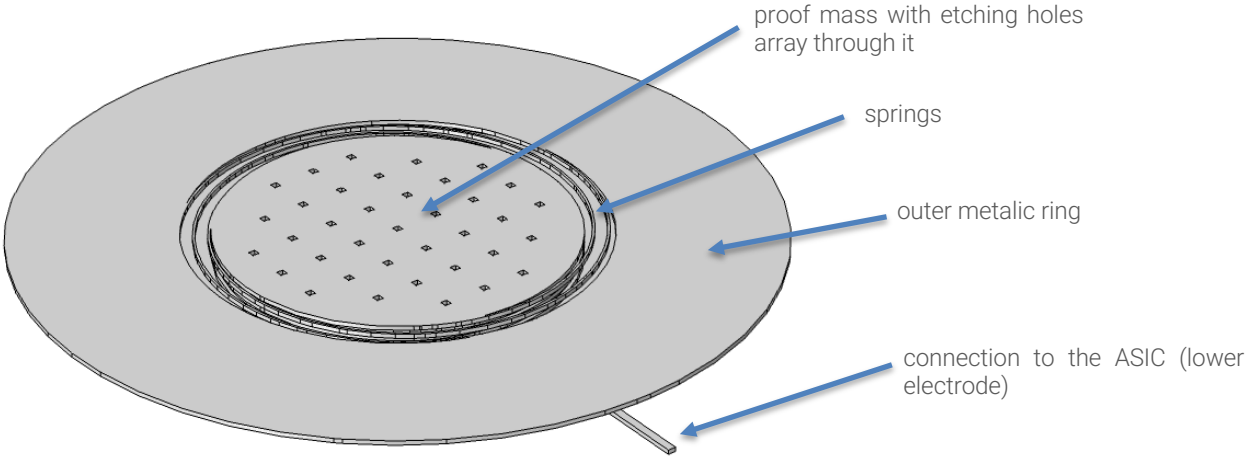


Figure 30. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3, V3 and M4.

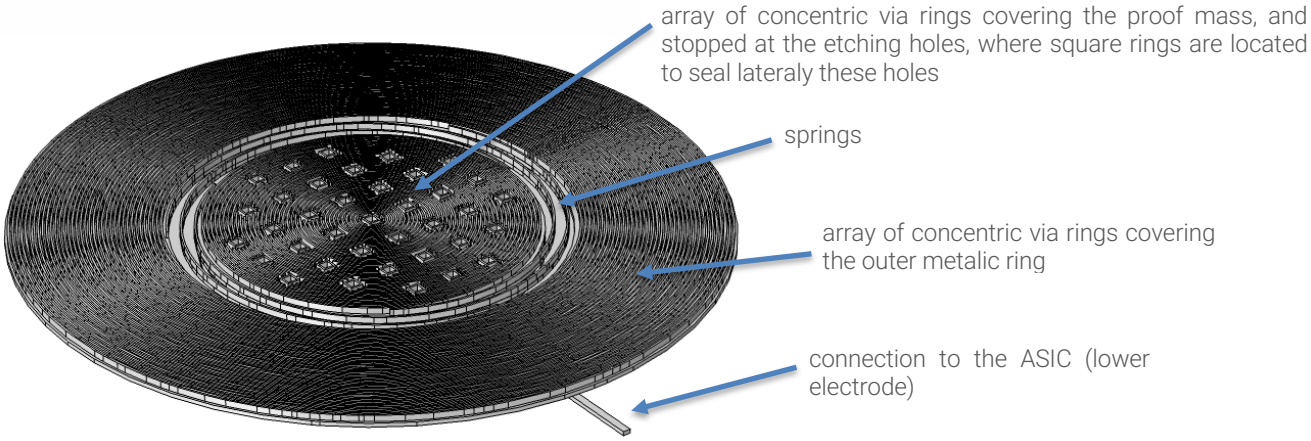


Figure 31. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3, V3, M4 and V4.

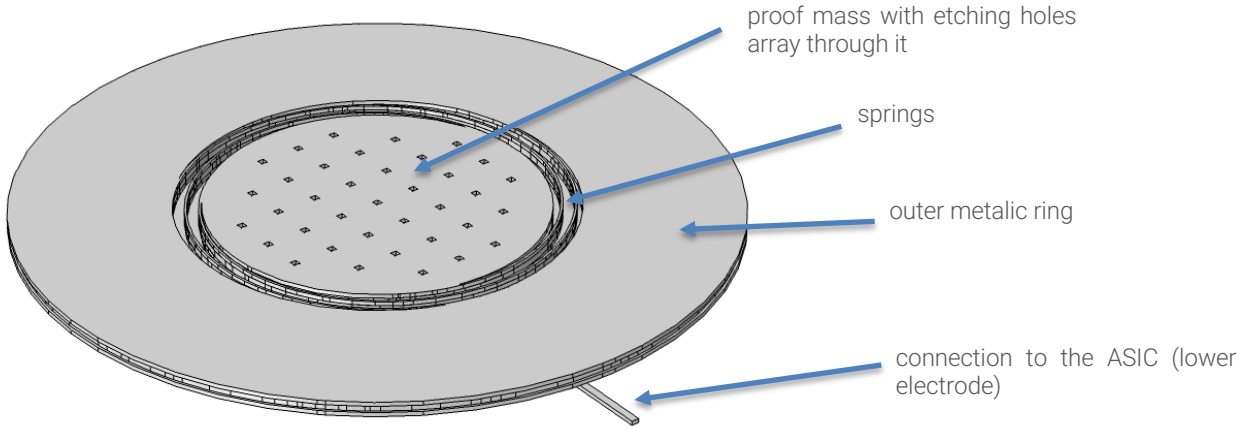


Figure 32. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3, V3, M4, V4 and M5.

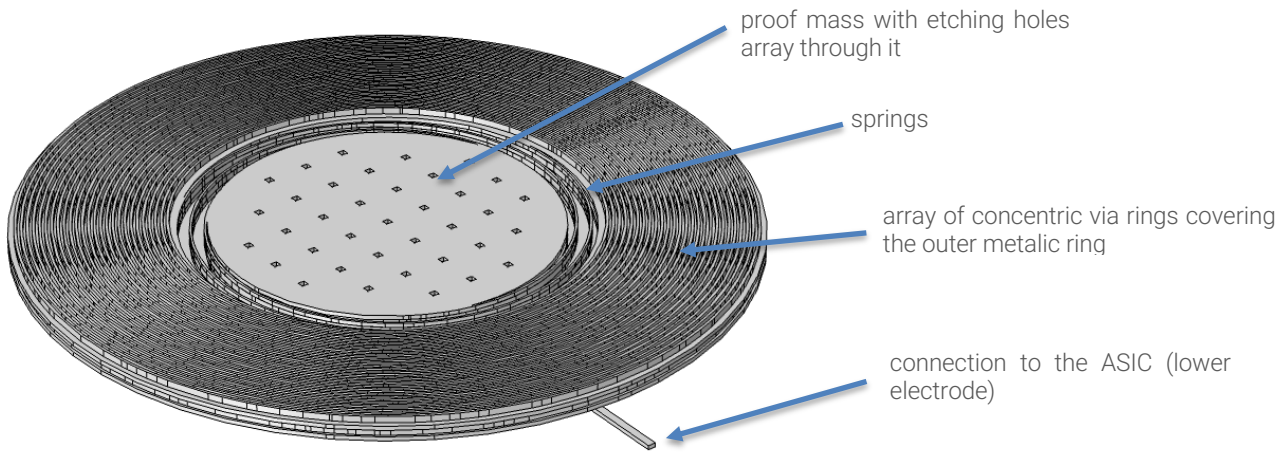


Figure 33. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3, V3, M4, V4, M5 and V5.

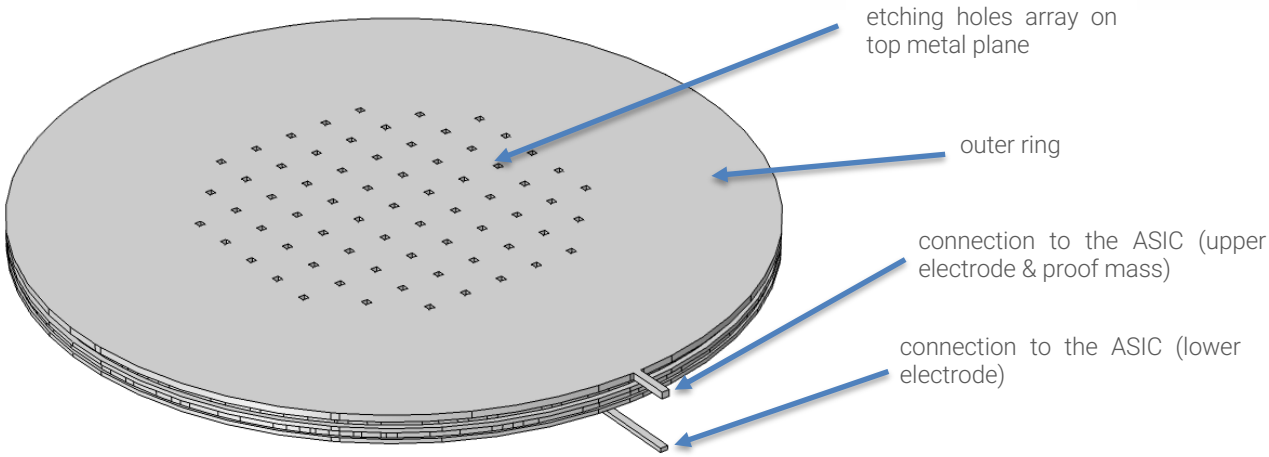


Figure 34. 3D CAD drawing of the metal layers without transparency. All metal layers: M1, M2, V2, M3, V3, M4, V4, M5, V5 and M6.

The following figure shows an exploded view of the 3D cad.

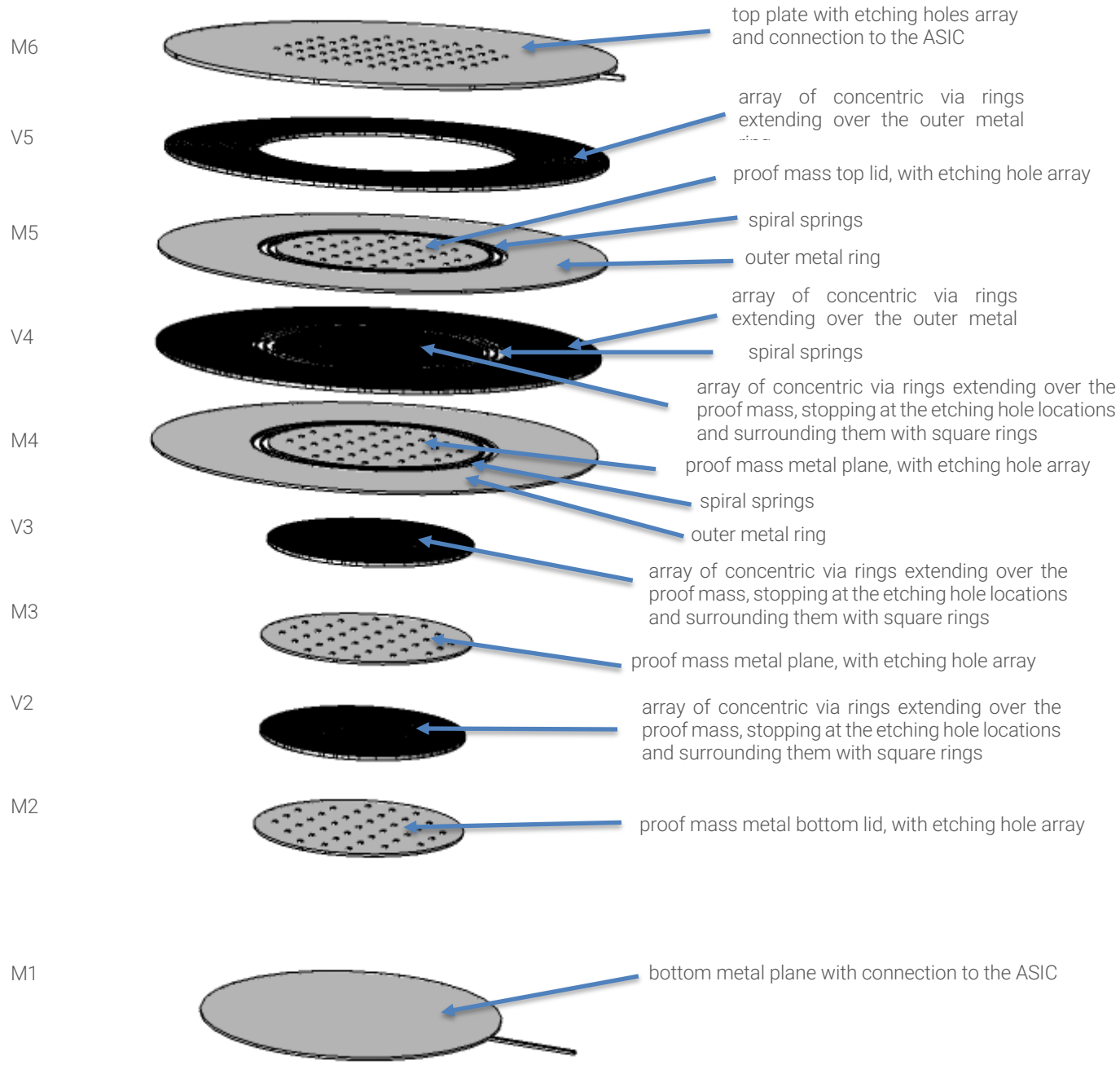


Figure 35. Exploding view of the 3D CAD of the metal and via layers only without showing SiO2, passivation and substrate.

The following figures shows a zoom in of M5, so that we can better see the spiral springs design.

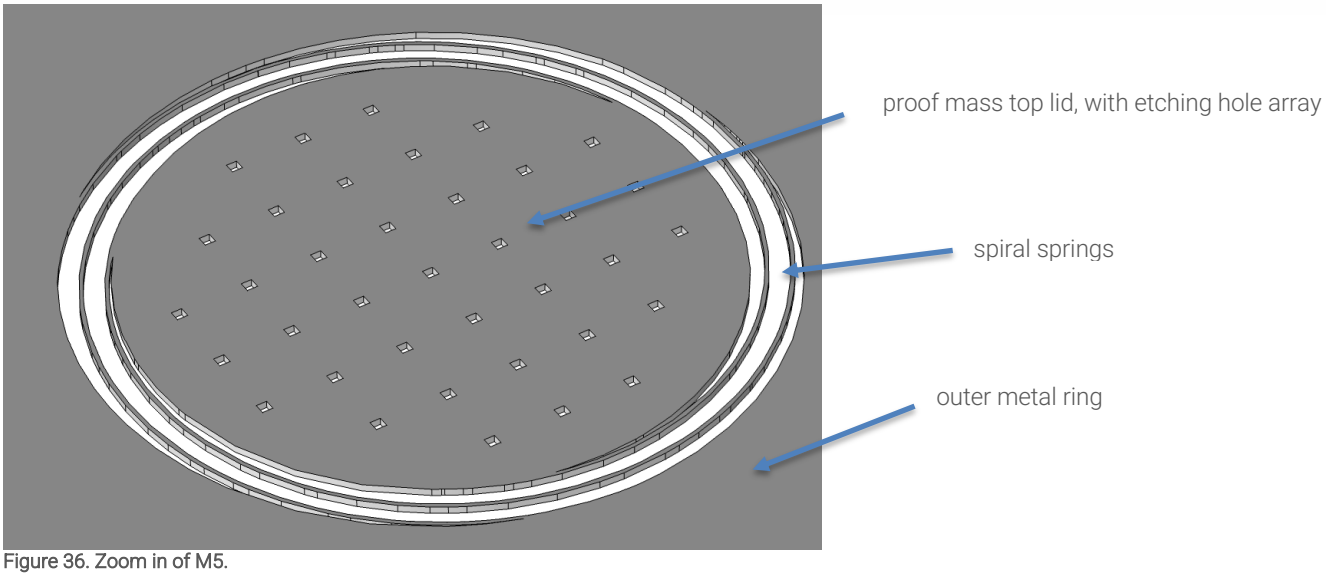


Figure 36. Zoom in of M5.

The following figure shows a zoom in of V4, so it is easier to see the details in it.

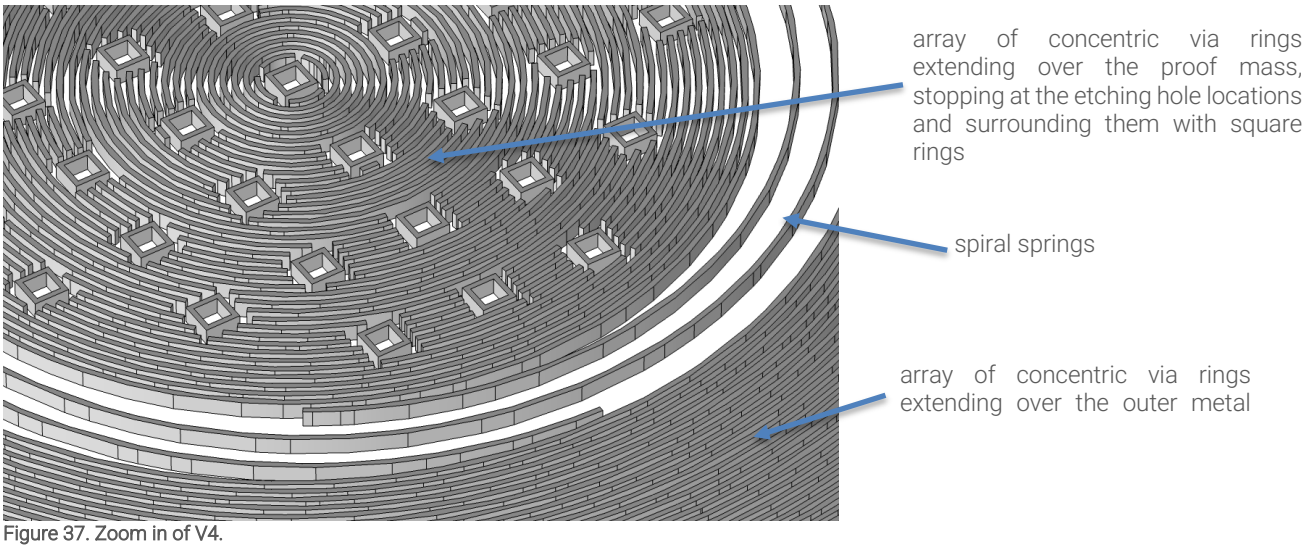


Figure 37. Zoom in of V4.

The following figure shows a schematic cross section of the device.

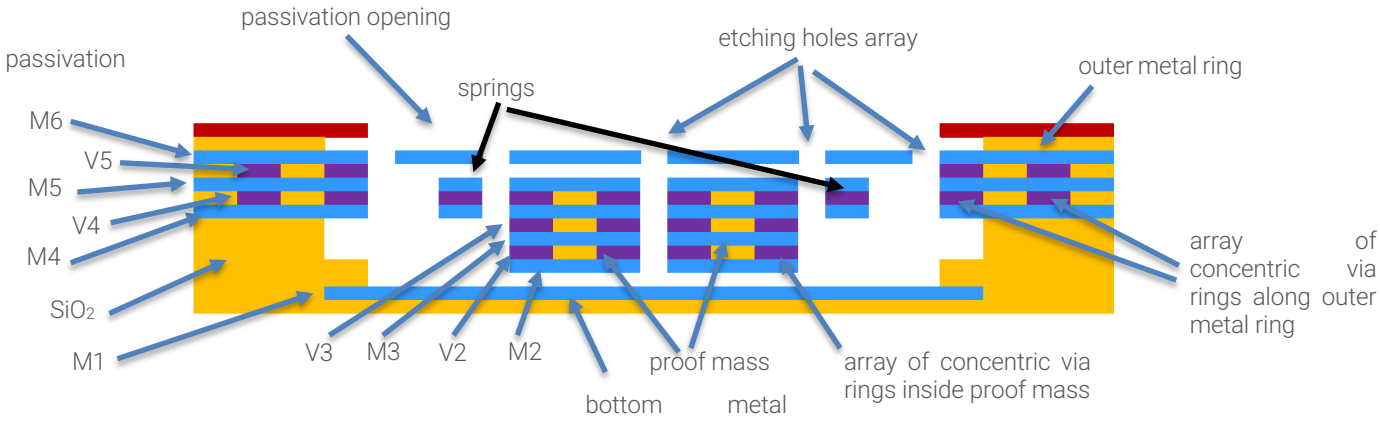


Figure 38. Schematic cross section of the device.

A variant of this design to sense in plane acceleration consists of adding lateral electrodes around the bottom part of the proof mass where there are no springs. In this case, the proof mass is made up with 4 metal layers, M2 to M5. The springs are made with metal layers M4 and M5, having outer metallic rings to support them. Therefore, the proof mass does not use the metals around it at layers M2 and M3. This way we can use layers M2 and M3 to build lateral electrodes for the proof mass. The shape of these lateral electrodes is essentially like the outer rings of the above metal layers (M4 and M5), but, instead of a whole ring, there are two half rings. Each of these half rings is made with the two available metal layers (M2 and M3) stacked up together, which means that there are plenty of vias inside to join them. These vias are shaped as an array of concentric half rings.

The external diameter of these lateral electrodes is shorter than the outer ring of the upper metal layers. In this design, it is also made shorter than the bottom metal plane, but an improvement would be to make the bottom metal plane diameter smaller than the outer diameter of these lateral electrodes.

Rather than repeating all the drawings, in the following figures we show the difference in the design.

Since we keep having the bottom metal plane, we can still sense out-of-plane acceleration as well. Therefore, the device has several electrodes, which allow for the sensing of 1, 2 or even the 3-axis altogether with the same device. This is possible if, instead of dividing these lateral electrodes in half rings, we divide them in quarter rings. Furthermore, for the X and Y axis (in-plane acceleration) we can have differential capacitances.

As usual, although this design is for an inertial sensor, we could apply the same design principle (electrodes, springs supports etc) to implement other types of capacitive sensors and actuators.

The following image shows the whole layout with all the layers

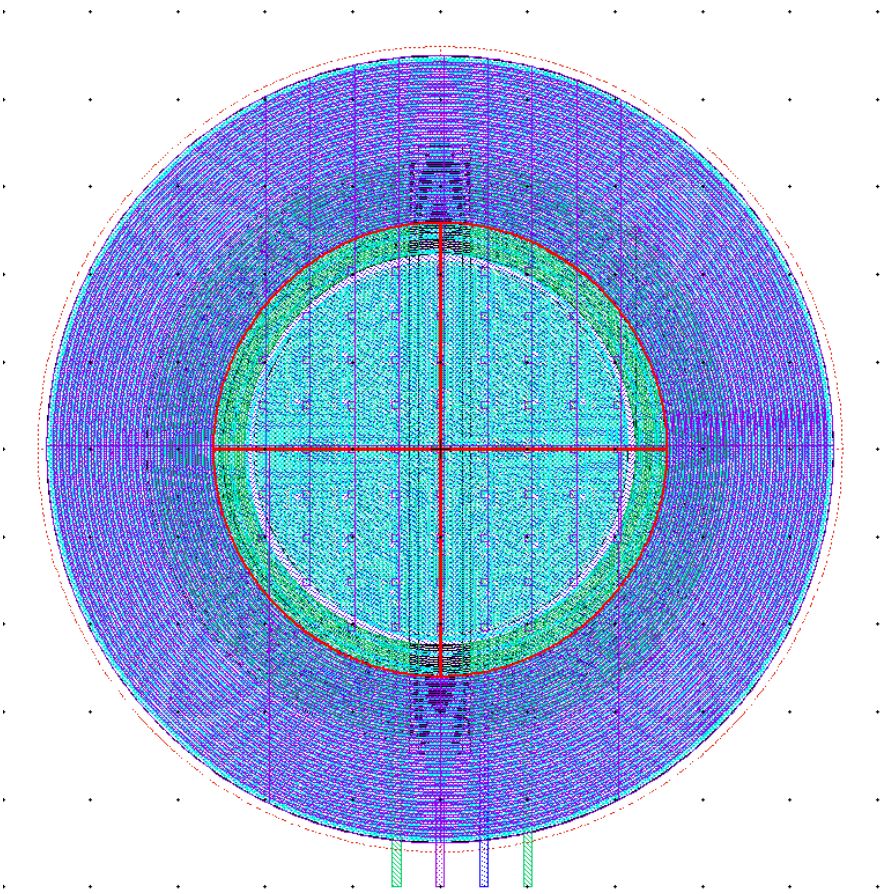


Figure 39. Complete layout for the in-plane inertial sensor.

The following images show the layout of each layer from M1 up to M3, without V1 because it is empty. The rest of layers are exactly the same as with the previous design sensing only out-of-plane acceleration.

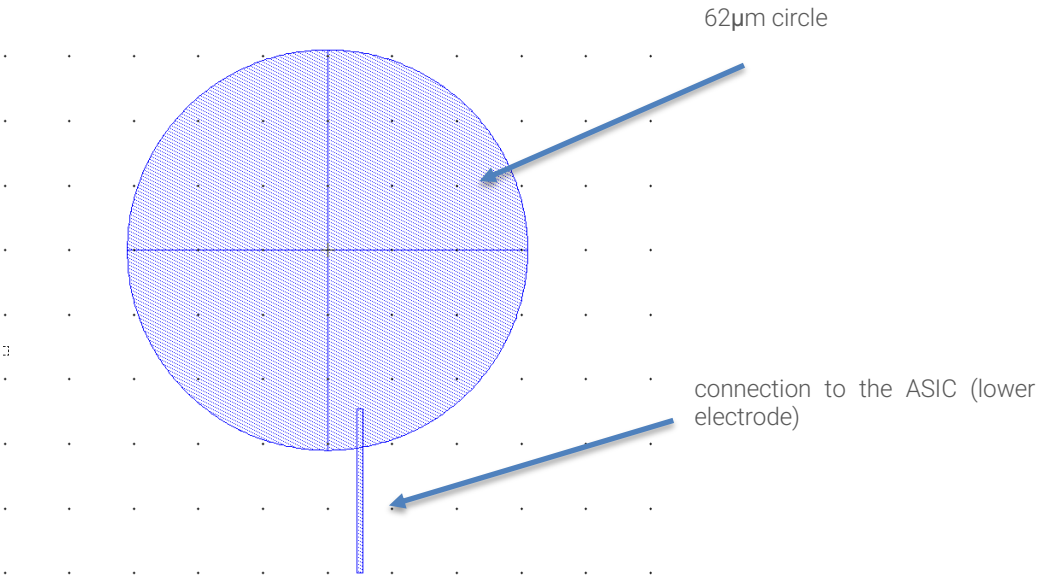


Figure 40. M1 layer. It could actually be made exactly the same as the out-of-plane inertial sensor. However, the connection has been displaced a bit to facilitate the connection of the other electrodes to the ASIC. In general, the electrical connections to each of the electrodes in all the designs can be placed as desired, with the goal to minimize the parasitic capacitance of this connection towards the ASIC so we will want to minimize its length.

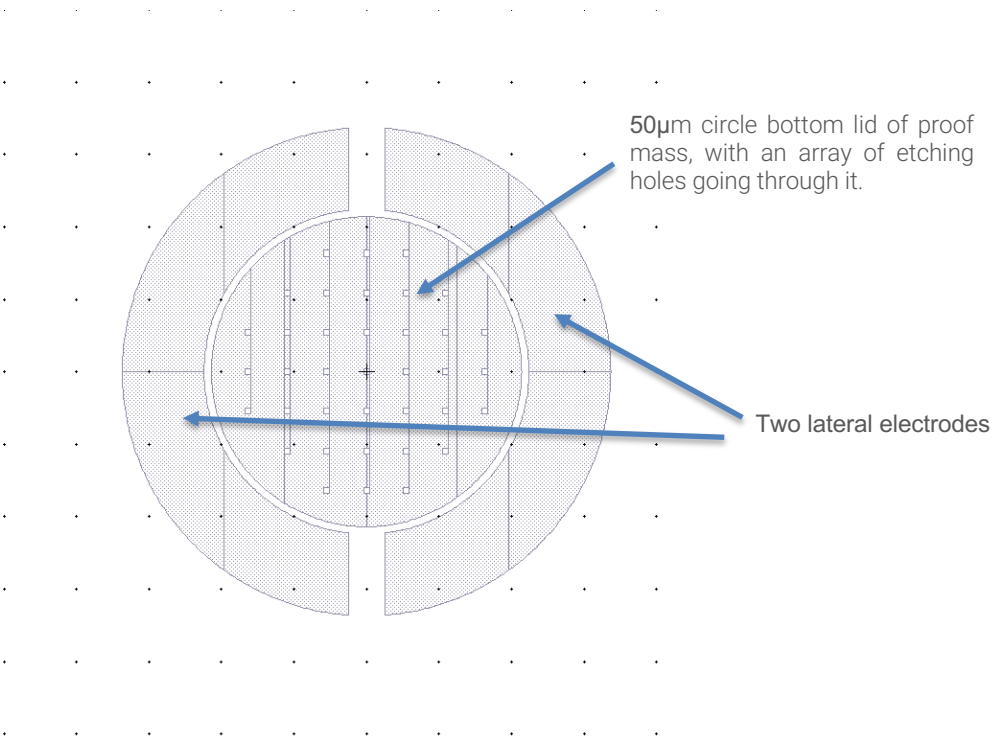


Figure 41. M2 layer.

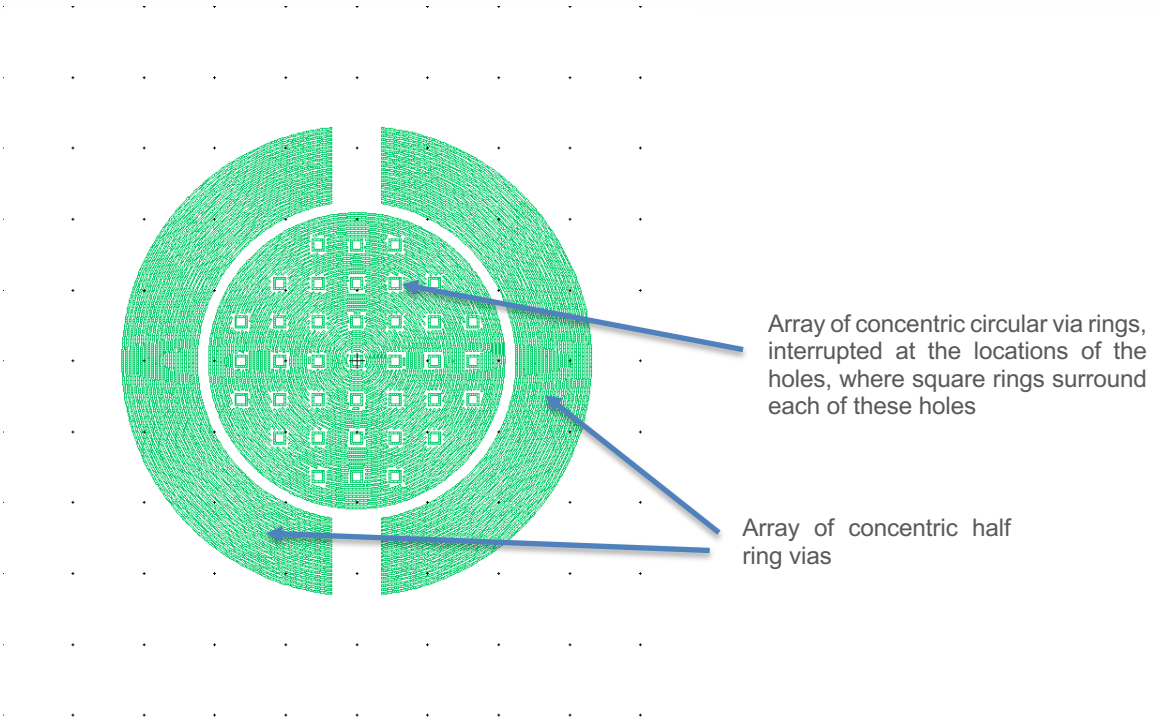


Figure 42. V2 layer.

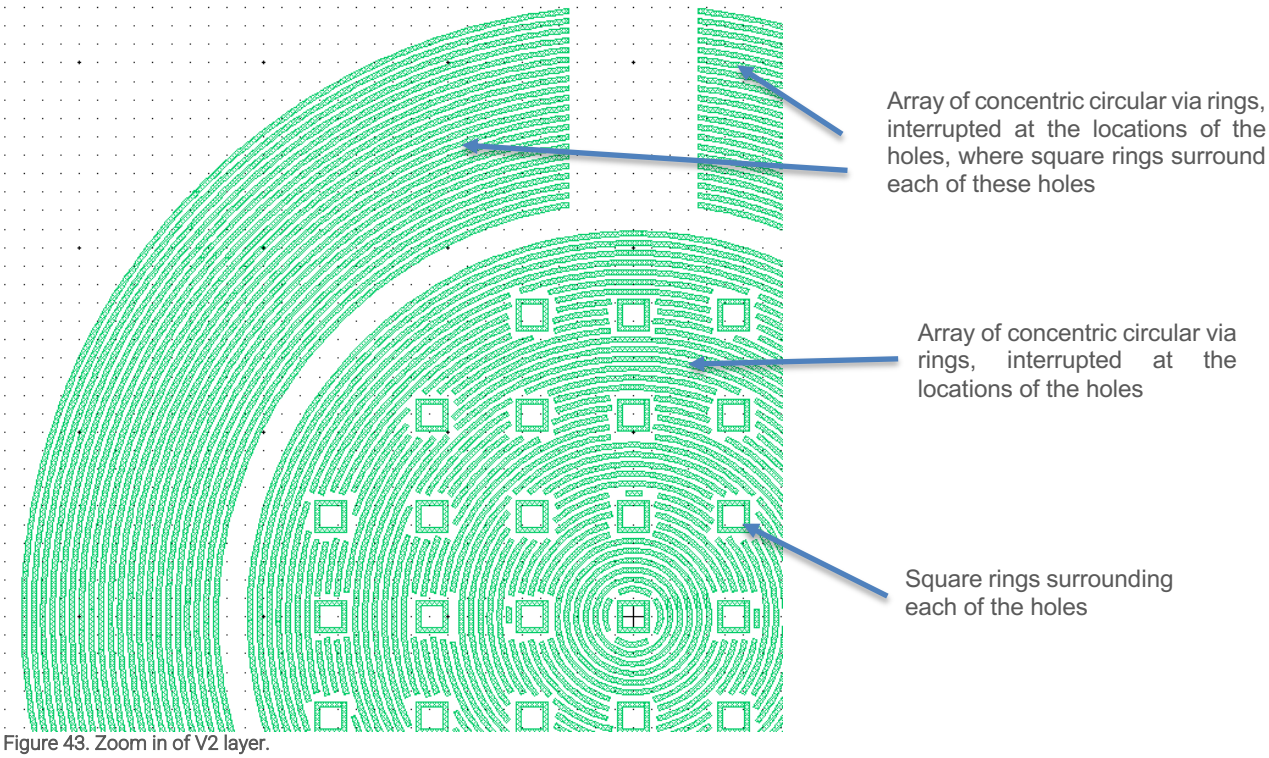


Figure 43. Zoom in of V2 layer.

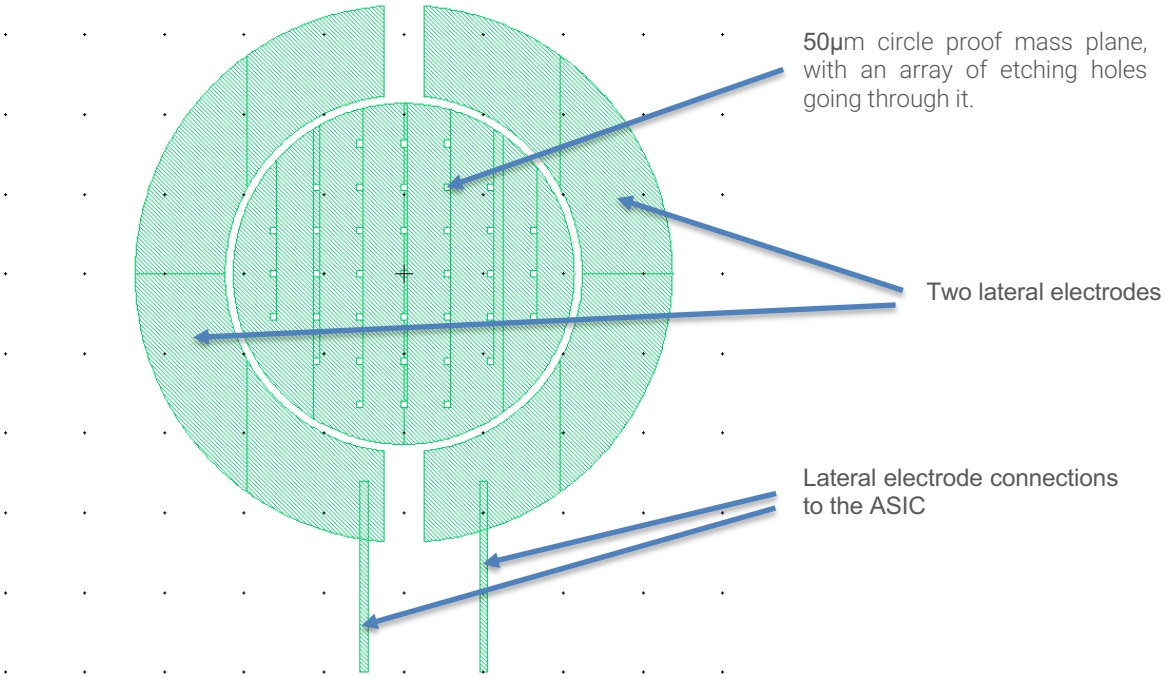


Figure 44. M3 layer.

The following figures shows a 3D CAD drawing of the metal layers of the device, without showing the silicon oxide, with and without transparency, seen from the top.

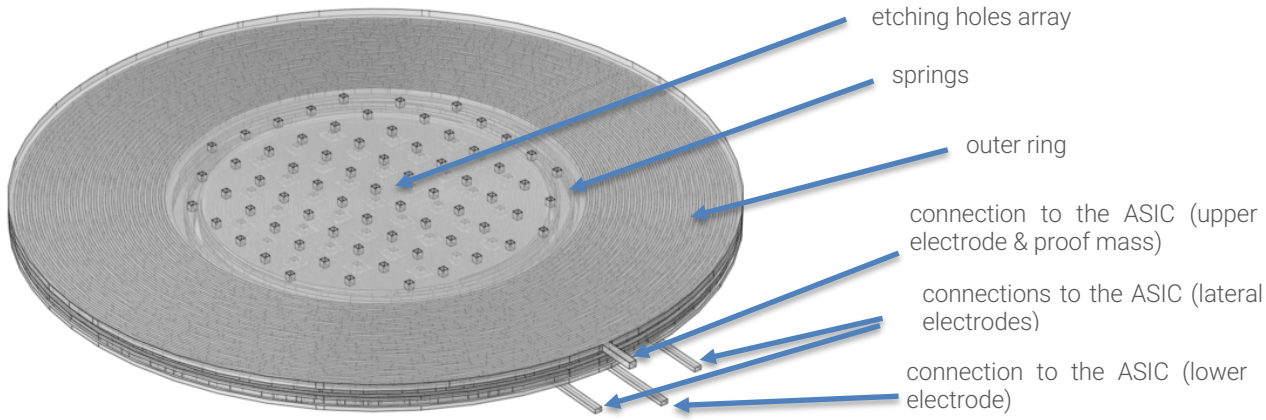


Figure 45. Full 3D CAD model of the metal layers with transparency. Top view.

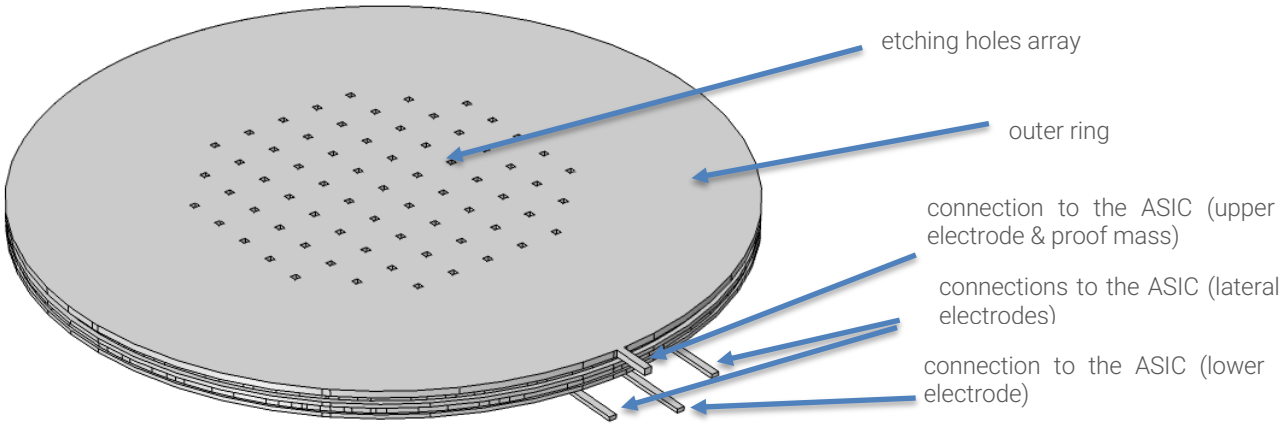


Figure 46. Full 3D CAD model of the metal layers without transparency. Top view.

The following images show the same but seen from the bottom.

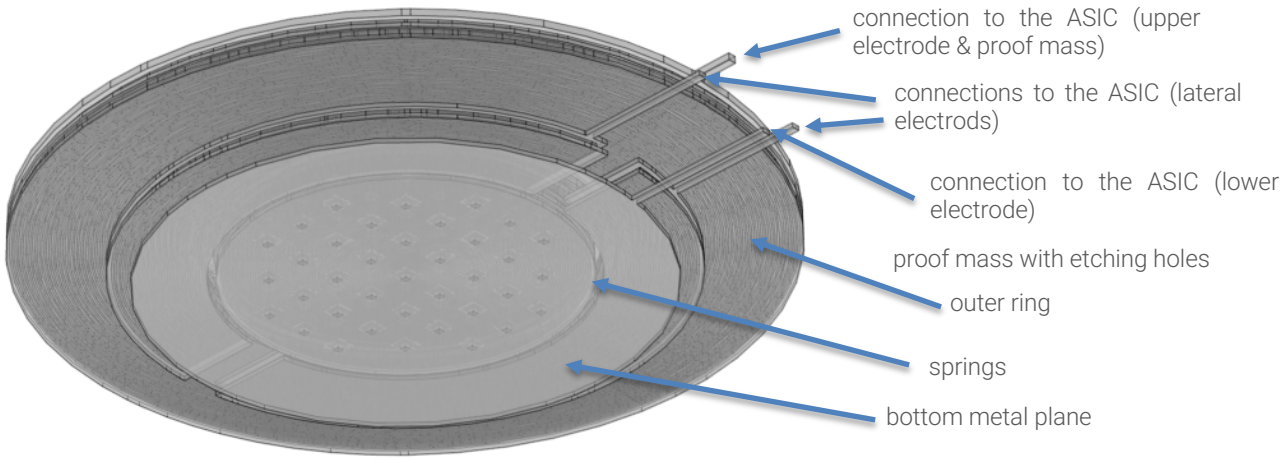


Figure 47. Full 3D CAD model of the metal layers with transparency. Bottom view.

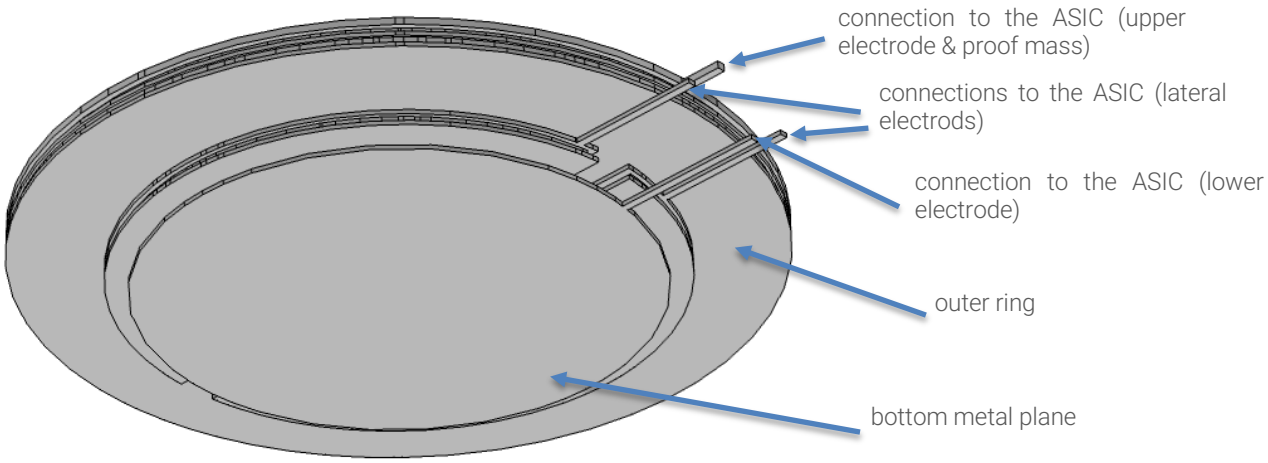


Figure 48. Full 3D CAD model of the metal layers without transparency. Bottom view.

The following figure shows the same seen from one side and without transparency.

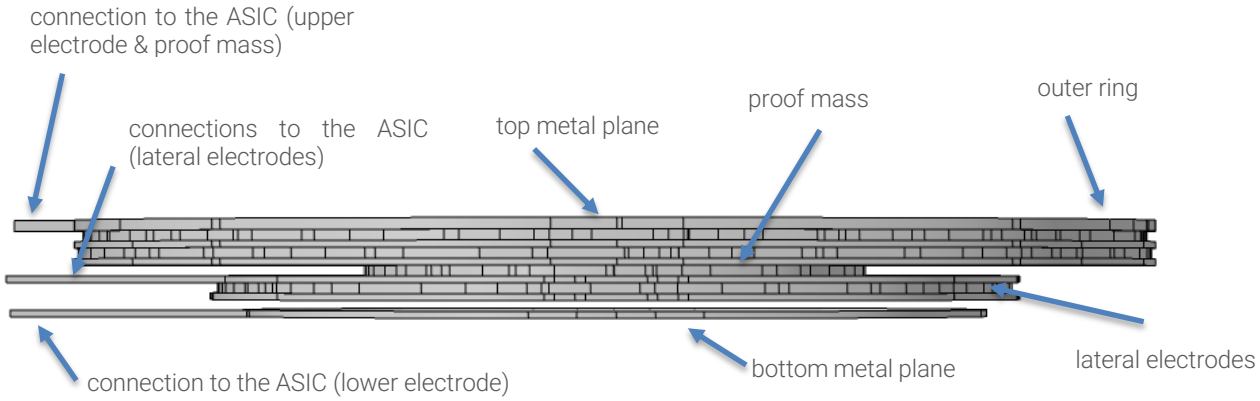


Figure 49. Full 3D CAD model of the metal layers without transparency. Side view.

The following figures show in a sequence of all the metal layers, starting with M1 and then adding one by one each metal layer (except V1 which is empty) up to M6. Although the sequence is identical to the previous description of the out-of-plane only inertial sensor beyond M3 it may help to see these 3D images for this multi-electrode implementation.

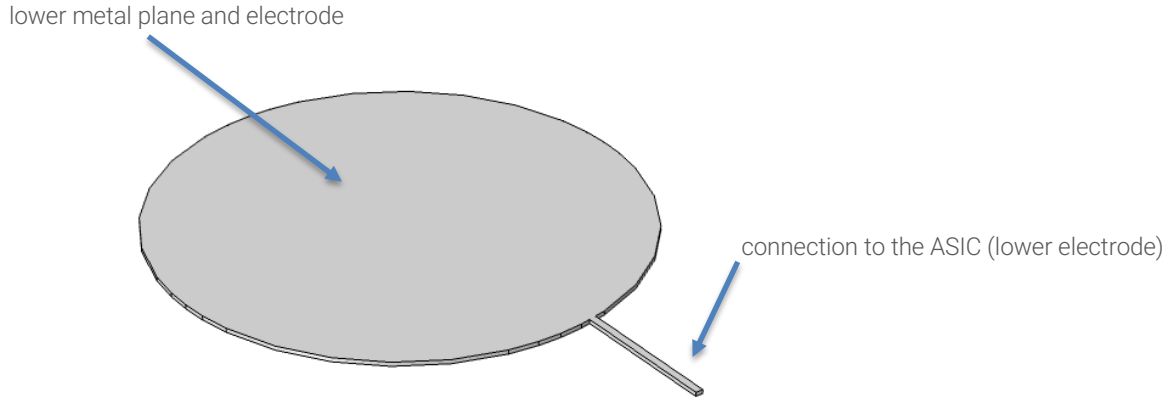


Figure 50. 3D CAD drawing of the metal layers without transparency. Only M1.

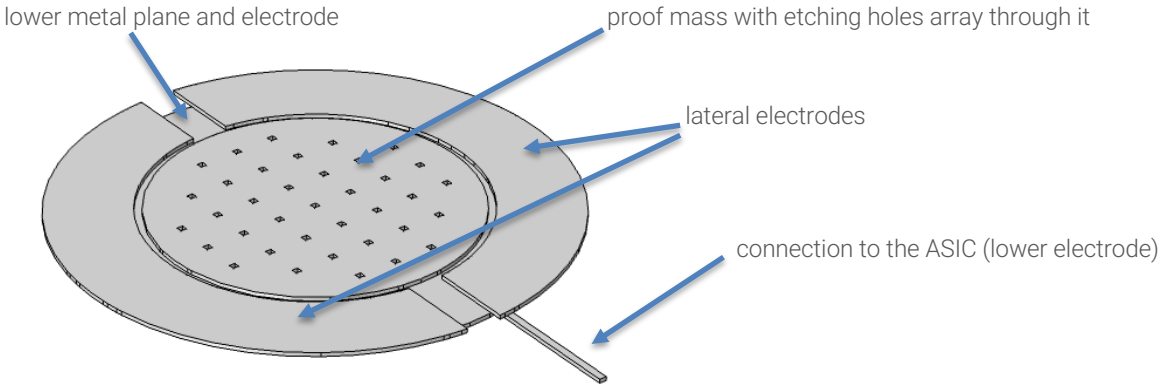


Figure 51. 3D CAD drawing of the metal layers without transparency. M1 and M2.

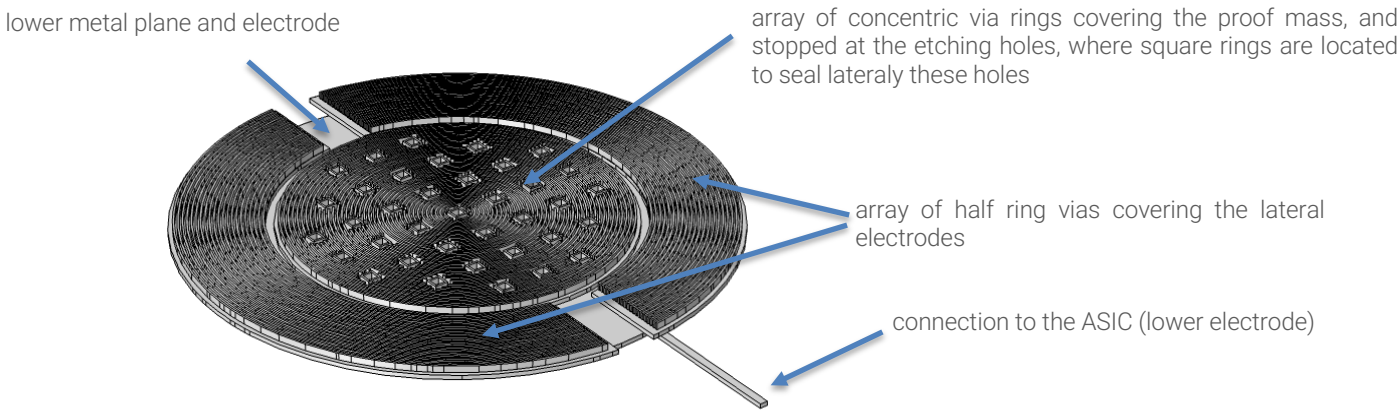


Figure 52. 3D CAD drawing of the metal layers without transparency. M1, M2 and V2.

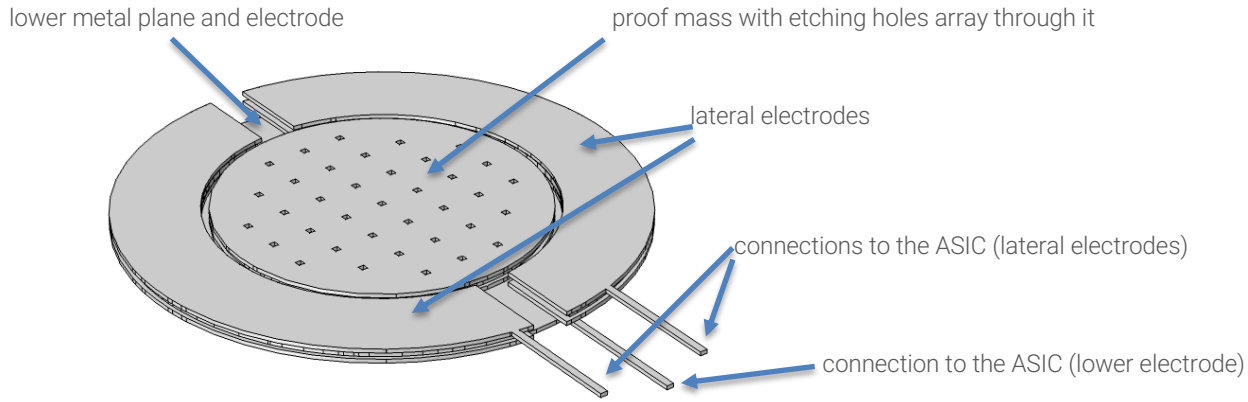


Figure 53. 3D CAD drawing of the metal layers without transparency. M1, M2, V2 and M3.

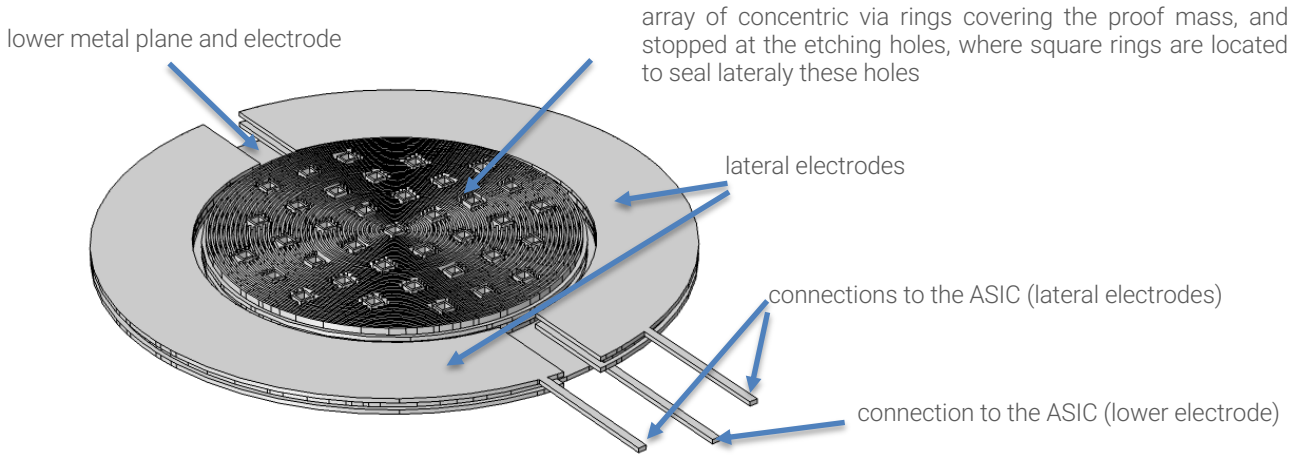


Figure 54. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3 and V3.

proof mass with etching holes array through it

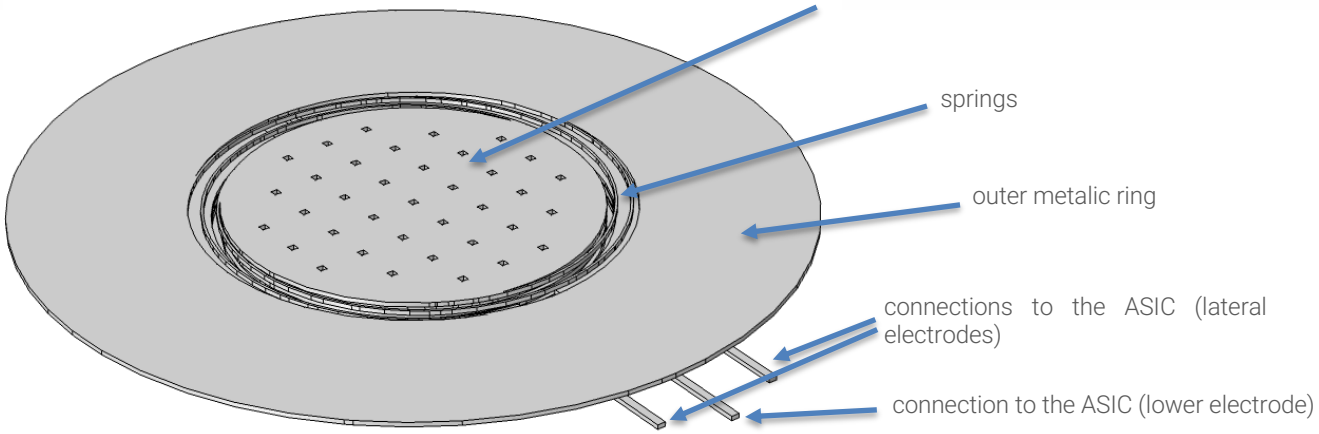


Figure 55. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3, V3 and M4.

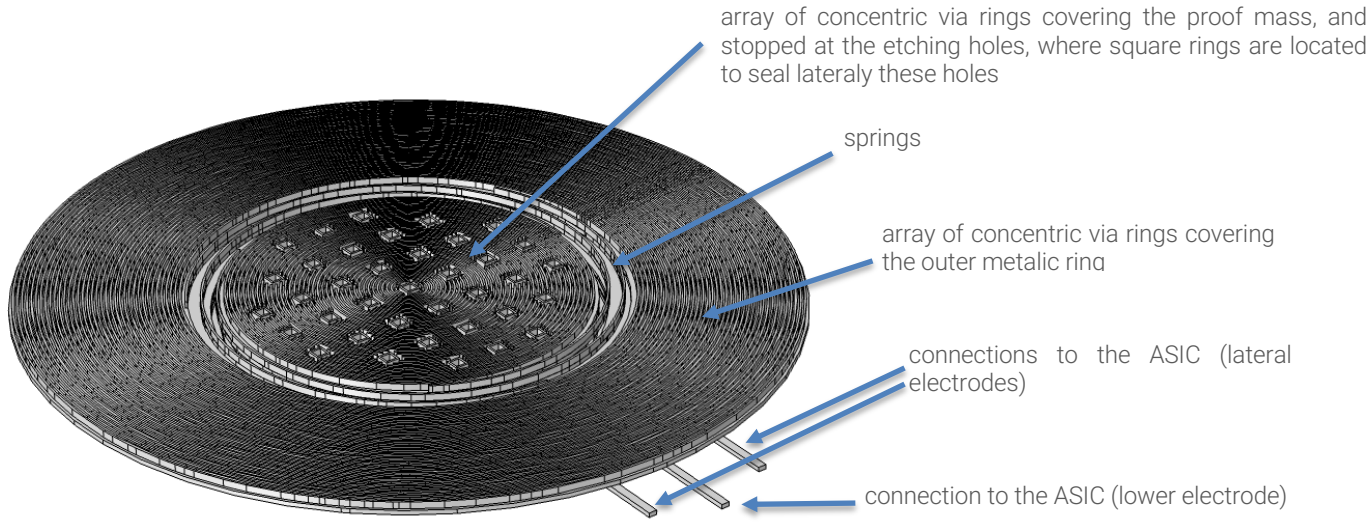


Figure 56. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3, V3, M4 and V4.

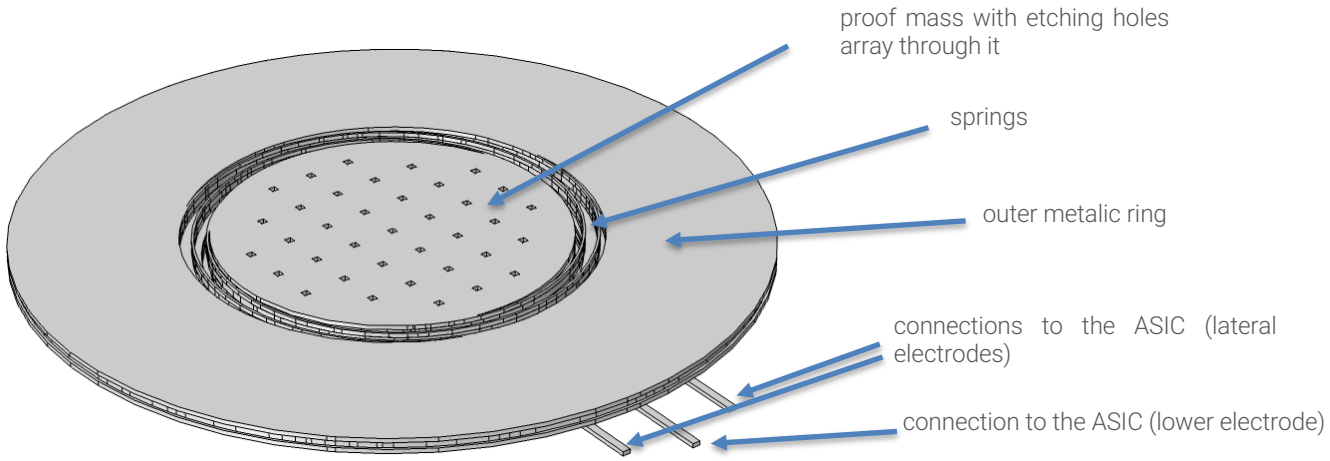


Figure 57. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3, V3, M4, V4 and M5.

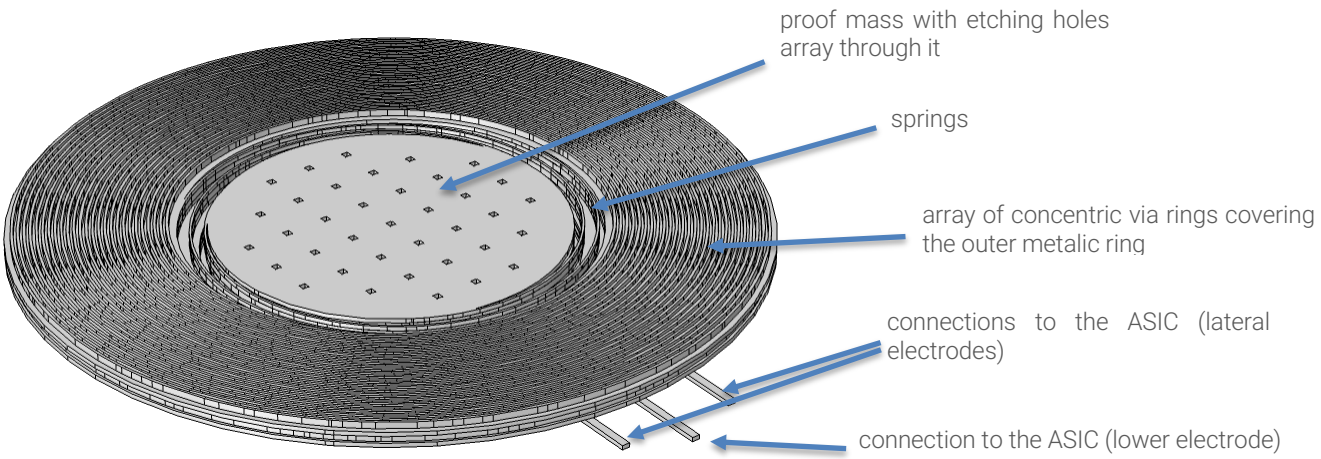


Figure 58. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3, V3, M4, V4, M5 and V5.

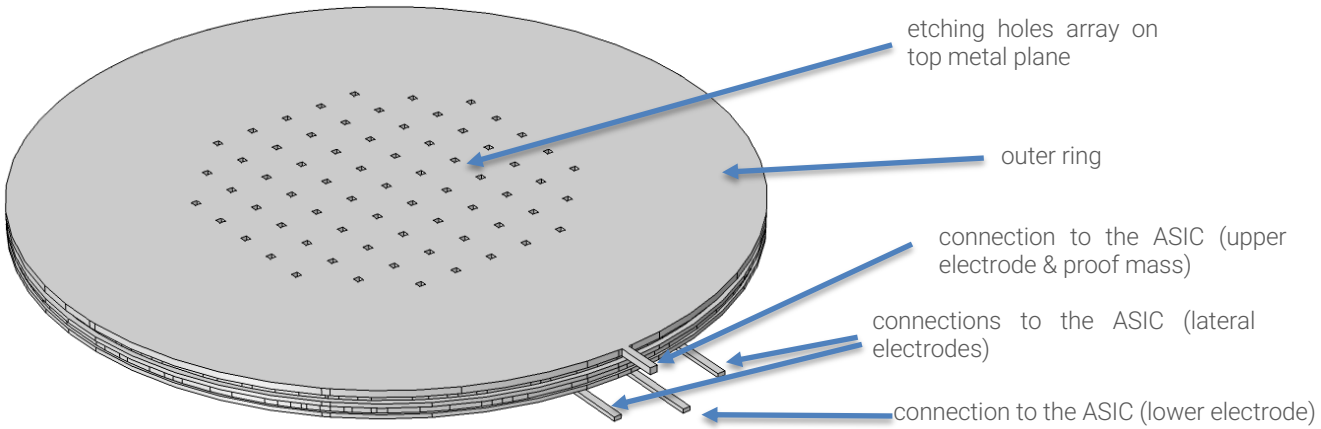


Figure 59. 3D CAD drawing of the metal layers without transparency. M1, M2, V2, M3, V3, M4, V4, M5, V5 and M6.

The following figure shows an exploded view of the 3D cad.

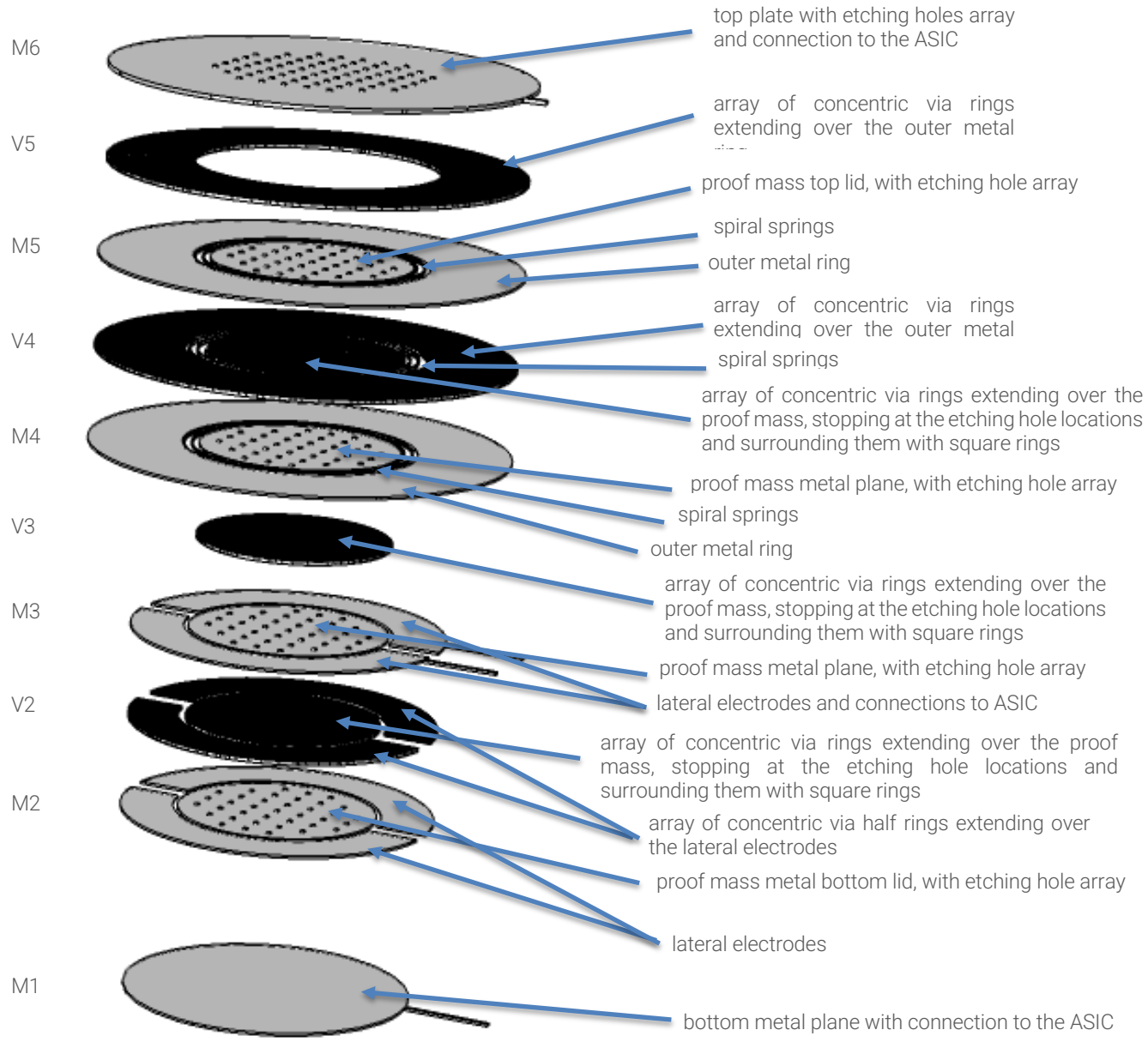


Figure 60. Exploding view of the 3D CAD of the metal and via layers only without showing SiO2, passivation and substrate.

The following figure shows how it could be layer M2 if we made 4 instead of 2 lateral electrodes, thus leading to X and Y differential sensing, in addition to single ended Z (out-of-plane):

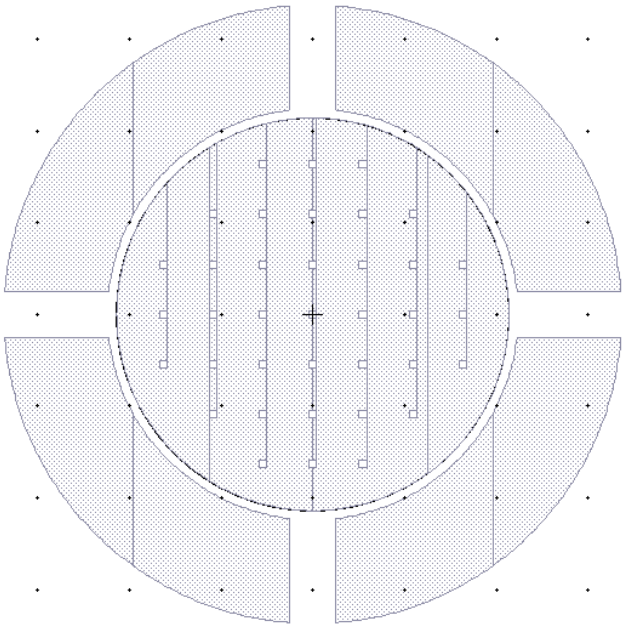


Figure 61. M2 metal layer for a 3D axis inertial sensor.

4. References

1. Ultra-low Cost MEMS Devices in Standard CMOS. SEMICON Europa 2012, 9-11 October- Dr. Josep Montanyà Silvestre
2. THE CMOS-MEMS 3-AXIS CAPACITIVE ACCELEROMETER TO MEET THE COMMERCIAL SPECIFICATIONS. IEEE MEMS 2016, Shanghai. Ming-Han Tsai, Pixar Imaging, National Tsing Hua University, Taiwan.
3. Implementation of a monolithic capacitive accelerometer in a wafer-level 0.18μm CMOS MEMS process. JOURNAL OF MICROMECHANICS AND MICROENGINEERING 2012. Sheng-Hsiang Tseng. National Tsing Hua University, Taiwan.
4. CMOS BEOL-embedded z-axis accelerometer. ELECTRONICS LETTERS. 28th May 2015. P. Michalik.
5. CMOS BEOL-embedded lateral accelerometer. 2015 IEEE. Piotr Michalik.
6. Experiments on the Release of CMOS-Micromachined Metal Layers. Journal of Sensors 2010. Daniel Fernández - UPC.
7. CMOS processing for MEMS devices and their applications. Josep Montanyà i Silvestre. Ph.D. Thesis – January 2012 – UPC