Capacitive Micromachined Ultrasonic Transducers for Gas Sensing

Mølgaard, Mathias J.G.

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Capacitive Micromachined Ultrasonic Transducers for Gas Sensing

Mathias J. G. Mølgaard
PhD Thesis June 2018

DTU Nanotech
Department of Micro- and Nanotechnology

DTU Nanotech
Department of Micro- and Nanotechnology

DTU

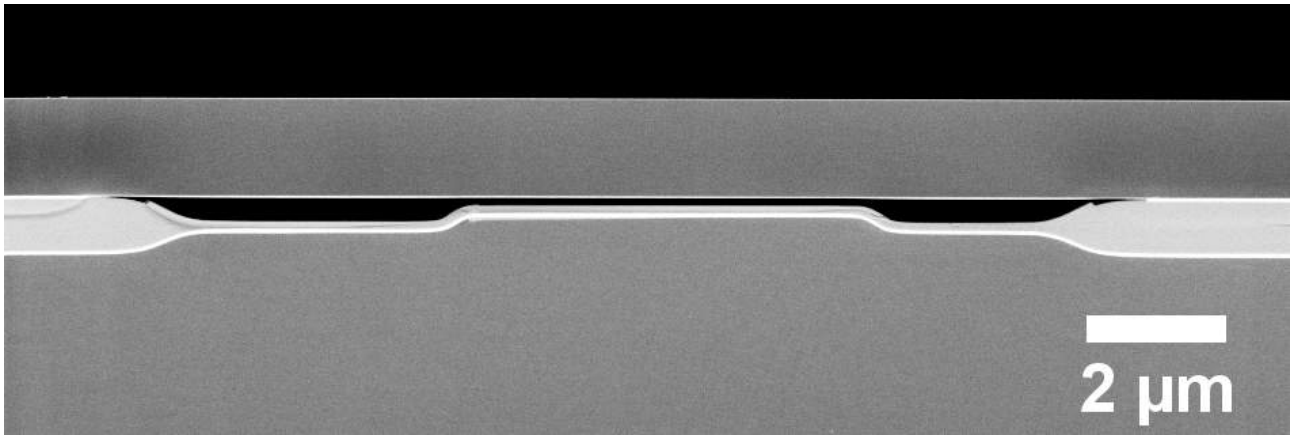
Technical University of Denmark

Ph.D. Thesis

Capacitive Micromachined Ultrasonic Transducers for Gas Sensing

Author:
Mathias J. G. Mølgaard

Supervisors:
Prof. Erik V. Thomsen
Assoc. Prof. Mogens H. Jakobsen



21st June 2018

Kgs. Lyngby, Denmark

Cover image: Scanning electron microscope image of the cross-section of a G1 CMUT cell before the plate has been thinned down.

Technical University of Denmark
Department of Micro- and Nanotechnology
Ørsteds Plads 345C
2800 Kgs. Lyngby
DENMARK
Tel: (+45) 4525 5700
Fax: (+45) 4588 7762
Web: www.nanotech.dtu.dk
E-mail: info@nanotech.dtu.dk
Author e-mail: migmol@nanotech.dtu.dk

Preface

This PhD thesis has been submitted to the Technical University of Denmark, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The work has been supervised by Professor Erik Vilain Thomsen and Associate professor Mogens Havsteen Jakobsen. This thesis consists of a recapitulation of the research work carried out from May 2015 to June 2018 at DTU Nanotech, Technical University of Denmark.

Capacitive Micromachined Ultrasonic Transducers (CMUTs) are currently one of the topics being researched in the group of Erik V. Thomsen but with a focus on the use of CMUTs as transducers for medical ultrasound imaging. Therefore, both practical knowledge of CMUT fabrication and characterization as well as a theoretical understanding of the CMUTs had already been established when my project began. This experience was the foundation and starting point of my project.

Mathias J. G. Mølgaard
Kgs. Lyngby, June 2018

Summary

Gas sensors are used in many fields for a wide range of applications from monitoring the indoor air quality control to detecting tiny amounts of illegal drugs. Different sensing methods exist, but in this thesis the focus is on gravimetric detection, that is measuring a perturbation (typically a resonance frequency shift) due to the addition of a mass on the sensor. Gravimetric sensors have a number of advantages over the other sensor types: small device footprint, low power consumption, and low cost. Capacitive micromachined ultrasonic transducer (CMUT) gas sensors are a subcategory of gravimetric gas sensors that offer a high sensitivity, low limit of detection per area, simple fabrication process, and an electrostatic actuation and detection scheme. Consequently, CMUTs are good candidates for being used as gas sensors. The aim of this thesis is to develop, fabricate, and apply state of the art CMUT gas sensors.

The theoretical background of CMUTs is given by presenting both a static and a dynamic model of the CMUT. Furthermore, design rules are developed for optimizing the mass sensitivity and limit of detection. Four generations of CMUT chips are designed, fabricated and characterized. The design space was investigated which resulted in an optimal choice of materials and geometry for CMUTs with small radii (radii $< 5\,\mu$m). The best CMUTs showed a record high distributed mass sensitivity of 0.83 Hz/ag and a record low limit of detection of 1.16 ag, which in part is due to an optimized design and fabrication process. The CMUTs were used for detection of a precursor molecule for the synthesis of (meth)amphetamine called Benzyl Methyl Ketone (BMK). This was done in conjunction with a colorimetric chip which showed a high selectivity towards BMK. Finally, an alternative method of reading out the resonance frequency from the CMUT was presented. The top electrode of the CMUT is split in three: one for actuating the CMUT and two for detecting the capacitance change caused by the plate deflection.

Hence, the actuating and detection electrode and signals are decoupled and common mode noise can be suppressed by making a differential measurement of the signals from the two sensing electrodes.

Resumé

Gassensorer er udbredte indenfor mange felter, hvor de bliver brugt til forskellige formål, fra at overvåge den indendørs luftkvalitet til at detektere små mængder ulovlige stoffer. Der eksisterer forskellige måder til at måle gas, men i denne afhandling har fokus været på den gravimetriske målemetode, hvor man detekterer en perturbation af sin sensor (typisk et resonansfrekvens skift) efter masse er blevet tilføjet sensoren. Gravimetriske sensorer har en række fordele i forhold til andre gassensor-typer: lille sensor størrelse, lavt energiforbrug og lav enhedspris. Kapacitive Mikrofremstillede Ultraslydstransducere (CMUTs) er en underkategori af gravimetriske sensorer og tilbyder en høj følsomhed, god opløsning pr. areal, en simpel fremstillingsproces og en elektrostatisk aktuerings-mekanisme. Af disse grunde er CMUTs gode kandidater til at blive brugt som gas sensorer. Målet for denne afhandling er at udvikle, fabrikere og anvende state of the art CMUTs til brug som gas sensorer.

Den teoretiske baggrund for CMUTs er givet ved en præsentation af både en statisk og en dynamisk CMUT model. Ydermere er der udviklet designregler for at optimere masse-følsomheden og opløsningen. Fire generationer af CMUT chips er blevet udviklet, fabrikeret og karakteriseret. Grænserne for designet er undersøgt, og der blev fundet et optimalt valg af materialer og geometri for CMUTs med små dimensioner (radier $< 5\,\mu$m). Disse CMUTs udviste en rekord høj distribueret masse-følsomhed: 0.83 Hz/ag og en rekord lav detektionsgrænseværdi: 1.16 ag, grundet et optimeret design og fabrikationsmetode. Sensorerne blev anvendt til at detektere et molekyle som bliver brugt til at syntetisere (met)amfetamin kaldet Benzyl Methyl Ketone (BMK). Dette blev målt sammen med en kolorimetrisk chip, som udviste en god selektivitet overfor BMK. Til sidst blev en ny metode til at udlæse resonansfrekvensen præsenteret. Topelektroden splittes i tre: en bliver brugt til at aktuere, og de to andre bliver brugt til at udlæse en kapacitans-ændring,

som er proportional med pladeudbøjningen. Herved afkobles aktuerings- og udlæsnings-signalet, og eventuel støj kan blive undertrykt ved at foretage en differentiel måling.

Acknowledgements

First of all I would like to thank my main supervisor Professor Erik V. Thomsen for his great support and enthusiasm throughout the project on both a professional and personal level. Furthermore, I would like to thank Associate Professor Mogens Havsteen Jakobsen and his group for providing me with the know-how and infrastructure for the functionalization of my devices. Especially, thank you to Milan Laustsen who have helped me both fabricate colorimetric chips and functionalize my CMUTs.

Moreover, thanks to my colleagues for our valuable academic discussions, as well as our not so academic discussions. I have enjoyed the informal work environment you have helped create. A special thanks goes out to Jan Bagge for his tireless effort of teaching us about electrical engineering. Furthermore, thanks to Ole Hansen for always being willing to help with everything from fabrication problems to theoretical calculations. I would like to thank the master and bachelor students I have supervised, for their hard work in the cleanroom fabricating CMUTs. On the same note, a thanks goes out to the staff at DTU Danchip for sharing their practical fabrication knowledge, as well as keeping the cleanroom running. Finally, thanks to my aunt for proofreading the manuscript.

Contents

CHAPTER 1

Introduction

Gas sensing is the act of detecting and possibly determining the concentration of a specific gas using a sensor. These sensors have an active element which is perturbed by the gas and an output signal is generated. The final output signal is typically electrical and can be logged by a computer.

The value of the global market for gas sensors is estimated to reach ∼2.5 billion USD by the year 2020 [1, 2]. Table 1.1 shows a list of well established fields in which gas sensors are used and examples of corresponding gases of interest. Commercial gas sensors are available for these use cases but one could imagine new markets such as: gas sensors embedded in consumer electronics such as smartphones [3], tablets or wearable devices and as replacements for the current mass spectrometers in high vacuum systems [4, 5]. Clearly, these vastly different application areas result in different requirements for the gas sensors in each specific case. For example, a gas sensor replacing a mass spectrometer in a vacuum system will require an extremely low limit of detection while the unit cost may not be important, whereas the opposite would be true for a sensor in a consumer product such as a smartphone.

In the following chapter different sensor types are compared. This comparison forms the basis for the choice of technology of the gas sensor to be made.

Table 1.1: Examples of fields in which gas sensors are utilized [6–10].

Field	**Gas examples**
Indoor air quality	CO_2, CO and humidity
Safety in industry	Volatile Organic Compounds (VOCs), NH_3 and O_2
Automotive	Emissions: CO_2 and NO_X
Defence	Explosives, nerve gas and biological weapons
Law enforcement	Alcohol and potentially drugs
Environmental studies	Greenhouse gases
Food quality and safety	Complex mixture

1.1 Gas sensor specifications

As previously mentioned: the application area of the sensor determines the sensor specifications. The most important specifications are defined in Table 1.2. For real non-ideal sensors there will be a trade-off between some of the specifications. Therefore, a good sensor will be one where as many of the specifications can be optimal at the same time.

Apart from the specifications listed in Table 1.2 it may be desirable to have a gas sensor capable of detecting multiple gases at the same time. The inherent ability to do this varies among the different sensor types. Furthermore, remote distributed sensors may become important in the future for creating sensor networks. This requires remote connectivity, low power consumption, and a small footprint of the sensor.

In this thesis the focus will mainly be on the first three specifications from Table 1.2, namely: sensitivity, limit of detection, and selectivity.

1.2 Gas sensor types

A wide range of gas sensor types exists, some being commercially available, whereas others still are at the research stage. The common aspect shared by all gas sensors is that a perturbation of the sensing element occurs due to the surrounding gas. The actuation and detection scheme may be coupled or decoupled influencing the complexity of the sensor. Figure 1.1 gives an overview of detection techniques and examples of specific sensor types. In the following the focus is on gravimetric sensors as the devices treated in this thesis fall within this category. Table 1.2 showed a list of specifications which now will be discussed in the context of gravimetric sensors. Although some gravimetric gas sensors have reached the commercial market [11, 14], gravimetric sensors are primarily found in research projects in

Table 1.2: Qualitative definitions of gas sensor specifications. Based on [11–13].

Specification name	**Definition**	**Ideal sensor**
Sensitivity	Slope of the output vs. input curve	High
Limit of detection	Minimum detectable input	Low
Selectivity	Ability to distinguish between gases	High
Size	Physical dimensions	Small
Power consumption	Energy consumed per time unit	Low
Price	Unit cost	Low
Response time	Time before a stable output is reached	Short
Range	The interval of measurable input values	Wide
Linear range	The range in which the output is linear	Wide
Reversibility	Is the initial state reached again?	Yes
Stability	Does the characteristics change over time?	No

academia at the research and development stage. Gravimetric gas sensors have a number of inherent favorable traits: a small device size, low power consumption, low cost (batch cleanroom fabrication), possibly Complementary Metal–Oxide–Semiconductor (CMOS) compatible fabrication process which means that the sensor and the driving electronics can be integrated into a small and low power unit [3]. Furthermore, excellent volume sensitivities have been demonstrated, e.g. for dimethyl methylphosphonate (DMMP) of 34.5 pptv/Hz with a limit of detection of 50 pptv [3]. This combination of traits is unique among the sensor types listed in Figure 1.1 which makes gravimetric sensors an interesting choice as gas sensors.

A drawback of the gravimetric method is that, without modification of the surface, the sensors cannot distinguish between gasses, i.e. they have no selectivity. Therefore, gravimetric gas sensors commonly have a functionalization layer applied to the sensing element which selectively binds, adsorbs or absorbs the specific analyte gas molecules. However, these layers increase the mass of the sensor itself, hereby decreasing the sensitivity and affects the reversibility and response time of the sensor.

1.3 Resonant gravimetric sensing

A resonant gravimetric sensor is a mechanical resonator with a resonance frequency, f_0, that is determined by the geometry and materials of the device. When the resonator is loaded with a mass, the resonance frequency will decrease. For a resonant gravimetric gas sensor it is the gas molecules which adds the mass to the resonator, as shown in Figure 1.2. To the left in

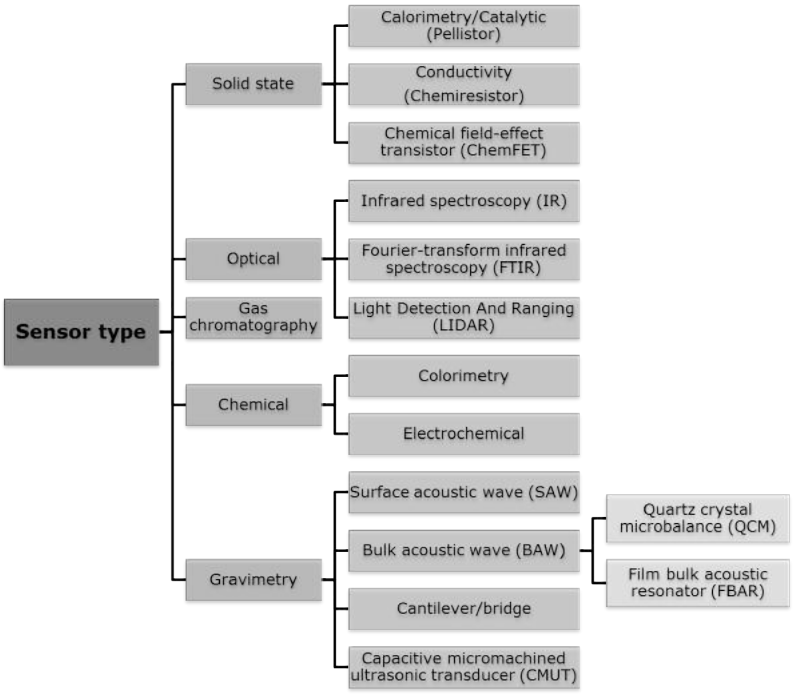


Figure 1.1: Tree diagram of different gas sensor types. The sensor types are grouped according to their detection technique. Based on [6–8, 11, 15].

the figure the gas is not in contact with the resonator but when the gas is introduced, in the figure to the right, some of the molecules will load the resonating element, thus decreasing the resonance frequency to f_1, such that $f_0 > f_1$. This shift in resonance frequency can be recorded and correlated with the concentration of the gas and consequently the sensor can give a quantitative value for the gas concentration.

The *gravimetric sensors* in Table 1.1 are all either MicroElectroMechanical Systems (MEMS) or NanoElectroMechanical Systems (NEMS) devices. Generally, as the dimensions of a device shrink the surface-area-to-volume ratio increases as it is proportional to the inverse of the characteristic device dimension. As will be evident later this relationship favors miniaturization of the devices in order to increase e.g. the sensitivity.

Various actuation and read-out principles have been used in gravimetric devices. However, common for almost all sensors is that the readout signal in the end is converted to an electric signal. A lot of exotic transduction techniques have been published (see [16] p.115) which not will be mentioned here. The most widely used actuation and detection schemes used in gravimetric sensors are [16] p.115-143:

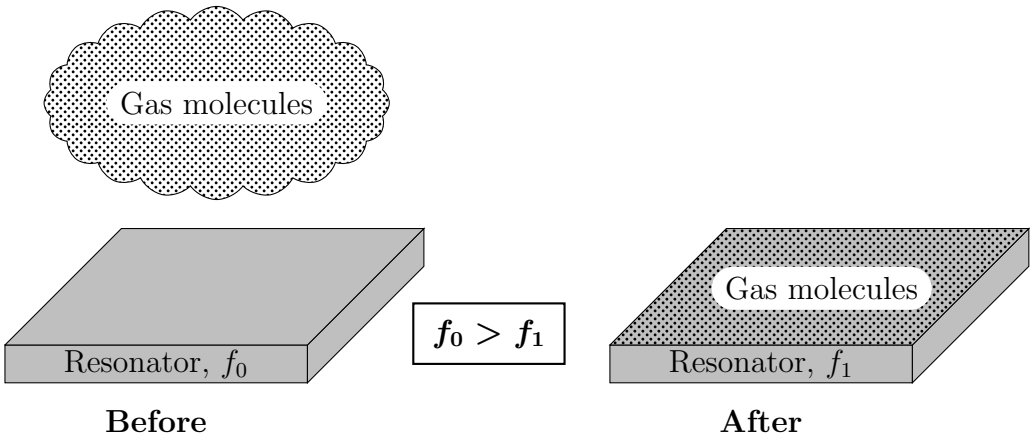


Figure 1.2: Sketch illustrating the principle behind gravimetric gas sensing. Left: resonator before exposure to a gas and right: resonator after exposure to a gas with the gas adsorbed on the surface. The resonance frequency of the resonator is seen to decrease after exposure to the gas.

Actuation: Electrostatic, Electrodynamic, Piezoelectric and Thermoelastic.

Detection: Electrostatic, Electrodynamic, Piezoelectric, Piezoresistive and Optic.

The electrostatic technique relies on having quasistatic charges on two electrodes typically separated by a small gap, where one of the electrodes is free to move. When a potential is applied between the two electrodes the charges experience an attractive electrostatic force. The electrostatic technique is widespread among MEMS transducers due to a low power consumption and since only simple external circuitry is needed for the actuation and detection. However, for NEMS devices the electrode areas are typically very small resulting in small signals which are difficult to measure.

The electrodynamic technique relies on the Lorentz force and thus requires *moving* charges in a magnetic field for the charges and hereby device to experience a force. The magnetic field may come from permanent magnets or electromagnets and the moving charges are typically supplied by running a current through a conductor in the device. Due to this current the power consumption is typically higher than for the electrostatic technique.

The piezoelectric technique relies on the fact that for piezoelectric materials the strain and electric field are coupled. Consequently, these materials can be used for both actuation and detection. The quality of the piezoelectric material determines the coupling between the mechanical and electrical domain and one of the challenges can be ensuring a high material quality.

In the gravimetric sensor literature much attention is given to three of the

specifications from Table 1.2, namely: sensitivity, limit of detection (LOD) and selectivity. In the following the gravimetric sensor types will be compared based on these three specifications.

1.3.1 Mass sensitivity

In Table 1.2 the distributed mass sensitivity was defined as: the slope of the output vs. input curve of the sensor. The sensitivity can be calculated for resonators modeled as single degree of freedom (1D) linear harmonic oscillators [11]:

$$S = \frac{\partial f}{\partial m} = \frac{\partial}{\partial m}\sqrt{\frac{k_{\text{eff}}}{m_{\text{eff}}}} = -\frac{1}{2}\frac{f_0}{m_{\text{eff}}}, \tag{1.1}$$

where k_{eff} is the effective spring constant, f_0 is the resonance frequency and m_{eff} is the effective mass of the resonator. Equation 1.1 is valid under the assumption that the mass added to the resonator is much smaller than the mass of the resonator: $\partial m \ll m_{\text{eff}}$. The equation shows that in order to increase the sensitivity a high resonance frequency and a low resonator mass is wanted. Which shows why MEMS and NEMS devices are often used.

To be able to better compare the sensitivities of the different sensor types a normalized sensitivity is used [17]:

$$S_{\text{norm}} = \lim_{\partial m \to 0} \frac{1}{f_0}\frac{\partial f}{\partial m/A} = \frac{\partial f}{\partial m}\frac{A}{f_0} = S\frac{A}{f_0}, \tag{1.2}$$

where A is the area of the resonator. The sensitivity is here normalized with the area and resonance frequency of the resonator. Since the operating principles differ between the gravimetric sensor types, the normalized sensitivity will have different dependencies on the geometrical parameters. Hence, the normalized sensitivities are expected to also differ between the sensor types. Figure 1.3 shows the highest normalized sensitivities I have found in the literature for each sensor type. A general trend is observed where the NEMS bridges and cantilever have the highest normalized sensitivities due to their small dimensions. The second highest values are found for the CMUTs where **G2-G4** are the CMUT generations designed and fabricated during this project. Finally, the FBAR, SAW and QCM devices have the lowest normalized sensitivities. The reason for this is elaborated more upon later.

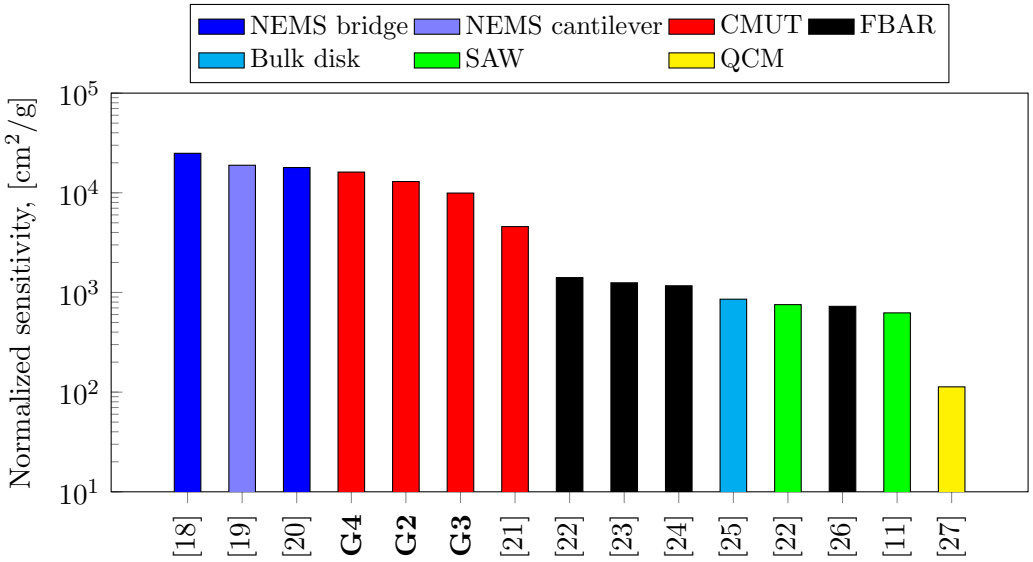


Figure 1.3: Normalized distributed mass sensitivities for different types of state of the art sensors. The CMUTs made in this project are labeled G2-G4 and for these the normalized sensitivity has been calculated using the measured (as opposed to the designed) geometrical values.

1.3.2 Limit of detection

The mass limit of detection was in Table 1.2 defined as: the minimum detectable input. The mass LOD can be defined more precisely as the minimum mass change which gives a detectable frequency shift distinguishable from the frequency noise in the system. Mathematically this minimum LOD can be written as [11]:

$$\mathrm{LOD} = 2m_{\mathrm{eff}}\sigma_{\mathrm{min}}, \tag{1.3}$$

where σ_{min} is the minimum Allan deviation, which is a measure for the frequency noise in a system. In order to achieve the lowest LOD a small resonator mass should be designed with a minimum of frequency noise which will reduce the Allan deviation, since a high frequency stability results in a low Allan deviation.

It can be shown that the *gas concentration* limit of detection is given by [11]:

$$\partial C \propto \frac{\mathrm{LOD}}{A} = \frac{2m_{\mathrm{eff}}\sigma}{A}, \tag{1.4}$$

where A is the surface area of the resonator. Equation 1.4 shows that the ratio $\frac{\mathrm{LOD}}{A}$ should be minimized for the best sensor performance and that this is accomplished by decreasing the LOD and increasing the surface area of the

Table 1.3: Resonance frequency, Allan deviation, surface area, and LOD/A for different sensor types. All devices are under atmospheric pressure. The values in the table are taken from [11].

Sensor type	f_r **[MHz]**	**Allan deviation,** σ_{min}	**Active surface area,** A **[mm²]**	**LOD/**A **[ng/cm²]**
BAW	1100	10^{-7}	0.25	1
SAW	500	10^{-8}	4	0.03
CMUT	8	10^{-8}	0.25	0.01
CMUT	47.7	3×10^{-8}	0.09	0.003
MEMS	3	10^{-7}	7.3×10^{-6}	0.1
NEMS	127	10^{-6}	8×10^{-8}	0.1

device. Table 1.3 shows an overview of typical values of the resonance frequency, Allan deviation, surface area and $\frac{\text{LOD}}{A}$ for different gravimetric sensor types. The values in the table are obtained at atmospheric pressure, which is important to note since many NEMS devices are operated in a vacuum which influences the Allan deviation. The devices with the smallest dimensions display the highest Allan deviations and are thus the least frequency stable. This can be seen by comparing the NEMS and MEMS devices with the other sensor types.

The lowest $\frac{\text{LOD}}{A}$ ratio is found for the CMUT due to a low LOD and relatively large surface area. The LOD is itself low due to a low Allan deviation and relatively low mass compared with the other sensors. The CMUT device has a combination of parameters which places it in the middle ground between the BAW and SAW on one side and MEMS and NEMS devices on the other, thus bringing together the best of both worlds. The CMUT maintains a surface area roughly as big as the BAW and SAW devices while having a much lower effective mass than these devices. This is due to the inherent parallelism of the CMUT structure which will be discussed more later. Equation 1.4 shows that the mass per area should be minimized by decreasing the amount of mass used per surface area. As will be shown later this is exactly what the CMUT structure allows for.

The lowest mass LOD is found for the NEMS devices due to the small size and hereby low mass, see Equation 1.3, but the surface area is for the same reason very small compared with the other devices. Also, the Allan deviation is approximately a factor 100 higher than for the CMUT. Therefore, the $\frac{\text{LOD}}{A}$ does not ends up being the lowest.

1.3.3 Selectivity

Resonant gravimetric sensors measure a response to a change of mass and they do therefore not exhibit any selectivity between gasses. In order to

make the sensors selective they are typically coated with a thin layer having an affinity towards the gas one is trying to detect. This thin layer is called the functionalization layer which enables the possibility of selective and label-free sensing.

The functionalization layer can be applied using different techniques. Sometimes the choice of technique is limited due to the geometry of the sensor. For example, it could be important not to coat the underside of a NEMS cantilever or bridge. It is generally easier to coat the sensors with a 'closed' surface such as BAWs, SAWs and CMUTs. Here several coating techniques can be used such as: spin coating, drop coating, dip coating, and spray coating.

1.3.4 Choice of sensor type

The choice of sensor technology was first narrowed down from the list of all potential gas sensing techniques shown in Figure 1.1, to only focus on gravmetric gas sensors. Within this subcategory the different types have been compared and a choice of technology can now be made.

Sensitivity. The normalized sensitivities for different resonant gravimetric sensors were compared in Section 1.3.1. The highest normalized sensitivities were found for the NEMS sensors followed by the CMUT sensors. The lowest normalized sensitivities were seen among the BAW and SAW sensors.
Limit of detection. It was shown in Section 1.3.2 that the important figure of merit (FOM) for the gas concentration LOD is the ratio $\frac{\text{LOD}}{A}$, which should be minimized. This ratio was lowest for the CMUT sensor, due to a favorable combination of Allan deviation, plate mass, and surface area compared with the other sensor types.
Selectivity. The point regarding selectivity came down to the ease of functionalization which is shared between the sensor types with a closed surface. Furthermore, these structures are also more robust towards the external environment.

Based on these criteria the NEMS and CMUT devices seem most promising but there are other points which should be considered:
Fabrication. The fabrication of CMUTs, even for high sensitivity sensors, can be done using standard cleanroom planar processing and UV lithography. The latter is not always possible for the fabrication of NEMS structures where lithography techniques with a higher resolution are needed, which in some cases (e.g. for electron beam lithography) are serial processes.
Actuation and detection. The actuation scheme in CMUTs is based on the electrostatic force while the detection is based on a capacitance change.

As stated earlier this allows for a low power consumption with an easy integration to the electronics, which can either be discrete or integrated [28]. Many different actuation and detection schemes have been used for NEMS sensors but many uses a variation of electrodynamic actuation or readout which requires an external magnetic field which in turn increases the complexity and size of the sensor.

The CMUT structure is inherently parallel where a single device is made up of many small resonators which in part explains the excellent noise performance but can also be used to tailor the electrical impedance seen by the external electronics. Furthermore, making several independent CMUT sensors on the same chip is straightforward which e.g. has been exploited for the detection of multiple gasses at the same time [29].

To conclude: for the reasons mentioned above the CMUT is seen as the best choice of technology for gas sensing. Therefore, the devices used in this thesis are CMUTs.

1.4 The CMUT

A capacitive micromachined ultrasonic transducer is an electromechanical resonator. The CMUT device can transform energy from the electrical domain to the mechanical domain to the acoustical domain and vice versa. In this thesis the focus will mainly be on the transduction between the mechanical domain and electrical domain. In this section the basic working principle of the CMUT is described, while Chapter 2 provides a more detailed theoretical description of the CMUT.

The CMUT has been used as a transducer in several areas. Much attention has been given to the use of CMUTs as transducers for medical ultrasound imaging [30–33]. This is a field in which both academia and industry have devoted much time and many resources to further develop the CMUT to suit the the requirements. The transducers currently in use typically utilize piezoelectric crystals as the oscillating part and compared with these the CMUT potentially offers a larger bandwidth, improved pressure sensitivity, increased design flexibility, reduced cost and a CMOS compatible fabrication process.

Airborne applications such as flow metering and range sensors has also been proposed as an area in which CMUTs could be used [34, 35]. CMUTs have been applied as transducers for ultrasonic non-destructive testing (NDT) [36–39]. NDT is e.g. used for thickness measurements in pipes, weld inspection or for detecting fatigue induced cracks.

Finally, CMUTs have, as already mentioned, successfully been used for

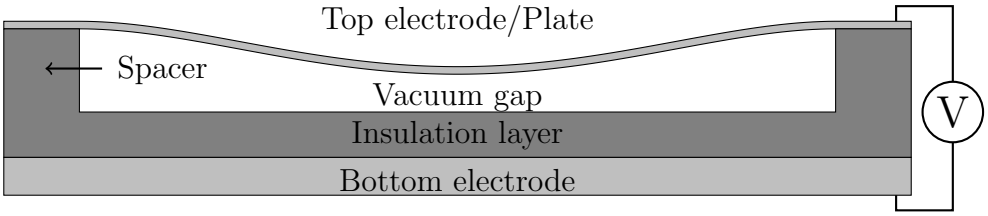


Figure 1.4: Cross-sectional sketch of a single CMUT cell. The plate is deflected into the vacuum gap due to an applied bias voltage between the top and bottom electrodes.

both immersion chemical/biological sensing and gas sensing [28, 40, 41]. A more thorough literature review of CMUT gas sensors is given in Section 1.4.3.

1.4.1 Working principle

Figure 1.4 shows a sketch of a cross-sectional view of a single CMUT cell, whereas a CMUT element typically consists of many cells electrically connected in parallel. Figure 1.4 highlights the main parts of the CMUT cell. The plate is the oscillating part of the CMUT and deflects due to an applied force. For some designs the plate itself also acts as the top electrode, which requires the plate material to be an electrical conductor. The geometry of the plate, as seen from above, can be chosen arbitrarily but circles [42], squares [43] and hexagons [44] are most commonly used. The spacer material separates the top electrode from the underlying layers and creates the vacuum gap. The insulating layer prevents the plate/top electrode from touching the bottom electrode, hereby creating a short circuit in the capacitor.

The CMUT is actuated by applying an electrical potential between the top electrode and bottom electrode. The difference in potential gives rise to electrical charges of opposite sign on the two electrodes. This separation of charge results in an attractive electrostatic/Coulomb force between the electrodes and consequently, the top electrode/plate is deflected down into the gap. The position at which the plate stops is determined by a equilibrium between the electrostatic force and the mechanical force stemming from the rigidity of the plate. Superimposing an AC voltage on the already applied DC voltage will make the plate oscillate with the frequency of the AC voltage. If the frequency of the AC voltage matches the mechanical resonance frequency of the plate, the deflection amplitude is maximized. The capacitance of the CMUT is affected by these oscillations of the plate since the distance between the two electrodes is varying as a function of time. Consequently,

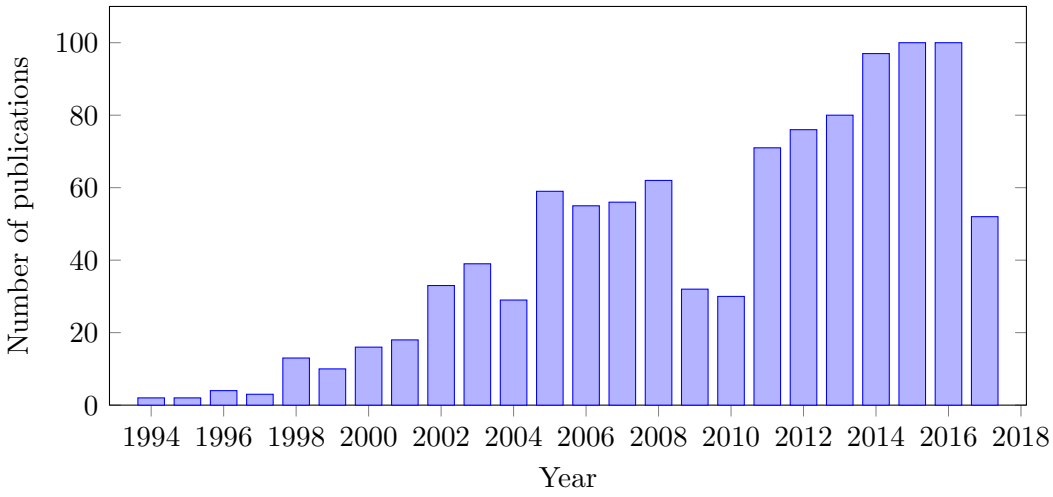


Figure 1.5: Number of publications as a function of publishing year for the entire CMUT field. The data for the chart is obtained from the database search site www.webofknowledge.com and includes publications in conference proceedings.

the capacitance change is at a maximum at resonance.

1.4.2 Brief historical account

This section presents an overview of the CMUT literature and the most important milestones, relevant for gravimetric sensing, in the development of the CMUT. The first CMUT was reported in a publication from 1994 by Matthew I. Haller and Butrus T. Khuri-Yakub [45]. Butrus T. Khuri-Yakub heads a research group at Stanford University and is a pioneer within the CMUT field. Since this first publication, the CMUT field has grown which is reflected in the increasing number of articles published each year, see Figure 1.5. This figure shows that the number of CMUT related articles published as a function of time is increasing.

The first CMUTs were fabricated using standard micromachining cleanroom processes and a sacrificial release method [46]. In the sacrificial release method the cavity is defined by a sacrificial layer which is selectively etched after a plate layer has been deposited on top. Figure 1.6 a) shows a cross-section of a CMUT cell fabricated using the sacrificial release method. In this example typical materials have been used such as Si for the bottom electrode, a Si_3N_4 plate and an Al top electrode. A new fabrication method was developed by the Stanford group in 2003 [47], where the cavities are etched out in a dielectric layer (e.g. SiO_2) and subsequently the plate is wafer bonded to this substrate wafer hereby encapsulating the cavities. The result of this wafer bonding fabrication technique is shown in Figure 1.6 b).

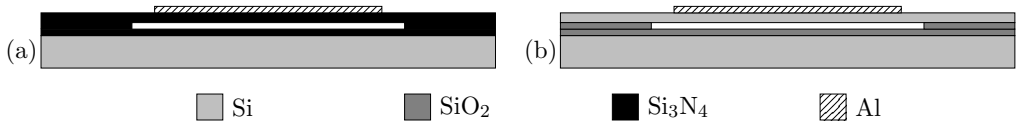


Figure 1.6: Cross-sectional sketches of single CMUT cells fabricated with the a) sacrificial release method and b) wafer bonding fabrication process. The choice of material is chosen to match what is commonly used.

Here the black line between the two SiO_2 layers denotes the bonding interface. Silicon on insulator (SOI) wafers have been used extensively to define the plate [33, 48, 49] due to the possibility of a low thickness variation and well controlled mechanical properties of the device layer [46], which is the layer comprising the plate.

The local oxidation of silicon (LOCOS) fabrication process was first published in [50] also by the Stanford Group. A journal paper was published by the same group in 2011 [42] which goes more into detail with the fabrication process and modeling. Figure 1.7 a) shows a sketch of a cross-section of a finished CMUT cell fabricated by the LOCOS process. The spacer between the top and bottom electrodes is called the post oxide which can be made thicker than the spacers in the processes shown in Figure 1.6 for the same gap height. This decreases the parasitic capacitance which improves the device performance [42]. Furthermore, the control and uniformity over a wafer of the vacuum gap thickness is approximately a couple of nanometer which enables the fabrication of small ($< 50\,\text{nm}$) vacuum gaps. In the LOCOS process the silicon substrate wafer is selectively oxidized to first form the Si bumps in the center of the cavities and then the oxide posts forming the cavities. The structure in Figure 1.7 b) is a CMUT fabricated by only using a single LOCOS process step forming cavities but not the Si bumps. Hence, the structure in 1.7 a) will here be called a double LOCOS structure, while the structure in 1.7 b) will here be called a single LOCOS structure. Both structures have been fabricated in this project along with the structure in Figure 1.6 b). The fabrication processes of these structures all rely on wafer bonding. It should be noted that the plate material for the wafer bonded structures also can be other materials such as Si_3N_4 as demonstrated in [51–54].

1.4.3 State of the art of CMUT sensing

In this section the state of the art for CMUT sensing is presented. The emphasis will be on CMUTs used for gas sensing but a short overview of CMUTs used as bio-sensors is presented in the end.

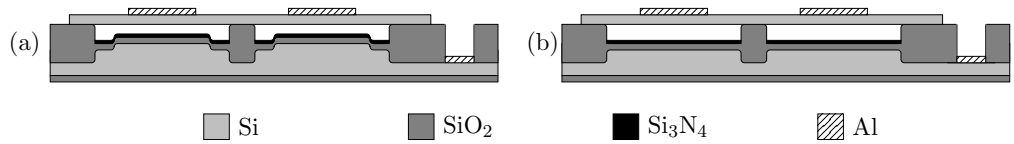


Figure 1.7: Cross-sectional sketches of single CMUT cells fabricated with the a) double LOCOS process b) single LOCOS fabrication process. Notice the central Si protrusions, called bumps, in the cavities in a) and the lack hereof in b).

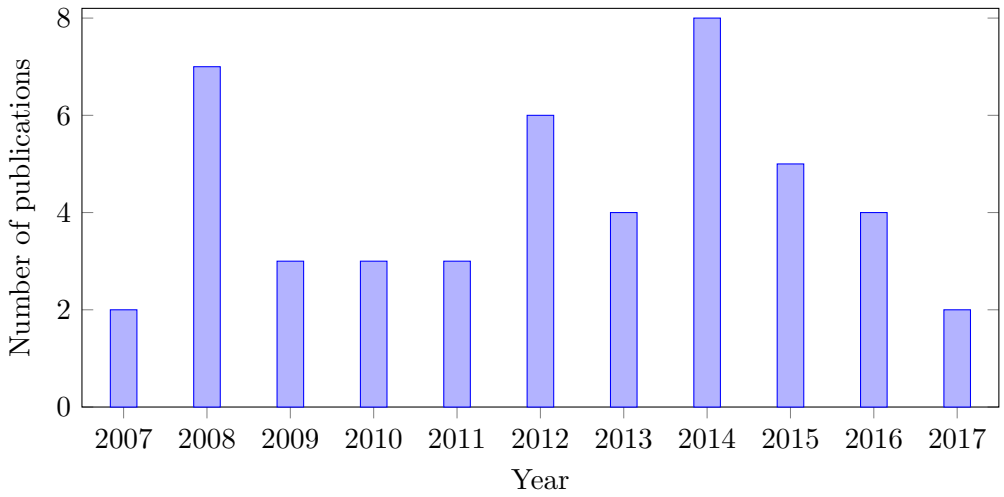


Figure 1.8: Number of publications as a function of time for CMUTs used as gas sensors and biological sensors, including papers in conference proceedings.

Figure 1.8 shows the number of publications as a function of time for CMUTs used as gas sensors or biological sensors. Compared with the plot for the entire CMUT field (Figure 1.5), the number of publications is rather constant with much fewer publications each year. The first articles were published in 2007 by the Stanford group [40, 55]. The same group has published almost half of the articles regarding bio-sensing and gas sensing, see Figure 1.9. Most of these articles are about gas sensing, whereas the publications from Kaunas University of Technology are mainly about biological sensing using CMUTs.

Gas sensing

The first articles concerning the use of CMUTs for gas sensing were published by the Stanford group. In 2007 Park et al. [40] described the use of a 6 MHz CMUT array, originally designed for medical ultrasonic imaging, for gas sensing. Specifically, four analytes were detected: water, isopropanol, acetone and methanol. The CMUT elements in the array were drop coated

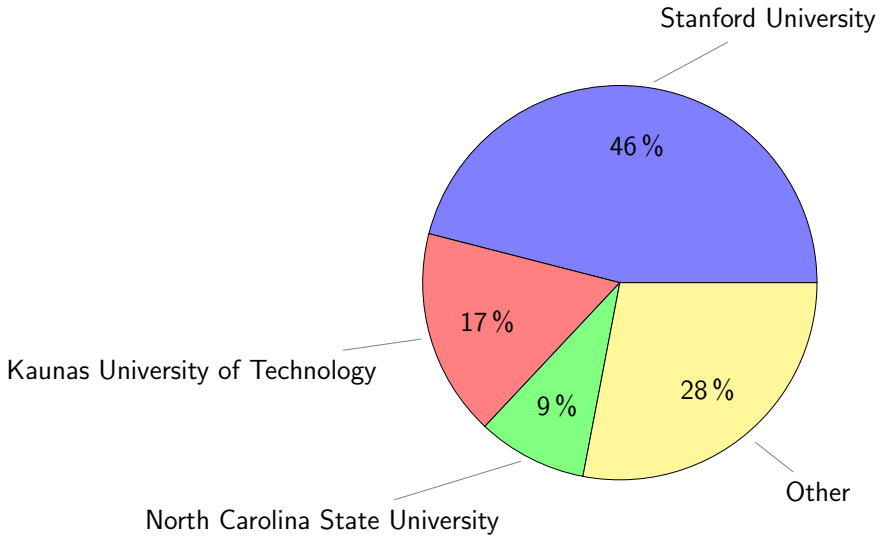


Figure 1.9: Publication distribution by institution. Note that almost half of the publications are published by an author in the Stanford group.

with three different polymers showing different sensitivities towards each of the analytes. This technique has in subsequent publications been employed in order to be able to distinguish between analytes.

Drop coating has been used to a great extent for the functionalization of CMUTs, e.g. [56] [29]. Drop coating consists of dissolving the functionalization layer material in a solvent. Droplets of this liquid solution is hereafter dispensed unto the CMUT element and when the solvent evaporates the functionalization layer remains on the surface. This technique is an inherently serial process but the deposition can be performed by a dedicated machine which can be programmed to functionalize entire wafers at a time at a fast pace. Li et al. [57, 58] demonstrated a novel 2-step coating technique where first a layer of nanoparticles is deposited to form a uniform substrate layer for the second chemical sensitive layer. The nanoparticle layer serves multiple purposes. First, it localizes the second layer to the spot where it is deposited. This is useful in the case where several different layers are to be deposited close to each other. Secondly, it ensures a more uniform layer which minimizes the decrease of the quality factor. Finally, the effective surface area is increased due to the spherical shape of the nanoparticles, at the expense of an increased mass of the plate, due to the weight of the particles.

Several articles have been published by the Stanford group regarding the detection of dimethyl methylphosphonate (DMMP) which is a common simulant for sarin gas. Such a gas sensor should have a gas concentration LOD

of in the order of parts per billion (ppb) or better [3]. In 2008 Lee et al. and Park et al. [59,60] functionalized CMUTs with a resonance frequency of 18MHz with a polyisobutylene (PIB) layer for the detection of DMMP. The mass sensitivity per area was found to be $130\,\frac{\mathrm{zg}}{\mathrm{Hz}\cdot\mu\mathrm{m}^2}$. The DMMP volume concentration sensitivity and limit of detection was found to be 37 ppbv/Hz and 21 ppbv, respectively. These already impressive values were further improved upon with a CMUT device with a resonance frequency of 47.7 MHz and a new functionalization layer [3, 61]. The layer was a polymer abbreviated DKAP, which is a poly(dimethylsiloxane) derivative. The mass sensitivity of the CMUT was $48.8\,\frac{\mathrm{zg}}{\mathrm{Hz}\cdot\mu\mathrm{m}^2}$, which is an improvement by a factor 2.7 compared with the previous device. The DMMP volume sensitivity was measured to be 34.5 pptv/Hz [3] which is about an order of magnitude better than the volume sensitivity reported previously in [59,60]. The same order of improvement is seen for the volume concentration LOD which was 50.5 pptv with gas flow [3] and 16.8 pptv without gas flow [61]. Both the volume sensitivity and gas concentration LOD were, at the time of publication, the best reported for a gravimetric DMMP gas sensor.

The detection of CO_2 using CMUTs has also been studied. Lee et al. [62] used mesoporous silica layers on top of CMUT plates in order to increase the volume sensitivity towards CO_2. The mesoporous layer increases the surface area of the sensor, thus more molecules can load the plate per area and a larger response/better sensitivity is expected. By using such a mesoporous layer with a (3-Aminopropyl)triethoxysilane (APTES) functionalization layer on top a volume sensitivity of 1.6 ppm/Hz and a volume LOD of 1.82 ppm was obtained. The volume sensitivity was 38 times higher than that of a mesoporous CMUT device with no APTES functionalization. In an article from the same year by the same first author [63] different functionalization layers were compared. The polymer guanidine showed the best volume sensitivity towards CO_2: 1.06 ppm/Hz. In 2015 and 2016 a conference and a journal paper were published by Barauskas et al. concerning the detection of greenhouse gasses, including CO_2 [41, 64]. A volume sensitivity of 0.25 ppm/Hz is obtained with a polymer functionalization layer made of polyethylenimine (PEI). This is the best volume sensitivity for CO_2 reported for CMUTs. However, the volume LOD was 330 ppm which is much higher compared with the volume LOD of 1.82 ppm reported by Lee et al.

One of the main challenges of any gravimetric sensor is how to obtain selectivity. Functionalization layers with a chemistry specific tailored to a target analyte can for gas mixtures with few confounding gasses be enough to detect the target analyte gas. However, when more complex gas mixtures, which can be encountered during operation outside of a laboratory,

are faced, other solutions to the selectivity challenge must be employed. One solution has been to use CMUT arrays with multiple separate elements that each can have a different functionalization layer [21, 29, 49, 65, 66]. The responses of the elements can then be analyzed e.g. using principal component analysis (PCA) [29]. Another approach to increasing the selectivity has been to combine the CMUT with another sensor type with a better selectivity. This approach was investigated by Mølgaard et al. [53, 56], where a colorimetric sensor was combined with a CMUT for detection of benzyl methyl ketone (BMK) which is a precursor molecule for the synthesis of (meth)amphetamine. The colorimetric sensor showed an good selectivity towards BMK, while the CMUT sensor can give a quantitative value for the concentration when calibrated.

Biosensing

A short review of the publications regarding the use of CMUTs as biological sensors is given here for completeness. Since CMUTs are flexural resonators with out of plane deflections, the use of them as immersion sensors is not an obvious choice since they will be damped more than sensors using bulk or surface/interface waves such as BAWs and SAWs [28]. However, CMUTs generally exhibit a higher mass sensitivity than e.g. FBARs and SAWs [28], also see Figure 1.3. Moreover, the previously mentioned advantages of the CMUT over the other gravimetric sensors still apply. Particularly, the vacuum filled cavity is a great advantage in immersion since the plate is only exposed and hereby dampened by the medium on one side, unlike the situation for cantilevers or double clamped bridges.

The CMUT has been used as an immunosensor for the detection of various antigens or antibodies (depending on what part is immobilized on the CMUT surface). The immune system of the human body makes antibodies that can react and partake in a very specific/selective bond with an antigen. Clearly, this selectivity is desired as it will translate to a selective immunosensor overcoming one of the major challenges of gravimetric sensors.

The first peer-reviewed paper of a CMUT used as a biosensor was published by Ramanaviciene et al. [67]. In this paper they demonstrate the use of a CMUT as an immunosensor by immobilizing bovine leukemia virus protein gp51 on the CMUT and detecting anti-gp51 in a solution. Both the resonance frequency and real part of the impedance was seen to change as a function of the concentration of anti-gp51. The same antigen antibody pair was used in subsequent publications [28, 68, 69], where the effect of the added mass and possible stress in the biofunctionalization layer was investigated. It was found that if sufficient stress is present in this layer the resonance

frequency can increase after the layer is deposited.

Onen et al. [70] studied the use of a CMUT for detection of ovarian cancer, more specifically, anti-apoptotic protein Bcl-2 was to be detected. The device consisted of two sets of inter-digital CMUT arrays for excitation and detection of Lamb waves. In between these two arrays was a delay line coated with the functionalization layer. The added mass after the antigen antibody reaction then causes a measurable decrease in resonance frequency. The device resembles a SAW sensor utilizing Lamb waves. By relaying on Lamb or surface waves in immersion a lower damping should in general be observed. CMUT sensors for other types of immersion measurements have previously used a similar detection principle [71–73].

1.5 Thesis aim

The overall aim of this thesis is to develop and apply state of the art gas sensors. This is to be achieved by using CMUTs as a technology platform from which different sensor designs can be realized. The CMUTs should be modeled, designed, and fabricated so a high mass sensitivity and low limit of detection is obtained. Furthermore, alternative ways of increasing the selectivity are to be explored. Finally, the use of the CMUTs as gas sensors should be demonstrated by applying them as such.

1.6 Academic contributions

The project resulted in two journal papers and one conference proceeding paper, see Appendix A, B, and C, respectively. Furthermore, two talks were given at the conferences International Workshop on Micromachined Ultrasonic Transducers (MUT) 2016 in Rome and MUT 2017 in Glasgow, respectively. In addition, a poster presentation was given at IEEE International Ultrasonics Symposium (IUS) 2017 in Washington D.C.

1.7 Thesis outline

The outline of the thesis is given below to provide the reader with an overview of the thesis.

Chapter 1 - Introduction: This chapter qualifies the choice of CMUTs for use as gas sensors by comparing their performance to other gas sensor types. Further, the history and state of the art of CMUTs used for sensing

is presented which provides the reader with a background knowledge of prior work in the field.

Chapter 2 - Theory: The theory chapter describes the static and dynamic behavior of the CMUT, leading to an expression of the electrical impedance which can be used to evaluate later impedance measurements. Finally, expressions for the mass sensitivity and mass limit of detection is given.

Chapter 3 - Design: The design of the four CMUT generations, G1-G4, made during this project is presented. The impact of the different fabrication methods on the design of the CMUT cell is shown. Specifically, the design space for the LOCOS CMUTs is investigated, showing how different design choices impact the final cell geometry. In addition, the numerical model that was developed to model the CMUT is presented along with a discussing of the choice of plate material.

Chapter 4 - Device fabrication: This chapter introduces the process flows used for fabricating the CMUTs in the cleanroom. The most critical process steps are highlighted and discussed. Finally, some of the process optimization performed during and before processing is presented.

Chapter 5 - Characterization: In this chapter the fabricated chips are presented through various measurements: optical and electron microscopy, impedance spectroscopy. The limit of detection and mass sensitivity are determined experimentally. These two important figures of merit are compared with the state of the art in the literature.

Chapter 6 - Gas sensing experiments - BMK: In this chapter the CMUT is applied as a gas sensor for the detection of a precursor molecule for the synthesis of (meth)amphetamine. The CMUT is coated with a functionalization layer and combined with a colorimetric chip. The colorimetric method showed a high degree of selectivity while the functionalization layer on the CMUT increased the selectivity towards the precursor molecule.

Chapter 7 - Alternative read-out method: In this chapter an alternative method for detecting the resonance frequency is presented. The method involves both a new top electrode design for the CMUT and an external circuit. The design of both is discussed and finally fabricated CMUT chips with this new design are shown.

Chapter 8 - Conclusion and outlook: In this final chapter the conclusions

from the previous chapters are collected. The chapter ends with an outlook where suggestions of improvements to the work already done and future work is given.

CHAPTER 2

Theory

In this chapter the behavior of the CMUT is sought described through mathematical models. First, the linear harmonic oscillator is presented which is the most basic dynamic model for resonators in general. Next, a static CMUT model is presented, from which analytical expressions of the resonance frequency and pull-in voltage are obtained. The dynamic behavior of the CMUT is modeled by an equivalent circuit model spanning the three energy domains: electrical, mechanical, and acoustical. This model provides insight into what affects the electromechanical coupling coefficient and gives an expression for the impedance which is useful for comparison with measurements. Next, the distributed mass sensitivity is calculated for the CMUT both with regular and stressed plates. The normalized sensitivities of different sensor types are compared which shows why the range of values is so large, as seen in Figure 1.3. Finally, an expression for the mass limit of detection is given which is shown to be affected by the noise in the system and thus by the quality factor of the CMUT.

2.1 Linear harmonic oscillator

In this section the most simple oscillator model is presented: the linear harmonic oscillator. From this model the concepts of eigenfrequency, resonance, damping and quality factor is derived. The derivation and results are valid for any resonator that oscillates and is therefore the most simple model of a CMUT operated as a resonator. The following section is based on [16]

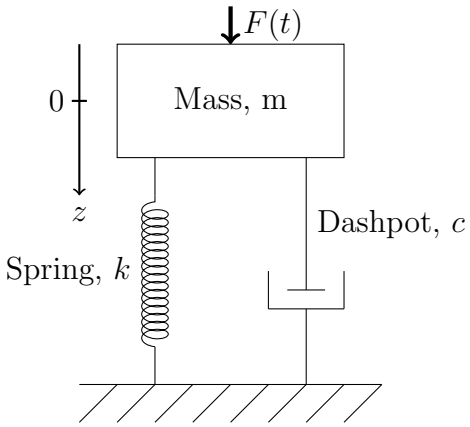


Figure 2.1: Mass, spring, dashpot linear single degree of freedom oscillator. The spring and dashpot are fixed at the ground and at the mass which is free to move along the z-axis.

(p.29-40).

Figure 2.1 shows a schematic representation of the single degree of freedom linear harmonic oscillator model. All relevant properties are lumped in to three discrete elements: the mass m, the spring with a spring constant k, and the dashpot with a coefficient of damping c. All mass is located in a single point and is only made up of the mass m, that is the spring and dashpot are massless. The resonator is assumed to be harmonically driven with am external driving force: $F(t) = F_0 \cos(\omega t)$. Small amplitudes are assumed resulting in linear behavior of all components.

The equation of motion is found by applying Newton's second law on the system. The left side of Equation 2.1 shows the mass times the acceleration while the right side shows the sum of all the forces acting on the mass.

$$m\frac{d^2z}{dt^2} = F(t) - kz - c\frac{dz}{dt}, \tag{2.1}$$

where the spring represents the restoring force obeying Hook's law $F_{\text{spring}} = -kz$ and the damping force is proportional to the speed of the mass $F_{\text{dashpot}} = -c\frac{dz}{dt}$. The terms and constants in Equation 2.1 can be rearranged to the more convenient form:

Table 2.1: Table showing the three damping regimes defined by the damping ratio ζ. The solution characteristic described the behavior of the mass which is exemplified in Figure 2.2.

Regime name	**Damping ratio ζ**	**Solution characteristic**
Over-damped	$\zeta > 1$	No oscillation, amplitude exponentially decreasing
Critically damped	$\zeta = 1$	No oscillation, amplitude exponentially decreasing
Under-damped	$\zeta < 1$	Oscillatory, amplitude exponentially decreasing

$$\frac{d^2z}{dt^2} + 2n_c\frac{dz}{dt} + \omega_0^2 z = \frac{F(t)}{m}, \tag{2.2}$$

$$n_c = \frac{c}{2m}, \tag{2.3}$$

$$\omega_0 = \sqrt{\frac{k}{m}}, \tag{2.4}$$

$$\zeta = \frac{n_c}{\omega_0} = \frac{c}{2\sqrt{km}}, \tag{2.5}$$

where n_c is called the coefficient of damping, ω_0 is the eigenfrequency of the resonator and ζ is the damping ratio.

Undriven oscillator

The simplest case is when there is no external force driving the oscillator: $F(t) = 0$ and consequently the right hand side of Equation 2.2 is zero and the equation becomes homogeneous. Table 2.1 shows the three damping regimes which can be identified by inserting a trial function in the homogeneous differential equation (for a detailed derivation see [16]). Solutions for these three regimes are plotted in Figure 2.2. The solutions in the figure can be seen as the step responses. The behavior described in Table 2.1 is reproduced in the plot where the over-damped solution decreases the fastest followed by the critically damped solution and finally the under-damped solution which oscillates. Since no external force is acting on the system the amplitude of the oscillations will decrease as the oscillator is being dampened as time passes.

Driven oscillator

We now turn the attention to the driven harmonic oscillator, that is when a harmonic external force $F(t)$ is applied to the mass. For the CMUT the driving force is electrostatic and is applied to the plate. The solution is a sum

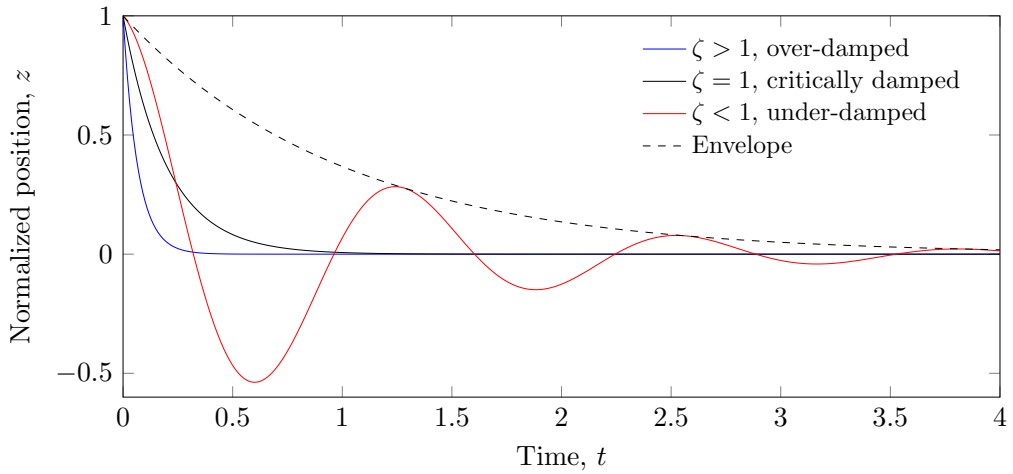


Figure 2.2: Position as a function of time for three solutions to the undriven linear harmonic oscillator. Parameter values: $\omega_0 = 5$, $\zeta_{\text{over}} = 3$, $\zeta_{\text{crit}} = 1$ and $\zeta_{\text{under}} = 0.2$. Note that the only oscillatory behavior is found for the under-damped oscillator.

of the transient solution found for the homogeneous equation and a steady state solution which is independent of the initial conditions. The steady state solution can take the form:

$$z(t) = z_0 e^{i\omega t}. \tag{2.6}$$

The amplitude z_0 can be written in polar form as the norm and phase:

$$|z_0| = \frac{F_0/m}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\zeta\omega_0\omega)^2}} = \frac{F_0/m}{\sqrt{(1 - (\frac{\omega}{\omega_0})^2)^2 + (2\zeta\frac{\omega}{\omega_0})^2}} \tag{2.7}$$

$$\phi = \arctan\left(\frac{2\zeta\omega_0\omega}{\omega^2 - \omega_0^2}\right) = \arctan\left(\frac{2\zeta\frac{\omega}{\omega_0}}{(\frac{\omega}{\omega_0})^2 - 1}\right). \tag{2.8}$$

The gain, G, of the steady state solution can now be extracted from the norm of the amplitude:

$$G = \left(\sqrt{\left(1 - \left(\frac{\omega}{\omega_0}\right)^2\right)^2 + \left(2\zeta\frac{\omega}{\omega_0}\right)^2}\right)^{-1}. \tag{2.9}$$

Figure 2.3 shows a plot of G as a function of the normalized frequency $\frac{\omega}{\omega_0}$ for three damping ratios ζ. The lower the damping (lower ζ) the higher the gain

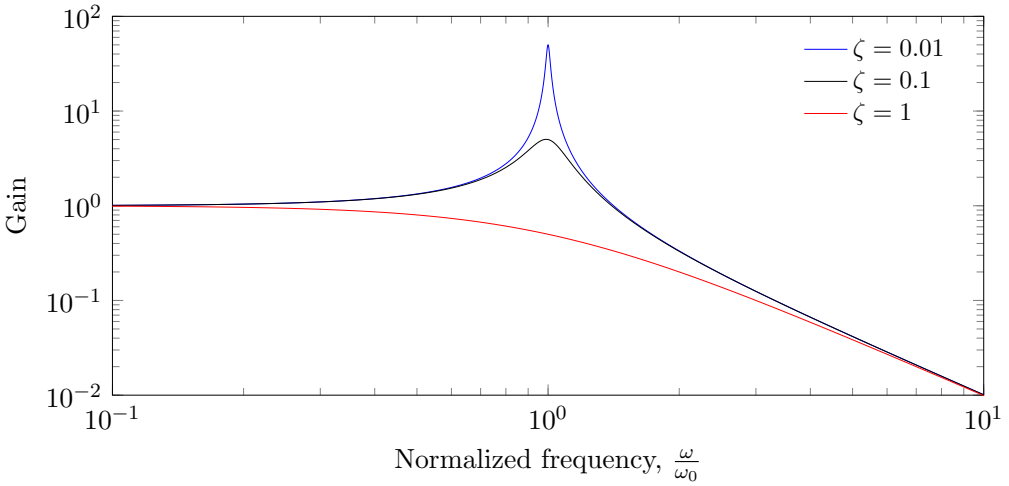


Figure 2.3: Gain as a function of normalized frequency for three damping ratios. The lower the damping ratio the more narrow and higher a peak is observed at resonance.

and more narrow peak. Intuitively this makes sense since a lower damping means less power is dissipated. The frequency at which the maximum gain is achieved is called the resonance frequency and is denoted ω_r. What is less apparent from the figure is the decrease of ω_r for increased damping. Figure 2.4 shows a zoom in on two of the peaks where this decrease of the resonance frequency relative to the eigenfrequency is seen. This phenomenon is described by:

$$\omega_r = \omega_0 \sqrt{1 - 2\zeta^2}, \tag{2.10}$$

which is what was observed in Figure 2.4. Equation 2.10 also shows that for oscillators with minor damping the resonance frequency and eigenfrequency will approximately be the same since $\lim_{\zeta \to 0} \omega_0 \sqrt{1 - 2\zeta^2} = \omega_0$.

For real physical resonators the concept of quality factor Q is used to describe the damping. The relationship between Q and the various physical damping mechanisms is given in Section 2.5. The quality factor Q is closely related to the damping ratio ζ:

$$Q = \frac{\sqrt{1 - 2\zeta^2}}{2\zeta}, \tag{2.11}$$

which means that a slightly damped resonator with $\zeta \ll 1$ will have a high quality factor. For many resonators, and especially for gravimetric mass sensors, a high quality is desired since it results in larger deflection amplitudes (see Figure 2.3), more narrow peaks (lower bandwidth) and a lower noise. For

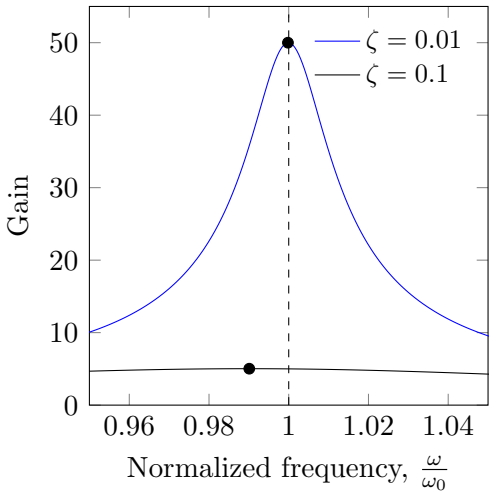


Figure 2.4: Zoom-in on two peaks in Figure 2.3 demonstrating the slight different resonance frequencies cause by the difference in damping.

gravimetric mass sensors a lower noise gives a lower LOD as will be shown in Section 2.5.

The quality factor describes the amount of damping of a resonator. The damping mechanism is solely described by the dashpot in the oscillator model (Figure 2.1) which is the only element capable of dissipating energy. The spring and mass represents the potential and kinetic energy, respectively, of the oscillator doing oscillation. The quality factor can be defined in the energy domain:

$$Q = 2\pi \frac{W}{\Delta W}, \tag{2.12}$$

where W is the total energy stored and ΔW is the energy which is lost in one oscillation period. This relation shows that a high Q is equivalent to having a resonator that loses energy more slowly which means that it will ring down over a longer period of time. Later it is shown how the quality factor can affect the limit of detection.

In this section a simple dynamic model of a linear resonator was presented. It was shown that three distinct damping regimes exist and how the resonator moves in each of these regimes. Lastly, the resonance frequency was determined and it was shown how the damping can be modeled by a single parameter called the quality factor.

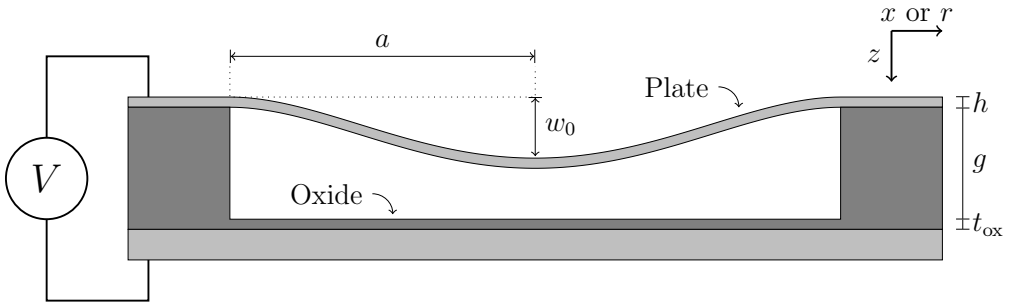


Figure 2.5: Cross-section sketch of a single CMUT cell defining the geometrical parameters and coordinate system.

2.2 Static CMUT model

In this section a static model of the CMUT is presented. The starting point is the plate deflection profile, for which the solution for anisotropic plate materials only differs by a numerical prefactor compared with the solution for isotropic plate materials. The deflection profile enables the calculation of the eigenfrequency, pull-in voltage, energies and capacitance of the CMUT, which are presented in the end of the section. Figure 2.5 shows a cross-section of a CMUT cell with various variables defined. In the remainder of the thesis it is assumed that the plate shape is circular, since circular plates exclusively have been used.

2.2.1 Plate deflection profile

The deflection profile of a plate can be found by solving the differential plate equation. This equation can in its simplest form be written as [74]:

$$\nabla^4 w = \frac{p}{D}, \tag{2.13}$$

where w is a function describing the deflection profile of the plate, p is the pressure difference over the plate and D is the flexural rigidity given by:

$$D = \frac{Eh^3}{12(1-\nu^2)}, \tag{2.14}$$

where E is Young's modulus, h is the thickness of the plate (see Figure 2.5) and ν is Poisson's ratio. Hence, the flexural rigidity contains all the plate material properties. The flexural rigidity given here is valid for isotropic materials. Later in this section the anisotropic version is presented.

Equation 2.13 is the governing equation determining the deflection profile and it is only valid when certain assumptions are met [75]:

- Small plate deflections relative to the plate thickness: $\frac{h}{w_{\max}} \gtrsim 5$. This is to avoid stress stiffening effects.
- Large plate aspect ratio: $\frac{a}{h} \gtrsim 40$. This is required so shear deformation can be ignored.
- No built-in material stress: $\sigma = 0$.

The pressure on the right side of Equation 2.13 corresponds to a uniform applied force on the plate. The boundary conditions are that of a perfectly clamped plate:

$$w(0) < \infty \tag{2.15}$$
$$w(a) = 0 \tag{2.16}$$
$$\frac{\partial w(0)}{\partial r} = 0 \tag{2.17}$$
$$\frac{\partial w(a)}{\partial r} = 0, \tag{2.18}$$

meaning that the deflection at the center of the plate is finite, the deflection at the edge is zero and that the slope is zero both the center and edge of the plate.

Equation 2.13 can for the isotropic material case be rewritten in cartesian coordinates as:

$$\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} = \frac{p}{D}. \tag{2.19}$$

Anisotropic materials are commonly used for plates in CMUTs, e.g. when SOI wafers are fusion bonded to the substrate wafer. In this case Si is the anisotropic material. It has been shown that the isotropic plate deflection profile can have an error of up to 10% as compared with finite element modeling while the anisotropic deflection had an error of $< 1\%$ [76]. This is the motivation behind using the anisotropic plate equation instead of the isotropic version.

The anisotropic generalized plate equation is given as [75, 76]:

$$\frac{\partial^4 w}{\partial x^4} + k_1\frac{\partial^4 w}{\partial x^3 \partial y} + k_2\frac{\partial^4 w}{\partial x^2 \partial y^2} + k_3\frac{\partial^4 w}{\partial x \partial y^3} + k_4\frac{\partial^4 w}{\partial y^4} = \frac{p}{D_a} \tag{2.20}$$

Table 2.2: Compliance constants for N-type crystalline silicon at 300 K for a low doping level ($150\,\Omega \cdot \text{cm}$, $\approx 2.8{\times}10^{13}\,\text{cm}^{-3}$) and high doping level ($3.26\ \text{m}\Omega \cdot \text{cm}$, $\approx 2.1 \times 10^{19}\,\text{cm}^{-3}$). Adapted from [76].

	Low doping	High doping
s_{11}^c	$7.691 \times 10^{-12}\,\text{Pa}^{-1}$	$7.858 \times 10^{-12}\,\text{Pa}^{-1}$
s_{12}^c	$-2.1420 \times 10^{-12}\,\text{Pa}^{-1}$	$-2.2254 \times 10^{-12}\,\text{Pa}^{-1}$
s_{44}^c	$12.577 \times 10^{-12}\,\text{Pa}^{-1}$	$12.628 \times 10^{-12}\,\text{Pa}^{-1}$

where k_1, k_2, k_3, k_4 are plate equation coefficients and D_a is the anisotropic flexural rigidity. These constants are material specific and depends on the orientation of the plate relative to the material crystal structure. In order to find the constants in Equation 2.20 the elastic constants for the plate must first be found. Only part of the derivation is given here, for the full derivation readers are referred to reference [75] and [76].

The classic stress σ, strain ϵ, relation is in the crystallographic coordinate system given by:

$$\epsilon^c = \mathbf{s}^c \sigma^c, \tag{2.21}$$

where $\mathbf{s}^c$ is the compliance matrix in the crystallographic coordinate system and the inverse of this $(\mathbf{s}^c)^{-1} = \mathbf{c}^c$ is the stiffness matrix. In general the compliance matrix is a 6x6 matrix using the 6-vector Voigt notation. However, the compliance matrix can be simplified by assuming plane stress, that is, no stress in the z direction. This assumption is justified by the requirement of a high aspect ratio plate. This assumption reduces the compliance and stiffness matrices to a size of 3x3. The plate constants can now be found for crystal types of different symmetries. For the silicon crystal structure the effective compliance matrix, in the crystallographic coordinate system, becomes:

$$\mathbf{S}_{\text{eff}}^{\text{c}} = \begin{bmatrix} s_{11}^c & s_{12}^c & 0 \\ s_{12}^c & s_{11}^c & 0 \\ 0 & 0 & s_{44}^c \end{bmatrix}, \tag{2.22}$$

where the constants in the matrix have been determined experimentally and is given in Table 2.2 for a low and a high Si doping level.

A coordinate system change is needed in order to transform the compliance and stiffness matrix 2.22 to the coordinate system of the plate. The compliance matrix coordinate system is aligned to $\langle 100 \rangle$ directions, while the most common substrate plane for silicon wafers is (100) and the plates are typically aligned to the primary flat of the wafer which is in the [110]

direction. Therefore, the compliance and stiffness matrices must be rotated (by 45°) to match the coordinate system of the plate. The resulting stiffness matrix is given by [76]:

$$\mathrm{C}^{\mathrm{eff}}_{\mathrm{Si(100),[110]}} = \begin{bmatrix} \frac{1}{s^c_{44}} + \frac{1}{2(s^c_{11}+s^c_{12})} & \frac{1}{2(s^c_{11}+s^c_{12})} - \frac{1}{s^c_{44}} & 0 \\ \frac{1}{2(s^c_{11}+s^c_{12})} - \frac{1}{s^c_{44}} & \frac{1}{s^c_{44}} + \frac{1}{2(s^c_{11}+s^c_{12})} & 0 \\ 0 & 0 & \frac{1}{2s^c_{11}-2s^c_{12}} \end{bmatrix}. \tag{2.23}$$

The plate equation constants for a Si (100) substrate with plates aligned to a primary flat in the [110] direction can now be written in terms of the elements in this stiffness matrix:

$$k_1 = \frac{4C^{\mathrm{eff}}_{13}}{C^{\mathrm{eff}}_{11}} \quad k_2 = \frac{2\left(C^{\mathrm{eff}}_{12} + 2C^{\mathrm{eff}}_{33}\right)}{C^{\mathrm{eff}}_{11}} \quad k_3 = \frac{4C^{\mathrm{eff}}_{23}}{C^{\mathrm{eff}}_{11}} \quad k_4 = \frac{C^{\mathrm{eff}}_{22}}{C^{\mathrm{eff}}_{11}} \tag{2.24}$$

$$D_a = \frac{1}{12}h^3 C^{\mathrm{eff}}_{11}. \tag{2.25}$$

Inserting values for a highly doped Si substrate (see Table 2.2) yields:

$$k_1 = k_3 = 0 \quad k_2 = 1.295 \quad k_4 = 1, \tag{2.26}$$

Inserting these values into the anisotropic plate equation (Equation 2.20) reduces it to:

$$\frac{\partial^4 w}{\partial x^4} + k_2 \frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} = \frac{p}{D_a}. \tag{2.27}$$

This reduced anisotropic plate equation can now be solved using the boundary conditions given earlier. The solution is the plate deflection profile for a plate aligned in the [110] direction on a Si (100) substrate. Equation 2.27 is also valid for circular plates due to symmetry [75]. The deflection profile is given in the next section for the isotropic and anisotropic case.

Solution: deflection profile

An exact solution exists for circular plates, while this is not generally the case, e.g. only approximate solutions exist for square plates. Only the solutions for circular plates will be considered here. The solution to the isotropic plate equation (Equation 2.19) is given in cylindrical coordinates by [75]:

$$w(r) = w_0 \left(1 - \left(\frac{r}{a}\right)^2\right)^2 \tag{2.28}$$

$$w_0 = \frac{pa^4}{64D} \tag{2.29}$$

$$D = \frac{E}{12(1-\nu^2)} h^3, \tag{2.30}$$

where $w_0 = w(0)$ is the deflection at the center of the plate. The corresponding solution to the reduced anisotropic plate equation (Equation 2.27) is given in cylindrical coordinates by [76]:

$$w_a(r) = w_{0,a} \left(1 - \left(\frac{r}{a}\right)^2\right)^2 \tag{2.31}$$

$$w_{0,a} = \frac{1}{8(3 + k_2 + 3k_4)} \frac{pa^4}{64D_a} \tag{2.32}$$

$$D_a = \frac{1}{12} C_{11}^{\text{eff}} h^3, \tag{2.33}$$

the anisotropic deflection profile is very similar to the isotropic deflection profile. The difference lies in a numerical prefactor on the center deflection.

Figure 2.6 shows a plot of the isotropic and anisotropic deflection profiles for a circular Si plate on a (100) substrate. The shape of the profiles is the same but the center deflection differs by approximately 10% at the center of the plate, which is in good agreement with what is found in [76].

Multilayered stressed anisotropic plates

The deflection profiles presented here are valid under the assumptions stated earlier but they are limited to plates with a single layer of material. In some CMUTs the plate consists of multiple layers of different materials, which in addition can have a built-in stress. An example of this could be a Si_3N_4 plate with a top electrode made of Al. To accurately model such systems multiple layers must be included that each can exhibit anisotropy and stress. A theory for this has been developed by Engholm et al. in [77]. The main points will be given here.

The plates can have $n \in \mathbb{N}$ layers, each with different anisotropic elastic properties and values of stress. The stress can either be compressive ($\sigma < 0$) or tensile ($\sigma > 0$). For circular plates the generalized multilayer plate

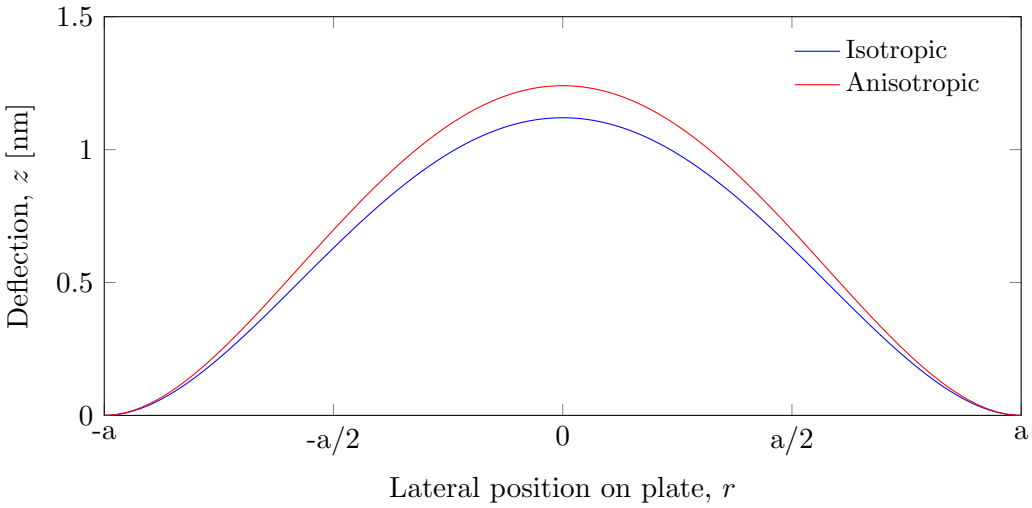


Figure 2.6: Deflection profiles for an isotropic and anisotropic circular Si plate on a (100) substrate. Parameters: $a = 10\,\mu$m, $h = 1\,\mu$m, $p = 1\,$atm, $E_{[110]} = 169\,$GPa and $\nu_{[110]} = 0.062$.

equation simplifies to:

$$\nabla^4 w - \frac{N}{C_d D_a}\nabla^2 w = -\frac{p}{C_d D_a} \tag{2.34}$$

$$C_d = \frac{3 + k_2 + 3k_4}{8}, \tag{2.35}$$

where N is called the stress resultant and is calculated by integrating the stress in a single layer over the thickness of the layer and subsequently summing over all layers. The plate equation constants k_m are in general not the same as presented earlier since they contain information about the materials in every layer. The structure of Equation 2.34 is similar to that of the isotropic plate equation (Equation 2.13) presented in the beginning of this section with the addition of an extra term containing the stress resultant.

The effect of the stress term differs depending on the sign of the stress. If the plate material has a compressive stress the plate will deflect more than if the stress was tensile. Furthermore, a plate under compressive stress will have a lower eigenfrequency than a plate with a tensile stress, which is shown in the next section.

2.2.2 Eigenfrequency

In the previous section the plate deflection profile was found. This profile will now be used to find the eigenfrequency of circular plates. As already

defined in Section 2.1, the eigenfrequency is the frequency at which an undamped plate with no external forces applied (such as a driving force) will have maximum deflection. The first eigenfrequency is found using a variational method, namely Rayleigh-Ritz's method, which is a method for finding eigenvalues. A system where no external forces are acting is a conservative system and energy is conserved. Consequently, the maximum kinetic energy and maximum potential energy are equal, $\max(K) = \max(U)$. Assuming harmonic oscillations this expression can give [74]:

$$\omega_0^2 = 2\frac{\max(U)}{\int\int_{\mathrm{A}} \rho h w^*(x,y)^2 \mathrm{d}x\mathrm{d}y} \tag{2.36}$$

$$= 2\frac{U_s}{\int\int_{\mathrm{A}} \rho h w(x,y)^2 \mathrm{d}x\mathrm{d}y}. \tag{2.37}$$

In Equation 2.36 the numerator represents the maximum potential energy and the denominator represents the kinetic energy. The mode shape function $w^*(x,y)$ is usually chosen to be the static deflection profile under an uniform load. Other approximative profiles can be used, e.g. when no exact solutions exist. In Equation 2.37 the maximum potential energy is substituted with the maximum potential energy for the CMUT which is the strain energy saved in the deflected plate. This energy can also be thought of as the potential energy saved in the spring in the linear harmonic oscillator model (see Figure 2.1). The potential strain energy U_s is given later in Section 2.3.1 but the result is used here to calculate the eigenfreuquency. Several deflection profiles have been presented: isotropic, anisotropic, and multilayer but for the sake of simplicity only the eigenfrequency for a isotropic deflection profile will be shown here.
For a circular isotropic mono layered plate the eigenfrequency calculated using Equation 2.37 gives:

$$\omega_0 = 2.949\sqrt{\frac{E}{\rho(1-\nu^2)}}\frac{h}{a^2}. \tag{2.38}$$

The eigenfrequency scales linearly with the plate thickness and inversely with the radius of the plate squared. The same scaling is seen for square plates [75].

The *resonance frequency* ω_r one can measure, e.g. for a CMUT with a circular plate will be lower than the eigenfrequency ω_0 for two reasons. First, Rayleigh-Ritz's method overestimates the eigenfrequency because the plate is stiffened due to the assumption of the modal shape (which here was chosen as the static deflection profile) [74]. Secondly, the resonance frequency for

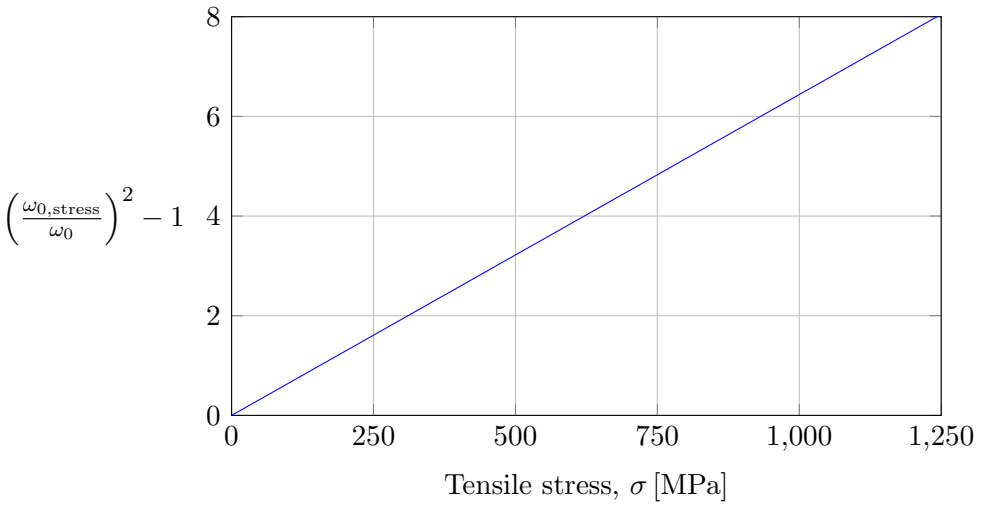


Figure 2.7: Plot showing the relative effect on the eigenfrequency of an added tensile stress to a stoichiometric Si_3N4 plate with realistic CMUT parameters: $a = 5\,\mu\text{m}$, $h = 100\,\text{nm}$.

any real resonator will be lower than the eigenfrequency due to damping as was shown in Equation 2.10. For CMUTs operated in air the eigenfrequency is a good approximation for the resonance frequency.

An important case for CMUTs is finding the eigenfrequency of a plate with internal stress. As mentioned in Section 1.4.2 some CMUTs have plates made of Si_3N_4 which can have a tensile stress $> 1\,\text{GPa}$ depending on the deposition conditions. The stress affects the eigenfrequency and Equation 2.38 needs to be expanded to include this effect. The eigenfrequency for a single layer stressed isotropic plate is given by [77]:

$$\frac{\omega_{0,\text{stress}}^2}{\omega_0^2} = 1 + \frac{1}{16}\frac{Na^2}{D}, \tag{2.39}$$

where ω_0^2 is the eigenfrequency without stress (Equation 2.38) and the stress resultant N for a single layer: $N = \int_0^h \sigma\,\mathrm{d}z$ is the integral of the stress over the plate thickness. Equation 2.39 shows that for tensile stressed materials ($N > 0$), the eigenfrequency is higher than for an unstressed plate and vice versa for compressive stressed materials ($N < 0$).

Figure 2.7 shows a plot of the effect of tensile stress on the eigenfrequency given in Equation 2.39. As the tensile stress σ increases the eigenfrequency increases. To give a sense of the magnitude of the effect, the eigenfrequency in Figure 2.7 is doubled for a tensile stress of $\sigma = 466\,\text{MPa}$ compared with a non-stressed plate.

The eigenfrequency and hereby the resonance frequency can be calculated using the equations presented in this section. It is important to be able to predict the resonance frequency as it, in part, determines the mass sensitivity of the CMUT.

2.2.3 Capacitance

In this section an expression for the electrical capacitance is found. The capacitance is an important parameter to control for the CMUT as a too small capacitance can result in a device with output signals below the noise floor. The capacitance of a parallel plate capacitor and a CMUT with a circular plate is here found and compared. The parallel plate capacitor model consists of two parallel plates that cannot bend (infinite rigidity). One of the plates is suspended in a spring and is free to move and one is fixed.

The capacitance is the constant of proportionality between a stored charge Q and applied voltage V between the top and bottom electrodes: $Q = CV$. The total capacitance can be calculated by regarding each dielectric layer as a capacitor connected in series. The total zero deflection capacitance is given by:

$$C_0 = \left(\sum_{n=1}^{N} \frac{1}{C_n} \right)^{-1} = \epsilon_0 A \left(g + \frac{t_n}{\epsilon_n} + \frac{t_{n+1}}{\epsilon_{n+1}} + \cdots \right)^{-1} = \epsilon_0 A g_{\text{eff}}^{-1}, \qquad (2.40)$$

where ϵ_0 is the vaccum permittivity, A is the electrode area, g is the vacuum gap height, t_n and ϵ_n are the thickness and relative permittivity of the n'th dielectric layer, respectively. Furthermore, an effective gap has been defined which is useful for CMUTs with multiple dielectric layers. As an example the LOCOS structure presented earlier in Section 1.4.2 will have a dielectric stack of: vacuum, Si_3N_4 and SiO_2 which results in an effective gap of: $g_{\text{eff}} = g + \frac{t_{Si_3N_4}}{\epsilon_{Si_3N_4}} + \frac{t_{SiO_2}}{\epsilon_{SiO_2}}$.

The total capacitance of a deflected plate is found by integrating the deflection profile over the area of the plate [76]:

$$C_{\text{t}} = \frac{1}{g_{\text{eff}}} \iint \frac{\epsilon_0}{1 - \eta w(x,y)} \mathrm{d}x\mathrm{d}y, \qquad (2.41)$$

where $\eta \equiv \frac{w_0}{g_{\text{eff}}}$ is the normalized center deflection. Therefore, the capacitance will vary for different geometries due to the differing deflection profiles. For the parallel plate model and circular plate model the integral can be solved

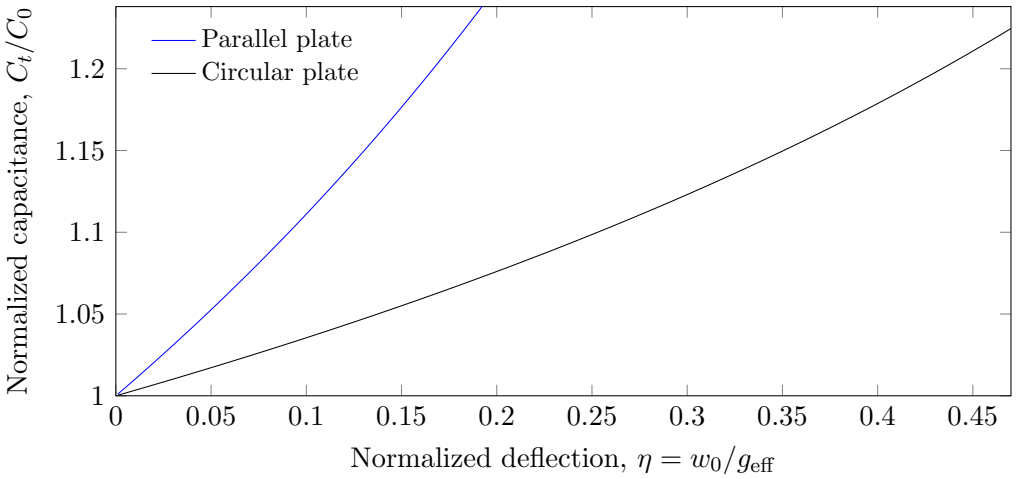


Figure 2.8: Total normalized capacitance (normalized to the total capacitance of the non-deflected CMUT) as a function of normalized deflection for the parallel plate capacitor and circular plate model.

analytically and the resulting capacitances are given by:

$$C_{\text{t,parallel}} = C_0 \frac{1}{1-\eta} \tag{2.42}$$

$$C_{\text{t,circular}} = C_0 \sqrt{\frac{1}{\eta}} \text{arctanh}\left(\sqrt{\eta}\right). \tag{2.43}$$

The two normalized total capacitances are plotted as a function of the normalized deflection in Figure 2.8. The normalized deflection has been plotted until the stable position of the circular plated CMUT. This position is derived later. The parallel plate capacitor deviates more and more from the circular plate model as the deflection is increased.

In practice, all CMUT devices consist of elements with multiple cells connected in parallel. The total measured capacitance is therefore a combination of the sum of cell capacitances with the addition of a contribution of a parasitic capacitance:

$$C_{\text{element}} = C_{\text{parasitic}} + \sum_{m=1}^{M} C_m, \tag{2.44}$$

where C_m is the cell capacitance of the M total cells in the element and the parasitic capacitance is the capacitance between the two electrodes were there are no cells. This is illustrated in Figure 2.9 where the two capacitance contributions are illustrated as lumped capacitors in the CMUT cell.

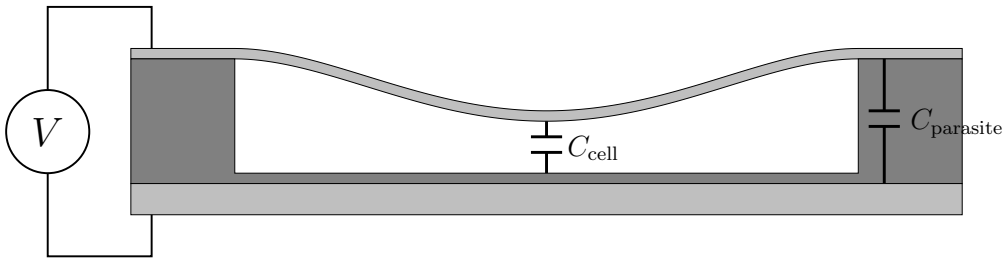


Figure 2.9: Cross-section of a CMUT cell illustrating the cell capacitance in the cavity and parasitic capacitance in the post.

Obviously, increasing the spacer thickness while decreasing its area and relative permittivity is desired in order to decrease the parasitic capacitance. Only looking at the structure in Figure 2.9 an increase of the spacer region thickness will increase the vacuum gap height and thus also decrease the cell capacitance. As will be shown later this is one of the reasons for using the LOCOS structure as the spacer/post region thickness can be decoupled from the vacuum gap thickness.

2.3 Dynamic CMUT model

In this section a dynamic CMUT model is presented which models the transfer of energy between three energy domains: electrical, mechanical, and acoustical. The model is derived from the energy relations of the CMUT which are transformed to state equations. The pull-in voltage is found from the equilibrium of the electrostatic force and mechanical force. A small signal equivalent circuit of the CMUT is presented and an expression for the impedance is derived.

The model presented here is a lumped element model (like the linear harmonic oscillator model) that uses effective variables, as opposed to the real distributed system. However, the lumped element model can provide valuable insight into the dynamics of the CMUT. The effective position variable is chosen to be the center deflection w_0. This is a natural choice due to the symmetry of the circular plate. The plate deflection can then be written as:

$$w(x, y, t) = w_0(t)\tilde{w}(x, y). \tag{2.45}$$

The time dependent center deflection will in the following be written as w_0 for simplicity even though it still explicitly depends on the time.

The starting point is the equation of motion of the CMUT in the me-

chanical domain, which is found by applying Newton's second law of motion:

$$m_0 \frac{\partial^2 w_0}{\partial t^2} = \sum_{n=1}^{N} F_n \tag{2.46}$$

$$m_0 \frac{\partial^2 w_0}{\partial t^2} = F_\mathrm{s} + F_\mathrm{e} + F_\mathrm{p} + F_\mathrm{d}, \tag{2.47}$$

where F_s is the restoring strain force, F_e is the electrostatic force, F_p is the force from the pressure on the plate, and finally F_d is the damping force. These forces are now to be found. For some of the forces it is easier to first calculate the energy and then differentiate this energy to get the force. The energy equation can be written as:

$$U_k = U_s + U_e + U_p + U_d, \tag{2.48}$$

where the energies from left to right are the kinetic, strain, electric, pressure and damping. In the following the energies are determined:

Kinetic energy

The kinetic energy of the plate is in the distributed model given by:

$$U_\mathrm{kin} = \frac{1}{2} \iiint_\Omega \rho \left(\frac{\partial}{\partial t} w(x,y,t) \right)^2 \mathrm{d}x\mathrm{d}y\mathrm{d}z. \tag{2.49}$$

Applying Equation 2.45 and evaluating the integral the expression is reduced to:

$$U_\mathrm{k} = \frac{1}{2} m \left(\frac{\partial w_0}{\partial t} \right)^2, \tag{2.50}$$

which is the classical expression for the kinetic energy of a point mass.

Strain energy

The strain energy is the potential energy stored in the plate when it is deformed during deflection. It is found by integrating the strain energy density over the volume of the plate. The strain energy of a plate in distributed variables and under the same assumptions as in Section 2.2.1 can be written as:

$$U_{\mathrm{s}} = \frac{1}{2}\iiint_{\Omega}(\sigma_1\epsilon_1 + \sigma_2\epsilon_2 + \sigma_6\epsilon_6)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{2.51}$$

$$\epsilon_1 = -z\frac{\partial^2 w(x,y)}{\partial x^2},\ \ \epsilon_2 = -z\frac{\partial^2 w(x,y)}{\partial y^2},\ \ \epsilon_6 = -2z\frac{\partial^2 w(x,y)}{\partial x \partial y}. \tag{2.52}$$

The strain energy will be different for plates with different geometries, but applying the assumption in Equation 2.45, the energy for the lumped plate reduces to:

$$U_s = \frac{1}{2}k_0 w_0^2, \tag{2.53}$$

where k_0 is a spring constant and the energy expression is that of a classical single degree of freedom spring obeying Hooke's law.

Electrical energy

The CMUT is a capacitor and can therefore store electrical charges which gives rise to a potential energy contribution:

$$U_{\mathrm{e}} = \frac{1}{2}V^2 C_{\mathrm{t}} = \frac{1}{2}V^2 \iint_S \frac{\epsilon_0}{g_{\mathrm{eff}} - w_0\tilde{w}(x,y)}\mathrm{d}x\mathrm{d}y \tag{2.54}$$

where C_{t} is the total capacitance of the deflected CMUT which was found in Section 2.2.3. Therefore, the electrical energy is dependent on the deflection profile and by extension so will the electrostatic force.

Energy from pressure

The pressure from an external medium can be seen as a distributed force on the CMUT plate doing work as the plate is deflected. The associated potential energy is given by:

$$U_{\mathrm{p}} = \iint_{\mathrm{S}} p w_0 \tilde{w}(x,y)\mathrm{d}x\mathrm{d}y = p w_0 A_0, \tag{2.55}$$

where the integral is over the surface of the plate. The area A_0 is the effective lumped area of the plate.

Damping energy loss

The damping force for the lumped model is based on the same expression as for the linear harmonic oscillator:

$$F_d = -b\frac{\partial w_0}{\partial t}. \tag{2.56}$$

Table 2.3: Effort and flow variables for the different energy domains. Based on Table 5.1 in [78].

Energy domain	Effort variable	Flow variable	Displacement variable
Electrical	Voltage, V	Current, $\dot{Q}$, i	Charge, Q
Mechanical	Force, F	Velocity, $\dot{w_0}$, v	Position, w_0
Acoustical	Pressure, p	Volumetric flow, $\dot{W}$	Volume, W

The energy dissipated by this force can be calculated by integrating the force:

$$U_d = \int F_d \mathrm{d}w_0 = \int -b\frac{\partial w_0}{\partial t}\mathrm{d}w_0 = \int -b\left(\frac{\partial w_0}{\partial t}\right)^2 \mathrm{d}t, \tag{2.57}$$

where the substitution $\mathrm{d}w_0 = \frac{\partial w_0}{\partial t}\mathrm{d}t$ has been used.

State equations

The CMUT is now considered in three energy domains: mechanical, electrical and the acoustical. In each domain flow $f(t)$, effort $e(t)$, and general displacement $q(t)$ variables are defined, see Table 2.3 [78]. The variables are defined in the time domain and are therefore each a function of time. The product of the flow and effort variables is the power, and the variables $f(t)$ and $e(t)$ are denoted conjugate power variables.

The total energy for the dynamic system (Equation 2.48) can now be given for each energy domain. The total energy is:

$$U_t = U_k + U_s + U_e + U_p + U_d. \tag{2.58}$$

In the following equations, the terms in the total energies with a positive sign are for energy stored/internal energy and the negative signs are for the energy terms where work is done *on* the system. The total energy is given by:

$$U_{t,m} = \frac{1}{2}m\left(\frac{\partial w_0}{\partial t}\right)^2 + \frac{1}{2}k_0 w_0^2 - \frac{1}{2}V^2 C(w_0) - p w_0 A_0 + \int b\left(\frac{\partial w_0}{\partial t}\right)^2 \mathrm{d}t \tag{2.59}$$

$$U_{t,e} = -\frac{1}{2}m\left(\frac{\partial w_0}{\partial t}\right)^2 - \frac{1}{2}k_0 w_0^2 + \frac{1}{2}V^2 C(w_0) - p w_0 A_0 - \int b\left(\frac{\partial w_0}{\partial t}\right)^2 \mathrm{d}t \tag{2.60}$$

$$U_{t,a} = -\frac{1}{2}m\left(\frac{\partial w_0}{\partial t}\right)^2 - \frac{1}{2}k_0 w_0^2 - \frac{1}{2}V^2 C(w_0) + p w_0 A_0 - \int b\left(\frac{\partial w_0}{\partial t}\right)^2 \mathrm{d}t. \tag{2.61}$$

The governing state equations can now be determined from these energy expressions by differentiating the energy with the appropriate displacement variable for the specific energy domain:

$$F = \frac{\partial U_{t,m}}{\partial w_0} = m\frac{\partial^2 w_0}{\partial t^2} + k_0 w_0 - \frac{1}{2}V^2\frac{\partial C(w_0)}{\partial w_0} - pA_0 + b\frac{\partial w_0}{\partial t} \tag{2.62}$$

$$Q = \frac{\partial U_{t,e}}{\partial V} = VC(w_0) \tag{2.63}$$

$$W = \frac{\partial U_{t,a}}{\partial p} = w_0 A_0. \tag{2.64}$$

The solution to these three coupled differential equations describes the behavior of the CMUT in the three energy domains. Before this full dynamic solution is found the pull-in voltage, stable position and effective spring constant is calculated in the next section.

2.3.1 Electrostatic analysis

In this section the pull-in voltage, stable position and effective spring constant is found from the static equations. The pull-in voltage is defined as the DC voltage at which the plate gets deflected so much it touches the bottom of the cavity. This happens when the electrostatic force and pressure force becomes greater than the spring force from the plate. In the static case all accelerations and velocities are zero and thus the inertial and damping term in Equation 2.62 are also zero. The total static force on the plate becomes:

$$F = k_0 w_0 - \frac{1}{2}V^2\frac{\partial C(w_0)}{\partial w_0} - pA_0. \tag{2.65}$$

The effective spring constant of the CMUT is found by differentiating the force with the deflection:

$$k_{\text{eff}} = \frac{\partial F}{\partial w_0} = k_0 - \frac{1}{2}V^2\frac{\partial^2 C(w_0)}{\partial w_0^2}. \tag{2.66}$$

The effective spring constant consists of two terms. The first is a constant that stems from the spring while the second term is due to the electrostatic force and acts to decrease the effective spring constant, hereby making the spring softer. Hence this effect is called spring softening. The second term can be controlled trough the voltage and is dependent on the plate geometry through the capacitance. When the voltage is increased the resonance frequency decreases due to the lowering of k_{eff}.

The voltage at which the device goes in pull-in is called the pull-in voltage V_{PI} and is the voltage at which the effective spring constant becomes zero.

Table 2.4: Lumped parameters, pull-in voltage and pull-in position of the parallel plate capacitor and circular plate capacitor. The applied pressure is set to zero. $D_{\text{eff}} = \frac{3+k_2+3k_4}{8} D_a$.

Parameter	**Parallel plate**	**Circular plate**
Lumped area, A_0	πa^2	$\frac{1}{3}\pi a^2$
Lumped mass, m_0	$\pi a^2 \rho h$	$\frac{1}{3}\pi a^2 \rho h$
Pull-in voltage, V_{PI}	$\sqrt{\frac{8kg_{\text{eff}}^2}{27C_0}}$	$\sqrt{\frac{89.45D_{\text{eff}}g_{\text{eff}}^2}{a^2C_0}}$
Pull-in position, w_{PI}	$\frac{1}{3}g_{\text{eff}}$	$0.463g_{\text{eff}}$

At this point the electrostatic force outbalances the spring force from the plate and the pull-in voltage becomes:

$$k_{\text{eff}} = 0 \Leftrightarrow k_0 - \frac{1}{2}V^2 \frac{\partial^2 C(w_0)}{\partial w_0^2} = 0 \Leftrightarrow \tag{2.67}$$

$$V_{\text{PI}} = \sqrt{\frac{2k_0}{\frac{\partial^2 C(w_0)}{\partial w_0^2}}}. \tag{2.68}$$

The pull-in voltage can now be found for the parallel plate capacitor and circular plate capacitor, by inserting in the expression above. The results are given in Table 2.4 alongside with some of the lumped parameters. The pull-in voltage in the table is given at zero applied pressure. The pull-in voltage is for both models seen to depend on the effective gap, capacitance and the flexural rigidity of the plate.

The pull-in position can now be found by inserting the expression for the pull-in voltage into the expression for the stable position. The pull-in position is the maximum deflection the plate can have before it goes into pull-in. First the stable position is found by setting the total force to zero so no net force is acting on the plate:

$$F = 0 \Leftrightarrow k_0 w_0 = \frac{1}{2}V^2 \frac{\partial C(w_0)}{\partial w_0} + pA_0, \tag{2.69}$$

inserting the expression for the pull-in voltage into this equation yields:

$$w_{\text{PI}} = \frac{\frac{\partial C(w_{\text{PI}})}{\partial w_0}}{\frac{\partial^2 C(w_{\text{PI}})}{\partial w_0^2}} + \frac{pA_0}{k_0}. \tag{2.70}$$

The pull-in position can now be calculated for the parallel plate capacitor and circular plate capacitor. The results are given in Table 2.4 in the case of

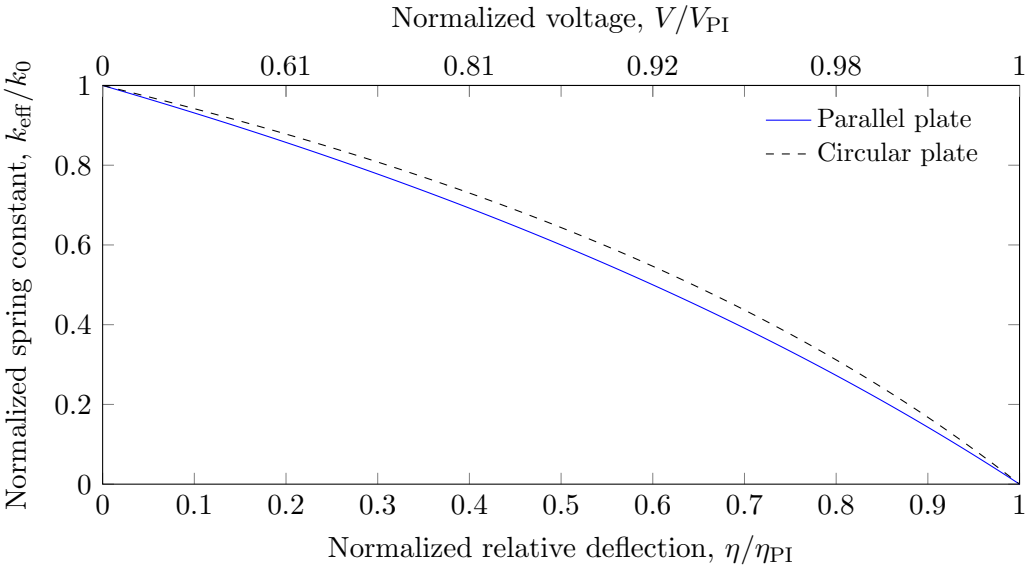


Figure 2.10: Normalized effective spring constant as a function of the normalized relative deflection and normalized voltage for the parallel plate and circular plate capacitor at zero applied pressure. Adapted from [76].

zero applied pressure. The center of the circular plate can be deflected more before going into pull-in than for the parallel plate model, but since the plate cannot bend, in the parallel plate model, it is not just the center position that can reach the pull-in position but every position along the plate.

Figure 2.10 shows the normalized spring constant as a function of the normalized relative deflection and normalized voltage. The spring constant decreases as the voltage or deflection is increased, exhibiting the spring softening effect. When the spring constant reaches zero the device is seen to be in pull-in. The non-linear nature of the CMUT is seen where a small change in voltage gives rise to a large deflection (notice the scale at the voltage axis). The maximum deviation between the parallel and circular plate is 0.047 at $\eta/\eta_{\mathrm{PI}} = 0.63$. This rather small deviation is impressive considering that the plate in the parallel plate model has an infinite flexural rigidity.

In this section the effective spring constant, pull-in voltage and pull-in position was found for the parallel plate capacitor and circular plate capacitor. The scaling of the results was seen to be similar for the two models. The results were derived by looking at the electrostatic case, that is all time dependent terms were excluded. In the next section we shall return to a dynamic study using the governing state equations 2.62 to 2.64. From the solution to these equations the dynamic behavior of the CMUT can be examined.

2.3.2 Dynamic equivalent circuit model

In this section the dynamic equivalent circuit of the CMUT is derived. Such a circuit can be used to model the dynamic behavior of the CMUT and the transduction between the electrical, mechanical and acoustical domains. Finally, the equivalent circuit is reduced to the simple Butterworth-Van Dyke circuit in the electrical domain from which the electrical impedance is found.

The starting point is the non-linear governing equations given previously in Equations 2.62 to 2.64. If only small signals are applied, the response can be assumed to be linear around an operation point. The CMUT can be thought of as having been deflected by a DC voltage and around this point w_{op} only small signals $V(t)$ or $p(t)$ are applied so that the CMUT operates in a small region around these points V_{op} and p_{op} in the electrical and acoustical domain as well. This linearization enables the use of the Laplace transform which transforms the equations from the time domain to the frequency domain. The assumption about small signals is good in the case of the CMUTs used for gravimetric sensing as the AC voltages typically are small and much lower than the DC voltages. As an example consider the following typical biasing scheme: $\frac{V_{\text{DC}}}{V_{\text{AC}}} = \frac{50\,\text{V}}{50\,\text{mV}} = 1000$, exemplifying the point.

Applying the linearization to the governing equations in matrix form they become:

$$\begin{bmatrix} \mathrm{d}Q \\ \mathrm{d}F \\ \mathrm{d}W \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathrm{d}V \\ \mathrm{d}w_0 \\ \mathrm{d}p \end{bmatrix}, \tag{2.71}$$

with the matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} \frac{\partial Q}{\partial V} & \frac{\partial Q}{\partial w_0} & \frac{\partial Q}{\partial p} \\ \frac{\partial F}{\partial V} & \frac{\partial F}{\partial w_0} & \frac{\partial F}{\partial p} \\ \frac{\partial W}{\partial V} & \frac{\partial W}{\partial w_0} & \frac{\partial W}{\partial p} \end{bmatrix} = \begin{bmatrix} C(w_{\text{op}}) & V_{\text{op}} \frac{\partial C(w_{\text{op}})}{\partial w_0} & 0 \\ -V_{\text{op}} \frac{\partial C(w_{\text{op}})}{\partial w_0} & k_{\text{eff}} + m_0 \frac{\partial^2}{\partial t^2} + b \frac{\partial}{\partial t} & -A_0 \\ 0 & A_0 & 0 \end{bmatrix}, \tag{2.72}$$

which is now linearized around the operation point: w_{op}, V_{op} and p_{op}. The linearization is marked by having a 'd' in front of the variables in Equation 2.71. The displacement variables (see Table 2.3 for the definitions) are transformed into their flow counterpart in the same energy domain:

$$\mathrm{d}Q = \int \mathrm{d}\dot{Q}\mathrm{d}t = \int \mathrm{d}i\mathrm{d}t \tag{2.73}$$

$$\mathrm{d}w_0 = \int \mathrm{d}\dot{w}_0\mathrm{d}t = \int \mathrm{d}v\mathrm{d}t \tag{2.74}$$

$$\mathrm{d}W = \int \mathrm{d}\dot{W}\mathrm{d}t. \tag{2.75}$$

Inserting these new transformed variables into Equation 2.71 gives:

$$\begin{bmatrix} \int \mathrm{d}i\mathrm{d}t \\ \mathrm{d}F \\ \int \mathrm{d}\dot{W}\mathrm{d}t \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathrm{d}V \\ \int \mathrm{d}v\mathrm{d}t \\ \mathrm{d}p \end{bmatrix}. \tag{2.76}$$

Equation 2.76, which is in the time domain, is now Laplace transformed into the frequency domain. Thus, the independent variable is changed: $t \rightarrow s = j\omega$, where j is the imaginary unit. Applying the Laplace transform to Equation 2.76 gives:

$$\begin{bmatrix} \frac{1}{s}\mathrm{d}i \\ \mathrm{d}F \\ \frac{1}{s}\mathrm{d}\dot{W} \end{bmatrix} = \begin{bmatrix} C(w_{\mathrm{op}}) & V_{\mathrm{op}}\frac{\partial C(w_{\mathrm{op}})}{\partial w_0} & 0 \\ -V_{\mathrm{op}}\frac{\partial C(w_{\mathrm{op}})}{\partial w_0} & k_{\mathrm{eff}} + ms^2 + bs & -A_0 \\ 0 & A_0 & 0 \end{bmatrix} \begin{bmatrix} \mathrm{d}V \\ \frac{1}{s}\mathrm{d}v \\ \mathrm{d}p \end{bmatrix} \Leftrightarrow \tag{2.77}$$

$$\begin{bmatrix} \mathrm{d}i \\ \mathrm{d}F \\ \mathrm{d}\dot{W} \end{bmatrix} = \begin{bmatrix} sC(w_{\mathrm{op}}) & V_{\mathrm{op}}\frac{\partial C(w_{\mathrm{op}})}{\partial w_0} & 0 \\ -V_{\mathrm{op}}\frac{\partial C(w_{\mathrm{op}})}{\partial w_0} & k_{\mathrm{eff}}/s + sm + b & -A_0 \\ 0 & A_0 & 0 \end{bmatrix} \begin{bmatrix} \mathrm{d}V \\ \mathrm{d}v \\ \mathrm{d}p \end{bmatrix} \Leftrightarrow \tag{2.78}$$

$$\begin{bmatrix} \mathrm{d}i \\ \mathrm{d}F \\ \mathrm{d}\dot{W} \end{bmatrix} = \mathbf{B} \begin{bmatrix} \mathrm{d}V \\ \mathrm{d}v \\ \mathrm{d}p \end{bmatrix}. \tag{2.79}$$

The different elements in the matrix $\mathbf{B}$ each represent either an impedance or transformation factor. The diagonal elements are the impedances, as these relate the flow and effort variables in the same energy domain. The first diagonal element B_{11} relates the current and voltage and is therefore the inverse of the electrical impedance $\mathrm{B}_{11} = 1/Z_{\mathrm{e,b}}$ when the mechanical domain is blocked, that is $\partial v = \partial p = 0$. The elements are inserted in Equation 2.80 below. The next diagonal element B_{22} relates the force and velocity for zero voltage and pressure ($\partial V = \partial p = 0$) which is equivalent to the situation where the electrical circuit is short circuited and the CMUT is operated in a vacuum. Therefore, this element is the mechanical impedance: $\mathrm{B}_{22} = Z_{\mathrm{m,s}}$. The last diagonal element B_{33} is zero which means that the impedance of the acoustic domain is not included in the model for the CMUT itself. The CMUT is merely in contact with the acoustic domain as seen in Figure 2.11 where the CMUT model is confined within the dashed box labeled CMUT. This accounts for the all the impedances.

$$\mathbf{B} = \begin{bmatrix} 1/Z_{\mathrm{e,b}} & \Gamma_{\mathrm{em}} & 0 \\ -\Gamma_{\mathrm{em}} & Z_{\mathrm{m,s}} & -\Gamma_{\mathrm{am}} \\ 0 & \Gamma_{\mathrm{am}} & 0 \end{bmatrix} \tag{2.80}$$

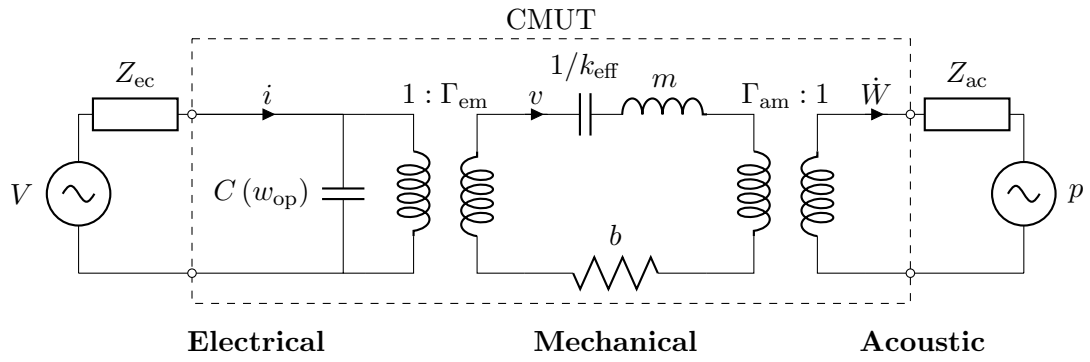


Figure 2.11: Equivalent circuit of the CMUT and surrounding acoustic medium. The CMUT itself is inside the dashed box, spanning all three energy domains.

The first pair of off-diagonal elements $\mathrm{B}_{12} = -\mathrm{B}_{21}$ gives the transduction Γ_{em} between the electrical and mechanical domains and vice versa, respectively, hence the sign difference. The next pair of transduction factors is the elements B_{13} and B_{31} which couples the electrical and acoustical domains but as this is not physically possible they have a value of zero. The last pair of transduction factors is $\mathrm{B}_{23} = -\mathrm{B}_{32}$ which couples the mechanical and acoustical domains.

The equivalent circuit of the CMUT in Figure 2.11 can be reduced to the classical Butterworth-Van Dyke (BVD) circuit for a resonator by assuming that the acoustic circuit is shorted, hereby effectively excluding this domain. This assumption is valid if the CMUT is operated in vacuum but the assumption is also approximately valid for operation in air at ambient pressure which is the case for most CMUT gas sensors. The discrete mechanical elements can be transformed to the electrical domain via the transduction factor Γ_{em} and by using that effort is shared for the electrical and mechanical domain. Whenever an effort variable is shared the components are connected in parallel. Consequently, the electrical impedance (the capacitor) and the three mechanical elements are to be connected in parallel, yielding the BVD circuit as seen in Figure 2.12. The component values in the mechanical domain have

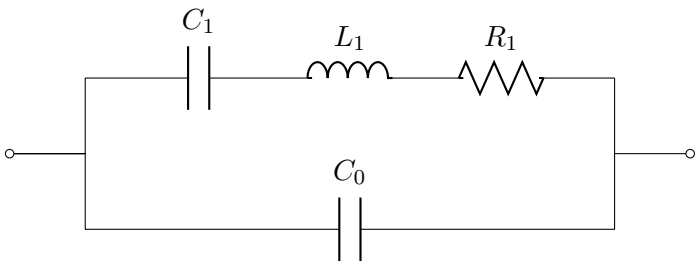


Figure 2.12: Butterworth-Van Dyke (BVD) equivalent circuit of the CMUT. The top branch is the electrical equivalent of the mechanical elements while the bottom branch is the electrical capacitance.

been transformed to their corresponding values in the electrical domain:

$$C_0 = C(w_{\text{op}}) \tag{2.81}$$

$$C_1 = \frac{\Gamma_{\text{em}}^2}{k_{\text{eff}}} \tag{2.82}$$

$$L_1 = \frac{m_0}{\Gamma_{\text{em}}^2} \tag{2.83}$$

$$R_1 = \frac{b}{\Gamma_{\text{em}}^2}. \tag{2.84}$$

The elements from the mechanical domain C_1, L_1, and R_1 are linked to the rigidity/spring constant, mass and damping of the plate, respectively. The transduction factor between the electrical domain and mechanical domain is an important parameter. Comparing Equations 2.80 and 2.78 this factor is given by:

$$\Gamma_{\text{em}} = V_{\text{op}} \frac{\partial C(w_{\text{op}})}{\partial w_0}. \tag{2.85}$$

The coupling between the two domains is greater for a voltage closer to the pull-in voltage. Furthermore, the factor $\frac{\partial C(w_{\text{op}})}{\partial w_0}$ is increasing the closer the plate is to the pull-in position. This can be seen in Figure 2.8 where the derivative is increasing for both the parallel plate and circular plate capacitor as a function of the deflection. Consequently, both factors in 2.85 are increasing with increasing DC voltage and the mechanical components in the electrical domain are expected to change with DC voltage.

The electromechanical coupling factor k^2 is a measure for how much energy is transferred between these two domains. One definition of this factor is the ratio between the energy stored in the mechanical domain E_m to the total energy E_t [79]:

$$k^2 = \frac{E_m}{E_t}. \tag{2.86}$$

Since the model in this reduced version only consists of the electrical and mechanical domains the total energy, $E_t = E_e + E_m$, can be substituted and the energies substituted with the corresponding powers:

$$k^2 = \frac{E_m}{E_m + E_e} = \frac{1}{1 + \frac{E_e}{E_m}} = \frac{1}{1 + \frac{P_e}{P_m}} = \frac{1}{1 + \frac{Z_m}{\Gamma_{em}^2 Z_e}}. \tag{2.87}$$

In order to increase k^2 the transduction factor Γ_{em} should be increased which, as just described, is achieved by operating close to pull-in. The mechanical impedance Z_m should be decreased while the electrical impedance Z_e should be maximized. The electrical impedance can be increased by decreasing the electrical capacitance, which in turn is achieved by either increasing the effective gap or decreasing the electrode area. These two changes will both work toward increasing the pull-in voltage (see Table 2.4). It is impractical to work with a too high pull-in voltage (e.g. $V_{\mathrm{PI}} > 300\,\mathrm{V}$), due to voltage ratings on components in external circuits. In addition, the dielectric layers can experience electrical breakdown if their breakdown voltage is exceeded. As a result the pull-in voltage is limited and hence the maximum capacitance. Another practical upper limit on the electrical impedance, and hereby k^2, is set by the noise floor of the measurement system including the CMUT. If the CMUT capacitance is made very small compared with the parasitic capacitance of the external circuit or the CMUT itself the signal will not be detectable.

The effect of parasitic capacitance is to decrease the coupling coefficient k^2 since the total capacitance is increased. In Section 2.2.3 it was shown that the parasitic capacitance from the areas around the cells are connected in parallel with the cell capacitance:

$$C_{\mathrm{total}} = C_{\mathrm{parasitic}} + C_{\mathrm{cell}}. \tag{2.88}$$

The increased total capacitance lowers the electrical impedance which consequently lowers the coupling coefficient k^2. An important objective in the design of a CMUT sensor is therefore to find ways of decreasing $C_{\mathrm{parasitic}}$. This can be achieved by increasing the thickness of the post regions around the cells or entirely eliminating the electrodes in these areas. Furthermore, the electrical field can be confined to the center of the cell by structuring the top or bottom electrode. These two strategies are applied in the double LOCOS structure, that was introduced in Section 1.4.2.

In summary, a dynamic model of the CMUT was derived and linearized to form an equivalent circuit representation. Furthermore, the mechanical coupling coefficient was introduced and it was shown how this was affected by the DC voltage and parasitic capacitance.

Table 2.5: Realistic lumped CMUT values for the BVD circuit. Adapted from [48].

Element	Value
C_0	21.7 pF
C_1	2.3 pF
L_1	36.1 μH
R_1	29.8 Ω

2.3.3 Impedance analysis

In this section the electrical impedance of the CMUT is introduced. The impedance is often measured for real devices as important parameters can be deduced, such as the resonance and antiresonance frequency, electromechanical coupling factor, and quality factor. The effect of damping and parasitic capacitance on the impedance spectra is presented.

The total impedance of the BVD circuit shown in Figure 2.12 is given by the electrical and mechanical impedance connected in parallel:

$$Z_t = (Z_e^{-1} + Z_m^{-1})^{-1} = \frac{Z_e Z_m}{Z_e + Z_m} \tag{2.89}$$

$$Z_e = \frac{1}{sC_0} \tag{2.90}$$

$$Z_m = \frac{1}{sC_1} + m_1 s + R_1. \tag{2.91}$$

It can be instructive to look at a plot of the impedance of a CMUT with typical parameter values. The impedance can be represented in a variety of ways in the frequency domain and Figure 2.13 shows the real and imaginary part of the impedance as a function of frequency together with a Bode plot.

Realistic lumped parameters for a CMUT are used in the plot in Figure 2.13. The parameters given in Table 2.5 are from [48]. Figure 2.13 shows a Bode plot which is a plot of the impedance magnitude as a function of frequency and the phase as a function of frequency. Furthermore, the figure also shows the real and imaginary part of the impedance. Two points are of special interest, namely where the imaginary part of the impedance becomes zero. As seen on the figure this happens twice and the corresponding frequency values are the resonance, ω_r, and antiresonance, ω_a, frequency. The resonance and antiresonance frequency approximately correspond to the minimum and maximum value of the impedance magnitude, respectively. The

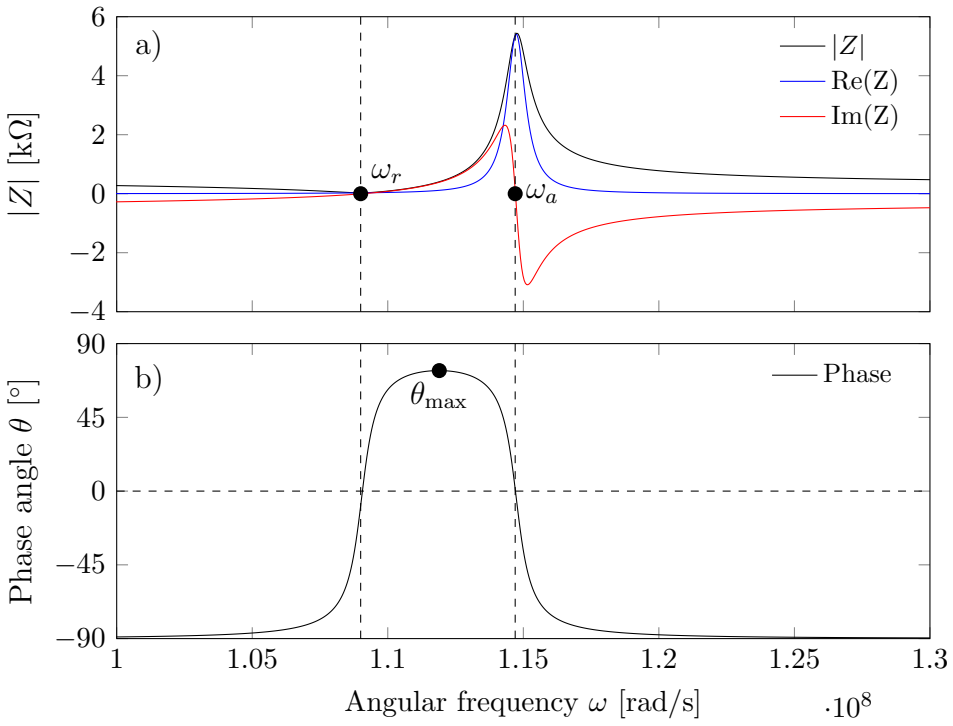


Figure 2.13: a) Plot of the magnitude, real and imaginary part of the impedance as a function of angular frequency. b) Phase as a function of angular frequency. The impedance has been calculated by Equation 2.89 and is based on the lumped values in Table 2.5.

reason why these two points are not exactly the same is due to the damping of the CMUT, but since the damping is relatively low the points are almost at the same frequency. In fact, in the case of the impedance of the CMUT in Figure 2.13 a) the relative difference between the antiresonance frequency where $\mathrm{Im}(Z) = 0$ and $\max(|Z|)$ is only 0.05%. This small relative difference shows that using the local extrema of the impedance magnitude only gives a negligible error and can therefore be used as an alternative way of determining ω_a and ω_r for CMUTs. As the damping increases the difference between the two points increase as well. Figure 2.13 b) illustrates that at ω_a and ω_r the phase is zero. At frequencies lower than the resonance frequency the CMUT behaves as a capacitor and the phase goes towards $-90°$. At resonance ω_r the impedance of the 'spring' C_1 and 'mass' L_1 cancel out (the magnitudes are equal but phase shifted 180°) and the BVD circuit is reduced to R_1 connected on parallel with C_0 and the total impedance decreases until it reaches the value of R_1. At antiresonance ω_a the impedance from the mechanical path (C_1, L_1 and R_1) and the electrical path (C_0) are equal in

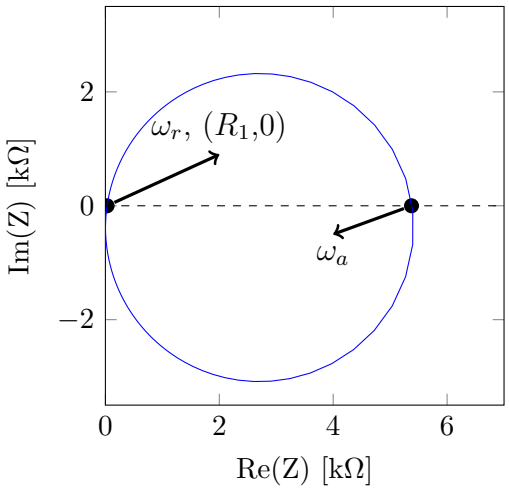


Figure 2.14: Nyquist plot of the impedance shown in Figure 2.13. The point at which the circle crosses the Im(Z)=0 line lies close to the second axis due to the low R_1 value.

magnitude and in phase and the total impedance is maximized. Finally, at frequencies higher than the resonance frequency the CMUT goes back to behaving as a capacitor and the phase tends towards $-90°$.

Figure 2.14 shows a Nyquist plot of the impedance. This is an alternative representation of the impedance where the real part of the impedance is plotted on the first axis and the imaginary part on the second axis. The first point at which the impedance is purely real is the resonance frequency and because of the small value for R_1 this point is close to the origin. The second purely real point is at antiresonance. In Nyquist plots each point correspond to a different frequency although not shown explicitly. In the following Bode plots ($|Z|$ and θ) are chosen for the impedance plots.

Effect of parasitic capacitance and damping

In the following the effects of varying the component values in the BVD circuit is investigated. Figure 2.15 shows the BVD circuit with a parasitic capacitor added in parallel with the electrical capacitance, increasing the total capacitance. Increasing C_1 leads to a decrease of the spring constant (making the spring softer) which will decrease the resonance frequency. Likewise, increasing L_1 will increase the mass which also decreases the resonance frequency. In a distributed system, as in Section 2.2.2, where the eigenfrequency for the distributed plate was given, these dependencies were found directly for the geometrical variables and material constants as opposed to the more generalized variables of spring constants and masses presented here.

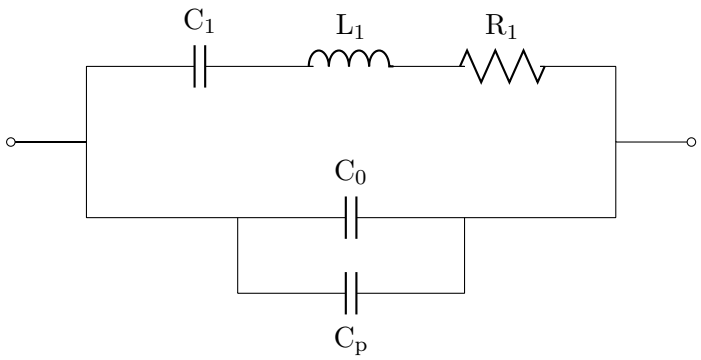


Figure 2.15: Butterworth-Van Dyke equivalent circuit of the CMUT with parasitic capacitance C_p included.

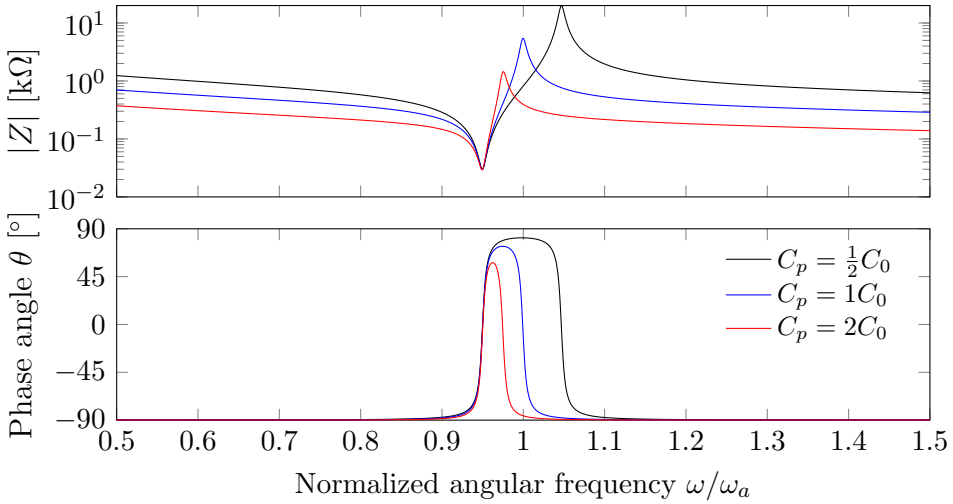


Figure 2.16: Bode plots for three values of parasitic capacitance C_p. The angular frequency is normalized to the antiresonance frequency for $C_p = C_0$.

However, the scaling of the resonance frequency is found to be the same.

Figure 2.16 demonstrates the effect of parasitic capacitance on the impedance response. The maximum impedance at antiresonance decreases approximately an order of magnitude for a quadrupling of the parasitic capacitance, in addition the antiresonance frequency is decreasing but the resonance frequency remains the same. This makes sense as the resonance frequency only depends on the mechanical branch of the BVD circuit. As mentioned earlier the coupling factor is increased by decreasing the parasitic capacitance. A phase response which is box-like with a phase shift close to 180° at resonance (such as the black line in Figure 2.16) is therefore indicative of a device with a high coupling factor.

The effect of damping on the impedance response is shown in Figure

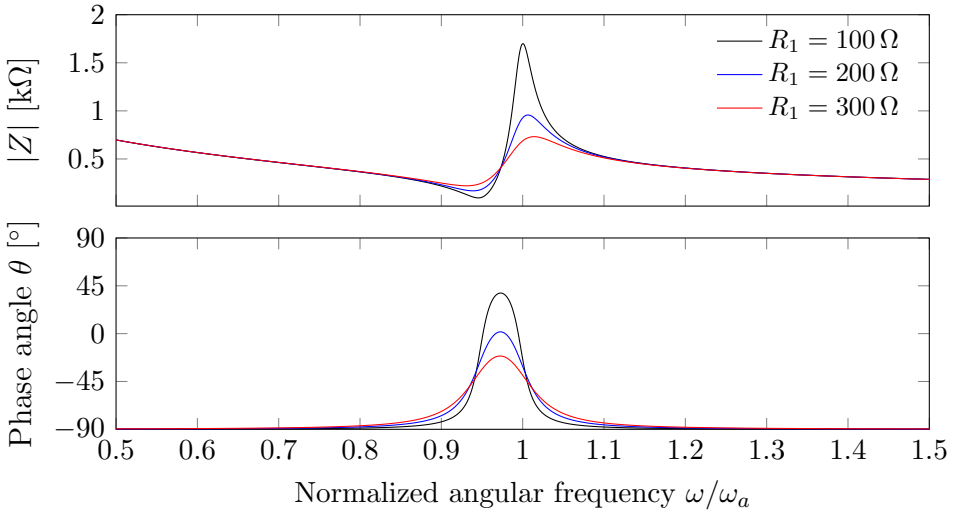


Figure 2.17: Bode plots for three values of damping R_1 showing the effect of an increased damping in the impedance spectrum.

2.17 which shows a Bode plot for different values of R_1. Here the maximum impedance and phase shift is decreasing for increasing damping which is the same effect as the parasitic capacitance has on the impedance response. Thus, some of the components in the BVD model are confounded in that they have a similar effect on the measured impedance. Hence, the analytic impedance function in Equation 2.89 is frequently fitted to the experimental impedance spectra giving the discrete component values. These values can then be used to calculate e.g. the coupling factor. Figure 2.18 shows the impulse response for different amount of damping corresponding to the Bode plot in Figure 2.17. The amplitudes are seen to decrease as a function of time which mirrors the behavior of the linear harmonic oscillator in Figure 2.2. For $R_1 = 0\,\Omega$ the amplitude is constant over time as the transducer is not damped. For all other values of R_1, in the range shown, the CMUT behaves as an under-damped oscillator.

This section presented a dynamic model of the CMUT resulting in a small signal equivalent circuit. From this circuit, important concepts such as the electromechanical coupling coefficient and impedance were derived. Moreover, it was seen how decreasing the parasitic impedance improved device performance by increasing the coupling coefficient and increasing the amplitude of the impedance signal.

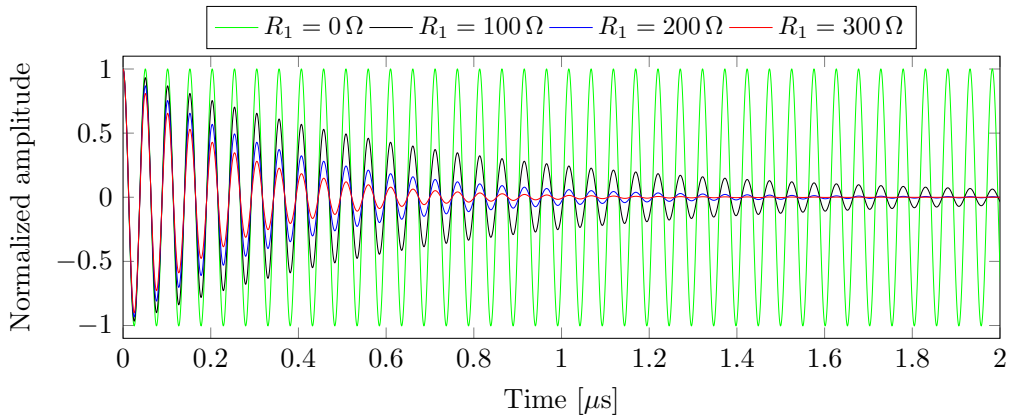


Figure 2.18: Impulse response of the lumped parameter CMUT with parameters given in Table 2.5 for different damping. For $R_1 = 0\,\Omega$ the oscillations are undamped and the amplitude does not decrease.

2.4 Mass sensitivity

In this section the mass sensitivity is investigated using the expressions for the resonance frequency found earlier. The normalized sensitivity is found for four gravimetric sensor types: BAW, SAW, FBAR, and CMUT. Finally, the sensitivity is plotted for a CMUT with an internally stressed plate.

The sensitivity is defined as the change of output for a given change of input. If not explicitly stated otherwise sensitivity is here referring to the *distributed mass sensitivity*, which is when a mass is uniformly covering the oscillator surface. For gravimetric sensors the distributed mass sensitivity was given in the introduction by Equation 1.1 and is repeated here for convenience:

$$S = \frac{\partial f}{\partial m} = \frac{\partial}{\partial m}\sqrt{\frac{k_{\text{eff}}}{m_{\text{eff}}}} = -\frac{1}{2}\frac{f_0}{m_{\text{eff}}}, \tag{2.92}$$

where ∂f is considered the output and ∂m the input. Equation 2.92 is based on the eigenfrequency of the undamped linear harmonic oscillator and is valid for added masses much smaller than the mass of the oscillator itself. The limits of this assumption are examined further in Section 5.5 through measurements and a finite element model. The expression in Equation 2.92 shows that to maximize the sensitivity a high resonance frequency and a low oscillator mass is required.

The sensitivity can for the CMUT be calculated by inserting expressions for the resonance frequency (Equation 2.38) and plate mass ($m = \pi a^2 \rho h$) in

Equation 2.92 or by evaluating $\frac{\partial f_r}{\partial m}$. For a circular isotropic plate this yields:

$$S = -\frac{f_r}{2m} = \frac{\partial f_r}{\partial m} = -0.0747\sqrt{\frac{E}{\rho^3(1-\nu^2)}}\frac{1}{a^4}. \tag{2.93}$$

The sensitivity is only a function of material properties and the radius of the plate. Thus, to increase the sensitivity of CMUTs the plate radius must decrease as this simultaneously increases the resonance frequency and decreases the mass of the plate. It is worth noting that the sensitivity is not dependent on the thickness of the plate. A practical limit exists for how small the radii can be made, both from a fabrication point of view (as will be discussed later in Section 3) and due to the fact that the total cell capacitance will become much smaller than the parasitic capacitance.

A common way of evaluating and comparing sensitivities between gravimetric sensors of the same or of different type is by normalizing with the sensor area and resonance frequency. In a paper by Wenzel and White [17] the normalized sensitivity is defined in general and expressions for different gravimetric sensor types derived. A slightly altered version of the normalized sensitivity given in [17] is given here:

$$S_{\text{norm}} = -\frac{1}{f_{\text{r}}}\frac{\partial f}{\partial m/A} = -\frac{A}{f_{\text{r}}}\frac{\partial f}{\partial m} = -S\frac{A}{f_{\text{r}}}, \tag{2.94}$$

where the added mass is normalized by the area A of the resonator and entire expression is normalized by the unloaded resonance frequency f_r. This normalized sensitivity is calculated for the CMUT by inserting expressions for S, A, and f_r into the equation:

$$S_{\text{norm}} = -\frac{1}{2\rho h}, \tag{2.95}$$

where ρ and h are the density and thickness of the plate, respectively. The same result was found in [80], although with a slightly different notation. This surprisingly simple result shows that a high normalized sensitivity is achieved by a thin plate made of a material with a low density.

Normalized sensitivities are calculated for other gravimetric sensor types in [17] and [26] and the results are given in Table 2.6. For all sensor types the normalized sensitivity is dependent on the inverse of the density of the resonator material. For both the BAW and SAW the normalized sensitivity is a function of the inverse acoustic wavelength λ in the sensor. Therefore, in order to increase S_{norm} the wavelength must be decreased resulting in an increased frequency of the wave. The scaling of S_{norm} for the FBAR and CMUT is similar in that it goes with the inverse of the thickness of

Table 2.6: Normalized sensitivities for different gravimetric sensor types. The expressions are derived in [17] and [26]. λ is the wavelength of the acoustic wave in the device and d is the thickness of the piezoelectric layer in the FBAR.

Sensor type	$\boldsymbol{S_{\text{norm}}}$
BAW	$-\frac{2}{\rho\lambda}$
SAW	$-\frac{2}{\rho\lambda}$
FBAR	$-\frac{1}{\rho d}$
CMUT	$-\frac{1}{2\rho h}$

the resonating element, specifically the piezoelectric layer thickness [26] and the plate thickness, respectively. Thus, the normalized sensitivities of the plate based sensors (FBAR and CMUT) are independent of the resonance frequency which allows for designs operating at lower frequencies, than for the BAW and SAW, while still maintaining the same normalized sensitivity.

2.4.1 Mass sensitivity - stressed plates

In this section the mass sensitivity of the CMUT is found for a plate with built-in stress. As was previously shown the resonance frequency of a plate with internal stress can be higher (tensile stress) or lower (compressive stress) than for the unstressed plate, see Equation 2.39. CMUT sensors with Si_3N_4 plates have been used for sensing [56] and these plates typically have an internal tensile stress. The sensitivity becomes a function of the stress as it influences the resonance frequency. The sensitivity of a stressed plate is calculated by inserting the resonance frequency for a stressed plate in Equation 2.92 for and Equation 2.94 for the normalized sensitivity. However, the analytic expressions become long and tedious and the results are thus only shown numerically through plots.

Figure 2.19 shows two contour plots of the resonance frequency of a plate for the model without (left) and with (right) stress included. As expected, the resonance frequency in the left figure is not dependent on the stress while the resonance frequency in the right figure is dependent on both the stress and the plate thickness. The resonance frequency of thin plates is seen to be more strongly influenced by the stress than for thicker plates. The dashed line is the contour line where the resonance frequency is 50 MHz. The maximum tensile stress in the figure is the stress measured for a Si_3N_4 film in this project, while the range of thicknesses is chosen to cover typical

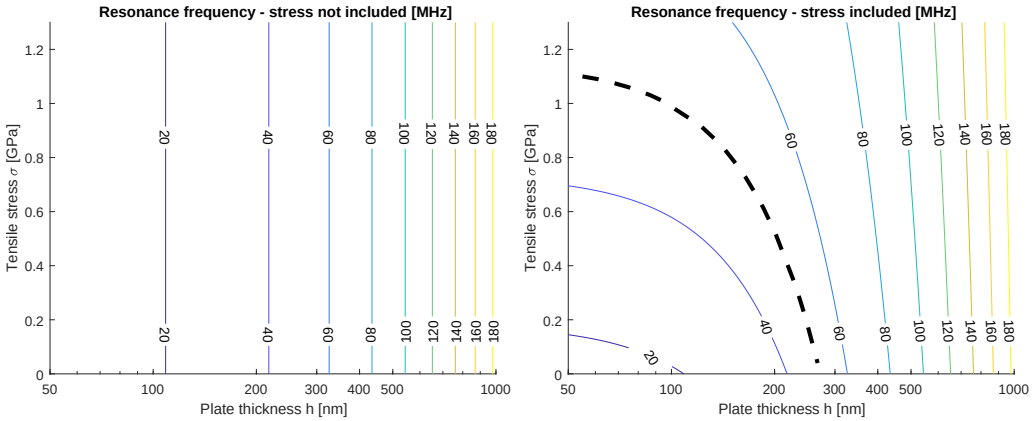


Figure 2.19: Resonance frequency of a Si_3N_4 plate, without (left) and with (right) stress as a function of plate thickness and tensile stress. The dashed line indicates a constant resonance frequency of 50 MHz. The plate radius is held constant at $a = 5\,\mu$m.

values for CMUT plates used for sensing. The plate radius is held constant at $a = 5\,\mu$m in this section.

Figure 2.20 (left) shows a plot of the sensitivity for a Si_3N_4 tensile stressed plate. In contrast to the non-stressed plates (Equation 2.93) the sensitivity now becomes a function of both the plate thickness and stress. The sensitivity is maximized by decreasing the plate thickness and increasing the stress. However, situations commonly arise where the resonance frequency is limited e.g. by external circuitry and a constant resonance frequency of 50 MHz is shown by the dashed line. The dashed line levels off for thinner plates and is almost perpendicular to the contour lines resulting in a rapid increase of sensitivity for decreasing plate thicknesses in this region. The sensitivity of a non-stressed plate is constant, as it is not a function of the plate thickness (Equation 2.93) and is therefore not plotted here.

Finally, Figure 2.20 (right) shows the normalized sensitivity which is the same for stressed and non-stressed plates as it is normalized with the resonance frequency. As Equation 2.95 dictates, the normalized sensitivity is a function of the inverse plate thickness. For a constant resonance frequency (dashed line) the sensitivity increases for a decreasing plate thickness and for these parameters the sensitivity begins to increase most rapidly below about $h = 100\,$nm. This exact threshold value is naturally only true for these exact parameters (plate radius and material properties) but the general behavior is valid for all plates.

In conclusion, in order to maximize the sensitivity of CMUTs, the plate

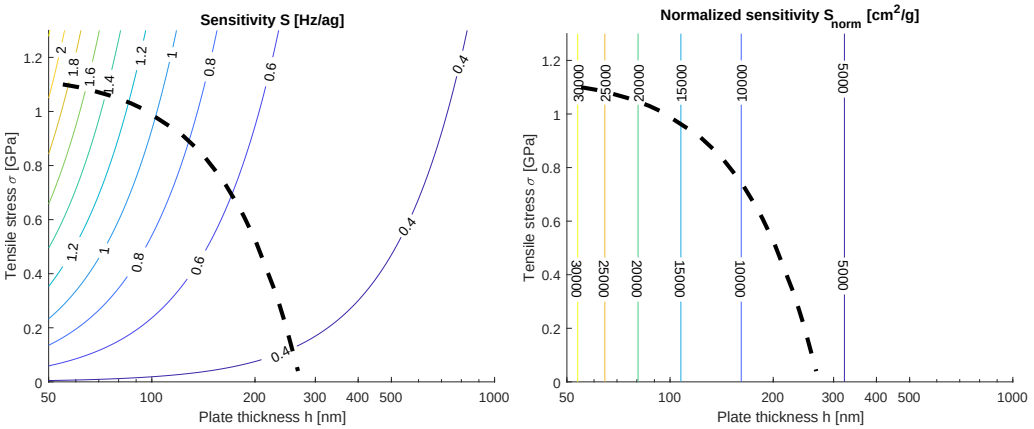


Figure 2.20: Left: Sensitivity as a function of plate thickness and tensile stress for a Si_3N_4 tensile stressed plate. Right: Normalized sensitivity as a function of plate thickness and tensile stress for a Si_3N_4 tensile stressed plate. The dashed line indicates a constant resonance frequency of 50 MHz. The plate radius is held constant at $a = 5\,\mu$m.

radius should be decreased while for stressed plates the sensitivity can be increased by either decreasing the radius or plate thickness while increasing the tensile stress. The normalized sensitivity scales in the same way for stressed and non-stressed plates and it is increased by decreasing the plate thickness. In the beginning of the section is was shown how gravimetric sensors with a flexural plate as the oscillating element (FBAR and CMUT) generally have a normalized sensitivity which scales with the inverse of the plate thickness while S_{norm} for BAWs and SAWs scales with the inverse of the acoustic wavelength and hence linearly with frequency. This is an advantage of the FBAR and CMUT as they can be operated at lower frequencies.

2.5 Limit of detection

In this section an expression for the mass limit of detection (LOD) is found for the CMUT. The concept of Allan deviation is presented and it is shown how this can be used to calculate the LOD. Finally, the quality factor is discussed in relation to the CMUT and how a higher frequency noise gives rise to a higher LOD.

The limit of detection was in the Introduction chapter introduced as the minimum detectable input. Hence, the mass LOD is the minimum mass that can be detected by the sensor due to frequency noise. Noise, in general,

can come from many sources in a measurement setup such as: the resonator itself, the transduction of the electrical signal and from the driving circuitry. All sources contribute to the overall noise of the system. Much research has been devoted to understanding the noise in resonator systems, but the most important aspect, for gravimetric sensors, is how the noise affects the LOD. The LOD can be defined using the so-called Allan deviation [11]:

$$\text{LOD} = 2m_{\text{eff}}\sigma(\tau) \tag{2.96}$$

The total noise of the resonator and external electrical circuit is represented by the Allan deviation $\sigma(\tau)$ which describes the frequency stability as a function of the averaging time τ [16]. In this thesis the *overlapping* Allan deviation is used, which is defined as [81]:

$$\sigma_y(\tau)^2 = (2m^2(M-2m+1))^{-1} \sum_{j=1}^{M-2m+1} \left[\sum_{i=j}^{j+m-1} \bar{y}_{i+m} - \bar{y}_i \right]^2, \tag{2.97}$$

where M is the total number of data points, averaging time $\tau = m\tau_0$ where m is an integer and τ_0 is the sample period. The variable y_i is the fractional frequency data: $(\omega(t) - \omega_r)/\omega_r$. The reason for not using the ordinary standard deviation is that this does not converge as a function of sample size for the noise types frequently encountered in oscillators. Figure 2.21 shows a so-called sigma-tau plot showing a theoretical Allan deviation as a function of τ. At different τ, different noise dominates. The combination of all the noise types gives the Allan deviation shown with a dashed line. The minimum value of the Allan deviation can be used to calculate the minimum LOD using Equation 2.96. The associated averaging time τ may be too slow for the sensor system and a compromise must be made between LOD and measurement speed [16] p.169.

For short averaging times where white additive noise typically is dominant the Allan deviation can be defined in terms of the quality factor [11, 83]:

$$\text{LOD} \propto \frac{m_{\text{eff}}}{Q}\frac{1}{\text{SNR}}, \tag{2.98}$$

where SNR is the signal to noise ratio. The expression is valid for white FM noise which is a function of $\tau^{-\frac{1}{2}}$, see Figure 2.21. This equation shows that a high quality factor and SNR are needed in order to decrease the limit of detection. As will be shown in the next section the overall quality factor of a M/NEMS resonator is affected by various damping terms which will be discussed.

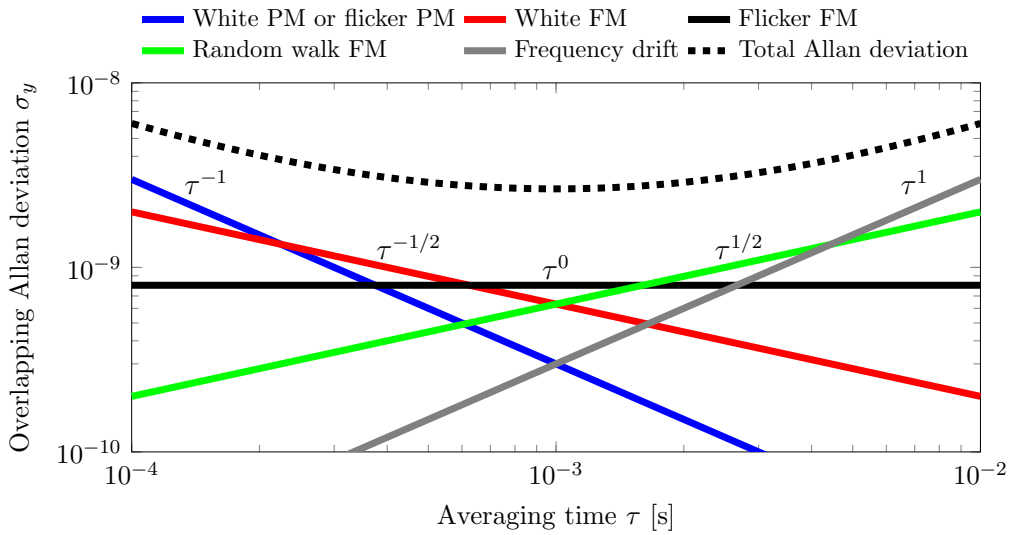


Figure 2.21: Allan deviation as a function of averaging time τ for different types of noise [82]. PM = phase modulation and FM = frequency modulation.

Equation 2.98 also showed that a high SNR is desired in order to decrease the LOD. Lee et al. [84] demonstrated that electrically connecting individual CMUT cells in parallel can decrease the phase noise and increase the SNR. Furthermore, it was found that increasing the number of cells decreases the motional impedance (R_1 in Figure 2.15) and increases Q. At a certain number of cells, Q begins to decrease again which most likely is due to nonuniformities from cell to cell coming from the fabrication process. The exact number of cells where this happens depends on the CMUT design and specific fabrication process. Increasing the number of cells increases the active area of the sensor and makes it more robust as it can still function even though some cells should become damaged etc. All in all, having multiple CMUT cells connected in parallel decreases the LOD and results in a higher SNR which makes it easier to detect the output signal.

2.5.1 Quality factor

The quality factor is important as it is a measure of the damping of the mechanical device and thus, in part, determines the limit of detection as shown in Equation 2.98. The following section is based on [16] chapter 2 and [85]. The quality factor was in Equation 2.12 defined as the ratio between the total energy and the energy dissipated during one oscillation cycle. The

total Q of a resonator can be written as:

$$\frac{1}{Q} = \sum_{i=1}^{N} \frac{1}{Q_i} = \frac{1}{Q_{\text{medium}}} + \frac{1}{Q_{\text{clamping}}} + \frac{1}{Q_{\text{intrinsic}}} + \frac{1}{Q_{\text{other}}}, \tag{2.99}$$

where Q_i are the quality factors of the individual loss mechanisms which will be presented here. For resonant gravimetric sensors the total Q is sought maximized and thus each of the Q factors should be maximized. The right side of Equation 2.99 lists the most important loss mechanisms for M/NEMS resonators and these are presented here. Specifically, Park et al. [86] found that the two most important loss mechanisms for CMUTs, designed for gas sensing, are the 1. medium loss (under ambient conditions) and 2. loss to the support of the plate through the clamping points. Nevertheless, the intrinsic damping mechanisms are included here for completeness and because these set the ultimate limit for the overall quality factor.

Medium losses

As opposed to CMUTs used for medical imaging, as little energy as possible should be coupled (and thus lost) to the surrounding medium. If the CMUT is operated in immersion (e.g. for bio-sensors) it will be heavily damped due to the much higher impedance of the water compared with air. The ideal case is if the surrounding medium is vacuum. Hereby, no energy is lost to the medium. In the following only gaseous media are considered.

Due to the plate structure of CMUTs the surface to volume ratio is high compared to other geometries and hence medium losses are important. When operating in a gas two loss regimes can be defined: the fluidic and ballistic regime. Whether the device is in one or the other is dependent on the relative size of the mean free path length of the gas λ and a typical length scale L of the device. The Knudsen number K is defined as the ratio between the two:

$$K = \frac{\lambda}{L}. \tag{2.100}$$

If $K < 1$ the medium is fluidic and can be thought of as a continuum of molecules with a damping due to a viscous flow at the sensor. If however, $K > 1$ the device enters the ballistic regime where the gas molecules are non-interacting and should be treated as individual molecules dissipating energy from the device through collisions. The mean free path length is given by:

$$\lambda = \frac{k_{\text{B}}T}{\sqrt{2}\pi d_{\text{gas}}^2}\frac{1}{p}, \tag{2.101}$$

where k_B is Boltzmann's constant, T is the temperature, d_{gas} is the diameter of the gas molecules, and p is the pressure. Consequently, the mean free path is dependent on the type of gas and the pressure. When the pressure is lowered the device will at some point enter the ballistic regime and energy dissipation will decrease and hence Q will increase. For ambient conditions in air the mean free path length is approximately [16] $\lambda \approx 70\,\text{nm}$. The diameter of the CMUT can be used as a typical length scale and for the CMUTs fabricated in this project it is about $2a \approx 10\,\mu\text{m}$, which gives a Knudsen number of:

$$K = \frac{70\,\text{nm}}{10\,\mu\text{m}} = 7 \times 10^{-3} \ll 1, \tag{2.102}$$

and as a result the CMUT in ambient conditions is in the fluidic regime and the medium will be the most important factor in lowering the overall Q. No closed-form exact analytical expression exists for Q_{medium} for the CMUT under ambient conditions but an approximation is given in [86]:

$$Q_{\text{medium}} \propto \frac{fh}{p}, \tag{2.103}$$

where f is the frequency, h is the thickness of the plate, and p is the pressure of the gas. To minimize the damping from the gas the frequency and plate thickness of the CMUT should be increased and it should be operated at a low gas pressure (n.b. the relation is only valid in the fluidic regime).

Squeeze film damping is a special case of viscous damping occurring when a beam or plate is in close proximity to a substrate with a gas layer in between. The amount of damping is highly dependent on the separation distance. One of the advantages of CMUTs is that the gap between the plate and bottom electrode can be made to contain a vacuum, thus eliminating squeeze film damping.

Clamping losses

Clamping or support losses are due to energy transferred as acoustic waves in the solid plate to the supporting structure. Generally the energy transfer is dependent on the geometry at the anchoring points and the acoustic impedances of the materials. In [86] Q_{clamping} for the CMUT is estimated from an expression for the clamping loss of a clamped-clamped beam:

$$Q_{\text{clamping}} \propto \left(\frac{a}{h}\right)^3, \tag{2.104}$$

which indicates a very strong dependency on the aspect ratio a/h. In [87] Q_{clamping} has been found for *square membranes* with a semi-infinite support

substrate. An expression for the circular plate is estimated in a similar way above, yielding:

$$Q_{\text{clamping}} \propto \frac{a}{h}, \tag{2.105}$$

which shows a less strong dependence on the aspect ratio.

Therefore, the power α ($(a/h)^{\alpha}$) is speculated to be in between these two extremes of the clamped-clamped beam and circular membrane: $\alpha \in [1, 3]$. Common for the two equations, is that Q_{clamping} is increased for a larger aspect ratio. Intuitively this makes sense as a smaller h results in a smaller contact area between the plate and support. Increasing the radius a increases the potential energy which can be stored in the plate, hereby increasing Q.

Intrinsic losses

Intrinsic losses are a group of loss mechanisms which can be divided into two subcategories: friction losses and fundamental losses. For CMUTs intrinsic losses are not limiting but in general they can put a limit on the maximum value of the quality factor. Here a couple of examples are given for the two subcategories of intrinsic losses.

Friction losses are caused by energy lost to friction at various places in the resonator: at defect sites in solids with a crystalline structure, boundary slipping between two materials (e.g. metal top electrode and plate in CMUTs), or molecular chain movement in e.g. polymers. These losses can be described by the Zener model where Young's modulus is expanded to contain an imaginary part modelling the dissipation of energy. The energy loss is at a maximum at a certain material specific frequency which can be modelled by so-called Debye equations.

Fundamental losses are caused by a heat flow in the resonator when it is oscillating, this could be thermoelastic damping (TED) or phonon-phonon interactions (Akhiezer damping) losses. TED occurs when the resonator is deflected during oscillations and the top and bottom part (for a CMUT) is under either compression or tension, respectively. This leads to a local strain in the material resulting in a temperature difference which consequently gives rise to an irreversible heat flow.

For the CMUT the most important damping mechanisms have been described together with simple geometrical scalings. Exact expressions are difficult or impossible to obtain for most of the damping losses and the expressions presented here are included to give a sense of the general behavior of the quality factor. The two most important contributions to the total quality factor Q_{medium} and Q_{clamping} was seen to show opposite scaling for the plate thick-

ness and radius. More specifically, increasing the plate thickness increases Q_{medium} and decreases Q_{clamping} while an increase of the radius decreases Q_{medium} and increases Q_{clamping}. Consequently, no simple scaling of the geometrical parameters exists for the maximization of the total quality factor.

In conclusion the LOD is affected by the noise in both the resonator itself and the external circuitry. The Allan deviation can be determined experimentally and directly used to calculate the LOD of a device. The Allan deviation is a measure for the total frequency noise in a system and is a function of the averaging time τ. Some noise types can be lowered by increasing the quality factor of the resonator and the two most important energy loss mechanisms for the CMUT was seen to be medium losses and clamping/support losses.

2.6 Gravimetric gas detection

In this section it is shown how adding a functionalization layer on the CMUT results in a concentration limit of detection which is proportional to the mass LOD per area. In Chapter 5 it will be shown that this scaling favors the CMUT as compared with the other gravimetric sensor types.

For gravimetric gas sensors a coating/functionalization layer is typically added on top of the plate in order to give selectivity. The gas molecules absorbed in or adsorbed on the functionalization layer provides the added mass which results in the resonance frequency shift. Under the assumption of linearity and equilibrium the gas concentration in the functionalization layer ∂C_f is proportional to the gas concentration ∂C in the surrounding ambient vapour phase [11]:

$$\partial C_f = K \partial C, \tag{2.106}$$

where the factor of proportionality is called the partition coefficient which is unique for every combination of gas and functionalization layer material. Obviously, K should be maximized to achieve the highest frequency shift for a given ambient gas concentration. The partition coefficient is dependent on a number of variables: "Chemical, vapour solubility properties, polarizability, dipolarity, hydrogen bond acidity, hydrogen bond basicity, between gas molecules and sorbent layer" [11]. The total added mass of the absorbed gas molecules can be written as:

$$\partial m = K(\pi a_f^2 h_f)\partial C = K V_f \partial C, \tag{2.107}$$

where $V_f = \pi a_f^2 h_f$ is the volume of the functionalization layer. The expression above is inserted in the relation relating the resonance frequency shift and added mass (Equation 2.92):

$$\partial f = S\partial m = -\frac{1}{2}\frac{f_r}{m} K V_f \partial C. \tag{2.108}$$

Hence, the largest frequency shift is obtained for a sensor with a high sensitivity S, large partition coefficient, large volume of the functionalization layer, and a high gas concentration in the vapour phase.

Typically the entire area of the sensor is coated with the functionalization layer with a uniform thickness. Thus, in order to increase the volume the thickness must be increased hereby increasing the total mass of the coating and plate. This increase of mass decreases the sensitivity of the sensor and hence there is a trade-off between increasing the volume of the functionalization layer and decreasing the sensor sensitivity. A thicker functionalization layer also increases the response time since it will take longer to reach equilibrium, that is reach the concentration ∂C_f in the layer.

In the Introduction chapter the figure of merit of the minimum detectable concentration was the mass limit of detection per area. The reason for this can be seen by combining Equation 2.107 and Equation 2.96, yielding:

$$K V_f \partial C = 2 m_{\text{eff}} \sigma_y \Leftrightarrow \tag{2.109}$$

$$\partial C_{\min} = \frac{2 m_{\text{eff}} \sigma_y}{K V_f} = \frac{\text{LOD}}{K V_f} \propto \frac{\text{LOD}}{A}, \tag{2.110}$$

where A is the coated active area of the resonator. The minimum detectable concentration of a gas is therefore proportional to the mass LOD over the resonator area (LOD/A). Compared with the other gravimetric sensor types the CMUT has a favorable combination of parameters [11] which results in a low LOD/A value, see Table 2.7 that is a repetition of Table 1.3. Even though the Allan deviation is the highest (worst) for the MEMS and NEMS sensor types the mass LOD is lowest (best) since the mass of the resonator scales with the volume. But as Equation 2.109 showed the important figure of merit for gas sensors is LOD/A which does not favor the M/NEMS sensors in terms of surface area. Consequently, the CMUT with its relatively large surface area, low Allan deviation, and medium effective mass offers a low LOD/A, which from a *concentration* LOD point of view is attractive.

Table 2.7: Resonance frequency, Allan deviation, active surface area A and LOD/A for different gravimetric sensor types. This table is based on [11].

Sensor type	f_r [MHz]	Allan deviation, σ_{min}	Active surface area, A [mm^2]	LOD/A [ng/cm^2]
BAW	1100	10^{-7}	0.25	1
SAW	500	10^{-8}	4	0.03
CMUT	8	10^{-8}	0.25	0.01
CMUT	47.7	3×10^{-8}	0.09	0.003
MEMS	3	10^{-7}	7.3×10^{-6}	0.1
NEMS	127	10^{-6}	8×10^{-8}	0.1

2.7 Sensitivity versus limit of detection

The two important figures of merit: sensitivity and limit of detection have now been introduced. Obviously, a high sensitivity and a low LOD is wanted but the design rules for these two figures of merit can affect each other, resulting in a compromise.

The LOD and especially the Allan deviation is difficult to calculate analytically. Therefore, it can be instructive to look at empirical values for the Allan deviation. Figure 2.22 shows a plot of the Allan deviation as a function of resonator mass for many different resonator types, from [16]. The tendency in the plot is clear: as the masses of the devices decrease the Allan deviation increases due to a decrease in frequency stability. The red line is proportional to $m^{-1/2}$ which for these empirical data fits well over many decades of mass. An approximate range of typical CMUT masses and Allan deviations has been inserted in the figure. Interestingly, the Allan deviations in the gray area in the figure, lie below the red line, demonstrating the high frequency stability of the CMUTs. This is in part due to the aforementioned parallelism of the CMUTs lowering the noise.

In Section 2.4 it was shown that in order to increase the CMUT sensitivity the mass of the plate should be decreased. This scaling is the opposite of what is wanted in order to obtain a low Allan deviation, as was just shown. But the limit of detection is not only given by the Allan deviation but also by the mass of the device itself (Equation 2.96): $\text{LOD} = 2m\sigma_y$. Inserting the empirical scaling for the Allan deviation in this equation gives:

$$\text{LOD} \propto m \cdot m^{-1/2} = m^{1/2}, \tag{2.111}$$

so an overall low LOD is obtained by decreasing the device mass which is the same that is required for a high sensitivity. From these considerations an optimal design, for both the LOD and sensitivity, is one where the mass of the plate is decreased. The mass of the plate can be decreased in three ways:

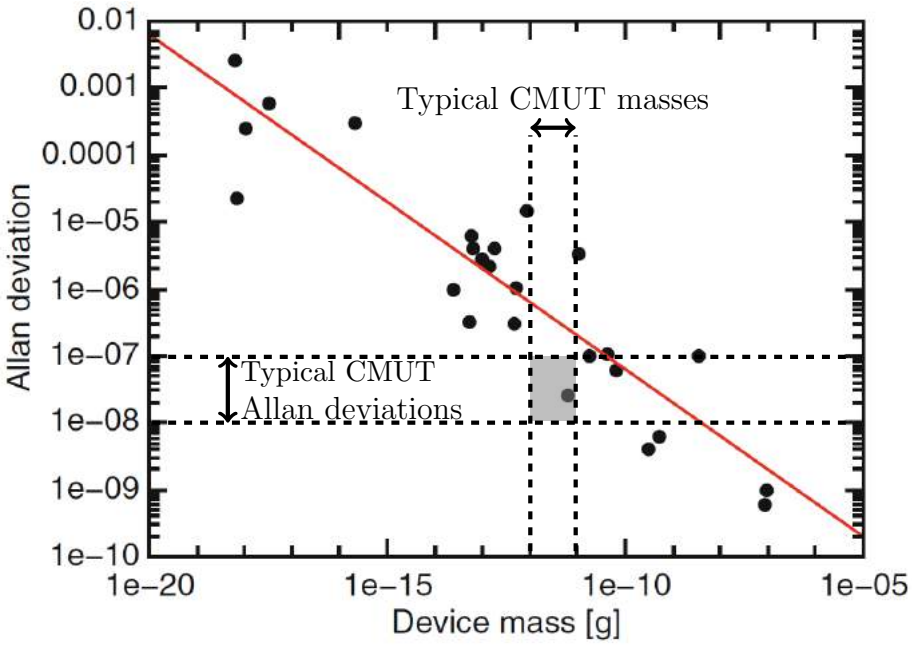


Figure 2.22: Empirical Allan deviations as a function of device mass for different resonator types. The range of typical CMUT masses is indicated by the two dashed lines. Figure taken from [16].

decreasing the radius, decreasing the thickness, or choosing a material with a lower density. For plates made of an identical material, designing a device with a thinner plate rather than decreasing the radius is attractive from a signal and SNR point of view. The reason being that a smaller radius results in a smaller plate area and thus a lower cell capacitance but the parasitic capacitance will be the same, and hence the SNR will drop. Furthermore, decreasing the plate thickness increases the normalized sensitivity.

2.8 Chapter conclusion

In this chapter the theoretical background of the CMUTs has been presented. The theoretical models were used to derive expressions for the important parameters such as: resonance frequency, pull-in voltage, capacitance, electromechanical coupling coefficient, and impedance spectrum. Furthermore, the two sensor figures of merit: mass sensitivity and limit of detection were investigated and discussed and CMUT design guidelines were derived.

CHAPTER 3

Design

In this chapter the design and modelling of the CMUT devices, fabricated during this project, are presented. First, the four generations of CMUTs that was made are introduced. In order to discuss the interplay between their designs and fabrication, their process flows are presented along with a discussion of the choice of plate material. Next, a section, based on my article [54], shows how the final CMUT cavities are affected by the choice of materials during fabrication. Furthermore, it is shown how the vertical lengths in the cell influence the final horizontal lengths in the cavity, which can lead to situations where the CMUT is not able to be fabricated. The Finite Element Method (FEM) model used to find the resonance frequency, pull-in voltage, and sensitivity, of a given cell design, is presented. In addition, a mesh convergence test is performed for the FEM model and the model is validated by comparing the results with the analytical expressions found in the previous chapter. Next, the cell designs are given for the four CMUT generations and the calculated resonance frequencies, pull-in voltages, and sensitivities are presented. Finally, the influence of a number of parameters, including: the number of cells in an element, top electrode design, cavity/cell type, and inter cell distance, on the capacitance is investigated.

N.B. the variable names of the vertical lengths such a the plate thickness is in this chapter denoted using t and not h.

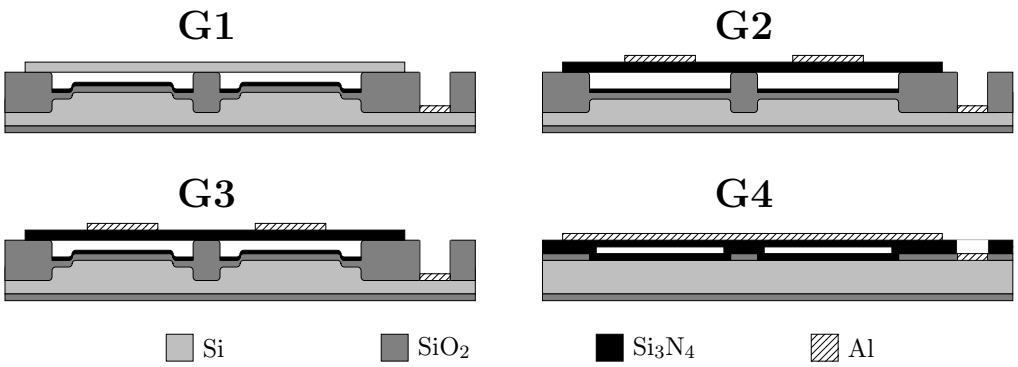


Figure 3.1: Cross sections of the four generations G1-4 of CMUTs designed and fabricated in this project. The generations are listed in chronological order as they were made. G1 is a double LOCOS cavity with a Si plate, G2 is a single LOCOS cavity with a Si_3N_4 plate, G3 is a double LOCOS cavity with a Si_3N_4 plate, and G4 is a RIE cavity with a Si_3N_4 plate.

3.1 CMUT generations

Four generations of CMUT devices have been designed and fabricated during this project. Figure 3.1 shows cross-sectional sketches of the CMUT cells from generation one to four: G1, G2, G3, and G4. The generations are listed in chronological order, hence the name. Generally, the mass sensitivity is sought increased from G1 through to G4, by decreasing the mass of the plate. Which at the same time decreases the LOD, which was shown in Section 2.7. Therefore, the two important figures of merit, mass sensitivity and LOD, are sought improved in the designs through the generations.

All devices in Figure 3.1 are fabricated using wafer bonding, which for CMUTs was first introduced by Huang et al. [47]. The advantages of the wafer bonding fabrication process over the sacrificial release process are: better gap height control, the possibility to fabricate small ($< 100\,\text{nm}$) gaps, and shorter turn-around time due to fewer process steps [46, 47]. The control and uniformity of the vacuum gaps are better for the wafer bonding process as the vertical dimensions are determined by thermal oxidations which both allows for very uniform layers over a wafer and excellent control of the height [42].

Furthermore, fabrication challenges associated with the sacrificial release of low frequency big cavities such as stiction/unwanted irreversible collapse is avoided. For small cavities ($a < 5\,\mu\text{m}$), typically used for gravimetric CMUT sensors, the sacrificial release process suffers from a low fill factor since the etch holes begin to take up an increasingly big fraction of the area when the cell radius is decreased [47]. A drawback of the wafer bonding process is

that voids, that is areas of poor or no bonding, are almost unavoidable when fusion bonding two wafers. Even when extreme caution is taken to perform the wafer bonding step in a clean and contaminant free environment, perfect bonds are difficult to achieve and thus the fabrication yield will decrease. This is in particular a problem for large CMUT chips where a single void can ruin an entire chip.

Two major differences can be seen between the devices in Figure 3.1: the plate material and the way the cavities are formed. The plate material is highly doped conducting Si for G1, while it is Si_3N_4 for G2-4. The Si plate is a thinned down device layer from a Silicon On Insulator (SOI) wafer, which were and still are used for defining the plates, in our research group. In Section 3.1.3 the choice of plate material is discussed in depth. The other major difference between the devices in the figure is the geometry of the cavities. The cavities in device G1 and G3 are fabricated using a double LOCOS process, G2 a single LOCOS process and G4 using a RIE process. The advantages and disadvantages of these processes will now be presented.

3.1.1 Cavity structure discussion

The advantage of the RIE process is a short fabrication time, in part due to the single oxidation step. In addition, the vertical dimensions are not coupled to the horizontal dimensions as is the case for the single and double LOCOS process. This only becomes an issue for small cavities and will be discussed more in Section 3.2.

The single LOCOS cavities (Figure 3.1 G2) have several advantages over the RIE cavities, including a lower parasitic capacitance due to the higher post oxide to vacuum gap ratio and a higher resilience towards electrical break-down in the post area.

Finally, the double LOCOS process offers an even lower parasitic capacitance (and thus higher coupling factor) than the single LOCOS structure due to the central Si bump. Also, the Si bump allows the vacuum gap and post oxide height to be chosen independently which gives a higher degree of design flexibility. A high capacitance and a high coupling factor can hereby be achieved at the same time because small gaps and high post oxides can be obtained simultaneously. The double LOCOS process requires an additional photolithographic mask compared with the single LOCOS and RIE process, which together with the additional LOCOS process step increases the fabrication time [42].

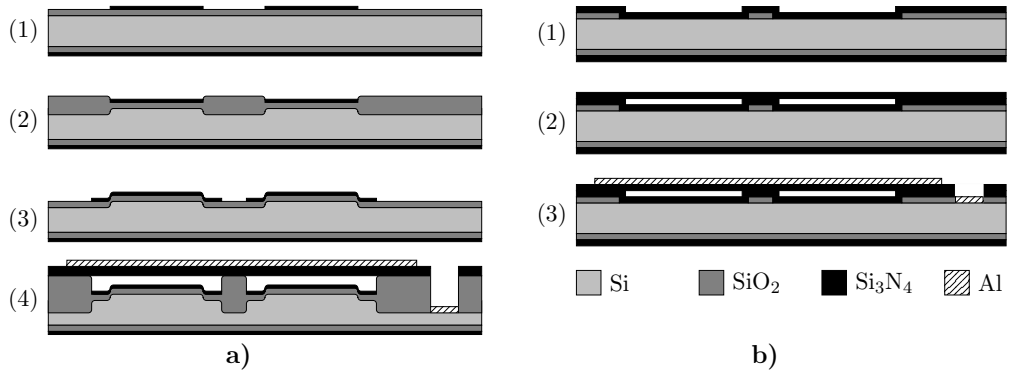


Figure 3.2: Process flow of the a) double LOCOS and b) RIE CMUT. In a) step (1) a thermal pad SiO_2 layer is grown and subsequently a LPCVD Si_3N_4 layer is deposited and patterned by a RIE etch. In step (2) the LOCOS is performed, forming the Si bumps. Step (3) is a repeat of step (1). Finally, in step (4) the second LOCOS is performed and the cavities are sealed under vacuum by fusion bonding a Si_3N_4 layer to the cavity wafer. Openings to the bottom electrode are made and Al is deposited and wet etched. In b) step (1) a thermal SiO_2 is grown and patterned by a RIE etch. Next a LPCVD Si_3N_4 layer is deposited and the subsequent steps are the same as in the LOCOS process.

3.1.2 Fabrication process flow

In this section an overview of the the process flows of the LOCOS and RIE CMUTs is presented. The full detailed review of the process flows is given in Chapter 4. Only two process flows need to be introduced, although three different CMUT structures were presented in Figure 3.1, due to the similarity of the single and double LOCOS process. Figure 3.2 shows the process flow for the a) double LOCOS and b) RIE process. The single LOCOS process is the same as in a) but with step (3) skipped.

Double LOCOS CMUT process flow

In the original LOCOS process by Park et al. [42] the first oxidation mask was a thermally grown SiO_2 layer, while in this process it is a Si_3N_4 layer. The choice of masking material becomes increasingly important as the dimensions of the CMUTs decrease, as will be explained in more detail later. In Figure 3.2 a) step (1) a thermal pad SiO_2 layer is grown by dry oxidation and Si_3N_4 is deposited by Low Pressure Chemical Vapor Deposition (LPCVD). The pad SiO_2 acts as a buffer layer between the tensile stressed Si_3N_4 layer and the silicon surface. Subsequently, the Si_3N_4 layer is patterned by RIE dry etching using an UV photoresist mask. In step (2) the first LOCOS is

performed by wet oxidation to form the Si bumps and afterwards all layers but the silicon are stripped by a buffered HF (BHF) etch and a H_3PO_4 etch. Another pad SiO_2 is grown in step (3) and Si_3N_4 is deposited using LPCVD and wet etched using a poly-silicon etching mask. The second LOCOS is performed in step (4) hereby forming the cavities. The substrate wafer is bonded, under vacuum, to a silicon wafer with a Si_3N_4 layer constituting the plate. The bonded wafer stack is annealed to strengthen the bond and the silicon handle wafer is selectively etched by KOH using the Si_3N_4 plate layer as an etch stop layer. Furthermore, openings to the bottom electrode are dry etched and Al is deposited by e-beam deposition and subsequently etched in a solution of $H_2O : H_3PO_4$, hereby patterning the metal electrodes.

RIE CMUT process flow

In Figure 3.2 b) step (1) a thermal SiO_2 layer is grown and patterned by dry etching, forming cavities. A Si_3N_4 layer is deposited using LPCVD, thus covering both sides of the wafer. In step (2) the cavity wafer is bonded to a silicon wafer with a Si_3N_4 layer, constituting the plate. All subsequent process steps are identical to the steps described for the LOCOS process.

3.1.3 Plate material

Historically, Si plates have been used in conjunction with the wafer bonding fabrication process and Si_3N_4 plates with the sacrificial release process. However, several groups have presented Si_3N_4 plate CMUTs fabricated using wafer bonding with the Si_3N_4 layer bonded to either a SiO_2 layer [51–54] or to another Si_3N_4 layer [54, 88]. Figure 3.1 showed that the first CMUT generation G1 in this project has Si plates while G2-4 have Si_3N_4 plates. In the following the considerations behind the choice of plate material for gravimetric CMUT sensors is given.

Availability and price

As already discussed (Section 2.4) thinner plates are desired with regard to the mass sensitivity. The Si plate is typically made up of the device layer from a SOI wafer but commercial 4″ SOI wafers with a device layer thickness of below about a couple hundred nm are not readily available for researchers [52]. Furthermore, the price is in the order of $200 - 300$ USD which is based both on our own experience and falls in line with the price cited in [52].

In contrast, fabricating Si_3N_4 plates on a Si carrier/handle wafer using a LPCVD process both ensures the availability in the desired diameter, and

layer thickness and a low price since cheap single side polished standard high resistivity Si wafers can be used, which cost a fraction of the SOI wafer price: ~ 20 USD per wafer.

Thickness uniformity

The wafer to wafer and intra-wafer thickness uniformity becomes increasingly important as the plate thickness decreases.

The thickness uniformity of the device layer of 4″ SOI wafers are typically between $\pm 0.3\,\mu$m and $\pm 0.5\,\mu$m which is greater than a typical desired plate thickness of e.g. 100 nm. This limits the usability of these standard SOI wafers. The wafer to wafer thickness uniformity of LPCVD (Low Pressure Chemical Vapor Deposition) deposited Si_3N_4 layers, as measured in our cleanroom, is maximum 10 nm over a boat of 25 wafers.

Figure 3.3 shows ellipsometer (Ellipsometer VASE, J.A. Woollam Co., Inc.) measurements, specifically a thickness map of a Si device layer for a SOI wafer a) and b) a LPCVD Si_3N_4 layer. These measurements are examples of the intra-wafer thickness uniformity. The red dots are measurements where the mean square error (MSE), from the fitting routine, is larger than 1.5 times the median MSE and thus have been rejected. The SOI wafer is bought from the company Okmetic and has a nominal device layer thickness of $2\,\mu\text{m} \pm 0.5\,\mu$m while the Si_3N_4 layer is deposited using an LPCVD furnace in our cleanroom. The relative measured thickness range for the SOI device layer and the Si_3N_4 layer is 42% and 1.9%, respectively. This shows that the intra-wafer thickness uniformity is much better for the Si_3N_4 layer.

Sensitivity

An expression for the sensitivity was given in Section 2.4 and in general a high mass sensitivity is obtained for a low plate mass and a high resonance frequency. For highly doped Si layers the resistivity can be low enough so that no metal electrode is needed on the plate over the cavity. This decreases the mass of the plate compared with the Si_3N_4 plate which is an insulator and thus non-conducting. Hence, the Si_3N_4 plates must have a conducting layer deposited which then increases the mass of the plate and hereby decreases the sensitivity.

However, the Si_3N_4 layers used in this project have a built-in tensile stress, which in Section 2.4 was seen to increase the sensitivity since the resonance frequency was increased. The normalized sensitivity on the other hand, was seen to be independent of the stress.

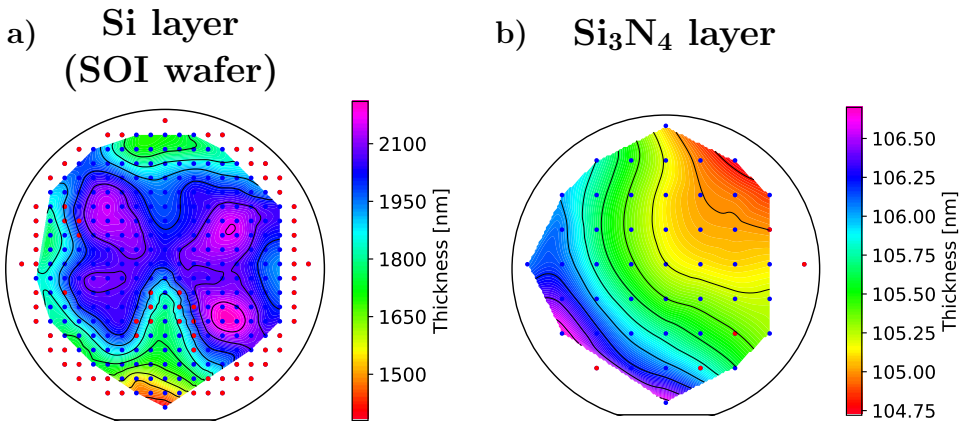


Figure 3.3: Ellipsometer thickness measurements over a wafer of a) the device layer of a SOI wafer and b) a LPCVD Si_3N_4 layer. The red dots are measurements that has been rejected on the basis of a too high mean square error from the fitting routine.

Parasitic capacitance

A compelling argument for using an insulating plate material such as Si_3N_4 is that the parasitic capacitance is reduced since the top electrode can be patterned independently of the plate. This means that a metal top electrode layer can be located primarily over the cells and removed in the area around the cells. To electrically connect the cells, thin metal connectors can be used in order to decrease the electrode area outside of the cells, hereby decreasing the parasitic capacitance. This kind of design is for example utilized in [52–54] and will be shown later in this chapter.

Manufacturability

A final important point is the difference in manufacturability between a CMUT with a Si or Si_3N_4 plate using the wafer bonding fabrication process. This boils down to how easy it is to bond the plate to the cavity wafer which ultimately is reflected in the yield of the process. In the classical LOCOS process [42] a Si plate is bonded to a SiO_2 surface. This has since been repeated several times, e.g. in [43]. Si_3N_4 plates have been bonded to Si_3N_4 [54, 88] and SiO_2 [51–54] cavity wafers. Chen et al. [52] found that a Chemical Mechanical Polishing (CMP) step was needed before a 450 nm LPCVD Si_3N_4 layer could be bonded to a SiO_2 layer. Mølgaard et al. [53,54] found that a CMP step was not required when bonding thin (100nm and 50nm) LPCVD Si_3N_4 layers to a SiO_2 and Si_3N_4 surface, respectively. This

agrees well with what is found by Reck et al. [89], who found that surface roughness is limiting the ability to bond LPCVD Si_3N_4 surfaces when the thickness exceeds about 300 nm.

We found that fusion bonding of two stoichiometric Si_3N_4 high stress layers empirically is the most difficult, resulting in the lowest yields. The second most difficult fusion bond is between a stoichiometric Si_3N_4 high stress layer and a thermal SiO_2 surface and finally the highest yields are obtained when bonding a Si surface to either Si or SiO_2.

In conclusion, Si_3N_4 plates are good alternatives to Si plates for resonant gravimetric CMUTs due to the Si_3N_4 plates being more readily available, cheaper, having a better thickness uniformity, higher sensitivity and having a lower parasitic capacitance. However, using Si_3N_4 plates can make fabrication more challenging.

3.2 Design of LOCOS cavities

In this section the design space of cavities made with the LOCOS process is presented. The section is based on the article [54], Appendix A. The reason for the special treatment of the LOCOS cavities is that the horizontal dimensions are coupled to the vertical dimensions, which is not the case for the RIE cavities. This makes designing the LOCOS CMUTs more challenging, and a design can in some cases be impossible to fabricate as will be shown. As an introductory example Figure 3.4 shows two simulated cavities made using a double LOCOS process but the oxidation masking material for the first LOCOS process, defining the Si bump, was a) SiO_2 and b) Si_3N_4. The Si bump made using the a) SiO_2 is seen to have a much less well-defined slope profile compared with the Si bump made with a Si_3N_4 mask in b). Figure 3.5 shows two AFM scans of Si bumps made with a) a SiO_2 mask and b) Si_3N_4 mask with the same dimensions for the masking layers as in Figure 3.4. The same tendency is seen for the bumps in the AFM scans as in the simulation profiles. The examples show that the choice of material can have an effect on the dimensions of the cavities which in turn affects the CMUT device characteristics. This warrants a further investigation which is presented below.

3.2.1 Introduction and motivation

In order to increase the sensitivity the radius of the CMUT must be decreased, as shown by Equation 2.93. For LOCOS cavities the smaller radii

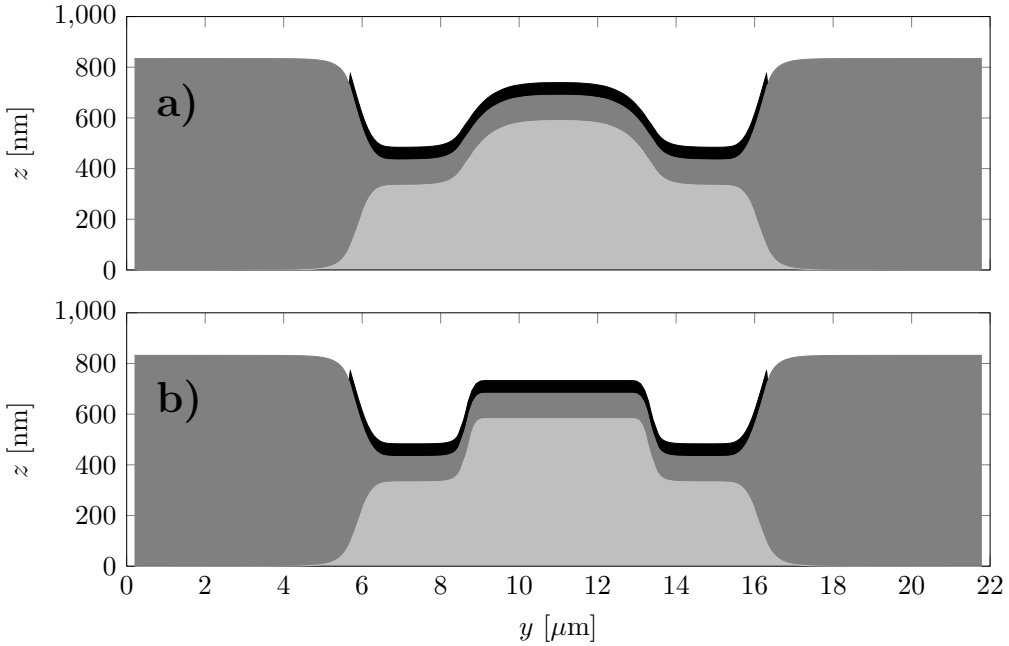


Figure 3.4: Cross sections showing simulation results of cavities made using the double LOCOS process where the Si bump in a) is defined using a SiO_2 oxidation mask with a height of 1 μm and b) a Si_3N_4 oxidation mask with a pad SiO_2 layer with thickness $t_{\text{pad}} = 10\,\text{nm}$. The cavities in a) and b) are made with an identical process and thus only the Si bump dimensions differ.

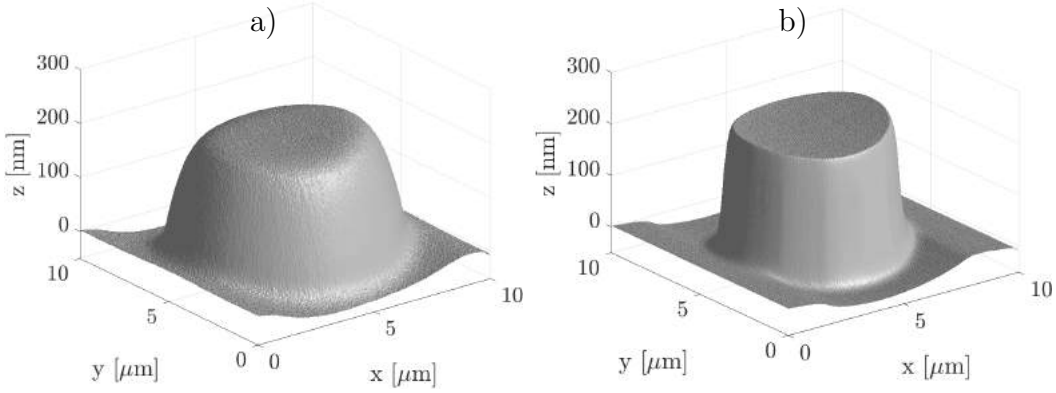


Figure 3.5: AFM profiles of two Si bumps with all dielectric layers removed. Figure a) is fabricated using a 1 μm thick SiO_2 mask. Figure b) is fabricated using a 50 nm thick Si_3N_4 mask with a pad oxide thickness of $t_{\text{pad}} = 10\,\text{nm}$.

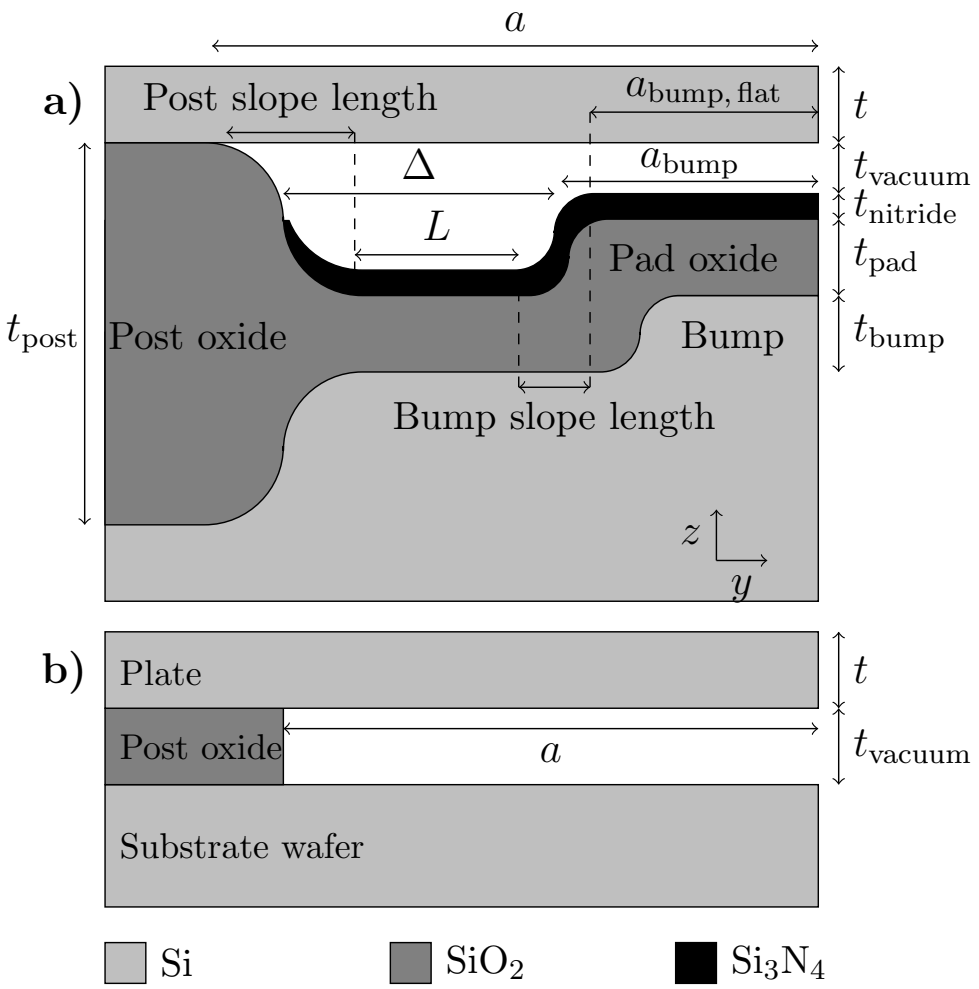


Figure 3.6: Cross-sectional sketches of half a) a LOCOS CMUT cell and b) a RIE CMUT cell. The parameter Δ denotes the designed distance on the lithography masks between the cavity edge and Si bump. L defines the length between the points at which the slopes are below 1%. Notice how L begins where the post slope length ends and ends where the bump slope length begins.

can become comparable with the lateral width of the bird's beak SiO_2 which in Figure 3.6 a) is defined as the 'slope length'. This sets a lower limit for the radius of the cavity and the sensitivity can hereby be limited by the process parameters that affect the slope length of the bird's beak. The bird's beak structure have previously been studied extensively due to its use as an isolation structure in semiconductor manufacturing [90–92], since reduction of the horizontal size has been especially important due to the decrease in device dimensions. However, for small radii ($a < 5\,\mu$m) CMUT cells the more complex geometry has yet to be investigated.

Furthermore, not all CMUT designs can be fabricated. The main reason being the inability to bond the plate to the cavity wafer. For the RIE process it has previously been shown by Sarioglu et al. [93] how oxidation of convex silicon corners can lead to protrusions in the SiO_2 at the corners which can hinder bonding. A study by Christiansen et al. [94] showed how the corners of a structured SiO_2 layer are lifted when the wafer is re-oxidized in order to e.g. grow an insulation SiO_2 layer, consequently making bonding impossible. Likewise, some LOCOS designs will hinder bonding but here the problem is slightly different as the solution depends on choosing the right combination of design and process parameters, where for the RIE process the problems are solved by adjusting the process flow. The main challenge, when designing the LOCOS device, from a fabrication point of view, is keeping the Si_3N_4 layer from protruding above the post SiO_2 at the cavity edge, thus inhibiting bonding. Figure 3.7 shows three simulated LOCOS devices where (a) is an example of a device that can be fabricated, (b) cannot be fabricated due to the protruding Si_3N_4 layer caused by an overlapping post SiO_2 and silicon bump, either as a result of a bad design choice or misalignment in the lithography step. Finally, in situation (c) the post SiO_2 is too low, which e.g. could happen if a very small gap is wanted. The validation of this simulation model is now shown, followed by a presentation of the design space for the LOCOS cavities.

3.2.2 Process simulation model validation

A numerical simulation model was made using the process simulator ATHENA (Silvaco, Inc., California) where part of the LOCOS process in Figure 3.2 a) was simulated. The mesh was optimized so the smallest mesh sizes were centered around the position where the bird's beak is formed. The size of the mesh in the SiO_2 has a maximum vertical length of 5 nm and a maximum horizontal length of 25 nm. These mesh sizes are chosen based on a mesh convergence study.

In the following the fabricated LOCOS structures used to verify the simu-

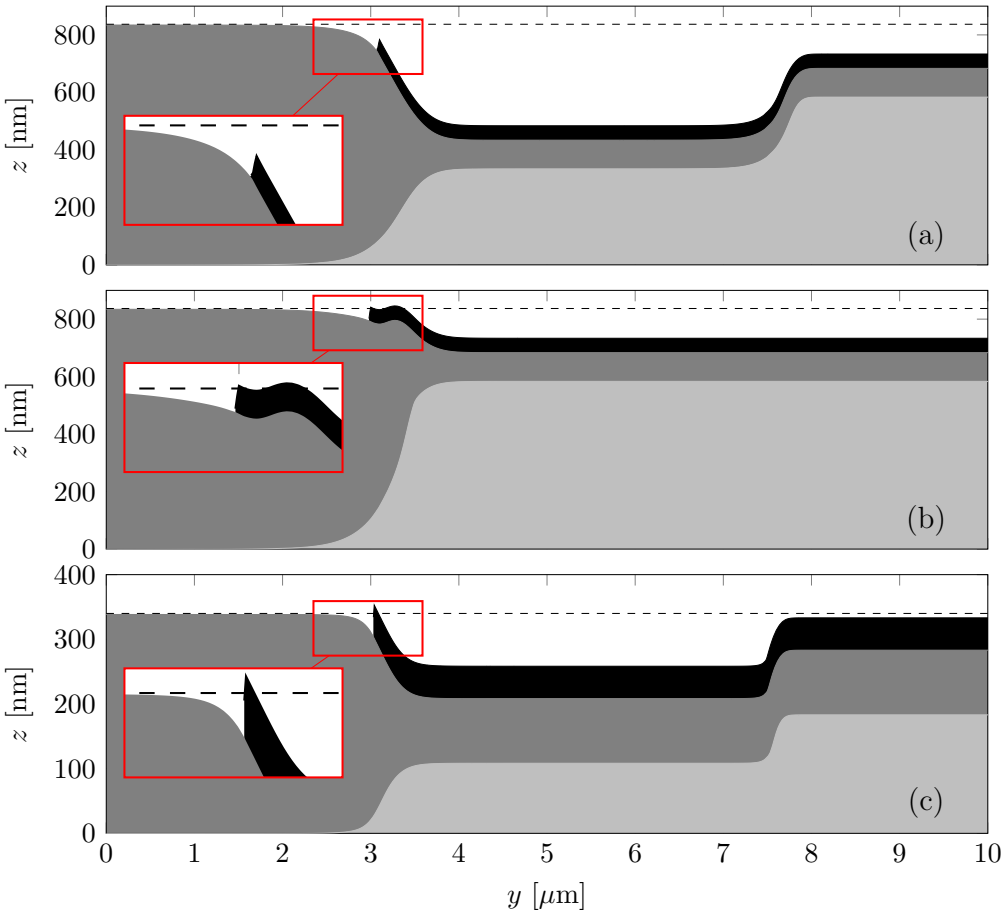


Figure 3.7: Simulation results of three different CMUT designs. **(a)** This design does not inhibit wafer bonding. The designs in **(b)** and **(c)** both inhibit wafer bonding since the Si_3N_4 layer protrudes above the post SiO_2 at $y \approx 3\,\mu$m. In **(b)** the bump is placed too close to the post SiO_2 and in **(c)** the post SiO_2 is too low.

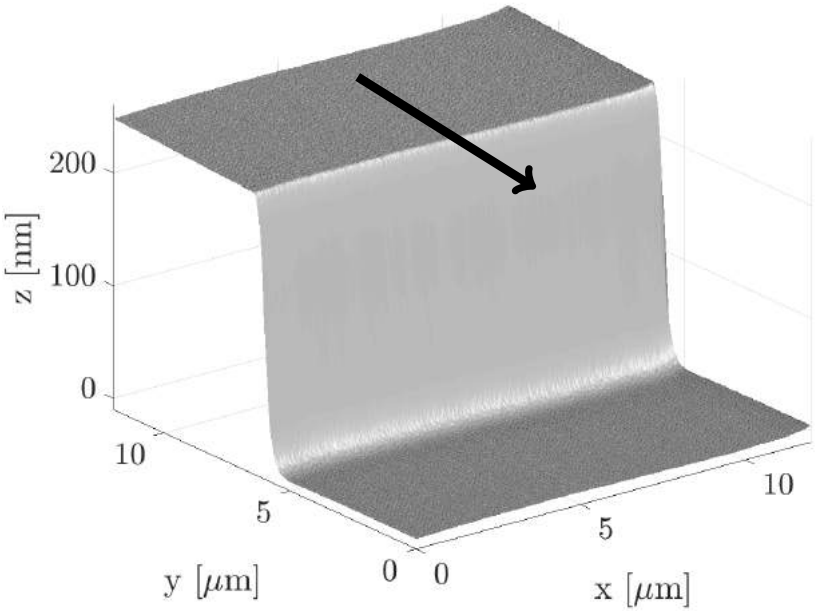


Figure 3.8: Example of an AFM scan of a square Si bump profile fabricated with a Si_3N_4 mask and a pad SiO_2 of $t_{pad} = 10\,nm$. The AFM scan lines are parallel to the arrow.

lation model are described. Specifically, square silicon bumps were fabricated using the LOCOS process following steps (1)-(2) in Figure 3.2 a), using both SiO_2 and Si_3N_4 as oxidation mask material. Table 3.1 shows an overview of the seven wafers oxidized in the experiment. The pad SiO_2 thickness and thickness of the SiO_2 mask layer were varied resulting in Si bumps of approximately the same height but with differing profiles. The wafers used were (100) oriented single side polished 4″ Si wafers with an electrical resistivity of 1-10 Ωcm. All thermal oxidations for the bump fabrication and simulation were performed in a wet atmosphere at a temperature of 1100 °C. Stoichiometric Si_3N_4 was deposited using a Low Pressure Vapor Deposition (LPCVD), resulting in a 50 nm thick layer on all the wafers with Si_3N_4 masks. Finally, all wafers underwent a LOCOS step with different oxidation durations, in order to reach a bump height of 250 nm. Subsequently, all layers were stripped from the wafer, leaving only the Si surface and the bare silicon bumps were characterized by AFM (Dimension Icon, Bruker). The maximum relative difference between the target and measured Si bump height was found to be 6.4%, which is low enough to allow for a comparison of the profiles.

The Si bump profiles were characterized using AFM and an example of a 3D surface of a Si bump can be seen in Figure 3.8. The AFM scans consist of

Table 3.1: Table of target and measured values for the seven wafers fabricated for the experimental validation of the simulation model. The first four wafers all have a Si_3N_4 mask with varying t_{pad} thicknesses, while the last three wafers have a SiO_2 mask of varying thicknesses. The largest relative difference between the target and measured target t_{bump} is 6.4%.

Parameter	**Wafer number**							
	1	2	3	4	5	6	7	Unit
t_{pad} target	0	10	50	100	-	-	-	[nm]
t_{pad} measured	0	10.6	56	99.4	-	-	-	[nm]
Mask SiO_2 thickness target	-	-	-	-	1000	2000	3000	[nm]
Mask SiO_2 thickness measured	-	-	-	-	963	1964	2950	[nm]
t_{bump} target	250	250	250	250	250	250	250	[nm]
t_{bump} measured	244	241	241	241	234	247	260	[nm]
Relative difference between target and measured t_{bump}	2.4	3.6	3.6	3.6	6.4	1.2	4	[%]

512 scan lines parallel (indicated with an arrow) to the y-axis and each scan line is made up of 512 data points. The scan area is a 10 μm x 10 μm square and no artifacts were observed due to the relatively slowly varying slope of the profiles. The tilt of the scan is removed by fitting and subtracting a first order polynomial plane from the scan.

Figure 3.9 shows an example of an experimental and simulated Si bump profile, for bumps fabricated using Si_3N_4 masks and SiO_2 masks. The measured and simulated profiles have been translated relative to each other along the y-axis to the position of minimum relative error. The relative errors between the measured and simulated profiles were found to be $\leq 5\%$. This consistently low error for both the Si_3N_4 and SiO_2 masks demonstrates the validity of the simulation model. Figure 3.9 also shows an example of the slope length, shown in Figure 3.6, of one of the profiles. The slope length is defined as the horizontal distance between the two points were the slope first becomes $< 1\%$, that is: where the profile becomes flat. The 2D AFM profiles are calculated by taking the mean of the scan lines in the direction orthogonal to the y-plane.

3.2.3 LOCOS cavity optimization and design space

It is trivial to calculate the sensitivity of a RIE CMUT as the lateral dimensions are not coupled to the vertical dimensions due to the dry etched cavities. That is, no matter the post SiO_2 height the radius will stay the same and consequently so will the sensitivity. The minimum RIE cavity ra-

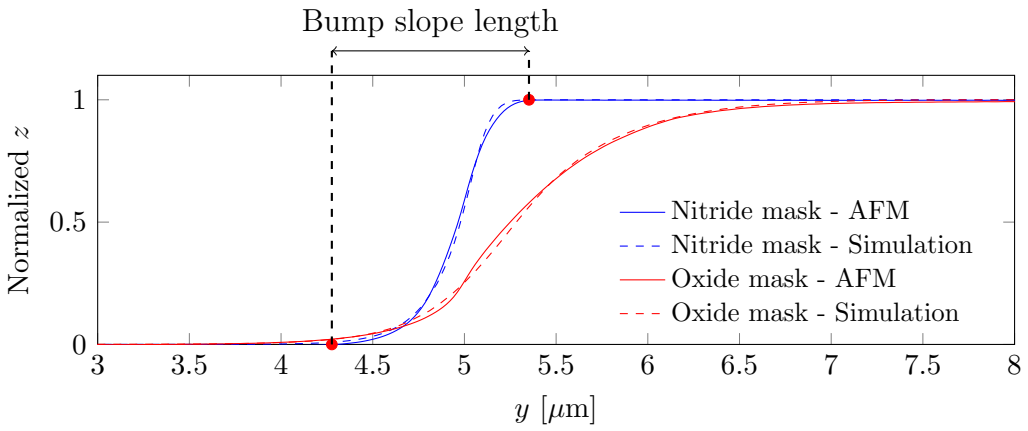


Figure 3.9: Silicon bump profiles made using either an SiO_2 mask (thickness: $1\,\mu$m) or a Si_3N_4 mask with a pad SiO_2 thickness of $t_{pad} = 10\,$nm. AFM measurements and simulation results are shown together. The distance between the two points, located at the positions where the slope is $< 1\%$, is denoted the bump slope length.

dius is therefore only limited by the minimum feature size of the lithography process.

For the LOCOS device the bird's beak profile results in a coupling between the vertical and lateral dimensions which ultimately limits the radius and hereby the sensitivity. In the following section the parameters influencing the LOCOS profile and thus sensitivity are investigated through AFM measurements and simulations. Furthermore, the limited design space is investigated due to the fabrication limitations exemplified in Figure 3.7.

To minimize the radius of the LOCOS CMUT cell and the Si bump radius, the slope length (Figure 3.6 and 3.9) should be minimized. Figure 3.10 shows that the slope length is approximately 3 times shorter for thin pad oxides ($t_{pad} = 0\,$nm to $10\,$nm) as compared with a SiO_2 mask to 2 times shorter for thicker pad oxides ($t_{pad} = 100\,$nm). In addition, Figure 3.10(a) shows that the slope length is roughly constant as a function of the thickness of the SiO_2 mask. Whereas, Figure 3.10(b) demonstrates an increasing slope length for a thicker pad SiO_2, since the diffusivity of the oxidizing species (O_2 or H_2O) is higher in SiO_2 than Si_3N_4 and a thicker pad SiO_2 layer allows for a higher lateral influx of H_2O under the mask. Likewise, increased lateral diffusion is the cause for the longer slope lengths when SiO_2 is used as the masking material compared with a Si_3N_4 mask. This agrees well with what is found in the literature [90–92]. The measured and simulated slope lengths are in agreement with each other, showing a maximum relative error of $\leq 10\%$ and demonstrating the same scaling tendencies.

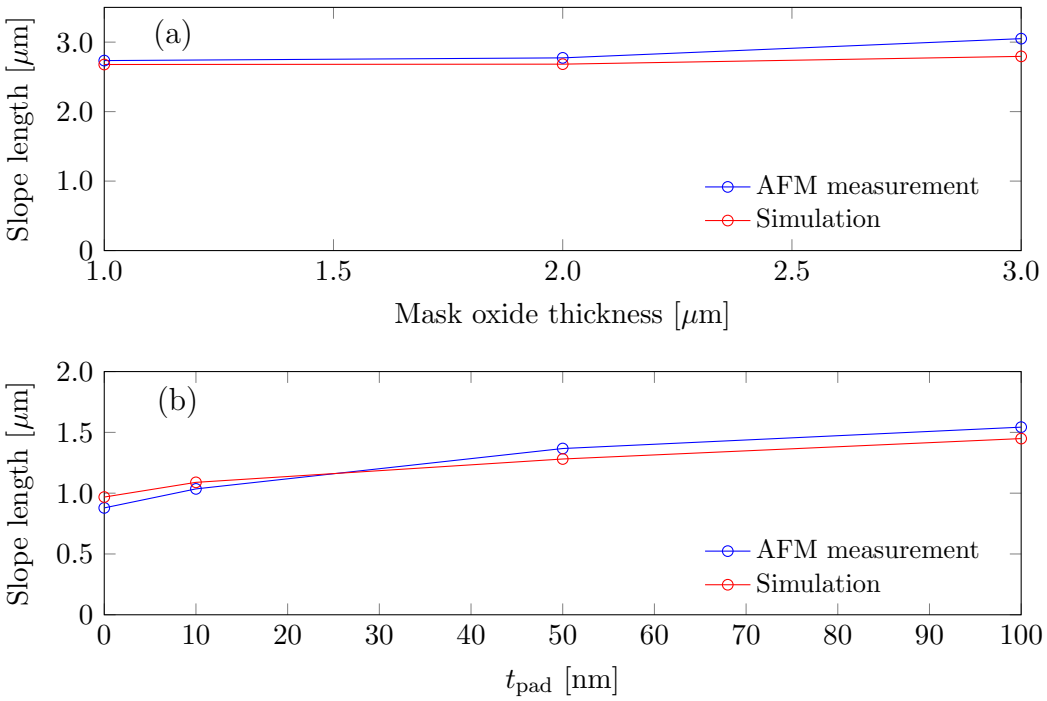


Figure 3.10: (a) Slope length as a function of the SiO_2 oxidation mask thickness. (b) Slope length as a function of the pad SiO_2 thickness, under the Si_3N_4 oxidation mask. The simulations and measurements agree with a maximum relative difference of 10 %.

When the radius of the CMUT cell decreases, the radius of the silicon bump must decrease as well and consequently the slope length of the bump will make up an increasing fraction of the total bump radius. As a result, the expected capacitance of the cell is decreased, the expected pull-in voltage is increased and wafer bonding can be rendered impossible if the slope of the bump overlaps too much with the cavity edge, as was the case in Figure 3.7(b). Thus, it is important to control the bump slope length and take it into consideration when designing the CMUT. In most cases it is desirable to minimize the slope length, hereby creating the most square-like corners for the bump. Figure 3.11 shows the ratio of the flat part of the bump to the bump radius as a function of the bump radius for different oxidation masking materials and bump heights. As the bump radius increases the ratio approaches 100%. The Si_3N_4 mask results in more well defined bumps with square-like corners where a larger fraction of the bump reaches the designed height. Furthermore, higher bumps lead to lower ratios and this difference is relatively larger for the SiO_2 mask. All in all this favors using a Si_3N_4 mask.

When designing bumps with radii in the sub $\sim 5\,\mu$m range it can therefore

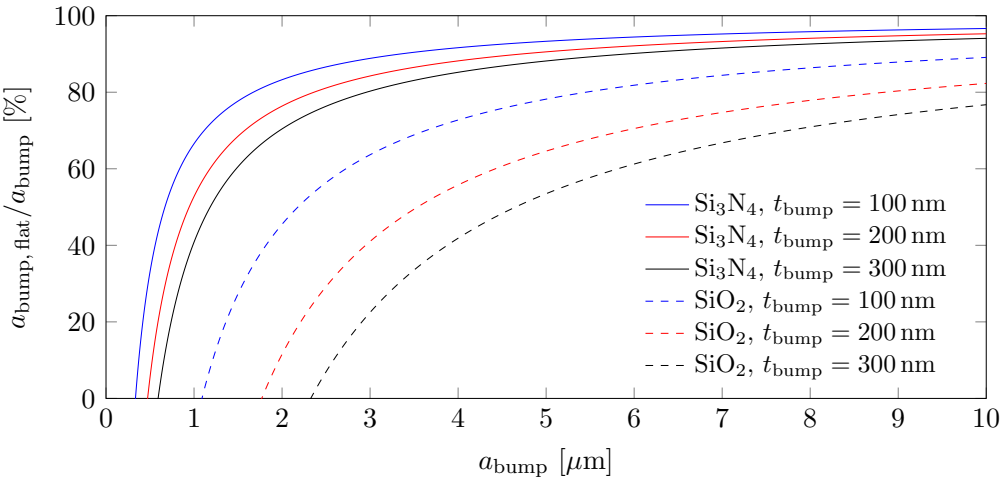


Figure 3.11: Simulation results: $a_{\text{bump, flat}}$ (see Figure 3.6) normalized to a_{bump} as a function of a_{bump} for Si_3N_4 ($t_{\text{pad}} = 10\,\text{nm}$) and SiO_2 (mask height: $2\,\mu\text{m}$) as masking material at three bump heights.

be advantageous to use Si_3N_4 as the masking material. In the article by Park et al. [42] the masking material for the first LOCOS process was SiO_2 but it was noted that Si_3N_4 could have been used at the cost of additional process steps. Therefore, the choice of LOCOS masking material is a trade-off between tight dimension control and reducing the number of process steps.

What is the minimum radius and hence sensitivity one can achieve with a LOCOS CMUT cell? In order to answer this question two limitations are imposed on the structures: the bump and the post SiO_2 should not overlap, that is $L \geq 0\,\text{nm}$ (see Figure 3.6) and the bump should at least reach the designed height in the center, that is $a_{\text{bump,flat}} \geq 0\,\text{nm}$. The minimum radius is obtained when both of these variables are zero. Figure 3.12 shows a plot of the minimum radius as a function of the vacuum gap height for three bump heights. As the vacuum gap height is increased the minimum radius increases since the post SiO_2 height increases, thus making the slope length of the post SiO_2 longer. The same effect is seen when the bump height is increased, resulting in larger minimum cavity radii. These minimum radii, given by Figure 3.12 for the LOCOS structure, directly determine the maximum achievable sensitivities.

In order to map out the design space for LOCOS cavities, simulations were made where the post SiO_2 height and bump height have been varied. Figure 3.13 shows a plot where the contour lines are the flat distance, L (see Figure 3.6), between the bump and the post SiO_2 as a function of t_{post} and t_{bump}. The plot is valid for a specific set of parameters, given in the figure caption, but the overall shape of the plot is general for all LOCOS designs.

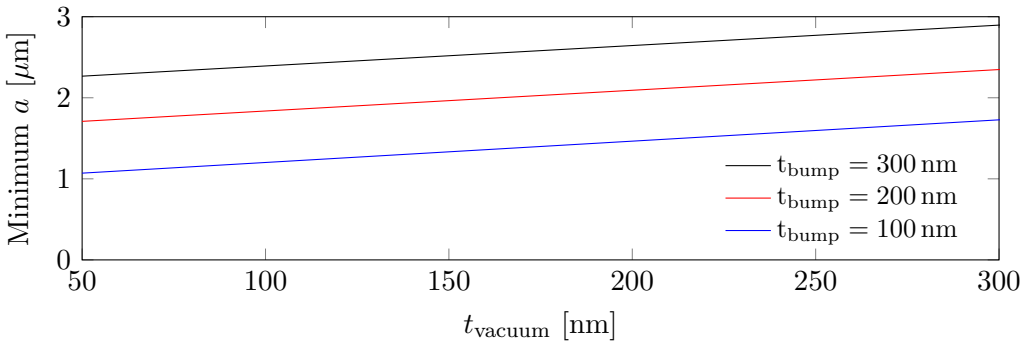


Figure 3.12: Minimum cell radius as a function of the vacuum gap height for three bump heights, when $L = 0\,\text{nm}$ and $a_{\text{bump,flat}} = 0\,\text{nm}$.

The parameter Δ is the designed distance on the photolithographic mask between the bump and cavity edge and in this plot it is $\Delta = 1.5\,\mu\text{m}$. The three colored regions (a), (b) and (c) correspond to the three situations in Figure 3.7. Region (a) is limited by the red line that denotes the minimum post SiO_2 height, the blue line at which the vacuum gap is zero and the $L = 0$-contour line which is the limit where the bump and post SiO_2 start to overlap. Thus, staying inside the green region (a) ensures a LOCOS design that can be successfully fabricated.

Figure 3.14 shows that as the bump to cavity edge distance, Δ increases, so does the design space, here shown as green areas. The solid lines are where $L = 0\,\text{nm}$, here plotted for several values of Δ. The smallest Δs are typically found for small radii cells which are here seen to be the most limited in their design space. Therefore, these design rules are especially important for these smaller radius CMUT cells, which are typical for CMUTs used for sensing, when a high mass sensitivity is wanted.

3.2.4 Partial conclusion

In conclusion, all variables that affect the LOCOS profile will ultimately affect the sensitivity, capacitance and pull-in voltage. Hence, being able to predict the effect on the final fabricated structure is important. The shortest slope length for the Si bump and post SiO_2 and hereby the smallest cell radii is obtained by using a Si_3N_4 mask with a thin pad SiO_2 (t_{pad}) and choosing a low Si bump height and a small vacuum gap, since this yields the lowest post SiO_2 height. The pad SiO_2 thickness for the first LOCOS step can be made thin as these layers are subsequently stripped. However, the pad SiO_2 thickness for the second LOCOS step is sometimes determined by the requirement for the CMUT cell to prevent electrical breakdown if pull-in

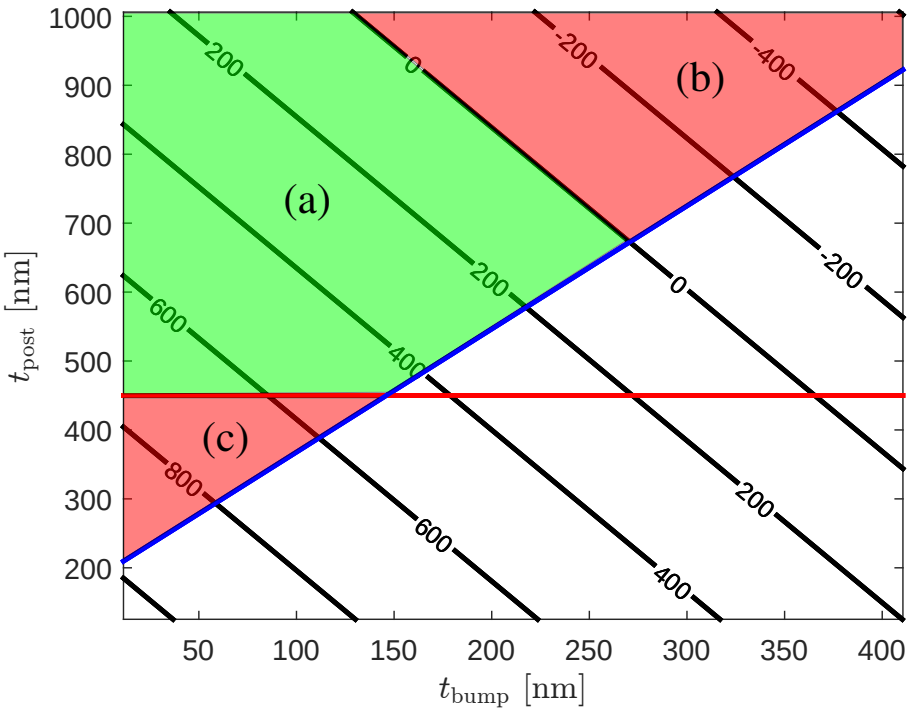


Figure 3.13: Contour plot showing the length of the flat region between the cavity and bump L (see Figure 3.6) as a function of the SiO_2 post thickness t_{post} and bump height t_{bump}. The red line gives the minimum required post SiO_2 height for successful bonding and the blue line are the points at which the vacuum gap height is 0 nm. The distance between the cavity and bump is set to $\Delta = 1.5\,\mu$m, bump and cavity Si_3N_4 thickness $t_{\mathrm{nitride}} = 50$ nm and bump and cavity pad SiO_2 thicknesses of $t_{\mathrm{pad,\,bump}} = 10$ nm and $t_{\mathrm{pad,\,cavity}} = 100$ nm, respectively.

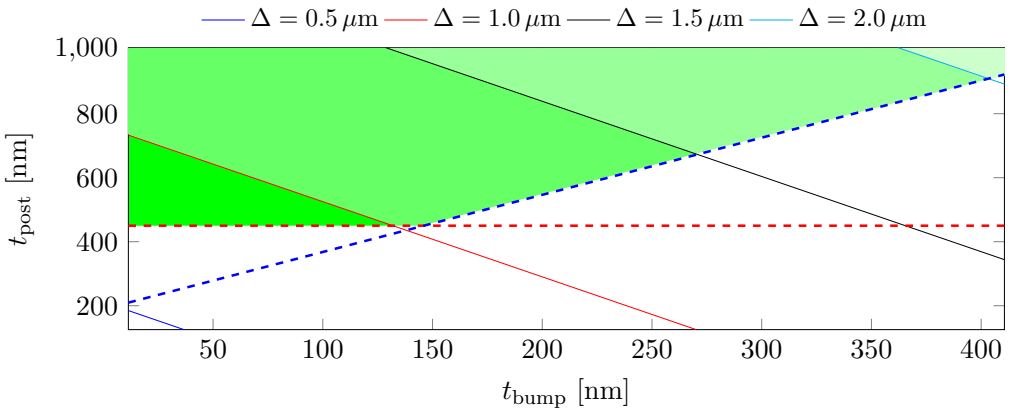


Figure 3.14: The solid lines show where $L = 0\,\text{nm}$ and are plotted for different values of Δ as a function of t_{bump} and t_{post}. The red dashed line gives the minimum required post SiO_2 height for successful bonding and the blue dashed line are the points at which the vacuum gap height is 0 nm. Si_3N_4 thickness $t_{\text{nitride}} = 50\,\text{nm}$ and bump and cavity pad SiO_2 thicknesses of $t_{\text{pad, bump}} = 10\,\text{nm}$ and $t_{\text{pad, cavity}} = 100\,\text{nm}$, respectively.

occurs. Finally, the design space of LOCOS cavities were investigated and it was shown that for high sensitivity, small radii cavities the design space is limited.

3.2.5 Vertical LOCOS cavity design

In this section the layer thicknesses in the cavity are related to each other to give two expressions: one to find the required SiO_2 thickness needed in order to get a specific Si bump height and one that can be used to find the required post SiO_2 for a given design.

The SiO_2 thickness needed in order to get a specific Si bump height is given by:

$$t_{\text{post, 1.LOCOS}} = 0.44 t_{\text{bump}}, \tag{3.1}$$

where the factor 0.44 is the fraction of SiO_2 below the original Si surface after oxidation.

The equation that gives the vacuum gap height t_{vacuum} as a function of the various layer thicknesses is now presented. The equation was first derived in [42] although with a small misprint and is here given in the corrected form:

$$t_{\text{vacuum}} = 0.56(t_{\text{post}} - t_{\text{pad}}) - t_{\text{nitride}} - t_{\text{bump}} \Leftrightarrow \tag{3.2}$$

$$t_{\text{post}} = \frac{1}{0.56}(t_{\text{vacuum}} + t_{\text{nitride}} + t_{\text{bump}}) + t_{\text{pad}}, \tag{3.3}$$

refer to Figure 3.6 for all variables. The factor 0.56 is the fraction of SiO_2 above the original Si surface after oxidation. Equation 3.2 can be used for designing the layer thicknesses in the cavity. The equation was used to determine when the gap was zero in Figure 3.13 and 3.14 shown as the blue line in the previous section. Furthermore, Equation 3.3 has been used to determine the required post oxide heights of the four CMUT generations.

3.3 Finite element modeling

In this section the finite element model (FEM) used to design the CMUTs is described. Specifically, the software COMSOL Multiphysics ®v5.3 (COMSOL AB, Stockholm, Sweden) was used with the Electromechanics Module which can couple an electrical domain to a mechanical domain. Thus, the acoustical domain is not included which, as per the previous discussion, corresponds to operating the CMUT in a vacuum. Furthermore, this implies that the damping of the CMUT in the model does not include the acoustical/medium damping. Since the CMUT is to be operated in air the effect on the resonance frequency and pull-in voltage due to the missing medium damping is minimal but the computational time for solving the model is greatly reduced. The main purpose of the model is to find the resonance frequency and pull-in voltage of a CMUT design, but any variable of interest can be extracted such as e.g. the deflection profile, electrical field, or capacitance. This can for simple geometries also be calculated using analytical expressions as the ones presented in Chapter 2, but as the geometry gets increasing complex so does the expressions and a FEM model is warranted. In addition, the mechanical clamping condition at the interface between the plate and supporting structure is more realistically modeled in the FEM model as a real plate is not perfectly clamped right at the edge of the cavity. Finally, the FEM model includes nonlinear effects such as spring stiffening where the lumped spring no longer follows Hooke's law. Shear deformation is also included, which becomes non-negligible when the aspect ratio of the plates decreases, as will be shown later in this section.

Figure 3.15 shows the structure of the CMUT COMSOL model. The model is two dimensional and axisymmetric around the red axis of rotation. As discussed above the plate is not clamped at the cavity edge but rather the entire device is clamped along the blue lines shown in the figure. This allows the post oxide near the plate at the edge of the cavity to be strained and the overall effect is that of a plate with a slightly increased radius [95]. Two electrodes are defined by green lines in Figure 3.15 for the top and bottom electrode, respectively. The extent of the top electrode follows the extent of

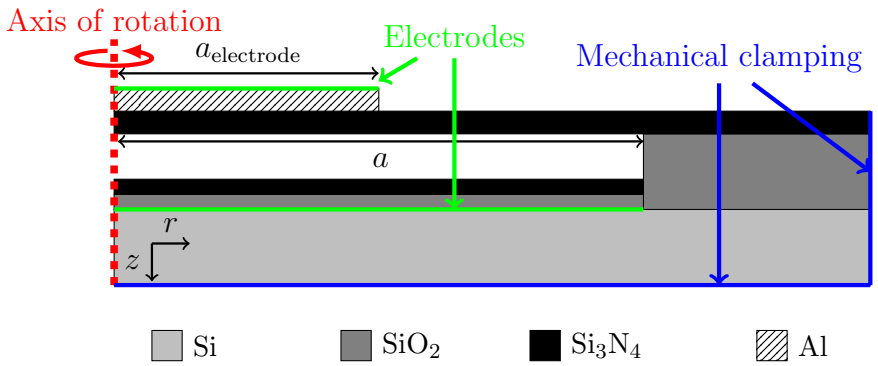


Figure 3.15: Cross-sectional sketch of half a CMUT cell, demonstrating the basic 2D axisymmetric structure of the FEM model. The red dashed line indicates the axis of rotation, the green lines show where the electrodes are defined, and the blue lines show where the structure is mechanically clamped.

the Al layer. The model is easily modified to encompass all the geometries shown in Figure 3.1.

The resonance frequency and pull-in voltage are found by running an Eigenfrequency study and Stationary optimization study, respectively. As discussed in Section 2.1 and 2.2.2 the eigenfrequency is a good approximation for the resonance frequency for slightly damped resonators, e.g. a CMUT operating in air. The spring softening effect is also included, as will be shown later, and so is the built-in stress of the Si_3N_4 layers. To find the pull-in voltage the model calculates the DC voltage that must be applied to the CMUT in order to achieve a certain center deflection. By varying the deflection set point the COMSOL Optimization module can determine a minimum voltage which is equal to the pull-in voltage. Finding the pull-in voltage in this way, instead of varying the DC voltage, ensures that pull-in does not occur which crashes the simulation model.

3.3.1 Mesh convergence study

The mesh is an important aspect of any finite element model. In this model each part of the structure (plate, vacuum gap, etc.) can be meshed with a different mesh density. The mesh used here is a so called *mapped mesh* with square mesh elements. The number of mesh elements along the sides of the rectangles in Figure 3.15 determines the total number of mesh elements in that specific rectangle. In the following a mesh convergence study is performed in order to determine the optimal mesh element size/number of mesh elements. The optimum is a trade-off between model accuracy and computation time. The two most important output variables: f_0 and V_{pi}

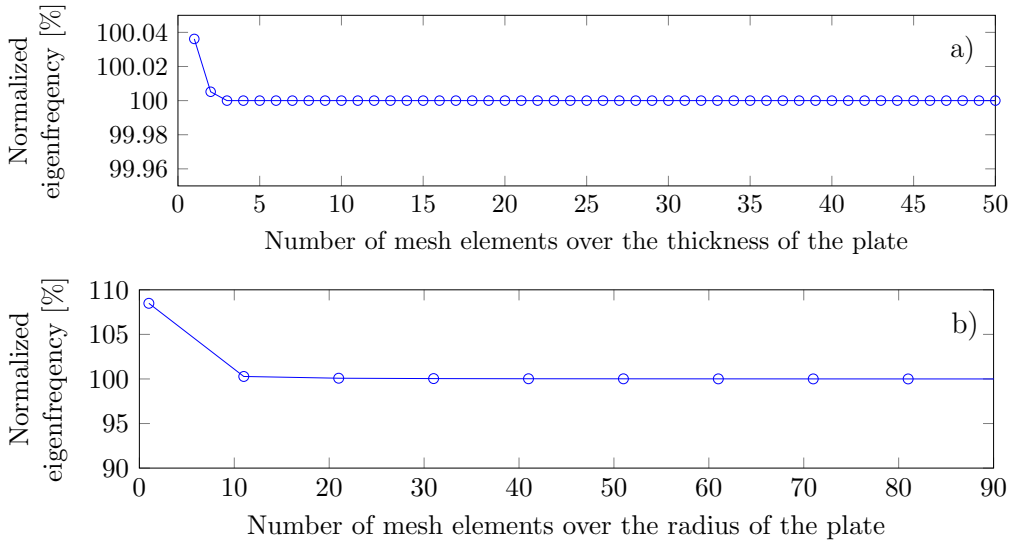


Figure 3.16: Simulated normalized eigenfrequencies as a function of a) the number of mesh elements over the thickness of the plate and b) the number of mesh elements along the radius of the plate.

are evaluated while the mesh is varied. Since the number of ways the mesh can be varied is large, only the study of the two most important domains is shown here, but similar studies have been conducted for all domains in the model.

Figure 3.16 a) shows the normalized eigenfrequency as a function of the number of mesh elements along the thickness of the plate. The eigenfrequency converges already at 3 mesh elements but to be on the safe side 5 mesh elements are chosen. Figure 3.16 b) shows the normalized eigenfrequency as a function of the number of mesh elements along the radius of the plate. The eigenfrequency has converged at 10 mesh elements but this number is increased since the pull-in voltage converges at a higher values as will be shown next. Figure 3.17 a) shows the normalized pull-in voltage as a function of the number of mesh elements along the thickness of the vacuum gap. At 10 mesh elements the line converges and this is chosen as the standard value. Finally, Figure 3.17 b) shows the normalized pull-in voltage as a function of the number of mesh elements along the radius of the plate. This is seen to converge quickly for about 10 mesh elements but then it increases very slightly again. The standard value is chosen as 100 mesh elements. Common for these convergence plots is that even for a very sparse mesh with only a couple of mesh elements along the side lengths, the output varies $< 10\%$ points.

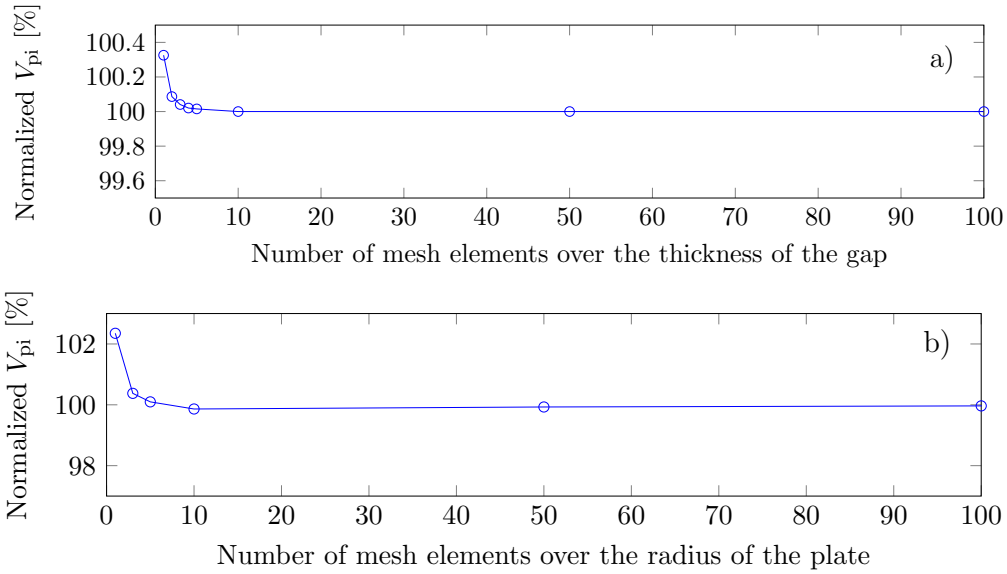


Figure 3.17: Simulated normalized pull-in voltage as a function of a) the number of mesh elements over the thickness of the gap and b) the number of the mesh elements over the radius of the plate.

3.3.2 Model validation

In this section the FEM model results are compared with the analytical expressions given in Chapter 2, in order to validate the FEM model. More specifically, it is checked whether the output parameters such as the resonance frequency and pull-in voltage scale as they should according to the analytical theory.

Figure 3.18 shows a plot of the eigenfrequency, found using the FEM model, as a function of the inverse radius squared, which from Equation 2.38 is expected to be a straight line. This is seen to be the case for all three plate thicknesses. Furthermore, the slopes are expected to be an evenly spaced multiple of each other since the thicknesses are evenly spaced. This is almost seen to be the case: $s_1 = 2.0s_2 = 2.9s_3$. Interestingly, the plate with the thickest plate has a slightly lower slope than expected. This is due to the fact that the assumption behind the analytical expression is broken, namely that: $a/t \gtrsim 40$, which was required so that shear deformation can be ignored. The worst case aspect ratios (smallest radius, $a = 10\,\mu$m) for the three plate thicknesses are: $a/t = 100$, 50, and 33.3 for $t = 100\,$nm, 200 nm, 300 nm, respectively. It is clear that for the plate with a thickness of 300 nm the assumption is violated and consequently the FEM model output deviates from the analytical prediction. As well as validate the FEM model this also

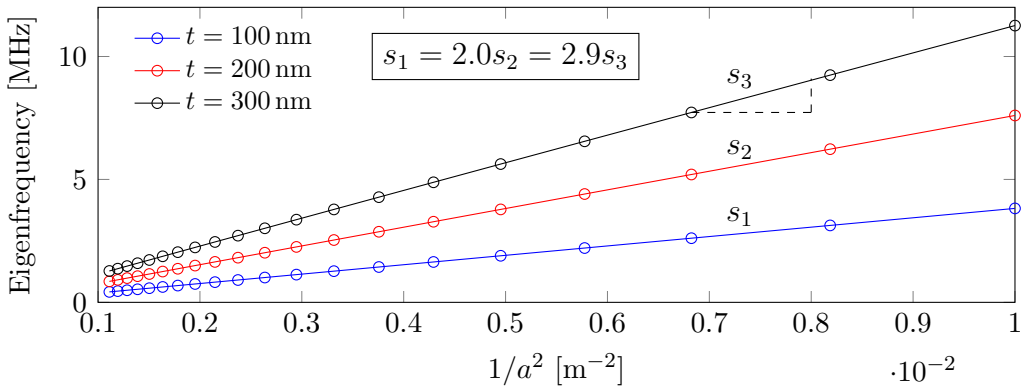


Figure 3.18: Simulated eigenfrequencies as a function of $1/a^2$ for three plate thicknesses, demonstrating that the FEM is capable of deviating from the simple analytical expressions when the underlying assumptions are broken.

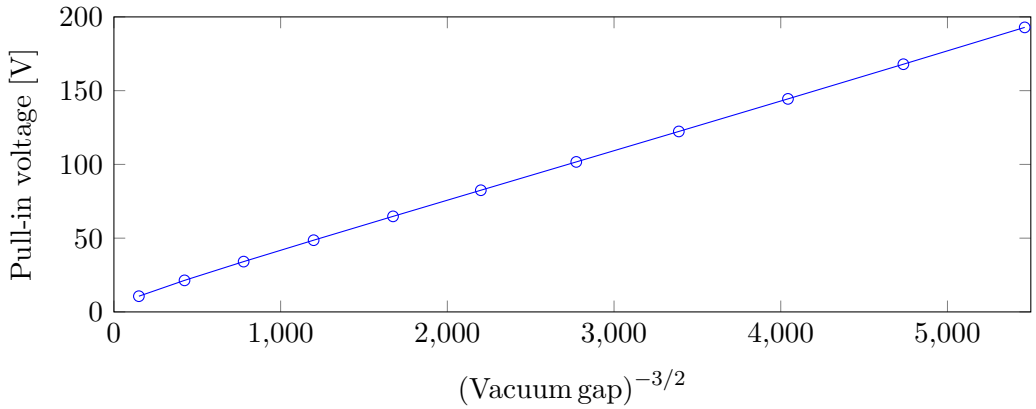


Figure 3.19: Simulated pull-in voltage as a function of $g^{-3/2}$ giving a straight line.

demonstrates the power of using numerical models that includes effects such as this which in the end, hopefully, results in a smaller deviation between the model output and the experimental findings.

Figure 3.19 shows the other main output of the FEM model, the pull-in voltage, as a function of the vacuum gap which has been scaled so that a graph, according to Table 2.4, should give a straight line. Indeed, this is seen to be the case. Since the two main outputs of the FEM model: the eigenfrequency and the pull-in voltage behaves as predicted by the theory the FEM model is validated and can be used to design the CMUT cells.

Figure 3.20 demonstrates the spring softening effect where the eigenfrequency decreases for increasing V_{DC}. The same effect was shown in Figure 2.10 where the effective spring constant decreased for increasing DC volt-

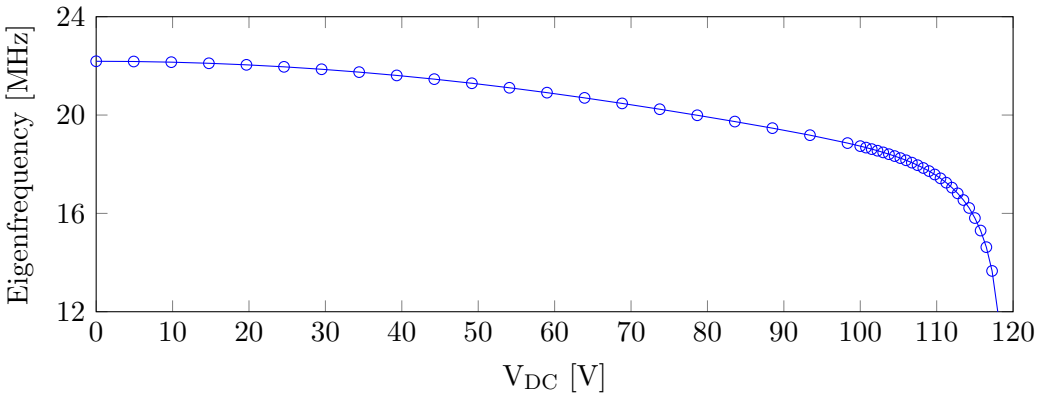


Figure 3.20: Simulated eigenfrequency as a function of bias voltage demonstrating the spring softening effect where the eigenfrequency decreases due to a decrease of the effective spring constant when V_{DC} is increased.

age. When the DC voltages reaches the pull-in voltage at about 118 V in Figure 3.20 the eigenfrequency goes toward zero because the effective spring constant goes toward zero.

Finally, Figure 3.21 shows a plot of the pull-in voltage as a function of the relative size of the top electrode. In this case the top electrode is an Al layer (as in Figure 3.15) and the pull-in voltage is seen to have a minimum for both plate thicknesses. The pull-in voltage is difficult to calculate analytically for a geometry such as this, since the Al layer is not covering the entire cell plate area. For small values of a_{elec}/a the rigidity of the plate is not increased much by the Al layer but at the same time the electrical field is confined to a small central area in the middle of the cell and the resulting V_{PI} is high. For increasing electrode coverage, V_{PI} decreases but at $a_{\text{elec}}/a > 0.9$ the rigidity from the Al layer begins to dominate over the larger spatial spread of the electrical field and the pull-in voltage increases again. This is because the electrostatic force between the electrodes near the edge of the cavity is smaller due to the larger separation of the two electrodes here caused by the smaller plate defection.

In conclusion, a FEM model was presented along with a mesh convergence study that determined the optimal mesh settings. Next, the model was validated by comparing the output with the analytical expressions of the CMUT. The FEM model is used for designing the CMUTs by finding the resonance frequency and pull-in voltage for a given geometry.

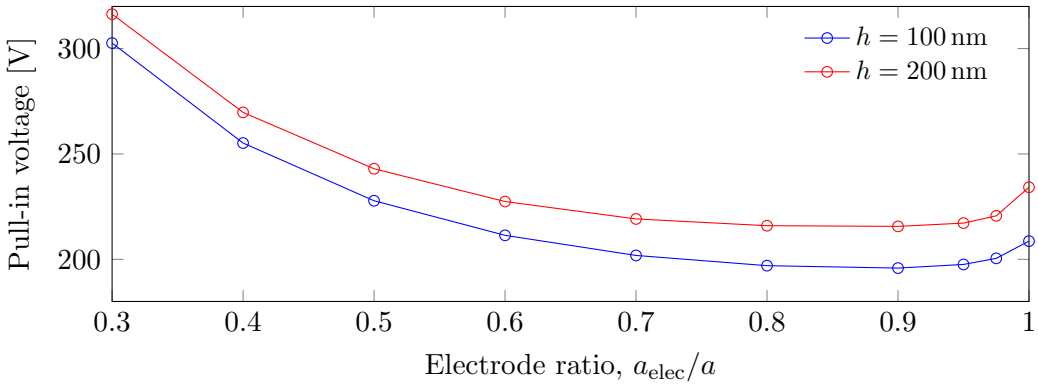


Figure 3.21: Pull-in voltage as a function of relative top electrode/Al layer radius for two plate thicknesses. The non-monotone behavior is due to the interplay between the flexural rigidity of the plate and Al electrode and the electrostatic force.

3.4 Cell designs

In this section the design of each CMUT generation is presented. Table 3.2 lists the dimensions of the four generations. The general trends and differences are discussed here. The first device G1, has a Si plate because our research group had previous experience with the LOCOS process using SOI wafers to define the plate. After some process development in the cleanroom, Si_3N_4 was successfully bonded to both SiO_2 and Si_3N_4. This enabled the use of Si_3N_4 plates giving the benefits listed in Section 3.1.3.

The plate thickness and cell radius has been decreased through the generations, resulting in decreased oscillating masses. This increases the sensitivity and could, according to the empirical Equation 2.111, decrease the limit of detection. The thickness of the top electrode is for G2-4 the same as the plate thickness. A thinner electrode decreases the mass of the plate but increases the electrical resistance of the top electrode as will be discussed later. Furthermore, a minimum metal layer thickness is needed for the wirebonding process. The vastly different post oxide heights are a consequence of the different cavity types: single LOCOS, double LOCOS, and RIE. The influence on the parasitic capacitance is shown in Section 3.4.1. The thickness of the insulation pad oxide is decreased for the later generations as it was shown in the Section 3.2 that the slope lengths are shorter for thinner oxide pad layers. The cell to cell distance or inter-cell distance is decreased in order to decrease the parasitic capacitance. The shortest length is found for G4 with $a_{\text{bond}} = 1\,\mu\text{m}$ which is feasible since this is a RIE device. For the LOCOS devices (both single and double) a_{bond} is limited by the bird's beak which

Table 3.2: Designed parameters of the four CMUT generations.

Geometrical parameter	Unit	Generation			
		G1	G2	G3	G4
Plate radius a	[μm]	7.5	5.4	3.75, 4.5, 5.4	5
Plate thickness t	[nm]	~ 500	100	100	50
Plate material	-	Si	Si_3N_4	Si_3N_4	Si_3N_4
Vacuum gap height t_{vacuum}	[nm]	100	255	50, 100	200
Electrode ratio $a_{\text{electrode}}/a$	-	0, 1	0.5, 1	0.5	1
Electrode thickness $t_{\text{electrode}}$	[nm]	400	100	100	50
Plate mass m_{plate}	[pg]	206, 397	35	17, 25, 35	23
Post oxide thickness t_{post}	[nm]	814	555	546, 635	200
Si bump radius ratio a_{bump}/a	[μm]	0.5	-	0.5	-
Si bump height t_{bump}	[nm]	250	-	200	-
Insulation pad oxide thickness t_{pad}	[nm]	100	10	10	0
Insulation nitride thickness t_{nitride}	[nm]	50	50	50	20
Cell to cell distance a_{bond}	[μm]	6	6	5	1, 3
Wirebond pad area A_{pad}	[μm^2]	6.25e4	3.75e4	3.75e4	1e4

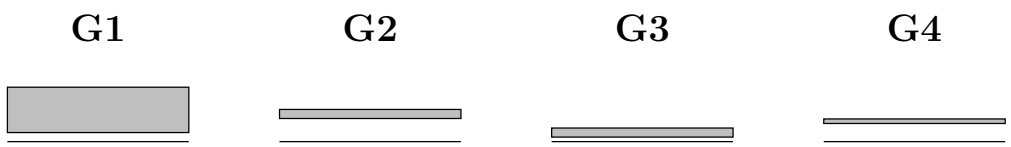


Figure 3.22: Vertical comparison of the plate thicknesses and vacuum gap heights between the four generations. The grey rectangle represents the plate and the white gap represents the vacuum gap. The vertical dimensions are to scale relative to each other. ($t_{\text{vacuum}} = 50\,\text{nm}$ for G3).

is discussed more in Section 3.4.1. Finally, the top electrode wirebond pad area is decreased since this only adds extra parasitic capacitance. However, the area must be big enough for an electrical probe to make contact and big enough that wire bonding can be performed.

Figure 3.22 shows a schematic view of the vertical scales of the four generations. The grey part is the plate while the white gap represents the vacuum gap. The vertical lengths are to scale relative to each other demonstrating the difference in plate thickness and vacuum gap height between the designs.

The dimensions in Table 3.2 are used to calculate the resonance frequency, pull-in voltage, and sensitivity of the devices. These are listed in Table 3.3. The FEM model presented previously has been used to calculate the resonance frequency and pull-in voltage, while the sensitivity is calculated using the simple expression given in Equation 2.92. The theoretical sensitivity is seen to be highest for G3 due the small radius ($S \propto a^{-4}$) and the theoret-

Table 3.3: Theoretical values for the four generations, calculated using the FEM model presented earlier. The configuration for each generation with the highest sensitivity is presented here. The sensitivity and normalized sensitivity are calculated using Equation 2.92 and 2.94, respectively.

Parameter	**Unit**	**Generation**			
		G1	G2	G3	G4
Plate radius a	[μm]	7.5	5.4	3.75	5
Electrode ratio $a_{\text{electrode}}/a$	-	0, 1	0.5	0.5	1
Resonance frequency f_r	[MHz]	31.3, 38.5	42.4	64.7	41.2
Pull-in voltage V_{PI}	[V]	95, 163	167	34	68
Cell area A	[μm^2]	177	92	44	79
Plate mass m_{plate}	[pg]	206, 397	35	17	23
Sensitivity S	[Hz/ag]	0.08, 0.05	0.60	1.90	0.89
Sensitivity normalized S_{norm}	[cm^2/g]	0.43e4, 0.22e4	1.30e4	1.30e4	1.70e4

ical sensitivity is predicted to be second highest for G4 even with a plate thickness half of G3, due to a lower resonance frequency and slightly higher mass. The normalized sensitivity is expected to be highest for G4 since the plate and top electrode is thinnest ($S_{\text{norm}} \propto t^{-1}$). Later it is shown that the calculated sensitivities in the table differ compared with the measured sensitivity values due to differences in geometry after fabrication and mass loading of the CMUT, which decreases the sensitivity.

3.4.1 Capacitance

In this section the theoretical capacitance of the designed CMUT elements is given for non-deflected plates, that is non-biased cells. The design choices affect the total capacitance but also the amount of parasitic capacitance which in turn will decrease the SNR. The parasitic capacitance is sought minimized by increasing the post oxide height to effective gap ratio and by structuring the top electrode. Figure 3.23 shows two elements with $4 \times 4 = 16$ cells with a) a top electrode covering all cells and the area in between and b) a structured top electrode only covering part of the cells and as small an area in between the cells as possible.

Number of cells

The effect on the capacitance of increasing the number of cells in an element is investigated. The designed dimensions and layout of a G2 chip with a non-structured top electrode (Figure 3.23 a)) is used as an example. Since the area

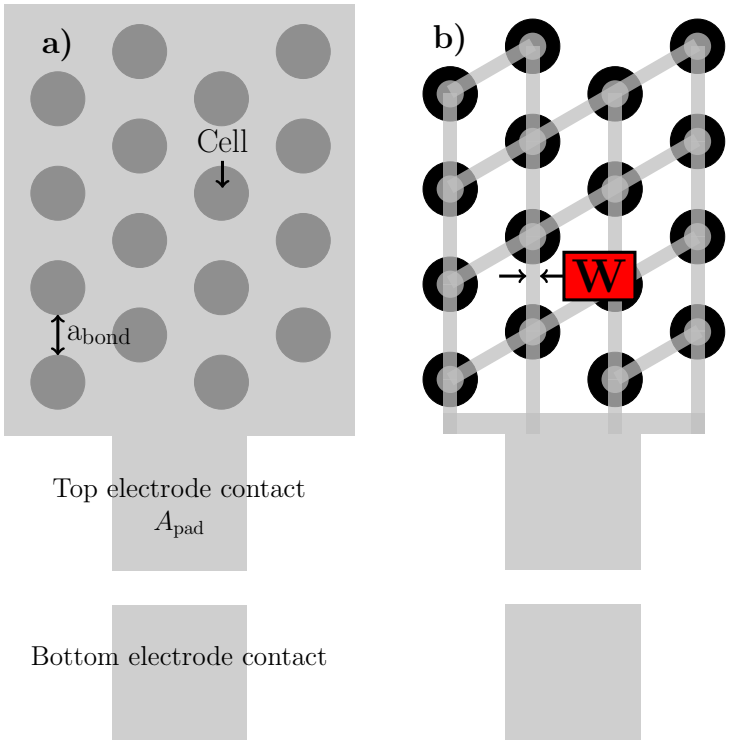


Figure 3.23: Top view of sketches of two CMUT elements where a) the top electrode covers the entire element and is thus *non-structured* and b) where the top electrode is *structured* and only placed over the cells and as thin connectors between the cells.

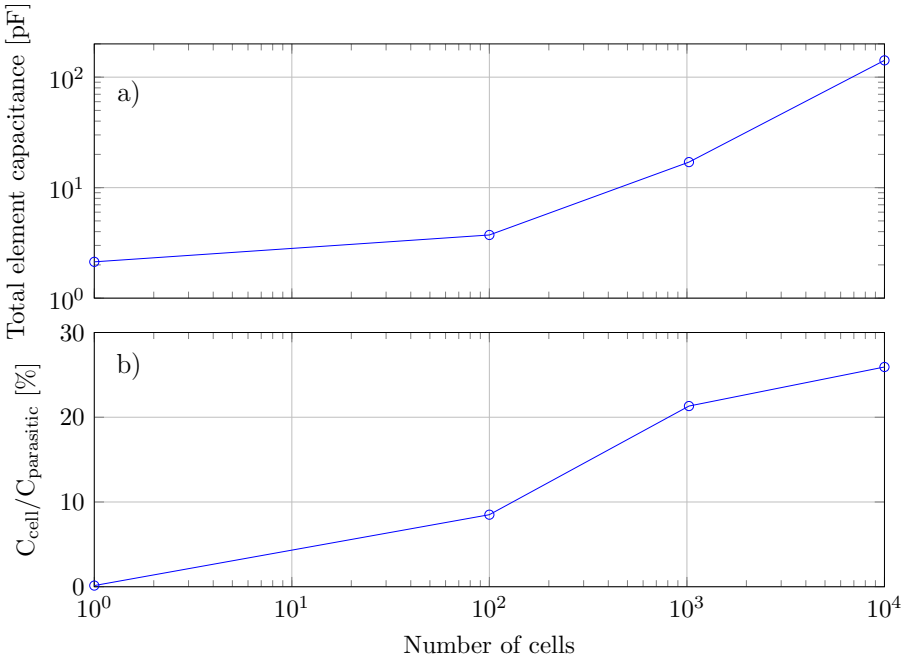


Figure 3.24: a) Total element capacitance (sum of the parasitic and cell capacitance) as a function of the number of cells in the elements for the layout of the G2 elements. b) Cell capacitance relative to the parasitic capacitance as a function of the number of cells in an element for the layout of the G2 elements.

of the top electrode contact pad is fixed for a given design, the total CMUT cell capacitance (C_{cell}) will be increased relative to the parasitic capacitance ($C_{\text{parasitic}}$) when the number of cells is increased. Figure 3.24 shows a plot of the total element capacitance for an increasing number of CMUT cells. The plot is based on the calculated capacitance of the designed values for a G2 design with a full top metal electrode with zero biasing. As the number of cells is increased the total element capacitance increases as well as the ratio $C_{\text{cell}}/C_{\text{parasitic}}$ because the ratio between the cell area and top electrode covered area in between the cells increases. Thus, the SNR is expected to be best for the elements with the most cells. However, increasing the number of cells worsens the effects from non-uniformities of the cells. Therefore, the number of cells should not be increased indefinitely.

Top electrode structuring

For designs with dielectric non-conducting plates (such as Si_3N_4) the structuring of the top electrode results in a decrease of the parasitic capacitance. Figure 3.23 showed a design with a non-structured and a structured top

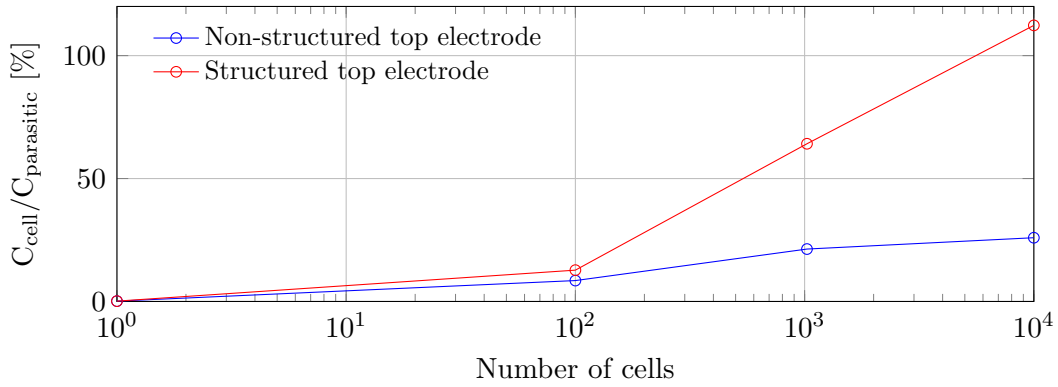


Figure 3.25: Cell capacitance relative to the parasitic capacitance as a function of the number of cells in an element of G2 design. The element with a structured top electrode has less parasitic capacitance than the element with a non-structured top electrode.

electrode. Also shown in the figure is the close packing of the circular cells giving the highest fill factor and hereby the lowest $C_{\text{parasitic}}$ value. Figure 3.25 shows a comparison between the ratio $C_{\text{cell}}/C_{\text{parasitic}}$ for two elements where the only difference is the structuring of the top electrode. That is, a comparison between the situation in Figure 3.23 a) and b). The cell areas are fully covered by the top electrode in both cases. As the number of cells increases the ratio $C_{\text{cell}}/C_{\text{parasitic}}$ increases faster for the element with the structured top electrode as the parasitic capacitance in between the cells becomes important compared with the initially large parasitic contribution from the top electrode pad contact. This is also the reason why $C_{\text{cell}}/C_{\text{parasitic}}$ is almost the same for elements with few cells; the parasitic capacitance from the top electrode contact pad dominates when the elements only have a few cells but when the number of cells increase the contribution to the parasitic capacitance from the area in between the cells becomes comparable and larger than the contribution from the contact electrode.

The structuring of the top electrode is, as already mentioned, only possible for designs with non-conducting plates. Thus, designs with highly doped Si plates lack this design freedom and the parasitic capacitance will be higher.

Cavity design

The effect of the cavity design (RIE, single or double LOCOS) on the capacitance is investigated. The real designs with different cavity types cannot directly be compared as the vacuum gaps, radii, and or dielectric layer thickness in the cavity are of different values. In order to compare the three cavity types three fictional devices are considered with the same vacuum gap

of $t_{\text{vacuum}} = 100\,\text{nm}$, no dielectric layers in the cavity, and a Si bump height of $t_{\text{bump}} = 200\,\text{nm}$ for the double LOCOS cavity. The post oxide heights are in this case given by:

$$t_{\text{post, RIE}} = t_{\text{vacuum}} = 100\,\text{nm} \tag{3.4}$$
$$t_{\text{post, single LOCOS}} = t_{\text{vacuum}}/0.56 = 179\,\text{nm} \tag{3.5}$$
$$t_{\text{post, double LOCOS}} = (t_{\text{vacuum}} + t_{\text{bump}})/0.56 = 536\,\text{nm}. \tag{3.6}$$
$$\tag{3.7}$$

The corresponding $t_{\text{post}}/t_{\text{vacuum}}$ ratios are:

$$t_{\text{post, RIE}}/t_{\text{vacuum}} = 100\,\text{nm}/100\,\text{nm} = 1 \tag{3.8}$$
$$t_{\text{post, single LOCOS}}/t_{\text{vacuum}} = 179\,\text{nm}/100\,\text{nm} = 1.79 \tag{3.9}$$
$$t_{\text{post, double LOCOS}}/t_{\text{vacuum}} = 536\,\text{nm}/100\,\text{nm} = 5.36, \tag{3.10}$$

which demonstrates one of the advantages of the double LOCOS cavities. The cell and parasitic capacitances are calculated and the ratio for the three devices is shown in Figure 3.26. All variables other than t_{post} are held constant, including the top electrode contact area etc. Furthermore, it is assumed that $a_{\text{bump}} = a$ for the double LOCOS device, in order to simplify the comparison. As expected, the lowest $C_{\text{cell}}/C_{\text{parasitic}}$ ratios are found for the RIE device as the $t_{\text{post, RIE}}/t_{\text{vacuum}}$ values are lowest. Because the vacuum gap height is decoupled from the post oxide height through the Si bump the highest $C_{\text{cell}}/C_{\text{parasitic}}$ ratios are found for the double LOCOS CMUT. Finally, the single LOCOS device is situated in between these two extremes.

Inter cell distance - $\mathbf{a}_{\text{bond}}$

The cell to cell distance or inter cell distance was in Figure 3.23 denoted a_{bond}. Decreasing a_{bond} increases the fill factor in an element which is the area of the cells relative to the total area of the element. The lower limit of a_{bond} is set by the requirement that the plate should be properly bonded to the substrate wafer. A too small bonding area will result in a plate that ultimately is not attached to the substrate wafer. As shown in Table 3.2, a_{bond} has been decreased going through the generations from G1: $a_{\text{bond}} = 6\,\mu\text{m}$ to G4: $a_{\text{bond}} = 1\,\mu\text{m}$. As the radius of the cavity decreases, the inter cell distance becomes increasingly important in order to uphold a high fill factor and hereby a low parasitic capacitance. It should be noted that when designing LOCOS cavities, extra care must be taken, since a_{bond} can become comparable in length to the slope length of the bird's peak structure. Figure 3.27 shows a plot of a) the flat bondable inter cell length as a function of the

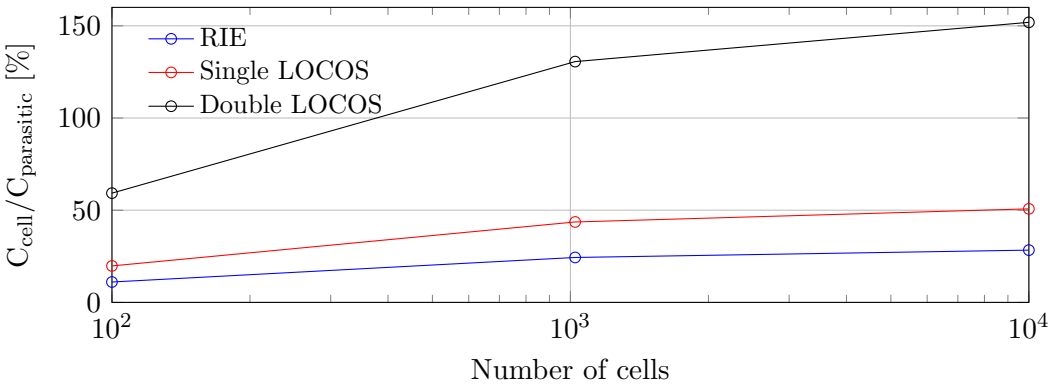


Figure 3.26: Capacitance of the cells in an element relative to the parasitic capacitance in that element as a function of the number of cells. All dimensions of the three cells and elements are the same, except the post oxide height t_{post}, which is varied to keep a vacuum gap height of $t_{vacuum} = 100\,nm$. It is assumed that no dielectric layers are in the cavities and that $a_{bump} = a$ for the double LOCOS cavity.

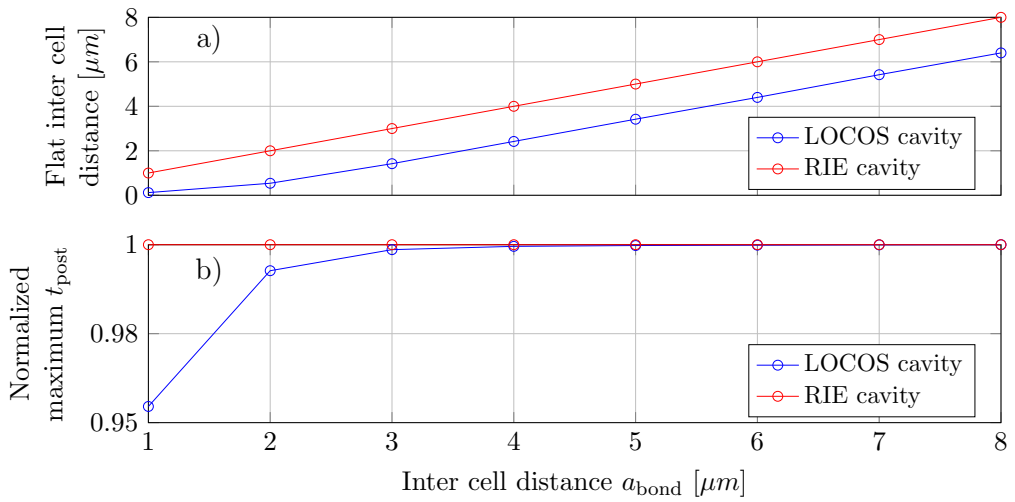


Figure 3.27: a) Distance between two cells which is flat, that is has a slope less than 1% as a function of the designed distance between the cells on the photolithographic masks a_{bond}. Note how the bonding area is consistently larger for the RIE cavities. b) Maximum post SiO_2 height t_{post} normalized to t_{post} for $a_{bond} = \infty$ as a function of a_{bond}.

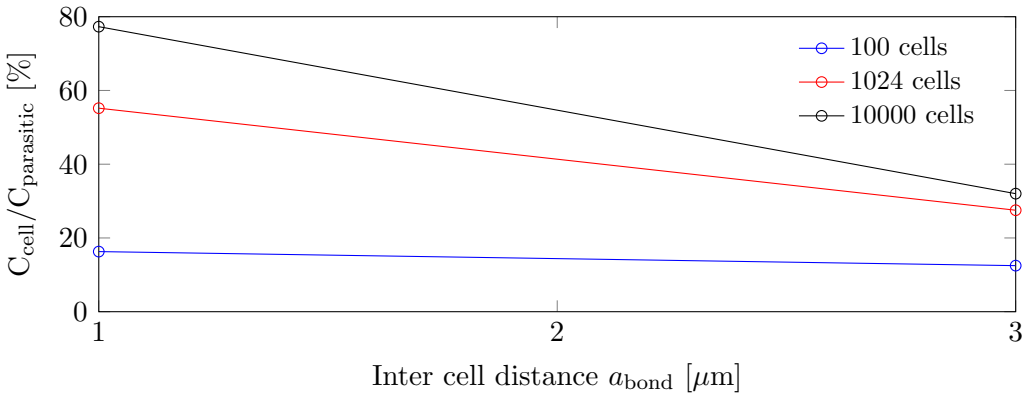


Figure 3.28: Cell capacitance relative to the parasitic capacitance as a function of the inter cell distance for three number of cells.

designed a_{bond} length for a LOCOS and RIE cavity. A linear relationship is seen for the RIE cavity, where the designed a_{bond} length is the same as the flat a_{bond} length. Further, in b) t_{post} is constant as a function of a_{bond} for the RIE device. For the LOCOS cavity the linear behavior in a) is also seen for $a_{\text{bond}} > 3\,\mu$m. Under this threshold value the flat inter cell distance becomes smaller than expected and deviates from the linear tendency and t_{post} does not reach its nominal height. Indeed, this demonstrates that placing LOCOS cavities closer to each other than $a_{\text{bond}} \approx 3\,\mu$m, results in a smaller bondable area than expected. In addition, the two lines in a) are seen to be parallel shifted relative to each other, showing a larger bondable area for the RIE cavities for the same a_{bond} length. In conclusion, LOCOS cavities can not be placed as close to each other as RIE cavities.

The effect of decreasing a_{bond} on the capacitance is shown in Figure 3.28, which is calculated for a G4 element using the designed values from Table 3.2. The figure shows that the parasitic capacitance is increased for increasing values of a_{bond}. Furthermore, this effect is larger for elements with more cells, as the constant parasitic capacitance from the top electrode contact is the same for all the elements. Thus, if the fabrication process allows for it, a_{bond} should be made as small as possible.

In conclusion, the effect of various design parameters on the capacitance has been studied. Increasing the number of cells in an element increased the $C_{\text{cell}}/C_{\text{parasitic}}$ ratio due to the fixed contribution to the parasitic capacitance of the top electrode contact pad. Structuring the top electrode was also seen to decrease the parasitic capacitance along with increasing the post oxide to vacuum gap ratio. Finally, if the fabrication process allows for it, the inter cell distance should be decreased in order to decrease the parasitic capacitance.

3.4.2 Top electrode design

When designing the top electrode the electrical resistance should not be too high compared with the impedance of the cell capacitance. For a CMUT element with a single cell the calculation of the electrical resistance is straightforward but for elements with multiple cells the resistance network becomes distributed. The most extreme case is found when the top electrode is structured as shown in Figure 3.23 b). Here the width w, height $t_{\text{electrode}}$, length l, and material choice of the metal bars connecting the cells, determine the potential drop over the element. The highest potential drop is found for the cells located furthest away from the top electrode contact, thus elements with more cells experience the largest potential drop. As an example the resistance of the top electrode is calculated for G2 for the element with most cells, namely 100x100 = 1e4, as these are geometrically the largest. The resistance of each metal connector is:

$$R_{\text{connector}} = \rho_{\text{Al}}(l/(wt_{\text{electrode}})) = 1.8\,\Omega. \tag{3.11}$$

The resistance can now be calculated between the top electrode contact and the cell furthest away. This calculation is carried out by a script developed by PhD. student Andreas Havreland and results in:

$$R_{\text{max}} = 10.7\,\Omega, \tag{3.12}$$

which, as expected, is lower than the resistance of 100 metal connectors connected in series ($100R_{\text{connector}}$). The electrical resistance found in Equation 3.12 can now be compared with the absolute impedance of the single cell at the operating frequency of 40 MHz:

$$|Z_{\text{cell}}| = |\frac{1}{j\omega C_{\text{cell}}}| = 5.44\,\text{M}\Omega. \tag{3.13}$$

As $R_{\text{max}} \ll |Z_{\text{cell}}|$ these parameters for the top electrode do not give rise to a too high electrical resistance.

3.4.3 Dielectric breakdown

It is important to know at which voltage dielectric breakdown is expected to occur when in pull-in. Since only SiO_2 and Si_3N_4 are in between the top and bottom electrode for G1-4, only the dielectric strengths of these two materials are needed. The dielectric strength of SiO_2 has, in our cleanroom, been measured to be approximately 0.7 V/nm (the theoretical value is typically stated as 1 V/nm). Taking a conservative approach and using the same value

Table 3.4: Breakdown voltage estimates of the four designs.

Parameter	Unit	Generation			
		G1	G2	G3	G4
Breakdown voltage $V_{\text{breakdown}}$	[V]	105	112	112	49

for the Si_3N_4 the dielectric breakdown voltage can be calculated for the four generations, remembering that the plate itself is a dielectric material for G2-4. The breakdown voltage is estimated using the simple relation:

$$V_{\text{breakdown}} = t_{\text{dielectric}} K, \tag{3.14}$$

where $t_{\text{dielectric}}$ is the total thickness of the dielectric layers and K is the dielectric strength constant of proportionality. Table 3.4 lists the calculated breakdown voltages. Indeed, some the breakdown voltages for some of the designs are lower than the pull-in voltages and breakdown is expected if pull-in occurs. The devices are not to be operated in pull-in nor with a bias voltage close to pull-in, thus the designs are acceptable for normal operation.

3.5 Chapter conclusion

In this chapter the designs of the four generations of CMUTs have been presented. Further, is was shown how the design space of LOCOS CMUTs is limited when the cell radius becomes small ($a < 10\,\mu$m). Adhering to the rules set up in the chapter, an example of a minimum radius of a double LOCOS cell is $a_{\text{min}} \approx 2.1\,\mu$m for a design with $t_{\text{bump}} = 200\,$nm, $t_{\text{vacuum}} = 200\,$nm, and a Si_3N_4 mask. A FEM model was presented, validated and used to calculate the resonance frequency and pull-in voltage of the designs. The designs of the four generations of CMUTs were presented along with their theoretical resonance frequencies, pull-in voltages, capacitances, and sensitivities. Table 3.2 showed the general trend of decreasing plate masses, partly due to decreasing plate thicknesses, for the later CMUT generations. This resulted in high theoretical mass sensitivities and normalized mass sensitivities (Table 3.3). Moreover, the parasitic capacitance was sought decreased by structuring the top electrode, decreasing the inter cell distance a_{bond}, and increasing the post SiO_2 height to vacuum gap ratio. This was done in an effort to increase the electromechanical coupling coefficient.

CHAPTER 4

Device fabrication

In this chapter the process flows and fabrication details of the CMUT devices are presented. Specifically, the process flow for each generation is reviewed. During the process flow review and in the end of the chapter, process optimization of selected process steps is presented. In particular, the wafer bonding step of a Si_3N_4 surface to a SiO_2 surface or Si_3N_4 surface is considered along with an investigation of the surface roughness of the wafers after being exposed to different processes. Finally, the finished chips are wire bonded to a chip carrier.

All fabrication has been carried out in DTU Danchip's 1350 m^2 class 10-100 cleanroom located at the campus of DTU. Detailed process flows for the four generations can be found in Appendix D, E, F, and G.

4.1 General fabrication aspects

In this section general information about the fabrication, shared among the four generations, is presented. Specifically, the substrate wafer type, photolithography process, and measurement of thin film thicknesses are discussed.

4.1.1 Wafers

Substrate wafer

For all the designs the substrate wafer also acts as the bottom electrode. This means that the electrical resistivity should be as low as possible, which is achieved by using highly doped Si wafers. All substrate wafers used in this project are single side polished 4″ Si with an electrical resistivity $< 0.025\,\Omega$cm highly doped with either boron or antimony making them either p or n doped. The Si crystal orientation is (100) with a flat in the <110> direction.

SOI wafer

A 4″ SOI wafer is used in the fabrication of G1. The device layer has a specified thickness of $2.00\,\mu\text{m} \pm 0.30\,\mu\text{m}$, Si crystal orientation (100), flat direction <110>, and an electrical resistivity of $0.001\,\Omega\text{cm} - 0.01\,\Omega\text{cm}$.

4.1.2 Lithography

The cleanroom at DTU Danchip offers a variety of lithography equipment: mask-less UV aligner, standard UV aligner, DUV stepper, nanoimprint lithography machine, and an e-beam lithography machine. In this project UV lithography was used as the higher resolution of the other techniques are not required. This reduces the cost of the photo masks as compared with e.g. DUV masks and is considerably faster than e.g. e-beam lithography. More specifically, the machine used was a Süss MicroTek Mask Aligner MA6 with a minimum resolution of $1.25\,\mu$m. The mask to wafer alignment can be done with an approximate precision of $1\,\mu$m. The masks were all $5'' \times 5''$ soda-lime glass with a Cr pattern. The photoresists used were the negative AZ nLOF 2020 and the positive AZ MiR 701 both at a thickness of $1.5\,\mu$m. Prior to applying the photoresist the wafers were typically coated with hexamethyldisilazane (HMDS) to increase the adhesion between the wafer and resist. After exposure the wafers and resist are subjected to a post exposure bake (PEB) and a puddle development using AZ 726 MIF which is a tetramethylammonium hydroxide (TMAH) based developer. Both the process of applying the photoresist and developing the exposed substrates are performed by fully automated machines, ensuring stable and consistent results. After each lithography process the resist pattern is inspected by optical microscopy.

The lithography masks are aligned to the patterns on the wafers using standard alignment marks. After the wafer bonding, the alignment marks on the substrate wafers are covered by the bonded top wafer. When the top

wafer has been thinned down, leaving only the plate, the alignment marks can be seen through the plate layer if this is sufficiently thin (e.g. 100 nm Si_3N_4). However, before the metallization step, openings through the plate over the alignment marks must be made, as the metal layer is not transparent to visible light. This is needed for the alignment of the top electrode mask. These openings to the alignment marks have been achieved in three ways: 1) For the double LOCOS process the Si bumps can be used as alignment marks and are placed outside of a cavity. This creates local voids at the location of the alignment marks, thus effectively removing the plate over the alignment marks. 2) After wafer bonding, the next process step is opening up to the bottom electrodes with an etch. If the alignment marks can be seen through the plate, the bottom electrode etch area can be included to cover the area over the alignment marks, hereby removing the plate. 3) Since the position on the wafer of the alignment marks is the same for all generations, a special mask has been used with two rather large (1 cm × 1 cm) square openings over the location of the alignment mark cluster. This special extra mask can be used to make a resist etching mask for the etching of the plate over the alignment marks, without the need of alignment.

4.1.3 Film thickness measurement

The film thickness of deposited dielectric films are measured after each deposition in order to check whether the desired film thickness has been achieved. When growing a thermal SiO_2 or depositing Si_3N_4 or polysilicon in a LPCVD furnace, a test wafer is included. The film thicknesses are measured for these test wafers so that the process wafers are exposed as little as possible to the environment. The film measurements are performed using an ellipsometer.

4.2 Process flow - G1

In this section the process flow of G1 is presented. A more detailed process flow can be found in Appendix D. This fabrication process is based on the one originally presented in [42] with a few minor differences.

Figure 4.1 shows the process flow for G1. The process begins in step **(1)** with a substrate wafer of the type described in Section 4.1.1. Subsequently, the wafer is RCA cleaned and a pad SiO_2 is grown by dry oxidation at 900 °C. The reason for choosing such a relatively low oxidation temperature is in order to have a higher degree of control of the process by increasing the processing time. This is done since the SiO_2 target thickness is only 10 nm. Next, a Si_3N_4 layer is deposited in a LPCVD furnace and then a

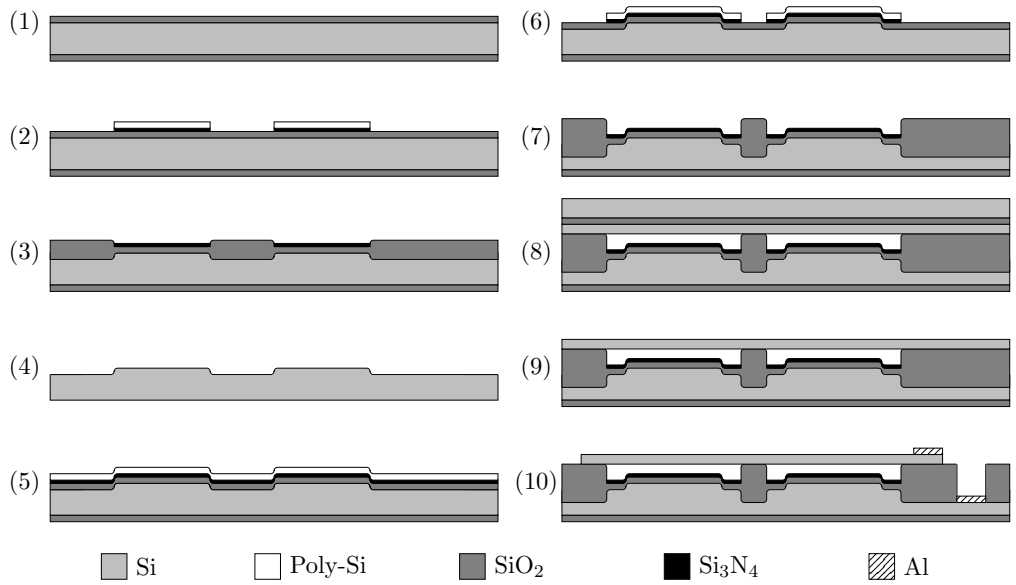


Figure 4.1: Process flow for the G1 CMUT. In step (1)-(4) the Si bumps are formed. Next, the cavities are made in step (5)-(7), and in step (8) the SOI wafer is wafer bonded to the cavity wafer. Then the handle and buried oxide layer of the SOI wafer are removed in step (9). Finally, openings to the bottom electrode are etched and metal is deposited on the wafer. Subsequently, the metal is etched and lastly the device layer/plate is etched to isolate the elements.

polysilicon layer is deposited in a different LPCVD furnace. This polysilicon layer will function as a mask for the wet etching of the underlying Si_3N_4 layer. Alternatives to this step are discussed in Section 4.6.2. A lithography step (described in Section 4.1.2) is performed on the polysilicon layer with the photo mask defining the Si bumps. After development of the resist, the polysilicon layer is etched in a DRIE machine. The overetch that will occur into the Si_3N_4 layer is not critical since this layer is to be removed at these positions in the next process step. The resist mask is stripped in an oxygen plasma.

Step **(2)** shows the wafer after the Si_3N_4 layer has been wet etched in 180 °C H_3PO_4 (85 wt%), hereby defining the oxidation masks for the first LOCOS oxidation. The selectivity between Si_3N_4 and SiO_2 is reported to be > 35 [96]. Therefore, only a small amount of SiO_2 is expected to be etched. Next, the wafers are cleaned in a solution of 80 °C H_2SO_4 (98%) and ammonium persulfate, a so-called 7-up clean. This is done to remove potential alkali ions and traces of organic matter from the resist. The remaining polysilicon mask on the now etched Si_3N_4 layer can either be left during the LOCOS oxidation or removed using the same DRIE machine used to pattern the polysilicon layer earlier.

Step **(3)** demonstrates what the wafer looks like after the first wet LOCOS oxidation at 1100 °C. The Si bumps have now been formed. Subsequently, in step **(4)** all dielectric layers are stripped: first the wafer is immersed in a buffered HF (BHF) etch to etch any oxidized Si_3N_4, then the Si_3N_4 is stripped in 180 °C H_3PO_4, followed by a 7-up cleaning step, and finally the remaining SiO_2 is stripped by a BHF etch. This leaves the bare Si surface with the Si bumps.

The processes used in steps **(5)** to **(7)** are identical to steps **(1)** to **(3)** and are therefore only briefly covered. In step **(5)** the wafer is RCA cleaned and oxidized by a wet oxidation process at 1100 °C in order to grow a pad SiO_2. This is followed by LPCVD depositions of Si_3N_4 and polysilicon and a lithography step to define the cavities. In step **(6)** the polysilicon is dry etched and the resist mask is stripped in an oxygen plasma. Next the Si_3N_4 is wet etched using the polysilicon as a mask and the wafer is cleaned in 7-up. The polysilicon must be stripped at this point (unlike at the 1. LOCOS step), since during the LOCOS process the polysilicon will oxidize and be located in what will be the center of the cavity. In step **(7)**, after a RCA clean, the second LOCOS process is performed in a wet atmosphere at 1100 °C, forming the post SiO_2 and hereby the cavities.

In step **(8)** the substrate wafer and SOI wafer are cleaned in a RCA cleaning step. This cleaning step is performed prior to the fusion bonding step. It is critical that both wafers are free from particles in order to ensure

that no voids are formed at the bonding interface. This process step typically limits the overall yield the most. The bonding process is performed under vacuum at a pressure of $5 \cdot 10^{-4}$ mbar. A piston applies a pressure to the wafer pair of 667 mbar for 5 min at a temperature of 50 °C. This is followed by an annealing step in a furnace at 1100 °C. The substrate wafer and SOI wafer are now bonded together.

In step **(9)** the Si handle layer of the SOI wafer is dry RIE etched. The process can be monitored by looking at the wafer during processing. The process is stopped when the handle layer is completely removed, which can be visibly seen by the naked eye. The buried oxide layer (BOX) is used as an etch stop layer and is partially, but not completely, etched by the dry etch. The remaining BOX layer is etched with a BHF etch. The wafer is RCA cleaned. The device layer is thinned down using a dry RIE etch since the available device layer thicknesses from the supplier did not match the design requirement. The device layer could also have been thinned by oxidizing the SOI wafer prior to wafer bonding and then removing the SiO_2 in a BHF etch. The advantage of using the oxidation method for thinning down the device layer is that the thickness of Si removed can be controlled much more precisely and the uniformity over the wafer is higher.

In step **(10)** openings to the bottom electrode/substrate wafer are made, the device is metallized and the elements are isolated. First, a lithography step is performed with the mask defining the bottom electrode openings, followed by a dry etch of the Si plate and a dry etch of the post SiO_2 layer. The resist is stripped in an oxygen plasma. Next, a thin (10 nm) Ti adhesion layer is deposited by e-beam deposition to ensure a good adhesion between the Si and Al layer which subsequently also is deposited by e-beam deposition. These two depositions are performed without exposing the wafer surface to the atmosphere, thus avoiding any oxidation of the metal layers. A lithography step is performed with the mask defining the top Al electrode pattern. The Ti and Al layers are dry etched and afterwards the resist is stripped in an oxygen plasma. A finial lithography step is carried out with the mask defining the CMUT elements. The Si plate is dry etched to electrically isolate the elements and the resist is stripped in an oxygen plasma.

In a final step, which is not shown in Figure 4.1, the wafers are diced into individual chips by a saw wafer dicer. The wafers are coated with a layer of 4.2 μm AZ 5214E resist prior to dicing in order to make sure that no conducting Si pieces (from the dicing process) lands on the wafer surface and cause short circuits.

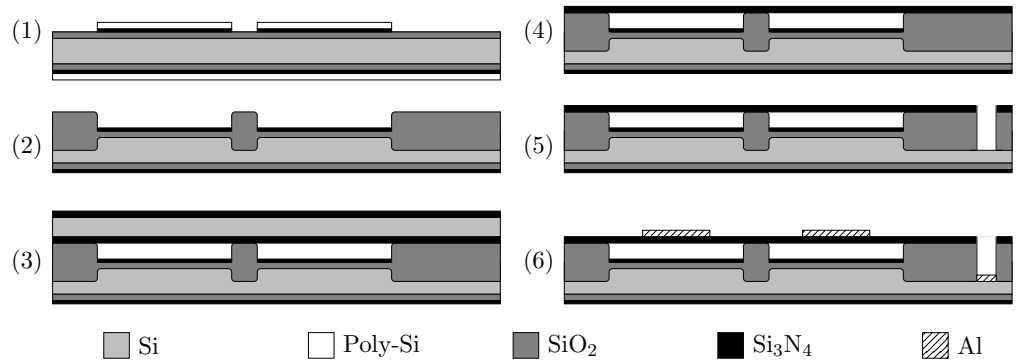


Figure 4.2: Process flow for the G2 CMUT. The cavities are formed in step (1)-(2) with a LOCOS process step. In step (3) a wafer with a Si_3N_4 layer on each side is wafer bonded to the cavity wafer. The top Si_3N_4 layer and Si handle layer are etched in step (4) leaving the Si_3N_4 plate. Openings to the bottom electrode are made in step (5). Finally, metal is deposited and patterned in step (6).

4.3 Process flow - G2

In this section the process flow of G2 is presented. A more detailed process flow can be found in Appendix E. Due to the similarity of the device structure for G1 and G2, many of the fabrication steps are the same and will therefore only be described briefly.

Figure 4.2 shows the process flow for G2. The single LOCOS cells are fabricated using a single LOCOS step and sealed with a Si_3N_4 plate. Step **(1)**-**(2)** are equivalent to what was described in the previous section. In step **(3)** Si_3N_4 has been deposited by LPCVD on a double side polished Si wafer. This wafer is subsequently bonded to the substrate wafer. Section 4.6.1 elaborates more on the bonding of Si_3N_4.

From step **(3)** to **(4)** the top Si_3N_4 layer is dry etched and the Si handle layer is first thinned down to about 1/3 of the original thickness in a dry etch and then completely removed in 80 °C KOH (28 wt%). The KOH etch has a very high selectivity between Si and Si_3N_4 [96] and therefore it effectively stops at the Si_3N_4 plate and on the back of the substrate wafer which has a protective Si_3N_4 layer. The polysilicon on the back of the wafer protected this back side Si_3N_4 layer in step **(1)** during the Si_3N_4 wet etch. The reason for first thinning down the Si handle layer in a dry etch is to reduced the etching time in the KOH etch. For wafers where the Si handle layer is only etched in KOH the edges of the wafer are attacked by the KOH, resulting in very rough edges. This makes further processing more difficult, although not impossible. Figure 4.3 shows a photograph of the edge of a) a wafer only etched in KOH and b) a wafer first dry etched and then KOH etched.

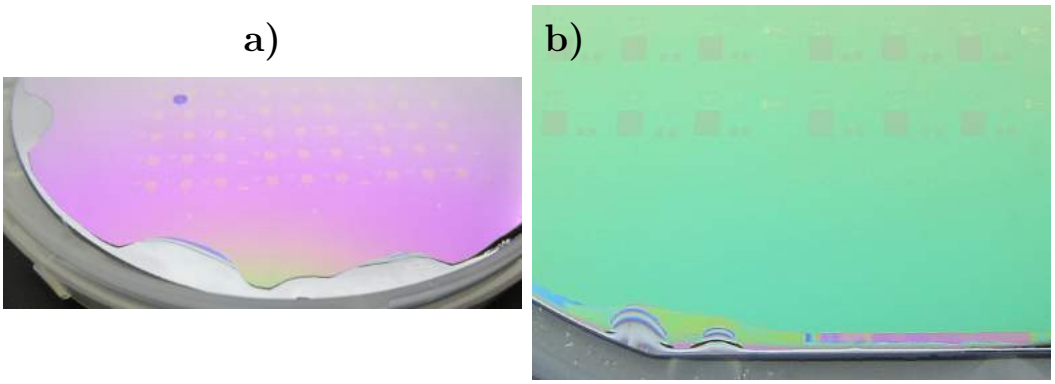


Figure 4.3: Photographs of two wafers after removal of the Si handle layer using a) only KOH and b) a combination of dry etching and KOH. The wafer edge is more damaged in a) than in b).

For wafer a) the KOH has had more time to etch the wafer edge and much less rough/damaged edges are seen for the wafer in b). If a thick oxide, resembling a burried oxide for a SOI wafer, had been grown prior to Si_3N_4 deposition on the plate wafer, a dry etch could have been used to remove the handle Si layer entirely. This process would almost be the same as when using SOI wafers in the process for G1.

In step **(5)** openings to the bottom electrode are made. First, a lithography step is performed with the mask defining the bottom electrode openings, followed by a dry etch of the Si_3N_4 plate and a dry etch of the post SiO_2 stopping on the Si layer. Finally in step **(6)**, the device is metallized with Al and annealed in a furnace at 425 °C to increase the adhesion between the Al and Si_3N_4 plate, as no intermediate adhesion layer is used. This reduces the weight of the top electrode. Previous experiments, performed in the group, have shown that by annealing the Al layer, the adhesion is increased. Next, a final lithography step is performed with the addition of an extra baking step on a hotplate at 120 °C before the wet etch of the Al layer. The extra baking step is included to reduce the risk of delamination of the resist on the Al layer. The Al etch is a solution of $H_2O : H_3PO_4$(1:2). A change is seen when the metal has been etched. During the etch H_2 bobbles are formed at the wafer surface. Some of the bubbles can stick to the surface hereby hindering fresh etching solution in reaching the surface and effectively stopping the etching process locally. Figure 4.4 shows an optical microscope image where two metal areas in the top electrode grid of G2 have not been completely removed. This is most likely due to the formation of bubbles that can get lodged easily in this patterned top electrode structure. To remove these bub-

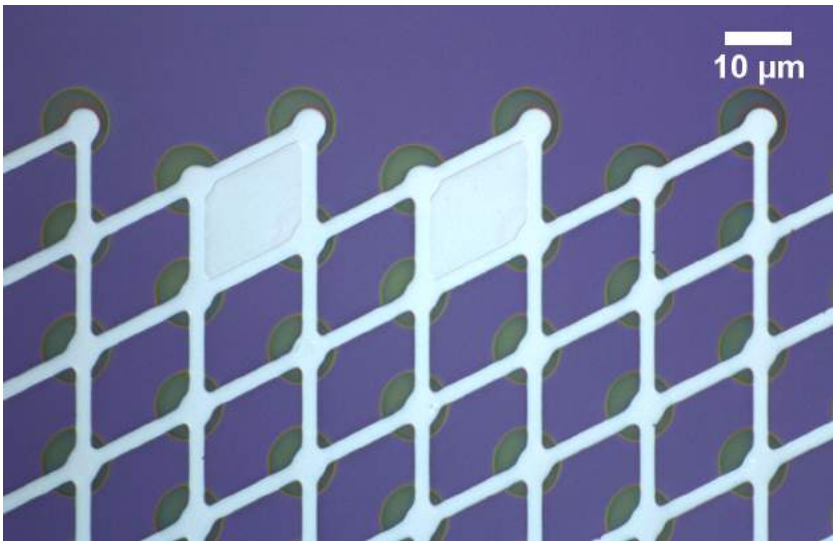


Figure 4.4: Optical microscope image of a finished G2 chip after the Al wet etch. Two areas of Al in between the connectors have not been etched, possibly due to the formation of H_2 bubbles blocking the flow of fresh etch reactants.

bles the wafers are lifted out of the solution up to every 5 s depending on the Al pattern. This eliminated the problem. Increasing the water concentration in the etch would decrease the etch rate and hereby the frequency at which the wafers should be lifted from the solution.

In this process for G2 the Al is wet etched where in the process for G1 it was dry etched. The reason being that no dry etching process was available with a high selectivity to Si_3N_4 (the plate) that here serves as the etch stop layer. Therefore, using a dry etch is not feasible since the plate thickness is a critical parameter to control. After the Al etch the resist is stripped in an oxygen plasma and the wafer is ready to be diced.

4.4 Process flow - G3

The process flow for G3 is shown in Figure 4.5. The fabrication of the substrate wafer follows that of G1 with a few minor differences, which will be highlighted in the following. After the wafer bonding step the processing is almost identical to that of G2.

The first difference between the fabrication process of G3 and G1 is that in Figure 4.5 step **(2)** the Si_3N_4 is dry etched using a resist mask. Patterning the Si_3N_4 in this way reduces the number of processing steps but could lead to an increased surface roughness of the SiO_2 where the Si_3N_4 has been

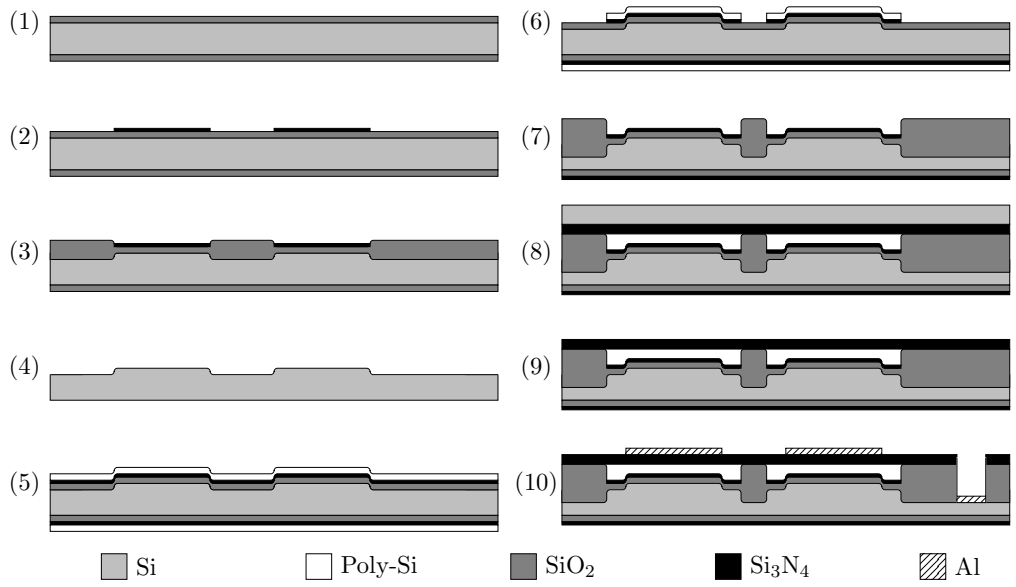


Figure 4.5: Process flow for the G3 CMUT. In step (1)-(4) the Si bumps are formed. Next, the cavities are made in step (5)-(7) and the plate layer is bonded to the cavity wafer in step (8). The handle Si layer is etched in step (9). Finally, in step (10), openings to the bottom electrodes are made and metal is deposited and patterned.

etched. If the surface roughness is too high ($R_{\mathrm{RMS}} > 0.5\,\mathrm{nm}$ [42]), the fusion bonding step can be rendered impossible. Before the G3 process was carried out, a small study investigating the surface roughness was conducted and is presented in Section 4.6.2. Based on these findings it was decided that dry etching the Si_3N_4 is feasible, and is used in step **(2)**. However, in step **(6)** a more conservative approach was chosen and the Si_3N_4 was wet etched using polysilicon as a mask. The remainder of the steps **(7)**-**(10)** are the same as for process G2.

Two AFM scans are shown in Figure 4.6 of the cavity before wafer bonding of a) an entire cavity and b) a zoom in on part of the Si bump and post SiO_2. In Figure 4.6 b) the Si_3N_4 edge is seen to be below the top of the post SiO_2 which is a requirement for a successful wafer bond. This is ensured by following the design rules presented earlier in Section 3.2.

4.5 Process flow - G4

The process for fabricating G4 is similar to a 'standard' process used to fabricate wafer bonded RIE CMUTs, with the exception of the Si_3N_4 plate and the Si_3N_4-Si_3N_4 bonding step.

Figure 4.7 shows the process flow for G4. In step **(1)** the thermal post SiO_2 is grown in a wet oxidation process at 1050 °C. Subsequently, a lithography process was performed and the cavities dry etched. The resist is stripped in an oxygen plasma and a RCA clean is performed. In step **(2)** a Si_3N_4 layer is deposited in a LPCVD furnace conformally coating the cavities. In step **(3)** a Si wafer with a Si_3N_4 layer, constituting the plate, is bonded to the substrate wafer and the handle Si layer is etched as in G2 and G3. Finally, in step **(4)** openings to the bottom electrode are dry etched and Al is deposited and patterned by a wet etch.

4.6 Process optimization

In this section the process optimization of selected process steps is presented. First, it is shown how thin Si_3N_4 surfaces can be wafer bonded, followed by an investigation of the effect of the RCA cleaning step, prior to bonding, on the chip yield. Then, the surface roughness is studied after several different processing steps, in order to determine if the Si_3N_4 oxidation mask can be dry etched. Finally, the minimum metal thickness required for wire bonding is found.

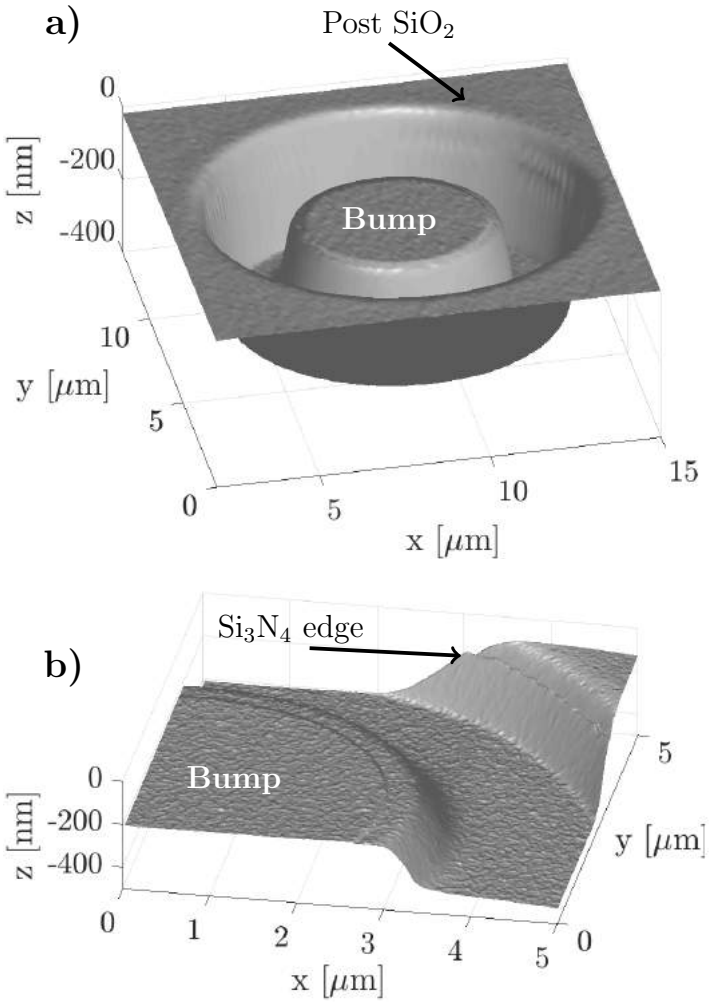


Figure 4.6: AFM scans of the cavity of a G3 device ($a = 5.4\,\mu$m) showing a) an entire cavity with the central Si bump and b) a close up of a cell (about a quarter) demonstrating that the Si_3N_4 is below the post SiO_2 and that the distance between the bump and post is flat.

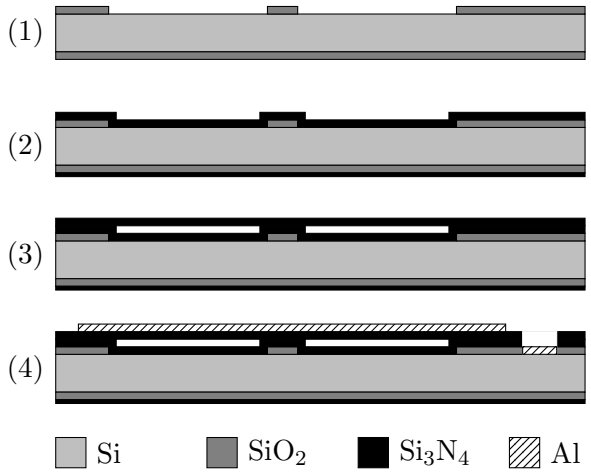


Figure 4.7: Process flow for the G4 CMUT. In step (1) the cavity is formed by a RIE etch. Subsequently, the wafer is coated with a LPCVD Si_3N_4 layer in step (2). A handle Si wafer with the Si_3N_4 plate is wafer bonded to the cavity wafer and the handle Si layer is removed in step (3). Finally, in step (4), openings to the bottom electrode are made, metal is deposited and patterned.

4.6.1 Wafer bonding

Wafer bonding of Si to SiO_2 was performed as part of the fabrication process of G1. These two surfaces are routinely bonded in our cleanroom and in the literature and will not be discussed further here.

Wafer bonding of Si_3N_4 to SiO_2 or Si_3N_4 is empirically found to be more difficult and is discussed here. The bond strength can be increased by oxidizing the surface of the Si_3N_4 layer [89]. This could be due to an increase of the number of silanol groups (Si-OH) on the Si_3N_4 surface which can form siloxane (Si-O-Si) bonds with the silanol groups on the SiO_2 surface, when annealed at a high temperature. Oxidized and non-oxidized Si_3N_4 surfaces were bonded to SiO_2 surfaces in order to check whether bonding could be achieved. The wafers with Si_3N_4 layers were oxidized at a temperature of 1100 °C for 3 h forming an oxy-Si_3N_4 layer [89]. Both the oxidized and non-oxidized Si_3N_4 wafers were successfully bonded to the SiO_2 surface of the substrate wafer. This means that the oxidization of the Si_3N_4 could be omitted, but if the surface roughness was increased, the oxidized layer helps with increasing the bonding strength [89].

In order to investigate the effect of the RCA cleaning step on the number of voids and hereby yield, an experiment was conducted where three wafer pairs were bonded (Si_3N_4 to SiO_2). Some of the wafers were RCA cleaned prior to wafer bonding and the chip yield was noted after the handle layer

Table 4.1: Table showing the chip yield after the wafer bonding step for wafer pairs exposed to different pre-cleaning procedures. For wafer pair 1 both the plate wafer and cavity wafer have been RCA cleaned prior to the wafer bonding process. For wafer pair 2 only the cavity wafer was RCA cleaned prior to wafer bonding, while no RCA cleaning was performed on either wafers for wafer pair 3. All wafer pairs were successfully bonded after the bonding process, although with different chip yields.

Wafer pair	RCA cleaned	Chip yield after bond
1	SiO_2 (cavity) and Si_3N_4 (plate)	100%
2	SiO_2 (cavity)	97%
3	None	92%

had been removed. Table 4.1 shows an overview of the experiment. The decrease in chip yield is a direct consequence of an increase in the number of voids, which increases when the wafers are not RCA cleaned. Even though the sample size is small, the tendency of an increased number of voids when the RCA step is skipped is observed in all processes. In conclusion, when the cleanliness of the handling of the wafers is worse the yield is reduced due to voids.

4.6.2 Dry etching of Si_3N_4

A study was conducted to investigate the evolution of the wafer surface roughness during processing when the Si_3N_4 oxidation mask, in the LOCOS process, is dry etched instead of wet etched. In the fabrication of G1 the Si_3N_4 layer was wet etched for both LOCOS processes and subsequent wafer bonding was possible. However, if the Si_3N_4 layer is dry etched the surface roughness could potentially exceed the limit of $R_{\mathrm{RMS}} = 0.5\,\mathrm{nm}$, which was stated earlier.

Table 4.2 shows the process steps of the two wafers. In the process for wafer 1 both the Si_3N_4 and SiO_2 layers are dry etched until the Si substrate is reached. For wafer 2 the Si_3N_4 layer is dry etched and the remaining SiO_2 layer is wet etched in BHF. In both processes a resist mask is used to pattern the surface. The theory is that finishing with a more gentle wet etch, as for wafer 2, the surface roughness is reduced. In between processing steps the surface roughness is measured with AFM where the bonding interface would be. That is, outside of the cavities.

Figure 4.8 shows a plot of the surface roughness for the two processes at different process steps. At step 10 the surface roughness of wafer 1, which

Table 4.2: Process steps of the two test wafers used to test the effect of the etching processes on the surface roughness. If a process step is skipped a "-" is shown. The structure at process step 13 is that of a single LOCOS cavity wafer just before wafer bonding.

Process step	Description	Wafer 1	Wafer 2
1	AFM	x	x
2	RCA	x	x
3	Oxidation	x	x
4	AFM	x	x
5	LPCVD Si_3N_4	x	x
6	Photolithography	x	x
7	Dry etch Si_3N_4	x	x
8	Dry etch SiO_2	x	-
9	BHF	-	x
10	AFM	x	x
11	Resist strip	x	x
12	Oxidation	x	x
13	AFM	x	x

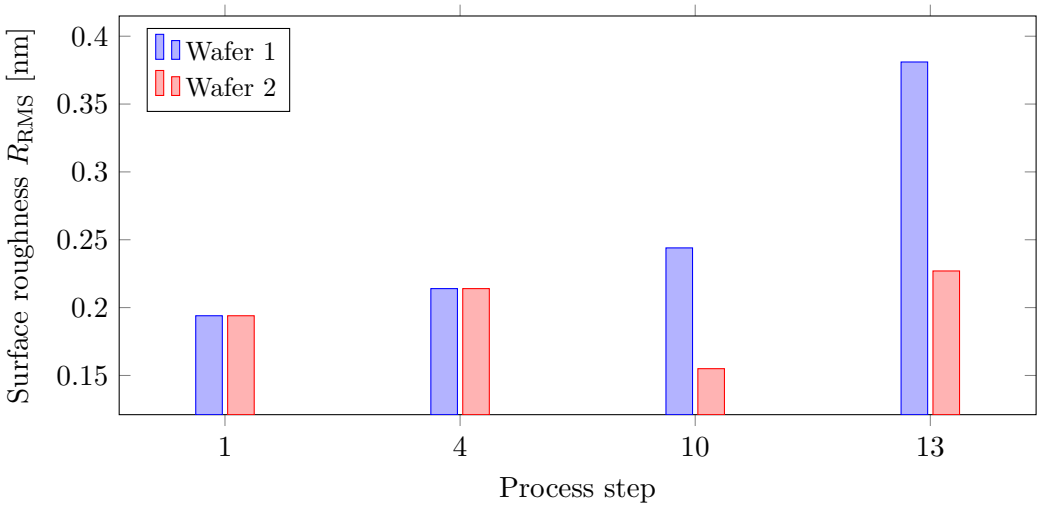


Figure 4.8: Surface roughness of the two wafers going through the processing steps in Table 4.2. Wafer 1, that is only dry etched, ends up having an increased surface roughness compared with wafer 2 that is both dry and wet etched.

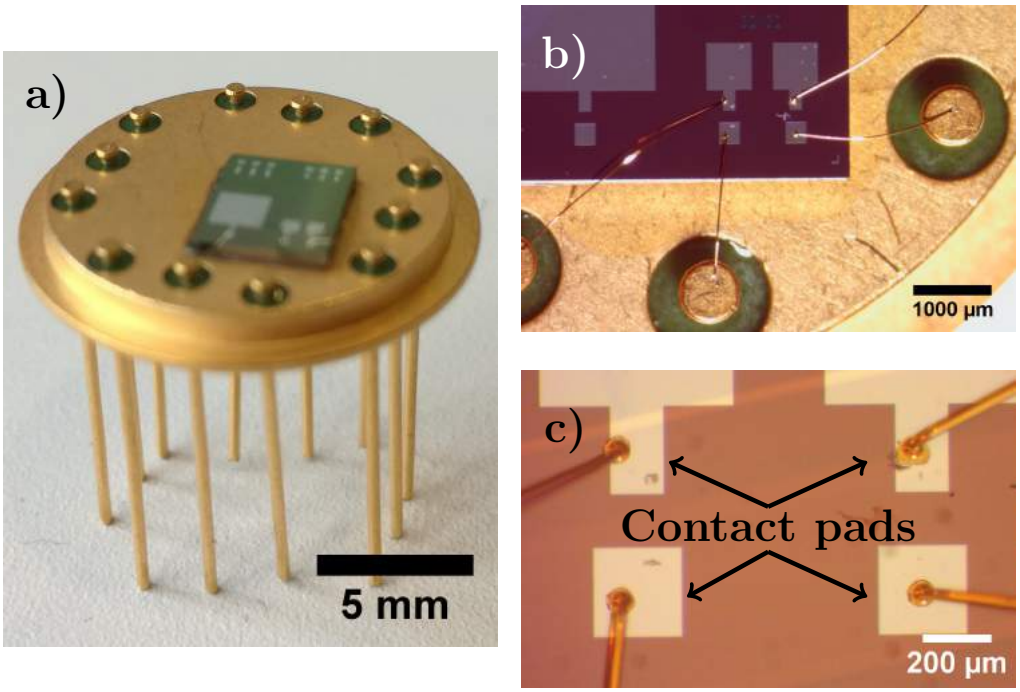


Figure 4.9: a) Photograph of the TO-8 chip carrier with a chip glued to the top. b) Optical microscope image of two CMUT elements wire bonded to the pins of the TO-8 substrate. c) Optical microscope image of the four wires contacting the top and bottom contact pads of the two CMUT elements.

had the Si_3N_4 and SiO_2 layers dry etched, is significantly higher than wafer 2 where the SiO_2 layer was wet etched. The tendency is the same after the resist has been stripped and the surfaces are oxidized in step 13. Furthermore, the overall surface roughness is increased but for both processes $R_{\mathrm{RMS}} < 0.5\,\mathrm{nm}$. Thus, both processes could potentially be used in a device fabrication process where bonding should be achieved.

4.6.3 Wire bonding

After fabrication is finished in the cleanroom, the wafers are diced into individual chips. The chips can be electrically contacted directly using a two-electrode probe or by wire bonding the chips to a carrier substrate. In this project the chips are ball wire bonded (K & S 4524 ball wire bonder) to a TO-8 12 pin chip carrier. Figure 4.9 a) shows a photograph of a TO-8 carrier with a G2 chip glued on top and wire bonded to the pads. Figure 4.9 b) shows a zoom in on the wire bonds going from the top and bottom electrode contact pads to the contact pads of the TO-8. Finally, Figure 4.9 c) shows

an even more zoomed in view of the top and bottom electrode contact pads each with a wire bond connected.

The Al thickness needed to successfully wire bond the contact pads were investigated in order to determine the lower Al thickness limit. Three wafers, with a Si_3N_4 surface, had Al deposited with thicknesses of 100 nm, 50 nm, and 10 nm, respectively. The wafers were annealed in order to increase the adhesion, as discussed previously. Wire bonds were successfully made on the wafers with 100 nm and 50 nm Al but not on the wafer with an Al thickness of 10 nm. Therefore, the minimum Al layer thickness used is 50 nm, which is the Al thickness used for the G4 chips.

4.7 Chapter conclusion

In this chapter the process flows and fabrication details of the four generations of CMUT chips have been presented. Furthermore, some of the process optimization carried out during the project is shown. The fabrication processes resulted in four generations of functioning CMUT chips and a better understanding of the limitations of the wafer bonding process.

CHAPTER 5

Characterization and results

In this chapter the CMUT generations are characterized with focus on the two figures of merit: mass sensitivity and limit of detection. First, the geometry of the chips is inspected by means of optical microscopy, secondary electron microscopy (SEM), and AFM. The measured geometries are used throughout the rest of the thesis when calculating e.g. the sensitivity or LOD. Next, impedance measurements provide the basis for a comparison of the resonance frequency and pull-in voltage with the simulated values from the FEM model. Furthermore, the capacitance and electromechanical coupling coefficient are found from the impedance measurements and used to compare the chip designs. The frequency noise is measured and the Overlapping Allan deviation is calculated which is then used to calculate the mass LOD. The LOD of the CMUTs fabricated in this project was found to be the lowest in the CMUT literature, to my knowledge. In addition, the LOD per surface area is also found to be the lowest compared with other CMUTs and compared with other gravimetric sensor types. Finally, the mass sensitivity is determined by three separate methods: analytically, using FEM, and experimentally. It is shown that the analytically determined sensitivity and the sensitivity found by the FEM model are in good agreement, which further validates the FEM model. The non-linearity of the sensitivity is observed for both the simulated sensitivities and measured sensitivities when a large mass ($m_{\text{added}}/m_{\text{plate}} > 1\%$) is added to the plate. This decreases the sensitivity as compared with the prediction from the analytical expression. The mass sensitivities are, to the knowledge of the author, the highest published for CMUTs and the normalized sensitivities are the highest among the CMUTs

but still lower than for the NEMS devices.

5.1 Chip designs

For each generation the chips have several CMUT elements where the number of cells and/or the top electrode design is varied. In the following, optical microscope images and Scanning Electron Microscopy (SEM) images are presented, representative of the chip and element designs fabricated. The top view optical microscope images can be used to determine the final horizontal dimensions of the device, while SEM microscopy is used to determine the vertical dimensions of the CMUT cells. It is not practical to use SEM for determination of the horizontal dimensions of a cell since the cells are all circular and the chips are cleaved by hand before inspection in SEM. Thus, the chance of the line of cleavage going through the middle of the cell is small. AFM can on the other hand give both the horizontal and vertical dimensions of a CMUT cell, but this has to be done prior to the wafer bonding step.

5.1.1 Optical microscopy

Figure 5.1 shows an optical microscope image of a finished G2 chip with a structured top electrode. The light areas are the Al surface and the green area is the Si_3N_4 plate on top of the SiO_2 post. The text is written in the cavity layer and the plate is most likely collapsed and broken in these areas, creating the non-uniform coloring inside the letters and numbers as seen in b) for the text "10X10". Figure 5.1 a) shows an element with 100 cells on each side of the square resulting in a total of $100 \times 100 = 10000$ cells. Further, in a) two elements each with $32 \times 32 = 1024$ cells are shown. Figure 5.1 b) shows three elements each with $10 \times 10 = 100$ cells. These two microscope images give a good indication of the relative size of the top electrode contact area to the active cell area.

Figure 5.2 shows two optical microscope images of the same 1024 cell element on a G4 chip. In this design the top electrode is not etched between the cells. In a) the microscope focus is on the metal layer while in b) the focus is in further down, in the cavity. This makes the individual cells visible through the metal and plate layer. This is possible since the metal and plate layers only have a thickness of 50 nm.

The element in Figure 5.2 demonstrates that a cell to cell distance of only $a_{\text{bond}} = 1\,\mu$m can be achieved, at least when the cells are fabricated with the RIE process (recall the reduced cell to cell distance for the LOCOS cavities, see Figure 3.27).

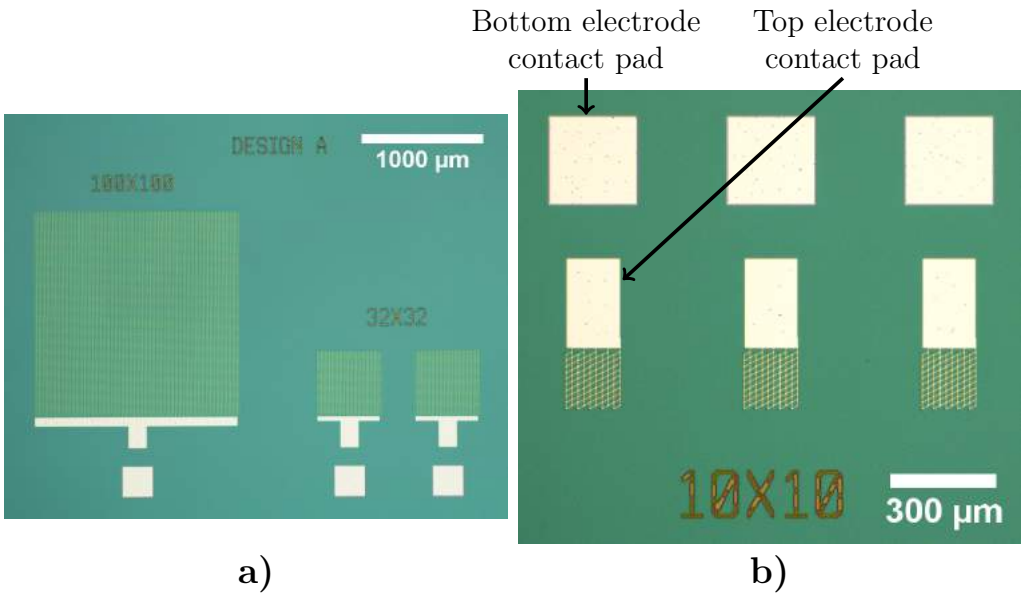


Figure 5.1: Optical microscope image of a G2 chip with a structured top electrode showing a) elements with 100x100 and 32x32 cells and b) three elements with 10x10 cells.

5.1.2 Scanning electron microscopy

In order to make cross-sectional SEM images of the CMUT cells, the CMUT chips must be cleaved. The cleaving process consists of making a scratch at the edge on the back of the wafer or chip along the cleaving direction. The scratch is made with a handheld wafer scriber. Designs with a conducting plate, G1, can readily be imaged but designs with a non-conducting plate, G2-4, must be coated with a thin ($10\,\text{nm} - 20\,\text{nm}$) Au layer which is sputtered on the cleaved chip. An inlens secondary electron (SE) detector is used as it is primarily the topography information that is wanted. The acceleration voltage is typically 5 kV with a working distance of $< 10\,\text{mm}$, and a vacuum of $< 10^{-4}\,\text{mbar}$.

Figure 5.3 shows cross-sectional SEM images of several CMUT cells in three elements for a) G1, b) G2, and c) G2. The image in a) is taken before the plate was thinned down and the metal electrode deposited. In contrast, the images in b) and c) are of finished functioning devices. These images provide another perspective on the massive parallelism of the cells in an element.

Figure 5.4 shows SEM images of single CMUT cells in an element for all four generations. Again, the image in a) is taken before the plate is thinned down and metal has been deposited. Thus, the plate is $1.3\,\mu\text{m} - 1.4\,\mu\text{m}$ thick

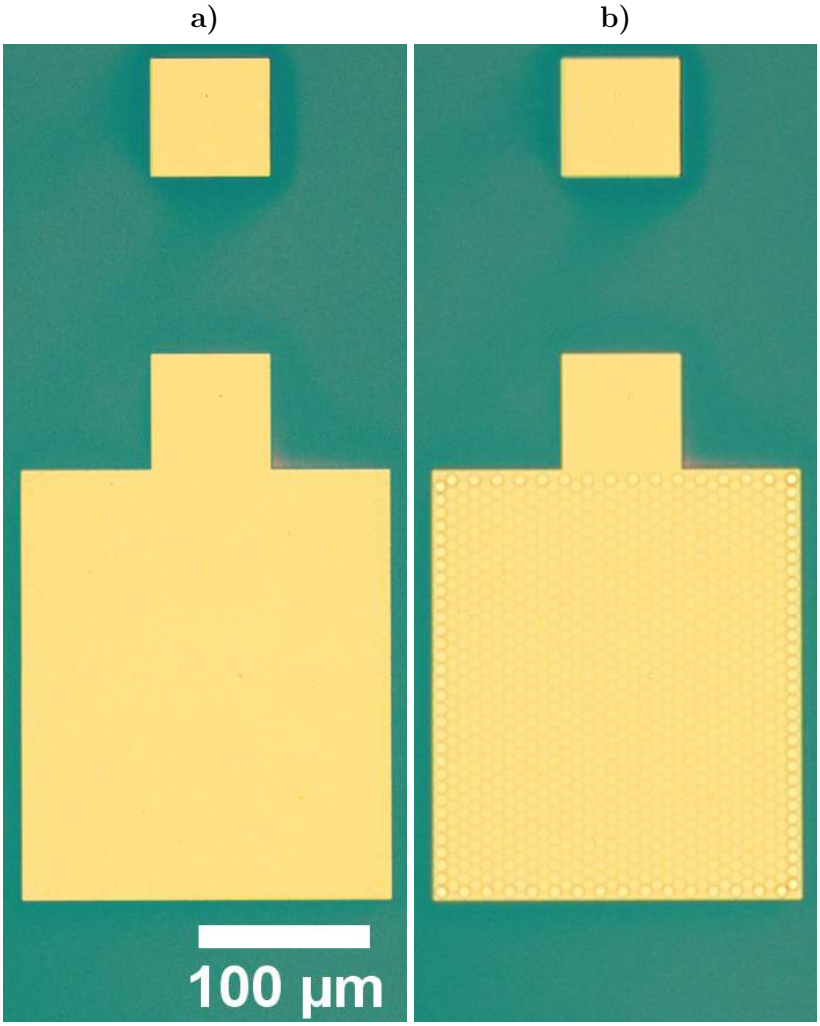


Figure 5.2: Optical microscope images showing a G4 element with 32x32 cells and a inter cell distance of $a_{\text{bond}} = 1\,\mu$m. The difference between a) and b) is the focus point with lies deeper into the plane for b) making the cells visible through the plate.

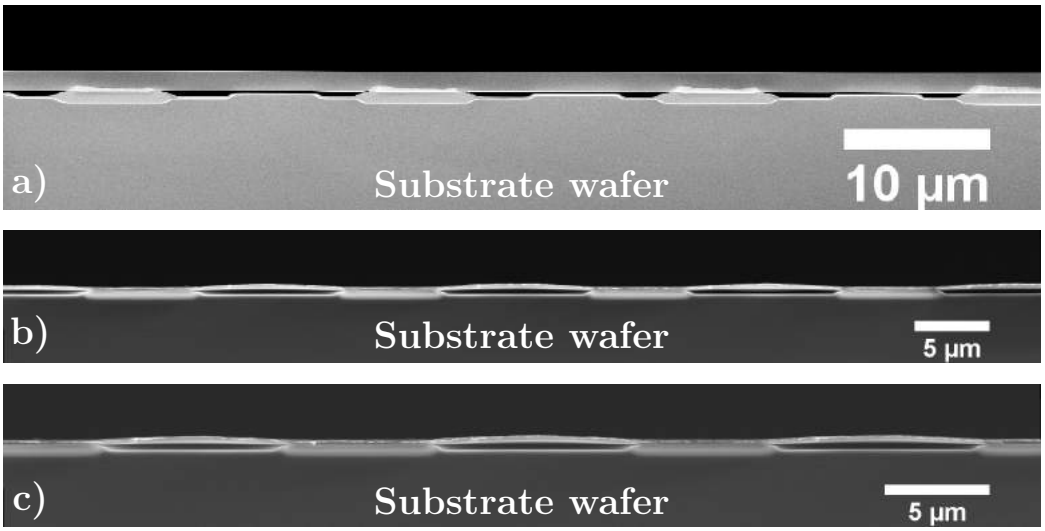


Figure 5.3: SEM images of chip cross-sections through elements showing multiple cells for a) G1 (before the plate has been thinned down), b) G2, and c) G2.

and no metal can be seen on the plate. At either side of each image the post SiO_2 layers are visible. The difference in the ratio $t_{\text{post}}/t_{\text{vacuum}}$ for the double LOCOS, single LOCOS, and RIE cavities can be seen when comparing the images. Looking at the images it is apparent why the highest ratios are expected for the double LOCOS cavity and by extension the highest cell to parasitic capacitance and hereby the highest electromechanical coupling coefficient.

The Al top electrode is visible at Figure 5.4 b), c) and d) on top of the plate. The edge of the Al layer is in all these images torn and rough, whereas the Si_3N_4 edge looks completely smooth and straight. This is most likely due to the Al being much more ductile than the brittle (and highly stressed) Si_3N_4 layer. Since the chip is cleaved the resulting interfaces of the Si_3N_4 and Al are to be expected.

The alignment of the Si bump to the cavity can be inspected using the SEM images in a) and c). A SEM image with a misaligned Si bump can give the *minimum* misalignment value since the direction in which the circular cells are cleaved is not controlled. The real misalignment value can be found from AFM scans or optical microscope images. It is clear that the Si bump in c) is misaligned relative to the cavity and the *minimum* misalignment value is from this image found to be $0.75\,\mu$m. This value is lower than the alignment precision of the operator and equipment of $1\,\mu$m stated in Section 4.1.2. This example shows that, for the double LOCOS structure, alignment becomes increasing important and demanding as the cell radius decreases.

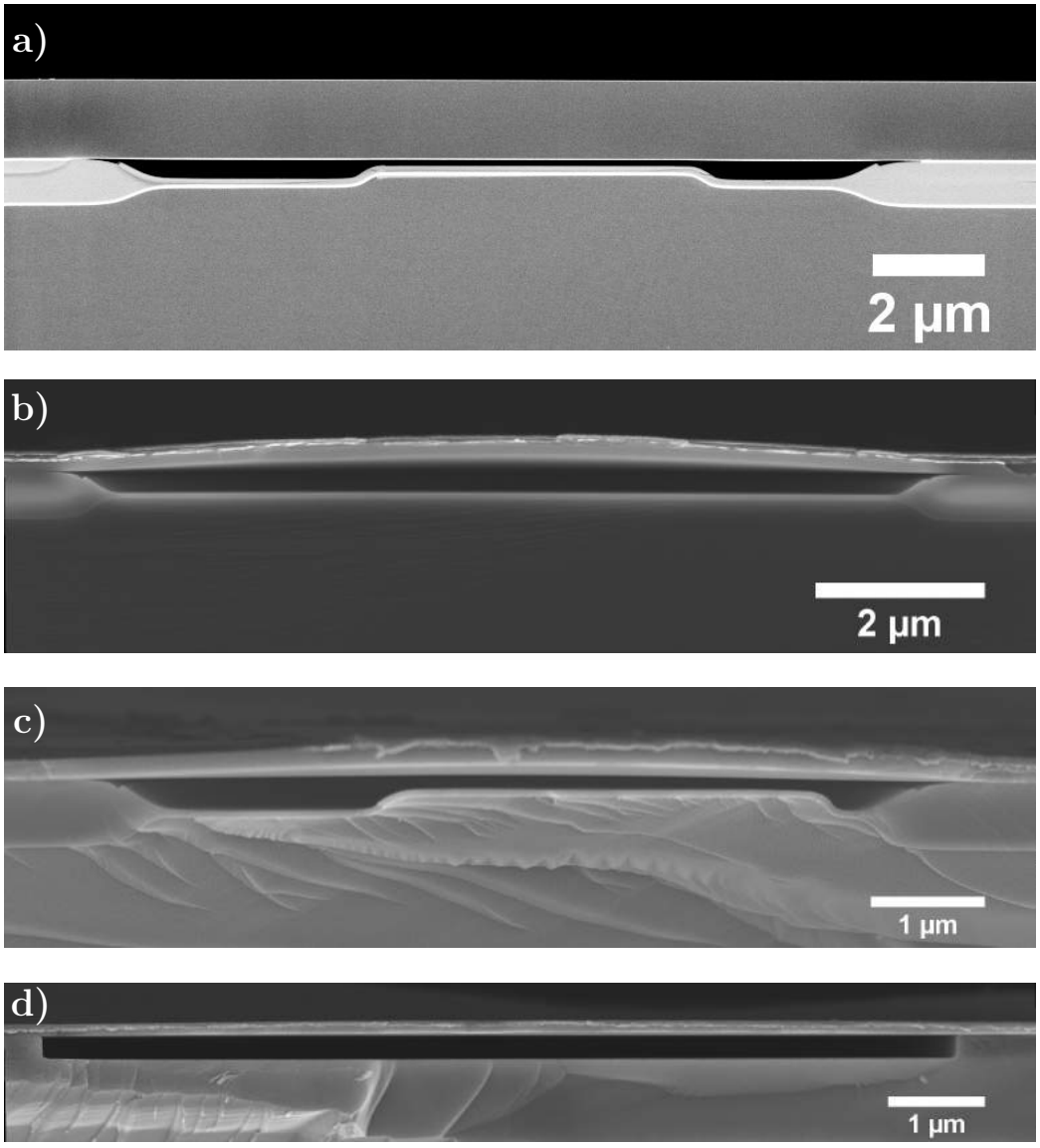


Figure 5.4: SEM images of cross-sections of single cells for all generations: a) G1 (before the plate has been thinned down), b) G2, c) G3, and d) G4. Note that the images are not on the same scale.

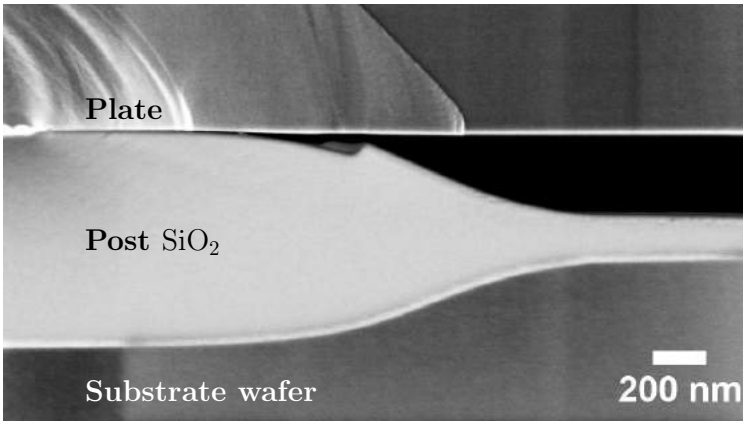
Several solutions exist: improving the alignment marks used by including Vernier alignment marks, that both help with the alignment and with assessing the amount of misalignment after exposure. Another solution is to switch technology to an aligner with a better alignment accuracy; e.g. a DUV stepper. As an example, the DUV stepper available in the Danchip cleanroom (Canon, FPA-3000EX4) has an alignment accuracy of 50 nm which is 20 times lower than the current UV aligner used. As an added benefit by switching to a DUV stepper, the minimum critical dimension could be reduced down to a about 250 nm if needed. A final solution to the alignment challenge is to used either a single LOCOS cavity or RIE cavity, thus eliminating the small tolerance alignment step. The downside of this is an increased parasitic capacitance.

Figure 5.5 shows two SEM images of the bird's beak structure for the two double LOCOS structures a) G1 and b) G3. Notice that the images in Figure 5.5 all are on the same scale. It is interesting to compare the cavity slope lengths for the bird's beak structures (introduced in Section 3.2). The slope length is expected to be longer for G1 a) compared with G3 b). This is due to the thicker t_{post} and t_{pad} for G1. The slope length is measured for a) and b) and is found to be 1.53 μm and 0.79 μm, respectively. Thus, the slope length is almost twice as long for G1 compared with G3. This example demonstrates that using the optimized parameters found in Section 3.2, decreases the slope lengths for real devices.

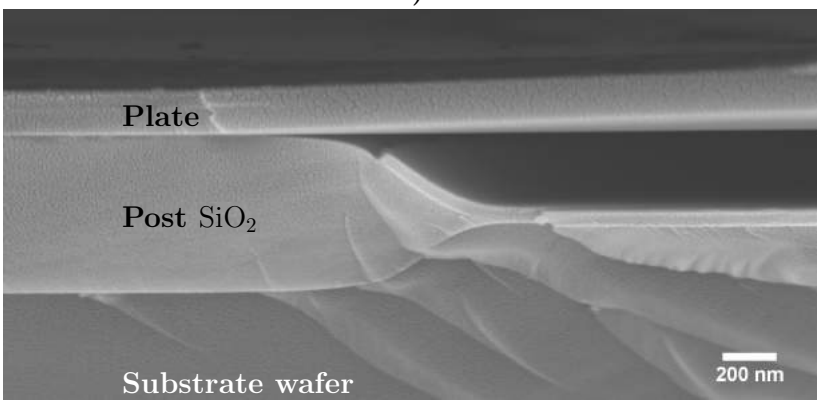
Finally, Figure 5.5 c) shows the cavity edge for G4, which compared with a) and b) takes up significantly less horizontal space since the slope length is zero. The scale of image c) is the same as a) and b) and they can therefore be compared directly. Being able to eliminate the slope length is important for cells with radii on the same length scale as the LOCOS slope lengths themselves.

5.2 Impedance analysis

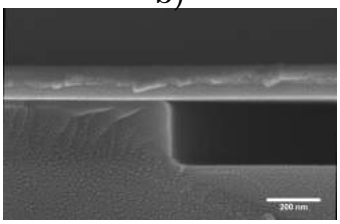
In this section the characterization of the CMUTs using impedance spectroscopy is presented. Impedance spectra contain a wealth of information, as discussed in Section 2.3.3, such as the resonance and anti-resonance frequency and capacitance. Furthermore, by fitting the lumped element model, all the lumped parameters can be extracted and derived values such as the electromechanical coupling coefficient can be calculated. Many of these quantities are functions of the applied bias voltage which in an experimental setting is easily varied. In the following, impedance spectra and the most interesting derived values are shown for selected devices.



a)



b)



c)

Figure 5.5: SEM images of the bird's beak structure at the cavity edge a) G1 and b) G3. c) Cavity edge for the RIE etched cavity G4. The images share the same scale, that is, the lengths can directly be compared. This illustrates the difference of space taken up by the cavity edge for the different designs.

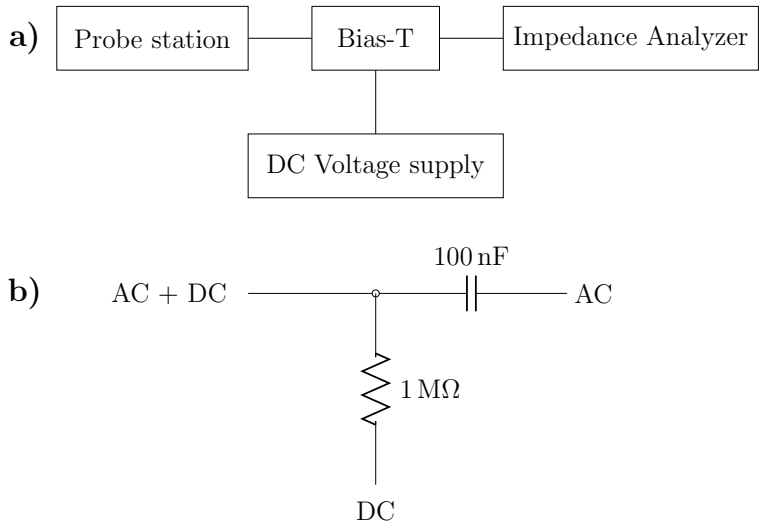


Figure 5.6: a) Diagram of the experimental impedance setup showing the probe station to the left connected to a bias-t which decouples the DC voltage from the AC signal to the impedance analyzer. b) Simplified schematic of the bias-t.

5.2.1 Measurement setup

Figure 5.6 a) shows a schematic overview of the impedance measurement setup. The probe station is a Cascade Microtech semi-automatic Summit 12000 wafer probe station fitted with either an ACP or Infinity series probe also from Cascade Microtech. The probes have a pitch slightly larger than the distance between the top and bottom electrode contact pads on the chips. Figure 5.7 shows a photograph of the probe station and impedance analyzer (Keysight, E4990A). A simplified electrical circuit of the bias-T is shown in Figure 5.6 b), where the CMUT is fed an AC signal and biased by a DC voltage but the impedance analyzer sees only the AC signal. The bias-T is implemented on a two layer PCB with surface-mounted components. The ground layer of the PCB is removed under the signal path from the impedance analyzer to the CMUT in order to decrease the capacitive coupling.

5.2.2 Impedance spectra

Impedance spectra contain a lot of information about the CMUT device and are therefore extremely useful for the device characterization. In this section impedance measurements are used to determine the: resonance and anti-resonance frequencies, pull-in voltage, capacitance, and electromechanical coupling coefficient. Results are shown for selected devices, demonstrating the trends shown in the Theory and Design chapters.

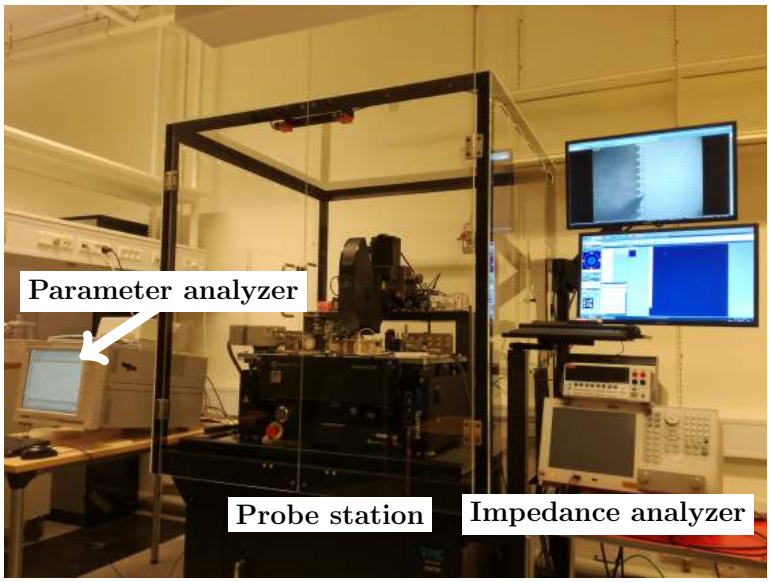


Figure 5.7: Photograph of the probe station next to the parameter analyzer and impedance analyzer.

Figure 5.8 shows a plot of the impedance magnitude and phase angle as a function of frequency for a G1 device. All the features seen in the theoretical impedance spectrum in Figure 2.13 are reproduced in the measured impedance spectra. From the spectra the resonance and anti-resonance frequencies can immediately be determined which in this case are 22.95 MHz and 23.72 MHz, respectively.

The resonance frequency and pull-in voltage, calculated using FEM with the designed geometrical parameters, can now be compared with the measured f_r and V_{PI}. Furthermore, the measured (optical microscopy and SEM) geometrical values are used as input in the FEM model and the resulting f_r and V_{PI} are compared with the measured values. Table 5.1 shows the designed and measured geometrical values for the G4 device. These values differ and the designed and measured f_r and V_{PI} are therefore expected to differ. Indeed, this is observed as shown in Table 5.2 where the pull-in voltage, calculated using the designed geometrical parameters, differs by 25 % compared with the measured value. However, when the measured geometrical parameters are inserted in the FEM model the relative difference in pull-in voltage is only 1 %, relative to the measured pull-in voltage. This is a testament to the predictive capabilities of the FEM model. Moreover, this example demonstrates the importance of being able to measure the geometry of the fabricated devices. Finally, the resonance frequency is only affected by the thicker plate of the fabricated device and the resulting relative difference

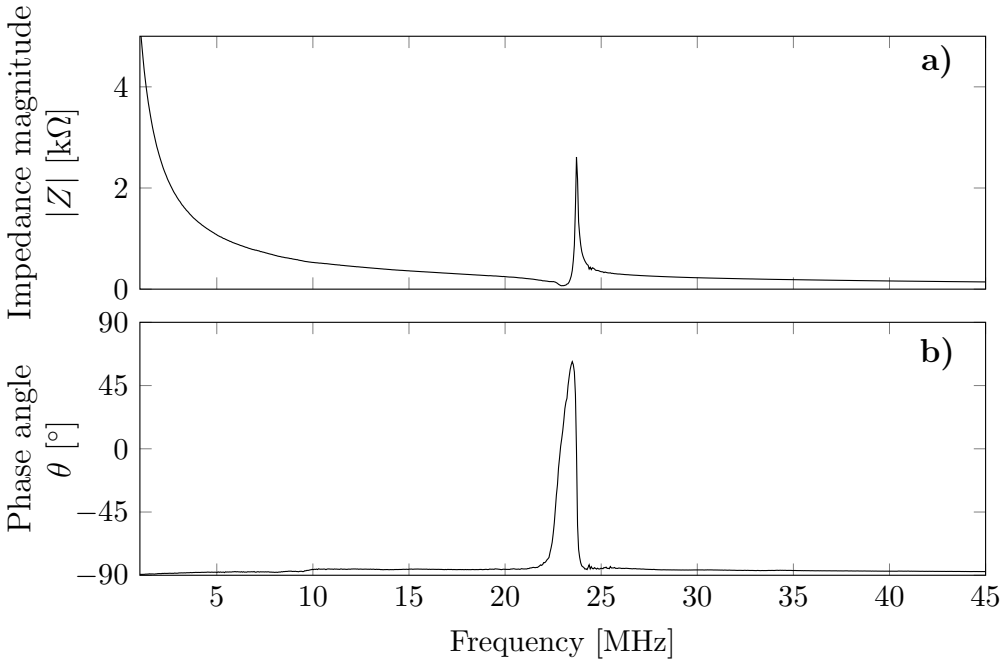


Figure 5.8: a) Impedance magnitude and b) phase angle as a function of frequency. The spectra are measured for a device on a G1 chip with no metal on the Si plate ($a_{\text{electrode}}/a = 0$). The CMUT is biased at $V_{\text{DC}} = 0.84 V_{\text{PI}}$.

of the resonance frequencies is only on the order of 1 %. Again the FEM model with the measured parameters as input results in a output closer to the measured value.

This example can be seen as a validation of the FEM CMUT model. Being able to predict the behavior of the devices is crucial for the design of new generations.

Table 5.1: Example of designed and measured geometrical values for G4.

Parameter	**Unit**	**Designed**	**Measured**
Plate radius, a	[μm]	5.0	-
Plate thickness, t	[nm]	50	55
Vacuum gap height, t_{vacuum}	[nm]	200	235
Insulation nitride thickness, t_{nitride}	[nm]	20	27.3

Table 5.2: Resonance frequency and pull-in voltage of G4 found by FEM using both the designed and calculated geometrical values and by impedance measurements. The results from the FEM model are in close agreement with the values from the impedance measurement when the measured geometry is used.

	Resonance frequency f_r **[MHz]** ($V_{DC} = 0\,V$)	**Pull-in voltage** V_{PI} **[V]**
FEM - designed geometry	41.2	68
FEM - measured geometry	42.4	92.2
Measurement - impedance	42.0	91

5.2.3 Effect of the bias voltage

A bias voltage is needed in order to increase the electromechanical coupling coefficient. As the bias voltage is increased the resonance frequency decreases, an phenomenon known as the spring softening effect which was introduced in the Theory chapter. Figure 5.9 shows an impedance spectrum for a G1 device biased at three different voltages. Increasing the voltage causes a shift in resonance frequency towards lower frequencies, demonstrating the spring softening effect. Further, increasing V_{DC} increases the maximum impedance peak and the maximum phase shift, and finally broadens the phase peak as the resonance and anti-resonance frequencies move further apart. This behavior is indicative of an increased coupling coefficient due to the increased V_{DC}.

Figure 5.10 shows a plot of the resonance frequency as a function of the bias voltage. The resonance frequency decreases until pull-in is reached. The last measurement point is just before pull-in occurs. The device is seen to be most non-linear closest to the pull-in voltage where the largest frequency shifts are observed. This is in agreement with the theory, particularly with Figure 2.10 showing the same tendency for the effective spring constant. An increased bias voltage and hereby decreased resonance frequency will decrease the mass sensitivity as will be discussed later in Section 5.5.

Figure 5.11 shows a plot of the normalized capacitance as a function of the normalized bias voltage for three designs. As the voltage is increased the capacitance increases as well since the plate deflects and the gap height decreases. The relative change in capacitance is determined by the amount of parasitic capacitance relative to the cell capacitance. This is reflected in the relative capacitance change for the two G2 devices where one has a patterned top electrode (as in Figure 3.23 b)) and the other does not (Figure 3.23 a)). The parasitic capacitance is expected to be lowest for the device with the

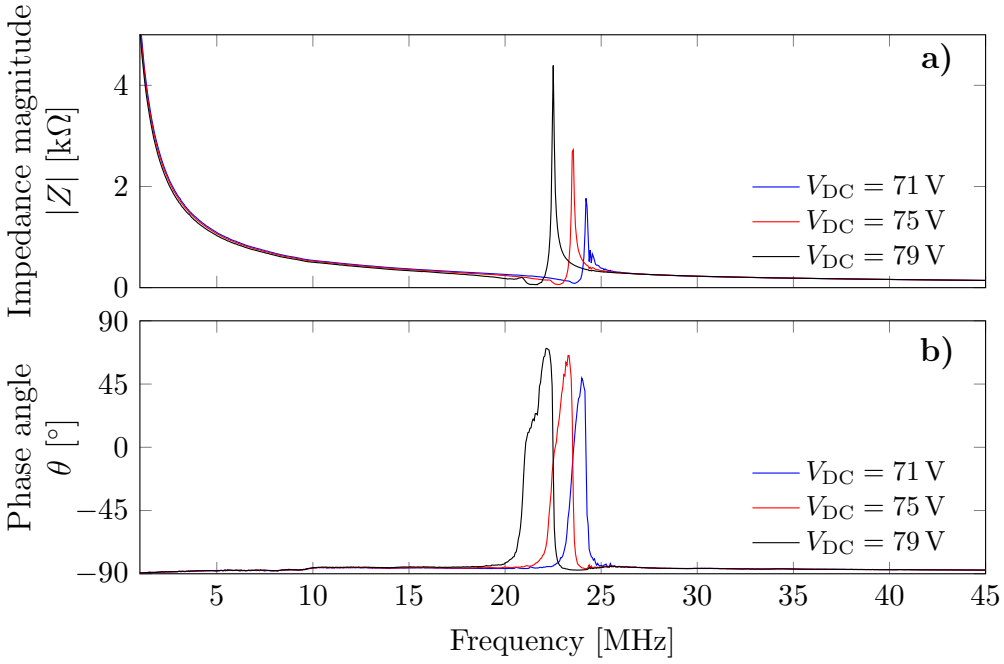


Figure 5.9: Impedance spectra of G1 for three bias voltages.

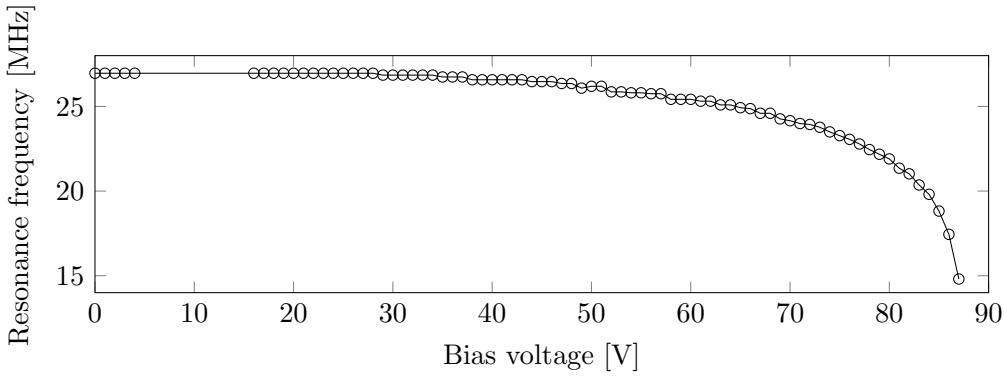


Figure 5.10: Resonance frequency as a function of bias voltage, showing the spring softening effect. G1 chip with no metal on the Si plate ($a_{\text{electrode}}/a = 0$).

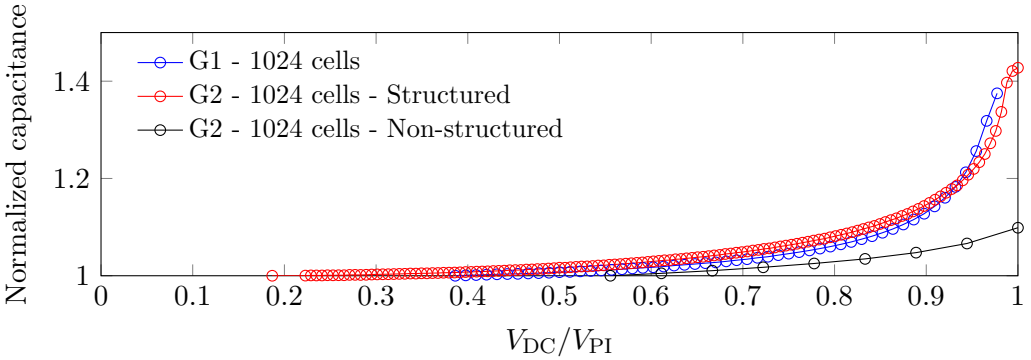


Figure 5.11: Normalized capacitance as a function of normalized bias voltage for three different CMUT designs.

patterned/structured top electrode and hence the relative change in capacitance is expected to be highest. This is exactly what is observed in Figure 5.11 where the capacitance of the G2 device with a structured top electrode increases more than the G2 device without a structured top electrode. The capacitance of the G1 device is seen to increase the most, indicating the lowest parasitic capacitance, despite of the conducting Si plate. This is most likely due to the G1 cavities being of the double LOCOS type where the G2 cells are of the single LOCOS type. Although this last comparison of G1 and G2 is complicated by different cell radii and top electrode contact areas.

Finally, Figure 5.12 shows a plot of the electromechanical coupling coefficient as a function of the bias voltage. These values are found by fitting the BVD equivalent circuit to the impedance spectra and subsequently calculating k^2 as described in the Theory section. As the bias voltage increases so does k^2, as it should according to the theory. The highest k^2 values are in Figure 5.12 found for the G1 device, which is in line with what was seen for the capacitance in Figure 5.11. Furthermore, comparing the two G2 designs, the highest k^2 is found for the device with a structured top electrode which is a direct consequence of the lower parasitic capacitance. The efforts of decreasing the parasitic capacitance can thus be directly observed as a higher electromechanical coupling coefficient, which was the purpose. The same effect is observed (not shown here) when increasing the number of cells, as was predicted in Section 3.4.1 due to the large contribution of parasitic capacitance from the top electrode pad. In conclusion, applying the theoretical ways of reducing the parasitic capacitance of real devices, results in an increase of the electromechanical coupling coefficient.

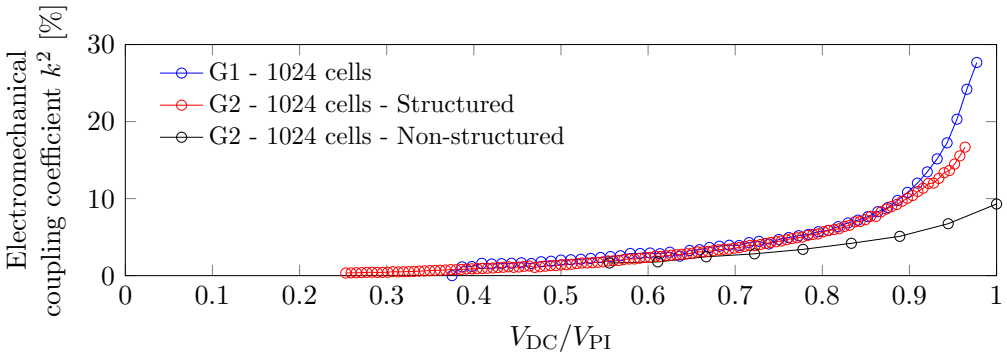


Figure 5.12: Electromechanical coupling coefficient as a function of normalized bias voltage for three CMUT designs. The designs with the highest relative parasitic capacitance shows the lowest coupling coefficient.

5.2.4 Partial conclusion

In this section the impedance of the CMUTs was measured and analyzed. First, the measurement setup was presented and then the measured resonance frequency and pull-in voltage was compared with the predicted values from the FEM model. These were found to be in excellent agreement with each other. Then, the spring softening effect was demonstrated and it was shown that a design with a lower parasitic capacitance displays a higher electromechanical coupling coefficient.

5.3 Dielectric charging

In this section the well known issue of dielectric charging is discussed. Due to the high electrical fields in the vacuum gap and over the dielectric layers, charges can be trapped in these layers or existing charges can be moved by the electrical field. The consequence of having charges in the CMUT is a different electrical field in the CMUT than what is expected from the applied external bias voltage. This will influence the behavior of the CMUT. The charges can either be mobile, that is able to move when an electrical field is applied, or static, that is, fixed to one position. The effect of the mobile charges will be observed as time dependent changes to the CMUTs characteristics, e.g. the capacitance. While the static charges will give an offset in the effective electrical field in the CMUT which will show as an offset in the characteristics of the CMUT.

In the following measurements are shown for CMUTs that do not show signs of charging and that do show signs of charging. Both cases are seen for CMUTs fabricated in this project.

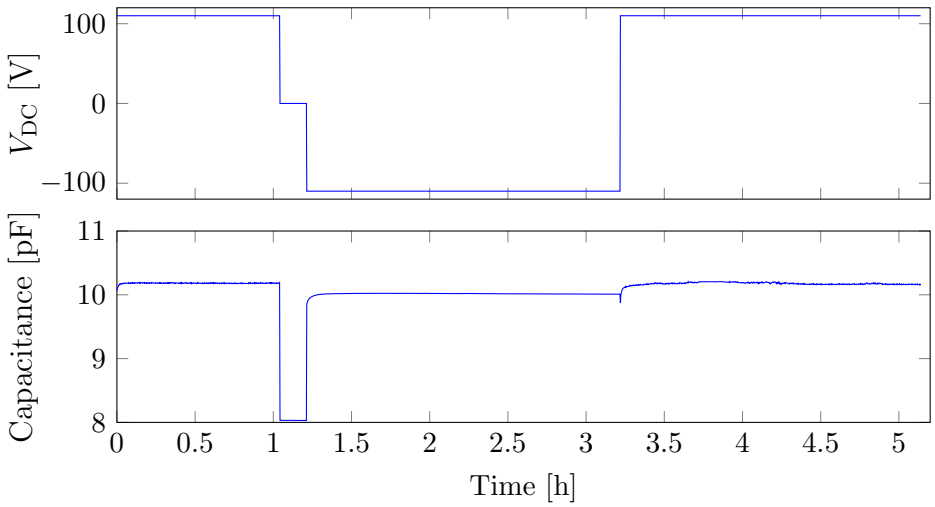


Figure 5.13: Bias voltage and capacitance as a function of time for a G2 element with 1024 cells. The capacitance is seen to be constant over the relative long period of time.

5.3.1 Experimental

Impedance and CV measurements are performed in order to investigate the charging behavior in the CMUTs. Figure 5.13 shows a plot of the bias voltage and capacitance as a function of time for an element on a G2 chip. The data are obtained by measuring the impedance and extracting the capacitance. The purpose of the measurement is twofold: does the capacitance change over a long period of time with a constant bias and how does the capacitance respond to sudden changes in the bias voltage? The capacitance is seen to be constant over periods of several hours, which answers the first question above. When changing the bias voltage the capacitance take some time to reach the maximum value, this is seen at $t = 0\,\text{h}$, $1.25\,\text{h}$, $3.25\,\text{h}$. However, as will be shown, the relative change in capacitance at these positions is small compared to CMUTs that display charging behavior. The capacitance is seen to be slightly higher when a positive bias is applied to the top electrode than when the negative bias is applied. This can be indicative of fixed charges in or on the dielectric layers.

Figure 5.14 shows a plot of the bias voltage and capacitance as a function of time for an element with 100 cells on a G4 chip. Again, the bias voltage is varied in steps and the resulting capacitance is measured. Many effects can be seen in this plot. First, when the negative voltage is applied the capacitance is actually below that of the capacitance at $0\,\text{V}$. The capacitance at positive bias voltages is seen to be the highest, which was also seen in Figure 5.13.

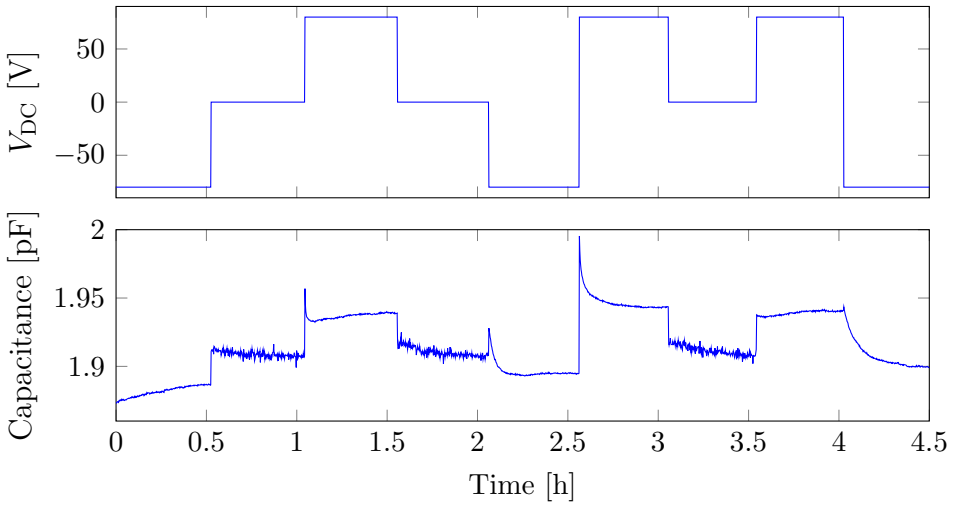


Figure 5.14: Bias voltage and capacitance as a function of time for a G4 element with 100 cells. The capacitance is seen to be time dependent and dependent on the bias in the previous step.

Second, the capacitance is seen to change as a function of time and it is dependent on the previous biasing voltage. Consequently, it is speculated that both fixed and mobile charges are present in this CMUT. The contrast to the behavior of the CMUT in Figure 5.13 is striking.

Capacitance-voltage measurements can also give insights into the charging behavior of the CMUTs. Figure 5.15 shows a CV curve for a G2 chip. The minimum of the capacitance curves are offset from $V_{DC} = 0\,V$ and the forward and backward sweeps show hysteresis in that they are not positioned on top of each other. The first effect of the minimum offset is usually seen for CMUTs with a fixed charge and the hysteresis behavior is seen for CMUTs with mobile charges. Thus, this CMUT demonstrate the same charging behavior as the one in Figure 5.14.

In conclusion, the CMUTs fabricated in this project both did and did not show signs of charging. The focus of this project has not been to solve this problem which sometimes appear. The precise origin of the charging and what causes the fixed and mobile charges is still unknown and more research is needed in order to investigate this phenomenon. However, it is our belief that the cleanliness of the process equipment and cleanliness of the general wafer handling in the cleanroom is important in order to minimize the effects of charging.

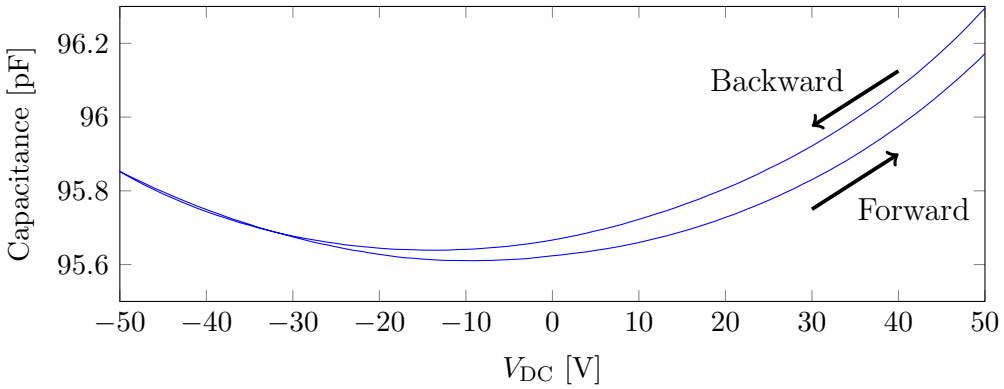


Figure 5.15: Capacitance as a function of bias voltage for a G2 element with non-structured top electrode. The CV curves show that both static and mobile charges are expected to be present due to the shift away from 0 V for the minimum and the hysteresis of the curves, respectively.

5.4 Mass limit of detection

In this section the mass limit of detection is determined through measurements of the frequency noise which is used to calculate the Allan deviation. The experimental setup for such measurements is presented and finally the mass limit of detection is compared with the state of the art.

In order to measure the Allan deviation an external circuit is required that can drive the CMUT at resonance while tracking this frequency. This can be achieved in a variety of ways but the two most common (all electronic) ways of doing this is by use of an oscillator circuit such as a Colpitts oscillator [21] or by using a phase-locked loop (PLL). For this project a PLL was used and the experimental setup, including the PLL, is described in this section.

5.4.1 Experimental setup

In this section the experimental setup used for determining the Allan deviation and tracking the resonance frequency is presented. An overview of the setup is shown in Figure 5.16(a). Here the CMUT is connected to a phase-locked loop (PLL) and a bias voltage is applied through a bias-T. The objective of the PLL circuit is to track the resonance frequency of the CMUT so that it can be driven at this frequency. This is done by maintaining a constant phase shift between the drive/output signal and the input signal. A resonance frequency shift of the CMUT will affect the position in frequency of the phase shift and the PLL will try to adjust the frequency of the drive signal so the target phase shift is maintained. The CMUT is hereby driven at resonance and the resonance frequency can be tracked.

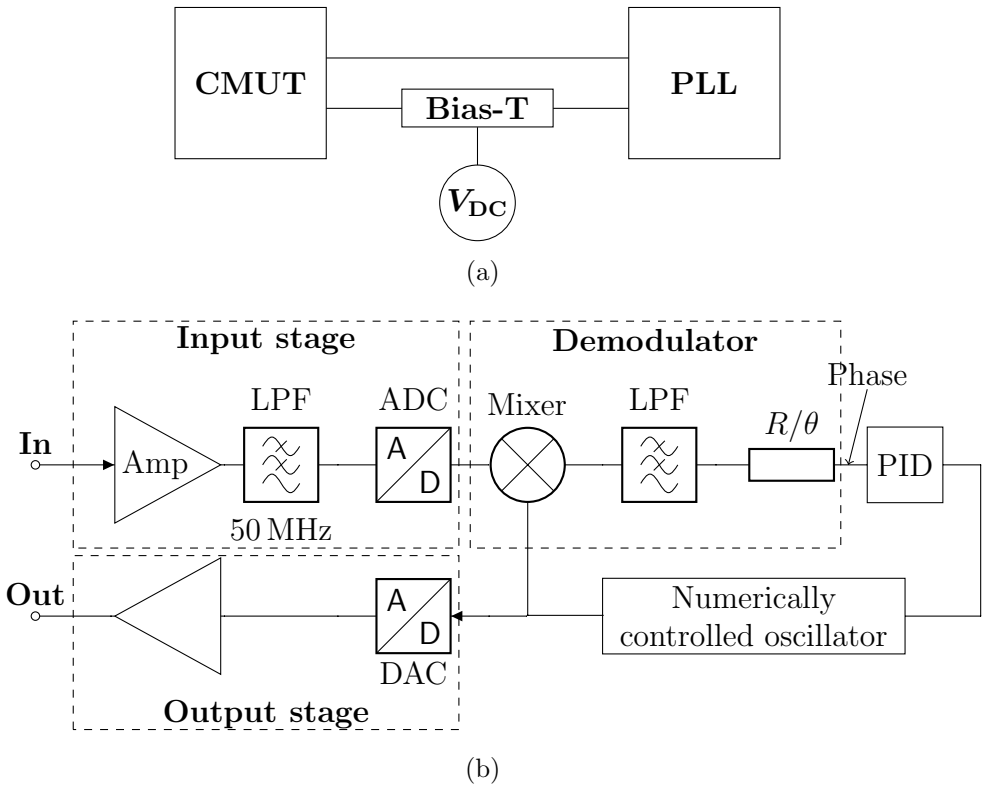


Figure 5.16: (a) Schematic of the experimental setup of the CMUT and PLL. (b) Block diagram illustrating the most important blocks of the PLL in the HF2LI-PLL (Zurich Instruments).

The PLL is shown in more detail in Figure 5.16(b). The PLL is digital and integrated in a lock-in amplifier HF2LI-PLL (Zurich Instruments). Following the signal from left to right in Figure 5.16(b) from the point marked "In" the signal is first amplified and then it passes through a low-pass filter (LPF) with a cutoff frequency of 50 MHz, since this is the maximum frequency of the lock-in amplifier. This is part of the reason why all the CMUT generations had designs which had resonance frequencies under 50 MHz. Next, the signal is converted from an analog signal to a digital signal in an analog-to-digital converter (ADC). This marks the end of the input stage. The signal then enters the demodulator where it is mixed with a signal from the internal oscillator. The mixer produces an output signal which describes the phase difference between the two input signals. This output signal is then passed through another LPF after which the signal is converted to polar coordinates. The phase signal is now passed to the proportional–integral–derivative (PID) controller. This controller maintains a constant phase difference, chosen by the user, between the input and reference signal. This is achieved by changing the frequency of the oscillator. Hence, when the resonance frequency of the CMUT is perturbed, the phase difference changes and the output signal must shift in frequency if the constant phase difference is to be upheld. Finally, the output signal is converted to an analog signal in a digital-to-analog converter (DAC) and possibly amplified before it is sent back to the CMUT.

The process of 'locking in' is as follows: first a bode plot is typically measured for the CMUT showing the resonance peak. The phase shift at resonance is then used as a set-point for the PID controller. Lastly, the bandwidth (BW) of the LPF in the demodulator, after the mixer, is chosen. A higher BW allows for more noise in the signal giving a lower SNR but at the same time a high BW increases the speed at which the oscillator can change its frequency. Hence, if too small a BW is chosen the rate at which the frequency can change might be too low in order to follow the actual resonance frequency changes of the CMUT due to external perturbations. This is discussed more in the next section.

Figure 5.17 shows a photo of the experimental setup that was schematically shown in Figure 5.16(a). At the bottom of the figure a bias voltage is supplied through a coaxial cable to the PCB which functions as a bias-T and mechanical holder for the CMUTs that are wirebonded to TO-8 carriers. The PCB is connected to the lock-in amplifier with the PLL through two to four coaxial cables depending on whether one or two channels are to be used.

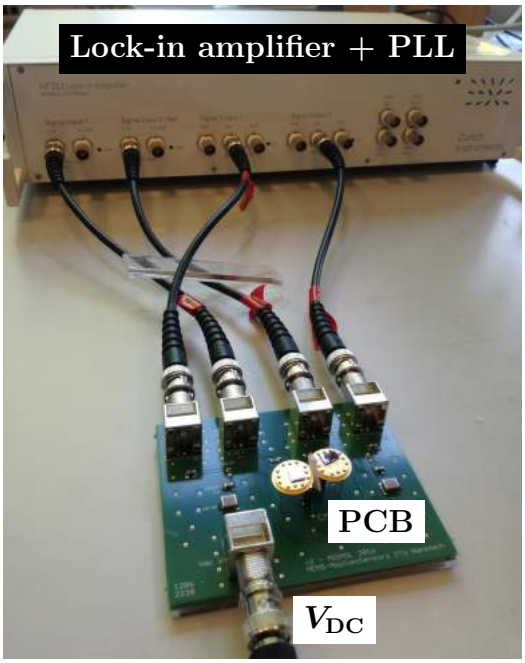


Figure 5.17: Photograph of the custom made PCB connected to a DC voltage supply (not shown) and the lock-in amplifier containing the PLL. The coaxial cables shown have a length of 0.25 m.

5.4.2 Allan Deviation

The Allan deviation gives a measure of the frequency noise in a system, as discussed in the theory chapter. An example of a plot of the overlapping Allan deviation as a function of the averaging time is given in Figure 5.18. Here a noise measurement is performed and subsequently the Overlapping Allan deviation is calculated using Equation 2.97. Figure 5.18 a) shows the measured Allan deviation for a G3 CMUT. The shape of this plot resembles that of the theoretical Allan deviation plot shown in Figure 2.21, with a distinct minimum and straight lines. The slope of the line at lower averaging times follows approximately the relation $\sigma \propto \tau^{-0.5}$, which is indicative of white FM noise (see Figure 2.21). Around the point of minimum flicker FM noise is expected and is indeed seen, which is shown by a line following $\sigma \propto \tau^{0}$. Finally, the straight line from the minimum to higher averaging times follows approximately $\sigma \propto \tau^{1}$, which is indicative of frequency drift. This can be due to a number of phenomena including aging or temperature changes of the CMUT or drift in the measurement equipment itself. Figure 5.18 b) shows a plot of the Overlapping Allan deviation using the same data as in a) but with fractional frequencies, that is, the relative difference of the

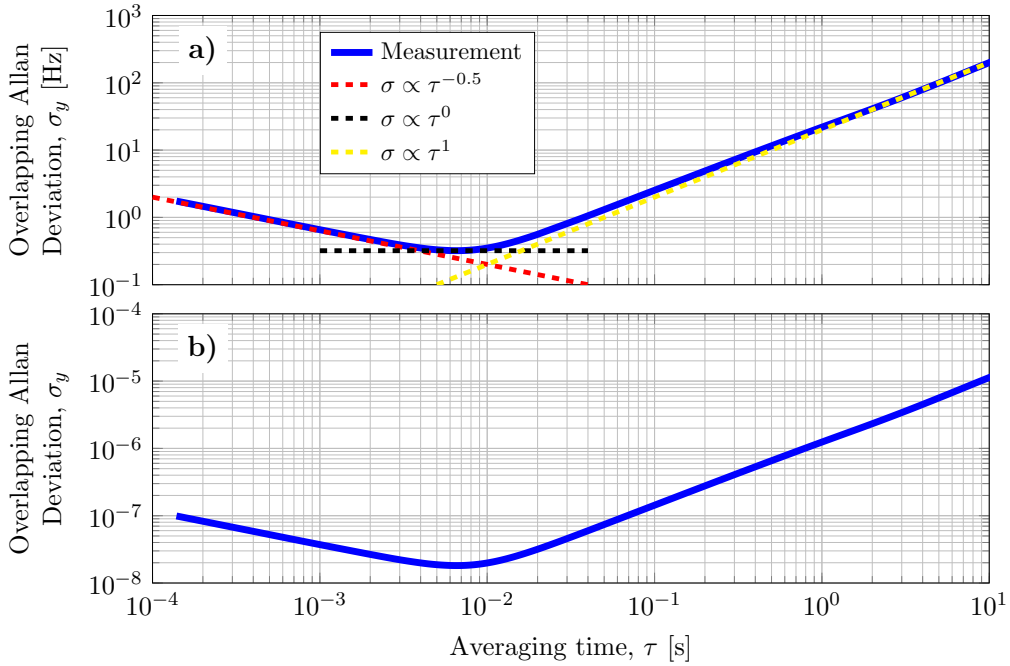


Figure 5.18: Overlapping Allan deviation measurement as a function of averaging time for a G3 device with a) non-fractional frequency input and b) fractional frequency input. The three lines are drawn to help approximate the slope values.

frequencies are used as input. The shape of the two plots is identical but the Allan deviation is now unitless.

In the previous section it was noted that the bandwidth of the PLL gives a trade-off between speed and noise. Since a lower bandwidth should result in less noise being measured, the lowest Allan deviations should be found for low bandwidths. Figure 5.19 shows a plot of the Overlapping Allan deviation for a G2 device for several bandwidths. Indeed, the lowest bandwidth gives the lowest Allan deviation, which ultimately results in the lowest mass limit of detection. Interestingly, this effect is seen to decrease for the highest bandwidths. Empirically it was found that the lowest bandwidth that is fast enough for experiments where the resonance frequency can change fast (relative to these noise measurements) is a BW = 10 kHz, which will be used for the rest of the experiments.

5.4.3 Mass limit of detection

The mass limit of detection (LOD) is an important figure of merit of any gravimetric sensor. Equation 2.96 showed how the LOD can be calculated from the device mass and Allan deviation and is here given in a slightly

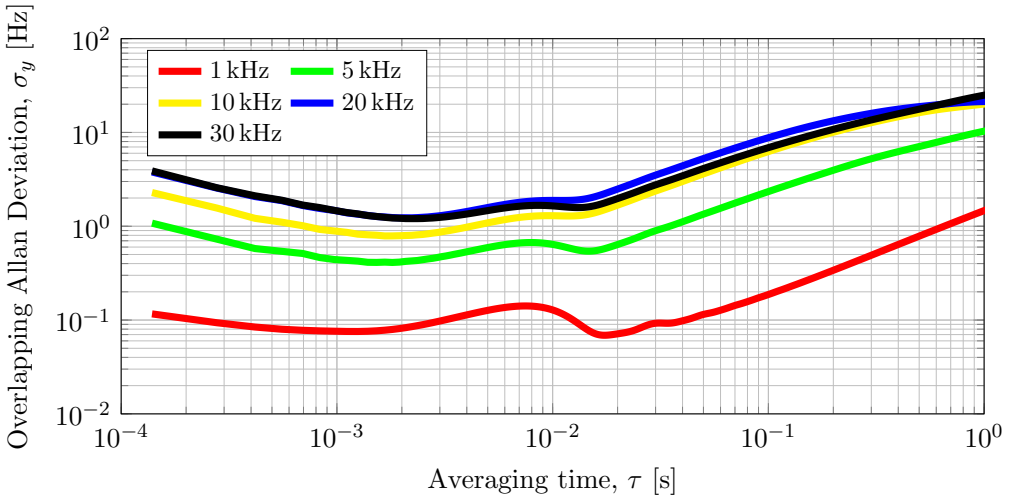


Figure 5.19: Measurements of the Overlapping Allan deviation as a function of averaging time for a G2 chip with different PLL bandwidths.

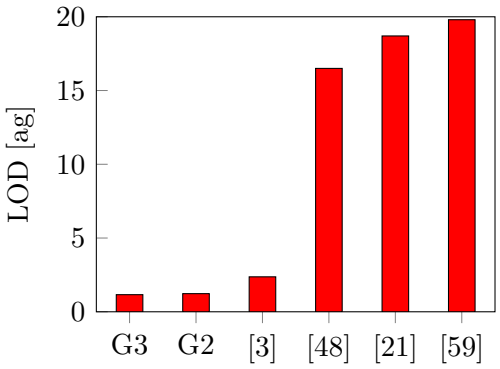


Figure 5.20: Experimentally determined LOD values for state of the art devices.

different version [21]:

$$\text{LOD} = \Delta m = 2m_{\text{plate}}\frac{\sigma}{f_r} = 2m_{\text{plate}}\frac{\Delta f}{f_r}. \tag{5.1}$$

Using Equation 5.1 the LOD can be calculated. Figure 5.20 shows a plot of the LOD for several CMUTs in the literature and G2/G3. All Allan deviation values used to calculate the LOD are 1σ (as opposed to e.g. 3σ) in order to make a fair comparison. The plate masses of G2 and G3 are calculated using the measured geometrical values, resulting in a higher LOD than if the designed geometries were used. Nevertheless, both G2 and G3 have a slightly lower LOD than the, to my knowledge, lowest published CMUT LOD found in [3], see Figure 5.20.

The minimum change in gas concentration was in Equation 2.109 shown

to be proportional to the LOD divided by the area of the sensor: $\partial C_{\min} \propto \frac{\mathrm{LOD}}{A}$. Thus, this figure of merit is just as or more important for gravimetric gas sensors than the mass LOD itself. The LOD/A can be written as:

$$\mathrm{LOD/A} = \Delta m/A = 2\rho t\frac{\sigma}{f_r} = 2\rho t\frac{\Delta f}{f_r}, \tag{5.2}$$

which favors a low frequency noise, high resonance frequency, low plate material density, and low plate thickness. This is expected to favor the CMUTs with a thin Si_3N_4 plate, despite the higher material density compared with Si and the need for a metal top electrode.

Figure 5.21 shows a comparison of the LOD/A for mainly CMUTs but also for the different gravimetric sensor types presented earlier in Table 2.7 taken from [11]. The CMUTs display the lowest LOD/A of all the sensor types, since the Allan deviation is among the lowest while the surface area is big relative to the other MEMS sensors and the NEMS sensors. At the same time the plate mass is much lower than the effective mass of the SAW and BAW sensors which have surface areas on the same order of magnitude as the CMUT. In Figure 5.20 the device in [21] has a lower LOD than the device in [59] but in Figure 5.21 the opposite is observed. This is due to the fact that the radius and hereby area of the CMUT plates in [59] are bigger ($a = 9\,\mu$m) than the radii of the devices in [21] ($a = 5.3\,\mu$m).

The lowest LOD/A is found for the G2 and G3 CMUTs as in Figure 5.20. Thus, these two devices also have, to my knowledge, the lowest published LOD/A of the CMUTs.

In this section the experimental setup tracking the resonance frequency of the CMUTs was shown. The setup consisted of a PLL connected to a custom-made PCB functioning as a bias-T. The Overlapping Allan deviation was calculated from measurements and used to determine the LOD and LOD/A. Both of these figures of merit were found to be the lowest published for CMUTs, to the knowledge of the author. The relative improvements (lowering) of the LOD and LOD/A relative to the best (lowest) published values are 51% and 11%, respectively.

5.5 Mass sensitivity

The mass sensitivity is one of the most important figures of merit for gravimetric sensors. In this section the mass sensitivity is studied using FEM and experimental measurements. In addition, the effect of the bias voltage on the sensitivity is investigated.

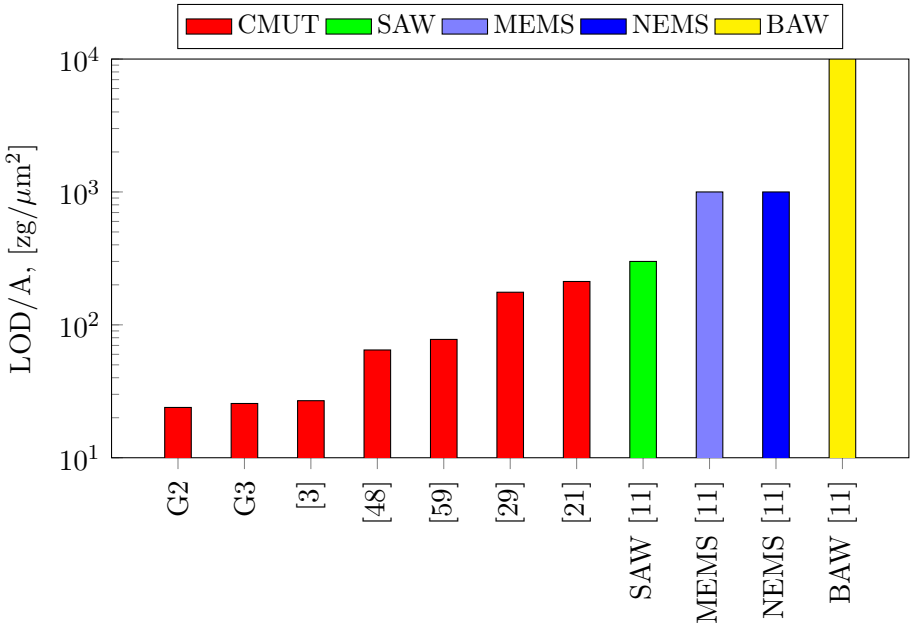


Figure 5.21: Experimentally determined LOD/A values for state of the art CMUT devices and four other sensor types from [11] which were also given in Table 2.7 at ambient pressure.

For masses added to the plate of the CMUT which are small compared with the mass of the plate, theory predicts a linear relationship between the frequency shift and amount of added mass [3]:

$$\partial f = \left(-\frac{1}{2}\frac{f_r}{m_{\text{plate}}}\right)\partial m = S\partial m. \tag{5.3}$$

In other words, the sensitivity can be found by measuring the resonance frequency shift for a known amount of mass added to the plate. This can be done as a calculation in a FEM model or experimentally for a device. In this section both methods are shown and the results are compared with the theoretical sensitivity calculated using Equation 5.3.

5.5.1 Sensitivity non-linearity

When the assumption behind Equation 5.3 is broken and the added mass is on the order of the mass of the plate itself, the relation between the frequency shift and added mass becomes non-linear. Specifically, the sensitivity decreases as more mass is added to the plate. This is seen in Figure 5.22 where in a) the resonance frequency decreases linearly as a function of added mass according to the simple theory following Equation 5.3. Whereas, the

resonance frequency decreases less and non-linearly for the two FEM results. The difference between the simple analytical model and FEM model increases as more mass is added to the plate. The FEM result named *boundary mass layer* is a boundary layer added to the top of the plate which can increase the plate mass without changing the flexural rigidity of the plate. The FEM result named *Au layer* is an Au layer added on top of the plate which both changes the mass and flexural rigidity of the plate. The reason for choosing Au is that this material will be used later for the experimental determination of the sensitivity. As expected, the resonance frequencies in Figure 5.22 a) all begin at the same resonance frequency when no mass is added ($m_{\text{added}}/m_{\text{plate}} = 0\%$) but then begins to deviate. The plot illustrates that the assumption of a linearly decreasing resonance frequency is only valid for $m_{\text{added}} \ll m_{\text{plate}}$, exemplified by the fact that the relative difference between the analytical and FEM result is 46% for $m_{\text{added}}/m_{\text{plate}} = 100\%$. Another interesting point is the relatively small difference between the two FEM models which is found to be a maximum of 3%, which also increases for more added mass, as the flexural rigidity of the plate becomes more dominated by the thicker Au layer. The geometry parameters used for the analytical and FEM calculation in Figure 5.22 are for the G4 device with the Al top electrode included.

Figure 5.22 b) shows the mass sensitivity as a function of normalized added mass. The sensitivity is found by taking the gradiant of the curves in figure a). The sensitivity is constant for the analytical expression while it decreases for the two FEM solutions. As expected, at $m_{\text{added}}/m_{\text{plate}} = 0\%$ the analytical and FEM boundary mass layer sensitivities are equal. The FEM sensitivities decrease for increasing added mass, and these types of curves can then be used for comparing the measured experimental sensitivity with the theoretical analytical one (which is only found if an infinitesimal mass is added to the plate).

As a rule of thumb the linear relationship between the resonance frequency shift and added mass of a QCM is valid up $m_{\text{added}}/m_{\text{QCM}} = 2\%$ [97]. The relative difference in Figure 5.22 b) between the analytical sensitivity and FEM mass loaded sensitivity of the CMUT at $m_{\text{added}}/m_{\text{plate}} = 2\%$ is 2.9%, which is a relatively small deviation.

5.5.2 Sensitivity and operating conditions

In this section it is shown how the bias voltage affects the mass sensitivity. As the bias voltage is increased the resonance frequency decreases due to the spring softening effect, as already discussed. Therefore, the sensitivity is expected to be affected by the bias voltage. Figure 5.23 shows a plot of

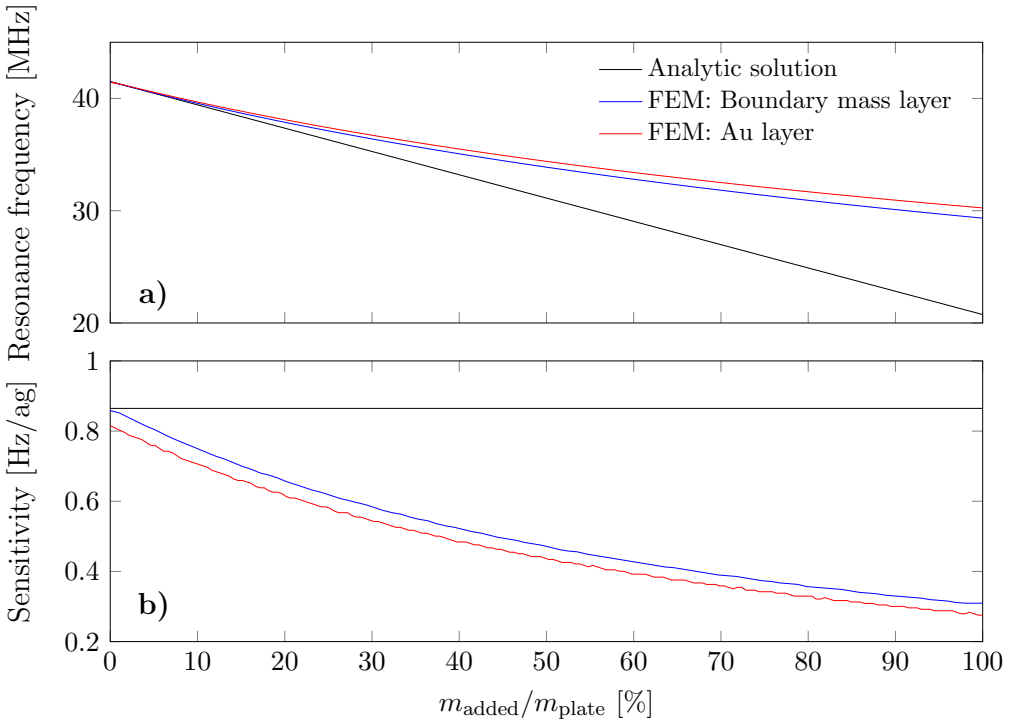


Figure 5.22: a) Resonance frequency and b) mass sensitivity as a function of mass added to the plate of a G4 device normalized with the mass of the plate itself. The analytical solution is shown together with the two FEM solutions: pure mass is added to the top of the plate for the boundary mass layer and an Au layer (with flexural rigidity) is added to the top of the plate for the Au layer solution.

a FEM calculation of the normalized sensitivity as a function of the normalized bias voltage. The curve is made by calculating the sensitivity at $m_{\text{added}}/m_{\text{plate}} \approx 0\%$ for each V_{DC} value. As the pull-in voltage is approached the sensitivity decreases as the resonance frequency decreases. This shows that the CMUT should be operated in the region where the curve in Figure 5.23 is approximately flat, which is the case up to about $V_{\text{DC}}/V_{\text{PI}} = 0.5$. The bias operating point will be a compromise between the sensitivity and the signal to noise ratio, which decreases as V_{DC} is decreased.

5.5.3 Experimentally determined sensitivity

The mass sensitivity can be experimentally determined by measuring the resonance shift after a known amount of mass is added to the plate. In the following, impedance spectra are used to determine the resonance frequency after thin (a couple of nm) Au layers have been sputtered on the CMUT

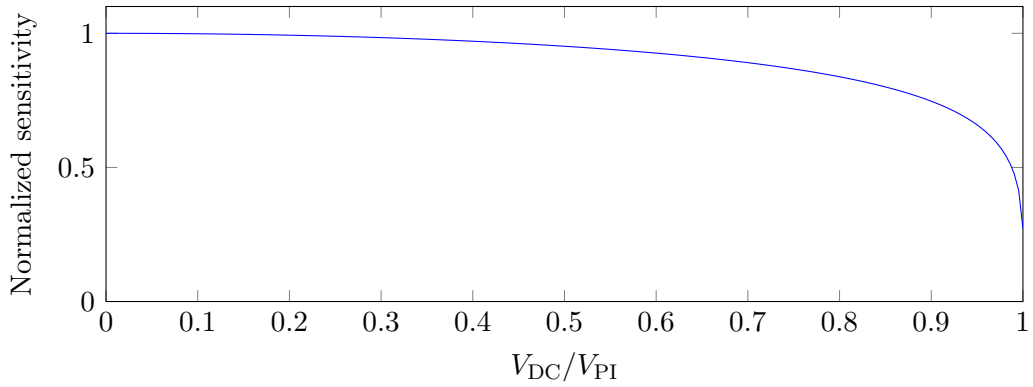


Figure 5.23: FEM simulation showing the normalized sensitivity as a function of normalized bias voltage for a G2 device. As the bias voltage increases the sensitivity decreases due to the decrease in resonance frequency caused by the spring softening effet.

plate. The added mass is determined from the area of the CMUT plate and thickness of the added Au layer, measured with AFM. Figure 5.24 shows an example of an AFM scan of the Au layer on a Si substrate, used to determine the Au layer thickness. For every Au deposition on the CMUT a reference wafer is made with the same sputtering time, which later has the Au layer thickness measured in an AFM. After each deposition the resonance frequency is found by measuring the impedance spectrum.

To avoid short circuits between the top and bottom electrode or between neighboring elements, a shadow mask is used that only allows the sputtered Au to cover the target element. Figure 5.25 shows a series of optical microscope images demonstrating the process of making the shadow mask. In a) the bare CMUT chip is shown and the bottom and top electrode of the target element is highlighted. The shadow mask is made up of blue tape (normally used when dicing wafers), covering everything but the top electrode. The most critical alignment is shown in b) which is the placement of the tape between the top and bottom electrode. The tape placement is done by hand under a microscope, and takes about 5 min for a chip. Figure 5.25 c)-e) shows the other tape pieces being placed. The chip in e) is now ready to have Au sputtered on the surface. After the sputtering process the tape can simply be peeled off and the impedance can be measured.

Figure 5.26 shows a plot of the measured resonance frequency as a function of the added Au mass, which was calculated from the measured Au layer thickness. As theory predicts the data points form nearly straight lines. The data are for devices G2, G3, and G4 and include linear fits. The slope of this fit gives the mass sensitivity, which can be be converted to the familiar

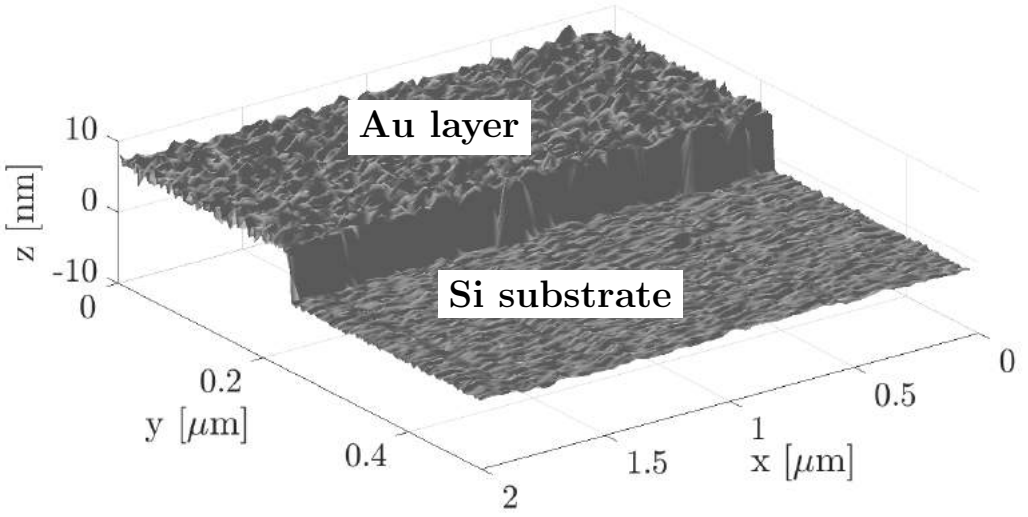


Figure 5.24: AFM scan of the edge of an Au layer used to determine the deposited Au layer thickness and ultimately the added mass from the Au layer on the CMUT plate.

units of Hz/ag. This experimental sensitivity S_{exp} is given i Table 5.3 along with the theoretical analytically calculated sensitivity (Equation 5.3) S_{theo} for G2-G4. The discrepancy between the theoretical sensitivity and experimental sensitivity, caused by the addition of the big masses, is discussed in the following.

To more easily compare the sensitivities of the generations, Figure 5.27 shows a plot of the normalized resonance frequency as a function of the normalized added mass. Note that the masses of the plates m_{plate} are different for the three generations. Looking at this plot it is clear to see the difference in the slopes and thus the sensitivities. The plot also reveals that the plates have been loaded with a different relative amount of mass. In Figure 5.22 b) it was shown how the sensitivity decreased as more mass was added to the plate. This point is further demonstrated in Figure 5.28 where two separate measurements were performed on two G4 chips loaded with a different amount of mass. It is clear that the sensitivity of the chip loaded with the least mass ($m_{\text{added}}/m_{\text{plate}} \approx 0.4$) is higher than for the chip with the most added mass ($m_{\text{added}}/m_{\text{plate}} \approx 1.3$). The relative difference of the sensitivities between the two measurements is 43%, showing the same tendency as was found in the FEM model: more mass decreases the sensitivity. Additionally, a decrease of the slope and hereby sensitivity, can be observed within each measurement as the center measurement point is always below the fitted line. This can be seen in Figure 5.27 and 5.28 and is observed for all measurements

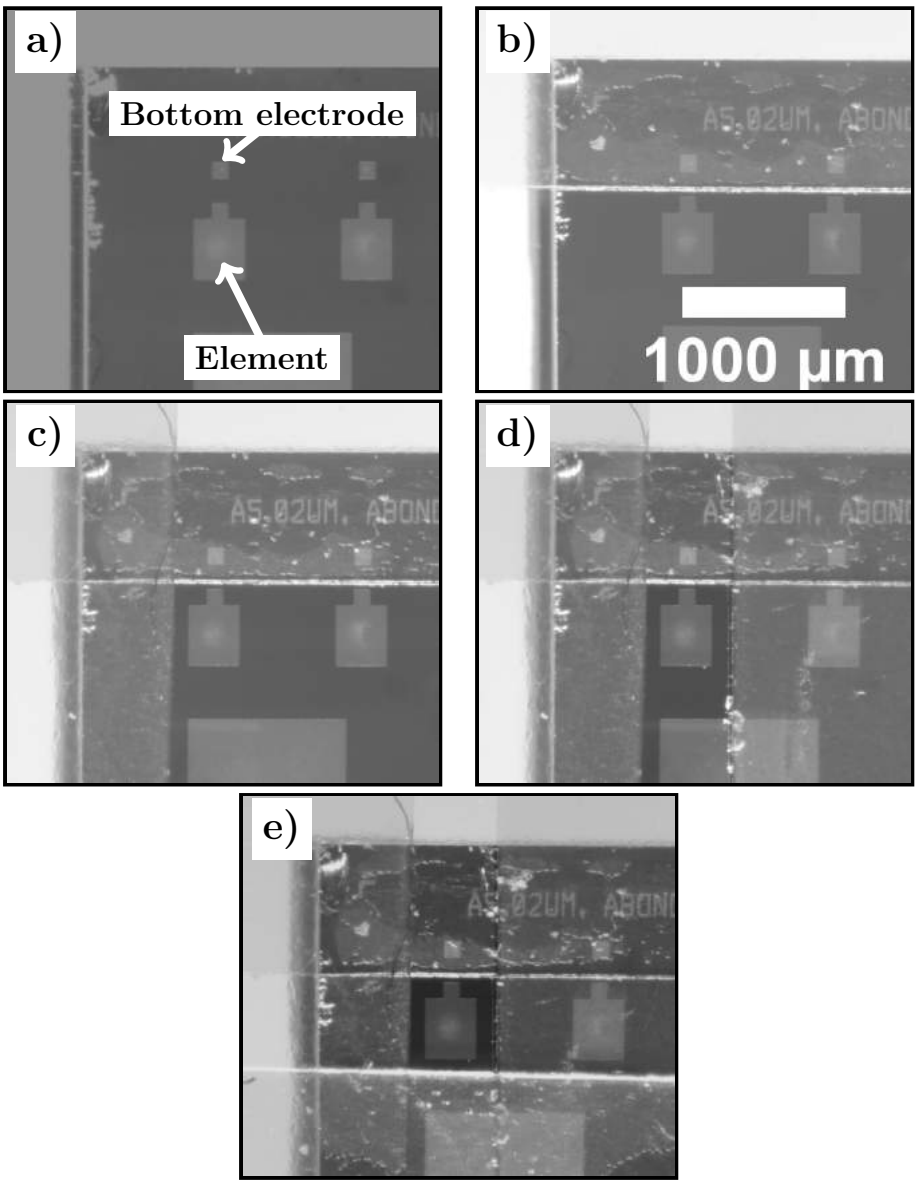


Figure 5.25: Optical microscope images showing the process of applying a tape shadow mask to a chip. The shadow mask is used to control where the Au layer is deposited. In a) no tape has been applied yet, while in e) four pieces of tape give a small window in the shadow mask over the target element.

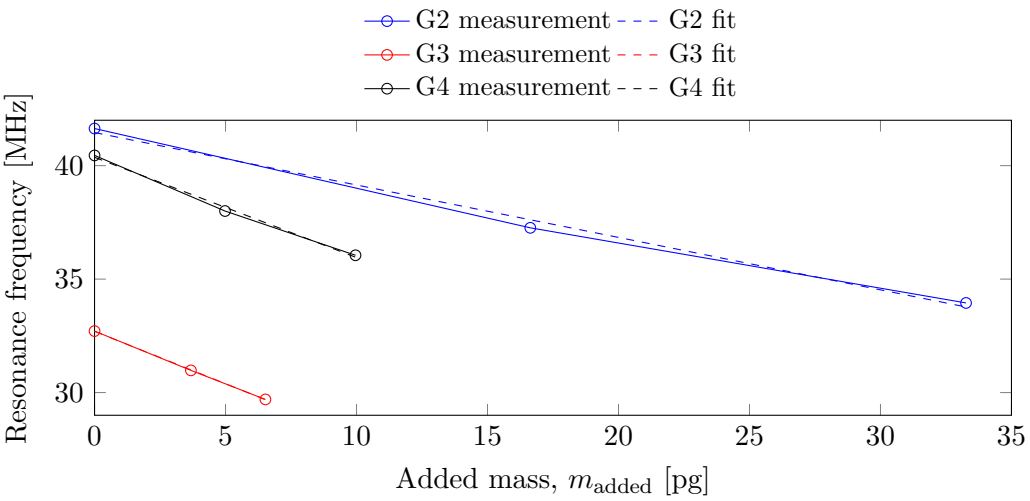


Figure 5.26: Measured resonance frequencies as a function of added mass for G2-G4 devices along with linear fits.

but most dominantly for the experiments with a large relative added mass.

The measured experimental sensitivities should be compared with predicted sensitivities of CMUTs loaded with the same amount of mass. This can be achieved by loading the CMUT with mass in the FEM model and finding the sensitivity. Table 5.3 shows the sensitivities of the CMUT generations, calculated in the FEM model when no mass is added (unloaded case) called $S_{\text{FEM, unloaded}}$ and when the same amount of mass is loaded as in the experiment (loaded case) called $S_{\text{FEM, loaded}}$. These data are also shown as a bar chart in Figure 5.29 for the three generations. The theoretical and FEM calculations are performed using the measured geometrical parameters and measured values for the resonance frequencies. Generation G3 behaves differently than G2 and G4 since the relative difference between S_{exp} and $S_{\text{FEM, loaded}}$ and the relative difference between S_{theo} and $S_{\text{FEM, unloaded}}$ is larger. This is due to a lower measured resonance frequency compared with the resonance frequency found by the FEM model. In contrast, the FEM model does a good job of predicting both the experimental and theoretical sensitivities of G2 and G3, for which the measured and simulated resonance frequencies match when the measured geometrical values are used as input. The general trend of the sensitivities for all three generations, shown in Figure 5.29, is that the measured/loaded sensitivities are lower than the theoretical/unloaded which is what is expected from the discussed above.

Lastly it should be noted that the use of Au as a layer for adding mass to the CMUTs for finding the experimental mass sensitivity is not ideal. First, since Au is an electrical conductor the top and bottom electrode can short

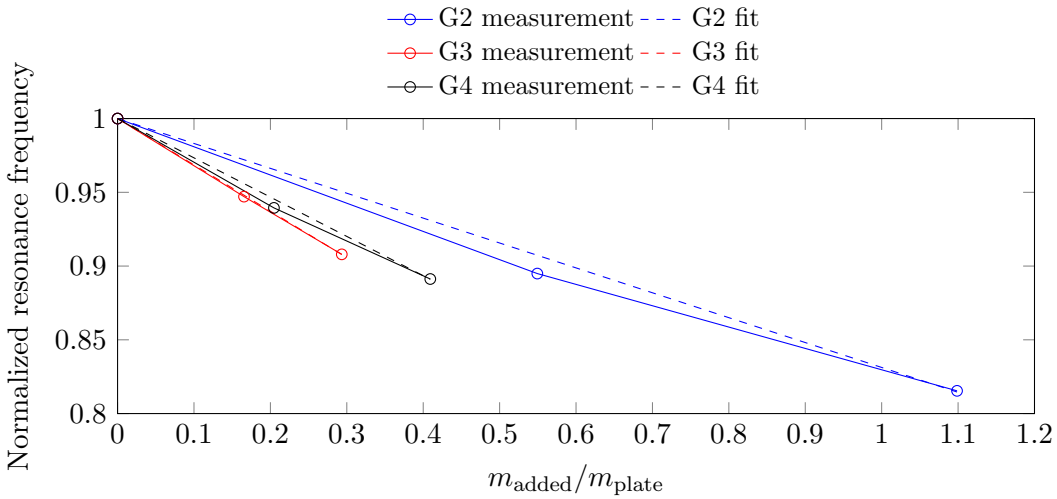


Figure 5.27: Normalized resonance frequencies measured as a function of normalized added mass. The plot demonstrates the difference in slope values and the relative amount of added mass.

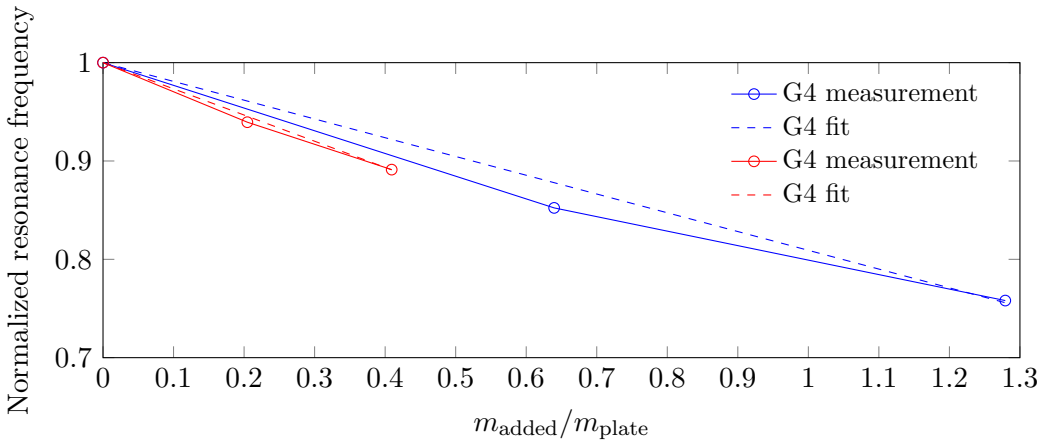


Figure 5.28: Normalized resonance frequencies as a function of normalized added mass measured for two G4 chips with a different amount of relative added mass. The sensitivity is seen to be highest for the device with the least amount of added mass, demonstrating the effect of the added mass on the sensitivity.

Table 5.3: Measured and simulated mass sensitivities for the G2-G4 devices. S_{exp} is the experimentally determined sensitivity, $S_{\text{FEM, loaded}}$ is the sensitivity found by FEM when the CMUT plate is loaded with the amount of mass listed in the row: $m_{\text{added}}/m_{\text{plate}}$. S_{theo} is the theoretically/analytically determined sensitivity (using the measured geometrical values). Finally, $S_{\text{FEM, unloaded}}$ is the sensitivity found in FEM when no mass is added to the plate (unloaded). The sensitivities above the central horizontal line should be compared and likewise for the sensitivities below this line. All values have the unit Hz/ag.

Method	**G2**	**G3**	**G4**
S_{exp}	0.24	0.46	0.44
$S_{\text{FEM, loaded}}$	0.18	0.74	0.48
$m_{\text{added}}/m_{\text{plate}}$	72%	21%	42%
S_{theo}	0.45	0.52	0.86
$S_{\text{FEM, unloaded}}$	0.44	1.02	0.82

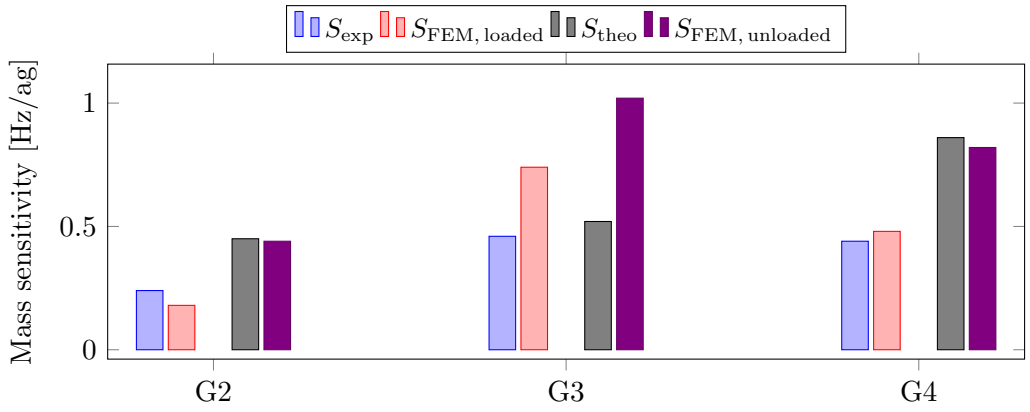


Figure 5.29: Measured and simulated mass sensitivities for the G2-G4 devices. The data are from Table 5.3. S_{exp} is the experimentally determined sensitivity, $S_{\text{FEM, loaded}}$ is the sensitivity found by FEM when the CMUT plate is loaded with the amount of mass listed in the row: $m_{\text{added}}/m_{\text{plate}}$ in Table 5.3. S_{theo} is the theoretically/analytically determined sensitivity (using the measured geometrical values). Finally, $S_{\text{FEM, unloaded}}$ is the sensitivity found in FEM when no mass is added to the plate (unloaded).

circuit if not protected during deposition. It would be easier to use a non-conducting material. Second, the high density of Au ($19.3\,\mathrm{g/cm^3}$) results in thin $\leq 10\,\mathrm{nm}$ layer thicknesses if $m_{\mathrm{added}}/m_{\mathrm{plate}}$ values are to be kept as low as possible in order to not decrease the sensitivity. The deposition of these thin Au layers is more difficult to control and a material with a lower density would allow for better control of the relative layer thickness and hence give a lower error on the final deposited mass. Despite all these drawbacks of using Au, it was the only material readily available at the time of experimentation.

5.5.4 Sensitivity comparison

In this section the sensitivities of the generations are compared with the sensitivities of other CMUTs and gravimetric sensors. The sensitivities, shown in this section, are the theoretical sensitivities calculated using the measured geometrical parameter values. The theoretical sensitivity values are chosen to make a fair comparison with the sensitivities reported in the literature, which typically are theoretical values.

CMUTs

Figure 5.30 shows a bar chart of the sensitivities for the CMUTs from this project and from the literature. The sensitivity in [3] and [21] are, to my knowledge, the highest found in the literature. These CMUTs are from the Stanford group and have a higher resonance frequency but also a higher mass than G2-G4 resulting in a lower mass sensitivity as shown in Figure 5.30. In addition, the highest sensitivity is found for G4 which is due to the low mass of the plate and relative high resonance frequency. The sensitivity of G4 is 257% higher than the state of the art: [3, 21].

Gravimetric sensors

Figure 5.31 shows a bar chart of the *normalized* sensitivities (Equation 2.94) of different types of gravimetric sensors. Note that the normalized sensitivity axis is logarithmic. A version of this plot with fewer entries was shown in the introduction chapter. In the version shown here the measured geometrical parameters are used, which is reflected in that fact that G2 and G3 have different normalized sensitivity values, due to the thicker measured plate of the G3 device compared with the G2. Remember that the normalized sensitivity is only dependent on the material density and plate thickness: $S_{\mathrm{norm}} \propto (\rho t)^{-1}$.

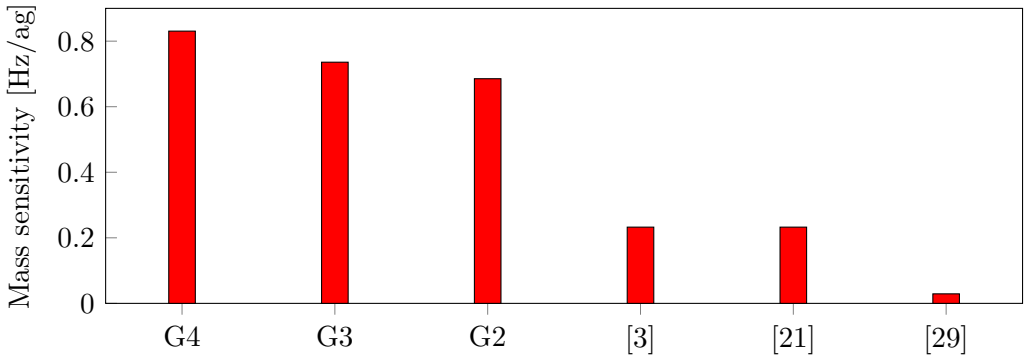


Figure 5.30: Comparison of the theoretical analytically calculated mass sensitivities for different state of the art CMUTs. The geometrical parameters used to calculate the sensitivities for G2-G4 are the measured values resulting in slightly lower sensitivities than the ones given in the design chapter.

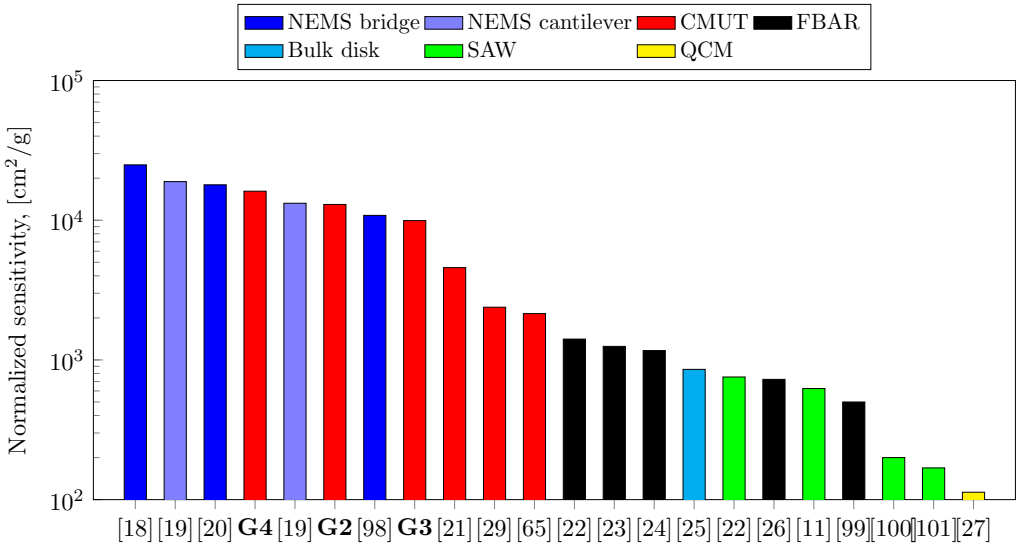


Figure 5.31: Normalized sensitivities for different sensor types. The CMUTs generally have lower sensitivities than the NEMS devices but higher sensitivities than all the other device types. G2-G4 show the highest normalized sensitivities of the CMUTs due to the thin plates.

The highest normalized sensitivities are found for the NEMS devices that with their tiny dimensions are expected to dominate such comparisons. Next, the CMUTs are found and as the theory predicts: the thinner the plate the higher the normalized sensitivity. This is exemplified by the CMUT with the highest S_{norm}, G4, which has the thinnest plate with a thickness of only 50 nm with an Al top electrode of 50 nm. This is possible due to the use of the Si_3N_4 LPCVD plates, which is the common trait of G2-G4 that all have higher normalized sensitivities than the other CMUTs in Figure 5.31. In fact, the normalized sensitivity of G4 is 253% higher than for the CMUT in [21], which has the highest published S_{norm} value for CMUTs, to my knowledge. Next, the FBARs are coming in with a lower S_{norm} than the CMUT, due to restrictions on the minimum thickness of the support/electrode/piezoelectric material stack. The second to last lowest normalized sensitivity is that of the SAW gravimetric sensors, as these require increasingly higher resonance frequencies to increase S_{norm} as this is inversely dependent on the wavelength of the acoustic wave, as shown in Table 2.6. Finally, the QCM/BAW gravimetric sensors are found last since the normalized sensitivity is inversely dependent on the wavelength of the acoustic wave and the wavelength is dependent on the thickness of the piezoelectric material of the QCM which is limited.

5.6 Chapter conclusion

In this chapter the fabricated CMUTs were characterized using different measurement techniques. The electromechanical coupling coefficient was seen to be highest for the CMUT designs with the lowest relative parasitic capacitance. The lowest mass limit of detection, for chip G3, was found to be 51% lower than the lowest published LOD [3], to the knowledge of the author. The values of the limit of detection per area were also found to be lower than the current state of the art for CMUTs, with a relative improvement of 11%. The mass sensitivity was determined by the analytical expression, the FEM model, and experimentally. In fact, it was found that for the unloaded CMUTs the analytical and FEM sensitivity were in good agreement while the same is true for the measurements and FEM calculations of the the mass loaded CMUTs. The highest theoretical mass sensitivity, G4, was 257% higher than the current highest sensitivity of a published CMUT [3]. Finally, the CMUTs have a lower S_{norm} than the NEMS devices, but it is higher compared with the other gravimetric sensor types. The highest CMUT normalized sensitivity is found for G4 which is 253% higher than the current state of the art for CMUTs [21]. In conclusion, the fabricated CMUT chips

show both a low mass limit of detection and high mass sensitivity compared with the current state of the art for CMUTs.

CHAPTER 6

Gas sensing experiments - BMK

Any gas can in principle be detected using gravimetric gas sensors. The challenge is finding a method of making the sensor selective towards the gas molecules of interest. Many different fields were listed in the introduction chapter, in which gas sensors are used, e.g.: indoor air quality, safety in industry, automotive, defence, law enforcement, environmental studies, food quality and safety. We have chosen to focus on the field of law enforcement, more specifically drug detection. The sensor should detect the precursor molecule Benzyl Methyl Ketone (BMK) (also called phenylacetone or P2P) which is used in the synthesis of the recreational drugs amphetamine and methamphetamine [102]. These drugs are controlled substances in many countries and so are the common precursors such as BMK. Therefore, the detection of the precursors are of interest for law enforcement agencies, both as stationary measurement stations e.g. at border crossings and as mobile sensors to be used in the field. The low LOD/A of the CMUTs means that only a small concentration of gas needs to be present in order for it to be detected.

Two common synthesis routes that begin with BMK as a precursor molecule are: the Leuckart reaction and reductive amination of BMK [102]. The latter reaction is shown in Figure 6.1. This reaction has relatively few process steps and this is one of the reasons why BMK is a popular precursor for (meth)amphetamine production in clandestine laboratories. In addition, BMK is easily stored as it is a liquid at ambient conditions, although not odorless.

This chapter is based on the article [53] and Appendix B. The chapter

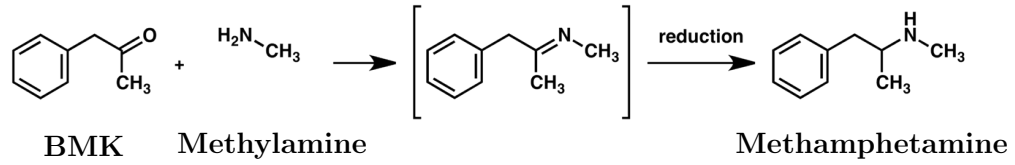


Figure 6.1: Example of the synthesis of methamphetamine: reductive amination of BMK. This is one of the possible synthesis routes.

describes experiments regarding the detection of BMK in the gas phase. First the challenge of selectivity is addressed, then an introduction to colorimetry is given. Next, the experimental setup is presented and finally the results are discussed.

6.1 Selectivity

The bare CMUT chip is not selective towards any gas molecules due to the inert surface. The most typical way described in the literature (see Section 1.4.3) of making the CMUTs selective has been to coat the plates with a thin functionalization layer showing an affinity, and hereby selectivity, towards the target gas molecules. This approach is also used for the experiments shown here, but in conjunction with a colorimetric detection method. The colorimetric sensor works by detecting a change of color of a dye when the target gas is present. *Hence, the sensor system for the following experiments consists of a gravimetric sensor (CMUT chip) coated with a functionalization layer and a colorimetric chip made up of colorimetric dyes.* Combining the two sensor types provides the overall sensor system with redundancy, in that one sub-sensor can fail without rendering the overall system completely useless. However, adding the colorimetric sensor increases the size and cost of the measurement system considerably, since an optical system is needed for the detection of the color change.

An important note about the two measurement techniques is that the CMUT gives a quantitative output if the mass sensitivity is known or the sensor has been calibrated. In contrast, the colorimetric sensor provides an qualitative output, detecting whether the analyte is present or not. Thus, the CMUT has the capability of detecting the gas concentration while the colorimetric sensor cannot.

6.2 Colorimetry

A colorimetric sensor typically detects a change of color of colorimetric dyes. The detection can be done by microscope or high quality consumer cameras, since the light typically is in the visible wavelength range. By making an array of several different dyes, each dye will react differently to an analyte and thus leave a "fingerprint" to be detected. This is the working principle behind the so-called colorimetric electronic noses [103].

In this project we have collaborated with the research group of Mogens Havsteen Jakobsen at DTU Nanotech who have developed and provided the colorimetric dyes used in the experiments. They have developed an all colorimetric array sensor capable of detecting several gas analytes e.g. [104]: BMK, H_2O_2, acetone, ethanol, and gasoline. To give an example of the working principle of a colorimetric sensor and what such a chip looks like, Figure 6.2 a) shows an optical image of a chip before being exposed to any analyte. Many different dyes have been spotted on the white polypropylene substrate. The dyes are also repeated throughout the array. Optical images are recorded before, while, and after the chip is exposed to the target analyte, in this case BMK. Figure 6.2 b) shows an optical image of the chip after BMK exposure. It is difficult to see the difference between the two images, hence difference maps are often used to highlight the difference between images before and after analyte exposure. Difference maps are simply made by calculating and displaying the difference of each pixel value for two images. Figure 6.2 c) shows the difference map for a) and b). Clearly, for some of the dyes a large change in pixel value has happened due to a color change, while other dyes are almost black and no color change has occurred. The intensity of the dye spots in the difference map is used as a measure for when a reaction has occurred, as will be shown.

For the experiments in this section two dyes were selected: one that showed a color change (coating A, Figure 6.2) and one that did not (coating B, Figure 6.2). Coating A is a dye called Reichardt's dye and coating B is a dye called Bengal rose B. For the colorimetric experiments to come, coating A is expected to change color in the presence of BMK and coating B is not. Furthermore, these two dyes are used to coat the CMUT surfaces. Thus, the dyes are also used as functionalization layers. The hypothesis is that the frequency shift, in the presence of BMK vapor, is greater for the CMUT coated with coating A than the CMUT coated with coating B. This hypothesis is tested at the end of this section.

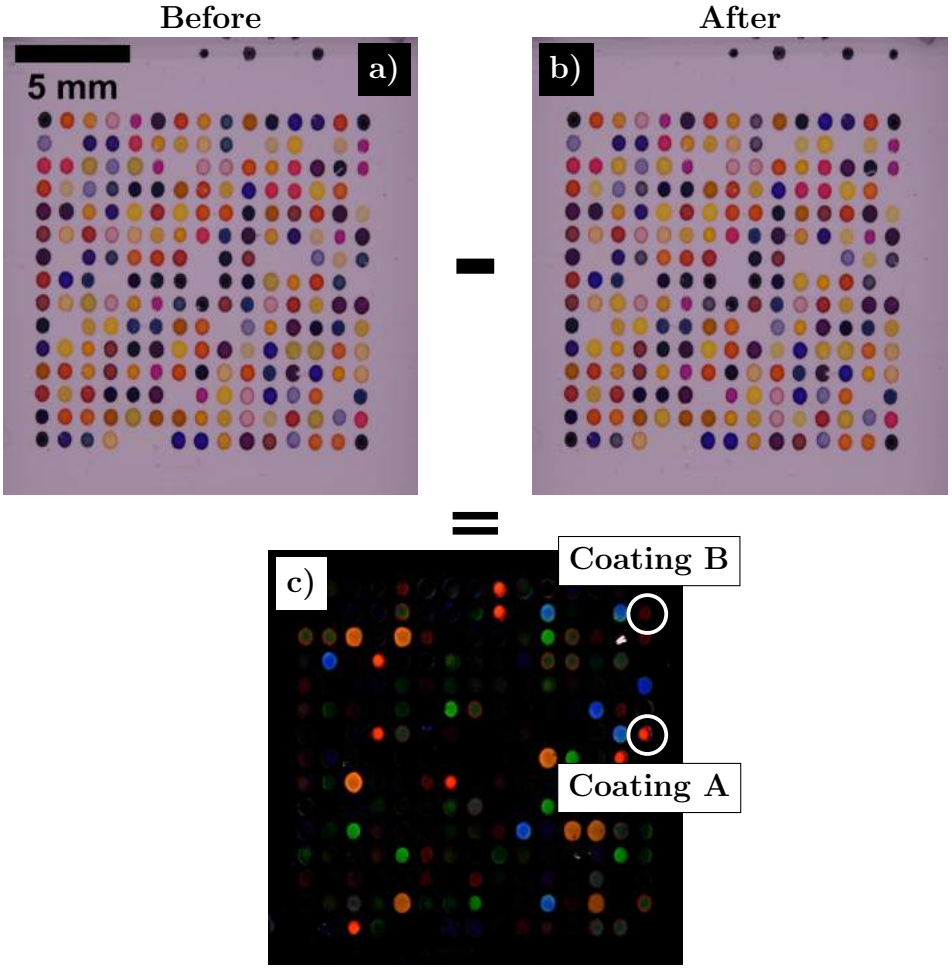


Figure 6.2: a) Optical image of a colorimetric chip with many different dye spots before exposure to BMK. b) The same colorimetric chip after exposure to BMK. c) Difference map of the two previous images showing color changes. The difference map can be understood as subtracting two images from each other.

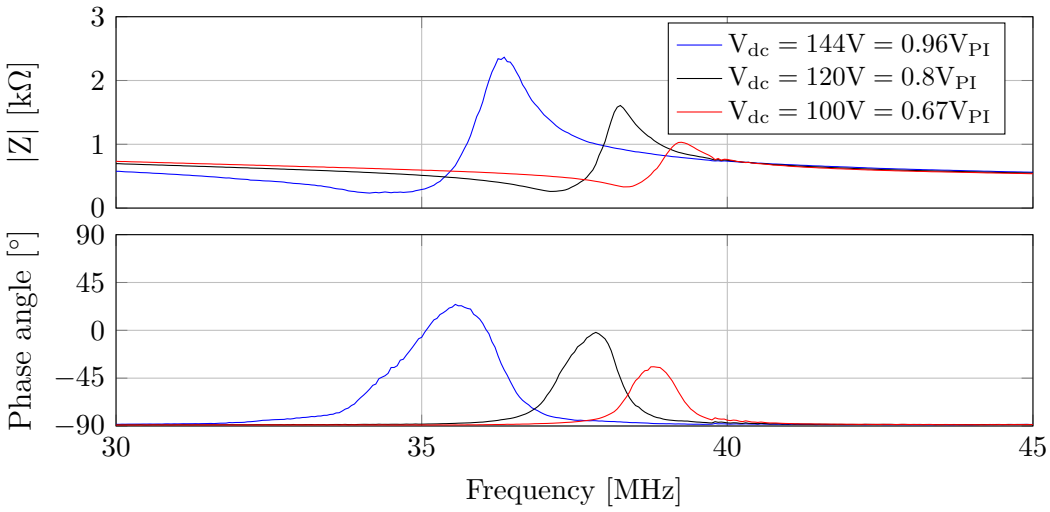


Figure 6.3: Impedance spectra of the G2 CMUT at three bias voltages.

6.3 Experimental setup and method

In this section fabrication/preparation of the colorimetric and CMUT chips is described. Lastly, the experimental setup is shown, which allows for simultaneous colorimetric and gravimetric measurements.

6.3.1 CMUT chips

The CMUT chips used for this experiment were the from the G2 generation. An impedance spectrum is shown in Figure 6.3 for such a chip. The chips are functionalized with either coating A, coating B, or with no coating. Figure 6.4 a) shows a microscope image of the CMUT chip wirebonded to two pins on the TO-8 carrier substrate. Furthermore, the coating on the CMUT element can be seen to the left in a). The coating is dissolved in a liquid solution and applied using a dispensing machine (GESIM, Nano-Plotter) with a piezoelectric actuated pipette. The alignment between the pipette and CMUT element is currently done manually, but this could be automated if many elements were to be functionalized.

6.3.2 Colorimetric chips

The colorimetric chips used in this experiment only have two different dyes (coating A and coating B), as compared with the colorimetric chips shown in Figure 6.2. Figure 6.4 b) shows a zoom-in on the colorimetric chip. Two columns of dye spots (left: coating B and right: coating A) are visible. The

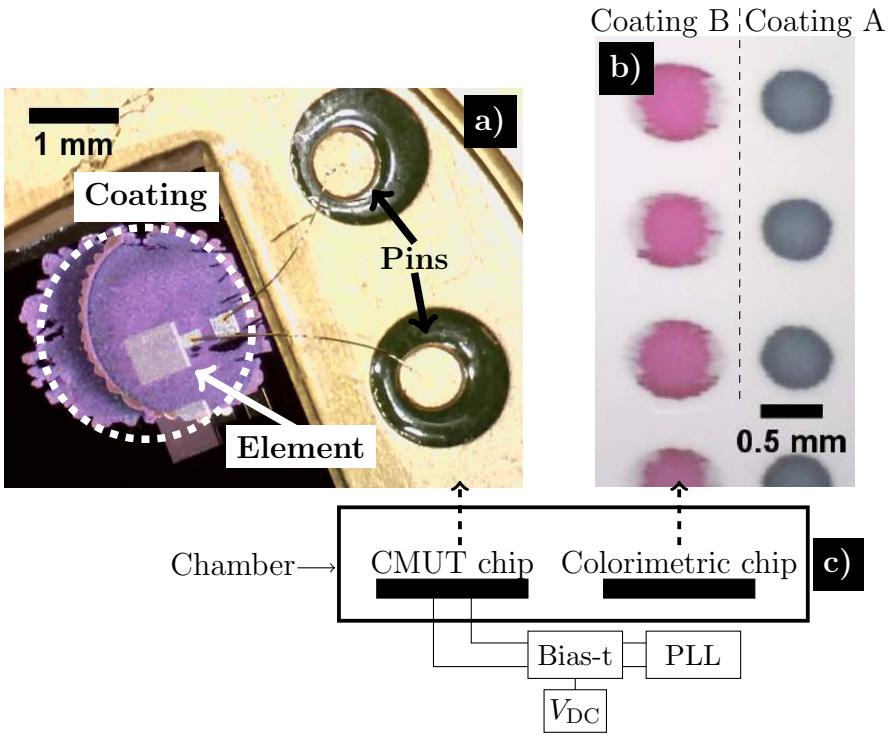


Figure 6.4: a) Optical microscope image of the CMUT chip wirebonded to two pins. A single element is seen, coated with a dye. b) Optical microscope image of the colorimetric chip showing two columns of dye. c) Schematic of the experimental setup showing the chamber enclosing the CMUT and colorimetric chip. The electrical actuation/readout circuit of the CMUT is also seen which consists of a phase locked loop, a bias-t, and a DC power supply.

dye spots are made by dispensing several droplets of dye in a liquid solution on a white polypropylene substrate. The dispensing and spot pattern are controlled by a computer and an automatic dispensing machine (GESIM, Nano-Plotter) using a piezoelectric actuated pipette. Since these chips consist of many dye spots the machine is operated using the automatic mode. The image in Figure 6.4 b) is an example of the images used to determine the color change and is taken with a microscope.

6.3.3 Experimental setup

A sketch of the experimental setup is shown in Figure 6.4 c). The purpose of the setup is to simultaneously measure the color change of the colorimetric chips and frequency shift of the CMUTs. The CMUTs and the colorimetric chip are placed in a small ($72\,\mathrm{cm}^3$) chamber. Through an inlet hole in the

chamber, analytes can be injected. The analytes are in the liquid phase when injected into the chamber where they evaporate and thus increase the gas phase concentration of the analyte until the vapour pressure is reached. A microscope records an image of the colorimetric chip every 30 s through the transparent chamber top, while the PLL keeps the CMUT at resonance and records this resonance frequency. The electrical circuit is identical to the one already shown previously in Figure 5.16 which was used to measure frequency noise.

6.4 Results and discussion

6.4.1 Colorimetry

The basic method of analyzing the images of the colorimetric chips have already been introduced with the idea of finding the difference map. Figure 6.5 shows two sets of images of a colorimetric chip a) before ($t = 0\,\text{min}$) and b) after ($t = 16\,\text{min}$) exposure to BMK. Comparing these two images it looks like coating A has changed color while it seems like the color of coating B has not changed. This is confirmed when looking at the difference map between the two in d), where the two dye spots of coating A clearly light up and have changed color. Figure 6.5 c) is the difference map at $t = 0\,\text{min}$, which is all black since no color changes could have happened yet.

In order to quantify these results the red-green-blue (RGB) values are found inside the spots at the difference maps as a function of time, using the first image as the reference image. The intensity is calculated by taking the mean of the RGB values inside the spots. Examples of these areas inside the spots on a difference map is marked by white dashed circles in Figure 6.5 d). The background signal is subtracted by subtracting the intensity of an area outside of the spots. This helps get rid of unwanted systematic changes of the signal. Figure 6.6 a) shows the intensity as a function of time for coating A and B when BMK is the analyte. In addition, the molecular structure of BMK is shown. A couple of minutes after the liquid BMK has been injected into the chamber the intensity of coating A increases until it reaches a constant plateau, while coating B does not change significantly. This is what was expected from the screening experiment in Figure 6.2 where the same behavior is observed.

To test the selectivity of the colorimetric dyes, this experiment is repeated with two other analyte molecules with new colorimetric chips. The first analyte is acetone, which is chosen due to the molecular structure which is similar to the structure of BMK with the exception of the aromatic ring.

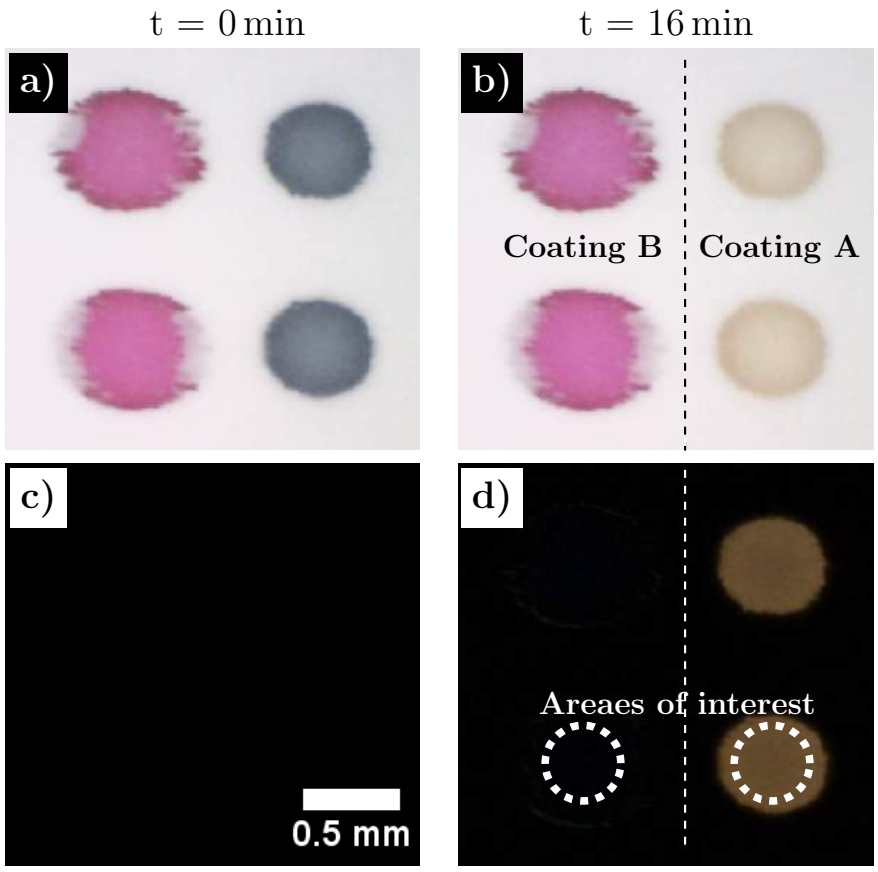


Figure 6.5: Optical microscope images of coating A and B, a) before and b) after reaction with BMK. Difference maps c) before and d) after reaction with BMK. BMK is injected at $t = 4\,\text{min}$.

Figure 6.6 b) shows a plot of the intensity as a function of time with acetone as the analyte. No color change was observed and thus no change in intensity is seen. Consequently, the aromatic ring may play a role in the color change seen for coating A in a). Finally, water is the analyte in Figure 6.6 c), chosen since water is always present in the ambient atmosphere. No color changes were observed for the water experiment.

Two conclusions can be drawn from the experiments in Figure 6.6: Firstly, coating A changes color when in contact with BMK while coating B does not and secondly, coating A shows selectivity between BMK and acetone/water. Whether this selectivity is also seen for the response of the CMUTs is investigated next.

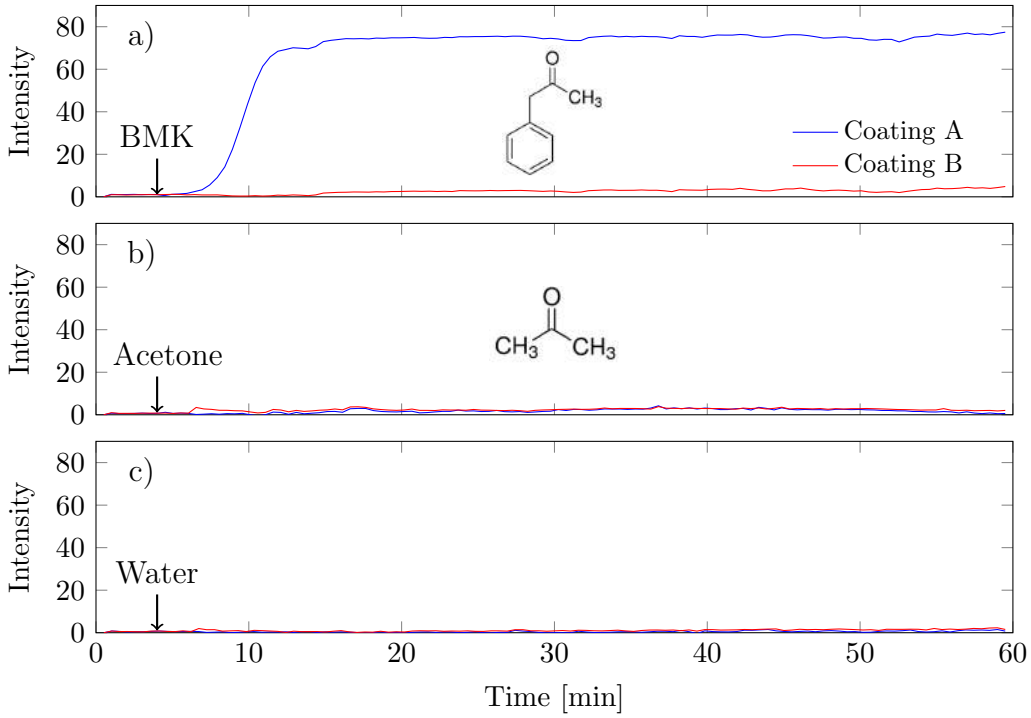


Figure 6.6: Mean intensity of the red, green and blue (RGB) values inside the dye spot area as a function of time. The analytes are: a) BMK, b) acetone and c) water. The analyte is introduced at $t = 4\,\text{min}$ indicated by an arrow.

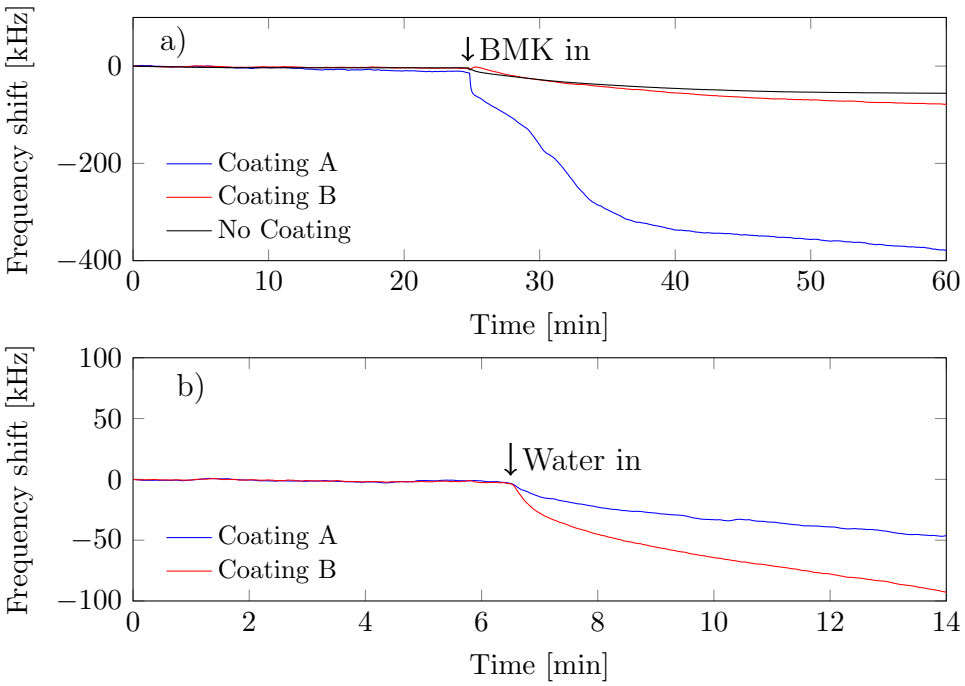


Figure 6.7: Frequency shifts as a function of time for a) BMK and b) water as analytes for CMUTs coated with coating A and coating B.

6.4.2 CMUT

Two CMUTs are functionalized with coating A and coating B, respectively. In addition, a CMUT is left without a coating. Subsequently, BMK is injected in the chamber and the resonance frequency is detected. Figure 6.7 a) shows a plot of the resonance shift as a function of time for these three CMUTs. The BMK is injected at the time indicated by the arrow and afterwards the resonance frequency is seen to decrease for all three CMUTs, as expected. The resonance frequency shift is largest for the CMUT with coating A, while the frequency shift for the CMUT with coating B and with no coating is almost the same. This, along with the colorimetric results, indicates that coating A has an affinity for BMK.

In order to investigate the selectivity of the functionalization layers water is again used as an analyte. Figure 6.7 b) shows a plot of the resonance frequency shift as a function time for a CMUT functionalized with coating A and B, respectively. After water is injected the largest frequency shift is found for the CMUT with coating B, which is the opposite of what was seen when BMK was injected. Thus, a selectivity is observed between BMK and water for CMUTs functionalized with coating A.

6.4.3 Discussion

The resonance frequency shifts in Figure 6.7 are all large compared to what is typically found in the literature for similar experiments. This could be due to a combination of a thick functionalization layer, a high gas concentration, and a high sensitivity. The functionalization layers were spotted on the CMUT device by manually aligning the piezoelectric actuated pipette and CMUT element. Subsequently, a fixed number of droplets (in which the dye was dissolved) were dispensed. This process was repeated until the entire CMUT element was covered with the dye/functionalization layer. If the alignment was off, the process had to be repeated a couple of times which would increase the layer thickness for some parts of the CMUT element. This thick functionalization layer will both decrease the mass sensitivity of the CMUT and result in large resonance frequency shifts when exposed to an analyte. This was discussed in the theory section (Equation 2.108) and is due to the increased volume of the functionalization layer. Hence, a more consistent functionalization method should be developed.

Another reason why the frequency shifts are large, is that the vapor concentration is not held at a constant low value (as in most other experiments) but rather the liquid analyte is allowed to evaporate until, potentially, the vapour pressure is reached. These, expected, high vapor concentrations will result in large frequency shifts as seen. This is e.g. observed for the uncoated CMUT in the experiment in Figure 6.7 a), with a maximum frequency shift of $\sim 65\,\mathrm{kHz}$ which shows that the large frequency shifts not only are attributed to the properties of the functionalization layer but also to the vapor concentration.

Comparing selectivities is difficult since it is not easily quantifiable, but the intensity signal in Figure 6.6 for coating A does not increase at all for the acetone and water experiments. Whereas, in Figure 6.7 the CMUT with coating A experiences a frequency shift when both BMK and water were the analytes. Thus, the selectivity can be seen as better for the colorimetric method. This is another reason to combine a gravimetric and colorimetric sensor; the colorimetric detection is the selective part while the gravimetric sensor can give a quantitative output.

In the future the colorimetric detecting scheme could be integrated more with the CMUT by detecting the color change of the functionalization layers/dyes directly on the CMUT surface instead of doing it on a separate chip. This is more difficult as the contrast between the dye and the background is lower on the chip surface than on the white polypropylene substrate.

6.5 Chapter conclusion

A sensor system consisting of a CMUT and a colorimetric chip was fabricated and used to detect BMK in the gas phase. The colorimetric chip displayed an excellent selectivity between BMK, acetone and water, while the CMUT, functionalized with coating A, had an affinity to BMK, as compared with water. In a future sensor system the color change could directly be detected for the coating on the CMUT which would decrease the sensor footprint. An unsolved challenge is the reversibility of the colorimetric dyes, which pull in the direction of keeping the CMUT and colorimetric chips separate and regarding the colorimetric chips as single use.

CHAPTER 7

Alternative read-out method

In this chapter an alternative method of measuring the resonance frequency and driving the CMUT at resonance is presented. For standard CMUTs having two electrodes (top and bottom), the drive signal will be coupled to the output signal. This also means that any noise in the drive signal can affect the output signal. By physically separating the drive electrodes from the sensing electrodes this can be circumvented. Furthermore, having two sensing electrodes the CMUT can give a differential output signal with reduced common-mode noise. The idea and design was conceived in collaboration with M.Sc. E.E. Jan Peter Bagge.

7.1 Working principle

The main idea consists of decoupling the driving electrode on the CMUT from the electrodes sensing the resonance frequency. This can be done by adding an additional electrode to the CMUT plate and using this as the sensing electrode. However, the close proximity of these two top electrodes will cause electrical crosstalk affecting the much smaller sensing signal. If two additional electrodes are added to the plate instead of one, the common mode driving signal (and any common mode noise) can be suppressed as will be shown. Figure 7.1 shows a sketch of a top and cross-sectional view of a CMUT cell with three top electrodes. These two top sensing electrodes are placed symmetrically on the CMUT plate so their capacitance and change in capacitance will be identical.

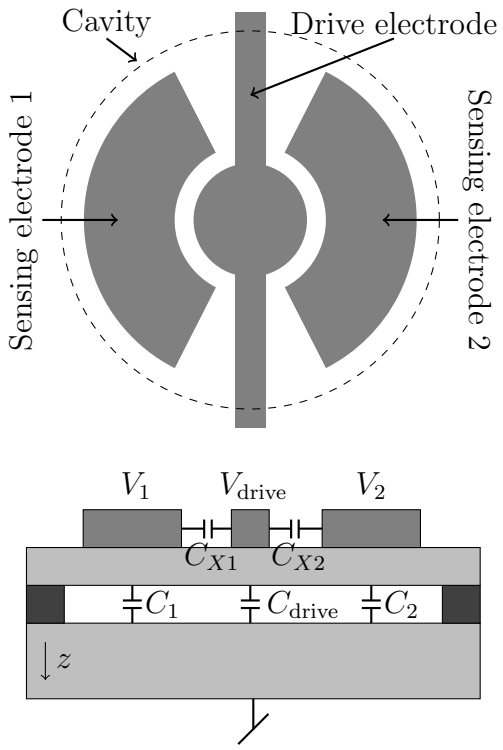


Figure 7.1: Top and cross-sectional view of a single CMUT cell with three top electrodes. In this arrangement the center electrode is the driving electrode, while the two side electrodes are the sensing electrodes.

The fundamental relation of a capacitor is: $Q = C \cdot V$, where Q is the total electrical charge on the capacitor, C is the capacitance, and V is the voltage across the capacitor. In the following the capacitance is for simplicity assumed to be that of a parallel plate capacitor with a vacuum gap and a movable top electrode. If the charges can be assumed to be fixed on the electrodes, the voltage at the two sensing electrodes can be written as:

$$V_1(t) = \frac{Q_1}{C(t)} = \frac{Q_1 z(t)}{\epsilon_0 A_1} \tag{7.1}$$

$$V_2(t) = \frac{Q_2}{C(t)} = \frac{Q_2 z(t)}{\epsilon_0 A_2}, \tag{7.2}$$

where $z(t)$ is the displacement of the top electrode into the gap and $A_1 = A_2$ are the areas of the top electrodes which are designed to be the same.

If the two sensing electrodes are biased with a DC voltage of the same magnitude but with opposite polarity, the charges on the electrodes will be of opposite sign. Since the displacements and areas of the two sensing electrodes

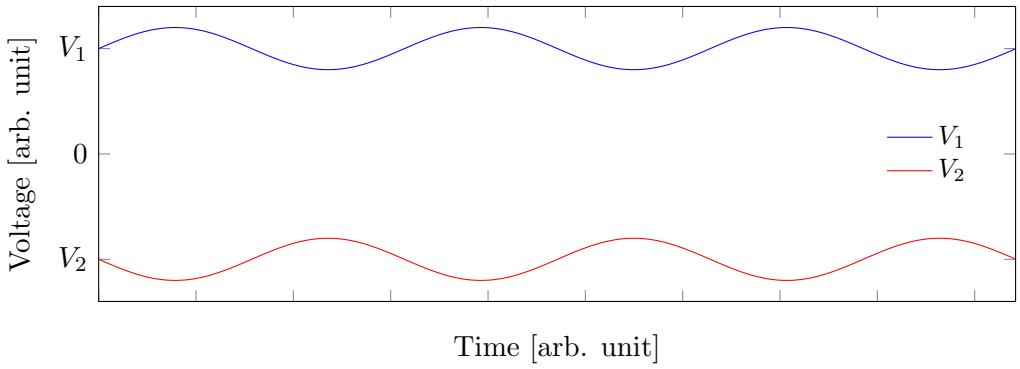


Figure 7.2: Example of the two potentials of V_1 and V_2 showing the opposite biasing and the out of phase AC voltage.

are the same, the AC voltage must (see Equation 7.1) be out of phase by π due to the opposite sign of the charges but have the same frequency, namely that of the mechanical plate oscillation. The two signals are qualitatively plotted in Figure 7.2 showing their out of phase AC component and DC offset.

The circuit design is given in Figure 7.3 a) where to the left the bias voltage is applied and the two sensing electrodes are represented with the two variable capacitance capacitors C_1 and C_2. The two out of phase voltages (V_1 and V_2) are also shown in the figure and passes through the DC blocking capacitor C_b in order to remove the DC component of the signal. The two signals V_{in1} and V_{in2} then enters an instrumentation amplifier which ideally outputs the difference of the two signals while reducing the common mode signal. A possible configuration of the instrumentation amplifier is shown in Figure 7.3 b) where the two amplifiers to the left function as differential buffer amplifiers with a high impedance. The gain of this circuit can be written as:

$$G = \frac{V_{out}}{V_{in2} - V_{in1}} = \left(1 + \frac{2R_A}{R_G}\right)\frac{R_C}{R_B}. \tag{7.3}$$

Ideally the common mode gain is zero, meaning that the common mode rejection ratio is infinite, but in reality this ration will be lower. Consequently, parasitic coupled noise can end up in the output signal V_{out}.

Going back and looking at Figure 7.3 a) the voltages V_{in1} and V_{in2} can be written as a sum of the voltage due to the capacitance change of the CMUT and parasitic coupling of the drive signal (or other noise sources):

$$V_{in1} = V(C_1) + V(C_{X1}) \tag{7.4}$$
$$V_{in2} = V(C_2) + V(C_{X2}). \tag{7.5}$$

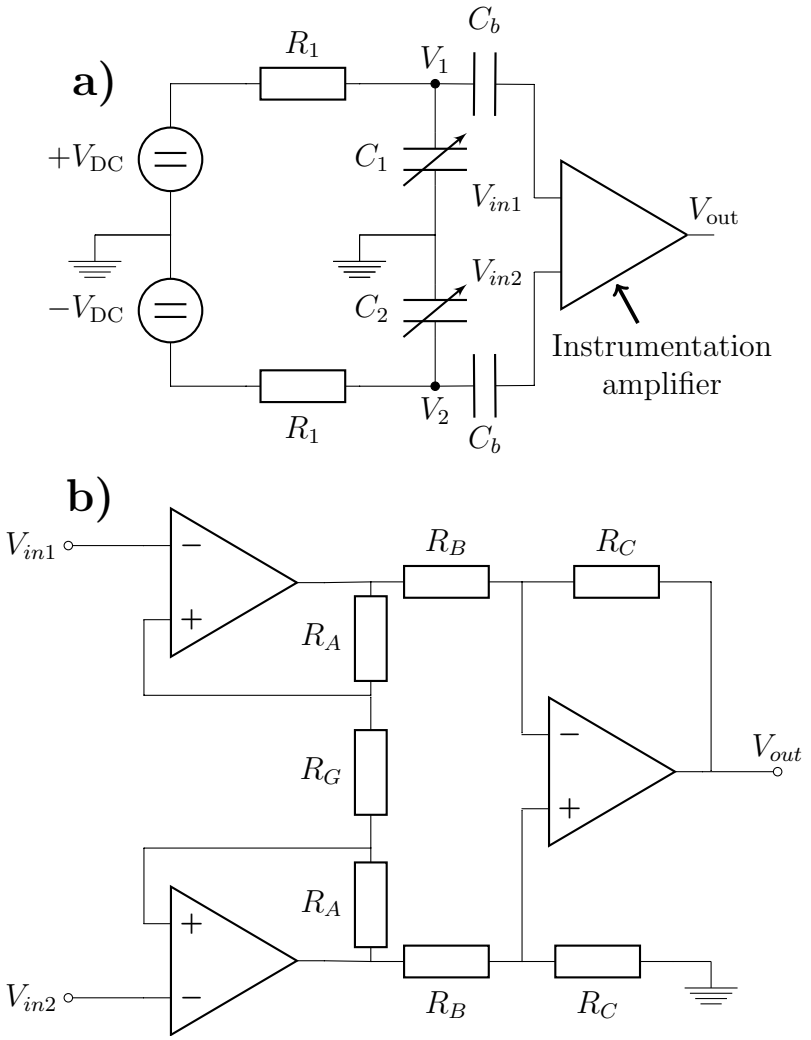


Figure 7.3: a) Read-out circuit for the two sensing electrodes. b) Instrumentation amplifier circuit. The two amplifiers to the left are buffer amplifiers with a high input impedance.

Since the phase and magnitude of the parasitic signals will be the same, these cancel out after the instrumentation amplifier and V_{out} can be written as:

$$V_{\text{out}} = G(V_{in1} - V_{in2}) \tag{7.6}$$
$$= G(V(C_1) + V(C_{X1}) - V(C_2) - V(C_{X2})) \tag{7.7}$$
$$= G(V(C_1) - V(C_2)), \tag{7.8}$$

which will yield an output signal since the input signals are out of phase by π, as illustrated in Figure 7.2.

7.1.1 Circuit design considerations

The component values in the circuit shown in Figure 7.3 a) must be matched carefully for the circuit to function properly. A simplified circuit is shown in Figure 7.4 a), which can be used to approximate the required component values. An additional capacitor C_{in} has been added to represent the input capacitance of the amplifier. Typical values in commercial instrumentation amplifiers are from 1 pF to 10 pF, which is not a lot but in relation to e.g. a 10 pF CMUT it can have a significance as will be shown. Furthermore, the drive signal is seen to couple through the parasitic capacitance C_{X1} to the the CMUT sensing electrode.

The capacitor C_b should block the DC component of the signal so the instrumentation amplifier circuit only sees the AC signal. Hence, the capacitance should be chosen such that C_b does not significantly affect the AC signal. By choosing a high capacitance value, relative to the other capacitors, e.g. 100 nF, the impedance will be much lower than the other components at a normal operating frequency of e.g. 40 MHz. Hence, the capacitor C_b can be omitted in the reduced circuit in Figure 7.4 b).

Next, the value of R_1 is considered. The purpose of this resistor is to keep the charges on the sensing electrode of the CMUT. The impedance of this resistor should be significantly larger than the impedance of the CMUT capacitance at the operating frequency. More formally, the value can be determined by considering the cutoff frequency (-3 dB point) of the parallel coupling of the resistor R_1 and two capacitors C_1 and C_{in} to ground. This makes up a high pass filter and since the two capacitors are connected in parallel their capacitances can be added to form a single equivalent capacitor. The maximum cutoff frequency $f_{c,max}$ should at maximum be a factor of 10 below the operating frequency f_r. The minimum resistance value R_1 can now be calculated:

$$f_{c,max} = f_r/10 = (2\pi R_1 C_1)^{-1} \leftrightarrow \tag{7.9}$$
$$R_{1,min} = ((f_r/10)2\pi C_1)^{-1}, \tag{7.10}$$

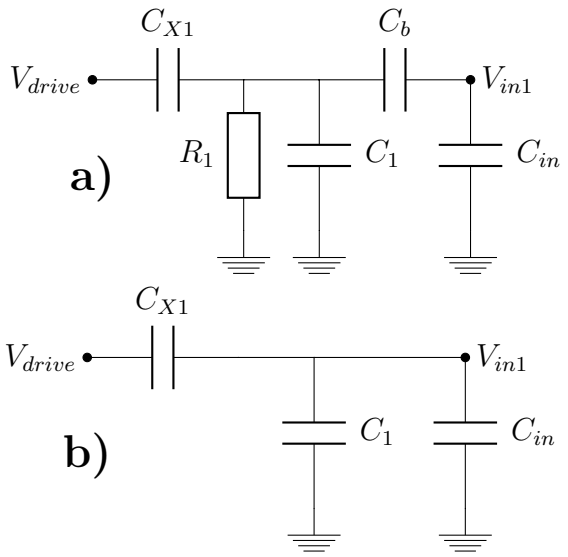


Figure 7.4: a) Circuit showing how the drive signal is coupled parasitically through C_{X1} to the sensing electrode. The circuit can be seen as part of the larger circuit in Figure 7.3 a). Figure 7.4 b) shows a reduced version of the circuit in a).

where only the CMUT capacitance C_1 is used for simplicity, since this typically is an order of magnitude larger than the amplifier input capacitance. Inserting typical values in Equation 7.10 gives the minimum value of R_1:

$$R_{1,\,min} = ((40\,\mathrm{MHz}/10)2\pi 10\,\mathrm{pF})^{-1} = 3979\,\Omega, \tag{7.11}$$

but since this is a minimum value, R_1 can be chosen to have a higher value which effectively decreases the cutoff frequency. A value of $R_1 = 1\,\mathrm{M}\Omega$ is chosen. By choosing resistance values higher than $R_{1,\,min}$, the resistor R_1 can be omitted in the reduced circuit in Figure 7.4 b).

The circuit in 7.4 b) is now analyzed. The circuit can be seen as a voltage divider between the capacitor C_{X1} and an equivalent capacitor with impedance $Z_{eq} = s(C_1 + C_{in})^{-1}$. The ratio between V_{in} and the parasitic capacitively coupled drive signal voltage V_{drive} can now be found using the standard voltage divider relation:

$$\frac{V_{in}}{V_{drive}} = \frac{Z_{eq}}{Z_{X1} + Z_{eq}} = \frac{C_{X1}}{C_1 + C_{in} + C_{X1}} = \frac{1}{\frac{C_1}{C_{X1}} + \frac{C_{in}}{C_{X1}} + 1}. \tag{7.12}$$

Equation 7.12 shows that in the limit where C_{X1} tends towards zero, the voltage V_{in} due to the parasitic coupled drive signal will be zero. That is, the drive signal is no longer coupled to the sensing electrodes. In contrast,

when C_{X1} becomes large relative to the CMUT capacitance C_1 (e.g. due to a poorly designed CMUT), more of the drive signal is coupled to the sensing signal. Thus, the CMUT design should minimize the capacitance C_{X1} which will be discussed further in the next section.

Next, the CMUT capacitance C_1 should be as large as possible relative to C_{X1} and C_{in}, respectively. Since, when the ratio C_1/C_{X1} increases, the ratio V_{in}/V_{drive} decreases. Further, the input capacitance C_{in} acts like a static parasitic capacitance with regard to the CMUT capacitance. This can be written as $\Delta C_1/(C_1+C_{in})$ since C_{in} is connected in parallel with the CMUT. Hence, the input capacitance C_{in} should be decreased in order to increase the relative capacitance change of the CMUT. This can be achieved by carefully choosing the amplifier model, but it cannot be completely eliminated.

7.1.2 Driving the CMUT

The CMUT should be driven at resonance. This can be achieved by creating a feedback loop going from the output of the instrumentation amplifier to the drive electrode of the CMUT. The feedback loop should have unity loop gain and a phase shift of zero at the mechanical resonance frequency of the CMUT.

7.1.3 CMUT design considerations

The CMUT must be designed to accommodate three top electrodes and a bottom electrode. This limits the geometry of the CMUT elements, as the cells can only be placed in 1D lines (and not 2D grids) using conventional processing. If the metal connecting the electrodes of the cells could be lifted above or below each other, without short-circuiting, conventional 2D elements can be made.

The placement of the electrodes relative to each other on the plate can be done in many ways. Figure 7.5 shows two possible configurations. In a) the drive electrode is centered and the two sensing electrodes are placed at the periphery of the cavity. In b) the drive electrode is split in two (but connected electrically) and the two sensing electrodes are centered on the plate. The central placement of the drive electrode, as compared with a drive electrode of the same area not centrally placed, will give a larger electrostatic force on the plate and thus a larger plate deflection resulting in an increased capacitance change. However, placing the sensing electrodes in the center of the plate will give the largest relative change in capacitance and thus signal magnitude.

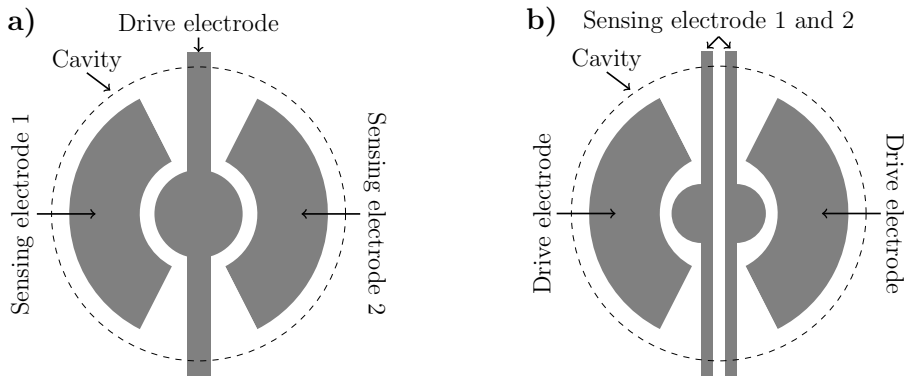


Figure 7.5: Top view of a CMUT cell showing two possible configurations of top electrode placement where a) the drive electrode is centered and the two sensing electrodes are at the periphery of the cavity. b) The two sensing electrodes are centered and the drive electrode is split in two but electrically connected (not shown in the figure).

The area of the electrodes should be maximized from a SNR point of view, in that the electrostatic force will increase when increasing the area of the drive electrode. Further, the amount of charge the sensing electrode can hold will increase for an increased electrode area. The disadvantage of increasing the areas of the electrodes is that they need to be placed closer to each other, hereby increasing the parasitic capacitance C_X between the drive and sensing electrodes. This is true for a plate/cavity of a fixed area which could be a design requirement as this determines the mass sensitivity. The parasitic capacitance can also be decreased by decreasing the height of the metal electrode layers. However, the fringe fields between the electrodes will lessen this effect. In addition, thinner metal layers increase the electrical resistance. The effect of these design considerations needs to be studied more e.g. using FEM modeling.

7.2 Experimental

Chips with three top electrodes have been fabricated. The design was added to the photolithographic mask set of the G4 batch. This resulted in some compromises for the dimensions of the three electrode CMUTs. The plate thickness (50 nm) and vacuum gap height was fixed by the G4 design. Consequently, simulations showed that the maximum radius of the three electrode CMUTs was limited since pull-in would occur when large plate radii were used due to the pressure from the atmosphere on the plate. Hence, the radii

and hereby the area of the plates were limited, resulting in relatively small top electrodes over the cavities. As a consequence the electrical signals are small and near the noise floor in the measurement system used to perform the initial measurements. However, a resonance peak was seen in the impedance spectra.

Figure 7.6 shows optical microscopy images of a single element having just 10 CMUTs placed in a straight line. The bottom electrode contact pad is not shown in this image but is located south of the element. This element has the electrode configuration shown in Figure 7.5 a) where the drive electrode is placed centrally on the plate. These elements are expected to have a high parasitic capacitance relative to the cell capacitance due to the three large contact pads.

Figure 7.7 shows optical microscope images of an element with 10 CMUTs where the sensing electrodes are placed centrally. This is equivalent to the configuration in Figure 7.5 b). In Figure 7.7 left, it is seen how the two driving electrodes are electrically connected at the bottom pad, thus effectively forming a single drive electrode. Elements with 1 and 100 CMUTs were also fabricated for both electrode configurations.

Unfortunately time did not allow for further testing of these devices. The detection principle can be demonstrated by driving a CMUT while measuring the output voltage at a biased sensing electrode. Changing the sign of this bias voltage should then result in a phase shift of π of the AC signal.

Future work could consist of redesigning the devices without the constraints on the plate thickness and vacuum gap. Moreover, the theoretical modeling of the device should be expanded, starting with finding the parasitic capacitance from the drive electrode to the sensing electrodes and hereby calculating the parasitic contribution from the drive signal. This capacitance calculation can e.g. be done using a FEM model to include the, for this geometry, important fringe fields.

7.3 Chapter conclusion

In this chapter an alternative method of measuring an output signal from the CMUTs was presented. Using the described alternative readout method, any noise or nonlinearities in the drive signal will be filtered by the mechanical resonance of the CMUT, thus providing an output signal that is proportional to the physical displacement of the CMUT plate. The idea is based on physically separating the electrode driving the CMUT and the electrodes used for detecting the resonance frequency. Both the CMUT design and electrical circuit design were discussed.

In order to decrease the parasitic coupled noise, the parasitic coupling capacitance C_X should be decreased. At the same time a higher CMUT capacitance increases the signal amplitude relative to any potential noise signals. Finally, the input capacitance C_{in} should be minimized as this acts as a static parasitic capacitance. CMUT elements with two electrode designs were fabricated. However, the main method still needs to be verified experimentally. This could be part of future work.

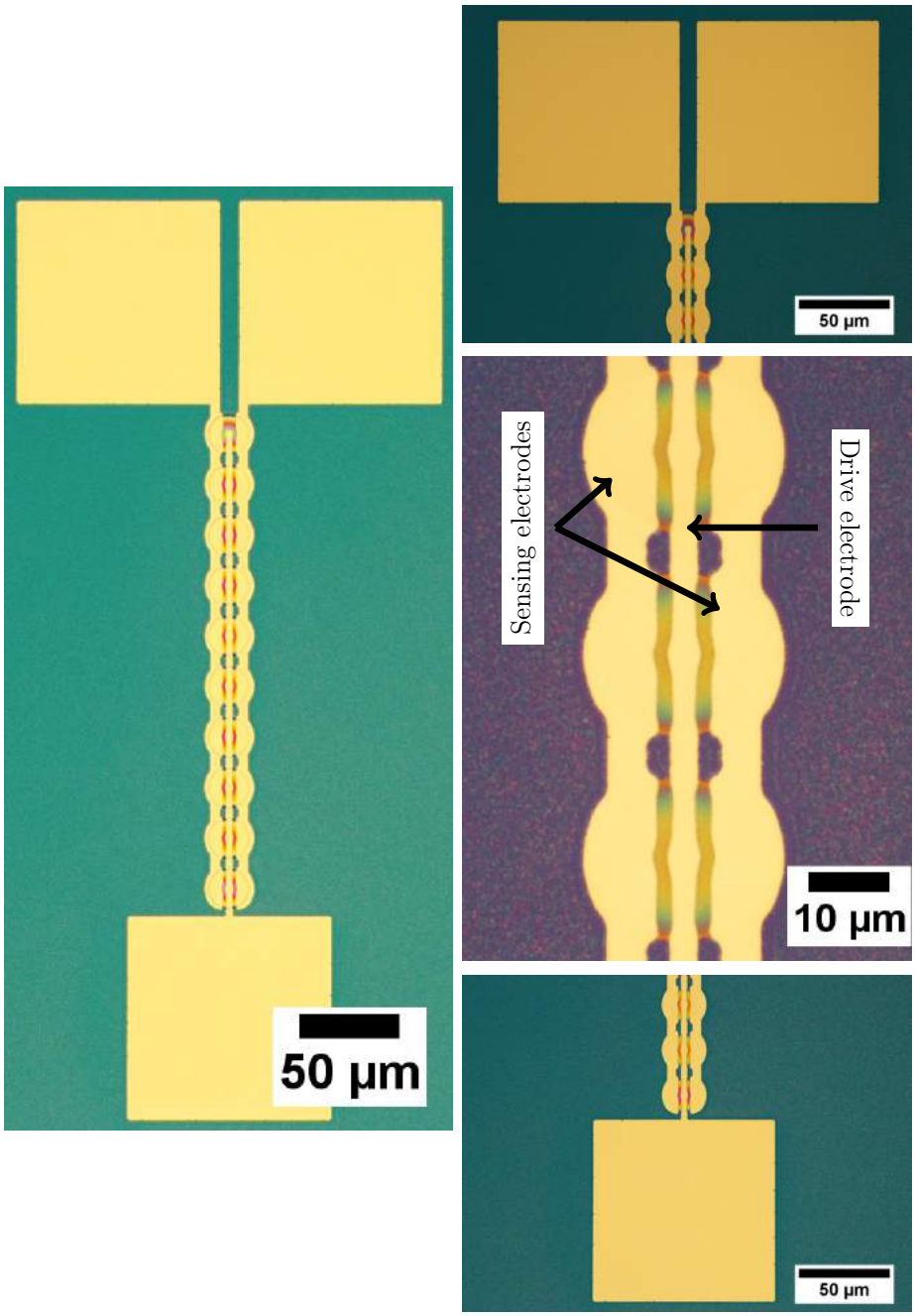


Figure 7.6: Optical microscope images of an element with 10 CMUTs placed in a column. The drive electrode is centrally placed. The circular cavities can be spotted under the electrodes. The bottom electrode contact pad is not in frame in these images.

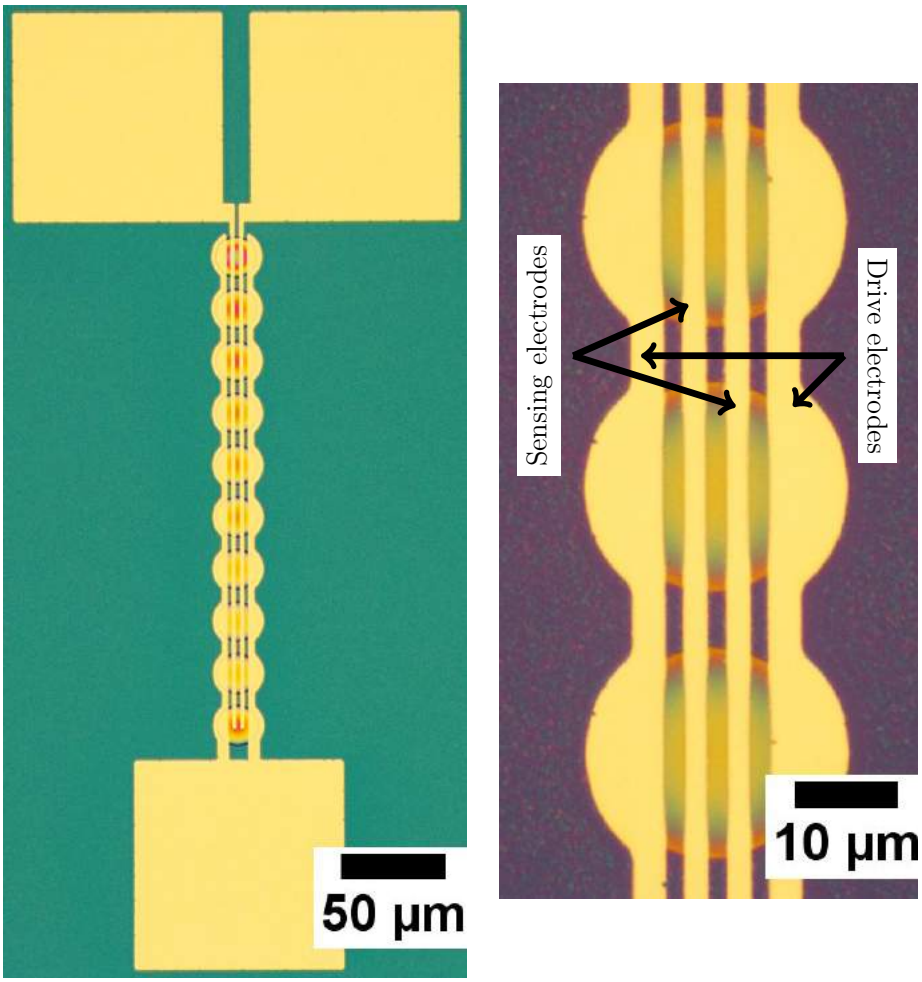


Figure 7.7: Optical microscope images of an element with 10 CMUTs placed in a column. The sensing electrodes are centrally placed. The circular cavities can be spotted under the electrodes. The bottom electrode contact pad is not in frame in these images.

CHAPTER 8

Conclusion and Outlook

The goal of this project was to develop and apply state of the art gas sensors. This goal has been met by modeling, designing, fabricating, and lastly applying CMUTs as gas sensors. The two figures of merit: mass sensitivity and limit of detection were both found to be better than the current state of the art. Furthermore, the first gravimetric detection of BMK was demonstrated.
Chapter 2. A static CMUT model was presented where first the plate deflection profile was found, from which the resonance frequency (with and without plate stress) and capacitance were calculated. Following this an electrostatic analysis resulted in an expression for the pull-in voltage. A dynamic CMUT model was derived, resulting in the classical small signal equivalent circuit. The important electromechanical coupling coefficient was introduced and it was shown how it is increased by operating close to pull-in or decreasing the parasitic capacitance. The sensitivity and normalized sensitivity were calculated analytically and it was seen that they are proportional to a^{-4} and t^{-1}, respectively. When the plates include a built-in tensile stress the sensitivity is also a function of the plate thickness, and the sensitivity is increased for thinner plates. Finally, the limit of detection was shown to be proportional to the mass of the plate and the frequency noise in the system expressed through the Allan deviation.
Chapter 3. In this chapter the design of the devices was presented. The four CMUT generations fabricated in this project were introduced along with an introduction to the fabrication processes. The choice of plate material was discussed by comparing a Si plate from a SOI wafer and a LPCVD Si_3N_4 plate. The latter plate material has the advantages of a reduced price, bet-

ter availability, higher thickness uniformity, reduced parasitic capacitance, increased sensitivity (due to stress), but at the cost of a potentially more challenging fabrication process. The design of double LOCOS cavities was studied and it was found that some combinations of parameters hinder wafer bonding and thus fabrication. Further, it was found that using Si_3N_4 oxidation masks decrease the slope lengths by a factor of 2-3 as compared with using SiO_2. In addition, the pad oxide thickness should be decreased in order to decrease the horizontal extent of the bumps. Finally, when the radius of the cavities is decreased the available design space becomes more limited and the allowed ranges of the parameters are decreased. Hence the sensitivity is ultimately affected if the minimum radius is limited.

The FEM model used to find the resonance frequencies and pull-in voltages of the CMUTs was presented. A mesh convergence test was shown which determined the minimum number of mesh elements, and the model itself was validated by comparing the results with analytical expressions.

The relative parasitic capacitance decreases when: the number of cells in an element increases, the top electrode is structured, a high post to gap ratio is achieved, and the cell to cell distance is decreased.

Finally, the designed dimensions of the four CMUT generations were presented and the theoretical sensitivities calculated.

Chapter 4. In this chapter the fabrication of the CMUT generations was described as well as process optimization of selected process steps. In general, all generations were fabricated using fusion/direct wafer bonding and UV lithography. The cavities are made using one of three processes: a RIE process, a single LOCOS process or a double LOCOS process. Besides the standard Si to SiO_2 bonding, Si_3N_4 to SiO_2 and Si_3N_4 to Si_3N_4 bonding was achieved without a chemical mechanical polishing (CMP) step.

For some of the fabrication steps, process optimization was necessary. A small study showed that the number of voids after wafer bonding and handle layer removal increased when the RCA cleaning step was omitted for one or both of the wafers to be bonded. More specifically, the wafer pair where both the top and bottom wafers were RCA cleaned showed a chip yield of 100%, while the same number was 97% and 92% when only the bottom wafer was RCA clean or no RCA cleaning was performed, respectively. Hence, a direct correlation between the wafer cleaning and formation of voids was observed.

The surface roughness was monitored for two wafers where the Si_3N_4 oxidation mask was dry etched. The wafer which had the pad oxide layer subsequently dry etched ended up with a surface roughness 68% higher than the wafer where the pad oxide was wet etched in BHF. However, for both the wafers the absolute surface roughness remained under the maximum threshold of $R_{\mathrm{RMS}} = 0.5\,\mathrm{nm}$.

Chapter 5. In this chapter the fabricated CMUTs were characterized using various measurement techniques. The resonance frequency and pull-in voltage were determined experimentally by measuring the impedance of the CMUTs. The relative difference between the measured and simulated resonance frequency and pull-in voltage was for a G4 element seen to be below 2%, demonstrating the predictive power of the FEM model. The electromechanical coupling coefficient was seen to be higher for designs that minimize the parasitic capacitance. This was exemplified by measurements of two G2 elements, where the element with a patterned top electrode had a k^2 at $V_{\mathrm{DC}}/V_{\mathrm{PI}} = 0.94$ that was 102% higher than for an element with a non-patterned top electrode, due to the difference of parasitic capacitance.

The limit of detection was determined by measuring the frequency noise and calculating the Allan deviation. It was found that the limit of detection for both G2 and G3 is lower than the previously lowest published value of a CMUT by 51% (for G3). Furthermore, the LOD per area was compared other sensor types and here it was found that the CMUTs excelled, in that they had the lowest LOD/A values.

The mass sensitivity was studied by analytical expressions, FEM modeling, and measurements. The non-linearity of the sensitivity was seen in the results for both the FEM model and measurements, when the amount of mass on the plate was increased. The sensitivity decreased by 46% for a mass loading of $m_{\mathrm{added}}/m_{\mathrm{plate}} = 100\%$. In addition, an excellent agreement was found between the analytically determined sensitivity and the sensitivity found by the FEM model with a mass boundary layer, with a relative difference of 0.7% at zero mass loading. The theoretical sensitivity was compared with the state of the art CMUTs in the literature and it was found that the G4 device has a sensitivity 257% higher than the highest previously published value, to the knowledge of the author. In order to compare the sensitivities of different sensor types the normalized sensitivity was calculated. Here the highest normalized sensitivities were found for the NEMS devices followed by the CMUTs.

Chapter 6. In this chapter the CMUT was applied as a gas sensor for detecting BMK, which is an illegal precursor molecule for the synthesis of (meth)amphetamine. The CMUT was functionalized with two different coatings and combined with a colorimetric chip to increase the selectivity of the sensor system and robustness/redundancy. Coating A on the colorimetric chip showed excellent selectivity between BMK, acetone, and water. The CMUT functionalized with coating A showed the largest frequency shift, demonstrating the affinity for BMK of coating A. In other words the CMUT also displayed selectivity between BMK and water due to the functionalization layers.

Chapter 7. In the last chapter a new method of sensing the plate movement was presented. By physically splitting the top electrode into three separate electrodes the signal driving the CMUT is decoupled from the signal sensing the plate movement. The two sensing electrodes are biased with voltages of different sign resulting in their AC signals being out of phase by π. The difference of these two out of phase signals are then to be taken, hereby eliminating any common mode noise that may be present.

The signal driving the CMUT was seen to capacitively couple to the sensing electrode through a parasitic capacitance C_X, which should be minimized. Furthermore, the capacitance of the CMUT should be maximized to increase the signal from the sensing electrodes relative to the parasitic coupled noise signal.

The CMUT design was considered and a new design trade-off was found between the electrostatic force and the relative change in capacitance of the sensing electrodes.

Finally, two different configurations of fabricated CMUTs with three top electrodes were presented. Their function as 'normal' CMUTs with two electrodes have been demonstrated experimentally but the new detection principle has yet to be experimentally verified.

Outlook

The mass sensitivity can, as demonstrated, be predicted, whereas the mass limit of detection is more difficult to theoretically calculate or simulate since absolute values of the frequency noise are challenging to estimate. Thus, a model is needed that can predict the noise of a given CMUT design making a theoretical prediction of the mass limit of detection possible.

The chip carriers used in this project are both expensive and cumbersome to handle. For future work flat and cheaper chip carriers should be employed.

The mass linear and dynamic range and hereby the concentration dynamic range can be increased by having a chip with multiple elements. Each element should have a different mass sensitivity. This could simply be achieved by changing the thickness or radius of the cells. In this way elements with cells with a high sensitivity can be used for sensing when a low gas concentration is present and elements with a lower sensitivity can be used for higher gas concentrations. The linear range of the low sensitivity cells is greater than the high sensitivity cells, which on the other hand give larger signals for low gas concentrations.

The current method of driving the CMUT and measuring the resonance frequency evolves the use of a phase locked loop. Currently, the setup is

bulky, and the equipment expensive and thus not suitable for use outside of a laboratory. Several alternatives exist: small relatively cheap PLL chips are available which would lower the cost and size of the setup, an oscillator circuit could be used as has been demonstrated extensively in the literature [3], and finally a read-out and drive method such as the one presented in Chapter 7 could be used, although the feasibility of this last option still needs to be experimentally verified.

Finally, the main challenge for most gravimetric sensors is how to increase the selectivity. The CMUTs shows both a high mass sensitivity (and normalized mass sensitivity) and a low LOD/A which is ideal for a gas sensor. However, the selectivity must be increased as typically has been done by applying a functionalization layer, which was also done in this project. Other methods of bettering the selectivity typically increase the complexity of the sensor, such as combining the CMUT with another sensor type which also was done in this project with the colorimetric chips or using several CMUT elements on the same chip with different functionalization layers as was done in [29]. Regardless of the method this is an area in which more research could be devoted.

Bibliography

[1] Business Wire. (2017) Global gas sensors market - growth, trends and forecasts to 2021 - research and markets. Accessed: 2018-01-09. [Online]. Available: https://goo.gl/hS3KV9

[2] Radiant Insights. (2015) Gas sensors market size, share, industry report, 2020. Accessed: 2018-01-09. [Online]. Available: https://goo.gl/n9opSz

[3] H. J. Lee, K. K. Park, M. Kupnik, Ö. Oralkan, and B. T. Khuri-Yakub, "Chemical Vapor Detection Using a Capacitive Micromachined Ultrasonic Transducer," *Anal. Chem.*, vol. 83, pp. 9314–9320, 2011.

[4] A. K. Naik, M. S. Hanay, W. K. Hiebert, X. L. Feng, and M. L. Roukes, "Towards single-molecule nanomechanical mass spectrometry," *Nature Nanotechnology*, vol. 4, pp. 445–450, 2009.

[5] E. Sage, A. Brenac, T. Alava, R. Morel, C. Dupré, M. S. Hanay, M. L. Roukes, L. Duraffourg, C. Masselon, and S. Hentz, "Neutral particle mass spectrometry with nanomechanical systems," *Nature Communications*, vol. 6, p. 6482, 2015.

[6] X. Liu, S. Cheng, H. Liu, S. Hu, D. Zhang, and H. Ning, "A Survey on Gas Sensing Technology," *Sensors*, vol. 12, pp. 9635–9665, 2012.

[7] K. Arshak, E. Moore, G. Lyons, J. Harris, and S. Clifford, "A review of gas sensors employed in electronic nose applications," *Sensor Review*, vol. 24, no. 2, p. 181–198, 2004.

[8] R. Bogue, "Detecting gases with light: a review of optical gas sensor technologies," *Sensor Review*, vol. 35, no. 2, p. 133–140, 2015.

[9] S. Matindoust, M. Baghaei-Nejad, and M. H. S. Abadi, "Food quality and safety monitoring using gas sensor array in intelligent packaging," *Sensor Review*, vol. 36, no. 2, pp. 169–183, 2016.

[10] G. M. Insights. (2017) Gas sensors market size by technology. Accessed: 2018-01-11. [Online]. Available: https://goo.gl/PY9GHE

[11] S. Fanget, S. Hentz, P. Puget, J. Arcamone, M. Matheron, E. Colinet, P. Andreucci, L. Duraffourg, E. Myers, and M. Roukes, "Gas sensors based on gravimetric detection - A review," *Sensors and Actuators B: Chemical*, vol. 160, pp. 804 – 821, 2011.

[12] National Instruments, "White paper: Sensor terminology," 2013, accessed: 2018-01-11. [Online]. Available: https://goo.gl/iBRsYH

[13] J. Irish, "Course note on instrumentation specifications," 2005, accessed: 2018-01-11. [Online]. Available: https://goo.gl/8E8Dr3

[14] R. A. McGill, V. K. Nguyen, R. Chung, R. E. Shaffer, D. DiLella, J. L. Stepnowski, T. E. Mlsna, D. L. Venezky, and D. Dominguez, "The "NRL-SAWRHINO": a nose for toxic gases," *Sensors and Actuators B*, vol. 65, pp. 10–13, 2000.

[15] Z. Yunusa, M. N. Hamidon, A. Kaiser, and Z. Awang, "Gas sensors: A review," *Sensors and Transducers*, vol. 168, no. 4, pp. 61–75, 2014.

[16] S. Schmid, L. G. Villanueva, and M. L. Roukes, *Fundamentals of Nanomechanical Resonators*. Springer, 2016.

[17] S. W. Wenzel and R. M. White, "Analytic comparison of the sensitivities of bulkuwave, surfacemwave, and flexural plate-wave ultrasonic gravimetric sensors," *Appl. Phys. Lett.*, vol. 54, no. 20, pp. 1976–1978, 1989.

[18] Y. T. Yang, C. Callegari, X. L. Feng, K. L. Ekinci, and M. L. Roukes, "Zeptogram-Scale Nanomechanical Mass Sensing," *Nano Lett.*, vol. 6, no. 4, 2006.

[19] M. Li, H. X. Tang, and M. L. Roukes, "Ultra-sensitive NEMS-based cantilevers for sensing, scanned probe and very high-frequency applications," *Nature Nanotechnology*, vol. 2, pp. 114–120, 2007.

[20] Y. T. Yang, C. Callegari, X. L. Feng, and M. L. Roukes, "Surface Adsorbate Fluctuations and Noise in Nanoelectromechanical Systems," *Nano Lett.*, vol. 11, pp. 1753–1759, 2011.

[21] H. J. Lee, K. K. Park, Ömer Oralkan, M. Kupnik, and B. T. Khuri-Yakub, "A Multichannel Oscillator for a Resonant Chemical Sensor System," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 10, pp. 5632–5640, 2014.

[22] G. Wingqvist, L. Arapan, V. Yantchev, and I. Katardjiev, "A micromachined thermally compensated thin film Lamb wave resonator for frequency control and sensing applications," *J. Micromech. Microeng.*, vol. 19, 2009.

[23] J. Weber, W. M. Albers, J. Tuppurainen, M. Link, R. Gabl, W. Wersing, and M. Schreiter, "Shear mode FBARs as highly sensitive liquid biosensors," *Sensors and Actuators A*, vol. 128, p. 84–88, 2006.

[24] Y. lu Feng, M. wei Liu, X. Li, and J. Li, "Micromachined mass sensor based on film bulk acoustic resonator," *Symposium on Piezoelectricity, Acoustic Waves, and Device Applications*, 2015.

[25] A. Cagliani and Z. J. Davis, "Ultrasensitive bulk disk microresonator-based sensor for distributed mass sensing," *J. Micromech. Microeng.*, vol. 21, 2011.

[26] H. Zhang and E. S. Kim, "Micromachined Acoustic Resonant Mass Sensor," *Journal of Microelectromechanical Systems*, vol. 14, no. 4, pp. 699–706, 2005.

[27] M. Pohanka, "The Piezoelectric Biosensors: Principles and Applications, a Review," *Int. J. Electrochem. Sci.*, vol. 12, p. 496–506, 2017.

[28] D. Virzonis, G. Vanagas, A. Ramanaviciene, A. Makaraviciute, D. Barauskas, A. Ramanavicius, W. Wen, and R. Kodzius, "Resonant gravimetric immunosensing based on capacitive micromachined ultrasound transducers," *Microchim Acta*, vol. 181, p. 1749–1757, 2014.

[29] Q. Stedman, K. K. Park, and B. T. Khuri-Yakub, "An 8-Channel CMUT Chemical Sensor Array on a Single Chip," *IEEE International Ultrasonics Symposium (IUS)*, 2017.

[30] F. L. Degertekin, R. O. Guldiken, and M. Karaman, "Annular-Ring CMUT Arrays for Forward-Looking IVUS: Transducer Characterization and Imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 53, no. 2, pp. 474–482, 2006.

[31] A. S. Savoia, G. Caliano, and M. Pappalardo, "A CMUT Probe for Medical Ultrasonography: From Microfabrication to System Integration," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 59, no. 6, pp. 1127–1138, 2012.

[32] B. T. Khuri-Yakub and Ö. Oralkan, "Capacitive micromachined ultrasonic transducers for medical imaging and therapy," *J. Micromech. Microeng.*, vol. 21, 2011.

[33] M. Engholm, T. L. Christiansen, C. Beers, J. P. Bagge, L. N. Moesner, H. Bouzari, A. Lei, M. Berkheimer, M. B. Stuart, J. A. Jensen, and E. V. Thomsen, "A hand-held row-column addressed CMUT probe with integrated electronics for volumetric imaging," *IEEE International Ultrasonics Symposium Proceedings*, 2015.

[34] B. T. Khuri-Yakub, C.-H. Cheng, F.-L. Degertekin, S. Ergun, S. Hansen, X.-C. Jin, and O. Oralkan, "Silicon Micromachined Ultrasonic Transducers," *Jpn. J. Appl. Phys.*, vol. 39, p. 2883–2887, 2000.

[35] N. Apte, K. K. Park, A. Nikoozadeh, and B. T. Khuri-Yakub, "Bandwidth and Sensitivity Optimization in CMUTs for Airborne Applications," *IEEE International Ultrasonics Symposium Proceedings*, pp. 166–169, 2014.

[36] A. Octavio, C. J. Martin, Y. Gomez-Ullate, O. Martinez, L. Gomez-Ullate, F. M. de Espinosa, P. Gatta, and M. Dominguez, "Design and Characterization of Air Coupled Ultrasonic Transducers Based on MUMPs," *IEEE International Ultrasonics Symposium Proceedings*, pp. 2373–2376, 2006.

[37] X. Wang, Y. Fan, W.-C. Tian, H.-J. Kwon, S. Kennerly, G. Claydon, and A. May, "An Air-Coupled Capacitive Micromachined Ultrasound Transducer for Noncontact Nondestructive Evaluation," *IEEE Sensors*, pp. 1464–1467, 2007.

[38] L. L. Wong, A. I. Chen, Z. Li, A. S. Logan, and J. T. Yeow, "A row–column addressed micromachined ultrasonic transducer array for surface scanning applications," *Ultrasonics*, vol. 54, p. 2072–2080, 2014.

[39] Aditi, R. Mukhiya, R. Gopal, and V. K. Khanna, "FEM simulation of CMUT cell for NDT application," *16th International Workshop on Physics of Semiconductor Devices*, 2012.

[40] K. K. Park and H. J. Lee and G. G. Yaralioglu and A. S. Ergun and Ö. Oralkan and M. Kupnik and C. F. Quate and B. T. Khuri-Yakub and T. Braun and J.-P. Ramseyer and H. P. Lang and M. Hegner and Ch. Gerber and J. K. Gimzewski, "Capacitive micromachined ultrasonic transducers for chemical detection in nitrogen," *Applied Physics Letters*, vol. 91, 2007.

[41] D. Barauskas, D. Pelenis, D. Virzonis, J. P. Baltrus, and J. Baltrusaitis, "Greenhouse Gas Molecule CO2 Detection Using a Capacitive Micromachined Ultrasound Transducer," *Anal. Chem.*, vol. 88, no. 13, pp. 6662–6665, 2016.

[42] K. K. Park, H. Lee, M. Kupnik, and B. T. Khuri-Yakub, "Fabrication of capacitive micromachined ultrasonic transducers via local oxidation and direct wafer bonding," *Journal of Microelectromechanical Systems*, vol. 20, no. 1, 2011.

[43] M. Engholm, H. Bouzari, T. L. Christiansen, C. Beers, J. P. Bagge, L. N. Moesner, S. E. Diederichsen, M. B. Stuart, J. A. Jensen, and E. V. Thomsen, "Probe development of CMUT and PZT row-column-addressed 2-D arrays," *Sensors and Actuators A: Physical*, vol. 273, p. 121–133, 2018.

[44] A. S. Ergun, Y. Huang, X. Zhuang, Ömer Oralkan, G. G. Yaralıoğlu, and B. T. Khuri-Yakub, "Capacitive Micromachined Ultrasonic Transducers: Fabrication Technology," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 52, no. 12, pp. 2242–2258, 2005.

[45] M. I. Haller and B. T. Khuri-Yakub, "A surface micromachined electrostatic ultrasonic air transducer," *IEEE Ultrasonics Symposium*, pp. 1241–1244, 1994.

[46] B. T. Khuri-Yakub and Ö. Oralkan, "Capacitive micromachined ultrasonic transducers for medical imaging and therapy," *J. Micromech Microeng.*, vol. 21, no. 5, 2011.

[47] Y. Huang, A. S. Ergun, E. Hæggström, M. H. Badi, and B. T. Khuri-Yakub, "Fabricating Capacitive Micromachined Ultrasonic Transduc-

ers With Wafer-Bonding Technology," *Journal of Microelectromechanical Systems*, vol. 12, no. 2, pp. 128–137, 2003.

[48] K. K. Park, H. Lee, M. Kupnik, Ö. Oralkan, J.-P. Ramseyer, H. P. Lang, M. Hegner, C. Gerber, and B. T. Khuri-Yakub, "Capacitive micromachined ultrasonic transducer (CMUT) as a chemical sensor for DMMP detection," *Sensors and Actuators B: Chemical*, vol. 160, p. 1120– 1127, 2011.

[49] M. M. Mahmud, M. Kumar, X. Zhang, F. Y. Yamaner, H. T. Nagle, and Ö. Oralkan, "A Capacitive Micromachined Ultrasonic Transducer (CMUT) Array as a Low-Power Multi-Channel Volatile Organic Compound (VOC) Sensor," *IEEE Sensors*, 2015.

[50] K. K. Park and H. J. Lee and M. Kupnik and Ö. Oralkan and B. T. Khuri-Yakub, "Fabricating capacitive micromachined ultrasonic transducers with direct wafer-bonding and locos technology," *IEEE MEMS*, pp. 339–342, 2008.

[51] K. Midtbø, A. Rønnekleiv, and D. T. Wang, "Fabrication and Characterization of CMUTs realized by Wafer Bonding," *IEEE Ultrasonics Symposium*, pp. 934 – 937, 2006.

[52] A. I. H. Chen, L. L. P. Wong, Z. Li, S. Na, and J. T. W. Yeow, "Practical CMUT Fabrication With a Nitride-to-Oxide-Based Wafer Bonding Process," *Journal of Microelectromechanical Systems*, vol. 26, no. 4, pp. 829 – 836, 2017.

[53] M. J. G. Mølgaard, M. Laustsen, M. H. Jakobsen, T. L. Andresen, and E. V. Thomsen, "Combined Colorimetric and Gravimetric CMUT Sensor for Detection of Benzyl Methyl Ketone," *Sensors and Actuators B: Chemical*, vol. -, no. -, pp. –, 2018, under review.

[54] M. J. G. Mølgaard, J. Fly, M. H. Jakobsen, T. L. Andresen, and E. V. Thomsen, "Sensitivity optimization of wafer bonded gravimetric CMUT sensors," *Journal of Microelectromechanical Systems*, vol. -, no. -, pp. –, 2018, submitted.

[55] B. T. Khuri-Yakub, K. K. Park, H. J. Lee, G. G. Yaralioglu, S. Ergun, Ö. Oralkan, M. Kupnik, C. F. Quate, T. Braun, H. P. Lang, M. Hegner, J.-P. Ramseyer, C. Gerber, and J. Gimzewski, "The Capactive Micromachined Ultrasonic Transducer (CMUT) as a Chem/Bio Sensor," *IEEE International Ultrasonics Symposium (IUS)*, pp. 472–475, 2007.

[56] M. J. G. Mølgaard, M. Laustsen, I. L. Thygesen, M. H. Jakobsen, T. L. Andresen, and E. V. Thomsen, "Combined Colorimetric and Gravimetric CMUT Sensor for Detection of Phenylacetone," *IEEE International Ultrasonics Symposium (IUS)*, 2017.

[57] Y. Li, R. Lucklum, and P. Hauptmann, "2-Step Surface Modification Technology for Acoustic Chemical Sensor Arrays based on CMUTs," *IEEE International Ultrasonics Symposium Proceedings*, 2008.

[58] Y. Li, R. Lucklum, A. Brose, and M. Fleischer, "A 2-step coating technique for CMUT based chemical sensors," *Sensors and Actuators B: Chemical*, vol. 161, p. 171– 177, 2012.

[59] H. J. Lee, K. K. Park, Ö. Oralkan, M. Kupnik, and B. T. Khuri-Yakub, "CMUT as a chemical sensor for DMMP detection," *IEEE International Frequency Control Symposium*, pp. 434–439, 2008.

[60] K. K. Park, H. J. Lee, M. Kupnik, Ö. Oralkan, and B. T. Khuri-Yakub, "Capacitive Micromachined Ultrasonic Transducer as a Chemical Sensor," *IEEE Sensors conference proceedings*, pp. 5–8, 2008.

[61] H. J. Lee, K. K. Park, M. Kupnik, Ö. Oralkan, and B. T. Khuri-Yakub, "Highly Sensitive Detection of DMMP Using a CMUT-based Chemical Sensor," *IEEE Sensors conference proceedings*, pp. 2122–2126, 2010.

[62] H. J. Lee, K. K. Park, M. Kupnik, N. A. Melosh, and B. T. Khuri-Yakub, "Mesoporous Thin-Film on Highly-Sensitive Resonant Chemical Sensor for Relative Humidity and CO2 Detection," *Anal. Chem.*, vol. 84, no. 7, pp. 3063–3066, 2012.

[63] H. J. Lee, K. K. Park, M. Kupnik, and B. T. Khuri-Yakub, "Functionalization layers for CO2 sensing using capacitive micromachined ultrasonic transducers," *Sensors and Actuators B: Chemical*, vol. 174, p. 87– 93, 2012.

[64] D. Barauskas, D. Pelenis, G. Sergalis, G. Vanagas, M. Mikolajunas, D. Virzonis, and J. Baltrusaitis, "CMUT for high sensitivity greenhouse gas sensing," *IEEE International Ultrasonics Symposium (IUS) proceedings*, 2015.

[65] Q. Stedman, K.-K. Park, and B. T. Khuri-Yakub, "Distinguishing Chemicals Using CMUT Chemical Sensor Array and Artificial Neural Networks," *IEEE International Ultrasonics Symposium Proceedings*, 2014.

[66] C. Seok, M. M. Mahmud, O. Adelegan, X. Zhang, and Ö. Oralkan, "A battery-operated wireless multichannel gas sensor system based on a capacitive micromachined ultrasonic transducer (CMUT) array," *IEEE Sensors*, 2016.

[67] A. Ramanaviciene, D. Virzonis, G. Vanagas, and A. Ramanavicius, "Capacitive micromachined ultrasound transducer (cMUT) for immunosensor design," *Analyst*, vol. 135, p. 1531–1534, 2010.

[68] D. Barauskas, G. Vanagas, D. Virzonis, A. Ramanaviciene, A. Makaraviciute, and A. Ramanavicius, "Application of CMUT as immunosensor," *IEEE International Ultrasonics Symposium (IUS)*, pp. 1708–1711, 2013.

[69] D. Barauskas, G. Vanagas, D. Virzonis, A. Makaraviciute, and A. Ramanaviciene, "Capacitive Micromachined Ultrasound Transducers (CMUT) for Resonant Gravimetric Immunosensing," *IEEE International Ultrasonics Symposium (IUS)*, pp. 2588–2590, 2014.

[70] O. Onen, P. Kruk, and R. Guldiken, "A MEMS ultrasonic sensor design for early detection of ovarian cancer," *Proc. of SPIE*, vol. 7929, p. 79290X, 2011.

[71] M. H. Badi, G. G. Yaralioglu, A. S. Ergun, S. T. Hansen, and B. Khuri-Yakub, "Lamb Wave Devices Based On Capacitive Micromachined Ultrasonic Transducers," *IEEE Ultrasonics Symposium Proceedings*, pp. 1031–1034, 2002.

[72] J. McLean and F. L. Degertekin, "Interdigital capacitive micromachined ultrasonic transducers for sensing and pumping in microfluidic applications," *The 12th International Conference on Solid State Sensors Actuators and Microsystems*, pp. 915–918, 2003.

[73] M. Thränhardt, P.-C. Eccardt, H. Mooshofer, P. Hauptmann, and L. Degertekin, "A resonant CMUT Sensor for Fluid Applications," *IEEE Sensors*, pp. 878–883, 2009.

[74] E. Ventsel and T. Krauthammer, *Thin Plates and Shells*. Marcel Dekker, 2001.

[75] E. V. Thomsen, K. Reck, G. Skands, C. Bertelsen, and O. Hansen, "Silicon as an anisotropic mechanical material: Deflection of thin crystalline plates," *Sensors and Actuators A: Physical*, vol. 220, pp. 347 – 364, 2014.

[76] M. F. la Cour, T. L. Christiansen, J. A. Jensen, and E. V. Thomsen, "Electrostatic and small-signal analysis of CMUTs with circular and square anisotropic plates," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 62, no. 8, pp. 1563–1579, 2015.

[77] M. Engholm, T. Pedersen, and E. V. Thomsen, "Modeling of plates with multiple anisotropic layers and residual stress," *Sensors and Actuators A: Physical*, vol. 240, pp. 70–79, 2016.

[78] S. D. Senturia, *Microsystem design.* Springer, 2001.

[79] T. L. Christiansen, "Micromachined ultrasonic transducers for 3-d imaging," Ph.D. dissertation, Technical University of Denmark, 2015.

[80] L.-F. Ge, "A Design of High-Sensitivity Micromachined Capacitive Ultrasonic Mass Resonators," *IEEE International Ultrasonics Symposium Proceedings*, pp. 1034–1037, 2008.

[81] W. J. Riley, "Handbook of frequency stability analysis," p. 16, 2008, NIST Special Publication 1065.

[82] IEEE, "IEEE standard definitions of physical quantities for fundamental frequency and time metrology - random instabilities," 1999, IEEE Std 1139-1999.

[83] M. Sansa, E. Sage, E. C. Bullard, M. Gély, T. Alava, E. Colinet, A. K. Naik, L. G. Villanueva, L. Duraffourg, M. L. Roukes, G. Jourdan, and S. Hentz, "Frequency fluctuations in silicon nanoresonators," *Nature Nanotechnology*, vol. 11, pp. 552–559, 2016.

[84] H. J. Lee, K. K. Park, P. Cristman, Ö. Oralkan, M. Kupnik, and B. T. Khuri-Yakub, "The effect of parallelism of CMUT cells on phase noise for chem/bio sensor applications," *IEEE International Ultrasonics Symposium Proceedings*, pp. 1951–1954, 2008.

[85] N. Lobontiu, *Dynamics of Microelectromechanical Systems.* Springer, 2007.

[86] K. K. Park, H. J. Lee, P. Crisman, M. Kupnik, Ö. Oralkan, and B. T. Khuri-Yakub, "Optimum design of circular CMUT membranes for high quality factor in air," *IEEE International Ultrasonics Symposium Proceedings*, pp. 504–507, 2008.

[87] I. Wilson-Rae, R. A. Barton, S. S. Verbridge, D. R. Southworth, B. Ilic, H. G. Craighead, and J. M. Parpia, "High-Q Nanomechanics via Destructive Interference of Elastic Waves," *Phys. Rev. Lett.*, vol. 106, p. 047205, 2011.

[88] A. Logan and J. T. W. Yeow, "Fabricating Capacitive Micromachined Ultrasonic Transducers with a Novel Silicon-Nitride-Based Wafer Bonding Process," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 5, pp. 1074–1084, 2009.

[89] K. Reck, C. Østergaard, E. V. Thomsen, and O. Hansen, "Fusion bonding of silicon nitride surfaces," *J. Micromech. Microeng.*, vol. 21, pp. 1–5, 2011.

[90] D. Chin, S.-Y. Oh, S.-M. Hu, R. W. Dutton, and J. L. Moll, "Two-dimenlsional oxidation," *IEEE Transactions on electron devices*, vol. 30, no. 7, pp. 744–749, 1983.

[91] T.-C. Wu, W. T. Stacy, and K. N. Ritz, "The Influence of the LOCOS Processing Parameters on the Shape of the Bird's Beak Structure," *J. Electrochem. Soc.: Solid-state Science and Technology*, vol. 130, no. 7, pp. 1563–1566, 1983.

[92] N. Guillemot, G. Pananakakis, and P. Chenevier, "A New Analytical Model of the "Bird's Beak"," *IEEE Transactions on electron devices*, vol. 34, no. 5, pp. 1033–1038, 1987.

[93] A. F. Sarioglu, M. Kupnik, S. Vaithilingam, and B. T. Khuri-Yakub, "Nanoscale topography of thermally-grown oxide films at right-angled convex corners of silicon," *Journal of The Electrochemical Society*, vol. 159, no. 2, pp. 79–84, 2012.

[94] T. L. Christiansen, O. Hansen, J. A. Jensen, and E. V. Thomsen, "Thermal oxidation of structured silicon dioxide," *ECS Journal of Solid State Science and Technology*, vol. 3, no. 5, pp. 63–68, 2014.

[95] E. V. Thomsen, K. Reck, G. Skands, C. Bertelsen, and O. Hansen, "Silicon as an anisotropic mechanical material: Deflection of thin crystalline plates," *Sensors and Actuators A: Physical*, vol. 220, pp. 347–364, 2014.

[96] K. R. Williams and R. S. Muller, "Etch rates for micromachining processing," *Journal of Microelectromechanical Systems*, vol. 5, no. 4, pp. 256–269, 1996.

[97] E. Benes, "Improved quartz crystal microbalance technique," *J. Appl. Phys.*, vol. 56, no. 3, pp. 608–626, 1984.

[98] K. L. Ekinci and M. L. Roukes, "Nanoelectromechanical systems," *Rev. Sci. Instrum.*, vol. 76, 2005.

[99] E. Anderås, L. Arapan, I. Katardjiev, and V. Yantchev, "Thin film plate wave resonant sensor for pressure and gravimetric measurements," *Procedia Engineering*, vol. 25, pp. 571–574, 2011.

[100] S. Tadigadapa and K. Mateti, "Piezoelectric mems sensors: state-of-the-art and perspectives," *Meas. Sci. Technol.*, vol. 20, no. 9, p. 092001, 2009.

[101] A. Abdollahi, Z. Jiang, and S. A. Arabshahi, "Evaluation on mass sensitivity of saw sensors for different piezoelectric materials using finite element analysis," *IEEE International Symposium on Applications of Ferroelectrics*, pp. 734–737, 2007.

[102] I. J. Forbes and K. P. Kirkbride, "The Origin of Alkenes in Illicit Amphetamine: An Examination of the Illicit Synthesis of Phenyl-2-Propanone," *Journal of Forensic Sciences*, vol. 37, no. 5, pp. 1311–1318, 1992.

[103] N. Kostesha, T. S. Alstrøm, C. Johnsen, K. A. Nielsen, J. O. Jeppesen, J. Larsen, A. Boisen, and M. H. Jakobsen, "Multi-colorimetric sensor array for detection of explosives in gas and liquid phase," *Proceedings Volume 8018, Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XII*, vol. 8018, pp. 80 181H–1 – 80 181H–12, 2011.

[104] L. L. Mølgaard, O. T. Buus, J. Larsen, H. Babamoradi, I. L. Thygesen, M. Laustsen, J. K. Munk, E. Dossi, C. O'Keeffe, L. Lässig, S. Tatlow, L. Sandström, and M. H. Jakobsen, "Improved detection of chemical substances from colorimetric sensor data using probabilistic machine learning," *Proc. SPIE 10183, Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XVIII*, vol. 10183, p. 1018307, 2017.

Appendices

The papers with me as first author are reproduced in the following along with the detailed process flows for each CMUT generation.

APPENDIX A

Paper - Sensitivity optimization of wafer bonded gravimetric CMUT sensors

This paper is submitted to: Journal of Microelectromechanical Systems (IEEE) [54].

Sensitivity optimization of wafer bonded gravimetric CMUT sensors

Mathias J. G. Mølgaard, Jesper M. F. Hansen, Mogens H. Jakobsen, and Erik V. Thomsen,

***Abstract*—Optimization of the mass sensitivity of wafer bonded resonant gravimetric Capacitive Micromachined Ultrasonic Transducers (CMUTs) is presented. Gas phase sensors based on resonant gravimetric CMUTs have previously been demonstrated. An important figure of merit of these sensors is the sensitivity which is increased by decreasing the radius of the CMUT cell. This paper investigates how to minimize the radius of CMUT cells fabricated using the wafer bonding process. The design and process parameters affecting the radius of the CMUT and hereby the sensitivity are studied through numerical simulations and Atomic Force Microscopy (AFM) measurements. An excellent fit was obtained between the simulations and measured profiles with a low relative error of $\leq 5\%$, thus validating the simulation model. Two types of CMUTs are designed and fabricated using the design and process rules determined herein, with experimentally determined mass sensitivities of $0.46\,\mathrm{Hz/ag}$ and $0.44\,\mathrm{Hz/ag}$, respectively. The two CMUT devices have cavities made using the Local Oxidation of Silicon (LOCOS) and Reactive Ion Etching (RIE) process. For the LOCOS process it was found that the smallest radius can be obtained by choosing a $\mathrm{Si_3N_4}$ oxidation mask and lowering the pad $\mathrm{SiO_2}$ thickness, vacuum gap height and Si bump height. For the RIE process the vertical dimensions do not influence the horizontal dimensions and consequently equivalent rules do not exist.**

***Index Terms*—CMUT, sensor, gravimetry, sensitivity, LOCOS.**

I. Introduction

Detection of chemical and biological analytes in the gas phase is important in several fields including homeland security, environmental monitoring, and in different branches of industry. Many different types of sensors exist but it has been demonstrated that Capacitive Micromachined Ultrasonic Transducers (CMUTs), used as resonant gravimetric mass sensors, can achieve a low Limit of Detection (LOD) [1] in the ag range [2] and a high mass sensitivity of up to $0.23\,\mathrm{Hz/ag}$ [3]. These properties have e.g. been utilized to detect small concentrations of dimethyl methylphosphonate (DMMP) [2] and greenhouse gasses such as CO_2 [4]. Furthermore, the flat and closed surface of the CMUT eases functionalization, all resulting in a desirable combination of features. Since the first CMUT was reported in 1994 [5] one of the main applications, for both researchers and companies, has been medical ultrasonic imaging [6–9]. Since then CMUTs have also been applied as e.g. biological or chemical sensors, typically based on the resonant gravimetric principle where a small change in the mass of the plate, due to an absorbed analyte, results in a resonance frequency shift [4], [10], [11].

The authors are with the Department of Micro and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark. Corresponding author: Mathias J. G. Mølgaard, email: migmol@nanotech.dtu.dk

In general, two fabrication methods have been used for fabricating CMUTs: a process based on sacrificial release and a process based on wafer bonding. The first CMUTs were fabricated using the sacrificial release process [5], that consists of depositing a sacrificial layer, which is selectively etched by an isotropic wet etch, after it has been covered by a plate layer, hereby forming the cavity of the device. In the wafer bonding process, first proposed by Huang et al. [12], cavities are defined in an insulating layer on a substrate wafer that is directly bonded to the plate.

In this paper the wafer bonding fabrication method is used where the cavities typically are defined by either one of two processes, namely the Reactive Ion Etch (RIE) process and Local Oxidation of Silicon (LOCOS) process [13]. In the RIE process cavities are dry etched in an insulating layer (typically SiO_2), while in the LOCOS process two consecutive LOCOS steps result in cavities with a central silicon bump. A cross-sectional sketch of half a LOCOS and RIE defined cavity can be seen in Fig. 1, along with all relevant geometrical variables. The characteristic bird's beak structure is also shown in Fig. 1 a). The RIE and LOCOS process differ on several points: the number of photolithography masks required is one less for the typical RIE process, thus decreasing the process cost and time. For the LOCOS device the vacuum gap height is decoupled from the post SiO_2 height and small gaps can therefore be made while still maintaining a thick post SiO_2, which decreases the parasitic capacitance between the top and bottom electrode. Hence, the electro-mechanical coupling coefficient is generally higher for the LOCOS device compared with a RIE device.

The limit of detection and mass sensitivity are the two most important figure of merits for any resonant gravimetric sensor, which in the ideal case should be as low and as high as possible, respectively. In this article the focus will be on maximizing the mass sensitivity of the CMUT sensor. The CMUT can be modeled as a 1-D linear harmonic oscillator and if the added mass on the plate is small compared with the mass of the plate itself, the sensitivity can be written as [10]:

$$S = \frac{\partial f}{\partial m} = -\frac{1}{2}\frac{f_{\mathrm{res}}}{m_{\mathrm{plate}}} \propto \frac{1}{a^4}, \tag{1}$$

where f_{res} is the resonance frequency, m_{plate} is the mass of the plate and a is the cell radius. The variable dependency is found by inserting the symbolic expressions for the resonance frequency and mass of a circular clamped plate. Therefore, in order to increase the sensitivity the radius must be decreased. For LOCOS cavities the smaller radii can become comparable with the lateral width of the bird's beak SiO_2 which in Fig. 1

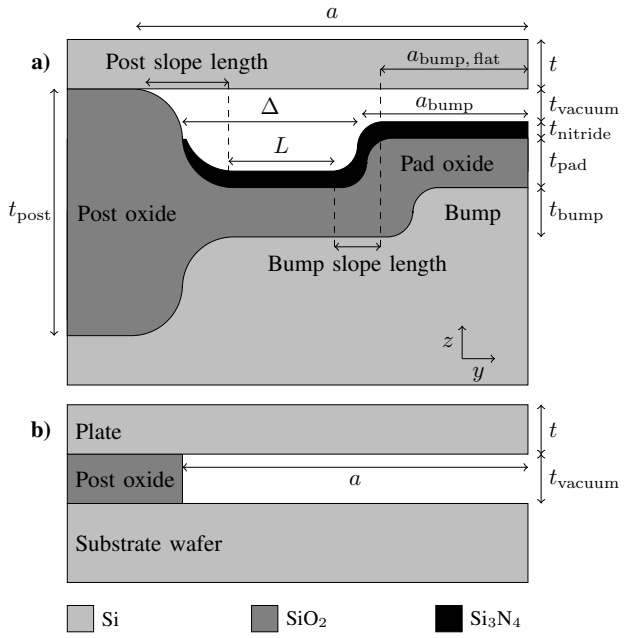


Fig. 1. Cross-sectional sketches of half a) a LOCOS CMUT cell and b) a RIE CMUT cell.

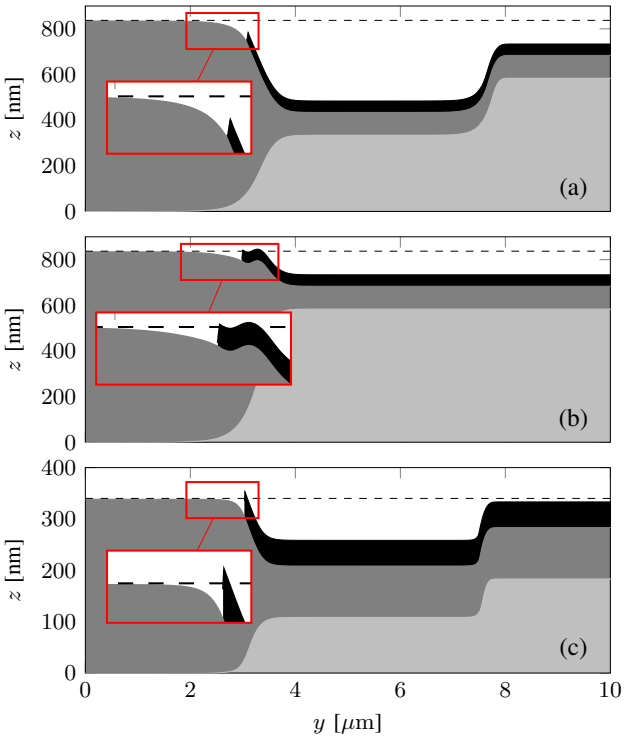


Fig. 2. Simulation results of three different CMUT designs. **(a)** This design does not inhibit wafer bonding. The designs in **(b)** and **(c)** both inhibit wafer bonding since the Si_3N_4 layer protrudes above the post SiO_2 at $y \approx 3\,\mu$m. In **(b)** the bump is placed too close to the post SiO_2 and in **(c)** the post SiO_2 is too low.

a) is defined as the 'slope length'. This ultimately limits the radius of the cavity and the sensitivity can hereby be limited by the process parameters that affect the slope length of the bird's beak. The bird's beak structure have previously been studied extensively due to its use as an isolation structure in semiconductor manufacturing [14–16], since reduction of the horizontal size has been especially important due to the decrease in device dimensions. However, for small radii ($a < 5\,\mu$m) CMUT cells the more complex geometry has yet to be investigated.

Not all CMUT designs can be fabricated. The main reason being the inability to bond the plate to the cavity wafer. For the RIE process it has previously been shown by Sarioglu et al. [17] how oxidation of convex silicon corners can lead to protrusions in the SiO_2 at the corners which can hinder bonding. A study by Christiansen et al. [18] showed how the corners of a structured SiO_2 layer are lifted when the wafer is re-oxidized in order to e.g. grow an insulation SiO_2 layer, consequently making bonding impossible. Likewise, some LOCOS designs will hinder bonding but here the problem is slightly different as the solution depends on choosing the right combination of design and process parameters, where for the RIE process the problems are solved by adjusting the process flow. The main challenge, when designing the LOCOS device, from a fabrication point of view, is keeping the Si_3N_4 layer from protruding above the post SiO_2 at the cavity edge, thus inhibiting bonding. Fig. 2 shows three simulated LOCOS devices where (a) is an example of a device that can be fabricated, (b) cannot be fabricated due to the protruding Si_3N_4 layer caused by an overlapping post SiO_2 and silicon bump, either as a result of a bad design choice or misalignment in the lithography step. Finally, in situation (c) the post SiO_2 is too low, which e.g. could happen if a very small gap is wanted. The simulation model is discussed and validated later.

In this paper the design of high mass sensitivity wafer bonded CMUTs is studied by means of AFM and numerical simulations. In addition, LOCOS and RIE CMUTs are fabricated and the feasibility of using a Si_3N_4 mask for both LOCOS steps is demonstrated. Furthermore, it is studied how the choice of oxidation masking material and process parameters influences the LOCOS device and hereby ultimately the minimum radius one can achieve and thus the maximum mass sensitivity. Finally, the experimental sensitivities of the fabricated LOCOS and RIE devices are compared with the theoretical values.

II. Methods

Standard cleanroom fabrication techniques were used to fabricate LOCOS profiles for experimental verification of the numerical simulation model, as well as LOCOS and RIE CMUT devices.

A. Simulation Model

A numerical simulation model was made using the process simulator ATHENA (Silvaco, Inc., California) where part of the LOCOS process in Fig. 3 a) was simulated. The mesh was optimized so the smallest mesh sizes were centered around the position where the bird's beak is formed. The size of the mesh in the SiO_2 has a maximum vertical length of 5 nm and

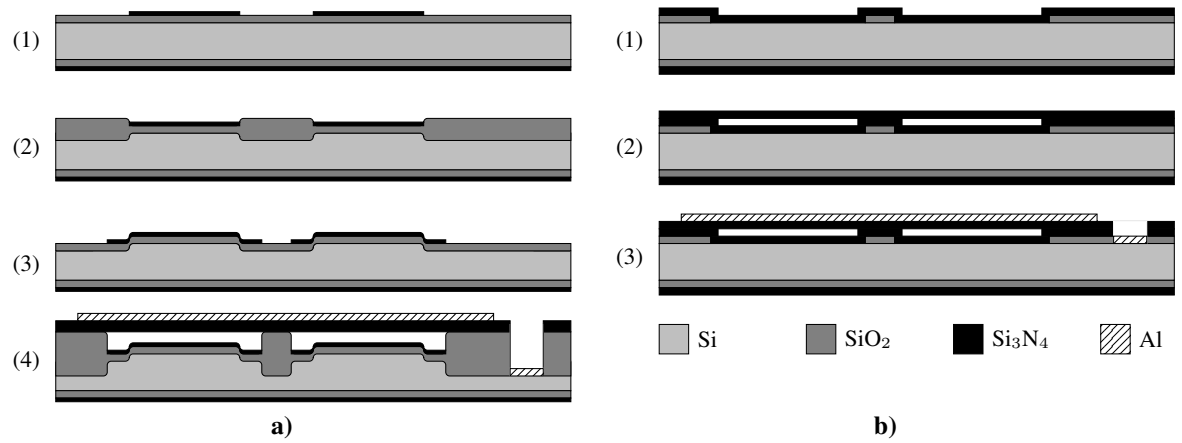


Fig. 3. Process flow of the a) LOCOS and b) RIE CMUT. In a) step (1) a thermal pad SiO_2 layer is grown and subsequently a LPCVD Si_3N_4 layer is deposited and patterned by a RIE etch. In step (2) the LOCOS is performed, forming the Si bumps. Step (3) is a repeat of step (1). Finally, in step (4) the second LOCOS is performed and the cavities are seal under vacuum by fusion bonding a Si_3N_4 layer to the cavity wafer. Openings to the bottom electrode are made and Al is deposited and wet etched. In b) step (1) a thermal SiO_2 is grown and patterned by a RIE etch. Next a LPCVD Si_3N_4 layer is deposited and the subsequent steps are the same as in the LOCOS process.

a maximum horizontal length of $25\,\mathrm{nm}$. These mesh sizes are chosen based on a mesh convergence study.

In the following the fabricated LOCOS structures used to verify the simulation model are described. Specifically, square silicon bumps were fabricated using the LOCOS process following steps (1)-(2) in Fig. 3 a), using both SiO_2 and Si_3N_4 as oxidation mask material. Table I shows an overview of the seven wafers oxidized in the experiment. The pad SiO_2 thickness and thickness of the SiO_2 mask layer were varied resulting in Si bumps of approximately the same height but with differing profiles. The wafers used were (100) oriented single side polished $4''$ Si wafers with an electrical resistivity of 1-10 Ωcm. All thermal oxidations for the bump fabrication and simulation were performed in a wet atmosphere at a temperature of $1100\,^{\circ}\mathrm{C}$. Stoichiometric Si_3N_4 was deposited using a Low Pressure Vapor Deposition (LPCVD), resulting in a $50\,\mathrm{nm}$ thick layer on all the wafers with Si_3N_4 masks. Finally, all wafers underwent a LOCOS step with different oxidation durations, in order to reach a bump height of $250\,\mathrm{nm}$. Subsequently, all layers were stripped from the wafer, leaving only the Si surface and the bare silicon bumps were characterized by AFM (Dimension Icon, Bruker). The maximum relative difference between the target and measured Si bump height was found to be 6.4%, which is low enough to allow for a comparison of the profiles.

Fig. 4 shows an example of an experimental and simulated Si bump profile, for bumps fabricated using Si_3N_4 masks and SiO_2 masks. The measured and simulated profiles have been translated relative to each other along the y-axis to the position of minimum relative error. The relative errors between the measured and simulated profiles were found to be $\leq 5\%$. This consistently low error for both the Si_3N_4 and SiO_2 masks demonstrates the validity of the simulation model. Fig. 4 also shows an example of the slope length, shown in Fig. 1, of one of the profiles. The slope length is defined as the horizontal distance between the two points were the slope first becomes $< 1\%$, that is: where the profile becomes flat. The 2D AFM

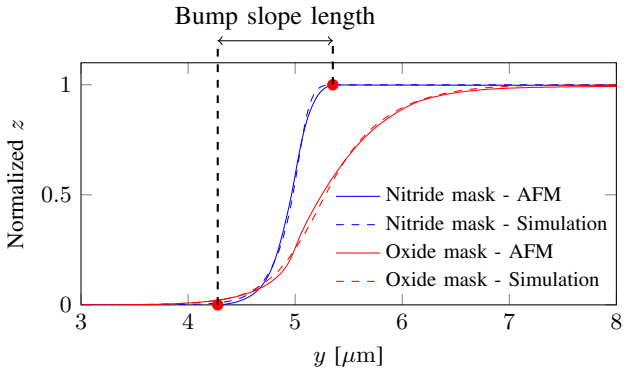


Fig. 4. Silicon bump profiles made using either an SiO_2 mask (thickness: $1\,\mu\mathrm{m}$) or a Si_3N_4 mask with a pad SiO_2 thickness of $t_{\mathrm{pad}} = 10\,\mathrm{nm}$. AFM measurements and simulation results are shown together. The distance between the two points, located at the positions where the slope is $< 1\%$, is denoted the slope length.

profiles are calculated by taking the mean of the scan lines in the direction orthogonal to the y-plane. The AFM scans consist of 512 scan lines parallel to the y-axis and each scan line is made up of 512 data points. The scan area is a $10\,\mu\mathrm{m}$ x $10\,\mu\mathrm{m}$ square and no artifacts were observed due to the relatively slowly varying slope of the profiles. The tilt of the scan is removed by fitting and subtracting a first order polynomial plane from the scan.

B. CMUT Fabrication Processes

In this section the fabrication processes for the LOCOS device (Fig. 3 a)) and RIE device (Fig. 3 b)) are presented. In the original LOCOS process by Park et al. [13] the first oxidation mask was a thermally grown SiO_2 layer, while in this process it is a Si_3N_4 layer. The choice of masking material becomes increasingly important as the dimensions of the CMUTs decrease as will be explained in more detail later. We used a Si_3N_4 layer as the plate, thus eliminating the need

TABLE I
TABLE OF TARGET AND MEASURED VALUES FOR THE SEVEN WAFERS FABRICATED FOR THE EXPERIMENTAL VALIDATION OF THE SIMULATION MODEL.

Parameter	**Wafer number**							
	1	2	3	4	5	6	7	Unit
t_{pad} target	0	10	50	100	-	-	-	[nm]
t_{pad} measured	0	10.6	56	99.4	-	-	-	[nm]
Mask SiO_2 thickness target	-	-	-	-	1000	2000	3000	[nm]
Mask SiO_2 thickness measured	-	-	-	-	963	1964	2950	[nm]
t_{bump} target	250	250	250	250	250	250	250	[nm]
t_{bump} measured	244	241	241	241	234	247	260	[nm]
Relative difference between target and measured t_{bump}	2.4	3.6	3.6	3.6	6.4	1.2	4	[%]

for a Silicon On Insulator (SOI) wafer, that are both more expensive and not easily attainable for thin ($< 200\,\mathrm{nm}$) device layers for $4''$ wafers. The silicon substrate wafers have a low electrical resistivity ($< 0.025\,\Omega\mathrm{cm}$, (100)) in order to decrease the electrical resistance in the bottom electrode.

In the LOCOS process in Fig. 3 a) step (1) a $10\,\mathrm{nm}$ thermal pad SiO_2 layer is grown by dry oxidation at $900\,^{\circ}\mathrm{C}$ and $50\,\mathrm{nm}$ Si_3N_4 is deposited by LPCVD. The pad SiO_2 acts as a buffer layer between the tensile stressed Si_3N_4 layer and the silicon surface. Subsequently, the Si_3N_4 layer is patterned by RIE dry etching using an UV photoresist mask. In step (2) the first LOCOS is performed by wet oxidation at $1100\,^{\circ}\mathrm{C}$ to form the Si bumps and afterwards all layers but the silicon are stripped by a buffered HF (BHF) etch and a $160\,^{\circ}\mathrm{C}$ H_3PO_4 etch. Another pad SiO_2 is grown (dry, $900\,^{\circ}\mathrm{C}$) in step (3) and Si_3N_4 is deposited using LPCVD and wet etched ($160\,^{\circ}\mathrm{C}$, H_3PO_4) using a poly-silicon etching mask. The second LOCOS (wet, $1100\,^{\circ}\mathrm{C}$) is performed in step (4) hereby forming the cavities. The substrate wafer is bonded, under vacuum, to a silicon wafer with a $137\,\mathrm{nm}$ Si_3N_4 layer constituting the plate. The bonded wafer stack is annealed at $1150\,^{\circ}\mathrm{C}$ and the silicon handle wafer is selectively etched by KOH ($28\,\mathrm{wt}\%$, $80\,^{\circ}\mathrm{C}$) using the Si_3N_4 plate layer as an etch stop layer. Furthermore, openings to the bottom electrode are dry etched and $100\,\mathrm{nm}$ Al is deposited by e-beam deposition and subsequently etched in a solution of $H_2O : H_3PO_4$ (1:2) at room temperature.

In the RIE process in Fig. 3 b) step (1) a thermal $225\,\mathrm{nm}$ SiO_2 layer is grown (wet, $1050\,^{\circ}\mathrm{C}$) and patterned by dry etching, forming cavities. A $27\,\mathrm{nm}$ Si_3N_4 layer is deposited using LPCVD, thus covering both sides of the wafer. In step (2) the cavity wafer is bonded to a silicon wafer with a $50\,\mathrm{nm}$ Si_3N_4 layer. All subsequent process steps are identical to the steps described for the LOCOS process, except that the Al layer thickness in this case was $50\,\mathrm{nm}$.

The structures are characterized in Section III-B where the dimensions are measured using Scanning Electron Microscopy (SEM) and the mass sensitivity is determined using impedance spectroscopy.

III. RESULTS AND DISCUSSION

The mass sensitivity was in Eq. 1 shown to be dependent on the radius of the CMUT cell, where smaller radii result in higher sensitivities.

The sensitivity of a RIE CMUT is trivial to calculate as the lateral dimensions are not coupled to the vertical dimensions due to the dry etched cavities. That is, no matter the post SiO_2 height the radius will stay the same and consequently so will the sensitivity. The minimum RIE cavity radius is therefore only limited by the minimum feature size of the lithography process.

For the LOCOS device the bird's beak profile results in a coupling between the vertical and lateral dimensions which ultimately limits the radius and hereby the sensitivity. In the following section the parameters influencing the LOCOS profile and thus sensitivity are investigated through AFM measurements and simulations. Furthermore, the limited design space is investigated due to the fabrication limitations exemplified in Fig. 2. Finally, the fabricated CMUT devices are experimentally characterized.

A. LOCOS cavity optimization

To minimize the radius of the LOCOS CMUT cell and the Si bump radius, the slope length (Fig. 1 and 4) should be minimized. Fig. 5 shows that the slope length is approximately 3 times shorter for thin pad oxides ($t_{\mathrm{pad}} = 0\,\mathrm{nm}$ to $10\,\mathrm{nm}$) as compared with a SiO_2 mask to 2 times shorter for thicker pad oxides ($t_{\mathrm{pad}} = 100\,\mathrm{nm}$). In addition, Fig. 5(a) shows that the slope length is roughly constant as a function of the thickness of the SiO_2 mask. Whereas, Fig. 5(b) demonstrates an increasing slope length for a thicker pad SiO_2, since the diffusivity of the oxidizing species (O_2 or H_2O) is higher in SiO_2 than Si_3N_4 and a thicker pad SiO_2 layer allows for a higher lateral influx of H_2O under the mask. Likewise, increased lateral diffusion is the cause for the longer slope lengths when SiO_2 is used as the masking material compared with a Si_3N_4 mask. This agrees well with what is found in the literature [14–16]. The measured and simulated slope lengths are in agreement with each other, showing a maximum relative error of $\leq 10\%$ and demonstrating the same scaling tendencies.

When the radius of the CMUT cell decreases, the radius of the silicon bump must decrease as well and consequently the slope length of the bump will make up an increasing fraction of the total bump radius. As a result, the expected capacitance of the cell is decreased, the expected pull-in voltage is increased and wafer bonding can be rendered impossible if the slope of the bump overlaps too much with the cavity edge, as was the case in Fig. 2(b). Thus, it is important to control the bump slope length and take it into consideration when designing the

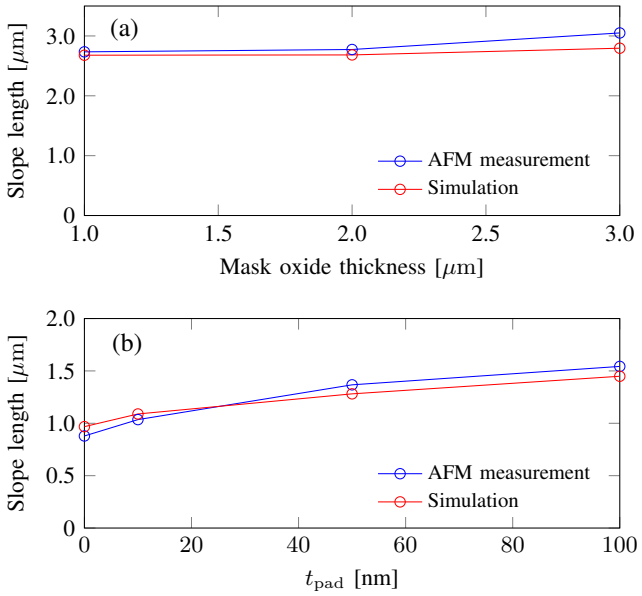


Fig. 5. (a) Slope length as a function of the SiO_2 oxidation mask thickness. (b) Slope length as a function of the pad SiO_2 thickness, under the Si_3N_4 oxidation mask. The simulations and measurements agree with a maximum relative difference of 10 %.

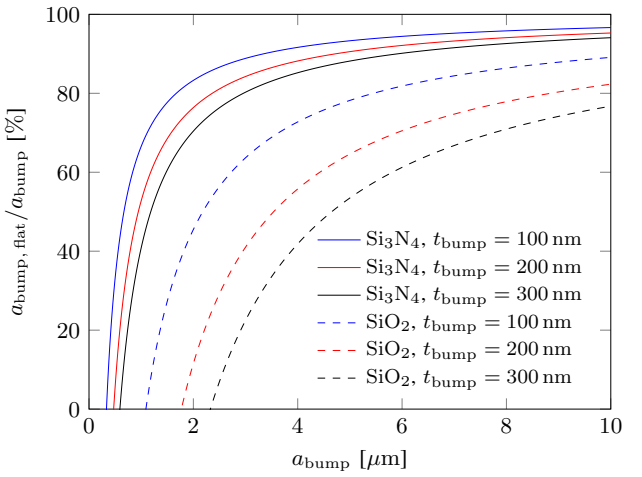


Fig. 6. Simulation results: $a_{\text{bump, flat}}$ (see Fig. 1) normalized to a_{bump} as a function of a_{bump} for Si_3N_4 ($t_{\text{pad}} = 10$ nm) and SiO_2 (mask height: 2 μm) as masking material at three bump heights.

CMUT. In most cases it is desirable to minimize the slope length, hereby creating the most square-like corners for the bump. Fig. 6 shows the ratio of the flat part of the bump to the bump radius as a function of the bump radius for different oxidation masking materials and bump heights. As the bump radius increases the ratio approaches 100%. The Si_3N_4 mask results in more well defined bumps with square-like corners where a larger fraction of the bump reaches the designed height. Furthermore, higher bumps lead to lower ratios and this difference is relatively larger for the SiO_2 mask. All in all this favors using a Si_3N_4 mask.

When designing bumps with radii in the sub $\sim 5\,\mu$m range

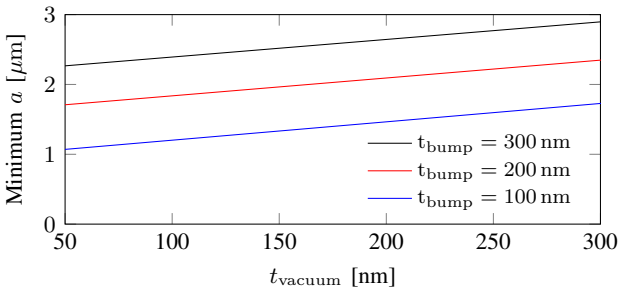


Fig. 7. Minimum cell radius as a function of the vacuum gap height for three bump heights, when $L = 0$ nm and $a_{\text{bump,flat}} = 0$ nm.

it can therefore be advantageous to use Si_3N_4 as the masking material. In the article by Park et al. [13] the masking material for the first LOCOS process was SiO_2 but it was noted that Si_3N_4 could have been used at the cost of additional process steps. Therefore, the choice of LOCOS masking material is a trade-off between tight dimension control and reducing the number of process steps.

What is the minimum radius and hence sensitivity one can achieve with a LOCOS CMUT cell? In order to answer this question two limitations are imposed on the structures: the bump and the post SiO_2 should not overlap, that is $L \geq 0$ nm (see Fig. 1) and the bump should at least reach the designed height in the center, that is $a_{\text{bump,flat}} \geq 0$ nm. The minimum radius is obtained when both of these variables are zero. Fig. 7 shows a plot of the minimum radius as a function of the vacuum gap height for three bump heights. As the vacuum gap height is increased the minimum radius increases since the post SiO_2 height increases, thus making the slope length of the post SiO_2 longer. The same effect is seen when the bump height is increased, resulting in larger minimum cavity radii. Eq. 1 shows that these minimum radii given by Fig. 7 for the LOCOS structure, directly determines the maximum sensitivity.

In order to map out the design space for LOCOS cavities, simulations were made where the post SiO_2 height and bump height have been varied. Fig. 8 shows a plot where the contour lines are the flat distance, L (see Fig. 1), between the bump and the post SiO_2 as a function of t_{post} and t_{bump}. The plot is valid for a specific set of parameters, given in the figure caption, but the overall shape of the plot is general for all LOCOS designs. The parameter Δ is the designed distance on the photolithographic mask between the bump and cavity edge and in this plot it is $\Delta = 1.5\,\mu$m. The three colored regions (a), (b) and (c) correspond to the three situations in Fig. 2. Region (a) is limited by the red line that denotes the minimum post SiO_2 height, the blue line at which the vacuum gap is zero and the $L = 0$-contour line which is the limit where the bump and post SiO_2 start to overlap. Thus, staying inside the green region (a) ensures a LOCOS design that can be successfully fabricated.

Fig. 9 shows that as the bump to cavity edge distance, Δ increases, so does the design space, here shown as green areas. The solid lines are where $L = 0$ nm, here plotted for several values of Δ. The smallest Δs are typically found

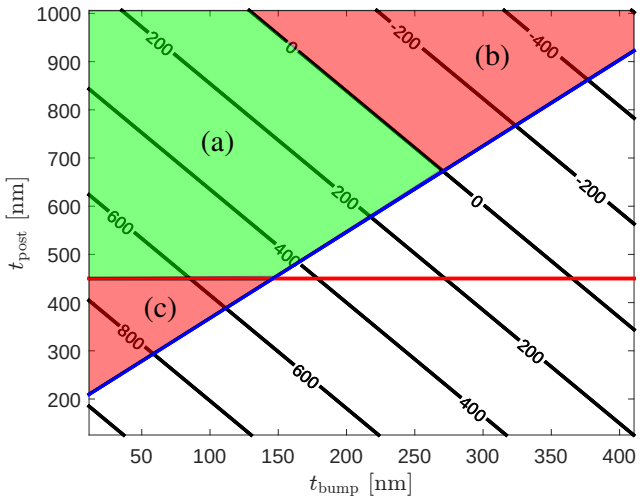


Fig. 8. Contour plot showing the length of the flat region between the cavity and bump L (see Fig. 1) as a function of the SiO_2 post thickness t_{post} and bump height t_{bump}. The red line gives the minimum required post SiO_2 height for successful bonding and the blue line are the points at which the vacuum gap height is 0 nm. The distance between the cavity and bump: $\Delta = 1.5\,\mu$m, bump and cavity Si_3N_4 thickness $t_{\mathrm{nitride}} = 50$ nm and bump and cavity pad SiO_2 thicknesses of $t_{\mathrm{pad,\,bump}} = 10$ nm and $t_{\mathrm{pad,\,cavity}} = 100$ nm, respectively.

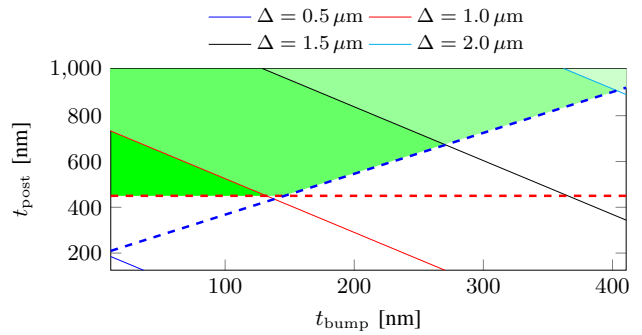


Fig. 9. The solid lines show where $L = 0$ nm and are plotted for different values of Δ as a function of t_{bump} and t_{post}. The red dashed line gives the minimum required post SiO_2 height for successful bonding and the blue dashed line are the points at which the vacuum gap height is 0 nm. Si_3N_4 thickness $t_{\mathrm{nitride}} = 50$ nm and bump and cavity pad SiO_2 thicknesses of $t_{\mathrm{pad,\,bump}} = 10$ nm and $t_{\mathrm{pad,\,cavity}} = 100$ nm, respectively.

for small radii cells which are here seen to be the most limited in their design space. Therefore, these design rules are especially important for these smaller radius CMUT cells, which are typical for CMUTs used for sensing, when a high mass sensitivity is wanted.

In conclusion, all variables that affect the LOCOS profile will ultimately affect the sensitivity, capacitance and pull-in voltage. Hence, being able to predict the effect on the final fabricated structure is important. The shortest slope length for the Si bump and post SiO_2 and hereby the smallest cell radii is obtained by using a Si_3N_4 mask with a thin pad SiO_2 (t_{pad}) and choosing a low Si bump height and a small vacuum gap, since this yields the lowest post SiO_2 height. The pad SiO_2 thickness for the first LOCOS step can be made thin as these layers are subsequently stripped. However, the

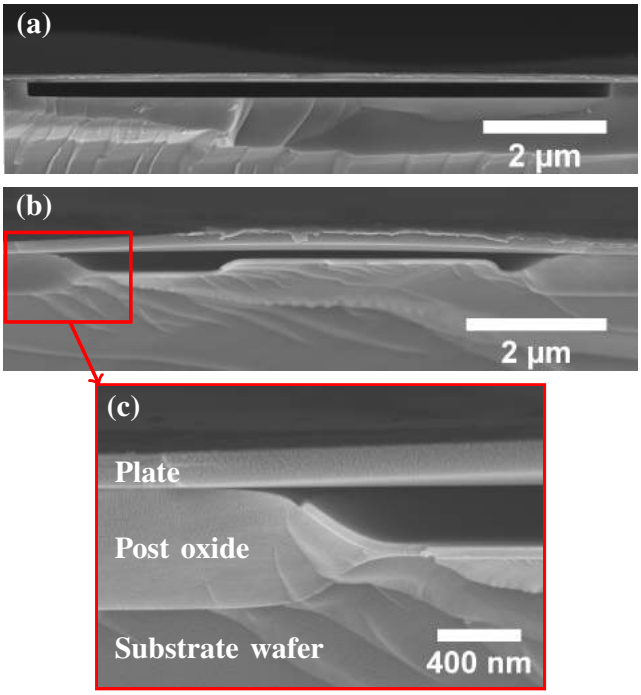


Fig. 10. Scanning electron microscope images of cross-sections of a single CMUT cell for (a) the RIE device, (b) the LOCOS device and (c) zoom in at the bird's beak.

pad SiO_2 thickness for the second LOCOS step is sometimes determined by the requirement for the CMUT cell to prevent electrical breakdown if pull-in occurs. Finally, the design space of LOCOS cavities were investigated and it was shown that for high sensitivity, small radii cavities the design space is limited.

B. CMUT sensors

Two CMUT devices have been fabricated: a LOCOS and a RIE device, following the process flows in Fig. 3. The design of the LOCOS CMUT utilized the findings above to minimize the slope lengths of the bump and post SiO_2. Namely, using Si_3N_4 as masking material in both LOCOS steps and having a pad SiO_2 thickness of only 10 nm. Furthermore, the vacuum gap is just 65 nm, resulting in a relatively low post SiO_2 height, hereby shortening the post SiO_2 slope length further. The RIE device has an insulation Si_3N_4 layer which both increases the breakdown voltage when in pull-in and prevents potential leakage currents from running during operation. The choice of a Si_3N_4 plate for this device results in a wafer bonding interface between two Si_3N_4 surfaces; two materials that empirically are more difficult to bond than e.g. Si-SiO_2 or Si_3N_4-SiO_2. The bonding strength is increased between the two Si_3N_4 surfaces by oxidizing the Si_3N_4 surfaces at a high temperature in a wet atmosphere, hereby creating a thin layer of oxy-Si_3N_4 [19].

Two SEM cross-sections of the devices can be seen in Fig. 10. Fig. 10(a) shows a RIE cell with a well-defined vertical post SiO_2 cavity edge. Fig. 10(b) shows the LOCOS cell with the central Si bump and (c) is a zoom in on the bird's beak

TABLE II
MEASURED DIMENSIONS OF THE LOCOS AND RIE DEVICES. THE VERTICAL DIMENSIONS ARE OBTAINED FROM THE SEM IMAGES. ALL VALUES ARE IN nm.

Dimension	**LOCOS**	**RIE**
Plate thickness	137(no Al)	105(with Al)
Vacuum gap height	65	235
Post SiO_2 thickness	582	235
Bump height	180	-

profile at the cavity edge. The Si_3N_4 layer is seen to end well below the bonding surface. The cracked surface is the result of a sputtered Au layer which was applied in order to prevent charge build up during imaging. The dimensions of the two cells are shown in Table II, where the thickness of the Au layer has been subtracted from the measured plate thickness. The Si bump is misaligned relative to the cavity and is therefore not placed completely centered in the cavity which highlights one of the challenges of making the cavities smaller, namely the increasing alignment tolerances. Indeed, if the Si bump is misaligned so much that it overlaps with the post SiO_2, bonding can be hindered which is exemplified in Fig. 2(b). Comparing the LOCOS and RIE device it is evident how the sloped part of the LOCOS device makes up a significant part of the total cavity area while the RIE device in unaffected by this.

The sensitivity of the devices can be calculated using Eq. 1 but the sensitivity can also be measured directly by measuring the resonance frequency shift when a known mass is added to the plate. In order for the CMUT to stay in the linear regime the inequality $m_{added} \ll m_{plate}$ must not be violated. The sensitivity was measured by depositing thin ($< 10\,nm$) layers of Au directly on the plate. The resonance frequency was found from impedance spectra measured after each deposition. The added mass was determined by measuring the step height of the Au layer using an AFM and calculating the mass based on the area of the plate and Au density. A plot of the resonance frequency as a function of mass added to the plate is given in Fig. 11 for both the LOCOS and RIE device. Eq. 1 predicts that these plots should be straight lines with a slope equal to the sensitivity. The lines in Fig. 11 are indeed seen to be straight and Table III lists the analytic and experimental sensitivities for the two devices. For both the LOCOS and RIE device the experimental sensitivity is lower than the theoretical value which is due to the assumption being broken stating that the added mass should be much smaller than the mass of the plate. Table III states the fraction m_{Au}/m_{plate} for both devices, showing that the assumption is indeed broken. The relative difference between the theoretical and experimental sensitivity is greater for the RIE device than the LOCOS device since the relative added mass is larger. The decrease in sensitivity between the mass loaded and unloaded case is studied with a simple Finite Element Method (FEM) model and the relative differences for the LOCOS and RIE devices are seen to agree. Nonetheless, both the theoretical and experimental sensitivity values, for both devices, are to the best of our knowledge the highest published for CMUTs.

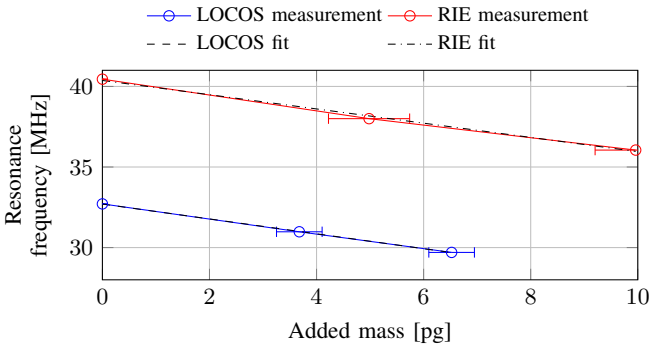


Fig. 11. Resonance frequency as a function of mass added on the plate for the LOCOS and RIE devices. The slopes of the two linear fits are the respective sensitivities. The error bars correspond to an estimated uncertainty of the Au thickness of $0.5\,nm$ from the measurement itself and spatial nonuniformity of the sputtering process.

TABLE III
THEORETICAL SENSITIVTY (S_{theo}), EXPERIMENTAL SENSITIVITY (S_{exp}), AND RELATIVE ADDED MASS ON THE LOCOS AND RIE DEVICES. IN ADDITION, S_{theo} AND S_{exp} ARE COMPARED ALONG WITH SENSITIVTIES CALCULATED IN A FEM MODEL.

	LOCOS	**RIE**	Unit
S_{theo}	0.74	0.84	Hz/ag
S_{exp}	0.46	0.44	Hz/ag
m_{Au}/m_{plate}	25	43	%
$1 - S_{exp}/S_{theo}$	38	48	%
$1 - S_{FEM,\,loaded}/S_{FEM,\,unloaded}$	28	48	%

IV. CONCLUSION

The minimization of wafer bonded CMUT cells was studied through a numerical process simulation model whose results were compared to AFM measurements of fabricated structures, giving a relative error of $\leq 5\%$. The smallest radius one can achieve using the LOCOS cavities is by using Si_3N_4 as the masking material which result in slope lengths that are 2 to 3 times shorter than using a SiO_2 mask. Further, the pad SiO_2 thickness should be decreased as well as the bump height. Lastly, it was demonstrated how the design space becomes increasingly limited for decreasing cavity dimensions.

CMUT senors were fabricated by the LOCOS and RIE processes and their mass sensitivities were both calculated and measured. The LOCOS CMUT device was designed using the findings from the simulation model while the RIE device was fabricated using Si_3N_4 to Si_3N_4 wafer bonding. The LOCOS and RIE devices both showed a high experimental mass sensitivity of $0.46\,Hz/ag$ and $0.44\,Hz/ag$, respectively, which for CMUTs, to our knowledge, are the highest mass sensitivities published.

REFERENCES

[1] S. Fanget, S. Hentz, P. Puget, J. Arcamone, M. Matheron, E. Colinet, P. Andreucci, L. Duraffourg, E. Myers, and M. Roukes, "Gas sensors based on gravimetric detection - A review," *Sensors and Actuators B: Chemical*, vol. 160, pp. 804 – 821, 2011.

[2] H. J. Lee, K. K. Park, M. Kupnik, Ö. Oralkan, and B. T. Khuri-Yakub, "Chemical vapor detection using a capacitive micromachined ultrasonic transducer," *Anal. Chem.*, vol. 83, p. 93149320, 2011.

[3] H. J. Lee, K. K. Park, Ö. Oralkan, M. Kupnik, and B. T. Khuri-Yakub, "A multichannel oscillator for a resonant chemical sensor system," *IEEE Transactions On Industrial Electronics*, vol. 61, no. 10, pp. 5632–5640, 2014.

[4] D. Barauskas, D. Pelenis, D. Virzonis, J. P. Baltrus, and J. Baltrusaitis, "Greenhouse gas molecule CO2 detection using a capacitive micromachined ultrasound transducer," *Anal. Chem.*, vol. 88, no. 13, pp. 6662–6665, 2016.

[5] M. I. Haller and B. T. Khuri-Yakub, "A surface micromachined electrostatic ultrasonic air transducer," *IEEE Ultrasonics Symposium*, pp. 1241–1244, 1994.

[6] F. L. Degertekin, R. O. Guldiken, and M. Karaman, "Annular-ring CMUT arrays for forward-looking IVUS: Transducer characterization and imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 53, no. 2, pp. 474–482, 2006.

[7] A. S. Savoia, G. Caliano, and M. Pappalardo, "A CMUT probe for medical ultrasonography: From microfabrication to system integration," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 59, no. 6, pp. 1127–1138, 2012.

[8] B. T. Khuri-Yakub and Ö. Oralkan, "Capacitive micromachined ultrasonic transducers for medical imaging and therapy," *J. Micromech. Microeng.*, vol. 21, no. 5, p. 054004, 2011.

[9] M. Engholm, H. Bouzari, T. L. Christiansen, C. Beers, J. P. Bagge, L. N. Moesner, S. E. Diederichsen, M. B. Stuart, J. A. Jensen, and E. V. Thomsen, "Probe development of CMUT and PZT row-column-addressed 2-D arrays," *Sensors and Actuators A: Physical*, vol. 273, p. 121133, 2018.

[10] K. K. Park and H. J. Lee and G. G. Yaralioglu and A. S. Ergun and Ö. Oralkan and M. Kupnik and C. F. Quate and B. T. Khuri-Yakub and T. Braun and J.-P. Ramseyer and H. P. Lang and M. Hegner and Ch. Gerber and J. K. Gimzewski, "Capacitive micromachined ultrasonic transducers for chemical detection in nitrogen," *Applied Physics Letters*, vol. 91, p. 094102, 2007.

[11] D. Virzonis, G. Vanagas, A. Ramanaviciene, A. Makaraviciute, D. Barauskas, A. Ramanavicius, W. Wen, and R. Kodzius, "Resonant gravimetric immunosensing based on capacitive micromachined ultrasound transducers," *Microchimica Acta*, vol. 181, pp. 1749–1757, 2014.

[12] Y. Huang, A. S. Ergun, E. Hæggström, M. H. Badi, and B. T. Khuri-Yakub, "Fabricating capacitive micromachined ultrasonic transducers with wafer-bonding technology," *J. Microelectromech. Syst.*, vol. 12, no. 2, p. 128137, 2003.

[13] K. K. Park, H. Lee, M. Kupnik, and B. T. Khuri-Yakub, "Fabrication of capacitive micromachined ultrasonic transducers via local oxidation and direct wafer bonding," *Journal of Microelectromechanical Systems*, vol. 20, no. 1, pp. 95–103, 2011.

[14] D. Chin, S.-Y. Oh, S.-M. Hu, R. W. Dutton, and J. L. Moll, "Two-dimenlsional oxidation," *IEEE Transactions on electron devices*, vol. 30, no. 7, pp. 744–749, 1983.

[15] T.-C. Wu, W. T. Stacy, and K. N. Ritz, "The influence of the LOCOS processing parameters on the shape of the Bird's Beak structure," *J. Electrochem. Soc.: Solid-state Science and Technology*, vol. 130, no. 7, pp. 1563–1566, 1983.

[16] N. Guillemot, G. Pananakakis, and P. Chenevier, "A new analytical model of the Birds Beak," *IEEE Transactions on electron devices*, vol. 34, no. 5, pp. 1033–1038, 1987.

[17] A. F. Sarioglu, M. Kupnik, S. Vaithilingam, and B. T. Khuri-Yakub, "Nanoscale topography of thermally-grown oxide films at right-angled convex corners of silicon," *Journal of The Electrochemical Society*, vol. 159, no. 2, pp. 79–84, 2012.

[18] T. L. Christiansen, O. Hansen, J. A. Jensen, and E. V. Thomsen, "Thermal oxidation of structured silicon dioxide," *ECS Journal of Solid State Science and Technology*, vol. 3, no. 5, pp. 63–68, 2014.

[19] K. Reck, C. Østergaard, E. V. Thomsen, and O. Hansen, "Fusion bonding of silicon nitride surfaces," *J. Micromech. Microeng.*, vol. 21, pp. 1–5, 2011.

[20] E. Benes, "Improved quartz crystal microbalance technique," *J. Appl. Phys.*, vol. 56, no. 3, pp. 608–626, 1984.

Mathias J. G. Mølgaard received his B.Sc. and M.Sc. degrees in engineering physics and nanotechnology from the Technical University of Denmark, Kgs. Lyngby, Denmark, in 2012 and 2015, respectively. He is currently pursuing a Ph.D. degree at the Technical University of Denmark, MEMS Applied Sensors group, focusing on gravimetric sensing using capacitve micromachined ultrasonic transducers.

Jesper M. F. Hansen received his B.Sc. and M.Sc. degrees in engineering physics and nanotechnology from the Technical University of Denmark, Kgs. Lyngby, Denmark, in 2015 and 2017, respectively. He is currently working at the company NIL technology.

Mogens H. Jakobsen is Associate Professor at DTU Nanotech, the Technical University of Denmark, where he is group leader of the Surface Engineering group. He received his M.Sc. degree and Ph.D. degree in Chemistry from University of Copenhagen, Denmark, in 1987 and 1990 respectively. In 1995 he left his job as Assistant Professor at University of Copenhagen to become CSO & VP R&D as co-founder of Exiqon A/S, Denmark. In 2003 he left Exiqon A/S for a position as Associate Professor at DTU Physics followed by a position as Senior Researcher at Cantion A/S, Demark. In 2006 he joined DTU Nanotech for his current position as Associate Professor.

Erik V. Thomsen is Professor at DTU Nanotech, the Technical University of Denmark, where he is heading the MEMS Applied Sensors group. He was born in Aarhus in 1964, received the M.Sc. degree in physics from Odense University, Odense, Denmark, and the Ph.D. degree in electrical engineering from the Technical University of Denmark, Lyngby, in 1998. His current research and teaching interests include MEMS multisensors, biomedicaldevices, small scale energy systems such as miniature fuel cells and energy harvesting devices, capacitive micromachined ultrasonic transducers, and piezoelectric MEMS. He teaches classes in solid state electronics, microtechnology, and nano- and microfabrication. Dr. Thomsen received the AEG Electron Prize in 1995 and has also received several teaching awards at DTU.

APPENDIX B

Paper - Combined Colorimetric and Gravimetric CMUT Sensor for Detection of Benzyl Methyl Ketone

This paper is submitted to: Sensors and Actuators B: Chemical, and is currently under review [53].

Combined Colorimetric and Gravimetric CMUT Sensor for Detection of Benzyl Methyl Ketone

Mathias J. G. Mølgaard*, Milan Laustsen*, Mogens H. Jakobsen*, Thomas L. Andresen*, and Erik V. Thomsen*
*Department of Micro and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

***Abstract*—The detection of benzyl methyl ketone (BMK) is of interest as it is a common precursor for the synthesis of (meth)amphetamine. Resonant gravimetric sensors can be used to detect the mass and hereby the concentration of a gas while colorimetric arrays typically have exceptional selectivities to a target analyte. We present a sensor system consisting of a Capacitive Micromachined Ultrasonic Transducer (CMUT) and a colorimetric array for detection of BMK. The CMUT is used as a resonant gravimetric gas sensor where the resonance frequency shift due to mass loading of the plate. A single Local Oxidation of Silicon (LOCOS) step was used to define the cavities which were sealed with a Si_3N_4 plate with a thickness of $100\,\mathrm{nm}$, resulting in a resonance frequency of $38.8\,\mathrm{MHz}$ and a mass sensitivity of $28.3\,\frac{\mathrm{zg}}{\mathrm{Hz}\cdot\mu\mathrm{m}^2}$ or $8.5\times10^3\,\frac{\mathrm{cm}^2}{\mathrm{g}}$. This sensitivity is, to our knowledge, the best reported as compared with the state of the art of CMUTs. The CMUTs were functionalized with the same dyes used to fabricate the colorimetric arrays. While both the CMUT and the colorimetric array showed selectivity to BMK, the best selectivity was achieved with the colorimetric array. Furthermore, the mass of the BMK was found as a function of time. Thus, the combination of the colorimetric array and the CMUT results in a good selectivity and a quantitative value for the mass.**

I. Introduction

Detection of illegal drug and explosive precursors is of interest for government agencies. Amphetamine and methamphetamine are recreational drugs and in most countries they are classified as controlled substances. A common synthesization route for these drugs involves the reductive amination of the illegal precursor molecule benzyl methyl ketone (BMK) [1]. Therefore, it is desirable to be able to detect BMK, especially in the gas phase.

The detection and identification of molecules in the gas phase can be done in a variety of ways, including: gas chromatography, infrared spectroscopy, nuclear magnetic resonance spectroscopy, mass spectrometry and by resonant gravimetric detection. Of these only the latter has the potential for miniaturization, low cost, low power consumption, and integration with integrated circuits, while still maintaining a high sensitivity and mass resolution [2]. Consequently, gravimetric sensors are strong candidates for mobile sensing systems or as distributed sensors. Examples of such sensors include: bulk acoustic wave (BAW) sensors: quartz crystal microbalances (QCMs) [3], and film bulk acoustic resonators (FBAR) [4], surface acoustic wave (SAW) sensors [5], micro- and nano-cantilevers/bridges [6], and capacitive micromachined ultrasonic transducers (CMUTs) [7]. The criteria for a good resonant gravimetric sensor for gas phase measurements are manifold: high mass sensitivity, low limit of detection (LOD), high signal to noise ratio (SNR), and ease of functionalization. Based on these criteria the CMUT is an attractive candidate.The normalized CMUT mass sensitivity is high compared with the other sensor types, as will be shown in Section V. Compared with the other gravimetric sensors the CMUT has a LOD per area which is among the lowest due to a combination of a low Allan deviation (10^{-8} to $3\cdot10^{-8}$) and a relatively large surface area ($0.09\,\mathrm{mm}^2$ to $0.25\,\mathrm{mm}^2$) [2] A single CMUT sensor is typically comprised of hundreds to thousands of CMUT cells connected electrically in parallel, lowering the overall impedance and noise [7], while increasing the SNR. The ease of massive parallelism of the CMUT cells is contrasted by the other gravimetric sensor types that often are comprised of a single resonating structure.

CMUTs have previously been used for sensing applications. In 2007 Park *et al.* demonstrated the use of CMUTs as sensors by detecting water, isopropanol, acetone, and methanol [8]. The sarin stimulant molecule dimethyl methylphosphonate (DMMP) has been detected by CMUTs with a volume LOD of $50.5\,\mathrm{pptv}$ and a volume sensitivity of $34.5\,\mathrm{pptv/Hz}$ [7]. These results show that unprecedented sensitivities and LODs can be obtained using CMUTs as gravimetric sensors. CMUTs have also been used as biosensors, specifically as an immunosensor where the antigen antibody pair interacts in a highly selective reaction causing the resonance frequency to decrease [9]. Finally, CMUTs have also been used for environmental monitoring where the concentration of CO_2 has been measured [10] [11].

A functional layer is needed in order to make the gravimetric sensors selective towards a specific analyte. Functionalization is typically done by applying a thin coating on top of the resonator. Due to the closed and planar surface of the CMUT device several methods can be employed including: spin coating, drop coating, dip coating, and spray coating. One method of improving selectivity between multiple analytes in a gas has been to employ several CMUT elements each coated with a different functionalization layer [12]. A single chip with several CMUT elements can easily be fabricated, thus minimizing the footprint and complexity.

In this paper we combine a CMUT sensor with a colorimetric array to detect BMK. Doing so, we benefit from the high

selectivity offered by the colorimetric array and the sensitivity and quantitative output of the CMUT. Furthermore, having two independent sensor types combined provides the total sensor system with redundancy. The article is structured as follows: first the fabrication and characterization of the CMUT and colorimetric array is presented. Then the experimental setup is described, followed by the results in the form of the CMUT and colorimetric response. Finally a short discussion and comparison of normalized sensitivities for different gravimetric resonators is given.

II. The Sensors

This section describes the fabrication and characterization of the CMUT and colorimetric chip.

A. CMUT

1) Fabrication: The CMUT process flow can be seen in Figure 1. In step (1) a SiO_2 layer was grown on a low electrical resistivity Si wafer ($< 0.025\,\Omega\mathrm{cm}$, $<100>$). A Si_3N_4 layer was deposited by LPCVD and subsequently patterned using UV lithography and a H_3PO_4 etch. The purpose of the SiO_2 layer is to separate the Si_3N_4 layer from the Si surface which otherwise could have defects introduced due to the tensile stress of $\sim 1.3\,\mathrm{GPa}$ in the Si_3N_4 layer. This value is found by measuring the curvature of a double side polished wafer with a layer of Si_3N_4 of known thickness on one side and then applying Stoney's equation. The tensile stress increases both the resonance frequency and pull-in voltage compared with a similar plate with no stress [13]. In step (2) local oxidation of silicon (LOCOS) was performed using the patterned Si_3N_4 as the oxidation mask. The cavities were sealed under vacuum in a wafer bonding process (step (3)) by bonding the cavity wafer to a Si wafer with a layer of Si_3N_4. In step (4) the top Si_3N_4 layer was etched in a dry etch and the Si in KOH, which was chosen due to the high etch selectivity between Si and Si_3N_4. Hence, the Si_3N_4 plate and the Si_3N_4 layer on the backside of the cavity wafer acted as etch stop layers. Bottom electrode openings were etched in step (5) with a dry etch and finally the Al top electrode was deposited and patterned in step (6) using a $H_3PO_4 : H_2O$ (2:1) etch at room temperature.

2) Design and characterization: The design parameters of the CMUT used for sensing is shown in Table I. Figure 2a) shows a top view optical microscope image of a CMUT element consisting of 100 individual cells. The dark circles are the individual CMUT cells, while the light-colored grid structure is the Al top electrode. The top electrode grid design is chosen to decrease the parasitic capacitance and the total mass of the plate and top electrode. A more narrow W decreases the parasitic capacitance but decreasing W increases the electrical resistance and ultimately W is limited by the lithography system used. The choice of W is therefore a trade-off between parasitic capacitance, sensitivity and electrical resistance. Figure 2b) shows a cross-sectional SEM image of a single CMUT cell. The figure shows the Bird's Beak formed by the growth of the SiO_2 at the edge of the cavity. Figure 2c) shows a zoom in which reveals that the Si_3N_4 edge is well below the bonding plane. The plate and metal layer is bending upwards, away from the substrate wafer. This is not expected since the stress in the plate is tensile but it may be an effect of the cleaving process.

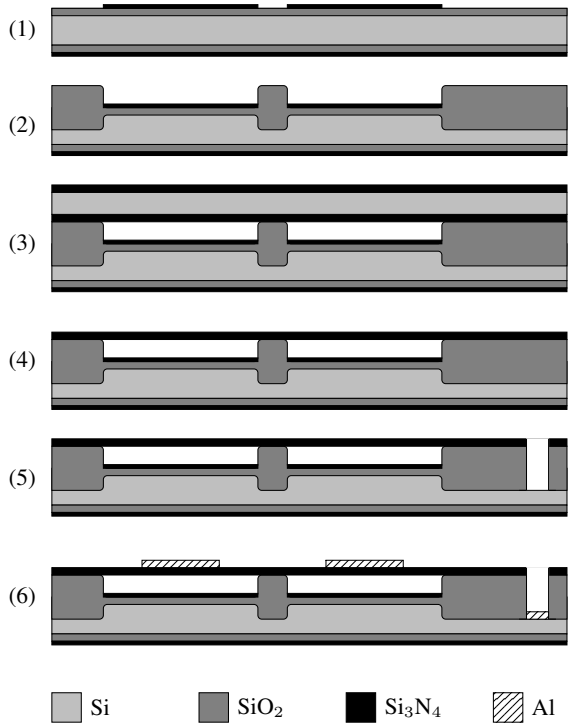


Fig. 1. Process flow of the CMUT.

TABLE I
Properties of the fabricated CMUTs.

Parameter	Value
Radius, a	$5\,\mu\mathrm{m}$
Plate thickness, t	$100\,\mathrm{nm}$
Metal thickness	$100\,\mathrm{nm}$
Gap height	$255\,\mathrm{nm}$
Metal line width, W	$2.5\,\mu\mathrm{m}$
Wire bond pad area	$(150\mathrm{x}250)\,\mu\mathrm{m}^2$
LOCOS Si_3N_4 thickness	$50\,\mathrm{nm}$
Si_3N_4 tensile stress	$1.3\,\mathrm{GPa}$
Number of cells (sensing element)	1024
Resonance frequency, f_0 ($V_{\mathrm{DC}} = \frac{2}{3}V_{\mathrm{PI}}$)	$38.8\,\mathrm{MHz}$
Pull-in voltage	$150\,\mathrm{V}$

The elements used for sensing in this study consist of 1024 cells connected in parallel. The theoretical zero voltage capacitance of the CMUT cells is $C_{\mathrm{CMUT}} = 0.64\,\mathrm{pF}$ while the theoretical parasitic capacitance is $C_{\mathrm{parasitic}} = 4.62\,\mathrm{pF}$. Hence, the useful capacitance C_{CMUT} only makes up 12% of the total capacitance of the element, which in part is due to the wire bond pad which accounts for 45% of the parasitic capacitance. In order to decrease this parasitic capacitance the area of the wire bond pad could be decreased. Another solution is to make a new design where the bond pad is raised above the plate by a polymer or dielectric material, thus effectively increasing the post height in this region, the

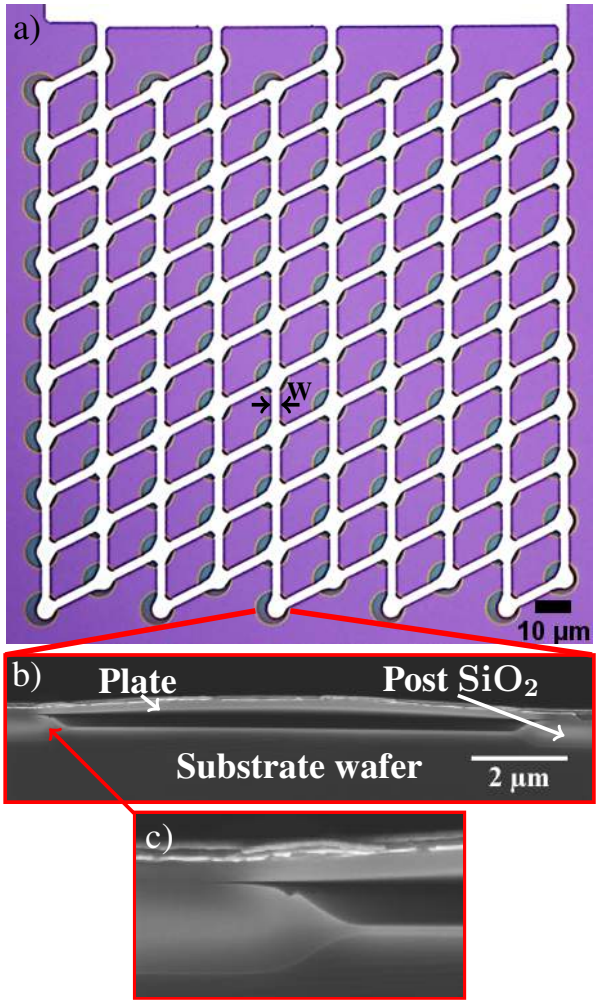


Fig. 2. a) Top view of a CMUT element of 10 by 10 cells showing the grid pattern of the metal layer. b) Scanning electron microscope image of a cross-section of a single CMUT cell. c) Scanning electron microscope image zoom in on the Bird's Beak structure and bonding interface.

downside being an additional lithography step. Figure 3 shows impedance spectra (Keysight, E4990A) of the CMUTs at three different DC voltages (Keithley, 2410) which demonstrates the spring softening effect. The resonance frequency and pull-in voltage is found from these spectra and was found to be $f_0 = 38.8\,\mathrm{MHz}\,(\mathrm{at}\,\frac{2}{3}V_{\mathrm{PI}})$ and $V_{\mathrm{PI}} = 150\,\mathrm{V}$, respectively (see Table I).

3) Sensitivity: Mass sensitivity is one of the key figures of merit of a gravimetric sensor. It can be calculated theoretically by approximating the CMUT as a 1-D linear harmonic oscillator and assuming that the mass added to the plate is much smaller than the mass of the plate. The theoretical distributed mass sensitivity, $S_{\mathrm{theoretical}}$, can then be expressed as [14]:

$$S_{\mathrm{theoretical}} = \frac{\partial f_0}{\partial m} = -\frac{1}{2}\frac{f_0}{m_{\mathrm{plate}}}, \qquad (1)$$

where f_0 is the resonance frequency and m_{plate} is the mass of the plate. The sensitivity is increased by having a plate with a

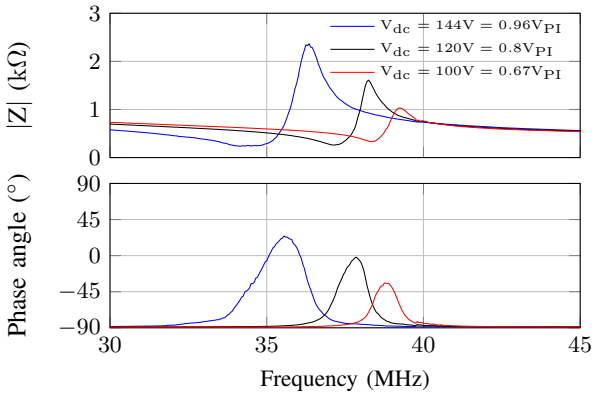


Fig. 3. Impedance spectra of a CMUT device at three bias voltages.

high resonance frequency and a low mass. In this paper this is achieved by having a low plate mass of 46 pg due to a plate thickness of 100 nm and Al top electrode thickness of 100 nm, while maintaining a resonance frequency of 38.8 MHz. The theoretical sensitivity $S_{\mathrm{theoretical}}$ for the CMUT device has been calculated using Equation 1: $S_{\mathrm{theoretical}} = 0.45\,\frac{\mathrm{Hz}}{\mathrm{ag}}$. Table II shows both the theoretical sensitivity and the measured sensitivity of $S_{\mathrm{measured}} = 0.24\,\frac{\mathrm{Hz}}{\mathrm{ag}}$. The measured sensitivity is found by determining the resonance frequency from impedance spectra and subsequently depositing a thin layer (11 nm) of Au on the CMUT element. Repeating the impedance measurement and deposition step, values of the resonance frequency as a function of added mass is found. Figure 4 shows such a plot for the device. The amount of mass added to the plate is calculated from the area of the plate and by measuring the step height of the deposited Au using AFM. Finally, the measured sensitivity is obtained by fitting a linear expression to the data and extracting the slope. Table II shows theoretical and measured sensitivities. The ratio $S_{\mathrm{measured}}/S_{\mathrm{theoretical}} = 0.53$ shows that the measured sensitivity is worse than the predicted theoretical value. This can be due to an overestimation of the added Au mass or a violation of the assumption for Equation 1 stating that $m_{\mathrm{plate}} >> m_{\mathrm{Au}}$ which is likely since $m_{\mathrm{plate}}/m_{\mathrm{Au}} \sim 3$. In order to decrease the discrepancy between the theoretical and measured value either a thinner Au layer or a layer of lower density should be used. In Table II the sensitivity normalized to the area of the CMUT cell is also given in order to ease comparisons with other CMUT sensors in the literature.

In order to compare sensitivities of different types of resonant gravimetric sensors a normalized sensitivity should be used. One way of calculating this normalized sensitivity is given by [15]:

$$S_{\mathrm{norm}} = \lim_{\Delta m \to 0} \frac{1}{f_0}\frac{\Delta f}{\Delta m/A} = \frac{\partial f}{\partial m}\frac{A}{f_0} = \frac{1}{2\rho t}, \qquad (2)$$

where ρ and t is the mass density and thickness of the plate, respectively. Equation 2 can be seen as a normalized version of Equation 1 with respect to the area and resonance frequency. For the CMUT the normalized sensitivity only depends on

TABLE II
TABLE OF THE THEORETICAL AND MEASURED SENSITIVITY AND LIMIT OF DETECTION FOR THE CMUT DEVICE.

Parameter	Value
$S_{\text{theoretical}}$	$0.45\,\frac{\text{Hz}}{\text{ag}}$
S_{measured}	$0.24\,\frac{\text{Hz}}{\text{ag}}$
$S^{-1}_{\text{theoretical}}/A$	$28.3\,\frac{\text{zg}}{\text{Hz}\cdot\mu\text{m}^2}$
$S^{-1}_{\text{measured}}/A$	$51.8\,\frac{\text{zg}}{\text{Hz}\cdot\mu\text{m}^2}$
LOD(1σ)	2.0 ag
LOD/A(1σ)	$25.8\,\frac{\text{zg}}{\mu\text{m}^2}$

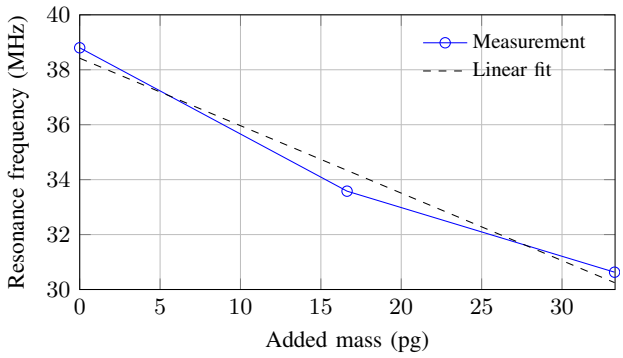


Fig. 4. Resonance frequency as a function of added mass from the Au deposition.

the inverse product of the mass density of the plate material and the thickness of the plate. The normalized sensitivity is therefore maximized, for a given material, by having a thin plate.

4) Limit of detection: The mass limit of detection (LOD) is a measure of the minimum change in mass a sensor can detect due to frequency noise. The LOD can be estimated by the following relation [2]:

$$\text{LOD}(\tau) = 2m_{\text{eff}}\sigma(\tau), \tag{3}$$

where m_{eff} is the effective mass of the plate and $\sigma(\tau)$ is the Allan deviation. The effective mass of the plate is lower than the total geometrical mass ($m_{\text{eff}} < m_{\text{total}} = \rho t A$) but in all calculations in this paper the total geometrical mass has been used in order to follow the established practice. The overlapping Allan deviation is shown in Figure 5 and is calculated from frequency noise data for the CMUT device. The minimum point on the graph is used to calculate the minimum LOD = 2.0 ag, using Equation 3. The results are gathered in Table II where also the normalized LOD is included.

B. Colorimetric chip

The colorimetric arrays consist of two columns of dye spots where each column comprises several spots, see Figure 6a). The spots in the left column, called coating A, is Reichardt's

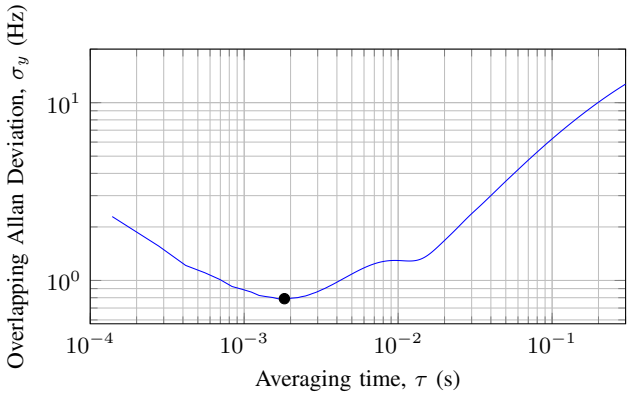


Fig. 5. Overlapping Allan deviation σ_y as a function of averaging time τ. Phase lock loop band width: 10 kHz.

dye and the dye in the right column, called coating B, is Bengal rose B. Coating A is expected to react with BMK, that is change color, while coating B is not expected to react. The spots are made by dispensing several droplets of dye in a liquid solution on a white polypropylene substrate. The dispensing and spot pattern are controlled by a computer and an automatic dispensing machine (GESIM, Nano-Plotter) using a piezoelectric actuated needle. The same setup is also used to coat the CMUT surfaces with dye. The two dyes used for the colorimetric arrays are used as functionalization layers on the CMUTs. The resonance frequency of the CMUT did not change significantly after the coating.

III. EXPERIMENTAL SETUP

The experimental setup is constructed so the CMUT frequency shift and colorimetric color change can be measured simultaneously. The CMUTs and the colorimetric chip are placed in a small ($\sim$ 72 cm^3) chamber, see Figure 6b). Through an inlet hole in the chamber, analytes can be injected. The analytes are in the liquid phase when injected into the chamber where they evaporate and thus increase the gas phase concentration of the analyte until the vapour pressure is reached. A microscope records an image of the colorimetric chip every 30 s through the transparent top of the chamber. An example of such an image can be seen in Figure 6a). The CMUT is glued to a TO-8 head and Au wirebonds connect the top and bottom electrodes to two pins, see Figure 6c). The coated CMUT element is also visible in this figure. Further, the two pins are connected to a bias tee which supplies the CMUT with a DC voltage (Keithley, 2410) and an AC voltage from a lock-in amplifier with a built-in phase locked loop (Zurich Instruments, HF2LI PLL). The PLL ensures that the CMUT is driven at resonance while the resonance frequency shifts are saved to a computer. The resonance frequency can be measured for two CMUT elements simultaneously.

IV. RESULTS

The color change of the colorimetric dyes is found by analysing the optical images recorded as time passes. Difference maps are created by subtracting all images from

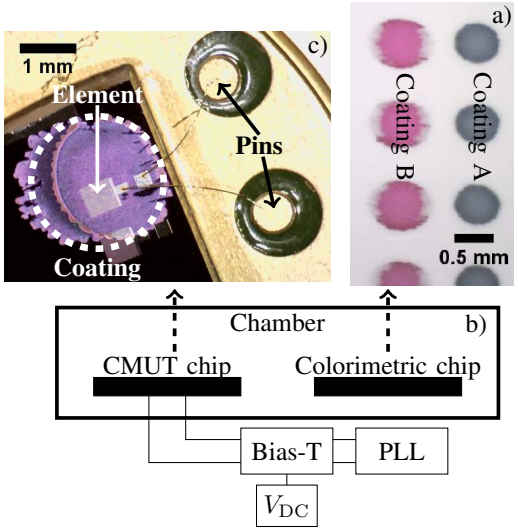


Fig. 6. a) Optical microscope image of the colorimetric chip showing two columns of dye. b) Schematic of the experimental setup showing the chamber enclosing the CMUT and colorimetric chip. The electrical actuation/readout circuit of the CMUT is also seen which consists of a phase locked loop, a bias-t, and a DC power supply. c) Optical microscope image of the CMUT chip wirebonded to two pins. A single element is seen, coated with a dye.

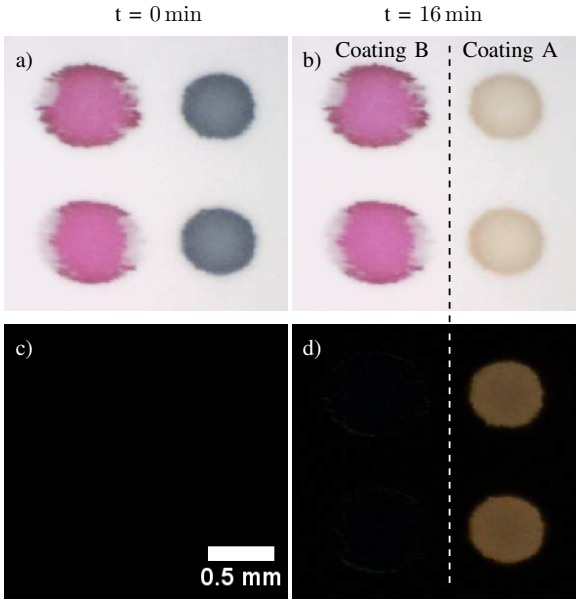


Fig. 7. Coating A and B, a) before and b) after reaction with BMK. Difference maps c) before and d) after reaction with BMK. BMK is injected at $t = 4\,\text{min}$.

a reference image, which here is chosen as the image at $t = 0\,\text{s}$. The difference maps highlight the *changes* in color and intensity of the dyes. Figure 7a) and b) show two microscope images before and after BMK has reacted with the dyes, respectively. Figure 7c) and d) show the difference maps at the corresponding times, where c) is all black since this was used as the reference image and d) clearly exhibits a color change for dye A.

Figure 8 shows the intensity of the difference maps inside

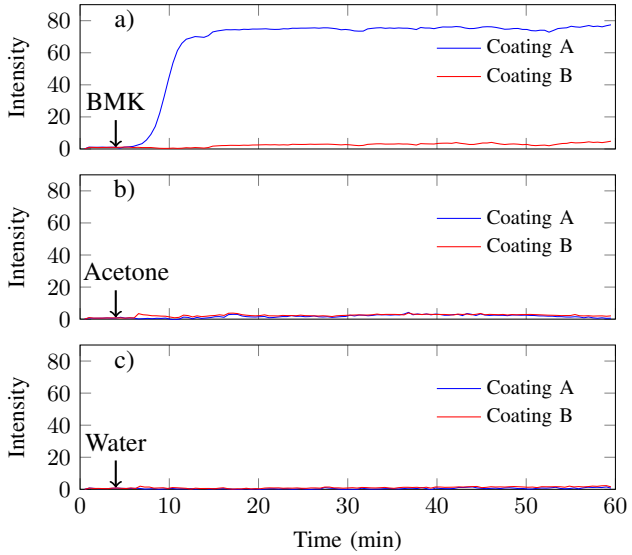


Fig. 8. Mean intensity of the red, green and blue channels as a function of time for a) BMK, b) acetone and c) water inside the dye spot area. The analyte is introduced at $t = 4\,\text{min}$ indicated by an arrow.

the dye spots A and B as a function of time. The intensity is calculated by taking the mean of the RGB values from the difference map in the area inside the dye spots. The three graphs each shows the intensity as a function of time when either a) BMK, b) acetone or c) water is introduced at $t = 4\,\text{min}$. In Figure 8 a) dye A shows an increase in intensity after BMK is injected and at about $t = 15\,\text{min}$ the intensity has reached a constant maximum value and it is assumed that all the dye has reacted with the BMK. The intensity of dye B does not change as a function of time since no color change has occurred. In addition, no color change is observed, for either coating, when acetone or water is injected, see Figure 8b) and c) respectively. Acetone is chosen as an analyte since the molecular structure is similar to that of BMK expect for an aromatic ring and water is chosen since water vapour is almost always present in the atmosphere in real life applications. Figure 8 demonstrates that coating A has an affinity towards BMK while coating B does not. Furthermore, coating A demonstrates a selectivity between BMK, acetone and water, since no color change is observed for the two latter cases.

Figure 9 shows two plots of the frequency shift as a function of time for two CMUTs coated with the two dyes. The added mass on the right axis is calculated using Equation 1 and the measured frequency shift. Two experiments are performed where in Figure 9, a) BMK and b) water are injected into the chamber, respectively. Consequently, the resonance frequency decreases due to the added mass of the evaporated analytes on the plate of the CMUT. The larger decrease in resonance frequency, in Figure 9a), for the CMUT with coating A, demonstrates that coating A has a stronger affinity towards BMK than the CMUT with coating B and no coating. Thus, the largest intensity change is observed for coating A and the

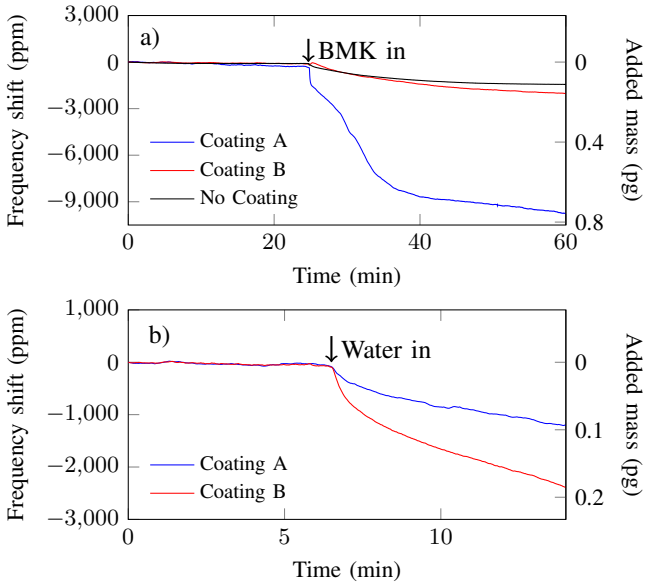


Fig. 9. Frequency shift and added mass as a function of time for a) BMK and b) water as analyte.

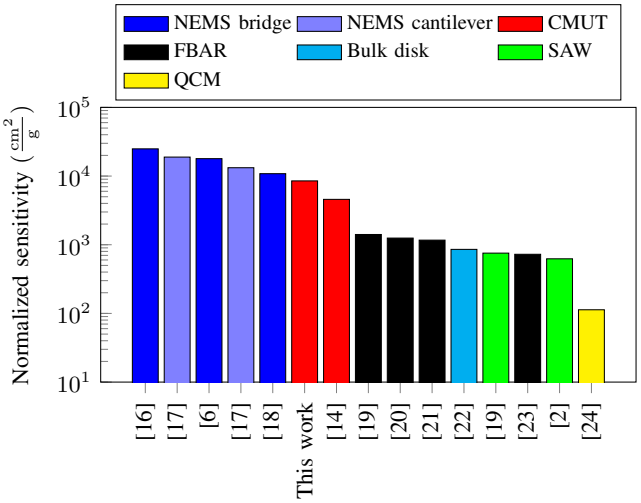


Fig. 10. Normalized sensitivity (see Equation 2) for different gravimetric sensor types.

largest frequency shift is observed for the CMUT with coating A. In order to test the selectivity of the coatings and whether the difference observed is due to the coatings and not due to differences of the specific CMUTs, the experiment is repeated with water as an analyte, see Figure 9b). The CMUT with coating A shows the smallest frequency shift. Hence, coating A is selective towards BMK both when used as a coating on the CMUTs and when used as a dye in the colorimetric chip.

V. Discussion

Normalized mass sensitivities for different gravimetric sensor types are plotted in Figure 10. The normalized sensitivities in the figure are examples of the state of the art for each sensor type. The plot shows that of the sensor types CMUTs have the second highest normalized sensitivity, with nanoelectromechanical system (NEMS) sensors being the highest, due to the small dimensions of these sensors. However, CMUTs are still desirable to use as gravimetric sensors due to a lower LOD [2], a higher SNR and an easier functionalization process. The highest normalized sensitivity of the CMUTs is obtained by the sensor with the thinnest plate, in agreement with Equation 2. This is the device presented in this work which has a plate thickness of $100\,\mathrm{nm}$.

As described in Section II-A, the CMUT cells were fabricated using a single LOCOS process and wafer fusion bonding to a Si_3N_4 plate. CMUTs have previously been fabricated by fusion bonding of SiO_2 and Si_3N_4 surfaces but the cavities were in these cases made with a reactive ion etch (RIE) process [25], [26]. In [26] a CMP step was required in order to reduce the roughness of the Si_3N_4 surface which was not needed in the process used in this paper. The advantage of using a LPCVD Si_3N_4 layer as a plate compared with using the Si device layer of a SOI wafer is a reduced cost and an increased thickness uniformity. In addition, it is difficult to obtain $100\,\mathrm{mm}$ SOI wafers with a device layer thinner than a couple hundreds nm, which, in the case of CMUTs for sensing, is desired for increasing the mass sensitivity.

VI. Conclusion

A sensor system consisting of a CMUT and a colorimetric array was fabricated and used to detect BMK in the gas phase. The CMUT frequency shift directly gave the mass of the absorbed BMK while the colorimetric array demonstrated excellent selectivity. When used as a functionalization layer on the CMUT, the dye also provided the CMUT with an affinity for BMK, compared with water. The CMUTs were fabricated using a single LOCOS step and wafer bonding to a thin ($100\,\mathrm{nm}$) Si_3N_4 plate, thus eliminating the need for expensive SOI wafers. Furthermore, the sensitivity of the CMUT was both measured and calculated which resulted in the lowest sensitivity reported: $28.3\,\frac{zg}{Hz\cdot\mu m^2}$. The normalized sensitivity was found to be $8.5\times10^3\,\frac{cm^2}{g}$ which is surpassed only by NEMS cantilevers/bridges when comparing with other resonant gravimetric sensor types, which highlights one of the advantages of using the CMUT as a gravimetric sensor. Future work could include making a setup where the color change is measured for the dye directly on the CMUT chip.

REFERENCES

[1] I. J. Forbes and K. P. Kirkbride, "The Origin of Alkenes in Illicit Amphetamine: An Examination of the Illicit Synthesis of Phenyl-2-Propanone," *Journal of Forensic Sciences*, vol. 37, no. 5, pp. 1311–1318, 1992.

[2] S. Fanget, S. Hentz, P. Puget, J. Arcamone, M. Matheron, E. Colinet, P. Andreucci, L. Duraffourg, E. Myers, and M. Roukes, "Gas sensors based on gravimetric detection - A review," *Sensors and Actuators B: Chemical*, vol. 160, pp. 804 – 821, 2011.

[3] S. Munoz-Aguirre, A. Yoshino, T. Nakamoto, and T. Moriizumi, "Odor approximation of fruit flavors using a QCM odor sensing system," *Sensors and Actuators B*, vol. 123, p. 11011106, 2007.

[4] M. Zhang, L. Du, Z. Fang, and Z. Zhao, "A Sensitivity-Enhanced Film Bulk Acoustic Resonator Gas Sensor with an Oscillator Circuit and Its Detection Application," *Micromachines*, vol. 8, no. 25, 2017.

[5] W. P. Jakubik, "Surface acoustic wave-based gas sensors," *Thin Solid Films*, vol. 520, p. 986993, 2011.

[6] Y. T. Yang, C. Callegari, X. L. Feng, and M. L. Roukes, "Surface Adsorbate Fluctuations and Noise in Nanoelectromechanical Systems," *Nano Lett.*, vol. 11, pp. 1753–1759, 2011.

[7] H. J. Lee, K. K. Park, M. Kupnik, Ö. Oralkan, and B. T. Khuri-Yakub, "Chemical Vapor Detection Using a Capacitive Micromachined Ultrasonic Transducer," *Anal. Chem.*, vol. 83, pp. 9314–9320, 2011.

[8] K. K. Park, H. J. Lee, G. G. Yaralioglu, A. S. Ergun, Ö. Oralkan, M. Kupnik, C. F. Quate, B. T. Khuri-Yakub, T. Braun, J.-P. Ramseyer, H. P. Lang, M. Hegner, C. Gerber, and J. K. Gimzewski, "Capacitive micromachined ultrasonic transducers for chemical detection in nitrogen," *Applied Physics Letters*, vol. 91, 2007.

[9] D. Virzonis, G. Vanagas, A. Ramanaviciene, A. Makaraviciute, D. Barauskas, A. Ramanavicius, W. Wen, and R. Kodzius, "Resonant gravimetric immunosensing based on capacitive micromachined ultrasound transducers," *Microchim Acta*, vol. 181, p. 17491757, 2014.

[10] H. J. Lee, K. K. Park, M. Kupnik, and B. T. Khuri-Yakub, "Functionalization layers for CO2 sensing using capacitive micromachined ultrasonic transducers," *Sensors and Actuators B: Chemical*, vol. 174, p. 87 93, 2012.

[11] D. Barauskas, D. Pelenis, D. Virzonis, J. P. Baltrus, and J. Baltrusaitis, "Greenhouse Gas Molecule CO2 Detection Using a Capacitive Micromachined Ultrasound Transducer," *Anal. Chem.*, vol. 88, no. 13, pp. 6662–6665, 2016.

[12] Q. Stedman, K.-K. Park, and B. T. Khuri-Yakub, "Distinguishing Chemicals Using CMUT Chemical Sensor Array and Artificial Neural Networks," *IEEE International Ultrasonics Symposium Proceedings*, 2014.

[13] M. Engholm, T. Pedersen, and E. V. Thomsen, "Modeling of plates with multiple anisotropic layers and residual stress," *Sensors and Actuators A: Physical*, vol. 240, pp. 70–79, 2016.

[14] H. J. Lee, K. K. Park, Ömer Oralkan, M. Kupnik, and B. T. Khuri-Yakub, "A Multichannel Oscillator for a Resonant Chemical Sensor System," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 10, pp. 5632–5640, 2014.

[15] S. W. Wenzel and R. M. White, "Analytic comparison of the sensitivities of bulk-wave, surface-wave, and flexural platewave ultrasonic gravimetric sensors," *Appl. Phys. Lett.*, vol. 54, 1989.

[16] Y. T. Yang, C. Callegari, X. L. Feng, K. L. Ekinci, and M. L. Roukes, "Zeptogram-Scale Nanomechanical Mass Sensing," *Nano Lett.*, vol. 6, no. 4, 2006.

[17] M. Li, H. X. Tang, and M. L. Roukes, "Ultra-sensitive NEMS-based cantilevers for sensing, scanned probe and very high-frequency applications," *Nature Nanotechnology*, vol. 2, pp. 114–120, 2007.

[18] K. L. Ekinci and M. L. Roukes, "Nanoelectromechanical systems," *Rev. Sci. Instrum.*, vol. 76, 2005.

[19] G. Wingqvist, L. Arapan, V. Yantchev, and I. Katardjiev, "A micromachined thermally compensated thin film Lamb wave resonator for frequency control and sensing applications," *J. Micromech. Microeng.*, vol. 19, 2009.

[20] J. Weber, W. M. Albers, J. Tuppurainen, M. Link, R. Gabl, W. Wersing, and M. Schreiter, "Shear mode FBARs as highly sensitive liquid biosensors," *Sensors and Actuators A*, vol. 128, p. 8488, 2006.

[21] Y. lu Feng, M. wei Liu, X. Li, and J. Li, "Micromachined mass sensor based on film bulk acoustic resonator," *Symposium on Piezoelectricity, Acoustic Waves, and Device Applications*, 2015.

[22] A. Cagliani and Z. J. Davis, "Ultrasensitive bulk disk microresonator-based sensor for distributed mass sensing," *J. Micromech. Microeng.*, vol. 21, 2011.

[23] H. Zhang and E. S. Kim, "Micromachined Acoustic Resonant Mass Sensor," *Journal of Microelectromechanical Systems*, vol. 14, no. 4, pp. 699–706, 2005.

[24] M. Pohanka, "The Piezoelectric Biosensors: Principles and Applications, a Review," *Int. J. Electrochem. Sci.*, vol. 12, p. 496506, 2017.

[25] K. Midtbø, A. Rønnekleiv, and D. T. Wang, "Fabrication and Characterization of CMUTs realized by Wafer Bonding," *IEEE Ultrasonics Symposium*, pp. 934–937, 2006.

[26] A. I. H. Chen, L. L. P. Wong, Z. Li, S. Na, and J. T. W. Yeow, "Practical CMUT Fabrication With a Nitride-to-Oxide-Based Wafer Bonding Process," *Journal of Microelectromechanical Systems*, vol. 26, no. 4, pp. 829–836, 2017.

APPENDIX C

Conference paper - Combined Colorimetric and Gravimetric CMUT Sensor for Detection of Phenylacetone

This paper is published in the conference proceeding: IEEE International Ultrasonics Symposium (IUS), 2017 [56].

Combined Colorimetric and Gravimetric CMUT Sensor for Detection of Phenylacetone

Mathias J. G. Mølgaard*, Milan Laustsen*, Ida L. Thygesen*, Mogens H. Jakobsen*, Thomas L. Andresen*, and Erik V. Thomsen*

*Department of Micro and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

***Abstract*—The detection of phenylacetone is of interest as it is a common precursor for the synthesis of (meth)amphetamine. Resonant gravimetric sensors can be used to detect the mass and hereby the concentration of a gas while colorimetric arrays typically have an exceptional selectivity to the target analyte if the right colorimetric dyes are chosen. We present a sensor system consisting of a Capacitive Micromachined Ultrasonic Transducer (CMUT) and a colorimetric array for detection of phenylacetone. The CMUT is used as a resonant gravimetric gas sensor where the resonance frequency shift due to mass loading of the plate. A single Local Oxidation of Silicon (LOCOS) step was used to define the cavities which were sealed with a Si_3N_4 plate with a thickness of $100\,\mathrm{nm}$, resulting in a resonance frequency of $38.8\,\mathrm{MHz}$ and a theoretical mass sensitivity of $28.3\,\frac{\mathrm{zg}}{\mathrm{Hz}\cdot\mu\mathrm{m}^2}$. The CMUTs were functionalized with the same dyes used to fabricate colorimetric arrays. While both the CMUTs and the colorimetric arrays showed selectivity to phenylacetone, the best selectivity was achieved by the colorimetric array. Furthermore, the mass of the phenylacetone was found as a function of time. Thus, the combination of the colorimetric array and the CMUT results in a good selectivity and a quantitative value for the mass.**

I. Introduction

Detection of illegal drug and explosive precursors is of interest for government agencies in order to prevent or stop this type of criminal activity. Amphetamine and methamphetamine are recreational drugs and illegal to possess in many countries as they are controlled substances. Multiple synthesization routes exists where a common one is the reductive amination of the illegal precursor molecule phenylacetone/BMK [1]. Therefore, it is desirable to be able to detect BMK, especially in the gas phase.

CMUTs have previously been used for sensing applications: In 2007 Park *et al.* demonstrated the use of CMUTs as sensors by detecting water, isopropanol, acetone, and methanol [2]. Detection of chemical weapons is another possible use case for the CMUT sensor as it can detect very small concentrations. Dimethyl methylphosphonate (DMMP) which is a common simulant for sarin has been measured in the gas phase with a CMUT sensor with a high volume sensitivity and low limit of detection while obtaining a good selectivity towards dodecane and 1-octanol [3] [4]. CMUTs have also been used as biosensors, specifically as an immunosensor where the antigen antibody pair interacts in a highly selective reaction causing the resonance frequency to decrease [5]. Finally, CMUTs have also been used for environmental monitoring where the concentration of CO_2 has been measured [6] [7].

A functional layer is needed in order to make the gravimetric sensors selective towards a specific analyte. Functionalization is typically done by applying a thin polymer coating on top of the vibrating part. Due to the 'closed' and planar surface of the CMUT device several methods can be employed including: spin coating, drop coating, dip coating, and spray coating. Despite of these layers, obtaining a very high selectivity still remains as a challenge for gravimetric sensors and therefore also CMUTs. One method of improving selectivity between multiple analytes in a gas has been to employ several CMUTs each coated with a different functionalization layer [8]. A single chip with several CMUT elements can easily be fabricated, thus minimizing the footprint and complexity.

In this paper we combine a CMUT sensor with a colorimetric array to detect BMK. By doing this we benefit from the high selectivity offered by the colorimetric array and the quantitative output of the CMUT which can be related to the concentration of the analyte. Furthermore, having two independent sensor types combined provides the total sensor system with redundancy. We present the fabrication and characterization of the CMUT and colorimetric array. The intensity as a function of time and resonance frequency shift as a function of time for the colorimetric array and CMUT respectively shows that the dyes provide both sensor types with a selectivity towards BMK.

II. Devices

In this section the fabrication and characterization of the CMUTs and colorimetric chips are described.

A. CMUT

1) Fabrication: The CMUT cells were fabricated using a single local oxidation of silicon (LOCOS) step and fusion wafer bonding. CMUTs have previously been fabricated by fusion bonding SiO_2 and Si_3N_4 surfaces but the cavities where in these cases made with a RIE process [9] [10].

The complete process flow can be seen in Figure 1. In step (1) a SiO_2 layer was grown on a low electrical resistivity Si wafer ($< 0.025\,\Omega\mathrm{cm}$, $<100>$). This SiO_2 layer separates the Si_3N_4 layer from the Si surface which otherwise could have defects introduced due to a tensile stress of $\sim 1.3\,\mathrm{GPa}$ in the Si_3N_4. Further, the Si_3N_4 layer was patterned using UV lithography and a H_3PO_4 etch. Subsequently, LOCOS was performed (step (2)) using the patterned Si_3N_4 as the

978-1-5386-3383-0/17/$31.00 ©2017 IEEE

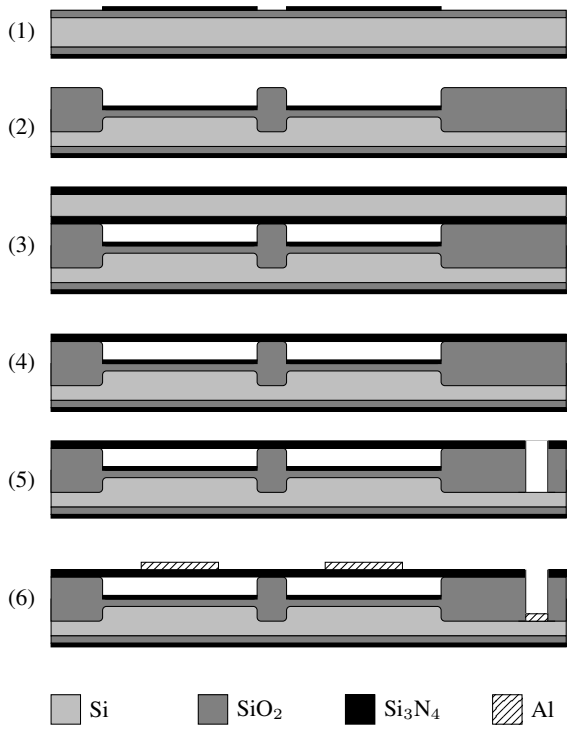


Fig. 1. Process flow for the CMUT.

TABLE I
PROPERTIES OF THE FABRICATED CMUTS.

Parameter	Value
Radius, a	$5\,\mu$m
Plate thickness, t	100 nm
Metal thickness	100 nm
Gap height	255 nm
LOCOS Si_3N_4 thickness	50 nm
Si_3N_4 tensile stress	1.3 GPa
Number of cells	1024
Resonance frequency, f_0 ($V_{DC} = \frac{2}{3}V_{PI}$)	38.8 MHz
Pull-in voltage	150 V

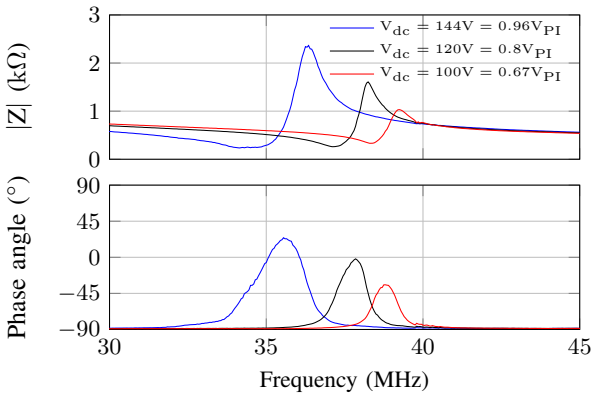


Fig. 2. Impedance spectra of a CMUT device at three different bias voltages.

oxidation mask. The cavities were sealed under vacuum in a wafer bonding machine (step (3)) by bonding the cavity wafer with a Si wafer with a layer of Si_3N_4. In step (4) the top Si_3N_4 layer was etched by a dry etch and the Si by a KOH etch. KOH was chosen due to the high etch selectivity between Si and Si_3N_4. Hence, the Si_3N_4 plate and the Si_3N_4 layer on the backside of the cavity wafer acted as etch stop layers. Bottom electrode openings were etched in step (5) with a dry etch and finally Al was deposited and patterned in step (6) using a $H_3PO_4 : H_2O$ etch at room temperature.

2) Dimensions and characterization: A single element consists of 1024 cells connected in parallel. This parallelism lowers the impedance and increases the signal to noise ratio. The dimensions and properties of the elements can be seen in Table I. The value of the tensile stress of the Si_3N_4 is found by measuring the curvature of a double side polished wafer with a layer of Si_3N_4 of known thickness on one side and then applying Stoney's equation. The tensile stress increases both the resonance frequency and pull-in voltage compared with a similar plate with no stress [11].

The resonance frequency and pull-in voltage is found from impedance spectra. Three such spectra can be seen in Figure 2 where also the well-known spring softening effect is observed.

3) Sensitivity: Mass sensitivity is one of the key figures of merit of a gravimetric sensor. It can be calculated theoretically by approximating the CMUT as a 1-D linear harmonic oscillator and assuming that the mass added to the plate is much smaller than the mass of the plate. The distributed mass sensitivity can then be expressed as:

$$S = \frac{\partial f_0}{\partial m} = -\frac{1}{2}\frac{f_0}{m_{\text{plate}}}, \tag{1}$$

where f_0 is the resonance frequency and m_{plate} is the mass of the plate. The sensitivity is increased by having a plate with a high resonance frequency and a low mass. The sensitivity is in this paper sought maximized by having a plate thickness of 100 nm, decreasing the mass of the plate, while still maintaining a resonance frequency of 38.8 MHz.

The theoretical sensitivity, $S_{\text{theoretical}}$, has been calculated using Equation 1 and can be seen in Table II. In addition, a measured sensitivity is given in the same table. The measurements are performed by finding the resonance frequency from impedance spectra and subsequently depositing a thin layer of Au on the CMUT element. By measuring the step height of the deposited Au the added mass on the plate is calculated. Repeating the impedance measurement and deposition step, values of the resonance frequency as a function of added mass is found. The sensitivity was directly found from the slope of such a graph. Table II gives the inverse of this slope value together with the theoretical value calculated by means of Equation 1. The ratio $S_{\text{measured}}/S_{\text{theoretical}} = 1.86$ shows that the measured sensitivity is worse than the predicted theoretical value. This can be due to an overestimation of the added Au mass or a violation of the assumption for Equation 1 stating that $m_{\text{plate}} >> m_{\text{Au}}$ since $m_{\text{plate}}/m_{\text{Au}} \sim 3$. In order to decrease the discrepancy between the theoretical and measured value either a thinner Au layer or a layer of lower density

978-1-5386-3383-0/17/$31.00 ©2017 IEEE

TABLE II
TABLE OF THEORETICAL AND MEASURED SENSITIVITIES AND LODS OF THE CMUT DEVICE.

Parameter	Value
$S_{\text{theoretical}}$	$2.2\ \frac{\text{ag}}{\text{Hz}}$
S_{measured}	$4.1\ \frac{\text{ag}}{\text{Hz}}$
$S_{\text{theoretical}}/A$	$28.3\ \frac{\text{zg}}{\text{Hz}\cdot\mu\text{m}^2}$
S_{measured}/A	$51.8\ \frac{\text{zg}}{\text{Hz}\cdot\mu\text{m}^2}$
LOD(1σ)	$2.0\,\text{ag}$
LOD/A(1σ)	$25.8\ \frac{\text{zg}}{\mu\text{m}^2}$

should be used. In Table II the sensitivity normalized with the area of the CMUT cell is also given in order to make comparisons with other sensors easier.

4) Limit of detection: The mass limit of detection (LOD) is a measure of the minimum change in mass a sensor can detect due to frequency noise. The LOD can be estimated by the following relation [12]:

$$\text{LOD}(\tau) = 2m_{\text{plate}}\sigma(\tau), \tag{2}$$

where m_{plate} is the mass of the plate and $\sigma(\tau)$ is the Allan deviation. The Overlapping Allan Deviation has been measured and the minimum limit of detection calculated, see Table II.

B. Colorimetric chip

The colorimetric arrays consist of two columns of dye spots where each column comprises several spots, see Figure 3 a). The spots in the left column, called coating A, is Reichardt's dye and the dye in the right column, called coating B, is Bengal rose B. Coating A is expected to react with BMK, that is change color, while coating B is not expected to react. The spots are made by dispensing several drops of dye in a liquid solution on a white polypropylene substrate. The dispensing and spot pattern are controlled by a computer and an automatic dispensing machine (GESIM, Nano-Plotter) using a piezoelectric actuated needle. The same setup is also used to coat the CMUT surfaces with dye. The two dyes used for the colorimetric arrays are used as functionalization layers on the CMUTs. The thickness of the dye on the CMUT was chosen so the resonance frequency did not significantly change compared with the uncoated state.

III. EXPERIMENTAL SETUP

The experimental setup is constructed so that the CMUT frequency shift and colorimetric color change can be measured at the same time. The CMUTs and the colorimetric chip are placed in a small ($\sim 72\,\text{cm}^3$) chamber, see Figure 3 b). Through an inlet hole in the chamber, analytes can be introduced. A microscope takes an image of the colorimetric chip every $30\,\text{s}$ through the transparent top of the chamber. An example of such an image can be seen in Figure 3 a). The

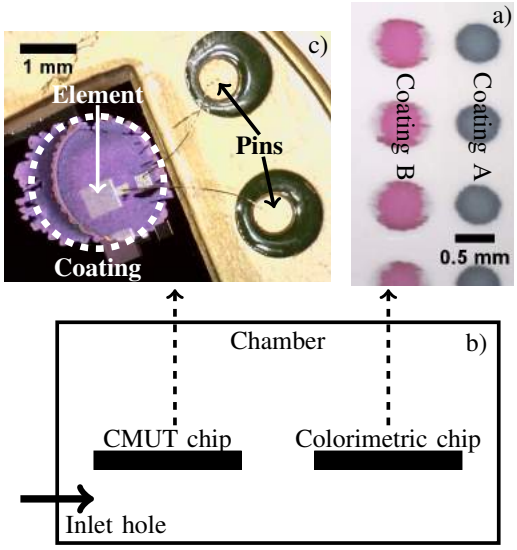


Fig. 3. a) optical microscope image of the colorimetric chip showing two columns of dye. b) schematic of the experimental setup. c) optical microscope image of the CMUT chip wirebonded to two pins. A single element is seen, coated with a dye.

CMUT is glued to a TO-8 head and Au wirebonds connect the top and bottom electrodes to two pins, see Figure 3 c). The coated CMUT element is also visible in this figure. Further, the two pins are connected to a bias tee which supplies the CMUT with a bias DC voltage (Keithley 2410) and an AC voltage from a lock-in amplifier with a built-in phase locked loop (Zurich Instruments, HF2LI PLL). The PLL ensures that the CMUT is driven at resonance while the resonance frequency shifts are saved to a computer. The resonance frequency can be measured for two CMUT elements simultaneously.

IV. RESULTS AND DISCUSSION

Figure 4 shows the intensity as a function of time for the two dyes. The intensity is calculated by taking the mean of the RGB channel intensities inside the dye spots for difference maps. In Figure 4 coating A shows an increase in intensity after BMK is injected at $t = 4\,\text{min}$. At about $t = 15\,\text{min}$ the intensity is constant and it is assumed that all the dye has reacted with the BMK. The opposite is true for coating B, here the intensity is more or less constant for the entire duration indicating no change in color. No color change is observed, for either coating, when acetone or water is injected. Acetone is chosen as an analyte since the molecular structure is the same as BMK expect for an aromatic ring and water is chosen since water vapour is almost always present in the atmosphere. Figure 4 demonstrates that coating A has an affinity towards BMK. Furthermore, coating A shows selectivity between BMK, acetone and water, since no color change is observed for the two latter analytes.

Figure 5 shows the frequency shift for three CMUTs: one with coating A, one with coating B and one without any coating. The added mass on the right axis is calculated using

978-1-5386-3383-0/17/$31.00 ©2017 IEEE

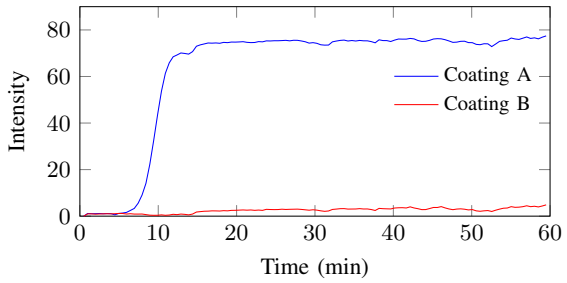


Fig. 4. Mean intensity of the red, green and blue channels as a function of time. BMK is injected into the chamber at $t \approx 4\,\mathrm{min}$. Background intensity has been subtracted.

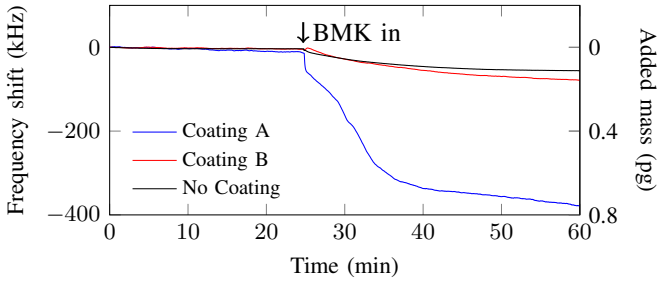


Fig. 5. Frequency shift as a function of time with BMK as the analyte.

Equation 1 and the measured frequency shift. At the time indicated in Figure 5, BMK is injected into the chamber. Consequently, the resonance frequency decreases due to the added mass of the evaporated analyte. The bigger frequency shift for the CMUT with coating A compared with coating B and the CMUT with no coating, shows, again, that coating A has an affinity towards BMK. Thus, coating A both gives the biggest change in intensity (due to a color change) (Figure 4) compared with coating B and a larger frequency shift than coating B with BMK as an analyte. Water was also used as an analyte in order to test the selectivity of the coatings and whether the difference observed, for BMK, is due to the coatings and not due to differences of the CMUTs themselves. With water as an analyte coating A had the smallest frequency shift. Coating A is therefore exhibits selectivity both when used as a coating on the CMUTs and when used as a dye in the colorimetric chip.

V. Conclusion

A sensor system consisting of a CMUT and a colorimetric array was presented. The CMUT was fabricated using a single LOCOS step and wafer bonding to a Si_3N_4 layer on a double side polished Si wafer, thus eliminating the need for expensive SOI wafers. Furthermore, the CMUT had a calculated sensitivity of $28.3\,\frac{\mathrm{zg}}{\mathrm{Hz}\cdot\mu\mathrm{m}^2}$ and a LOD(1σ) of $2.0\,\mathrm{ag}$. The colorimetric array and the CMUT both showed the largest affinity towards BMK in their respective responses due to the dyes which doubled as the functional coating on the CMUT.

References

[1] I. J. Forbes and K. P. Kirkbride, "The Origin of Alkenes in Illicit Amphetamine: An Examination of the Illicit Synthesis of Phenyl-2-Propanone," *Journal of Forensic Sciences*, vol. 37, no. 5, pp. 1311–1318, 1992.

[2] K. K. Park, H. J. Lee, G. G. Yaralioglu, A. S. Ergun, Ö. Oralkan, M. Kupnik, C. F. Quate, B. T. Khuri-Yakub, T. Braun, J.-P. Ramseyer, H. P. Lang, M. Hegner, C. Gerber, and J. K. Gimzewski, "Capacitive micromachined ultrasonic transducers for chemical detection in nitrogen," *Applied Physics Letters*, vol. 91, 2007.

[3] H. J. Lee, K. K. Park, M. Kupnik, Ö. Oralkan, and B. T. Khuri-Yakub, "Chemical Vapor Detection Using a Capacitive Micromachined Ultrasonic Transducer," *Anal. Chem.*, vol. 83, pp. 9314–9320, 2011.

[4] K. K. Park, H. Lee, M. Kupnik, Ö. Oralkan, J.-P. Ramseyer, H. P. Lang, M. Hegner, C. Gerber, and B. T. Khuri-Yakub, "Capacitive micromachined ultrasonic transducer (CMUT) as a chemical sensor for DMMP detection," *Sensors and Actuators B: Chemical*, vol. 160, p. 1120 1127, 2011.

[5] D. Virzonis, G. Vanagas, A. Ramanaviciene, A. Makaraviciute, D. Barauskas, A. Ramanavicius, W. Wen, and R. Kodzius, "Resonant gravimetric immunosensing based on capacitive micromachined ultrasound transducers," *Microchim Acta*, vol. 181, p. 17491757, 2014.

[6] H. J. Lee, K. K. Park, M. Kupnik, and B. T. Khuri-Yakub, "Functionalization layers for CO2 sensing using capacitive micromachined ultrasonic transducers," *Sensors and Actuators B: Chemical*, vol. 174, p. 87 93, 2012.

[7] D. Barauskas, D. Pelenis, D. Virzonis, J. P. Baltrus, and J. Baltrusaitis, "Greenhouse Gas Molecule CO2 Detection Using a Capacitive Micromachined Ultrasound Transducer," *Anal. Chem.*, vol. 88, no. 13, pp. 6662–6665, 2016.

[8] Q. Stedman, K.-K. Park, and B. T. Khuri-Yakub, "Distinguishing Chemicals Using CMUT Chemical Sensor Array and Artificial Neural Networks," *IEEE International Ultrasonics Symposium Proceedings*, 2014.

[9] A. I. H. Chen, L. L. P. Wong, Z. Li, S. Na, and J. T. W. Yeow, "Practical CMUT Fabrication With a Nitride-to-Oxide-Based Wafer Bonding Process," *Journal of Microelectromechanical Systems*, vol. 26, no. 4, pp. 829–836, 2017.

[10] K. Midtbø, A. Rønnekleiv, and D. T. Wang, "Fabrication and Characterization of CMUTs realized by Wafer Bonding," *IEEE Ultrasonics Symposium*, pp. 934–937, 2006.

[11] M. Engholm, T. Pedersen, and E. V. Thomsen, "Modeling of plates with multiple anisotropic layers and residual stress," *Sensors and Actuators A: Physical*, vol. 240, pp. 70–79, 2016.

[12] S. Fanget, S. Hentz, P. Puget, J. Arcamone, M. Matheron, E. Colinet, P. Andreucci, L. Duraffourg, E. Myers, and M. Roukes, "Gas sensors based on gravimetric detection - A review," *Sensors and Actuators B: Chemical*, vol. 160, pp. 804 – 821, 2011.

978-1-5386-3383-0/17/$31.00 ©2017 IEEE

APPENDIX D

Process flow - G1

Process flow title			Revision
CMUT – 2xLOCOS nitride 1G			**1G**
DTU Danchip National Center for Micro- and Nanofabrication	Contact email migmol@nanotech.dtu.dk	Contact person Mathias Mølgaard	Contact phone 22 73 0936 0-56310
	Labmanager group MEMS(3313)	Batch name 2xLOCOS 1G	Date of creation 23-Feb-16	Date of revision 23-Jun-16

Objective

Batch name: 2xLOCOS 1G

The purpose of this project is to fabricate Capacitive Ultrasonic Transducers using a double LOCOS process to define the cavities.

Substrates

Substrate	Orient.	Size	Doping/type	Polish	Thickness	Box	Device layer thickness	Purpose	#	Sample ID
Si	<100>	4”	p (boron)	SSP	525±25µm			Bottom electrode		OP479
SOI	<100>	4”	n (arsenic)	DSP	380-400µm	1um	0.5+-0.5um (thinned down beforehand)	Plate		

Step	Heading	Equipment	Procedure	Comments
1	Preparation and protective oxidation			**Bottom electrode wafers**
1.1	Wafer selection	Wafer box	Use the dedicated vacuum tweezer across from the A-stack furnaces	OP479
1.2	RCA			Remember 1 test wafer
1.3	SiO2 dry oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Dry900, time: 28 min Target thickness: 10 nm	Remember 1 test wafer
1.4	Measure thickness	Filmtek	Measured Thickness: 9.4 nm	Measure on test wafer
2	Nitride deposition – 1.LOCOS			**Bottom electrode wafers**
2.1	Si3N4 deposition	Nitride furnace – LPCVD 6” E3	Deposit nitride Recipe: 4nitdan Time: 20 min (@ 2.55nm/min) Target thickness: 50 nm	Remember 1 test wafer
2.2	Measure thickness	Filmtek	Measured Thickness: 52.6 nm	Measure on test wafer
3	PolySi deposition – 1.LOCOS			**Bottom electrode wafers**
3.1	Polysilicon deposiiton	Polysilicon furnace – LPCVD 6” E2	Recipe: Poly Time: 13 min (@7.23nm/min) Target thickness: 100 nm	Remember 1 test wafer
3.2	Measure thickness	Filmtek	Measured Thickness: 92.8 nm	Measure on test wafer
4	Lithography 1.5 µm – 1.LOCOS			**Bottom electrode wafer**
4.1	Prepare + coat wafer	HMDS + Spin track	AZ nLOF 2020 negative resist Recipe: 1.5 µm with HMDS	Faster to use HMDS oven than Spin track with many wafers. Run 2 dummy wafers to ensure resist coverage (without HMDS)
4.2	Exposure	KS aligner	Align to flat (first print). Hard contact Exposure time: 19 sec Mask: 1LOCOS	
4.3	PEB + Develop	TMAH UV developer	PEB 110C 60s + 60s puddle	
4.4	Inspection	Optical microscope	Check pattern, alignment marks and exposure test-structures	
5	Polysilicon mask etch – 1.LOCOS			**Bottom electrode wafer**
5.1	Polysilicon etch	DRIE-Pegasus	2 cycles of LF SOI recipe. Recipe name: Poly etch (20°C), use MACS	
5.2	Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 30 min.	
6	Nitride wet etch – 1.LOCOS mask nitride			**Bottom electrode wafer**
6.1	Nitride etch	Wet bench	Wet silicon nitride etch H3PO4 @ 180°C (85 Å/min) Time: ~6 min + a little over etch = 8 min	All wafers
6.2	7up	7up bath	Clean wafers after wet nitride etch due to	

			possible contamination by potassium ions from people stripping nitride after KOH	
6.3	Inspection		Optical microscope	Check pattern and alignment marks
7	Polysilicon strip – frontside only			**Bottom electrode wafer**
7.1	Strip polysilicon mask	DRIE-Pegasus	1 cycle of LF SOI recipe. Recipe name: 100nmPolySi_SOI etch (20°C)	
8	Oxidation - 1.LOCOS			**Bottom electrode wafer**
8.1	RCA	RCA bench		Remember test wafer
8.2	SiO2 wet oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Wet1100, time: 0h53min00s Target thickness: 573 nm (on test wafer) Measured thickness on test: 553.7 nm	1 test wafer
9	Strip oxide and nitride – 1.LOCOS			**Bottom electrode wafer**
9.1	Wet oxide etch	BHF – clean (across from HMDS oven)	Etch rate: 78.19 nm/min Etch time: ~8 min + over etch = 11 min	Compare the target thickness to the measured thickness – if much higher, adjust the etch time.
9.2	Nitride etch	Wet bench	Wet silicon nitride etch H3PO4 @ 180°C (85 Å/min) Time: ~6 min + a little over etch = 8 min	
9.3	7up	7up	Clean wafers after wet nitride etch due to possible contamination by potassium ions from people stripping nitride after KOH	
9.4	Wet oxide etch	BHF – clean (across from HMDS oven)	Etch rate: 78.19 nm/min Etch time: 5 min (3 min should be enough)	
9.5	Inspection	Dektak or AFM or Sensofar	Measure the bump height Target height: 250 nm Measured height: (249±2) nm	
10	Oxidation – Insulation oxide			**Bottom electrode wafer**
10.1	RCA	RCA bench		Remember 1 test wafer
10.2	SiO2 wet oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Wet1100, time: 0h03min30s Target thickness: 100 nm (on test wafer) Measured thickness on dummy: 134.9 nm	1 test wafer
11	Nitride deposition – 2.LOCOS			**Bottom electrode wafer**
11.1	Si3N4 deposition	Nitride furnace – LPCVD 6” E3	Deposit nitride Recipe: 4nitdan Time: 20 min (@ 2.55nm/min) (tidligere: 57.5nm) Target thickness: 50 nm	Remember 1 test wafer

11.2	Measure thickness	Filmtek	Measured Thickness: 49.7 nm	
12 PolySi deposition – 2.LOCOS				**Bottom electrode wafer**
12.1	Polysilicon deposiiton	Polysilicon furnace – LPCVD 6” E2	Recipe: Poly Time: 13 min (@7.23 nm/min) (tidligere: 81 nm) Target thickness: 100 nm	Remember 1 test
12.2	Measure thickness	Filmtek	Measured Thickness: 83.0 nm	
13 Lithography 1.5 µm – 2.LOCOS				**Bottom electrode wafer**
13.1	Prepare + coat wafer	HMDS + Spin track	AZ nLOF 2020 negative resist Recipe: 1.5 µm with HMDS	Faster to use HMDS oven than Spin track with many wafers. Run 2 dummy wafers to ensure resist coverage (without HMDS)
13.2	Exposure	MA6 aligner	Hard contact Exposure time: 8.5 sec Mask: 2LOCOS	Due to the alignment marks alignment can be difficult
13.3	PEB + develop	TMAH UV developer	PEB 110C 60s + 60s puddle	
13.4	Inspection	Optical microscope	Check pattern, alignment marks and exposure test-structures	
14 Polysilicon mask etch (+backside strip) – 2.LOCOS				**Bottom electrode wafer**
14.1	Polysilicon etch	DRIE Pegasus	3 cycles of LF SOI recipe. Recipe name: Poly etch (20°C), use MACS	10 min O_2 clean before start to avoid leak errors
14.2	Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 30 min.	
14.3	Coat wafer	Spin track	Mir 701 positive resist Recipe: 1.5 µm with HMDS	
14.4	Polysilicon etch	Wet Poly Etch	Strip PolySi of backside for 2 min. Frontside is protected by resist.	Check that the resist is OK right after the etch, before the water rinse.
14.5	Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 30 min.	
15 Nitride wet etch – 2.LOCOS mask nitride				**Bottom electrode wafer**
15.1	Nitride etch	Wet bench	Wet silicon nitride etch H3PO4 @ 180°C (85Å/min) Time: ~6 min + a little over etch = 8 min	All wafers
15.2	7up	7up	Clean wafers after wet nitride etch due to possible contamination by potassium ions from people stripping nitride after KOH	
15.3	Inspection		Optical microscope	Check pattern and alignment marks
16 Polysilicon mask strip (front side) – 2.LOCOS				**Bottom electrode wafer**
16.1	Strip polysilicon mask	DRIE	2 cycles of LF SOI recipe. Recipe name: Poly etch (20°C)	
17 Oxidation - 2.LOCOS				**Bottom electrode wafer**
17.1	RCA	RCA bench	**HF only first time and half the time (15s)**	Remember test wafer

17.2	SiO2 wet oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Wet1100, time: 2h 00min 00s Target thickness: 906nm (on test wafer, remember 4% correct correction) Measured thickness on dummy: 862.3 nm	1 test wafer This corresponds to a post oxide thickness of: 0.897µm (as there already is an insulation oxide present) The time is corrected for 4%
17.3	Cavity height	Dektak	Measure cavity height on device wafer test structures	
17.4	Roughness measurement	AFM	On a bottom electrode wafer which is not to be bonded Scan size: 15 µm Scan rate: 0.0501 Hz	
18	Wafer bonding			**All wafers**
18.1	RCA	RCA bench	**HF only first time and for half the time (15s)**	SOI wafers + bottom electrode
18.2	Fusion bonding	EVG NIL	Recipe: CMUT Transport wafers in dedicated box from RCA. Minimize ambient exposure and handling time. Use RCA cleaned tweezers	
18.3	Annealing	Anneal-bond furnace (A2)	Recipe: Ann1100 Time: 1h 10min	
19	Handle layer and box oxide etch			**Bonded wafers**
19.1	Oxide etch	BHF	Etch rate: 78.19 nm/min Etch time: 15 min Thickness: 1 µm	
19.2	Si etch	ASE	Etch handle layer away recipe: (prototyping /)etchaway ~1h 05 min to etch 380 µm	Look through window to check during the etch and when etching is finished
19.3	Oxide etch	BHF	Etch box oxide away, Etch time: 15 min, check oxide is gone.	
19.4	RCA	RCA bench		
20	Lithography - Access to bottom electrodes			**Bonded wafers**
20.1	Prepare + coat wafer	Spin Coater Gamma UV	Mir 701 positive resist Recipe: 1.5 µm with HMDS	Be aware that another resist is used! Run 2 dummy wafers to ensure resist coverage (without HMDS)
20.2	Exposure	MA6 aligner	Hard contact Exposure time: 14 sec Mask: Bottom electrode	
20.3	PEB + develop	TMAH UV developer	Recipe: PEB 110C 60s + 60s puddle	
20.4	Inspection	Optical microscope	Check pattern and alignment marks	
21	Etch - Access to bottom electrodes			**Bonded wafers**
21.1	Etch Si	ASE	Etch all device layer Recipe: Shallolr (2 µm trench) Temp: 20°C Time: 00h01min09s (6 cycles)	

21.2 Oxide etch	AOE	Recipe: SiO2_res Time: 4 x 2 min (@1356nm oxide) Temp: 0°C	Wait 30 min for temp to stabilize. The wafer is removed from the camber between the cycles to avoid heating
21.3 Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 45 min.	
22 Metallization			**Bonded wafers**
22.1 Deposit Al+Ti	Alcatel	Target Ti thickness: 10nm Target Al thickness: 400nm	
23 Lithography – Metal electrode etch			**Bonded wafers**
23.1 Prepare + coat wafer	Spin track	AZ nLOF 2020 negative resist Recipe: 1.5 µm with HMDS	Run 2 dummy wafers to ensure resist coverage (without HMDS)
23.2 Exposure	KS aligner	Hard contact Exposure time: 15 sec Mask: Metal etch	
23.3 PEB + develop	TMAH UV developer	Recipe: PEB 110C 60s + 60s puddle	
23.4 Inspection	Optical microscope	Check pattern and alignment marks	
24 Etch metal electrode			**Bonded wafers**
24.1 Carrier wafer	Crystal bond555	Glue device wafers onto 6” dummy wafer using crystal bond. Should be soluble in water!	Buy 6” carrier wafer
24.2 Al etch	ICP metal	Recipe: Al etch Temp: 20°C Breakthrough: 20 s Etch time: 80 s	Check in microscope if all metal is gone
24.3 Si etch	ICP	Temp: 0°C Etch time: 20 min	To thin down the device layer
24.4 Detach carrier wafer	Water	Remove carrier wafer using DI water	
24.5 Strip resist	**Plasma Asher 1**	O2/N2: 400/70 ml/min, 1000 W, 30 min.	
25 Lithography – Plate etch			**Bonded wafers**
25.1 Prepare + coat wafer	Spin track	AZ nLOF 2020 negative resist Recipe: 1.5 µm with HMDS	Run 2 dummy wafers to ensure resist coverage (without HMDS)
25.2 Exposure	MA6 aligner	Hard contact Exposure time: 8.5 sec Mask: Plate etch	
25.3 PEB+ develop	TMAH UV developer	Recipe: PEB 110C 60s + 60s puddle	
25.4 Inspection	Optical microscope	Check pattern and alignment marks	
26 Plate etch			**Bonded wafers**
26.1 Carrier wafer	Crystal bond555	Glue device wafers onto 6” dummy wafer using crystal bond. Should be soluble in water!	
26.2 Etch Si	ICP	Etch all device layer Recipe: pxsimicrotry 3 Time: 20s Temp: 0°C	

26.3 Detach carrier wafer	Water	Remove carrier wafer using DI water	
26.4 Strip resist	Plasma Asher 1	O2/N2: 400/70 ml/min, 1000 W, 30 min.	
27 Dicing – Can be skipped			**Bonded wafers**
27.1 Coat wafers	Spin Coater Gamma UV	AZ 5214E Recipe: 40 µm without HMDS	This type of resist is easier to dissolve in acetone
27.2 Dice out chips	Disco Saw	Dicing out chips	Disco Saw

Contents

APPENDIX E

Process flow - G2

Process flow title				Revision
CMUT – 1xLOCOS NitridePlate v1				**V1**
DTU Danchip National Center for Micro- and Nanofabrication	Contact email migmol@nanotech.dtu.dk		Contact person Mathias Mølgaard	Contact phone 22 73 0936 0-58153
	Labmanager group MEMS(3313)	Batch name 1xLOCOS NitridePlate v1	Date of creation 11-Jul-16	Date of revision 11-Jul-16

Objective

Batch name: 1xLOCOS NitridePlate

The purpose of this project is to fabricate Capacitive Ultrasonic Transducers using LOCOS process

Substrates

Substrate	Orient.	Size	Doping/type	Polish	Thickness	Box	Device layer thickness	Purpose	#	Sample ID
Si	<100>	4”	p (boron)	SSP	525±25µm	-		Bottom electrode		ON405
Si + nitride	<100>	4”	???	DSP	350µm + 94nm (nitride)	-	-	Plate		ON517

Step	Heading	Equipment	Procedure	Comments
1	Preparation and protective oxidation			**Bottom electrode wafers**
1.1	Wafer selection	Wafer box	Use the dedicated vacuum tweezer across from the A-stack furnaces	ON405
1.2	RCA			Remember 1 dummy wafer
1.3	SiO2 wet oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Dry900, time: 27min Target thickness: 10nm Measured thickness on dummy: 10nm	Remember 1 dummy wafer
1.4	Measure thickness	Ellipsometer	Measured Thickness: nm	Measure on dummy wafer
2	Nitride deposition – 1.LOCOS			**Bottom electrode wafers**
2.1	Si3N4 deposition	Nitride furnace – LPCVD 4” B2	Deposit nitride Recipe: NITRIDE4 Time: 12:30min (@ 4.08nm/min) Target thickness: 50nm Measured thickness: 51nm	Remember 1 dummy
2.2	Measure thickness	Filmtek	Measured Thickness: 51nm	
3	PolySi deposition – 1.LOCOS			**Bottom electrode wafers**
3.1	Polysilicon deposiiton	Polysilicon furnace – LPCVD 4” B4	Recipe:Poly620 Time: 11min (@8.5nm/min) Target thickness: 100nm	Remember 1 dummy
3.2	Measure thickness	Filmtek	Measured Thickness: 93nm	
4	Lithography 1.5 µm – 1.LOCOS			**Bottom electrode wafer**
4.1	Prepare wafer	HMDS		Faster to use HMDS oven than Spin track with many wafers. Run 2 test wafers to ensure resist coverage
4.2	Coat wafer	Gamma	1.5 µm Mir 701 positive resist	
4.3	Exposure	MA6-2 aligner	Align to flat(first print). Hard contact Exposure time: 13s Mask: 1LOCOS	
4.4	PEB + Develop	TMAH UV developer	PEB 110C 60s + 60s puddle	
4.5	Inspection	Optical microscope	Check pattern, alignment marks and exposure test-structures	
5	Polysilicon mask etch – 1.LOCOS			**Bottom electrode wafer**
5.1	Polysilicon etch	ASE or **DRIE**	**DRIE**: 2 (1 could do it) cycle(s) of LF SOI recipe. Recipe name: Poly etch (20degC), use MACS ASE: 8 seconds using a continuous deep etch process (The same as for removing the handle)	

5.2	Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 45 min.	
6	Nitride wet etch – 1.LOCOS mask nitride			**Bottom electrode wafer**
6.1	Nitride etch	Wet bench	Wet silicon nitride etch H3PO4 @ 180 C (84Å/min) Time: ~7min + a little overetch = 8min	All wafers
6.2	7up	7up	Clean wafers after wet nitride etch due to possible contamination by potassium ions from people stripping nitride after KOH	
6.3	Inspection		Optical microscope	Check pattern and alignment marks
7	Polysilicon strip – frontside only			**Bottom electrode wafer**
7.1	Strip polysilicon mask	**DRIE**	**DRIE**: 2 cycles of LF SOI recipe. Recipe name: 100nmPolySi_SOI etch (20degC) ~~ASE: 8 seconds using a continuous deep etch process (The same as for removing the handle)~~	
8	Oxidation - 1.LOCOS			**Bottom electrode wafer**
8.1	RCA	RCA bench	**First HF dip: 15s. Second HF dip 0s.**	Remember dummy wafer
8.2	SiO2 wet oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Wet1100, time: 0h55min00s (53min) Target thickness: ~560nm (on test wafer) Measured thickness on dummy: 564nm	1 dummy wafer Note time in logbook
9	Wafer bonding			**All wafers**
9.1	RCA	RCA bench	**First HF dip: 15s. Second HF dip 0s.**	Si-Si3N4 + bottom electrode
9.2	Fusion bonding	New bonder	Recipe: CMUT_standard Transport wafers in dedicated box from RCA. Minimize ambient exposure and handling time. Use RCA cleaned tweezers	
9.3	~~Inspection~~	~~Infrared camera~~	~~Check for pre-anneal voids~~	
9.4	Annealing	Anneal-bond furnace	Recipe: Ann1100 Time: 1h 10min	
9.5	Inspection	PL mapper	Inspect post anneal voids	
10	Handle layer and box oxide etch			**Bonded wafers**
10.1	Nitride etch	ICP or **AOE**	ICP: ? **AOE**: 1min SiN_res (~435nm/min)	
10.2	Si etch	ASE then KOH	ASE etchaway15: 45min KOH3: ~69min	Maybe better with pegasus due to uniformity
10.3	7up	7up bench	10min	
11	Lithography - Access to bottom electrodes			**Bonded wafers**
11.1	Coat wafer	Gamma	Mir 701 positive resist Recipe: 1.5 µm with HMDS	
11.2	Exposure	MA6-2 aligner	Hard contact Exposure time: 14 sec Mask: Bottom electrode	
11.3	Develop	TMAH UV	Recipe: PEB 110C 60s + 60s puddle develop	

	developer		
11.4Inspection	Optical microscope	Check pattern and alignment marks	
12 Etch - Access to bottom electrodes			**Bonded wafers**
12.1Etch nitride-plate	AOE	Recipe: SiN_res Temp: 0C Time: 30s (~435nm/min)	
12.2Oxide etch	AOE	Recipe: SiO2_res Time: 3 min (~230nm/min) Temp: 0C	
12.3Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 45 min.	
13 Metallization + annealing			**Bonded wafers**
13.1Deposit Al	Alcatel	Target Al thickness: 100nm	
13.2Anneal Al	C4	Time: ~30min, Temp: 425C	
14 Lithography – Metal electrode etch			**Bonded wafers**
14.1Coat wafer	Gamma	Mir 701 positive resist Recipe: 1.5 µm with HMDS	
14.2Exposure	MA6-2 aligner	Hard contact Exposure time: 13 s Mask: Metal etch	
14.3Develop	TMAH UV developer	Recipe: PEB 110C 60s + 60s puddle develop	
14.4Inspection	Optical microscope	Check pattern and alignment marks	
15 Etch metal electrode			**Bonded wafers**
15.1Cure resist	Hot plate	2min at 120C	To prevent delamination and undercut in the wet Al etch
15.2Al etch	Wet Al etch	Solution: 1:2 (H20:H3PO4) Temp: 50C Time: 1min:30s (~100nm/min) Note when the metal is gone	Move the boat around in the solution in order to mix it before the etch and during the etch. **Very important to take the boat out approx. every 5s (or even more) in order to burst the H2 bubbles.**
15.3Rinse wafers	DI water	5min	
15.4Strip resist	Plasma Asher 1	O2/N2: 250/70 ml/min, 1000 W, 45 min.	
15.5Anneal	Furnace: Al anneal (C4)	30 min @ 475C	
16 Dicing			**Bonded wafers**
16.1Coat wafers	Gamma UV	2 x 4.2µm AZ5214E (1x is enough)	
16.2Dice out chips	Disco Saw	Dicing out chips	Disco Saw

Contents

APPENDIX F

Process flow - G3

Process flow title				Revision
CMUT – 2xLOCOS NitridePlate v1				**V1**
DTU Danchip National Center for Micro- and Nanofabrication	Contact email migmol@nanotech.dtu.dk		Contact person Mathias Mølgaard	Contact phone 22 73 0936 0-58153
	Labmanager group MEMS(3313)	Batch name 2xLOCOS NitridePlate v1	Date of creation 8-Sep-16	Date of revision 8-Sep-16

Objective

Batch name: 2xLOCOS NitridePlate

The purpose of this project is to fabricate Capacitive Ultrasonic Transducers using a double LOCOS process

Substrates

Substrate	Orient.	Size	Doping/type	Polish	Thickness	Box	Device layer thickness	Purpose	#	Sample ID
Si	<100>	4”	p (boron)	SSP	525±25µm	-		Bottom electrode		ON405
Si + nitride	<100>	4”	-	DSP	350µm + 100nm (nitride)	-	-	Plate		ON517

Step	Heading	Equipment	Procedure	Comments
1	**Preparation and protective oxidation – 1.LOCOS**			**Bottom electrode wafers**
1.1	Wafer selection	Wafer box	Use the dedicated vacuum tweezer across from the A-stack furnaces	ON405
1.2	RCA			Remember 1 dummy wafer
1.3	SiO2 dry oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Dry900, time: 27min Target thickness: 10nm Measured thickness on dummy: 10nm	Remember 1 dummy wafer
1.4	Measure thickness	Ellipsometer	Measured Thickness:	Measure on dummy wafer
2	**Nitride deposition – 1.LOCOS**			**Bottom electrode wafers**
2.1	Si3N4 deposition	Nitride furnace – LPCVD 4” B2	Deposit nitride Recipe: NITRIDE4 Time: 12:30min (@ 4.08nm/min) Target thickness: 50nm	Remember 1 dummy
2.2	Measure thickness	Filmtek	Measured Thickness:	
3	**Lithography 1.5 µm – 1.LOCOS**			**Bottom electrode wafer**
3.1	Prepare wafer	HMDS		Run 2 test wafers to ensure resist coverage
3.2	Coat wafer	Gamma	1.5 µm Mir 701 positive resist	
3.3	Exposure	MA6-2 aligner	Align to flat(first print). Hard contact Exposure time: 13s Mask: Locos Bumps	
3.4	PEB + Develop	TMAH UV developer	PEB 110C 60s + 60s puddle	
3.5	Inspection	Optical microscope	Check pattern, alignment marks and exposure test-structures	
4	**Nitride dry etch – mask nitride 1.LOCOS**			**Bottom electrode wafer**
4.1	Nitride etch	AOE	Etch nitride Recipe: Nitr_res Time: 1min Etch rate: 60nm/min	All wafers
4.2	Inspection	Optical microscope		Check pattern and alignment marks
4.3	Strip resist	Plasma Asher 2	O2/N2:400/70 ml/min, 1000W, 30 min	
5	**Oxidation - 1.LOCOS**			**Bottom electrode wafer**
5.1	RCA	RCA bench	**First HF dip: 15s. Second HF dip 0s.**	Remember dummy wafer
5.2	SiO2 wet oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers.	1 dummy wafer Note time in logbook

			Recipe: Wet1100, time: 31 min 30 s Target thickness: ~ 438 nm (on test wafer) Measured thickness on dummy:	
6	Strip oxide and nitride – 1.LOCOS			**Bottom electrode wafer**
6.1	Wet oxide etch	BHF – clean (across from HMDS oven)	Etch rate: ~75 nm/min Etch time: ~_min + over etch = 10 min	Adjust the time according to the oxidation time
6.2	Nitride etch	Wet bench	Wet silicon nitride etch H3PO4 @ 160°C (2.5 nm/min) Time: ~20 min + a little over etch = 30 min	
6.3	7up	7up		Clean wafers after wet nitride etch due to possible contamination by potassium ions from people stripping nitride after KOH
6.4	Wet oxide etch	BHF – clean (across from HMDS oven)	Etch rate: ~75 nm/min Etch time: 15 min	
6.5	Inspection	Dektak, AFM or Sensorfar	Measure the bump height Target height: 200 nm Measured height: 205 nm	
7	Oxidation – Insulation oxide – 2.LOCOS			**Bottom electrode wafers**
7.1	RCA			Remember 1 dummy wafer
7.2	SiO2 wet oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Dry900, time: 27min Target thickness: 10nm Measured thickness on dummy:	Remember 1 dummy wafer
7.3	Measure thickness	Ellipsometer	Measured Thickness:	Measure on dummy wafer
8	Nitride deposition – 2.LOCOS			**Bottom electrode wafers**
8.1	Si3N4 deposition	Nitride furnace – LPCVD 4” B2	Deposit nitride Recipe: NITRIDE4 Time: 12:30min (@ 4.08nm/min) Target thickness: 50nm Measured thickness:	Remember 1 dummy
8.2	Measure thickness	Ellipsiometer	Measured Thickness: __nm	
9	PolySi deposition – 2.LOCOS			**Bottom electrode wafers**
9.1	Polysilicon deposiiton	Polysilicon furnace – LPCVD 4” B4	Recipe:Poly620 Time: 11min (@8.5nm/min) Target thickness: 100nm	Remember 1 dummy
9.2	Measure thickness	Filmtek	Measured Thickness:	
10	Lithography 1.5 µm – 2.LOCOS			**Bottom electrode wafer**

10.1	Prepare wafer	HMDS		Faster to use HMDS oven than Spin track with many wafers. Run 2 test wafers to ensure resist coverage
10.2	Coat wafer	Gamma	1.5 µm Mir 701 positive resist	
10.3	Exposure	MA6-2 aligner	Align. Hard contact Exposure time: 13s Mask: Cavities	
10.4	PEB + Develop	TMAH UV developer	PEB 110C 60s + 60s puddle	
10.5	Inspection	Optical microscope	Check pattern, alignment marks and exposure test-structures	
11	Polysilicon mask etch – 2.LOCOS			**Bottom electrode wafer**
11.1	Polysilicon etch	**Pegasus**	**DRIE**: 2 (1 could do it) cycle(s) of LF SOI recipe. Recipe name: Poly etch (0°C) ~~ASE: 8 seconds using a continuous deep etch process (The same as for removing the handle)~~	
11.2	Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 45 min.	
12	Nitride wet etch – mask nitride 2.LOCOS			**Bottom electrode wafer**
12.1	Nitride etch	Wet bench	Wet silicon nitride etch H3PO4 @ 160 C (2.5nm/min) Time: ~20min + a little overetch = 30min	All wafers
12.2	7up	7up	Clean wafers after wet nitride etch due to possible contamination by potassium ions from people stripping nitride after KOH	
12.3	Inspection		Optical microscope	Check pattern and alignment marks
13	Polysilicon strip (frontside) – 2.LOCOS			**Bottom electrode wafer**
13.1	Strip polysilicon mask	**Pegasus**	**DRIE**: 2 cycles of LF SOI recipe. Recipe name: 100nmPolySi_SOI etch (20degC) ~~ASE: 8 seconds using a continuous deep etch process (The same as for removing the handle)~~	
14	Oxidation - 2.LOCOS			**Bottom electrode wafer**
14.1	RCA	RCA bench	**First HF dip: 15s. Second HF dip 0s.**	Remember dummy wafer
14.2	SiO2 wet oxidation For 100nm gap	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Wet1100, time: 1 h 6 min Target thickness: 643 nm (on test wafer) Measured thickness on dummy:	1 dummy wafer Note time in logbook
14.3	SiO2 wet oxidation For 50 nm gap	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Wet1100, time: 49 min 30 sek	1 dummy wafer Note time in logbook

		Target thickness: 555 nm (on test wafer) Measured thickness on dummy:	
15 Wafer bonding			**All wafers**
15.1 RCA	RCA bench	**First HF dip: 15s. Second HF dip 0s.**	bottom electrode
15.2 Fusion bonding	Süss bonder	Recipe: CMUT_standard Transport wafers in dedicated box from RCA. Minimize ambient exposure and handling time. Use RCA cleaned tweezers	Bottom electrode + Top plate straight from the furnace
15.3 ~~Inspection~~	~~Infrared camera~~	~~Check for pre-anneal voids~~	
15.4 Annealing	Anneal-bond furnace	Recipe: Ann1100 Time: 1h 10min	
15.5 Inspection	PL mapper	Inspect post anneal voids	
16 Handle layer and box oxide etch			**Bonded wafers**
16.1 Nitride etch	ASE	ASE: 2min EtchA15	Should kill the oxynitride and remove nitride. Check for color change
16.2 Oxide etch	BHF	Etch rate: ~75 nm/min Etch time: 21 min	Only if Oxide is grown beneath the nitride
16.3 Si etch	ASE	ASE etchaway15: approx 55min. Look for circle.	Approx. 35 if KOH after
16.4 Si Etch	KOH		
16.5 7UP	7UP	Clean wafers after wet nitride etch due to possible contamination by potassium ions from people stripping nitride after KOH	Only if KOH is done before
16.6 Oxide etch	BHF	Etch rate: ~75 nm/min Etch time: 21 min	Only if Oxide is grown beneath the nitride
17 Lithography - Access to bottom electrodes			**Bonded wafers**
17.1 Coat wafer	Gamma	Mir 701 positive resist Recipe: 1.5 µm with HMDS	
17.2 Exposure	MA6-2 aligner	Hard contact Exposure time: 13 sec Mask: Bottom electrode	
17.3 Develop	TMAH UV developer	Recipe: PEB 110C 60s + 60s puddle develop	
17.4 Inspection	Optical microscope	Check pattern and alignment marks	
18 Etch - Access to bottom electrodes			**Bonded wafers**
18.1 Etch nitride-plate	AOE	Recipe: SiN_res Temp: 0C Time: 30s (~435nm/min)	
18.2 Oxide etch	AOE	Recipe: SiO2_res Time: 3 min (~230nm/min) Temp: 0C	This etch should also open to the alignment marks (feature of the mask)
18.3 Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 45 min.	
19 Metallization + annealing			**Bonded wafers**
19.1 Deposit Al + anneal	Alcatel + annealing(furnace?)	Target Al thickness: 100nm	
20 Lithography – Metal electrode etch			**Bonded wafers**

20.1Coat wafer	Gamma	Mir 701 positive resist Recipe: 1.5 µm with HMDS	
20.2Exposure	MA6-2 aligner	Hard contact Exposure time: 13 s Mask: Metal etch	
20.3Develop	TMAH UV developer	Recipe: PEB 110C 60s + 60s puddle develop	
20.4Inspection	Optical microscope	Check pattern and alignment marks	
21 Etch metal electrode			**Bonded wafers**
21.1Cure resist	Hot plate	2min at 120C	To prevent delamination and undercut in the wet Al etch
21.2Al etch	Wet Al etch	Solution: 1:2 (H20:H3PO4) Temp: 50°C Time: 1min:30s (~100nm/min) Note when the metal is gone	Move the boat around in the solution in order to mix it before the etch and during the etch. **Very important to take the boat out approx. every 5s (or even more) in order to burst the H2 bubbles.**
21.3Rinse wafers	DI water	5min	
21.4Strip resist	Plasma Asher 1	O2/N2: 250/70 ml/min, 1000 W, 45 min.	
21.5Anneal	Furnace: Al anneal (C4)	30 min @ 475C	
22 Dicing			**Bonded wafers**
22.1Coat wafers	Gamma UV	1 x 4.2µm AZ5214E	
22.2Dice out chips	Disco Saw	Dicing out chips	Disco Saw

APPENDIX G

Process flow - G4

Process flow title				Revision
CMUT – 1xLOCOS NitridePlate v1				**V1**
DTU Danchip National Center for Micro- and Nanofabrication	Contact email migmol@nanotech.dtu.dk		Contact person Mathias Mølgaard	Contact phone 22 73 0936 0-58153
	Labmanager group MEMS(3313)	Batch name 1xLOCOS NitridePlate v1	Date of creation 11-Jul-16	Date of revision 11-Jul-16

Objective

Batch name: 1xLOCOS NitridePlate

The purpose of this project is to fabricate Capacitive Ultrasonic Transducers using LOCOS process

Substrates

Substrate	Orient.	Size	Doping/type	Polish	Thickness	Box	Device layer thickness	Purpose	#	Sample ID
Si	<100>	4”	p (boron)	SSP	525±25µm	-		Bottom electrode		ON405
Si + nitride	<100>	4”	???	DSP	350µm + 50nm (nitride)	-	-	Plate		ON517

Step	Heading	Equipment	Procedure	Comments
1	Preparation and oxidation			**Bottom electrode wafers**
1.1	Wafer selection	Wafer box	Use the dedicated vacuum tweezer across from the A-stack furnaces	ON405
1.2	RCA			Remember 1 dummy wafer
1.3	SiO2 wet oxidation	Phosphorus Drive-in (A3)	Place a test wafer in the center of the boat and place device wafers and eg. test wafers equally distributed on each side of the test wafer. No spacing between wafers. Recipe: Wet1050, time: 20min Target thickness: 10nm Measured thickness on dummy: 225nm	Remember 1 dummy wafer
1.4	Measure thickness	Ellipsometer	Measured Thickness: 225nm	Measure on dummy wafer
2	Lithography 1.5 µm – Cavities			**Bottom electrode wafer**
2.1	Prepare wafer	HMDS		Faster to use HMDS oven than Spin track with many wafers. Run 2 test wafers to ensure resist coverage
2.2	Coat wafer	Gamma	1.5 µm Mir 701 positive resist	
2.3	Exposure	MA6-2 aligner	Align to flat(first print). Hard contact Exposure time: 13s Mask: Cavity	
2.4	PEB + Develop	TMAH UV developer	PEB 110C 60s + 60s puddle	
2.5	Inspection	Optical microscope	Check pattern, alignment marks and exposure test-structures	
3	Dry oxide etch - Cavities			**Bottom electrode wafer**
3.1	Oxide etch	AOE	Recipe name: SiO2_res at 0°C Time: 2min	
3.2	Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 45 min.	
4	Nitride deposition – Insulation			**Bottom electrode wafer**
4.1	RCA			
4.2	Si3N4 deposition	Nitride furnace – LPCVD 4” B2	Deposit nitride Recipe: NITRIDE4 Time: 6min Target thickness: 20nm Measured thickness: 27nm	All wafers
5	Wafer bonding			**All wafers**
5.1	RCA	RCA bench		
5.2	Fusion bonding	New bonder	Recipe: CMUT_standard Transport wafers in dedicated box from RCA. Minimize ambient exposure and handling time. Use RCA cleaned tweezers	
5.3	Annealing	Anneal-bond	Recipe: Ann1100	

	furnace	Time: 1h 10min	
5.4 Inspection	PL mapper	Inspect post anneal voids	
6 Etch handle Si and top nitride layer			**Bonded wafers**
6.1 Si3N4 and Si dry etch	ASE	Recipe: Etcha15 Time: 47min	The Si3N4 layer is gone after about 2 min and about 2/3 of the Si handle layer is etched at the end.
6.2 Selective Si etch	KOH	Time: 1h:16min	Maybe better with pegasus due to uniformity
6.3 7up	7up bench	10min	
7 Lithography - Access to bottom electrodes			**Bonded wafers**
7.1 Coat wafer	Gamma	Mir 701 positive resist Recipe: 1.5 µm with HMDS	
7.2 Exposure	MA6-2 aligner	Hard contact Exposure time: 14 sec Mask: Bottom electrode	
7.3 Develop	TMAH UV developer	Recipe: PEB 110C 60s + 60s puddle develop	
7.4 Inspection	Optical microscope	Check pattern and alignment marks	
8 Etch - Access to bottom electrodes			**Bonded wafers**
8.1 Etch nitride-plate	ASE	Recipe: Etcha15 Time: 1min	Using ASE instead of AOE due to clamping issues
8.2 Oxide etch	BHF	Time: 5min:30s	
8.3 Strip resist	Plasma Asher 2	O2/N2: 400/70 ml/min, 1000 W, 45 min.	
9 Metallization			**Bonded wafers**
9.1 Deposit Al	Alcatel	Target Al thickness: 50nm	
10 Lithography – Metal electrode etch			**Bonded wafers**
10.1Coat wafer	Gamma	Mir 701 positive resist Recipe: 1.5 µm with HMDS	
10.2Exposure	MA6-2 aligner	Hard contact Exposure time: 13 s Mask: Metal etch	
10.3Develop	TMAH UV developer	Recipe: PEB 110C 60s + 60s puddle develop	
10.4Inspection	Optical microscope	Check pattern and alignment marks	
11 Etch metal electrode			**Bonded wafers**
11.1Cure resist	Hot plate	2min at 120C	To prevent delamination and undercut in the wet Al etch
11.2Al etch	Wet Al etch	Solution: 1:2 (H20:H3PO4) Temp: 25C Time: 7min Note when the metal is gone	Move the boat around in the solution in order to mix it before the etch and during the etch. **Very important to take the boat out approx. every 5s (or even more) in order to burst the H2 bubbles.**

11.3 Rinse wafers	DI water	5min	
11.4 Strip resist	Plasma Asher 1	O2/N2: 250/70 ml/min, 1000 W, 45 min.	
11.5 Anneal the Al	Furnace: Al anneal (C4)	30 min @ 475C	Increases adhesion
12 Dicing			**Bonded wafers**
12.1 Coat wafers	Gamma UV	2 x 4.2µm AZ5214E (1x is enough)	
12.2 Dice out chips	Disco Saw	Dicing out chips	Disco Saw

Contents

Published by:
DTU Nanotech
Department of Micro- and Nanotechnology
Technical University of Denmark
Ørsteds Plads, building 345C
DK-2800 Kgs. Lyngby