Force Sensing in Robotic Surgery

A force sensing strategy for 3 mm shaft actuated tip articulated instruments in RAMIS

M. Suurveld

TUDelft

Force Sensing in Robotic Surgery

A force sensing strategy for 3 mm shaft actuated tip articulated instruments in RAMIS

by

M. Suurveld

to obtain the degree of Master of Science
in Mechanical Engineering and Biomedical Engineering
at the Delft University of Technology,
to be defended publicly on Tuesday October 4, 2022 at 12:30 PM.

This thesis is confidential and cannot be made public until September 20, 2025.

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TUDelft

Acknowledgements

I would like to express my gratitude to my supervisors Dr. Ir. Hans Goosen, Dr. Ir. Tim Horeman and Ir. Tomas Lenssen for giving me the opportunity to work on this interesting topic and for their guidance. Special thanks go to Tomas who, as the daily supervisor, has always been available for questions and feedback. I am also grateful to Ir. Jo Spronck, who started as one of my supervisors but is currently enjoying retirement, for his help in steering the project at its early stages.

In the latter part of the thesis period, when it was possible to go back to university, it was a group of fellow regular coffee-drinkers that helped motivate me for the final push to finish the thesis. Every break, at exactly $12:30$, they were available to put my mind of the thesis or to complain to about code or experiments not working. For this I am thankful.

I also want to express my deepest gratitude towards my family, friends and girlfriend for their continued support throughout the process. They also helped by giving me pressure and motivation to keep going by continually asking when I would be done.

Finally, it would be remiss not to acknowledge my colleagues at the Haga Hospital, where I have been working part-time, for their flexibility in allowing me to take the time I needed to finish this thesis.

M. Suurveld
Delft, September 2022

Summary

Over the past decades Minimally Invasive Surgery (MIS) has become a standard in abdominal surgery. Due to limited dexterity and bad ergonomics the field of MIS is shifting towards Robot-Assisted MIS (RAMIS). Simultaneously, the field is using increasingly small and steerable instruments to further reduce invasiveness. Surgeons now work with fragile instruments and tissues and are not able to feel the forces they apply. This can lead to unwanted damage to patient and instrument. In an effort to reduce the risk of accidental damage this work presents a force sensing strategy that is able to measure tip forces in RAMIS using 3mm shaft actuated tip articulated instruments. The tip force data can be used for force limiting or force feedback to the surgeon.

Analysis of the problem revealed that the trocar forces are large enough to significantly distort the measurement of the tip force and damage patient and instrument. The requirements imposed for the sensor to be reusable through autoclave sterilization are also stringent. Most solutions proposed in literature are either invasive and expensive or fail to effectively account for the trocar forces. Reusability was often neglected for both cases. The subject was either not mentioned or the proposed solutions were stated to be single-use regardless of cost. Consequently, a sensing strategy has been sought that is able to provide a solution that is both accurate and affordable by being reusable.

Because the trocar forces be sufficiently accurate the trocar forces have to be accounted for or eliminated. A solution was found which is able to eliminate the influence of the trocar forces by connecting the trocar to the same frame as the instrument and measuring the bending in the instrument due to the applied load. This solution was preferred over solutions that required multiple or invasive and expensive sensors. The sensors used to measure the bending in the instrument are Hall sensors. These were selected because of their good sensitivity, low cost and complexity and sterilizability. To improve sensitivity and eliminate the influence of temperature changes and parasitic displacements four Hall sensors are used which are placed in a differential configuration. To provide a package that is easy to produce, cheap and sterilizable a layer of soft silicone rubber is used as a transduction element.

A prototype was constructed in order to validate the compliance of the design with the requirements. A representative experimental setup was constructed for calibration and to compare the output of the prototype sensor to a reference sensor. The performance results show that the sensor is able to reach the required range of 5 N with a bandwidth of 120 Hz. The calibrated sensor has a maximum relative error less than 8.8% and root mean square error less than 3.4% and is able to reject parasitic displacements. The reusability of the sensor was also verified through several autoclave sterilization cycles. After the first cycle the sensor still functions though with a decreased sensitivity and increased noise. Both changes in performance remain stable after subsequent sterilization cycles.

In conclusion, this work presents and describes a sensing strategy that is able to accurately and affordably measure the forces on the tip during RAMIS. It meets almost all requirements with room for improvement on the others. The sensor and strategy provide ample reason and opportunity for further research and development.

Contents

List of Figures

List of Tables

1

Introduction

Over the past decades Minimally Invasive Surgery (MIS), also known as endoscopy, laparoscopy or keyhole surgery, has become the preferred approach to abdominal surgery because of the reduced invasiveness compared to conventional, open surgery. Advantages include decreased tissue trauma, blood loss, post-operative pain and risk of infection as well as a shorter hospital stay. As depicted in Figure 1.1, MIS is performed with long slender instruments generally $5-10$ mm in diameter. The instruments are inserted into the inflated abdomen through small incision. The incisions are fitted with a trocar to keep the inflation gases from escaping. An additional incision is required to pass the camera, called an endoscope, to gain vision of the surgical field [1, 2].
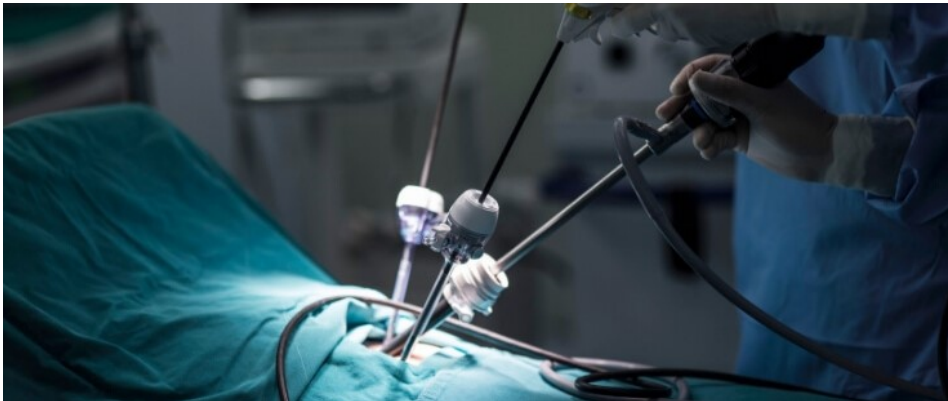
Figure 1.1: Laparoscopic surgery is performed with long slender instruments through small incisions in the abdomen. Image taken from `Surgical-Solutions.com`.

Despite all the advantages of MIS the use of long slender instruments also has some downsides. The long instruments and the awkward body positioning involved in maneuvering them cause severe ergonomic problems for the surgeon [3]. Especially during longer surgeries this is problematic. This and the promise of tele-operations has lead to the development of Robot-Assisted Minimally Invasive Surgery (RAMIS) [3]. In RAMIS the surgeon is removed from the operating table and instead sits behind a controller from which she steers robot arms that control the instruments and endoscope.

RAMIS provides several other advantages besides improved ergonomics. The system is able to reject hand tremors and remove the fulcrum effect experienced in MIS caused by rotating around the abdomen. Additionally, robotic systems have been able to improve dexterity by providing extra degrees of freedom with steerable instruments. These instruments are able to pitch and yaw the grasper tip in order to approach the tissue from the optimal angle [3]. Most steerable instruments are actuated by pulling on several cables. Figure 1.3a shows an example of such a cable-driven steerable instrument. The braided cables and mechanism with hinges and pulleys in such instruments are difficult to clean.
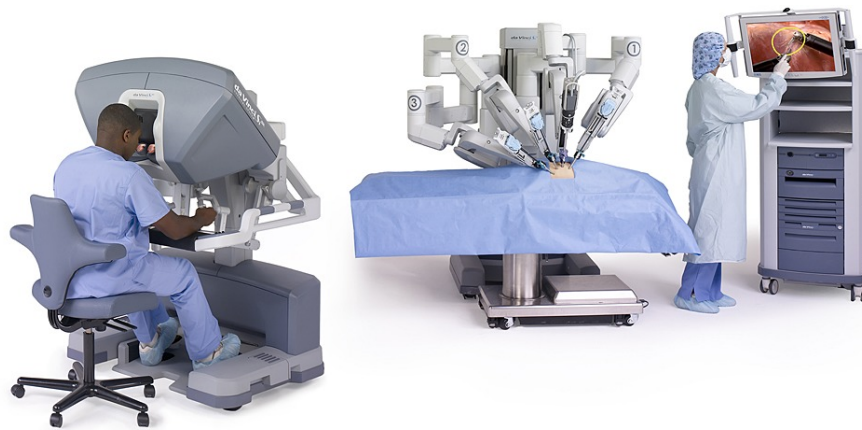
Figure 1.2: Da Vinci robotic surgery platform by surgical intuitive. Image taken from [3].

Therefore, most steerable instruments are only semi-reusable which makes them relatively expensive to use [4]. As a solution to this problem a new type of steerable instrument has been proposed. The so-called Shaft Actuated, Tip Articulated instruments (SATA), shown in Figure 1.3b, consist of concentric shafts and are steered by relative translation and rotation of these shafts [4]. SATA instruments are able to provide a yawing, rotating and grasping motion of the gasper tip. Due to their simple construction they are easily taken apart and cleaned. As a result these instruments are both modular and fully reusable.

(a)   (b)
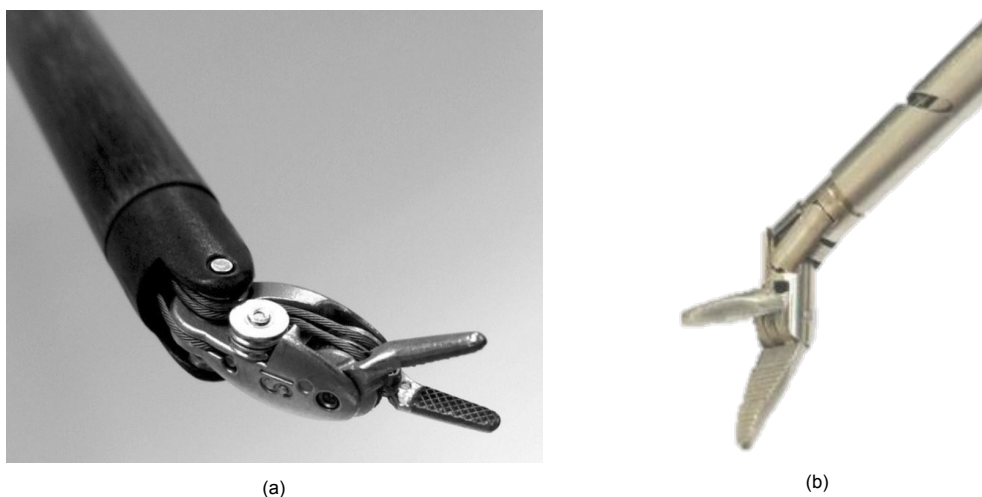
Figure 1.3: (a) Cable driven EndoWrist grasper used with the da Vinci robotic surgery platform. Image taken from [1]. (b) shaft actuated instrument. Image adapted from [4].

While the design proposed by Hardon et al. [4] is intended for conventional MIS, SATA instruments can also be used in RAMIS. A specific drive unit has been designed to actuate the instrument in all degrees of freedom. The instrument with drive unit are meant to be used on a new robotic surgery system being developed in the MISIT group at the TU Delft. Like the instrument and driver unit the robotic system is designed to be modular, reusable and cost-effective.

A second problem with the long and slender MIS-instruments is that they have limited stiffness and strength in bending, as well as fragile tips. This is especially true for the increasingly popular 3mm instruments which offer a further reduced invasiveness. Moreover, because of the way they operate, the SATA instruments no longer function properly after a permanent bend is introduced. The likelihood of this problem occurring is exacerbated by the fact that in RAMIS the surgeon is not able to feel the forces acting on the instrument tip. As a result the surgeon might not be able to prevent damage to
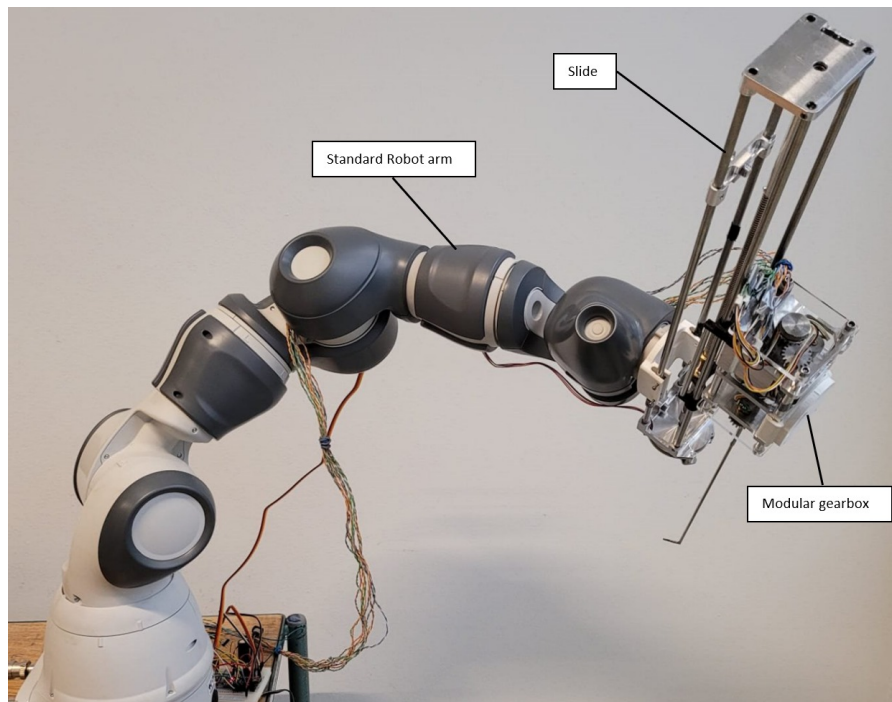
Figure 1.4: Robotic surgery platform under development by the TU Delft MISIT group.

the smaller, more fragile instruments. Breaking or bending an instrument during surgery and inside the patient can be dangerous and costly and must therefore be prevented. The lack of force sensation can also lead to unintentional tearing of internal tissues or the incision in the abdomen. Damage to both patient and instrument can be prevented by automatically limiting the force or through force feedback to the surgeon. This work will not focus on the implementation of force limiting or force feedback. Instead, it provides a study into a force sensing strategy that can be used to measure tip forces on 3mm shaft actuated instruments during RAMIS accurately and affordably. The strategy is meant to be suitable for force feedback in the future.

By analyzing the kinematics, mechanics and deformation, Chapter 2 explains that when measuring forces during surgery one has to account for the forces that occur at the trocar. Moreover, in order to provide a cost-effective solution the sensing system has to be cheap and reusable. A lot of research has already been performed on the topic of force measurements in (RA)MIS. A summary of the literature review performed on this topic is given in Chapter 3. The literature review identifies either a lack of accuracy or reusability and affordability in previously proposed sensing solutions. None of the solutions provided in literature are able to account for the trocar forces while also being affordable and reusable. Tackling these shortcomings has consequently become the premise of this work. By taking inspiration from several sources found in literature this work presents a sensing system that has been designed to meet all three requirements. Chapter 4 outlines the three-step approach used to come to a solution. First a global solution is found to account for the trocar forces by fixing the trocar to the robot arm and measuring the forces between instrument and trocar. Next, it is decided to use Hall sensors with a silicone layer as transduction element. The chapter ends by discussing the detailed design choices. The prototype system that was constructed, calibrated and tested is described in chapter 5. After discussing the value and limitations of this work in Chapter 6 this work ends with concluding remarks and recommendations for future research in Chapters 7 and 8.

2

Analysis

This chapter will discuss the theoretical background needed to understand the complexities of measuring forces in (RA)MIS in an usable and affordable way. It starts with a general description of the kinematics of RAMIS. This is followed by an analysis of the forces acting on instruments during RAMIS. In the next section, an analytical estimation of the instrument deformation in several loading conditions is presented. The final section discusses the cleaning and sterilization process and the implications on sensor design for a reusable sensor.

2.1. Kinematic Analysis

As explained in the introduction, MIS is performed with long slender instruments. The instrument is supported somewhere along its length by the trocar fitted in the abdominal incision. On one end the instrument is either held by the surgeon, or attached to a robotic arm in the case of RAMIS. On the other end the instrument ends in a grasper that is used to manipulate tissue. The grasper end is commonly referred to as the tip.

A non-steerable MIS instrument as depicted in Figure 2.1 has four kinematic Degrees of Freedom (DoF) excluding the grasping motion. One rotation around the longitudinal axis (γ), two rotations (α and β) pivoting around the incision point, and a translation along the longitudinal axis (t). A steerable MIS instrument has hinges located near the instrument tip that allow for rotations around the pitch and jaw axis of the tip. This brings the number of DoF for steerable instruments to six.
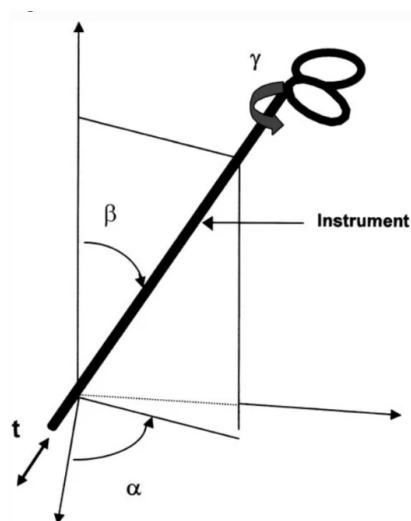
Figure 2.1: Degrees of freedom of a non-steerable instrument. Image taken from [5]

It is important to note that the incision point around which rotations α and β occur does not act as a normal spherical joint which would fix the instrument in the abdominal plane. Movement and loading from the instrument at this point should be limited to protect the incision. During conventional surgery the forces are limited by the surgeon who is able to roughly feel the amount of force applied to the incision and adjust the movements to rotate around this point. During RAMIS unwanted movement and loading at the incision can be prevented by implementing a Remote Center of Motion (RCM) [6]. The RCM ensures that the robot arm automatically rotates around a specific point in space; the incision. The RCM can be implemented in two distinct ways; passively and actively. Passive solutions use a robot arm designed such that all axes of rotations intersect at the same point. This point should be aligned with the incision point to prevent movement and forces at the incision. An example of such a robot is shown in Figure 2.2a. For robot arms that do not have intersecting axes of rotation, like the one in Figure 2.2b, it is possible to implement the RCM actively. In the active implementation the robot arm is programmed, through kinematic models in the software, to always move around a fixed point in space [6].
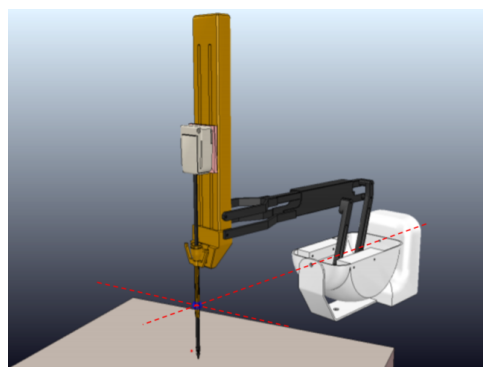
(a)

(b)

Figure 2.2: (a) Robot design with passive RCM. Image adapted from [7]. (b) 7-DoF serial robot without passive RCM. Image taken from [6]

In practice however, the incision is not completely fixed in space. It moves up and down slightly due to respiratory movement of the abdomen. It can also shift in other directions when the patient is moved or moves, for instance when only local anesthetics are used. For both conventional surgery and RAMIS the misalignment of incision and RCM will lead to unwanted loading on the incision. This loading can result in tearing of the incision and can also contribute to damaging the instrument. While the surgeon is generally able to notice the added force and adjust, the robotic system is unaware and the surgeon doesn't have any feedback. With an active RCM it is possible to continuously adjust the RCM position to align with the incision point. This requires force and position measurements to detect and correct the misalignment. For passive RCM systems this is more difficult because it is not as simple to move the RCM after the robot has been put in place.

2.2. Mechanical Analysis

Before analyzing the force components acting on the instrument a number of reference frames have to be defined. The first frame lies at the end of the robot arm at the interface with the instrument. Since this is usually the location of the instrument driver that steers the instrument it is referred to as the driver frame. For MIS instruments, like in Figure 2.3, this would be referred to as the handle frame. It has the z-axis pointing upwards along the longitudinal axis of the instrument and the y-axis pointing upward with respect to the driver. The trocar frame is oriented in the same way as the driver frame but the x- and y-axes are located on the abdominal plane when the instrument is perpendicular to the abdomen. The final reference frame is located at the grasping joint at the tip. It has the z-axis pointing

upwards along the longitudinal axis and the y-axis perpendicular to the grasping surface when the jaws are closed. This reference frame is referred to as the tip frame. The three reference frames are shown in Figure 2.3 for a MIS instrument. These frames can be considered a standard throughout literature.
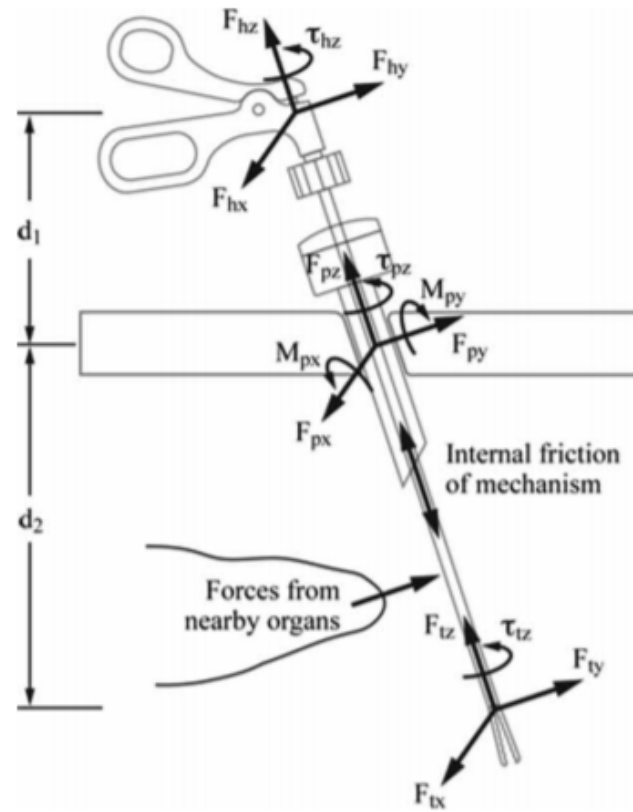
Figure 2.3: Forces acting on instrument during (RA)MIS. Image taken from [2].

As Figure 2.3 shows, there are generally three locations where forces are applied to the instrument: at the end of the robot arm, in the case of RAMIS, at the tip and at the trocar. It is also possible that organs exert a force on the instrument shaft somewhere between the trocar and tip. Since the shaft is not used to manipulate organs or tissues these forces do not occur frequently. Moreover, when they do occur the forces will be applied with a shorter moment arm. Since both impact and occurrence are negligible these forces will not be considered. Also the grasping force is omitted in this analysis because it is not likely to cause the instrument to break and it is decoupled from the other forces [2].

For the non-steerable instrument shown in Figure 2.3 there are only four force components occurring at the tip and handle. These forces correspond with the four DoF that the instrument has. They include three translational forces along the Cartesian axes (F_x,F_y and F_z) and a rotational force (torque) around the z-axis (τ_z). Two studies have found that these tissue interaction forces range from 0 to about 10 N and torques from 0 to 0.1 Nm during MIS [5, 8]. With steerable instruments two additional torques result from the ability to pitch and jaw the tip. These torques (M_x and M_y) align with the x- and y-axis respectively.

An additional set of forces can occur at the trocar. There are three different factors that can result in significant forces occurring at the trocar. These factors include movement, pivoting or tilting and (mis-)alignment. Moving the instrument along and rotating around the z-axis will result in corresponding forces F_{pz} and τ_{pz} due to friction between the instruments and the seals in the trocar. The linear force due to this friction can exceed 3 N [5]. When the instrument is rotated along the α and β directions described in Section 2.1 torques in the abdominal plane occur due to deformation of the abdomen. These

torques (M_{px} and M_{py}) can reach a magnitude of 0.7 Nm [5]. The final set of forces occurs when the incision point and RCM are not aligned. When this misalignment occurs the robot arm will effectively be pulling on the incision in the abdominal plane. This results in the forces F_{px} and F_{py}. The magnitude of these forces is generally unknown. However, the stiffness of the abdomen is about 0.5 N/mm in the anterior abdomen and 1 N/mm in the lateral abdomen [9] so a relatively small misalignment can lead to significant forces. Naturally, this stiffness will vary from location to location and patient to patient.

The forces at the instrument tip are the most relevant towards the goal of protecting the instrument and tissue from damage. Therefore, each possible sensing strategy proposed in this work is aimed at measuring these forces accurately. However, not all six force components occurring at the tip of steerable instruments are equally relevant. Since the instrument is weakest in bending the forces F_{tx} and F_{ty} are most important. Additionally, the tissue is most likely to tear when subjected to linear pulling forces (F_{tx}, F_{ty} and F_{tz}). It can thus be concluded that the three translational forces, especially F_{tx} and F_{ty}, are the main interest in preventing instrument and tissue damage.

2.3. Deformation analysis

In the previous section it was discussed which loads can occur on the instrument and under which conditions. This section will investigate how the instrument deforms under these loading conditions and what influence the trocar forces have on deformation.

The main goal of this analysis is to gain a basic insight into the forces and deformations and compare different loading conditions. Therefore, a simplified model of the system can be used to analytically calculate the deformation. As shown in Figure 2.4 the trocar is omitted in this model and the resistance to misalignment of the abdomen is represented by linear springs with a stiffness of 0.5 N/mm for a total stiffness of 1 N/mm [9]. For this model a pure moment of 0.3 Nm (Mpz) is also added at the RCM to simulate the effect of a rotation of the instrument with respect to the abdomen. A force F is applied to the tip. The instrument is modeled as a solid 3 mm 304 stainless steel rod with associated stiffness EI and a length L of 300 mm. In order to simplify the solution a different reference frame from those defined before is used. The reference frame is shown in Figure 2.4.
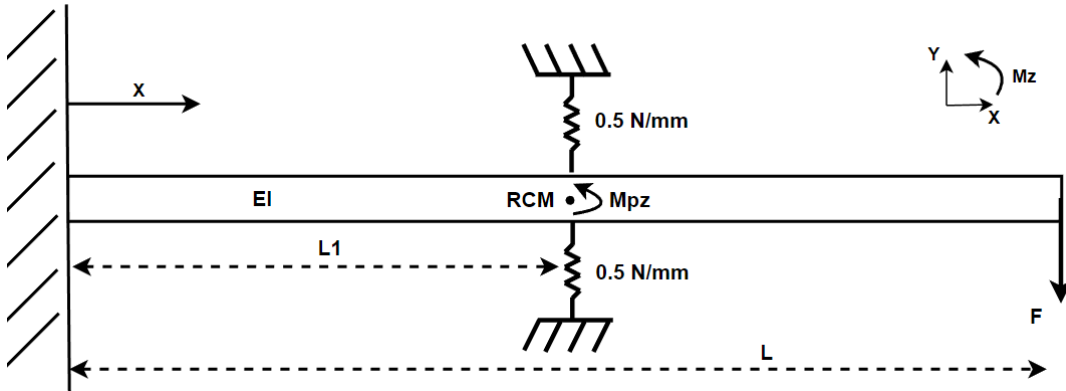
Figure 2.4: Simplified model of the instrument system and forces.

This problem is solved through the principle of superposition. The total deformation of the instrument is determined by summing the y-displacements due to each of the loads. The first includes only the tip force. The second includes the spring force and the third includes only the moment. While the sum of the displacements leads to a relatively large displacement, geometric non-linearity was not included in the solution. This is acceptable because the expected axial displacement of the tip is very small compared to the overall length. Moreover, the tip is not constrained which means that no stresses will be introduced by the axial displacement. Consequently, the displacement for each case can be found by using standard solutions for linear mechanics. The equations describing the y-displacements $V_1(x)$, $V_2(x)$, $V_3(x)$ and $V_{tot}(x)$ are given below.

$$V_1(x) = -\frac{Fx^2}{6EI}(3L - x) \qquad\qquad 0 \leq x \leq L \qquad\qquad (2.1)$$

$$V_{2a}(x) = \frac{KV_{RCM}x^2}{6EI}(3L_1 - x) \qquad\qquad 0 \leq x \leq L_1 \qquad\qquad (2.2)$$

$$V_{2b}(x) = \frac{KV_{RCM}L_1^2}{6EI}(3x - L_1) \qquad\qquad L_1 \leq x \leq L \qquad\qquad (2.3)$$

$$V_{3a}(x) = \frac{M_{pz}x^2}{2EI} \qquad\qquad 0 \leq x \leq L_1 \qquad\qquad (2.4)$$

$$V_{3b}(x) = \frac{M_{pz}L_1}{2EI}(2x - L_1) \qquad\qquad L_1 \leq x \leq L \qquad\qquad (2.5)$$

$$V_{tot}(x) = V_1(x) + V_2(x) + V_3(x) \qquad\qquad 0 \leq x \leq L \qquad\qquad (2.6)$$

The equations describing the displacement due to the spring force depend on the displacement of the RCM (V_{RCM}). This displacement can be found by solving: $V_{RCM} = V_1(L_1) + V_2(L_1) + V_3(L_1)$ for V_{RCM}. The result is the following equation:

$$V_{RCM} = \frac{\frac{M_{pz}L_1^2}{2EI} - \frac{FL_1^2}{6EI}(3L - L_1)}{1 - \frac{KL_1^3}{3EI}} \qquad\qquad (2.7)$$

When equation 2.7 is inserted into equations 2.2 and 2.3 the total displacement can be plotted. Figure 2.5 shows the results for three different load cases. In the first case the torque M_{pz} is zero. In the second case the moment is 0.3 Nm, and in the third it is -0.3 Nm. For each case a tip force of 3 N is applied. The moment and tip force used are chosen such that they are in the range of the forces expected to occur on a regular basis [5, 8].
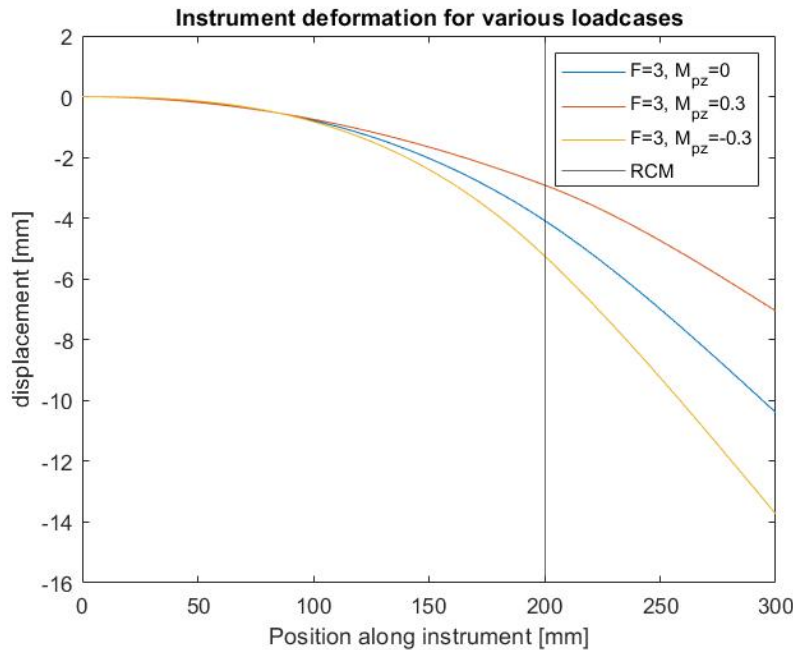
Figure 2.5: Deformation of instrument under several loading conditions.

The results from the equations and deformation plots show a few phenomena interesting for the development of a suitable sensing strategy. The first is found when the spring forces are calculated for the different load cases. The calculated spring force is 4.1 N for case 1, 2.9 N for case 2 and 5.2 N for case 3. This indicates that the deflection of the instrument under moderate load is large enough

that a significant force is added by the abdominal stiffness. This means that the abdominal forces are coupled to the tip forces. It also reaffirms that the trocar forces can cause damage to the patient and instrument. Therefore, even when the alignment of the RCM is perfect and there is no tilt with respect to the abdomen, forces at the trocar still have to be taken into account. It has to be added that this effect is particularly pronounced for the relatively compliant 3 mm instruments

The second conclusion arises when one compares the different plots. The maximum displacement for case 1 is 10 mm whereas the displacement for case 2 is about 7 mm and for case 3 it is 14 mm. It must be noted however that the deformations calculated here are larger than they would be in reality since the trocar would constrain bending over a section of the instrument. Apart from the large difference in maximum displacement there is also a small but visible difference in deformation shape. This means that moments that occur due to the angular position of the instrument can be large enough to alter the deformation of the instrument. Both observations lead to the conclusion that trocar forces and moments will have to be considered when measuring or calculating the tip force.

With the deformation known it is possible to calculate the stress in the instrument to get a feeling for the maximum allowed load with 3 mm SATA instruments. To calculate the stress, first the internal moment $M(x)$ has to be found using the following equation [10]:

$$M(x) = EI \frac{d^2 V_{tot}(x)}{dx^2} \tag{2.8}$$

Because of the complex set of equations the derivatives can be calculated using the discrete finite difference approximation. From the internal moment it is possible to determine the maximum stress σ_{max} in the instrument with the next equation [10]:

$$\sigma_{max} = \frac{M_{max} C}{I} \tag{2.9}$$

Here, M_{max} is the highest moment found along the length of the instrument, C is the radial distance from the central axis and I the second moment of inertia. Using this equation for the worst case, with the weakest stainless steel (316L) [11] and a solid 3 mm instrument, a tip force of 5.5 N is enough to reach the yield limit in the instrument.

2.4. Sterilization

Surgical instruments need to be clean and sterile when they are used during surgery. To achieve this instruments go through an extensive cleaning and sterilization procedure which involves hand or ultrasonic cleaning, exposure to cleaning agents (soap, detergents etc.) and finally a sterilization process. While many different sterilization methods exist, steam sterilization is the preferred method [12, 13] and will thus be considered in this work. Steam sterilization or autoclave sterilization is performed at high temperature (134°C), vacuum, and 100% humidity which is maintained for 3 minutes [13]. The fact that the instrument has to be cleaned and sterilized imposes a number of requirements on any medical device. Firstly, in order be able to remove all contaminants from the product it should not include small open crevices. It should also be robust enough to be cleaned by hand or ultrasonically without breaking. The materials used should not react with cleaning agents, have a large permeability to water vapour or degrade at the sterilization temperature. Finally, since the device could be exposed to salt solution, blood and cleaning agents it should be both corrosion resistant and sealed.

In the case of a sensor not only should it remain intact and functioning after a cleaning and sterilization cycle, it would also be preferred if the sensor is repeatable afterwards. This means that the sensor does not have to be re-calibrated after every sterilization cycle. Not having to re-calibrate the sensor before every use could save a lot of time in preparation before surgery and improve usability.

From here on out steam sterilization and underlying requirements are implied whenever the terms sterilization or sterilizability are used.

<div style="text-align: right; font-size: 3em;">3</div>

State of the art

In the last few decades a lot of literature has been published on force sensing strategies for both MIS and RAMIS. In this time the field has become very diverse with a lot of different sensing technologies and even more different implementations being proposed. This chapter gives a brief overview of the literature. The overview is a summary of the comprehensive literature review performed on the topic by the author [14]. The overview starts by listing the different force sensing technologies that have already been proposed to measure or estimate forces in MIS. Afterwards it will compare the advantages and disadvantages of the different types of sensors for this application. The review of sensing technologies is followed by a short discussion of structures that can be used to implement the different types of sensors. In the final section some general remarks about trends and limitations in the state of the art are discussed.

3.1. Force sensing technologies

Numerous different sensing technologies have already been applied to measure or estimate forces in MIS. In the review by the author [14] the technologies were divided into the following categories: Resistive, Capacitive, Optical, Magnetic, Input force and Image based. This categorization is largely based on earlier reviews by Puangmali et al. [1] and Trejos et al. [2]. The list of categories used here excludes the Acoustic and Pressure based categories included in the literature review because they are not compatible with the goal of this project. In the next part the sensing technologies will be described briefly and the most common sensor types from each category will be addressed. Afterwards, the sensor types will be compared by listing the advantages and disadvantages for this application.

Resistive

The resistive sensors form the largest group of sensors used to measure forces in MIS [14]. These sensors measure force through the modulation of resistance when subjected to strain. The way in which they modulate the resistance is what distinguishes the two most common types of resistive sensors. Strain gauges change their resistance by changing length and cross-section under strain. This is referred to as the geometric effect. Piezoresistors mainly use the piezoresistive effect which changes the resistivity of the material under applied strain [15, 16]. Piezoresistors are smaller and a lot more sensitive to strain than strain gauges but also more fragile and much more sensitive to changes in temperature [14, 15]. Their high sensitivity and small size makes them very suitable for use in pressure sensors.

The relative change in resistance is usually not more than a few percent for resistive sensors. To improve the sensitivity and reduce the Signal to Noise Ratio (SNR) most resistive sensors are implemented with a Wheatstone bridge circuit (also referred to as bridge circuit) [16]. With a bridge circuit the relative change in resistance of the gauge is translated into an absolute change in voltage from zero at the output of the bridge [17]. Implementing two gauges, subjected to opposite strain, in the same branch of the bridge circuit allows for effective temperature compensation [16, 17]. Such a configuration

is commonly referred to as a half-bridge and shown in Figure 3.1 for capacitive sensors.

Capacitive

The capacitive force sensors are, on average, the most sensitive sensors proposed [14]. Most capacitive sensors consist of two flat plate electrodes. They can be used to measure forces through a change in the capacitance between the two electrodes. The change in capacitance results from a change in distance or overlap between the plates. Because the gap between the electrodes is generally very small these sensors are very sensitive to the change in distance. However, the output of the variable distance sensors is non-linear for large displacements [16].
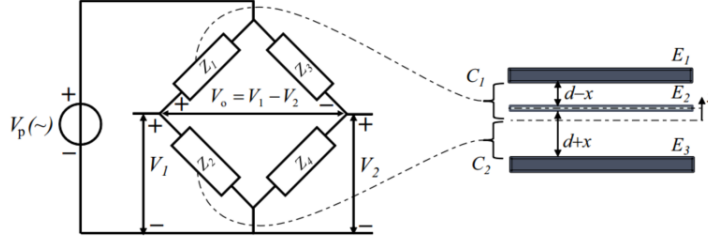
Figure 3.1: A variable distance capacitive sensor linearized through a differential setup with bridge circuit. Figure taken from [17].

Capacitive sensors are often placed in a differential configuration with a bridge circuit [17]. Doing so has a few advantages similar to the use of bridge circuits with resistive sensors. It increases the SNR and sensitivity while reducing the influence of temperature. An additional advantage specific to the variable distance sensor is that the differential configuration can be used to obtain a linearized output [17]. One important limitation of capacitive sensors that remains is a very high sensitivity to electromagnetic interference and stray capacitance.

Optical

In the past decade optical sensors have become very popular in the field of MIS. There are many different ways in which light can be modulated to construct a sensor. To date only two different methods have been applied to measure forces in MIS. The first type of sensor modulates the intensity of the light when a displacement occurs and is referred to as a Light Intensity Modulating (LIM) sensor. The second type modulates the wavelength of light when strain is applied. This type of sensor is called a Fiber Bragg Grating (FBG) sensor [14].
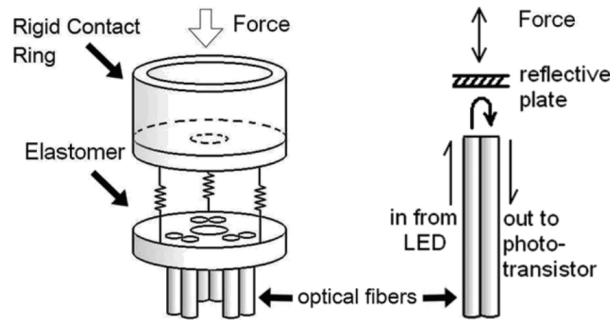
Figure 3.2: Example showing the design and operation of a variable distance light intensity modulating sensor. Figure taken from [18].

There are many different ways in which LIM sensors can influence the light intensity under a displacement. The most common way is to change the distance that the light travels [14]. This can be done by moving a mirror that reflects the light from the emitter back to the detector. The design and operation of such a sensor is shown in Figure 3.2. The advantage of LIM sensors compared to other

optical sensors is their simplicity and relatively low cost. LIM sensors can work with low-coherence light, simple light intensity detectors and simple transduction mechanisms like the moving mirror. Their sensitivity however, is not as high as other optical sensors [16].
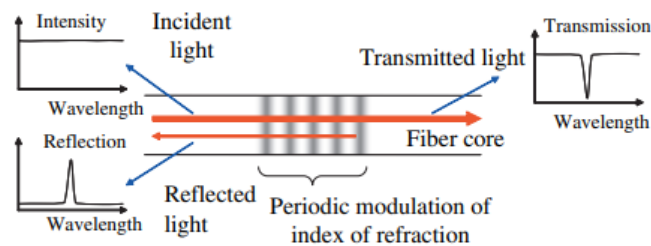
Figure 3.3: Operation of a Fiber Bragg grating sensor. Figure taken from [19].

FBG sensors consist of a periodic variation of the index of refraction, etched into an optical fiber using UV-light, called a grating [19, 20, 21]. The grating causes the light to scatter at each refraction interface. The scattered light adds up through interference to reflect only a specific wavelength which depends on the grating pitch. The other wavelengths pass through the grating [19, 20, 22]. This principle is shown in Figure 3.3. When the sensor is subjected to strain the grating pitch changes and so does the reflected wavelength which allows for the measurement of forces [17, 20]. One point of attention is that FBG sensors are also highly sensitive to changes in temperature. This means that a temperature reference sensor has to be included almost always [22]. The FBG sensor is completely contained in the optical fiber. With optical fibers being as small as 0.25 mm in diameter this makes the sensors very small. The fibers are biocompatible and sterilizable. FBG sensors are also insensitive to parasitic changes in intensity and electromagnetic interference [20].

Magnetic

Magnetic sensors have not been applied frequently for force sensing in MIS. However, because of the general popularity and commonality of magnetic sensors they are still included here. Two popular types of magnetic sensors are Hall/magnetoresistive sensors and inductive (LVDT) sensors.
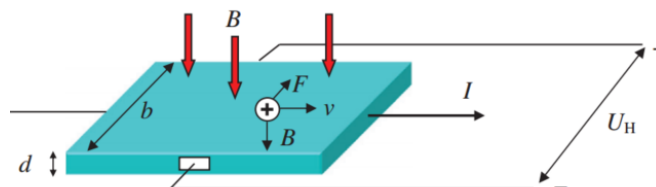
Figure 3.4: Illustration of a Hall plate sensor. Figure taken from [16].

Hall and magnetoresistive sensors are both used to measure the distance from a permanent magnet by detecting changes in the magnetic field density through the Hall effect [21, 23]. The Hall effect shown in Figure 3.4 is a phenomenon which causes a voltage when a plate, generally made of silicon, conducting a current is subjected to a magnetic field perpendicular to its surface. The effect is the result of Lorentz forces acting on the traveling electrons [21, 23]. Not only does the effect cause a voltage across the plate, it also increases the resistance on the current passing through the plate. This change in resistance is what magnetoresistors use as an output whereas Hall sensors measure the voltage. Both types of sensors are generally small with simple electronics and relatively good sensitivity. They are also sterilizable [24].

The most common type of inductive sensor is called a Linear Variable Differential Transformer (LVDT). As Figure 3.5 shows, the LVDT sensor consists of three stacked coils and a concentric ferromagnetic core. The middle core is supplied with an alternating current resulting in an alternating magnetic field. The two other sensing coils are inductively coupled to the magnetic field through the core like in a transformer. When the core is displaced up or down the induced voltage in both sensing
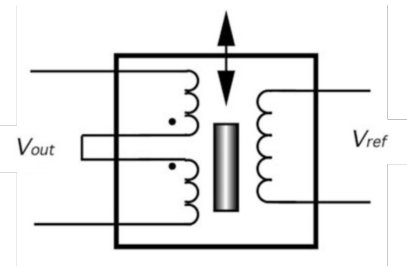
Figure 3.5: Schematic of the construction and operation of a LVDT. When the core is in the middle the output is zero. When the core moves the voltage at one coil will increase. Figure adapted from [21].

coils changes. The voltage decreases in one and increases in the other sensing coil which leads to a relatively high sensitivity to displacement.

Input force
This category of sensing technology does not cover a specific type of sensor. Instead, it concerns a distinct approach to obtaining the forces occurring at the tip and is therefore covered separately. The input force sensing strategy generally uses no additional sensors to determine the forces at the tip. The method uses input force data from actuators that steer the robot arm, or the instrument in the case of a steerable instrument. Relating the input force data to tip forces involves a relatively complicated kinematic model. The model also has to include losses due to friction and trocar forces to provide an accurate result [14]. While most construct a complete model including friction it is not uncommon to model the complete system or only the friction using neural networks.

Image based
Like the input force category the image based category does not cover a specific type of sensor. The image based category uses a sensor that is already present during any laparoscopic surgery; the endoscope. The method uses images of the surgical field from the endoscope to detect differences related to the application of a force. Figure 3.6 clearly shows how it is possible to detect the magnitude of force being applied to a tissue by looking at the indentation depth. Most implementations of this method use neural networks to detect the differences in the images over time [14]. The method does suffer from some significant limitations. The images of the indentation can be easily obstructed by reflected light, blood, smoke and other instruments [14, 25].
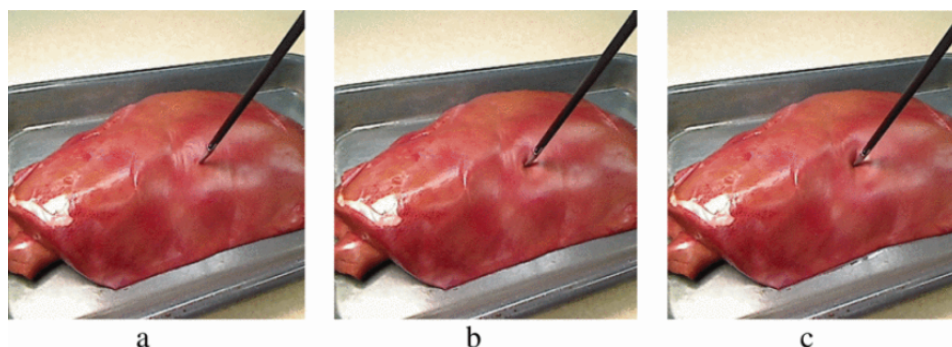
Figure 3.6: Three images to indicate visual changes resulting from increasing instrument-tissue interaction forces. Figure taken from [25].

Comparison
Table 3.1 gives a summary and comparison of the above-mentioned sensor types. The advantages and disadvantages were collected from review papers [1, 2, 14, 22, 26], books on sensor technology [16, 17, 21, 23] and a Master thesis report [24]. In the case of contradicting claims about sensor performance (not uncommon) the application specific source was leading.

Table 3.1: Comparison of the different sensor types previously used in MIS. Entries are based on literature on sensors in general and review papers focussed on medical applications [1, 2, 14, 16, 17, 21, 22, 23, 24, 26]

Sensor type	Modulated parameter	Advantages	Disadvantages
Strain gauge	Resistance	• Large sensing range • Relatively low cost	• Limited sensitivity • Temperature sensitivity • Hysteresis • High power consumption • Limited sterilizability
Piezoresistor	Resistance	• High sensitivity • Low noise • Small • Simple electronics	• Temperature sensitivity • Hysteresis • High power consumption • Not sterilizable • Low repeatability
Capacitive	Capacitance	• High sensitivity • Large sensing range • Large bandwidth • Robust • Small	• Stray capacitance • EM sensitivity • Non-linear • Complex electronics • Not sterilizable
LIM	Light intensity	• EM insensitive • Low cost • Sterilizable • Biocompatible	• Limited sensitivity • Parasitic intensity change • Microbending sensitivity
FBG	Light wavelength	• Large sensing range • High sensitivity • EM insensitive • Small • Sterilizable • Biocompatible	• Temperature sensitivity • Complex interrogation • Expensive
Hall	Voltage	• Low noise • Small • Low cost • Sterilizable • Simple electronics	• Limited sensitivity • EM sensitivity • Offset voltage • Magnet required
LVDT	Induced voltage	• Large sensing range • Good sensitivity • Linear output	• Low frequency response • Poor reliability • Large size • Not sterilizable
Input force	Actuator torque	• No instrument sensors • Cheap • Infinitely reusable	• Complex model • Inaccurate • Lot of data required
Image based	Endoscope image	• No instrument sensors • Cheap • Infinitely reusable	• Complex algorithm • Inaccurate • Image obstructions

3.2. Sensing structure

Choosing the appropriate sensing technology is usually only half the problem. Deciding how to implement the sensor also comes with considerable difficulty and variation. Some sensors can be directly applied to the surface, others require carefully designed structures to optimize their sensitivity. This section will briefly discuss the different methods and structures used to implement sensors in MIS.

Strain gauges, piezoresistors and FBG sensors can be implemented by bonding them directly to the substrate. This is possible because the material these sensors are made of deforms with the substrate to produce a signal. In short, they measure strain instead of displacement. The force resolution of these sensors depends largely on the amount of strain in the substrate under a certain load. When the substrate is very stiff it will deform very little under load and it will be difficult to measure small differences in applied load. In the opposite case, when the substrate is not stiff, the resolution can be very high. However, since most substrates and sensors only allow a limited deformation or displacement this reduces the force range. Moreover, stiffness can be different in each direction resulting in a different resolution in each direction.

To improve the resolution of strain sensors and to facilitate multi-DoF force sensing they are often implemented with specific sensing structures [2, 14, 27]. The structures are designed such that the different force components can be decoupled and sensed separately. They are often monolithic structures which allows for a compact sensor design. The most common designs of sensing structures are used for 6 DoF force sensing and include the Maltese Cross which has many derivatives and the Stewart platform [2]. The dissertation by Seibold [27] includes a study of six different monolithic sensing structures (shown in Figure 3.2) that can be useful for measuring tip forces in RAMIS. It must be noted that the structures are mostly suitable for strain gauges or piezoresistors and may be not compatible with FBG sensors. Although, FBG sensors also benefit from a higher sensitivity and might therefore not require a specific sensing structure.
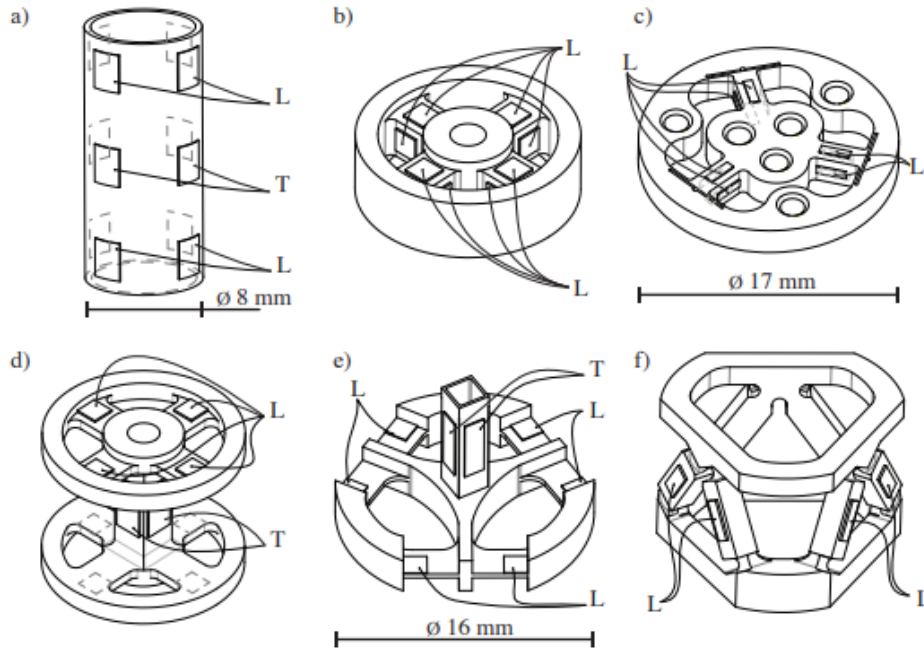
Figure 3.7: Monolithic sensor structures for 6 DoF force/torque sensing. 'L' Denotes longitudinal strain sensors and 'T' denotes transverse strain sensors. Structures include: Tubular shape (a), Maltese cross bar (b), ATI nano 17 (c), DLR fingertip sensor (e), and Stewart platform sensor(f). Figure taken from [27].

For sensors that measure displacement instead of strain, like the capacitive, LIM, Hall and LVDT sensors, the trade-off between resolution and range is similarly relevant. The difference is that the required displacement or deformation for these sensors is generally larger. This means that a spring or structure that facilitates or even amplifies displacement is almost always required. Structures used

for these sensors are application specific and vary greatly [14]. Some simple sensors use an elastic layer between emitter and detector [28], others use a standard helical spring [29] and some design more complicated spring/deflection structures for their application [30, 31]. Figure 3.8 includes two examples of compliant structures designed for displacement sensors. Many more designs are possible for both cylindrical and planar structures.

(a)

(b)
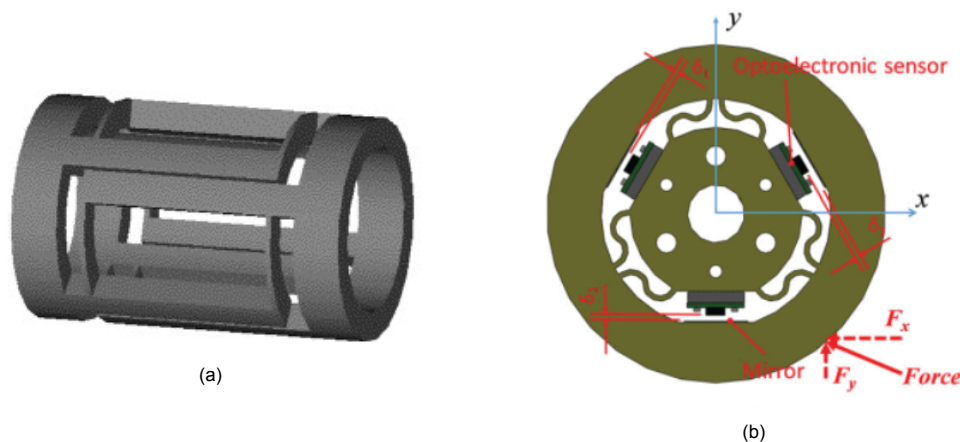
Figure 3.8: Examples of compliant structures for displacement sensing with a cylindrical (a) and a planar (b) structure. Figures taken from [30] and [31] respectively.

3.3. Discussion on State of the Art

Some global observations and remarks about the State of the Art of sensing in MIS were discussed in [14]. This section will briefly discuss observations made about the DoF, sensor location, reusability and validation in literature. These specific topics are discussed because they are relevant to the objective of this work.

Degrees of Freedom

In Section 2.2 it was already discussed that the three translational forces at the tip are most relevant for preventing damage to the instrument or tissue. Even for force feedback it is not strictly necessary to measure all 4 or 6 DoF and feed them back [14]. These assertions are also reflected in the literature. More than 80% of included sensors do not measure all available DoF. Most sensors measure forces in 3 DoF ($> 35\%$) from which almost all measure the linear forces in x-, y-, and z-direction. Only about 11% of sensors measure in 6 DoF, almost all off-the-shelf sensors.

Sensor Location

There are many different locations from which forces can be measured to obtain the forces acting at the tip. Possible locations include: the tip, the wrist, the shaft (proximal or distal), the trocar, the interface between instrument and robot arm or none of the above in the case of input force sensing [14]. Because of the forces occurring at the trocar choosing between these locations is a trade-off between accuracy and the sterility and size requirements. In the literature published to date there is a clear preference towards sensors placed at distal locations (tip, wrist or shaft). Almost two-thirds of the included sensors were placed in these locations with most placed at the tip [14]. This suggests that sensing accuracy was considered most important in this trade-off.

Reusability and Cost

As is often the case in the medical field, reusability and cost are important aspects to consider. Whether a sensor is reusable is a matter of being able to clean and sterilize the sensor without damage or a loss of performance. For a sensor to be cleanable it should be robust and not contain crevices from which residue cannot be removed. When sensors are very small because they are placed in invasive locations this can become a difficult problem.

Less than half of the distal sensors proposed to date claim to be sterilizable [14]. Most of those are the relatively to very expensive optical sensors. Only two papers discuss cheaper strain gauge sensors that are sterilizable through autoclave sterilization, albeit for a limited number of times [32, 33]. When the sensor is not sterilizable it should be cheap enough to be disposed of, preferably separate from the relatively expensive instrument.

Validation

One important observation about the state of the art is that there is a lack of representative validation concerning a number of different aspects. The aspects include the determination of the required force range, the accuracy of proximal sensing strategies, and the reusability of certain sensors.

Many papers cite a force range during surgery between 0 and about 10 N. The value for the force range has been presented in studies by Picod et al. [5] and Toledo et al. [8]. However, those measurements of the force range were obtained using proximal sensors. This means that the measurements were likely influenced by forces acting through the trocar. Although there is no reason to assume that the forces are higher in reality (in fact, the real range might be smaller) this is something that has to be considered. The lack of an accurate figure for the force range also exemplifies the difficulty of the problem.

The forces occurring at the trocar are also the reason to claim there is a lack of representative validation for proximal sensing strategies. A lot of proximal sensing strategies have been proposed. Especially the input force strategy has been popular in the past 10 years [14]. This method however uses measurements performed far from the instrument tip to calculate the forces at the tip. While some of these methods also include models that take into account the dynamics of the trocar and abdomen none of them have been tested using an actual or realistic trocar and abdomen setup. Similarly, the image based strategies which are easily obstructed have not been tested in a setting with fluids or other instruments. Consequently, it is not certain whether these sensing strategies meet stated performance requirements in real use.

Finally, in the part *Reusability and Cost* it was mentioned that about half of the distal sensors claim to be reusable. Most of these claims are based on the fact that the materials and adhesives they use are rated for the required temperature ($134°C$) and humidity. Nonetheless, problems could still occur when these materials are combined with materials with a large difference in thermal expansion. There are however only two studies that have actually demonstrated sterilizability specifically for strain gauges [32, 33]. Although there is no reason to assume that the other sensors are not sterilizable it would still be relevant to test for repeatability and deterioration of the materials after several cleaning and sterilization cycles.

4

Design

After discussing all the required theoretical background in the previous chapters this chapter will cover the design of the sensor. First, based on the information from the previous chapters a list of requirements is given that will be used to compare different concepts. Next, three global concepts are discussed. These concepts propose different configurations of sensors one could use to measure the tip forces while taking into account the trocar forces. The choice of concept is made by a simple comparison of the advantages and disadvantages of each concept. After selecting the global concept a choice is made on which sensor type and associated sensing structure should be used. Both are selected using Harris profiles. In the final part more detailed design choices like magnetic system, sensor model, materials and dimensions are discussed.

4.1. Requirements
The goal of this project is to measure forces acting on the instrument tip during RAMIS with 3mm SATA instruments. The force measurements will be used to prevent damage to the instrument and internal tissues either through force limiting or force feedback. This section provides a list of requirements the sensing system needs to meet in order to achieve this goal.

- **Degrees of Freedom:** In order to effectively protect the instrument and tissues from damage the sensing strategy should at least be able to measure the forces at the tip that result in bending (F_x and F_y). In order to also protect the tissue from damage by pulling it is preferred to also be able to measure the axial force (F_z).

- **Sensing Range:** According to literature, forces during MIS with conventional instruments range up to about 10 N at most [5, 8]. However, the 3 mm instruments considered in this project are significantly less stiff compared to conventional $5 - 10$ mm instruments. As described in Section 2.3, when the tip force is increased to 5.5 N the bending stresses already exceed the yield limit of stainless steel alloys used in the medical field. Therefore, taking into account the allowable error, a more conservative 5 N is used as the required force range.

- **Bandwidth:** Most surgical gestures during MIS occur within the $0 - 2$ Hz bandwidth range [34]. However, for force limiting it is useful to also take reflexive movements into account since those are most likely to result in accidental damage. Human reflexive action is limited to a bandwidth of 10 Hz [35]. In order to prevent aliasing the minimum sampling frequency should be 20 Hz.

- **Accuracy:** In order to effectively prevent damage to the instrument or tissue the system should have a maximum relative error (MRE) no larger than 10%. Together with the 5 N force range a maximum error of 10% will ensure the yield limit is not reached. This maximum error means that the trocar forces discussed in Section 2.2 have to be accounted for since they have a contribution much larger than 10%.

- **Size:** Regardless of where a sensor is actually placed it should fit on/in the robot arm, instrument or trocar. The dimensions of the instrument and trocar should preferably not change and result in a larger incision.

- **Biocompatibility:** When a sensor is placed somewhere it will contact internal or external tissues the sensor has to be biocompatible. This means that materials that come into contact with tissue or blood must not cause adverse reactions or an immune response.

- **Usability:** Included sensor(s) should be integrated in such a way that it does not require long or complicated steps to set up the system right before surgery. It should not hinder the surgeon or other staff in any way during surgery. It should remain simple to remove the instrument from the gearbox or the gearbox from the robot arm.

- **Cost:** The proposed sensing system should be as affordable to use as possible. This takes into account both initial investment cost and operational costs. As a consequence a solution that is not expensive to purchase as well as being reusable is required. In order to be reusable the sensor has to be sterilizable through a standard process of cleaning and autoclave sterilization as described in Section 2.4.

- **Repeatability:** The sensing system should preferably not require a calibration procedure before every surgery.

4.2. Three global concepts

The generation of concepts is performed in a two step process, starting with global concepts and later choosing the sensing technology to be used with the global concept. The global concepts consider the configuration of sensors that have to be used to be able to meet the requirements. The two step process is used because each global concept can be implemented using many different sensing technologies and structures. Considering all combinations would lead to a very long list of detailed concepts to choose from.
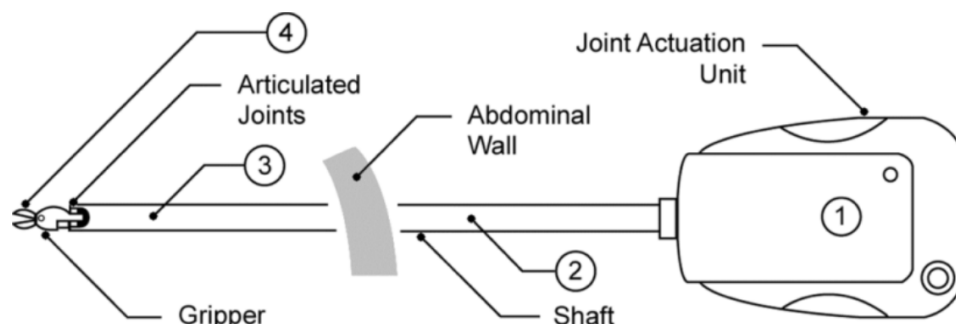
Figure 4.1: Schematic drawing of a RAMIS instrument with annotated sensor locations. Figure taken from [1]

Chapter 2 has taught that the instrument system is influenced by forces and torques that occur at the trocar because of the stiffness of the abdominal wall. These forces and torques are large enough to severely influence the measurements and cause damage to the patient and instrument. These forces are also coupled to the tip force being applied and the orientation of the instrument while also depending on a lot of different environmental and patient conditions. Because of their magnitude, near unpredictability and the fact they are coupled, these forces have to be taken into account in order to meet the requirements. This greatly reduces the number of possible solutions. Figure 4.1 shows a schematic drawing of a robotic instrument with annotated locations one could measure forces from. All sensors placed in locations inside the abdomen are not influenced by the trocar forces. Sensors placed in these locations however are severely constrained in size and compatibility with SATA instruments. All sensors placed to the right are influenced by the forces from the trocar but are less constrained in terms of size. After excluding all solutions that cannot account for the trocar forces, or are placed in locations that cannot feasibly accommodate a sensor, only three global concepts remain. Each of the remaining concepts, their various implementations and the advantages or drawbacks will be discussed in the following sections. After introducing all three concepts a choice is made based on their respective advantages and disadvantages.

Distal strain sensing

The first global concept measures the strain in the instrument shaft past the trocar, annotated with "3" in Figure 4.1. Placing a sensor at this location means that the measurement is not influenced by the forces from the trocar. As a result the concept requires no measurement or estimation of the trocar forces. Other locations like the joint or the gripper are not considered a feasible option due to the very limited size and geometry of the SATA instruments. As already mentioned in Chapter 3 distal sensing is the preferred solution to obtain the tip forces in literature. However, the implementation of distal sensors is more complicated in shaft actuated instruments. This is mainly because of the small size of the instrument and the thin shafts it consists of. The small size is also the reason why displacement sensors are not suitable for this solution since they generally require more space to allow for a sufficiently large displacement. The sensor technologies that remain for this solution are strain gauges and FBG sensors.

Both sensors can be integrated by milling grooves in one of the shafts to run the wires or fibers. The most suitable shaft to integrate the sensors in is the middle shaft because the forces from the tip are directly transferred to the middle shaft. A drawback of placing the sensors on the middle shaft is that it can only measure the bending forces and not the axial force. Because of the way the steering mechanism works the axial force is transferred to the outer shaft and center rod. To measure the bending in two directions with sufficient accuracy four strain gauges are required in two half-bridge configurations. To fit all four gauges they have to be less than 1.5 mm wide. Strain gauges of this size are not very common and relatively expensive (€200). Measuring forces in 2 DoF would require only two FBG sensors. Moreover, each FBG sensor requires only one fiber whereas each strain gauge requires two wires to connect them.

A second reason the implementation of distal sensors is more complicated results from the way the instrument is actuated and the fact that the instrument has to be separated from the gearbox. The strain sensors on the shaft have to be connected to an interrogation system placed at the base of the robot arm or separately. The wires or fibers that run along the length of the instrument have to be fed into the gearbox in a way that does not limit the range of motion nor complicate the assembly before surgery. This is particularly difficult for the optical fibers since their fragility causes them to break easily when bent into a small radius. Moreover, connecting and adapting the optical fibers inside the gearbox is rather complicated because of the large connectors required to ensure proper alignment. Connecting strain gauge wires from the instrument to gearbox is less complicated because electrical connections are not affected by alignment. However, connecting eight wires inside the gearbox in a practical manner that also allows for the required motion is still challenging.

Interface and trocar sensor

The second global solution uses two separate sensors to indirectly obtain the forces on the tip. One sensor measures the input forces provided by the robot arm, location "1" in Figure 4.1. If no interactions occurred between the trocar and abdomen these measurements alone would be sufficient to determine the forces at the tip. However, a second sensor is required because of the interaction forces at the trocar. The second sensor is located at the trocar, the abdominal wall in Figure 4.1 and measures those interaction forces. Together, the two measurements can be used to calculate the forces at the tip.

As described in Chapter 3 it is possible to determine the input forces at the robot arm side by measuring the motor torques and using a kinematic model. However, this system is actuated by stepper motors. Stepper motors are excellent for position controlled system but not for force controlled systems. To obtain the output torque from stepper motors, torque sensors would have to be added to the output shaft. Doing so would defeat the advantages of using cheap and standardized components. Therefore, this solution would use sensors placed either at the interface between the instrument actuation system and the robot arm, or between the actuation system and instrument. The former has the advantage that it is a permanent part of the robot arm. Replacing, removing and cleaning/sterilizing are not of concern here. However, when placed at the gearbox-instrument interface the sensor should be easy to connect to and remove from the system.

The second sensor will measure the forces acting between the instrument and trocar. Together with the input forces these forces are necessary to calculate the tip forces through the equations of motion. A trocar sensor can be located at the bottom, invasive side of the trocar as proposed by Fontanelli et al. [36] or at the top, non-invasive side as proposed by Zemiti et al. [37]. For this project the non-invasive side would be preferred mainly because the incision size would not have to be increased to fit the sensor. The sensor could for example be integrated in a custom removable sealing unit such that it can, if necessary, be removed and replaced without disposing of the entire trocar sleeve. The fact that there is a second sensor at the trocar means that the forces at the incision will also be known during surgery which allows for limiting of these forces.

Fixed trocar

The third global solution is based on the force sensing systems proposed by both Zemiti et al. [37] and Fontanelli et al. [36]. These solutions use measurements from a single sensor fitted on the trocar to estimate the forces on the tip. They are able to do so because the trocar and instrument are both fixed to the robot arm, hence the name 'Fixed trocar' for this solution. The result of fixing both to the same frame is that the measured forces are no longer influenced by the trocar forces exerted by the abdomen [37]. The forces exerted on the trocar by the abdomen are transferred trough the trocar to the robot arm instead of to the instrument. Zemiti et al. use an off the shelf 6 DoF piezoresistive sensor mounted at the top of the trocar [37]. Fontanelli et al. use optical intensity displacement sensors located at the bottom of the trocar [36]. As already mentioned in the previous section, placing the sensor in the top, non-invasive part benefits from reduced size and sterility limitations. Similar to the previous concept it would be preferred if the sensor for this solutions is a part that can be added to and removed from conventional trocar sleeves. The most applicable sensing technologies for this concept include displacement sensors that measure some displacement due to the bending of the instrument. Here the contact with the trocar and the sealing have to be taken into account.

Selection of global concept

Three different global concepts have been presented. Each of the concepts is in theory able to measure or estimate the tip forces while taking into account the trocar forces to meet the required accuracy. The different solutions however have different advantages and disadvantages which are listed in Table 4.1 and are used to make a choice on the most suitable global concept.

Table 4.1: Comparison of the proposed global solutions

Global solution	Advantages	Disadvantages
Distal strain sensing	• Direct measurement • Sensitivity • Single sensor	• Only 2 DoF • Expensive • Connecting wires/fibers • Reusability • Size
Interface and trocar sensor	• Non-invasive • Fully determined • Reusability • 3 DoF possible • Size	• Two sensors • Sensitivity • Solving EOMs
Fixed trocar	• Non-invasive • Cheap • Reusability • Direct measurement • Size	• Sensitivity • Only 2 DoF • Trocar fixed to robot

Based on the previous sections and Table 4.1 there are a number of reasons to support a choice of concept. Firstly, it is concluded that it is very difficult to devise a distal sensor that is affordable, reusable and usable considering the invasive location, expensive sensors and connecting wires or fibers. Secondly, when comparing the other options, it can be argued that both are essentially able to achieve the same goal. However, the 'interface and trocar sensor' concept measures 3 DoF but uses two sensors to do so and requires modeling and calculation to find the required forces. The 'Fixed trocar' concept uses only one sensor which requires no back-end calculations but the trocar is fixed to the robot arm. Since the third degree of freedom is not essential and the robot-trocar connection should not result in a loss of function or usability it was decided to proceed with the 'Fixed trocar' concept.

4.3. Concept outline

Before continuing with the choice of sensing technology, structure and more detailed design choices this section provides a more detailed outline of the fixed trocar concept. First, it discusses the way in which the forces between trocar and instrument will be measured. Next, the design and functioning of trocars in general and specific to this project are discussed.
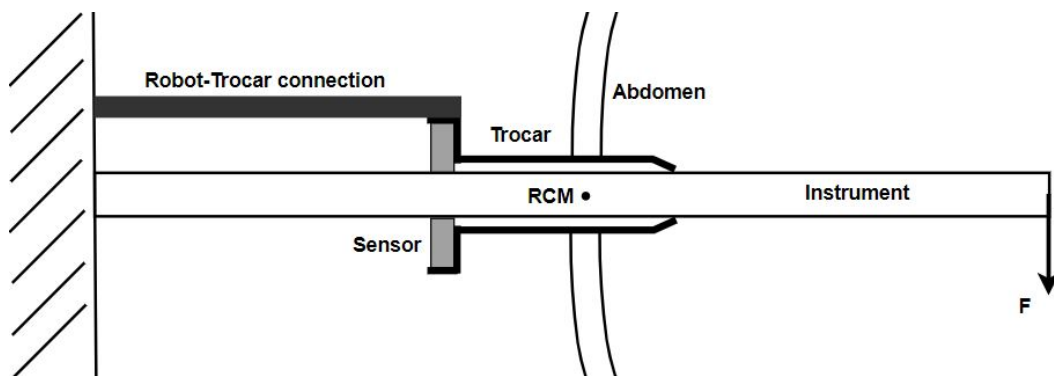
Figure 4.2: Schematic representation of the fixed trocar system

Figure 4.2 gives a schematic visualization of the fixed trocar concept. On the left it shows where the instrument is connected to the driver unit. On the far-right end the tip force is applied. In the middle it shows the trocar situated in the abdomen. As is essential for the concept the trocar is connected to the same 'frame' as the instrument. The instrument is supported by the trocar at the invasive side on the right and contacts the sensor on the non-invasive side on the left like in the design by Zemiti et al. [37]. Different from the system proposed by Zemiti et al. [37] this concept measures the strain or displacement caused by bending from the tip force. What sensor type is used to do this and what structure is needed to do this is discussed in sections 4.4 and 4.5.

The function of the trocar is to protect the incision and maintain the pressure established to inflate the abdomen referred to as the pneumoperitoneum. In its most common shape a trocar is a hollow tube with two seals to prevent the deflation of the abdomen and a port with a valve to further inflate when needed. The first seal, in the shape of a flap, is needed when there is no instrument in the trocar. The second seal, a rubber O-ring or membrane is used to seal when the instrument is inserted.

Trocars however, come in many more different types and various sizes. They range from plastic disposable trocars to reusable metal trocars. Some have double seals others have only one. Most, but not all, include a valve and port. Some have external threads or expandable arms or balloons to secure the trocar in the abdomen. Also many different shapes of the abdomen penetrating tip exist. In robotic surgery very minimalistic trocars are used in the access ports for the instruments. These trocars consist of a simple reusable metal body without valve and a single disposable plastic seal [38]. Examples of robotic trocars are shown on the top right and bottom in Figure 4.3.

While the trocar used in the sensing system will be custom in order to accommodate the sensor, the trocars used in robotic surgery will be used as an example. This means that similar to the sensor the custom trocar has to be reusable. In order to maintain its most important functionality the trocar also

Figure 4.3: Example of the different types of trocar used in robotic surgery. Two plastic disposable trocars and two metal reusable trocars are used. Image taken from [38].

has to be fully sealed, although like the standard trocars used in robotic surgery a single seal would suffice. The trocar also needs to include a way to fix it to the robot arm. It is important that the trocar itself and the connection with the robot arm are much stiffer than the instrument. When this is not the case the deflection of trocar and connection will influence the force measurements.

4.4. Selection of sensor type
Now that the global concept has been chosen the sensor type most suited for the concept has to be chosen. Chapter 3 already listed the different sensing strategies that have been used in the past to measure forces in MIS. However, not all of those sensing strategies are suitable for use in the 'fixed trocar' concept. Consequently, the 'Input force' and 'Image based' sensing strategies are not considered here. The remaining sensor types are compared using a Harris profile in Table 4.2. The Harris profile grades the sensor types on some of the most important qualitative and quantitative performance indicators based on the requirements from section 4.1. The grading is based on the information from Chapter 3 and the literature referenced there. A short summary of the Harris profile will be given before selecting the sensor type that will be used.

Although they are the most common sensor types the resistive sensors are relatively costly at the scale required for this application. Moreover, they are difficult to apply properly and generally not sterilizable. On the other hand, the resistive sensors as well as FBG sensors can be applied directly to the substrate and do not necessarily require a displacement structure. It is however not uncommon that strain sensors are also applied with a flexible structure to improve sensitivity and specificity. The capacitive sensors are the most sensitive sensors included. Because they are very sensitive they are also more difficult to apply. To limit the effect of EM-interference the analog to digital conversion has to happen very close to the electrodes. This complicates the electronic design, increases the size and limits sterilizability. The optical sensors, and especially the FBG sensors, are very sensitive while also being sterilizable and very small. However, both require expensive optical fibers and interrogation hardware. Additionally, the fibers are fragile and very sensitive to temperature. Hall sensors are not the most sensitive sensors included in the table. They are however small, low cost, sterilizable and easy to implement. They do require permanent magnets to be included in the system and are also relatively sensitive to interfering magnetic fields. Finally, the LVDT sensors are relatively insensitive, large in size and also not sterilizable. Together with the capacitive, LIM, and Hall sensors the LVDT

sensors require a structure that allows for the right displacement in the right direction.

From the Harris profile in Table 4.2 and summary it can be concluded that the Hall sensor is the most suitable sensor for this project. It is the only sensor type without any minuses in the table. The only negative notes listed in the final column are the EM sensitivity, the need for a displacement structure to measure forces and the fact that magnetic fields decay non-linearly. One of the main advantages of the Hall sensor is the fact that it is sterilizable and still very cheap. This is not the case with for example FBG sensors. With its simple electronics and physical implementation it also benefits from high usability and robustness.

4.5. Selection of sensing structure

Hall sensors respond to a change in magnetic field density caused by the displacement of a field source, permanent magnets in this case. Because the sensors are sensitive to displacement instead of strain a structure is required that facilitates displacements in the right amount and directions. This means that it should be relatively soft in the direction one wants to measure and stiff in the direction that is not of interest. The sensing structure and sensor arrangement have to be designed in a way that optimizes the range and resolution of the sensor. The sensor also has to fit on or in the trocar without compromising the functionalities described in Section 4.3.

From the outline of the fixed trocar concept in Section 4.3 it is clear that the expected forces and deformations at the sensor are in the plane perpendicular to the long axis of the instrument. Consequently we will only consider structures that allow for planar displacements. Table 4.3 provides a Harris profile on five different structures that could be used to measure those planar forces with Hall sensors. Before discussing the result of the Harris profile each structure will be briefly introduced.

The first structure was proposed by Noh et al.[31] to measure the radial contact force on a flexible manipulator. It uses pairs of radially mounted S-flexures that constrain all movements except for the planar translations. The result is a compact design with a stiffness and sensitivity easily tunable by altering the S-flexure parameters [31]. The optical sensors used by Noh et al. can be easily replaced by Hall sensors. However, with the open design, some attention has to be paid to cleanability and sealing capability of the sensor.

The second structure is based on the Hall sensor Jones et al. [28] proposed to measure the grip force. In their design Jones et al. separate the Hall sensor from a permanent magnet with a compliant layer of rubber. Translating the concept to this project results in magnet-sensor pairs placed radially with a compliant layer in between. Stiffness of this structure can be altered by choosing a different rubber material or compound, or by changing the geometrical parameters. A sensor with this structure using rubber can be cheap and easy to produce since most rubber materials come in two parts which can be mixed and poured into any desired shape. With the right rubber the sensor can also be fully cleanable and reusable trough autoclave sterilization. Because there are no gaps in the design of the sensor it is possible to fully seal the trocar without any further changes to the sensor.

The third and fourth structures are similar designs. Both consist of two perpendicular flexures placed in series. This is a concept that is common in planar positioning mechanisms. There are two main differences between the two designs. The third structure is less stiff in the sensitive direction compared to the fourth and therefore more sensitive. Though more sensitive, any displacement of this structure is coupled to a small unwanted tilt of the magnet with respect to the Hall sensor. The fourth structure is stiffer but not affected by any magnet tilt. To achieve a similar stiffness the sensor will become larger in size.

The fifth and final structure was developed by SIMTech and included in a paper by Teo et al. [39]. Its flexures are configured in a triangle shape leading to a slightly stiffer but compact design. Finally, it is also a more complex design compared to the previous two structures. Due to their open design each of the latter three designs has a limited cleanability and does not seal the trocar.

After considering the descriptions and Harris profile it was decided to use the second structure for this project. Compared to the other structures it has three main advantages. Because of the solid rubber layer the design is fully sealed. This also means that it has no places for residue to collect and
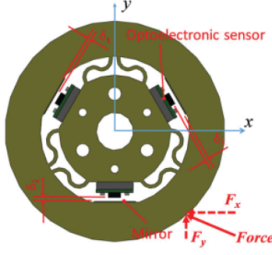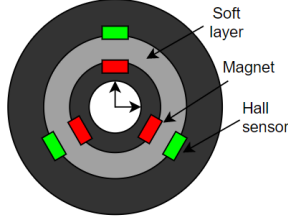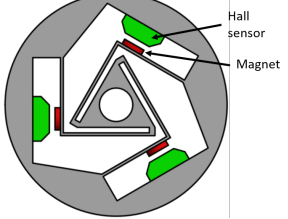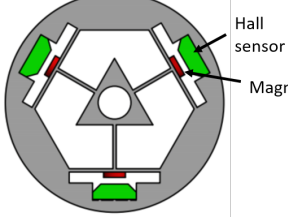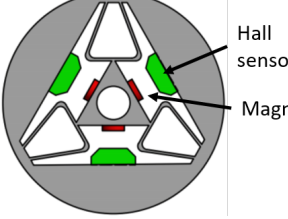
therefore the sensor is cleanable and sterilizable when the right rubber is used. None of the other designs are inherently sealed and cleanable in the same way. The third advantage is the low production cost of the structure. The second structure can be produced simply by pouring the rubber into the desired shape. The other mechanisms need some (laser-)cutting and/or assembly of the flexures which can be relatively expensive. Also the placement of sensor and magnet is relatively easy since this can be done before the rubber is poured.

The sensitivity of this structure is limited compared to the others because of the solid rubber layer. However, very soft rubber compounds are available. Softer rubber will mean that the deformations are less well constrained. However, solutions like differential measurements will allow for the rejection of some of the unwanted displacements. Additionally, the deformation of rubber materials is non-linear and could suffer from hysteresis. These factors can however be mitigated by choosing the proper type and shape of compliant material and reducing the relative deformation of the compliant layer. An important property of rubber that has to be considered is the fact that the mechanical properties are sensitive to changes in temperature. With the patient at around 310K and the ambient air at 293K it is probable that heat will be transferred from the patient to the instrument and into the sensor. An increase in temperature will result in stiffening of the rubber and an increase of the hysteresis [40]. However, the relative change in temperature and the change in stiffness per degree are limited. Therefore, the prototype developed in this work does not include a reference sensor to compensate for temperature effects. Other simpler measures to account for temperature changes are discussed in the next section. If it is found to be necessary due to a large amount of drift a reference sensor can be included in the design. The main reason this concept is still used is that usability and cost are very important requirements for the sensor giving them a higher weight over sensing performance.

Table 4.2: Harris profile of applicable sensor types

Sensor type	Sensitivity	Repeatability	Sterilization	Cost	Size	Usability	Notes
Strain gauge	+/-	+	-	-	++	-	• Difficult installation • Temperature sensitivity • Semi-reusable
Piezoresistor	+	+	–	–	++	-	• Difficult installation • Temperature sensitivity • Expensive • Fragile
Capacitive	++	+	–	+	++	+	• EM sensitivity • Non-linear • Displacement structure
LIM	+	+	++	+	-	-	• EM insensitive • Displacement structure
FBG	++	+	++	–	++	-	• EM insensitive • Very Expensive • Fragile fibers • Temperature sensitivity
Hall	+	++	++	++	+	++	• EM sensitivity • Non-linear • Displacement structure • Permanent magnets
LVDT	-	+	–	++	–	-	• EM sensitivity • Large size

Table 4.3: Harris profile of applicable sensor structures

Structure	Size	Sensitivity	DoF	Stiffness (OOP)*	Notes
	++	+	2	+	• Torque constrained • Compact • Cleanability
	++	-	3	+	• Sealed design • Compact • Simple production • Non-linear • Temp. sensitivity • Reusable
	-	++	3	-	• Low OOP stiffness • Magnet tilting • Simple • Cleanability
	-	+	3	-	• Simple • Cleanability
	+	-	3	+	• Relatively complex • Compact • Cleanability

Out Of Plane

4.6. Detailed design

In the previous sections it was decided to use a trocar fixed to the robot and measure the forces between the trocar and instrument. These forces will be measured using Hall sensors which measure a displacement. A compliant structure is required to translate the applied force to a displacement. The compliant structure will consist of a compliant material layer. After these general selections more detailed design choices still need to be made. These choices include the design of the magnetic system (size and

configuration of magnets) and the sensor geometry. The materials also have to be selected. Making these choices however will require a trade-off between sensor performance and practical requirements.

The first performance requirement to consider is the overall sensitivity of the sensor. The overall sensitivity is defined as the change in output in milliVolt for each input change in Newton (mV/N). The goal is to achieve a sensitivity that is as high as possible which results in a high resolution on the force measurement. The overall sensitivity is influenced by multiple sub-sensitivities. Firstly, the sensitivity of the Hall sensor to changes in magnetic field density (mV/Gauss) determined by the sensor model chosen. Secondly, the sensitivity of the magnetic system to displacements (Gauss/ mm displacement). This sensitivity is influenced by the magnetic configuration, magnet size and initial distance between magnet and sensor. Finally, the stiffness of the sensor, instrument and constraint system determine the deformation sensitivity (mm displacement/N).

A second important performance requirement to consider is the signal to noise ratio (SNR). A high SNR implies that the output (and change in output) are much larger than any signal changes that are caused by noise. Two main sources of noise can be identified for this sensor. Firstly, there can be small random changes of the magnetic field or fluctuating external magnetic fields that are measured and appear as noise. The effect of this noise can be minimized by using relatively strong magnetic fields in the sensor such that small random changes will not significantly influence the measurement. Secondly, there can be electronic noise, for instance from the power source or from any amplifiers in the Hall sensor. Both sources of noise can be classified as white noise (zero mean Gaussian noise) and can therefore easily be reduced through signal conditioning like filtering or averaging.

The final requirements involve more practical considerations like production and cost. The design of the sensor should take into account that both the final sensor has to be affordable to both buy and use. Therefore any geometries that are difficult to produce or include exotic materials should be avoided. Moreover, since the compliant material will be poured into the sensor there should be sufficient space between sensor and magnet for the material to flow. Additionally, a thin layer of rubber is a lot stiffer than a thicker layer of rubber which will reduce the force sensitivity. Considering that magnetic field densities are much stronger and sensitive close to the source this is a complicated trade-off.

4.6.1. Expected displacement

To design the sensor with the highest possible sensitivity and resolution it is important to know what size of displacements are to be expected. To get an estimation of the expected displacements a deformation analysis similar to the one in Section 2.3 was performed. Because of the fixed trocar concept the model used for this deformation analysis is different from the one used in 2.3. Figure 4.4 shows the model used for this analysis.
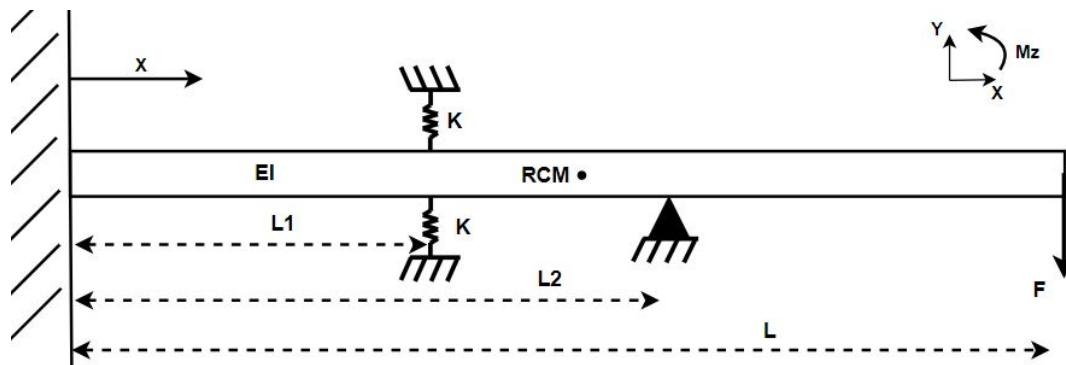
Figure 4.4: Deformation model of the fixed trocar system

The model in Figure 4.4 includes a force on the tip, a support where the instrument contacts the bottom tip of the trocar, and springs to represent the stiffness of the sensor located at the top of the trocar. The exact stiffness of the sensor is not known at this point but will be addressed in Section

4.6.5 and Appendix C. For this analysis it is assumed that the stiffness of the trocar and the frame that connects the trocar to the robot arm are much higher than the stiffness of the instrument and sensor. The distance between the springs and support is fixed by the length of the trocar at 80 mm. The distance L_1 represents the insertion depth and can be varied. The length of the instrument is 300 mm. Since the trocar and base of the instrument are connected to the same rigid frame the stiffness of the abdomen is not included in this model. For the same reasons described in Section 2.3 geometric non-linearity was not accounted for in the solutions.
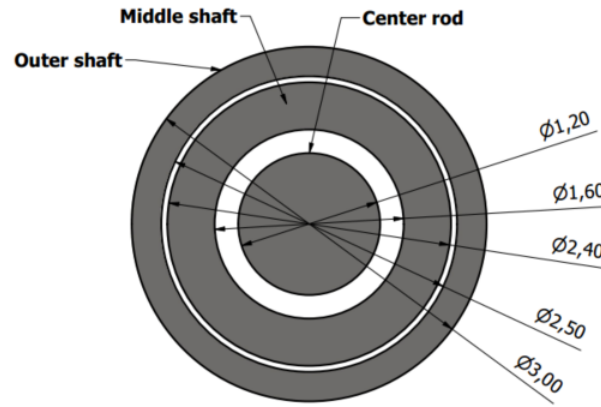
Figure 4.5: Cross-section of the 3 mm shaft actuated instrument.

Since the shaft actuated instrument consists of three concentric tubes calculating the bending stiffness is slightly more involved compared to the simplified model in Section 2.3. The cross-section of the instrument is shown in Figure 4.5. Because the shafts are concentric the second moment of inertia (I) is given by summing the moments of inertia for each of the shafts. In the cross-section it can be seen that there are gaps between the shafts. These gaps will in reality lead to some additional deformation. This effect is not included in this model, the friction that occurs between the shafts during deformation is also neglected.

Similar to the simplified model in Section 2.3 the deformation of the instrument in the fixed trocar model is found using the principle of superposition. Here, the first case only includes the force from the springs $F_b = KV_b$, the second includes only the force from the support F_c and the third the tip force F. The resulting equations which describe the displacements $V_1(x)$, $V_2(x)$, $V_3(x)$ and $V_{tot}(x)$ are given below.

$$V_{1a}(x) = -\frac{KV_b x^2}{6EI}(3L_1 - x_1) \qquad 0 \le x \le L_1 \qquad (4.1)$$

$$V_{1b}(x) = -\frac{KV_b L_1^2}{6EI}(3x - L_1) \qquad L_1 \le x \le L \qquad (4.2)$$

$$V_{2a}(x) = \frac{F_c x^2}{6EI}(3L_2 - x) \qquad 0 \le x \le L_2 \qquad (4.3)$$

$$V_{2b}(x) = \frac{F_c L_2^2}{6EI}(3x - L_2) \qquad L_2 \le x \le L \qquad (4.4)$$

$$V_3(x) = -\frac{F x_1^2}{6EI}(3L - x) \qquad 0 \le x \le L \qquad (4.5)$$

$$V_{tot}(x) = V_1(x) + V_2(x) + V_3(x) \qquad 0 \le x \le L \qquad (4.6)$$

In the equations above, Vb and Fc are still unknown. They can however be solved with the use of boundary conditions on the displacement. First it is recognized that the displacement at the support must be zero. This leads to the equality $V_1(L_2) + V_2(L_2) + V_3(L_2) = 0$ which can be solved to obtain an expression for F_c:

$$F_c = \frac{F}{2}(3\frac{L}{L2} - 1) - \frac{KV_bL_1^3}{2L2^3}(L_1 - 3L_2)$$
(4.7)

The expression for F_c together with the equality $V_b = V_1(L_1) + V_2(L_1) + V_3(L_1)$ can be used to solve for V_b. The solution is given in the following equation:

$$V_B = \frac{\frac{FL_1^2}{4EI}(L - \frac{LL_1}{L_2} - L_2 + L_3)}{\frac{KL_1^4}{12EIL_2^3}(6L_1L_2 - L_1^2 - 9L_2^2) + \frac{KL_1^3}{3EI} + 1}$$
(4.8)

The deformation that results from the equations above is shown in Figure 4.6. For this plot a sensor stiffness of 5 N/mm and a tip force of the required maximum 5 N is used. A length of 100 mm is used for L_1. With these parameters the displacement at the sensor (V_b) is 0.7 mm. The displacement of the tip is about 8.5 mm. Considering that the applied force is larger than in Figure 2.5 this means that the fixed trocar system is stiffer resulting in a smaller tip displacement.
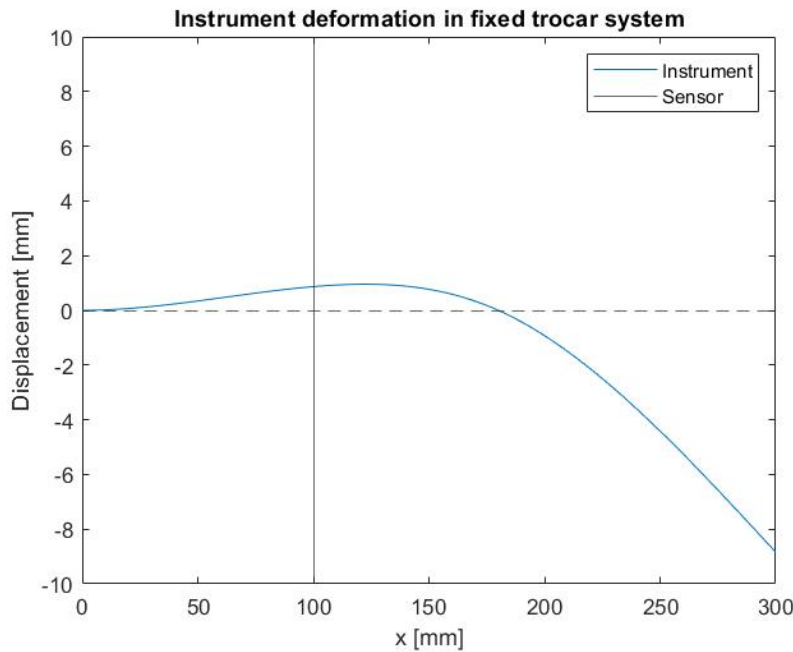
Figure 4.6: Estimated deformation of the instrument in the fixed trocar system.

To investigate the effect of the sensor stiffness and insertion depth on the measured displacement a surface plot is used. The surface plot in Figure 4.7 presents the insertion depth (L_1) and stiffness (K) on the x- and y-axis and the displacement V_b on the z-axis. A number of things can be observed in the surface plot. Firstly, the plot shows that increasing the stiffness leads to a reduction in displacement over the entire range. This reducing effect seems to approach an asymptote resulting in very small reductions in displacement at higher stiffness. Secondly, increasing the insertion depth does not lead to a monotonous increase or decrease in displacement. When the insertion depth is small, the moment arm between the tip force and the trocar is short leading to a stiffer system. Similarly, when the insertion depth is large, the sensor is close to the fixed constraint reducing the measured displacement. Both the size of the measured displacement and the insertion depth at which the maximum occurs are influenced by the sensor stiffness. Both shift down with increasing stiffness. The observations above lead to the conclusions that firstly, the sensor should be as soft as possible to achieve the highest possible resolution. Secondly, the insertion depth has a large influence on displacement and should therefore be included as a second calibration parameter. The resolution achieved is highest between 100 and

150 mm distance from trocar to instrument base. When the length of the instrument is increased this window also shifts up. The maximum displacement increases steeply for longer instruments. Changing the length of the trocar has a smaller but also significant effect. Shorter trocars constrain a smaller length of the instrument and therefore allow for a larger displacement.
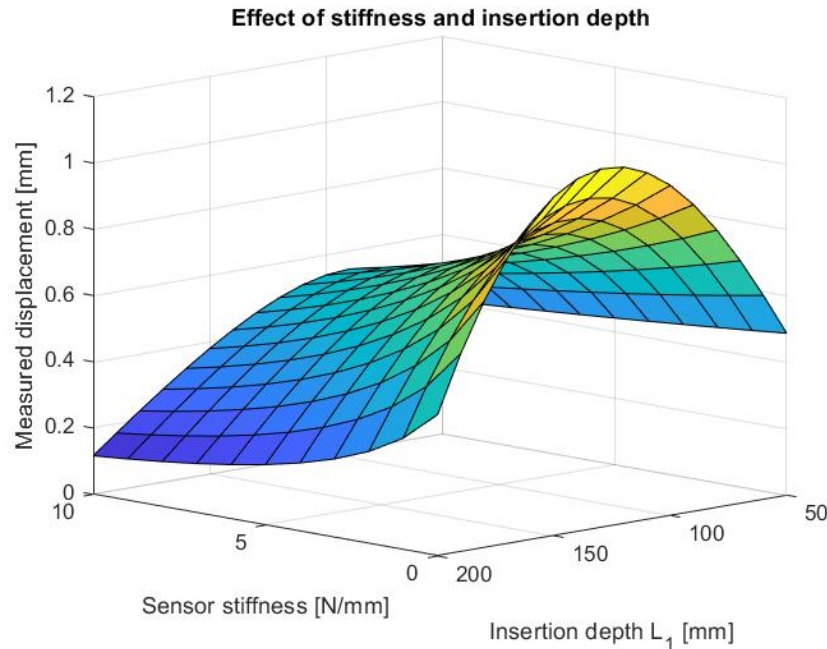
Figure 4.7: Surface plot showing the effect of sensor stiffness K and insertion depth L_1 on measured displacement V_b.

For the stiffest sensor included in the analysis the displacement ranges from 0.12 to 0.55 mm. At zero stiffness the displacement ranges from 0.43 to 1.08 mm. It must be noted that a relatively wide range of insertion depth is used for the surface plot. In reality the insertion depth will be more concentrated in the middle of the range. Additionally, the sensor will be made from the softest possible rubber. Therefore, the displacement in the final sensor is likely to be around 0.5 mm. This estimation of the displacement will be used to determine the appropriate magnetic system, magnet size and distance from the sensor. When the design is finished and the sensor stiffness is known a more accurate displacement will be calculated. If the displacement considerably deviates from the current estimation a design iteration will have to be made.

4.6.2. Magnetic configuration

Several different magnetic configurations can be used to measure the displacements using Hall sensors. It is possible to use pole pieces, pieces of iron that concentrate the field, in several different ways or not at all. It is also possible to measure with four instead of three magnet-sensor pairs and perform differential measurements. Since all of these measures influence the magnetic field and potentially improve sensor performance they have to be investigated. The configuration is investigated first because, while it does influence the other remaining parameters, it can be decided on without knowing the exact sensor, magnet and compliant material being used. The main goal in selecting the magnetic configuration is to achieve a change in output that is as large as possible for each unit of displacement i.e. sensitivity. A COMSOL analysis was performed to investigate and compare the available configurations. An extensive description of the model setup and results can be found in Appendix A. This section will summarize the results and draw conclusions.

First it was investigated whether adding a sensor-magnet pair and performing differential measurements results in a significant improvement of measurement performance. The results described in Appendix A indicate that performing differential measurements yields an almost 50% larger change in sensor output. This is less than the theoretical doubling of the output with differential measurements.

From the results it is concluded that this is the result of the non-linear relation between field density and displacement as well as field attenuation which occurs when the magnets are placed closely together. Differential measurements also allow for the rejection of errors induced by parasitic displacements. Because of the improved sensitivity, linearized output, temperature compensation and error rejection that differential measurements provide it is decided to use this in the final design.
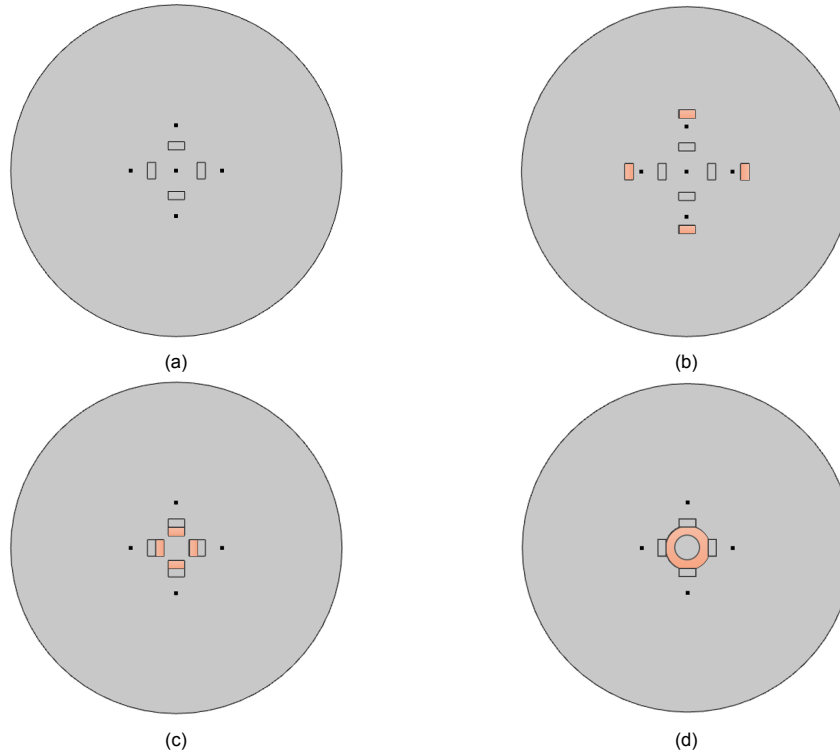
Figure 4.8: Geometries of the different differential configurations investigated in COMSOL. With (a) differential without pole pieces, (b) differential with pole pieces at the outer perimeter, (c) differential with magnet and pole piece stacked, and (d) differential with a ring shaped pole piece. Pole pieces are shown in red.

In the next step three different configurations with various implementations of pole pieces were compared to the normal differential configuration. One configuration uses pole pieces placed behind the hall sensors, one has pole pieces stacked with the magnets and the last one uses a ring shaped pole piece around which the magnets are placed. Only this last configuration appears to result in a significant increase in the field density and output, although at the cost of a less linear output curve. Because of the reduced linearity and the added complexity in terms of production and sterilization of adding an iron pole piece it was decided not to use pole pieces in the prototype sensor.

4.6.3. Reducing field attenuation

In the previous section it was concluded that placing four magnets close to each other results in a relative reduction in the field measured by the sensor. To improve the SNR and sensitivity however, it is the goal to have the field as strong as possible. In Appendix A two different measures to reduce the field attenuation were investigated. The first involves placing the magnets further from each other by increasing the magnet-center distance. Simulations show that increasing the distance has a large but diminishing effect on the field density measured by the sensor. Based on the results it was decided to increase the magnet-center distance to 4 mm by placing a solid ring at the inner diameter. Not only does this increase the field density, improving the SNR, it also helps in production by offering something to align and hold the magnets before pouring the compliant material.

The second strategy to reduce field attenuation flips the magnetic field of one pair of opposing magnets. Figure 4.9 shows the magnetic field plot that results from doing so and compares it to the

normal situation. It shows that in the normal situation (Figure 4.9a) the four magnets create eight field loops. When the magnetic field on one pair is flipped (Figure 4.9b) only four field loops are created. As a result the field measured by the Hall sensor is much stronger when the magnets are flipped. In fact, the field density is more than double measured perpendicular to the sensor. It does however not improve the relative output of the sensors. Nonetheless, because of the much higher field density and improved SNR it was decided to use this concept in the final design.

(a)   (b)
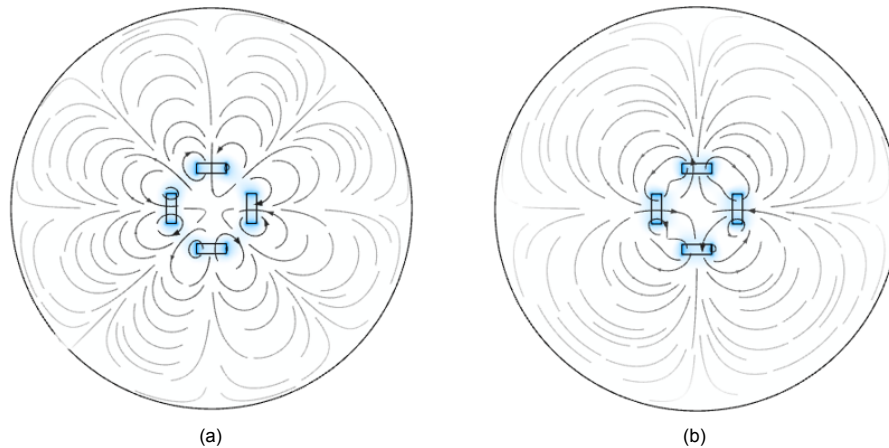
Figure 4.9: Plot of the magnetic fields when (a) all the magnets are magnetized in the same direction (north pole pointing inwards) and (b) one pair of magnets are flipped (north pole pointing outwards).

4.6.4. Sensor, Magnet and geometry

After selecting the magnetic configuration and finding ways to reduce the field attenuation it is possible to determine the sensor, magnet and geometry required. Before describing the final choices the general types of sensors and magnets available are discussed.

There are two main types of Hall sensors: analog and digital. Analog sensors, also known as linear or ratiometric sensors, have an output voltage that is proportional to the magnetic field density. Digital sensors have an output that jumps from zero to a certain voltage when a threshold value of the field density is exceeded [41]. For this project an analog Hall sensor is required. Among the analog sensors there are still more different products. Hall sensors are available with different input and output voltages. For this project a 5V input and output will be used for easy integration with an Arduino. The other differences between different models are the sensitivity and saturation field density. The saturation field density refers to the magnetic field density above which the sensor output is no longer linear. Sensitivity of the Hall sensor is defined as the change in output in millivolt for a change in magnetic field density of one Gauss. The product of these two values (saturation and sensitivity) generally remains constant from one model to the other. This means that a higher sensitivity comes with a proportionally lower saturation field density. The different saturation densities available are 420, 670 and 840 Gauss. The corresponding sensitivities are 5, 3.125 and 2.5 mV/Gauss respectively. Whatever the saturation field density of the sensor may be it is important that as much of the available range is being used to maximize the final force resolution.

For the magnets there is also a wide array of options. Magnets are available in different materials, shapes, sizes, and magnetization patterns and grades. Neodymium magnets are the strongest and smallest commercially and cheaply available magnets which is why they are the only type of magnets considered for the sensor [41, 42]. Neodymium magnets come in several magnetization grades. These grades range from N30 to N52 and indicate the increasing magnetization strength of the magnet. The most common magnetization grade in commercially available magnets is the N45 grade. Therefore the N45 grade is used for the magnetic analysis in Appendix A.

Appendix B describes a study performed to investigate the sterilizability of Neodymium magnets. It is important that the magnetic field of the magnets does not change significantly after sterilization

cycles. It was found in literature that magnets do incur permanent losses of magnetic field density after a thermal cycle. These permanent losses however, should stabilize after subsequent thermal cycles of similar heat and duration. A simple experiment was performed to confirm these behaviours. For the experiment ten Neodymium cube magnets were exposed to a thermal cycle similar to that of a steam sterilization using a kitchen oven. The experiment showed about a 10% loss in field density after a single cycle. The field density however stabilized at the same level after subsequent thermal cycles. This confirms that, when preheated, Neodymium magnets can be sterilized without further changes in field density.

The main problems in settling on the required sensor, magnet and distance are the following. Firstly, the largest change in magnetic field density occurs when the magnet and sensor are initially very close together (around 1 mm distance). This is clearly shown in Figure 4.10 which plots the calculated field density over distance for various magnet sizes. It also shows that at short distances the effect is strongest for the smallest magnets whereas at larger distances the changes are larger with bigger magnets. It would therefore be desirable to have the distance as small as possible and use small magnets to achieve the highest sensitivity. The sensor however has to be produced with a compliant layer between sensor and magnet. When the compliant layer is very thin its effective stiffness increases. This leads to a reduction in deformation sensitivity. Moreover, putting together and placing the magnets and pouring the compliant material into a gap of only 1 mm will be challenging. To be able to produce the sensor cheaply and achieve a high sensitivity a larger gap is required. It was decided that a gap of at least 3 mm should be enough to pour the material effectively and still maintain a sufficiently high field density. Based on Figure 4.10 it is also possible to conclude that a 3 mm gap requires bigger magnets in order to have an adequate field density and sensitivity at the Hall sensor. Therefore, only the ⌀2x1 mm, ⌀3x1 mm and ⌀4x1 mm will be considered.
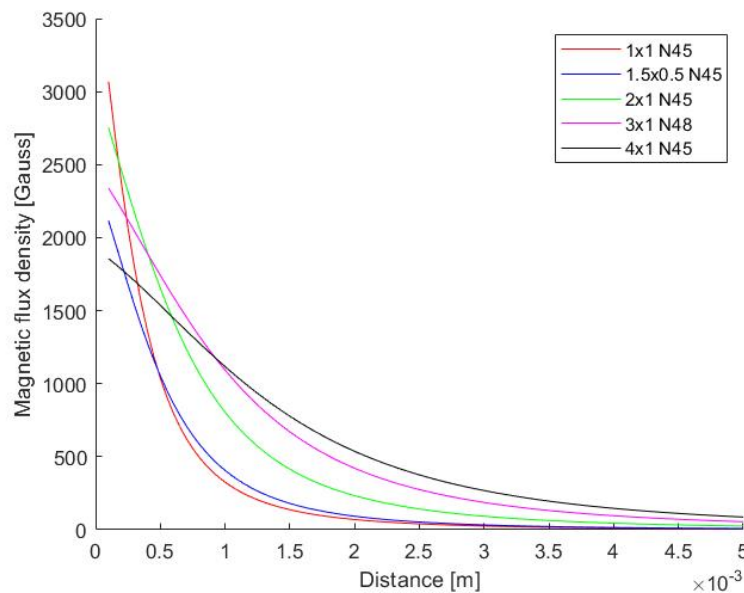
Figure 4.10: Analytical estimation of magnetic field density over distance for commercially available disk magnets.

To settle on a final magnet size, sensor and distance, combinations of the three magnets listed and distances of $3, 3.5$ and 4 mm were simulated over a 0.5 mm displacement. The resulting minimum and maximum field densities were used to match a suitable sensor. For the decision a number of things had to be taken into account. Firstly, the sensitivity and SNR have to be as high as possible. For this the difference between minimum and maximum had to be as large as possible and the maximum has to be close to the range of the sensor. Secondly, assuming that 3 mm is sufficiently wide for the production of the sensor, smaller magnets and distances are preferred to keep the sensor as compact as possible. The optimal combination found consists of a ⌀3x1 mm magnet at a distance of 3 mm from

the SS495A sensor by Honeywell that has a magnetic field range of ± 670 Gauss and a sensitivity of 3.125 mV/Gauss. The sensor is also temperature compensated resulting in a very small change in null voltage and sensitivity.

4.6.5. Compliant material and sensor stiffness

What remains for the detailed design is the choice of composition and stiffness of the compliant material that will be placed between the magnet and Hall sensor. The most used medical grade elastomers are silicone and polyurethane rubbers, although many more specialized compounds exist [43, 44]. Both of these materials can be moulded into shape and are available in a spectrum of stiffness. For silicone the spectrum is shifted more towards the soft rubbers and for polyurethanes the spectrum is shifted towards harder plastics. Both materials are sterilizable through autoclave sterilization [45, 46, 44]. Although Polyurethane has a better fatigue endurance it is also more susceptible to degradation compared to Silicone rubber [43, 44]. Finally, polyurethane is very 'sticky' during and after moulding which requires the use of releasing agents. Additionally, polyurethane is highly sensitive to humidity in the production process [43]. Finally, silicone rubber is generally more expensive compared to polyurethane rubber.

After evaluating the two options it was decided to use silicone rubber. The most important reason is the fact that silicone rubber is available in softer compounds. This is based on the conclusion in Section 4.6.1 that the sensor should have the lowest possible stiffness. The softer the sensor the higher the overall sensitivity. The stiffness of rubbers is indicated by the Shore hardness scale. The softest part of the scale is indicated with the prefix 00 and runs from shore 00-10 (comparable to a gummy bear) to shore 00-100 (comparable to the heel of a shoe). Since the silicone rubber also has to keep together the sensor it was decided not to go for the softest rubber available but instead use a shore 00-20 rubber.

With the material known it is possible to determine the overall sensor stiffness and calculate the expected deformation for the prototype. The sensor stiffness was determined through a second COMSOL analysis which is described in Appendix C. The COMSOL analysis uses material parameters for shore 00-30 silicone because the parameters were not available for 00-20. The stiffness found in Appendix C equals 1.9 N/mm. Since this is for a slightly stiffer rubber a stiffness of 1.75 N/mm is assumed for 00-20. Also relevant is the fact that the force-displacement relation is highly linear. This means that the linear spring assumption used in Section 4.6.1 is valid. It also helps to produce a more linear input-output relation of the complete system.
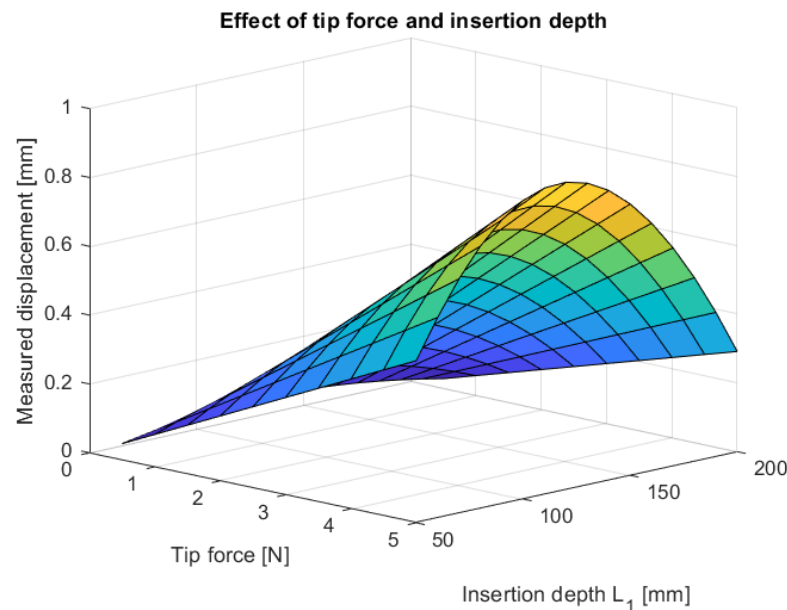
Figure 4.11: Surface plot showing the effect of tip force F and insertion depth L_1 on measured displacement V_b.

Now that the stiffness of the sensor is known more accurately it is possible to determine the expected displacement as a function of insertion depth and tip force. Figure 4.11 shows the measured displacement V_b for different combinations of tip force F and insertion depth L_1. The maximum displacement occurs at $F = 5$ N and $L_1 = 120$ mm and is 0.89 mm. The average displacement over the insertion depth at $F = 5$ N is 0.69 mm. The force-displacement relationship observed in the figure is linear with a slope/stiffness depending on the insertion depth. This allows for the determination of an overall system stiffness for a given insertion depth. At $L_1 = 120$ mm the system is the softest and has a stiffness of 5.6 N/mm. The system stiffness can be related to the insertion depth through a fourth-order polynomial with $R^2 = 0.99$. Since all the force-displacement curves intersect at zero this polynomial is sufficient to determine the measured displacement for any tip force and insertion depth.

4.6.6. CAD model

Figures 4.12 and 4.13 show the design of the trocar and sensor. In a true final design the Hall sensor leads would not extend up but rather into a PCB that combines source, ground and the four output leads into a single cable. The design shown in these figures uses a snap-fit mechanism to constrain the sensor to the trocar. The cross-section view in Figure 4.13 also clearly shows the seal used to make the sensor trocar system air-tight. The seal consists of a thin rubber disk that is squeezed between the trocar and sensor. The hole that passes the instrument has a slightly smaller diameter to provide a seal around the instrument which remains under deformation. The sensor and seal are both separate from the trocar in order to ensure reusability through steam sterilization. Figure 4.13 also shows that the inner and outer ring include some steps. These steps are used to let the silicone flow over and under and provide some additional axial stiffness to keep the sensor together. Finally, it must be noted that the design presented here does not clearly include a way in which to connect it to the robot arm as is needed for the fixed trocar strategy. Ways to fix the trocar include clamping around the 'stem' of the trocar or by including extrusions on the trocar which can be used to attach as shown in Figure 4.3.

(a)   (b)
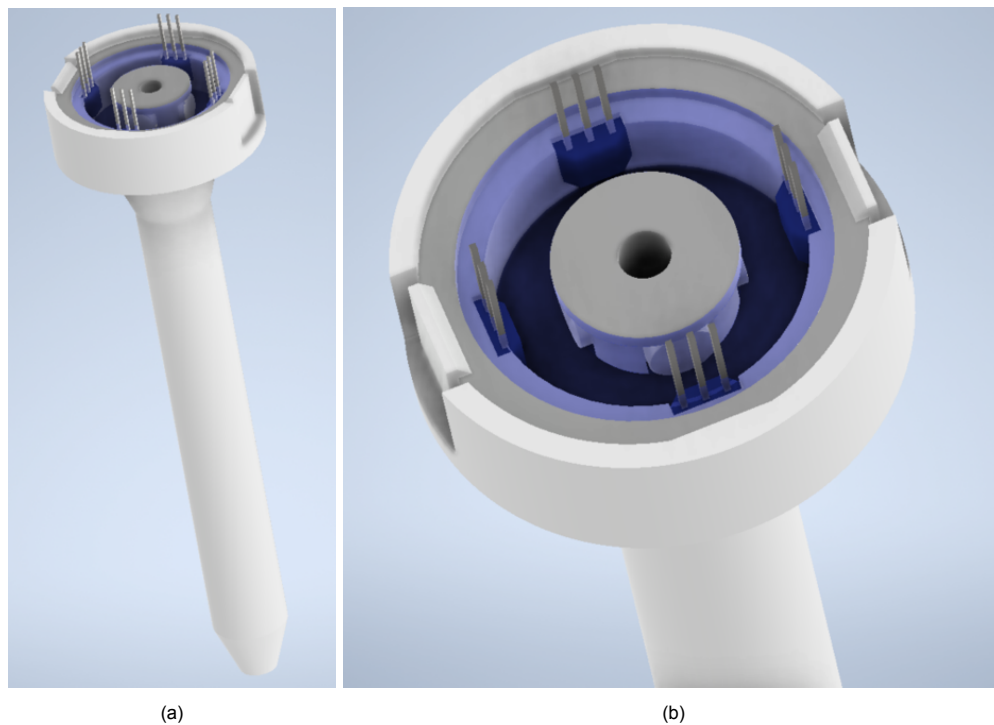
Figure 4.12: CAD model of proposed sensor and trocar. (a) Full view (b) Close up of sensor.
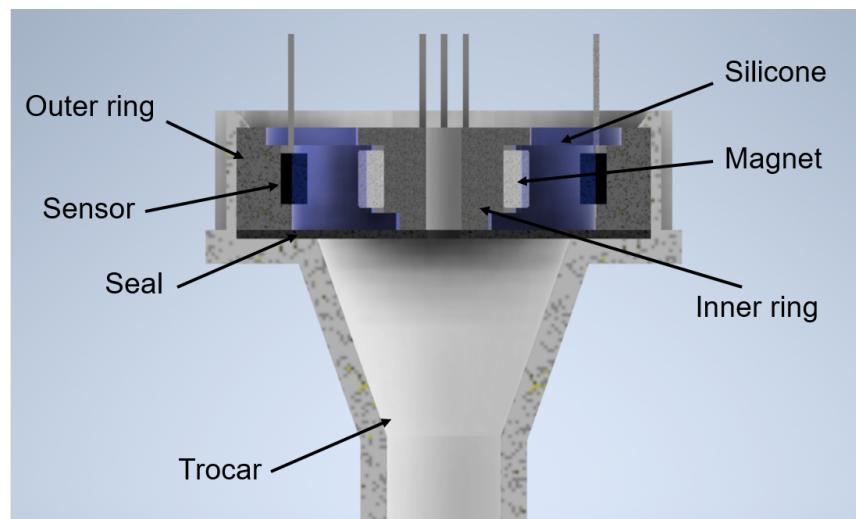
Figure 4.13: Cross-section of the sensor trocar system with annotated components

5

Prototype & Experimentation

A prototype was constructed as a proof of concept for the design presented in chapter 4. This chapter fist covers the electronics, software and mechanical construction of the prototype. Next, the experimental setup and method of data acquisition and processing are described. Then the method of calibration is explained and the performance is evaluated. Finally, the sterilizability of the sensor is investigated.

5.1. Software and Electronics

The sensor prototype is controlled from an Arduino Uno. The electronic circuit necessary to read the output from the analog Hall sensors using an Arduino Uno is simple. The sensors are connected to the regulated 5 Volt direct current output and ground on the Arduino as indicated in Figure 5.1 with the red and black wires. A third wire connects the sensor output to an analog input pin on the Arduino. When four Hall sensors are used their input and ground are wired in parallel and they occupy four separate analog pins as shown in Figure 5.1.
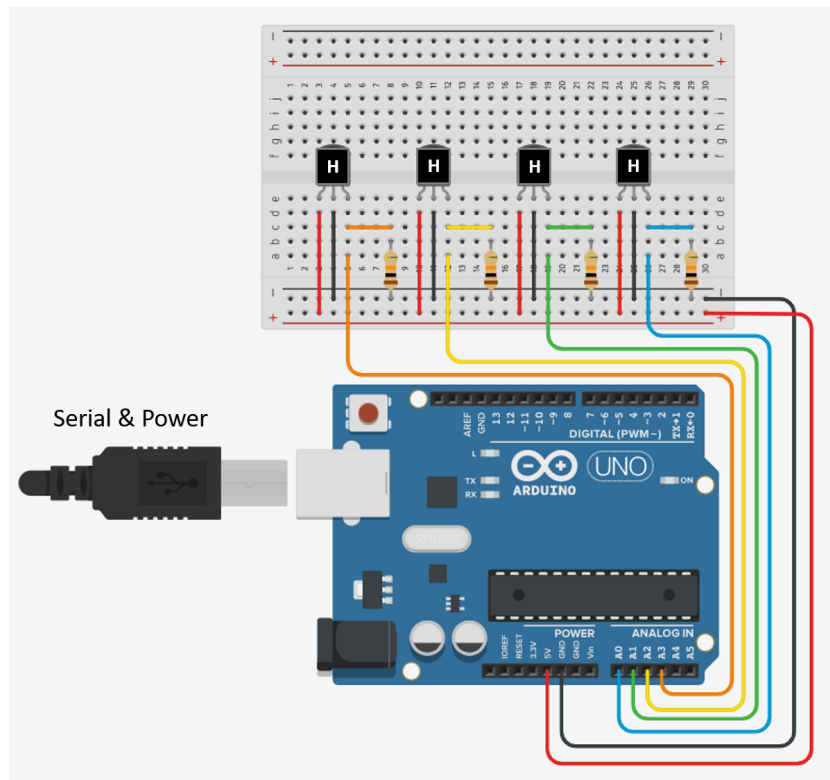
Figure 5.1: Schematic drawing of the sensor circuit

The figure also includes pull-down resistors wired parallel to the output pin. This is a noise reducing measure which will be discussed in Section 5.1.2. What Figure 5.1 doesn't show is that the Hall sensors are connected through jumper cables from the breadboard in order to reach the sensor. The Arduino is programmed to acquire data through the Arduino compiler using C++. The data is retrieved and processed in Matlab.

5.1.1. Bandwidth

At first, four Hall sensors were connected and programmed to send their data directly to Matlab through the serial connection using the *Matlab support package for Arduino hardware*. However due to the relatively large delays in the support package code the achieved bandwidth was only 8Hz. Since this rate is below the required bandwidth (Section 4.1) a different approach was taken to achieve an acceptable bandwidth.

The second approach does not use the support package but instead establishes a direct serial connection using the *serial* function in Matlab. On the Arduino side the C++ code used is very simple. It opens the serial port and waits for a command. Whenever a specific command is sent from Matlab, the Arduino reads the four analog pins and sends the results back through the serial connection. After establishing the connection with the Arduino, Matlab cycles for a specific amount of time, sending and receiving every cycle. The time is recorded after receiving each sample and stored in the fifth column of the matrix using the *tic toc* function in Matlab. Using this method a bandwidth of around 140 Hz can be achieved. When every tenth sample is plotted in a live plot in Matlab a bandwidth of about 120 Hz remains. This bandwidth is more than sufficient for the prototype. The code described here is included in Appendix D.

5.1.2. Noise mitigation

To determine if any noise mitigating measures are necessary an investigation was performed to identify the type and level of noise. This investigation was performed using only a single Hall sensor instead of four. First, the Allan deviation was computed to verify the absence of any noise other than white noise and investigate the influence of drift. The result is shown in Figure 5.2.
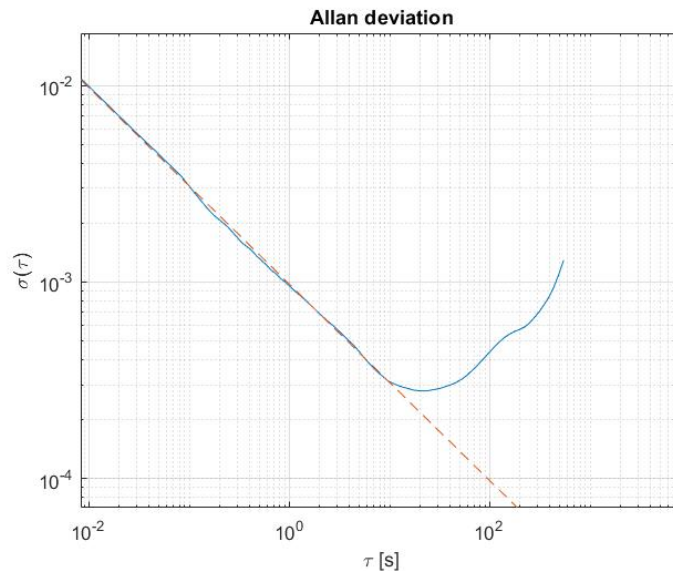
Figure 5.2: Allan deviation plot of a static Hall sensor measurement obtained over 20 minutes. Dashed line indicates theoretical slope for white noise. Y-axis displays Allan deviation in Volt.

Figure 5.2 shows the Allan deviation plot for a static measurement of 20 minutes (1200 seconds) obtained without a magnetic field source placed near the sensor. The x-axis represents the size of the averaging window (τ) and the y-axis of gives the standard deviation between subsequent windows in Volt. The plot includes a dashed line with a slope of $-\frac{1}{2}$ which is the slope that corresponds to white

noise [47]. The plot indicates that up to windows of about 10 seconds the deviation between subsequent windows will decrease following the dashed line. This signifies that up to that point the only present noise is white noise. For windows larger than 25 seconds the variation will start to increase again due to drift. This drift could be caused by several different factors. It could be due to changes in room temperature (the room was not temperature controlled), changes in background magnetic field or heating of electronics. At the maximum bandwidth the standard deviation between individual measurements is 11 mV. If the signal were to be averaged to just meet the minimum bandwidth of 20 Hz the standard deviation would be about 4.5 mV.

Although the noise level for a single sensor is already relatively low, two simple methods of reducing the effect of noise were investigated. The first method is to apply a Moving Average Filter (MAF). This noise mitigating measure is implemented completely in code and does not require any hardware. For the MAF a specified number of previous samples are used to compute a current, averaged output. Although this type of averaging does not necessarily reduce the rate at which a new output is provided it does smooth the signal. This smoothing means that very short duration and low amplitude events become more difficult to register. The more previous samples are used the more smoothing occurs. It is also possible to use a Central MAF. This filter uses both previous and 'future' samples to compute a current output. Generally CMAF gives a better estimate of the true value. However, the consequence is a slight delay in the output because future values have to be included in the average. Figure 5.3 shows the standard deviation of the entire measurement for different numbers of samples used by the moving average.
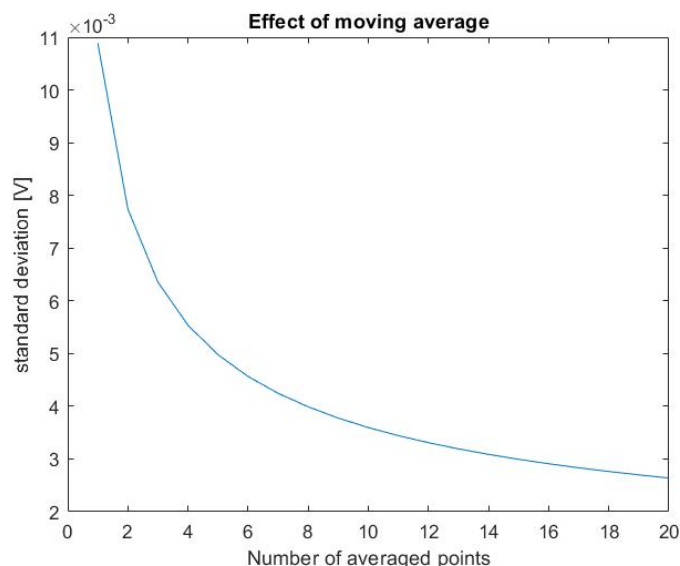
Figure 5.3: Plot showing the effect of moving average filter on standard deviation.

Without averaging, the standard deviation is 11 mV. If three previous samples are used to compute the current average the standard deviation is halved to 5.5 mV. Afterwards, the returns of adding more samples diminishes quickly. To get a sense of the Signal to Noise Ratio (SNR) of the Hall sensor with and without MAF they are plotted together in Figure 5.4. To obtain the plot a magnet was moved away from the Hall sensor in steps of 0.5mm and 300 samples were collected at each distance. The data obtained is plotted 'raw' and after a four point moving average was applied. The plot clearly shows that a four point MAF is very effective at improving the SNR. Because of this ability, and the fact that no changes of force above 20 Hz are expected, the MAF is a feasible option to reduce noise if deemed necessary. Whether it is necessary and the number of averaging points will be determined during final experimentation.

The second option for removing noise is called a pull-down resistor. A pull-down resistor is used to connect the sensor output to the ground parallel to the connection with the analog pin on the Arduino. Since the input pin has a very high impedance the pull-down resistor is able to conduct any unwanted
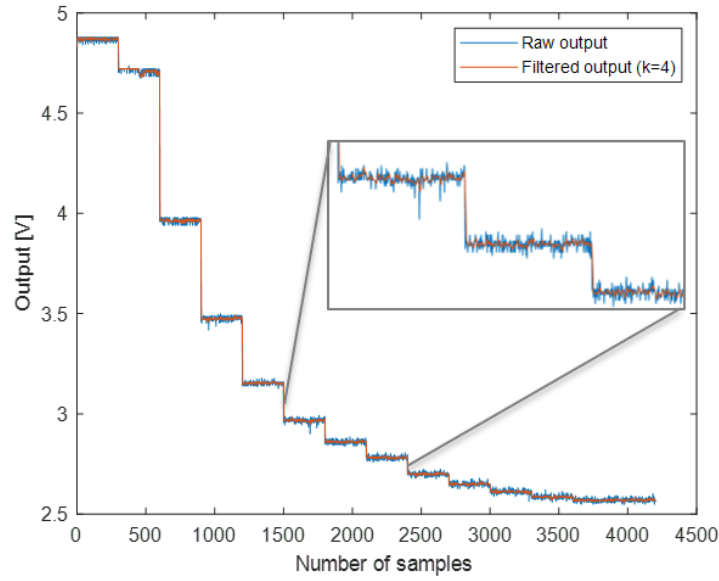
Figure 5.4: Plot showing the effect of moving average filter with a 4 point average on signal to noise ratio.

charges caused by noise so that they do not significantly alter the voltage on the input pin. Figure 5.5 shows the Allan deviation plot for a 20 minute measurement performed with a pull-down resistor installed. Comparing Figure 5.5 to Figure 5.2 it is clear that the pull-down resistor helps reduce the noise. The entire Allan deviation plot in Figure 5.5 is shifted down compared to the previous one. The maximum deviation for this case is 6.3 mV compared to the 11 mV before. Since the pull-down resistor provides a big drop in noise at little cost or added complexity it was decided to implement it in the final circuit for the sensor as shown in Figure 5.1.
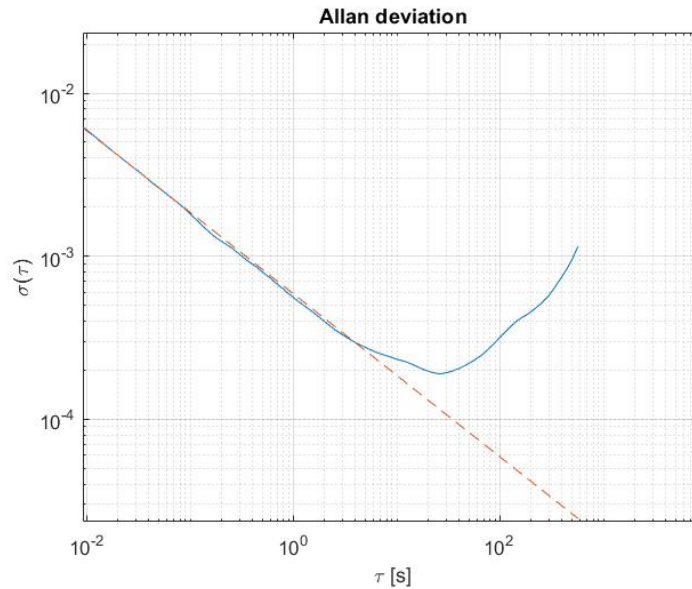
Figure 5.5: Allan deviation plot of a static Hall sensor measurement obtained over 20 minutes with a pull-down resistor. Dashed line indicates theoretical slope for white noise. Y-axis displays Allan deviation in Volt.

5.2. Magnet selection

Appendix A describes a COMSOL analysis performed to determine the initial choice of magnet size and sensor geometry. Apart from other design considerations described in Section 4.6 it was mainly this analysis that was used to determine that a ⌀3x1 mm magnet at a 3 mm distance from the sensor would result in the highest sensitivity. A simple experiment however revealed that the field density measured by the sensor in reality is different from both analytic and finite element predictions. For the experiment the magnetic field density from the magnet was measured at several distances from the Hall sensor in steps of 0.5 mm. The results are plotted together with the analytical and finite element prediction in Figure 5.6.
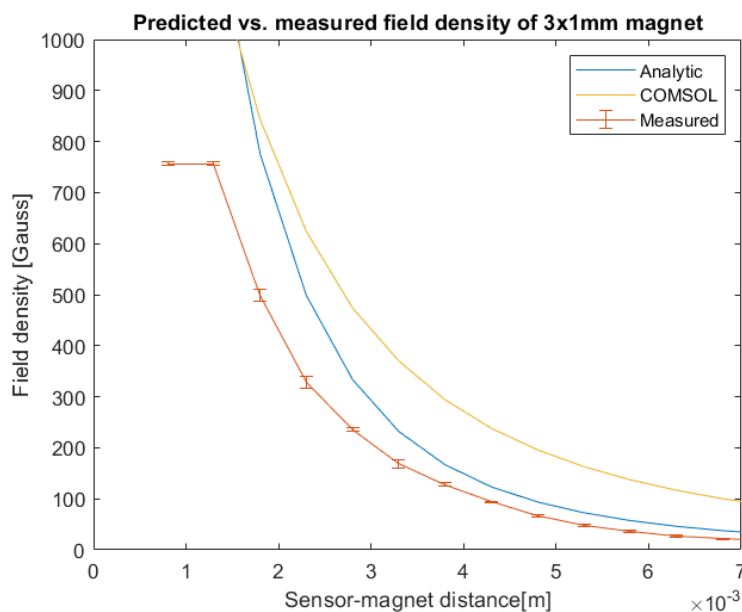
Figure 5.6: Comparison of the magnetic field density measured with the field density predicted analytically and through finite element for a ⌀3x1 mm magnet.

The figure clearly shows that the two predictions do not correspond with each other nor with the measurements. Looking at the shape of the curves it appears that the finite element prediction is shifted with respect to the measurements whereas the analytical prediction is scaled. A scaling factor of 1.5 is needed to make the measurement overlap with the analytical prediction. There are a number of possible causes for the large differences. Firstly, the COMSOL study was performed in a 2D approximation to save computation time. However, the magnets are disk shaped and therefore have a smaller volume in 3D compared to what was modeled in COMSOL. Secondly, in both the analytic and FE prediction the material is modeled as pure Neodymium with a certain remanent magnetization. In reality however, the magnet is coated with a Nickel-Copper coating. It is possible that the coating affects the measured field. Finally, the sensor also includes some margins. The null-voltage can vary with a ±3% margin around 2.5 V and the sensitivity can also differ from 3.125 mV/Gauss with a ±3% range according to the datasheet.

Such a big loss in field density compared to the predictions would translate to a big drop in sensor resolution. Therefore it was decided a different magnet size had to be selected. To make an appropriate choice of magnet, four different magnet sizes were measured using the same procedure used for Figure 5.6. The data obtained was used to calculate the differential output for each magnet at various initial sensor-magnet distances for a displacement of 0.5 mm. The calculated differential outputs are shown in Figure 5.7. The reducing output towards the left side of the figure for all but the ⌀2x1 mm magnet might appear to be incorrect. However, the drop in differential output is due to the fact that at small initial distances the Hall sensors become saturated. These points were kept in the plot because it is important to ensure that the sensors are not saturated. From Figure 5.7 it is clear that the ⌀4x1.5

mm magnet results in the highest output at the intended 3 mm sensor magnet distance. At a 3.5 mm distance it still has a higher output than the other magnets have at 3 mm.
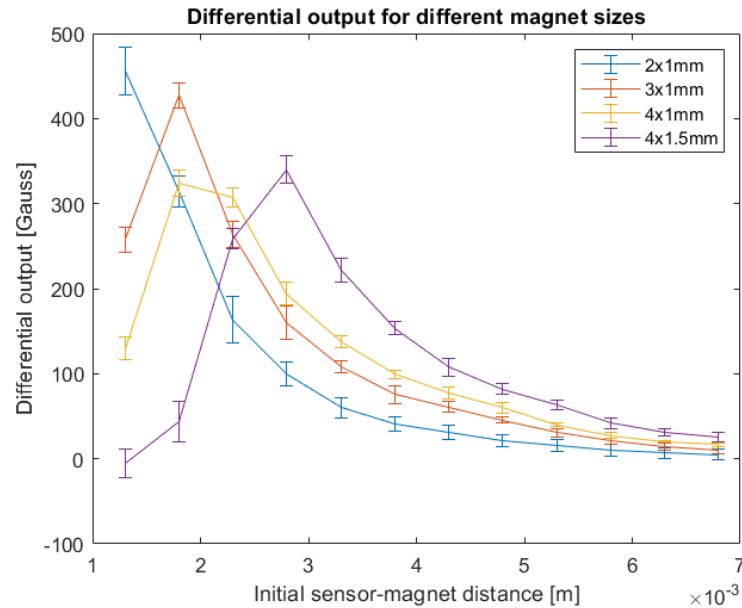
Figure 5.7: Estimated differential output for different magnet sizes with a 0.5 mm displacement.

5.3. Prototype design and production

Apart from the change in magnet size the design of the prototype is different from the design presented in Section 4.6.6 in two ways. Firstly, for simplicity, the seal that would be squeezed between sensor and trocar is not included. If it had been included it would have led to a slightly higher stiffness of the sensor which can be accounted for in calibration. Since the expected displacement is larger than the 0.5 mm designed for the higher stiffness would not lead to a problematic reduction in resolution. The second difference are the snap-fit elements which would be a convenient way of securing the sensor in real life. In the prototype however, the choice of material and production made this specific feature less desirable. Why that is the case will be explained shortly.

Including the back-end electronics a single sensor system consists of the following parts:

- Trocar

- Outer ring

- Inner ring

- Silicone layer

- Magnets (4x)

- Hall sensors (4x)

- Arduino uno

- USB A to B cable

- 10 kΩ resistor (4x)

- various lengths of wire and jumpers

The first three parts on the list were made from the same material and manufactured in the same way. They were 3D-printed using a stereolithography (SLA) printer. This printing method is able to print small features with high accuracy. The accuracy is important to allow for accurate form-closed alignment of the magnet and sensor as well as ensuring the four sensor-magnet distances are as close to equal as possible. The material used was the standard clear resin from Formlabs. This resin consists of Methacrylate monomers which link under exposure to UV-light to form an Acrylate-like polymer. After post-curing as prescribed by Formlabs, this material should be able to withstand the temperature of sterilization. This specific material however is skin-safe but not biocompatible. Formlabs does sell biocompatible material resins but the additional expense was not deemed necessary for this prototype. The downside of the SLA material and production is that products are relatively brittle and break easily in the print-direction. Because of the high likelihood of breaking it was decided not to include the snap-fit elements in the final prototype. A final product could use a different material or production method like ABS injection moulding that allows for an effective snap on placement of the sensor.

After printing and curing the inner and outer ring the magnets were glued to the inner ring and sensors to the outer ring using cyanoacrylate glue. As discussed in Section 4.6 two magnets were placed with their field facing out and two with their field facing in. The sensors were all mounted with the 'narrow' side facing in.

The next step in production involves the silicone rubber layer between the inner and outer ring. The silicone is bought as two liquid components which are mixed in equal parts and poured into the gap between the inner and outer ring using a syringe. As mentioned before the alignment between the magnets and sensors is very important. Since the silicone material is the only thing keeping the inner and outer rings together it is important they are aligned properly before pouring. To ensure this is the case a simple jig was made from lasercut PMMA. The jig is shown in Figure 5.8. By using flat faces on the outer ring and holes in the inner ring the jig makes for a form closed alignment tool. Needles are used to align the holes in the jig and inner ring. Additionally, the PMMA is a smooth surface which the silicone does not stick to. After pouring, the silicone is trimmed to be flush with the rings on both sides.
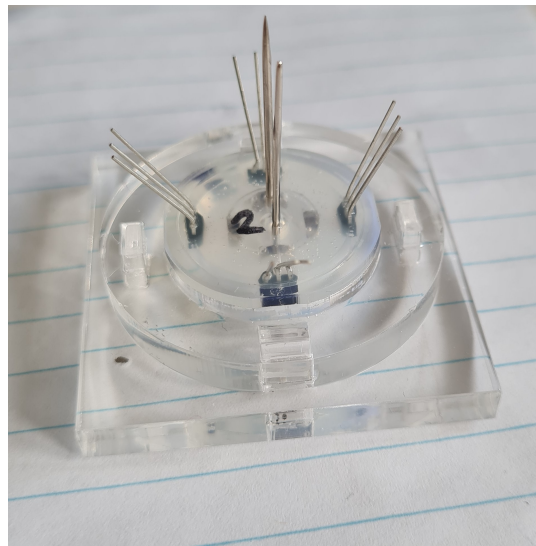
Figure 5.8: Image of the jig used for sensor production.

After producing the first few prototypes it was concluded that, due to a small amount of play in the jig, an additional alignment step had to be taken during production. After the outer ring and inner ring are placed in the jig the sensors are connected to the Arduino. The sensors are read-out to determine whether and how much the inner ring is off-center. Using live plotted data the inner ring is adjusted until the opposing hall sensors give approximately the same output. Due to differences in magnets and Hall sensors it is not feasible to aim for exactly the same output. Only when the outputs are approximately equal the silicone is poured.

Using the described production method four prototypes were made. The low-viscosity silicone rubber nicely flows into all the small corners of the enclosed volume. This also ensures that the sensor stays together well under axial load. The low viscosity also allows large air bubbles to escape. Only very small bubbles remain in the final product. These bubbles are distributed homogeneously throughout the silicone and should therefore have no significant influence on prototype performance.

5.4. Experimental setup

An experimental test setup was constructed to validate the proposed sensor and test it's performance. This setup is designed to represent the configuration depicted in Figures 4.2 and 4.4 in an experimental setting. The experimental setup and how it relates to the schematic drawings is shown in Figure 5.9.
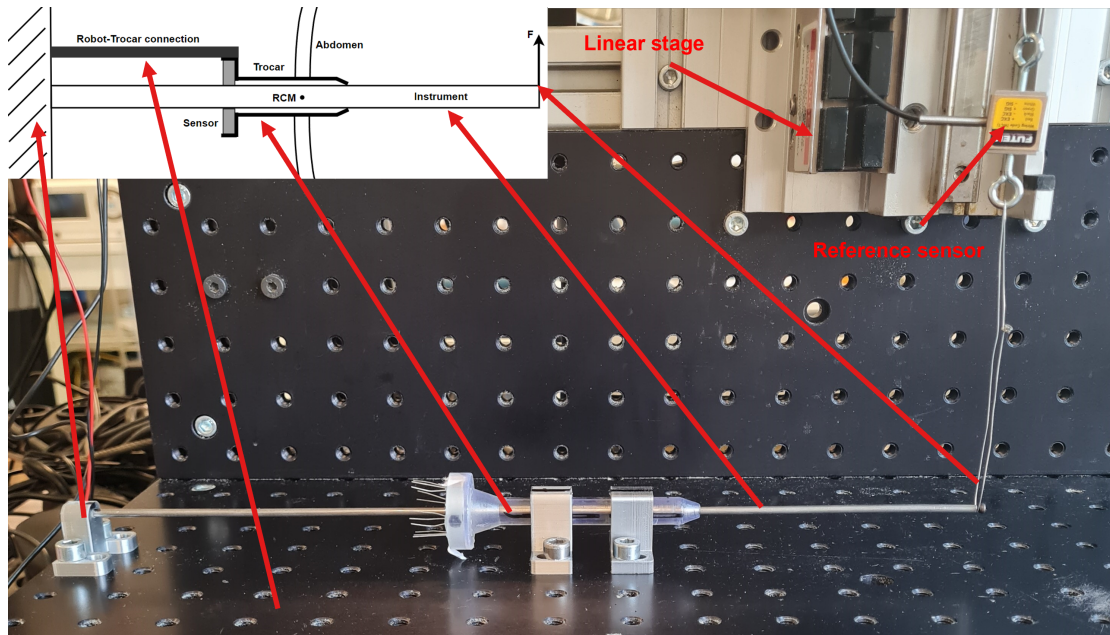
Figure 5.9: Image of the experimental setup used for sensor calibration and validation

The setup consists of the sensor placed in the custom trocar. The trocar is fixed to a base plate using two holders. The instrument is placed through the sensor and trocar and supported on the left side by a third clamp which is also fixed to the base plate. In the setup the base plate serves as the rigid connection between the 'robot arm' and trocar with sensor. The clamps used in the setup are printed using a standard Fused Deposition Modeling (FDM) 3D printer. To ensure a sufficiently high stiffness of the clamps a fill factor of 50% was used. The instrument used in the setup is not an actual SATA instrument. Instead, a 3 mm tube with a wall thickness of 0.6 mm was chosen to achieve the same stiffness as the actual 3 mm SATA. The tube is made from AISI 304 stainless steel which, similar to AISI 316, is non-magnetic and has a similar stiffness and yield limit.

To calibrate and validate the system an external force is applied using a linear motion stage. The linear stage moves in the direction perpendicular to the instrument and pulls on the tip of the instrument. The linear stage has a resolution of $0,01$ mm. To determine the actual applied force a reference sensor is used. This is done using a LSB200 series FUTEK load cell that is placed between the linear stage and the instrument. The reference sensor has a range of 2 lbs and a resolution of 14.8 noise free bits at a sampling rate of 2 kHz. The displacement of the stage and the reference sensor output are recorded together at the same sampling rate. To ensure an almost perfectly perpendicular force throughout the displacement range long wires connect the instrument to the reference sensor and the reference sensor to the stage. With this long connection (±250 mm) any axial displacement due to geometric non-linearity does not translate into a significant change in pulling angle.

5.5. Data acquisition and processing procedure

Because of the tools used and the way they were programmed to acquire data a specific procedure had to be devised to perform measurements and acquire data. As explained in the previous section, the linear stage and reference sensor are read out together. The system was programmed such that it only acquires data when the stage is moving up or down. As a consequence, each measurement consists of a trapezoidal type plot. First the system is static for five seconds, then it moves up with constant velocity of 0.5 millimeter per second. At the top it holds position for 10 seconds before moving down with the same speed. Finally it rests another five seconds before ending the measurement. The resulting outputs for both the prototype and reference sensor are shown in Figure 5.10. To get the prototype sensor output the voltage at the analog pin is converted to a 10 bit value which is then represented as an integer between 0 and 1024. Two of these integer values are subtracted to achieve the differential output plotted in Figure 5.10. This is repeated for two differential pairs to get an X and Y output.
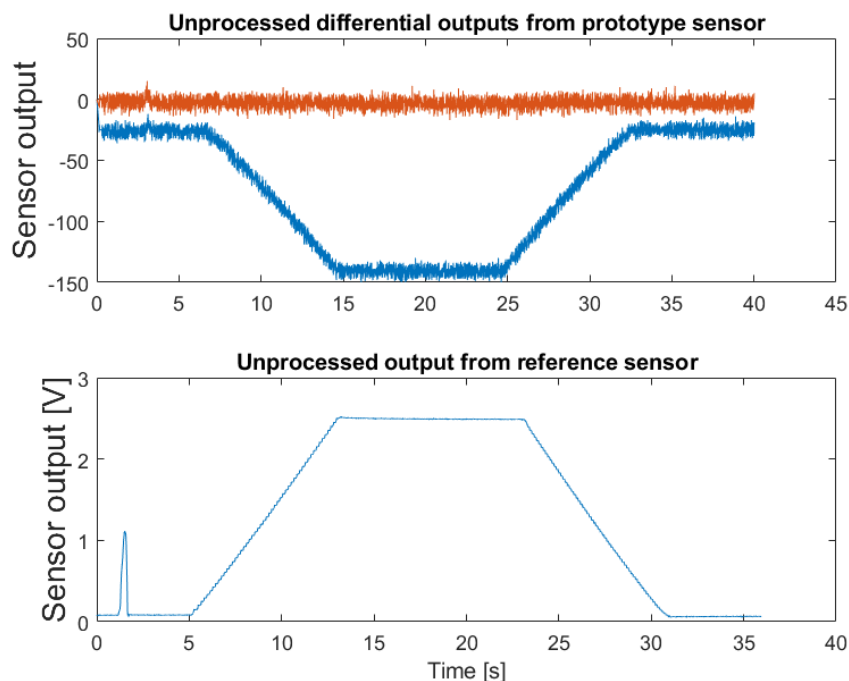
Figure 5.10: Raw, unprocessed data from single measurement for both the prototype and reference sensor for a 4 mm displacement at a insertion depth of 100 mm.

Since the prototype and reference sensor are sampled on two different devices the data has to be aligned in time. To this end, a single tap is given on the end of the instrument in the first five seconds of the measurement. The resulting peaks can be seen in Figure 5.10. The times of the peaks in both datasets are noted and used to trim the beginning and end of the datasets to end up with a measurement that is aligned and equally long in time.

After aligning the data in time the data is filtered using a twenty point central moving average filter. Note that with a bandwidth of 120 Hz this introduces a 0.083 second delay when plotting live data. After filtering, the offset in the data is corrected and the reference sensor data is multiplied with it's predetermined calibration constant. The resulting data is plotted in Figure 5.11. This figure provides a number of relevant insights. Firstly, the shape of the prototype data nicely follows the reference data, though with visibly more noise. This confirms that the deformations are indeed linear which was assumed in Section 2.3. Secondly, the differential output of the axis perpendicular to the pulling direction shows no visible change. This means there is no cross-talk when pulling perfectly perpendicular. Finally, there is no visible hysteresis in the prototype data.
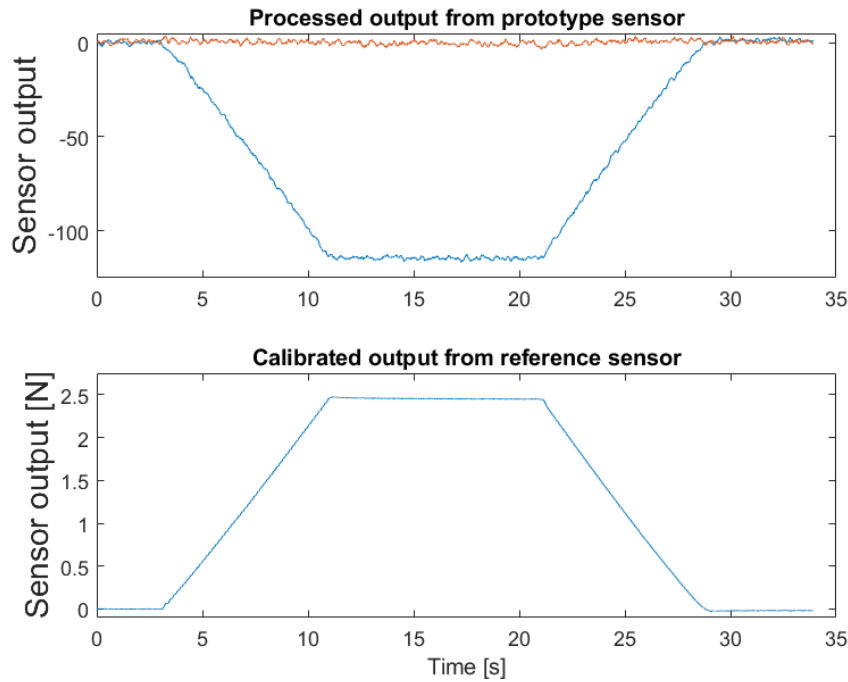
Figure 5.11: Processed, uncalibrated data from single measurement for both the prototype and reference sensor.

The final part of data processing involves resampling the reference sensor data. This is necessary to calculate the error of the calibrated data. To calculate the error, each point in the calibrated prototype data has to be compared to the point at the same time in the reference data. When the datasets are not of equal length this does not work. Since the reference sensor is sampled more often its dataset is about 40% longer and thus has to be shortened by resampling.

To determine the repeatability of the procedure five measurements were performed to the same displacement at the same insertion depth. When the height of the plateau for the five measurements is compared a standard deviation of 0.35% is found. This means that the measurement procedure is sufficiently repeatable for the intended use and permissible error.

5.6. Calibration & Performance

Before calibrating the prototypes there is a physical phenomenon that has to be considered when using silicone rubber. This phenomenon is called stress-softening or the Mullins effect. Stress-softening is defined as the reduction in stress caused by the same strain. This effect is most pronounced for rubbers undergoing their primary, large deformation [48]. For the sensor this means that the force-displacement relation will change slightly in the first few deformation cycles. In order to eliminate the influence of this effect each sensor was manually deformed a number of times before use.

In section 4.6.1 it was assumed that the force-displacement relation at a given insertion-depth is linear. The visibly linear output in Figure 5.11 shows this assumption is valid. This means that, having already corrected the offset, only a single parameter has to be found to relate the output of the sensor to the force applied. It was also found that this parameter varies approximately following a fourth-order polynomial when the insertion depth is changed. Because of the way in which the measurements are obtained a specific method is used to determine the calibration constant for each insertion-depth. Normally, one would measure the output at different levels of applied force and calculate a line through those points by linear regression. However, because of the way the linear stage is programmed this would take a lot of time and is also not necessary. Since both sensors are recording the force as the linear stage is moving to a certain position they are already capturing the output at a large number of

different force/displacement levels. By determining the slopes, up and down, for both the prototype and reference sensor through linear regression and calculating the ratio between the two it is possible to give an average calibration constant for many different displacements. To avoid any acceleration effects the beginning and end of each slope are not included in the linear regression. The final calibration constant is the average of the ratios found for the slope down and slope up.
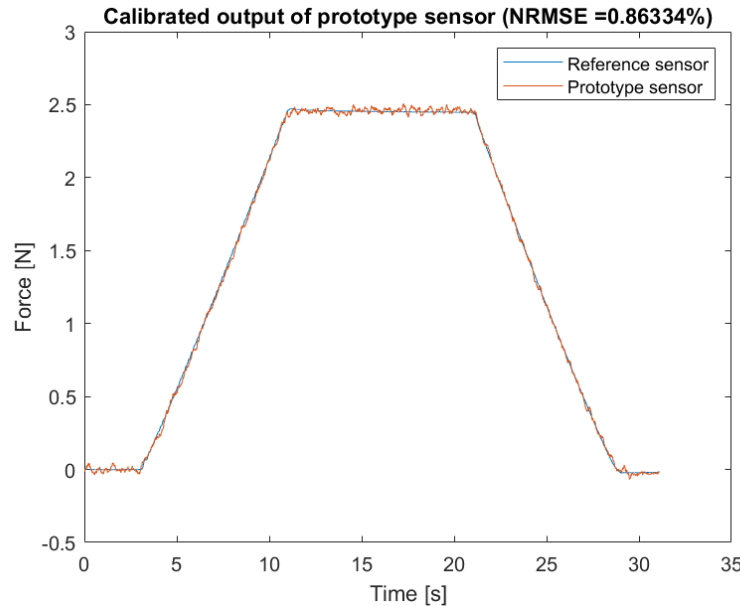
Figure 5.12: Calibrated sensor outputs from single measurement with a normalized RMSE percentage of 0.86%

The result of the described procedure on the data used in Figures 5.10 and 5.11 is shown in Figure 5.12. It shows that the prototype data accurately follows the reference sensor data with a Normalized Root Mean Square Error (NRMSE) of 0.86%. The Maximum Relative Error (MRE) is 3.1%. The MRE is evaluated only for the middle half of the samples to avoid outliers caused by dividing by values close to zero. The noise of this dataset was evaluated for the ten second period between upslope and downslope. In this period the standard deviation of the noise is 0.02 N leading to a maximum Signal to Noise Ratio (SNR) of $\frac{5N}{0.02N} = 250$.

The procedures of acquiring data and calculating the calibration constant were repeated for increasing levels of tip displacement building up to the maximum range of 5 N. With an integer output of around 240 for the maximum range a noise-free resolution of 0.021 N is achieved. The exact value found can vary slightly between measurements and prototypes. While finding that the maximum range of 5 N can be reached the measurements also revealed that the calibration constant found for increasing displacements decreases with between 2% and 8% depending on the insertion-depth. This is likely the result of the limited stiffness of the setup becoming increasingly pronounced with higher loads. The goal of the calibration is to determine a single parameter at each insertion-depth. Therefore it was decided to calculate every calibration constant at a load of approximately 2.5 N on the reference sensor. By fitting in the middle of the range the potential fitting error is minimized.

To capture the full behaviour of the sensor the calibration constants were determined for ten different insertion depths between 50 and 140 mm measured from the instrument tip to the trocar. A fourth-order polynomial was fitted to these points to be able to get a constant for every insertion depth. This process was conducted for both the X and Y axis of the sensor. Figure 5.13 shows that a fourth-order polynomial is indeed able to provide a good fit on the data for both axes with $R^2 > 0.999$. The data depicted in the figure also shows that there is significant difference in calibration constants between the X- and Y-axis. The Hall sensor outputs at zero displacement are quite similar for both axes. This means that production misalignment or magnetic field differences are not the cause. Therefore the
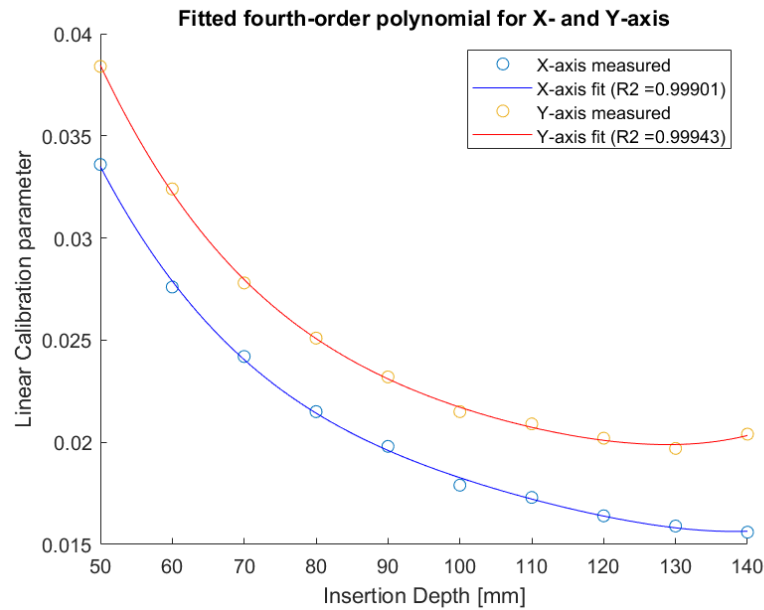
Figure 5.13: Plot showing the fourth-order polynomials fitted to measured calibration constants

difference is most likely the result of a difference in stiffness. Since the silicone was poured simultaneously and uniformly for both axes it is unlikely to have a different stiffness in both directions. The other factor which contributes to the stiffness is the 3 mm tube that represents the instrument. The tube had a slight permanent bend on arrival. This bend is most likely the cause of the difference in stiffness between the axes. To avoid errors introduced by this factor the instrument was not roted with respect to the sensor for all subsequent measurements.

A number of measurements were performed to verify the performance of the sensor after calibration. The measurements were taken at various force levels, insertion-depths and angles different from the calibration data. Table 5.1 lists the performed measurements and the errors found for the measurements. The angle is zero when aligned with the X-axis and 90 when aligned with the Y-axis. The data in Table 5.1 shows some important factors related to sensor performance. When comparing the results for the MRE and NRMSE one finds a large difference in error value found. This suggests that the MRE values represent mostly outliers in the data. Something not included in the the table but observed when processing the data is the fact that the location of almost all MRE locations is in the sloping sections of the data. This suggests that the MRE values are the combination of noise, fitting error and some lag in the data with respect to the reference sensor. Nonetheless, all but one of the verification datasets shows an error less than 10%. This one dataset has a deviatory shape with bellying in the sloping parts instead of a straight line. Because of the deviatory shape and statistically unlikely error values this dataset can be excluded. The reason for this deviation is unknown. When focussing on the NRMSE the maximum error found is less than 5%. All but three datasets have an error less than 2.5%. When the outlier is excluded the highest MRE is 8.80% and the highest NRSME is 3.41%.

Although not explicitly stated as a requirement the dynamic behaviour of the system was also briefly investigated. For this analysis a procedure similar to previous datasets was used. However, this time the speed was increased from 0.5 mm/s to 100 mm/s for a 4 mm displacement at a insertion-depth of 100 mm. Figure 5.14 shows the result using the same calibration parameters used for Table 5.1. Although with a slight and visible lag the prototype sensor is able to follow the quick displacement. The slight lag combined with the 20 point moving average filter does cause the sensor to not capture the overshoot. There is also a fitting error of about 5% causing the prototype sensor to underestimate the maximum force. It does appear however that the output of the prototype is slowly increasing to more closely match the reference sensor before dropping down again. Finally, the quicker action also seems to introduce a slight hysteresis visible at the end of the measurement.

Table 5.1: Performance results after sensor calibration. Angle defined as 0 when aligned with X-axis and 90 when aligned with Y-axis.

Angle [deg]	Insertion-dept [mm]	Displacement [mm]	Max load [N]	MRE [%]	NRMSE [%]
0	75	2	2.09	8.80	1.32
0	75	3	3.31	7.04	1.90
0	100	3	1.75	6.06	1.57
0	100	6	3.69	3.83	1.48
0	125	6	2.31	3.96	0.86
0	125	8	3.10	3.42	0.95
45	75	2.5	2.64	5.74	1.82
45	100	5	3.18	5.91	2.43
45	125	7	2.50	7.82	3.41
90	75	2	2.06	5.09	1.45
90	75	3	3.24	3.66	0.97
90	100	3	1.88	4.26	1.19
90	100	6	3.84	5.26	2.35
90	125	6	2.33	10.93	4.81
90	125	8	3.08	7.35	3.18

The last performance experiment performed considers the influence of parasitic displacements. The sensor is designed to measure forces in the radial direction. It is however possible that in a dynamic setting friction with the instrument will cause the sensor to move in the axial direction or create a planar rotation. Such parasitic movements could create unwanted errors in the measurement. Simple manual manipulation of the sensor to cause axial and rotational displacements revealed that these parasitic movements do not change the output. An out-of-plane rotation could not be accurately tested without accidentally causing in-plane translations. The insensitivity to parasitic displacements is the result of the differential measurement strategy. In differential measurements all changes that occur equally to all sensors are not registered. This is clearly the case for perfect axial translations and planar rotations and should also be the case for out-of-plane rotations.
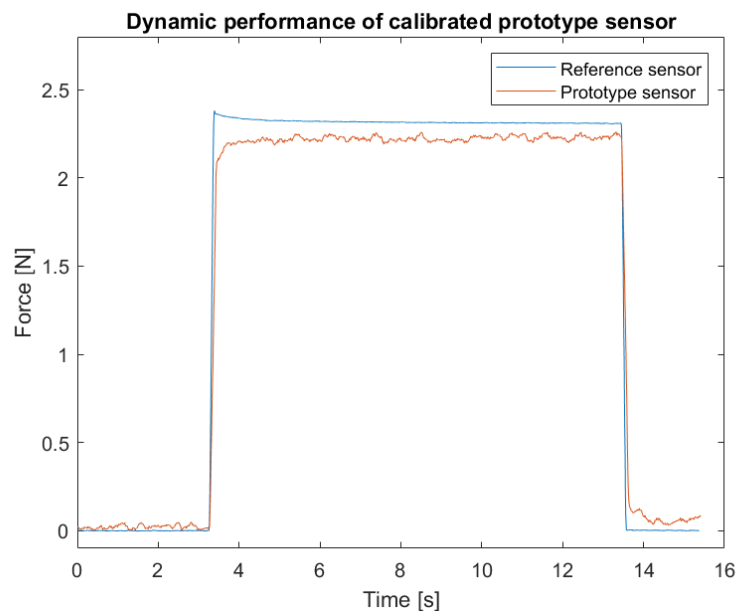
Figure 5.14: Plot showing the dynamic performance of calibrated prototype sensor

5.7. Sterilization

The goal of the project is to provide an affordable and reusable sensor. To this end all the separate components were chosen to be sterilizable by themselves. A simple experiment was performed to qualitatively verify whether the proposed sensor is indeed sterilizable. The experiment performed is similar to the one described in Appendix B. However, this time also the humidity is taken into account. Based on the results in Appendix B it is possible to make some predictions about the results. Since the magnets were not preheated a drop in static output is expected after the first thermal cycle. The drop however should be approximately equal for all sensor-magnet pairs and not result in a significant change in offset. Since the Hall sensors have a linear output this change in magnetic field will lead to a reduced range but should not lead to a significant change in the force-displacement relation represented by the calibration constant. For the second cycle it is expected that the static output and calibration parameters remain the same.

The experiment procedure was as follows. Two sensors were evaluated. First, before any sterilization, the calibration constants were calculated and averaged for both sensors at $2, 4$ and 6 mm displacement with an insertion-depth of 100 mm. Next, the sensors were removed from the trocar and Arduino and placed in the autoclave to undergo a sterilization cycle. The experiment uses the MELAtronic 15-EN+ class S autoclave suitable for sterilization of packaged and unpackaged loads according to EN 13060. The standard $134°$C program was used for all cycles. This cycle does not use a vacuum. The sensors were placed directly on the autoclave tray without any packaging material. After sterilization the sensors were let to cool down back to room temperature. After cooling down the calibration constants were determined again. This process was repeated three times.

Figure 5.15: Visual difference before and after sterilization

Figures 5.15 and 5.16 provide the results of the sterilization cycle(s). Figure 5.15 allows for a visual inspection of Sensor 2 after one sterilization cycle. It shows that the sensor is completely intact. There is however visible 'yellowing' of the 3D-printed rings but no visible difference in the silicone after sterilizing. Figure 5.16 shows the calibration constants that were found before and after sterilization for both sensors. It shows that the calibration constant more than doubles for both sensors after a single sterilization cycle equating to a halving of the output for the same load. After the first cycle the constants stay approximately the same for the second and third cycle. This large change cannot be explained by changes in the magnetic field. The magnetic field strength does drop with about 18% but does so uniformly for all sensor-magnet pairs. This will lead to a loss of range but not to such a change in calibration constant. The likely cause of the change is stiffening of the silicone rubber in the sensor. It is possible that the silicone was not fully cured and some post-curing or other structure changes occurred in the autoclave. A sign which points to incomplete curing is the fact that the silicone was slightly sticky before being sterilized. Alternatively it is possible that some small amount of water vapour diffused into the rubber which is slightly permeable to water.

Not shown in the figure is the fact that the standard deviation of the noise also increased from around 0.020 N to around 0.046 N, almost 2.5 times its original value. The noise also stays approximately con-

stant after the first sterilization cycle. The cause of this increase is more difficult to characterize. It is unlikely that the stiffening of the silicone results in higher noise. One possible cause of the noise is oxidation on the Hall sensor leads resulting in a noisy contact with the jumper-cable. Another explanation, which was not investigated in Appendix B, could be that the magnetic field becomes less constant and more noisy after sterilization.
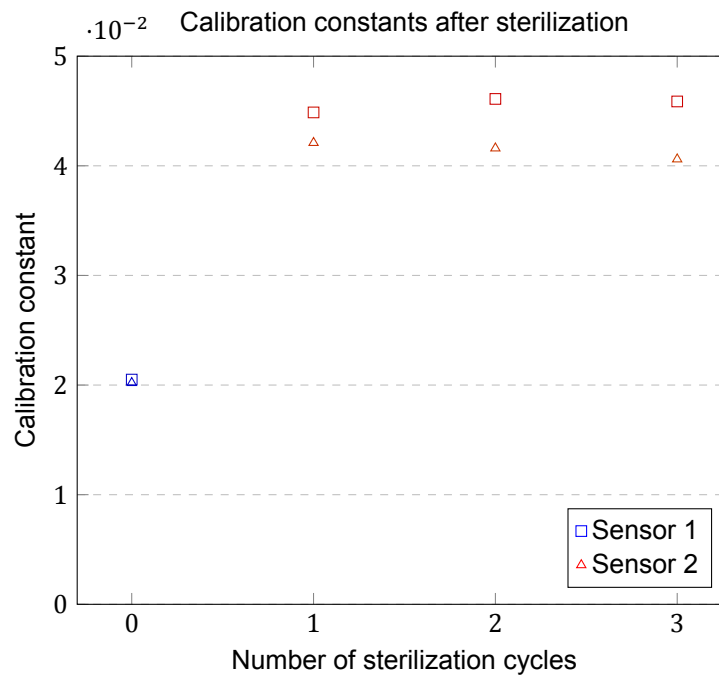
Figure 5.16: Calibration constants before and after sterilization for two sensors.

6

Discussion

This work aimed to design, prototype and validate a force sensing strategy for RAMIS using 3 mm instruments that is affordable and reusable. The sensing strategy is to be applied to limit forces during surgery or for feedback to the surgeon. In Section 4.1 a number of requirements were stated to ensure a design that is able to perform in the proposed application.

From the analysis of the problem in Chapter 2 it was concluded that special attention has to be paid to the trocar forces. These forces are large enough to significantly distort the measurement of the tip force and damage patient and instrument. The requirements imposed for the sensor to be reusable through autoclave sterilization are also limiting. In chapter 3 and the literature review on the topic [14] it was concluded that most solutions proposed in literature are either invasive and expensive or fail to effectively account for the trocar forces. Reusability was also often neglected. This subject was either not mentioned or the proposed solutions were stated to be single-use regardless of cost [14].

Based on solutions proposed by Fontanelli et al. [36] and Zemiti et al. [37] it was found that trocar forces can be effectively eliminated by fixing the trocar to the same rigid frame as the instrument and measuring the forces in or on the trocar. The sensor designed to be fitted to the trocar uses Hall sensors to measure changes in magnetic field density caused by bending of the instrument under load. Four sensor magnet pairs are used to perform differential measurements in 2D resulting in linearized output with a higher sensitivity and some error rejection. A silicone layer is cast between the Hall sensors and permanent magnets to act as both a transduction and structural element that keeps the sensor together. The Hall sensors and silicone rubber were chosen specifically for their low cost, simplicity and ability to be sterilized. The detailed design of the sensor is such that it optimizes the achievable range and sensitivity.

Prototypes and an experimental setup were constructed to quantitatively and qualitatively validate the performance of the proposed design and verify compliance with the requirements. The results from this experimentation indicate that the sensor is able to measure to the full required range of 5 N in 2D with a bandwidth between 100 and 120 Hz. This bandwidth is well above the required 20 Hz bandwidth which allowed for more aggressive filtering to reduce noise. The accuracy of the sensor was validated in terms of Maximum Relative Error (MRE) and Normalized Root Mean Square Error (NRMSE). To calculate the error the calibrated prototype sensor output was compared to the reference sensor output. Among the validation samples only one outlier had a MRE larger than 10% and all samples had a NRMSE of less than 5%. Neglecting the outlier for its deviatory shape it can be concluded that the accuracy requirement is met with the MRE less than 9.0% and NRMSE less than 3.5%. Although not listed among the requirements it was calculated that the sensor has a noise-free resolution of around 0.021N. It was also verified that because of the differential configuration the sensor is insensitive to parasitic displacements. The reusability of the of the sensor was tested trough repeated cycles in the autoclave. The results from this experimentation indicated that the proposed sensor is sterilizable although with a significant drop in range and resolution after the first cycle.

From the results summarized above it can be concluded that the prototype of the proposed sensor design is able to meet most imposed requirements including accuracy and reusability. It shows that tip forces are measured effectively by fixing the trocar to the same frame as the instrument which avoids influence of trocar forces on the measured tip force. The latter was also proven by Fontanelli et al. [36] and Zemiti et al. [37]. However, the sensor proposed by Fontanelli et al. is relatively affordable but not reusable and used inside the abdominal cavity [36]. The sensor by Zemiti et al. is not used inside the abdominal cavity but is expensive and not reusable. As Chapter 3 discusses, a sensor for RAMIS that is both accurate irrespective of the trocar forces as well as reusable has to date not been proposed.

While most requirements listed in Section 4.1 have been met there are two the design fails to meet. Firstly, the sensor is not able to measure forces in the axial direction. Rather, it is completely insensitive to axial forces and displacements. Since the most relevant forces for preventing damage to the instrument are the radial forces it was decided that not measuring the axial force is acceptable. It would however be possible to integrate an axial force sensor in another location such as the instrument driver-unit. The second requirement that is not fully met is the size requirement. Since the sensor measures bending in the instrument the custom trocar is designed with a certain gap to the instrument to allow it to bend. By including this gap the trocar has a larger diameter than standard 3mm instrument trocars. The design is however still smaller than a trocar for 5 mm instruments and much smaller than 8, 10 or 12 mm trocars. Moreover, decreasing the wall thickness by constructing a final production trocar from stainless steel instead of plastic and choosing a less conservative gap will allow the trocar to slim down to a smaller diameter.

With respect to the prototype and experimental setup there are a few noteworthy limitations. Firstly, as mentioned in section 5.3, the seal that would be necessary in a final product was not included in the prototype. If the seal is included the system stiffness would slightly increase leading to a slightly lower resolution. It is however possible to compensate for the increased stiffness in the sensor design. Secondly, the trocar used for the prototype is made from printed plastic. The printed plastic trocar has a limited stiffness. Any bending of the trocar during testing will have lead to a reduction of the measured displacement at the sensor. This is an additional reason to construct a final production trocar from stainless steel. Thirdly, as briefly mentioned in Section 5.6 the replacement instruments that were used in the setup had a slight bend. This resulted in significantly different calibration curves found for the two axes. It likely also caused some other inaccuracies. Another limitation related to the instrument is the fact that the actual instrument consists of three concentric shafts with small gaps between them. It is likely that the force-displacement curve will look different when using the actual instrument due to the gaps and friction occurring between the shafts. It is also important to recognize that the stiffness of the instrument is included in the calibration of the sensor. Consequently, when a different instrument with a significantly different stiffness is used, different calibration parameters will have to be used. Finally, it was mentioned in Section 4.5 that silicone rubber is sensitive to changes in temperature. All testing was performed at uncontrolled room-temperature. Therefore, it is possible that changes in temperature introduced additional error between calibration and validation datasets. This work does also not investigate the true variance in temperature during surgery and the effect it would have on the sensor.

All of the limitations listed above will have collectively contributed to a reduced performance and larger error due to reduced repeatability. This is especially the case for the sterilization experiments where the sensors were removed and reattached for the sterilization cycles. Other limitations warrant further research and development into the system to make it more robust and usable. Nonetheless, considering the proof of concept nature of this work, the results show that the proposed sensing strategy and sensor design are able to measure the required forces with good performance and sterilizability.

7

Conclusion

This work aimed to design, prototype and validate a force sensing strategy for RAMIS using 3 mm instruments that is affordable and reusable. The sensing strategy is to be applied to limit forces during surgery or for feedback to the surgeon to prevent damage to instrument and patient. Based on an analysis of the problem and a review of the available literature it was concluded that designing and constructing a sensor that is both accurate and reusable under the given requirements is both challenging and novel. This is mainly due to the forces that occur at the trocar. These forces are difficult to predict or estimate and large enough to distort measurements and damage patient and instrument.

To achieve a sufficiently accurate system a solution was found that fixes the trocar to the same rigid frame as the instrument to eliminate the distorting forces at the trocar. The force sensor, located in the trocar, uses Hall sensors to measure changes in magnetic field density caused by bending of the instrument under load. A cast silicone layer is used as a transduction element that fully encloses the sensitive parts and holds together the sensor. This combination of parts is affordable, simple and reusable.

Prototypes and an experimental setup were constructed using rapid prototyping methods to validate the claims and compliance with the requirements for the proposed design. The performance results show that the sensor is able to measure the full range of 5 N in 2D with a bandwidth between 100 and 120 Hz. After calibration the sensor is able to measure with a MRE of less than 9% and a NRMSE of less than 3.5%. A sensing resolution of 0.021 N is achieved. By using a differential measurement configuration the sensor is able to reject a number of errors. It was also shown that, although with a large initial drop in performance, the sensor remains stable and reusable through repeated autoclave sterilization cycles.

Based on the validated prototypes it can be concluded that the proposed sensing strategy and sensor design provide a viable and reusable way to measure forces during RAMIS using 3 mm instruments to prevent damage to the instrument. Further steps can be taken to improve the performance and robustness of the sensor to forge a path towards active force limiting or force feedback to the surgeon. These steps are outlined in the final chapter.

8
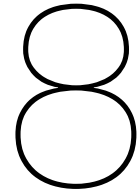
Recommendations

Considering the limitations mentioned in the discussion and ambitions for future applications there are a number of subjects suitable for further research. These topics and recommendations for their approach are presented here.

The first and most important topic for further research is to integrate the sensor proposed here in the robotic system to perform live measurements in a more realistic setting. To this end the design of the trocar and robotic system have to be altered to attach the trocar to the robot with a sufficiently stiff connection. How this could look on the trocar is shown by the 'arms' on the trocars in Figure 4.3. For the experiments described in this work it was acceptable to measure the insertion-depth using a ruler. When the sensor is integrated in the robot arm and used in a more dynamic way it is necessary to measure the insertion-depth automatically to adjust the calibration constant accordingly. This can be done with a separate position sensor or by drawing data from the position controlled robot arm. The second would suffice because different from conventional RAMIS the trocar is fixed to the robot in this case. When integrating the sensor into a usable system it is also relevant to further investigate the occurrence and influence of temperature variations on the sensor and include mitigation measures like a reference sensor when needed.

For the first topic it is theoretically possible to reuse the sensor design without any changes. A second topic however could be to improve the design of the sensor on a number of points. It is possible to improve the robustness of the sensor by constructing the two 3D-printed plastic rings from injection-moulded plastic or machined stainless steel. The electronics could be integrated and soldered on a single PCB. The PCB can be fixed to the outer ring on the sensor which remains static. Doing so will improve the signal quality with respect to using a breadboard and jumper cables. Additionally, if the analog Hall sensor voltages can be read-out using a 12 or 16 bit pin this would significantly improve resolution when the noise is reduced by the changes mentioned. Finally the sensor could be designed to account for the stiffening of the silicone and field-loss of the magnets after sterilization by intentionally choosing softer silicone and stronger magnets. This would probably lead to sensor saturation before sterilization meaning that the sensor is only usable after the first sterilization cycle.

This work focused on the use of 3mm instruments and the sensor was designed and optimized specifically for this size. However, it is theoretically possible to scale the sensor to be used for 5mm instruments. One important thing has to be considered when scaling the sensor. When design parameters are kept the same both the instrument and sensor will be stiffer for 5 mm instruments. Since the instrument cannot be made less stiff the sensor design has to be adjusted to reduce stiffness and increase sensitivity. Stiffness can be reduced by choosing Shore 00-10 silicone instead of Shore 00-20 and by increasing the radial thickness of the silicone layer. Both will also result in reduced axial stiffness and robustness of the sensor. When the stiffness and displacement changes the Hall sensor and magnet size also have to be re-evaluated to optimize the sensitivity. Here it is again important to maximize the magnetic field strength and change over the displacement range to maximize the SNR.

A fourth topic suitable for further research is designing a sensor to measure the axial forces during surgery. This sensor will be used in conjunction with the sensor proposed here to measure all the translational forces. With the third force component and discussed performance improvements it is possible to use the sensors for live force feedback instead of force limiting.

The final suggested topic for further research is to investigate the possibility to measure the trocar forces using the existing system. When the trocar forces are known it is possible to control the robot arm in a way that minimizes the forces exerted on the incision in the abdomen. In literature methods have been proposed that use robot actuator data to calculate the force that the robot is adding to the system. In a (quasi-)static situation the input force from the robot is equal to the sum of the tip forces measured by the sensor and the forces acting on the trocar. Consequently, when the input force and the tip force are known it should be possible to calculate the forces that are exerted on the trocar and incision.

A
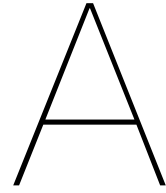
Finite Element Analysis of the magnetic system

A COMSOL Multiphysics study of the magnetic system was performed in order to determine the appropriate magnetic configuration, magnet size and geometrical parameters for this project. Here, the magnetic configuration refers to how the magnets are placed with respect to the sensor, whether pole pieces (pieces of high magnetic permeability material) are used to concentrate the magnetic field and whether differential measurements are performed. This Appendix will describe the model and settings used for the simulations of the magnetic field of the sensor. It will also explain the simulations and present the results used to make the design choices that are discussed in 4.6.

Model implementation

The stationary study of the magnetic fields is performed in a 2D approximation using the *Magnetic Fields, No Current* physics in COMSOL Multiphysics. The magnetic fields in the study are generated through the magnetization of the magnetic material using the Magnetization model *Remanant flux density* under the *Magnetic Flux Conservation* domain. The magnetic configurations are placed in a larger circle that represents the medium that the magnetic field travels through. A *Magnetic Insulation* boundary is defined at the edge of this circle. The size of this circle is large enough to not significantly constrain or influence the magnetic field. A *Magnetic Flux Conservation* domain with the Magnetization model *Relative permeability* is applied to the medium and pole pieces when they are included.

The goal of this study is not to investigate the performance of the Hall sensors themselves. Therefore it is not necessary to model actual Hall sensors for this study. Instead, the Hall sensors are represented by point evaluations of the magnetic field density in the locations the Hall sensors would be. The point evaluations are shown in Figure A.2b with the black dots.

The different configurations include up to three different materials. The magnets in this study are modelled as N45 Neodymium which is a predefined material in the COMSOL material library. This material choice is based on the fact that the smallest and strongest magnets commercially available are made from Neodymium. N45 neodymium has a remanant flux density of $13,500$ Gauss. The pole pieces included in some of the configurations are represented by a custom material with a relative permeability μ_r of 5000. In this study the medium is defined as air. In reality the medium will be a combination of the compliant material that surrounds the Hall sensors and magnets and the air around the sensor. Since the compliant material has a higher relative permeability than air it is likely that most of the magnetic field will remain in the sensor resulting in a higher field density than modelled here. The instrument that would be located in the center is not included in this model. Medical instruments are generally made from austenitic 3-series stainless steels (304 or 316). These materials are non-magnetic which means they have a relative permeability close or equal to that of air [11]. The physical parameters relevant for this study are listed in Table A.1.

Table A.1: Physical parameters used in the COMSOL analysis of the magnetic system

Parameter	Value	Description
μ_0	$1.257 \times 10^{-6}\,\mathrm{H\,m^{-1}}$	Vacuum permeability
μ_{ra}	1	Relative permeability of air
μ_{ri}	5000	Relative permeability of iron
μ_{rec}	1.05	Recoil permeability of Neodymium
B_r	1.35 T	Remanent flux density of N45 Neodymium
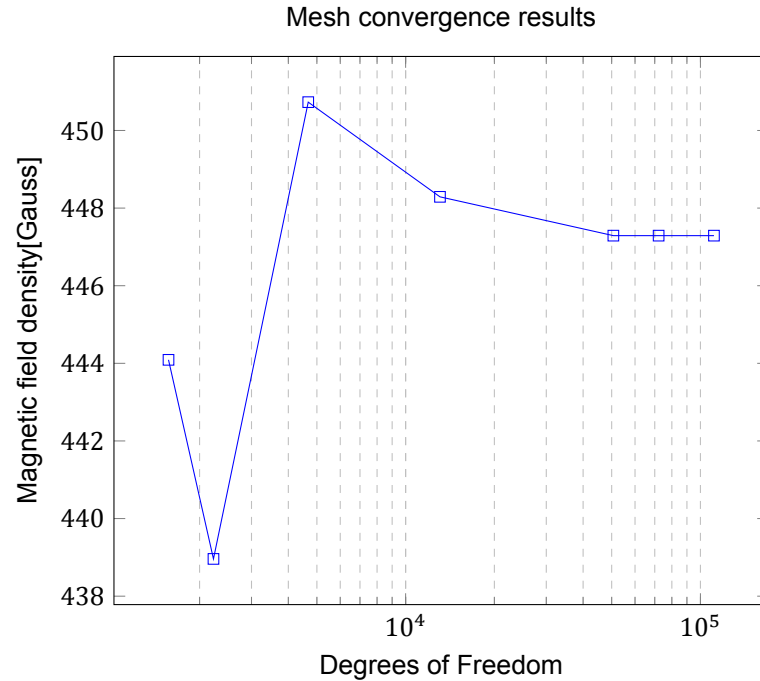
Figure A.1: Results from the mesh convergence study of the magnetic model performed on the simple configuration in centered position.

To determine the appropriate mesh size a mesh convergence study was performed on the simplest configuration. The configuration uses ⌀2x1 mm magnets at 2 mm from the sensor and 3 mm from the center. To perform the mesh convergence study the configuration was kept in the undisturbed position. Figure A.1 shows the results obtained by decreasing the mesh size step by step and evaluating the magnetic field density at the sensor. The figure clearly shows that after 50,000 Degrees of Freedom the measured field density does not change significantly. Therefore this mesh setting will be used for all subsequent simulations of the magnetic system.

Differential vs. non-differential measurements

The first thing investigated using the model described in the previous section is whether to perform differential measurements instead of standard non-differential measurements. As can be seen in Figure A.2 the differential configuration uses two sets of opposing sensor-magnet pairs. The differential measurement is obtained by subtracting the output from the sensor-magnet pair at the top from the one at the bottom and similarly the output of the left from that of the right. This means that in the undisturbed situation the output of the sensor is zero. When in this situation the magnets are moved up the field increases at the top sensor and decreases at the bottom sensor both contributing to a change in output. In linear situations this would lead to a doubling of the output and sensitivity. Other advantages of differential measurements include temperature compensation and linearization of the output.

To compare the two configuration the same parameters are used as before: ⌀2x1 mm magnets at 2 mm from the sensor and 3 mm from the center. The magnetic field density measured with the different
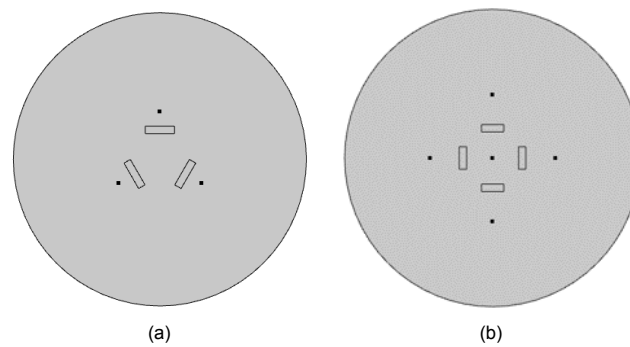
Figure A.2: Magnetic field density results for the different magnetic configurations. With (a) simple configuration, (b) Outer Pole piece, (c) Inner Pole piece (d) Inner ring Pole piece and (e) differential configuration.
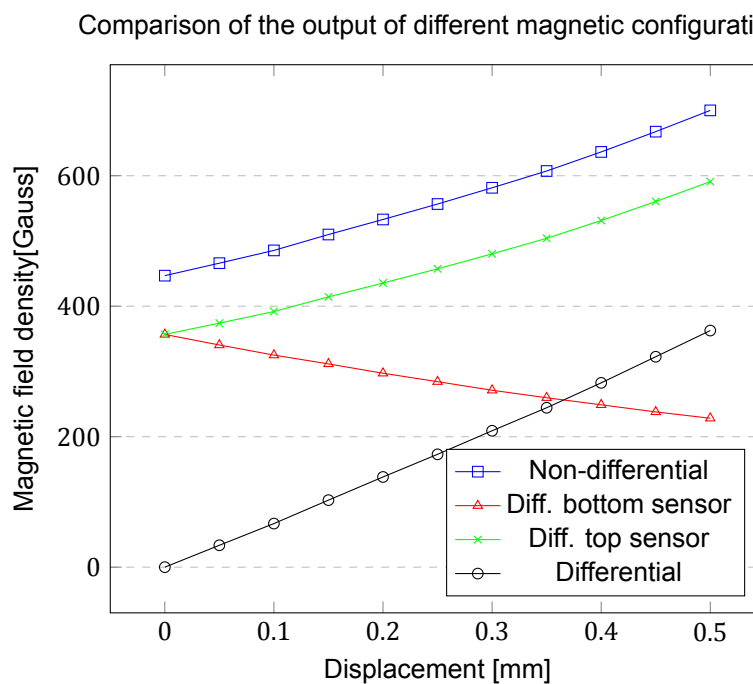
Figure A.3: Comparison of the output of a differential and non-differential magnetic configuration over 0.5mm displacement.

configurations over a displacement of 0.5 mm upwards is shown in Figure A.3. The results in the figure show that the differential configuration produces a significantly larger output. The change in field density however is not twice as large. The change in field density is 253 Gauss for the non-differential configuration and 363 Gauss for the differential configuration. The fact that the difference is smaller is caused by two factors. First, the magnetic field density from a source does not change linearly with distance. Therefore the change in field is different at the top sensor compared to the bottom sensor which can be seen in Figure A.3 as well. The second cause is the fact that adding a fourth magnet to the system at close proximity to the other three leads to field attenuation. The magnets are cancelling out some of each others fields. This can be seen in Figure A.3 by observing that the field density at zero displacement is lower for the sensors in the differential configuration compared to the non-differential configuration. However, since the output and sensitivity are still significantly increased it is decided to use differential measurements in the final design.

Pole pieces

Pole pieces are pieces of material, generally iron, that have a much higher magnetic permeability. Because field lines take the path of least resistance the high permeability pole pieces attract and concentrate field lines. When the Hall sensor and pole pieces are placed in the right location this should

result in a higher field density to be measured by the Hall sensor.

There are many different ways in which pole pieces can be added to the system. For this work only three different configurations with pole pieces were considered. The different configurations are shown in Figure A.4. The pole pieces in Figures A.4b and A.4c have the same size as the magnets (⌀2x1mm) and are shown in red. In Figure A.4b the pole pieces are located at the outer perimeter behind the Hall sensors giving it the name 'outer pole piece configuration'. The configuration in Figure A.4c has the pole pieces placed directly behind the magnets at the inner perimeter and is named 'stacked pole piece configuration'. The final configuration in Figure A.4d uses a ring shaped pole piece to concentrate the field.
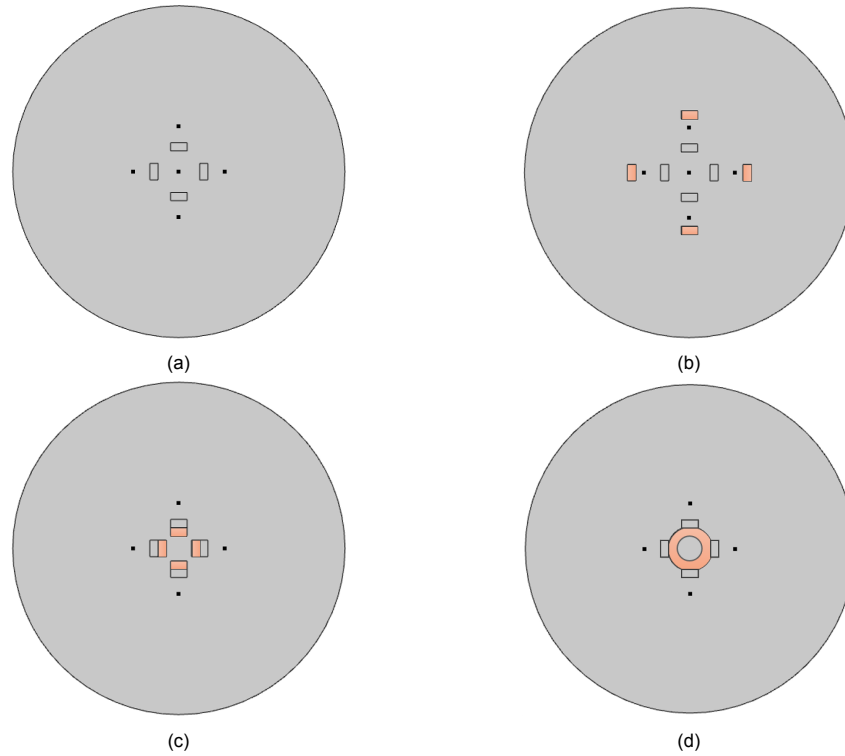
Figure A.4: Geometries of the different differential configurations investigated in COMSOL. With (a) differential without pole pieces, (b) differential with pole pieces at the outer perimeter, (c) differential with magnet and pole piece stacked, and (d) differential with a ring shaped pole piece. Pole pieces are shown in red.

To compare the different configurations their magnetic field was simulated over a displacement of 0.5 mm. The resulting differential outputs are plotted in Figure A.5. The figure legend includes the field density measured at the sensor in the undisturbed position for each of the configurations. The results in the Figure show that the outer pole piece and stacked pole piece configurations result in only a slight increase in initial field and differential output. The ring pole piece configuration has a bigger effect. It's initial field and differential output are significantly larger than the standard configuration. Looking at Figure A.5 the output curve for the ring pole piece configuration is noticeably more non-linear compared to the other three.

In Section 4.6 it was explained that producability and performance are important factors in the design of the sensor which are often at odds. This is also the case in the choice of whether to use pole pieces or not. Figure A.5 has shown that using the ring pole piece configuration will result in a higher sensitivity and SNR. In terms of production however it is difficult to include the iron ring that would be needed. Moreover, adding 'pure' iron to the design could also come with biocompatibility and sterilization problems. Therefore it was decided to not use pole pieces in the prototype that will be constructed for this work. For a final product which is produced in larger numbers it could be more feasible to include the ring pole piece if deemed necessary.

Comparison of the output of different magnetic configurations
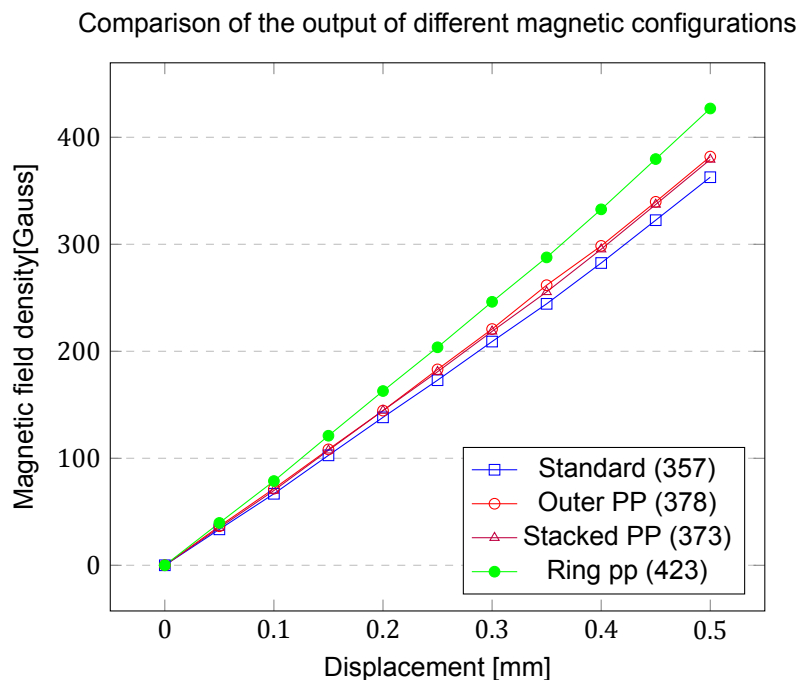
Figure A.5: Comparison of the differential output of different differential configurations with pole pieces (PP) over 0.5mm displacement. The initial magnetic field density in Gauss at the sensor is included in the legend

Reducing field attenuation

In the section comparing a differential to a non-differential configuration it was mentioned that adding a fourth magnet results in field attenuation. This section explores some ways in which the field attenuation could be reduced. The first solution is the most straightforward and involves increasing the distance between the magnets by placing them at a larger radius from the center. The second solution is to reverse the magnetic field of one pair of magnets to face outward instead of to the center.

To investigate the effectiveness of the first solution the standard differential configuration with ⌀2x1 mm magnets and a distance to the sensor of 2 mm is used. The magnetic field density at the Hall sensor is determined for configurations where the magnets are located between 2 and 5 mm from the center while keeping the distance to the sensor at 2 mm. The results are presented in Figure A.6. It shows that increasing the distance between the magnets by locating them further from the center significantly improves the magnetic field density at the sensor. There are however diminishing returns with increasing distance.

A simulation was ran with the magnetic fields of the top and bottom magnets pointing outwards instead of inwards to investigate the effectiveness of field reversal. Figure A.7 allows for a comparison of the magnetic field between the standard and reversed situation. Figure A.7a plots the magnetic field lines when four magnets are located close together and magnetized in the same direction. It shows that the four magnets create eight field loops. The magnetic field when the two magnets are flipped is shown in Figure A.7b. In this situation only four field loops are created. Consequently, the magnetic field density perpendicular to the magnet face is a lot higher. In the normal situation the field density is 357 Gauss at the sensor and with the reversed field it is 727 Gauss, which is about twice as high. Because of these improvements it was decided to also flip one magnet pair in the final prototype.

Geometry, magnet size and sensor

The final simulations were used to determine the optimal combination of geometry, magnet size and sensor. The geometrical parameters are selected mainly on the basis of producability rather than per-
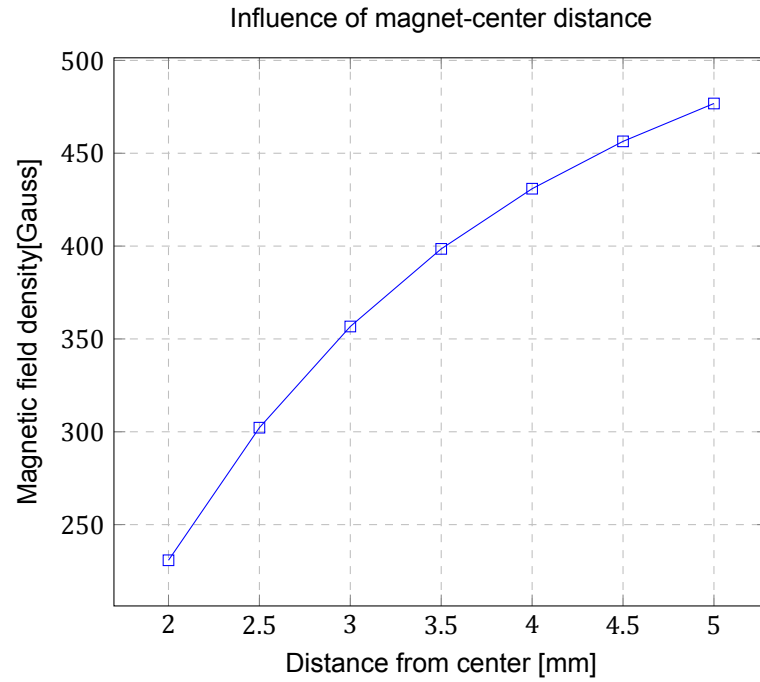
Figure A.6: Magnetic field density at the sensor with magnets at different distances from the center.

formance as explained in Section 4.6. In order to be able to assemble the sensor and place the magnets accurately it was decided to locate the magnets at 4 mm from the center. Also the sensor-magnet distance was set to a minimum of 3 mm to improve producability and allow for a sufficient displacement.

With a relatively large sensor-magnet distance only the larger magnets are worth considering. Therefore only the following magnets are included in the final simulations: ⌀2x1, ⌀3x1 and ⌀4x1 mm magnets. The simulations were used to find the combination of magnet, sensor and distance that is able to use as much of the sensing range of a sensor. Doing so will yield the highest sensitivity and SNR. As explained in Section 4.6 the available sensing ranges include $\pm 420, 670$ and 840 Gauss. Table A.2 includes the ranges of magnetic field density over a 0.5 mm displacement for different magnets and distances. By looking at the results in the table it was decided to use the ⌀3x1 at a distance of 3 mm. This combination fits closest to the sensing range of one of the available sensors and also has a high change of field. In fact, the combination will slightly exceed the stated range of the sensor but considering that in a non-ideal situation the field density is probably lower this should not be a problem. This combination was chosen over the ⌀4x1 at a distance of 3.5 mm to keep the sensor more compact.

Table A.2: Results from simulations varying magnet size and sensor-magnet distance

Comparison of magnet size and distance						
	3 mm		3.5 mm		4 mm	
	min	max	min	max	min	max
⌀2x1 mm	387	508	304	387	243	303
⌀3x1 mm	539	686	430	538	349	429
⌀4x1 mm	657	814	534	655	441	533
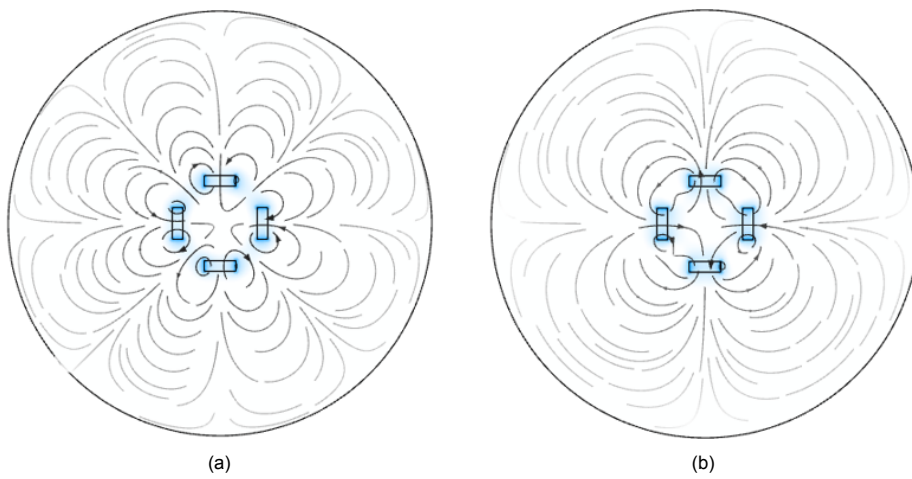
Figure A.7: Plot of the magnetic fields when (a) all the magnets are magnetized in the same direction (north pole pointing inwards) and (b) one pair of magnets are flipped (north pole pointing outwards).

B
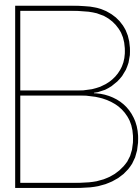
Thermal losses in Neodymium magnets

An important goal of this project is to design a reusable sensor. To achieve this the sensor has to withstand a cycle in a steam sterilizer. As described in Section 2.4, during sterilization the sensor will be subjected to steam at $134°$C in a vacuum at 100% humidity for a few minutes [13]. Because the sensors are copper-nickle plated it is assumed that the humidity during sterilization has no effect. For the temperature however, it is not yet clear how they will be affected by this. Although the Curie temperature of Neodymium magnets is $315°$C [49] most Neodymium magnets have a stated maximum operation temperature of $80°$C. This appendix will describe the theory behind thermal behaviour of neodymium magnets and a simple experiment performed to investigate the field strength after several thermal cycles.

Theory

The changes in magnetic performance that magnets experience at higher temperatures can be split into two categories: reversible and irreversible losses. Reversible losses are completely recovered upon return to the initial temperature, irreversible losses are not. Because of the low curie temperature of neodymium magnets the reversible losses increase quite quickly with temperature [49, 50]. Since the temperature of the magnets will be relatively constant during surgery the reversible losses are not of concern. Irreversible losses however could pose a problem because the magnets have to go through several sterilization cycles.

Irreversible losses are the result of reversals of domain magnetization direction caused by increased thermal energy at higher temperatures. Since the domain is more stable in the reversed direction it does not flip back upon decreasing the temperature. A higher temperature and higher thermal energy will cause more domains to reverse their magnetization direction and a longer time will increase the chance that the domains actually flip. Irreversible losses can only be recovered by remagnetization at room temperature [49, 50]. When a magnet is reheated to the same temperature as the first thermal cycle the additional reversible losses will be negligible. This is due to the fact that the most unstable domains will have already reversed directions. To cause more reversals either the thermal energy has to be increased with a higher temperature or the energy level has to be maintained for a longer time [49, 50]. Consequently, it is possible to stabilize the magnetic performance over thermal cycles by preheating the magnets to the working temperature for some longer amount of time [50].

Method

A simple experiment was performed to verify the behaviour predicted by the theory outlined in the previous section. The experiment used ten 5mm cube magnets of the N45 grade. The magnetic field density was measured using the SS4495A Hall sensor by Honeywell. The Hall sensor was glued to a digital caliper. The caliper was made from plastic to have minimal influence on the magnetic field. The magnets were placed in a holder in order to ensure the same alignment and distance for each measurement. The magnets were placed at 4mm from the sensor. This distance was chosen because it resulted in a field density in the middle of the sensor range. For each data point 500 samples were

collected and averaged.

First the repeatability of the experiment was tested by placing and removing the same magnet ten times. This eliminates the difference in magnetization between the different magnets. Next, a baseline was established by measuring the field density of the ten magnets before undergoing any thermal cycle. For the thermal cycles the magnets were placed in a standard kitchen oven at around 135 degrees Celsius for 5 minutes. A separate oven thermometer was used to ensure there was no significant deviation from the intended temperature. Afterwards the magnets were allowed to cool down back to room temperature for at least 10 minutes. This step is important to not include any reversible losses of magnetic field density in the measurements. After cooling down the magnets were measured in the same manner explained before.

Results

The results obtained using the procedure described in the previous section are presented in Figure B.1. The data is presented using boxplots to clearly show the relative difference and spread in the datasets.

The first boxplot of Figure B.1 gives the result from the repeatability study. The flat box and relatively short whiskers indicate that the experiment has a good repeatability.
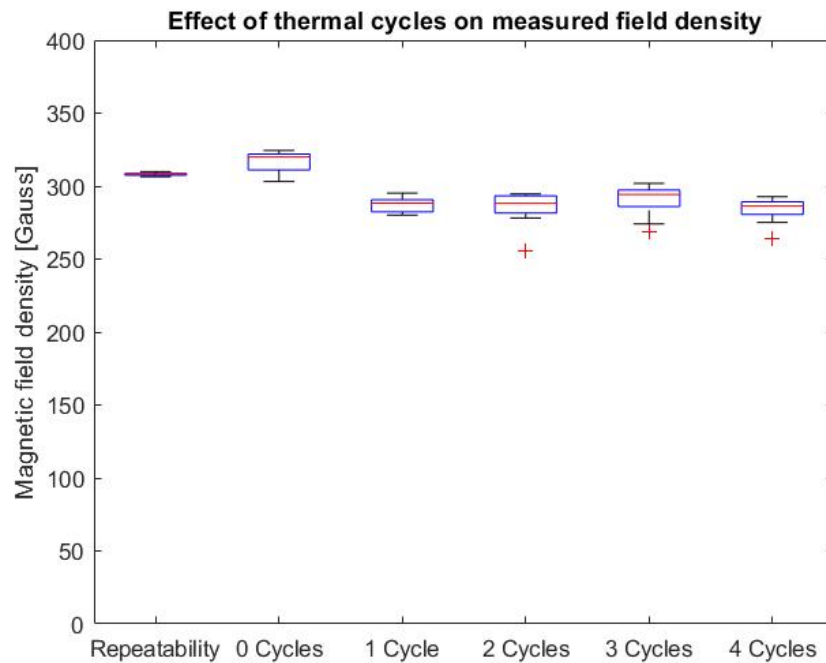
Figure B.1: Effect of thermal cycles on magnetic field density

The second boxplot in Figure B.1 presents the magnetic field density of the ten magnets before any thermal cycle. The plot shows that there is some spread in magnetic field density between the different magnets, with the minimum and maximum separated by 20 Gauss. This spread is important to consider since the sensor uses differential measurements.

The final four boxplots show the results after each thermal cycle. After the first thermal cycle there is a clear drop of almost 10% in average magnetic field density. After the second thermal cycle the average field density is almost exactly the same as after the first cycle. After the third cycle the average field density is even higher than the two before followed by a lower field density after the fourth cycle. Notably, the last three datasets show a slightly larger spread than expected based on the repeatability data and variation before thermal cycle. The spread of the data after the third cycle is also bigger than

those before and after. Finally, it must be noted that the last three datasets include a significant outlier. Presumably this is the same magnet but it is not possible to determine whether this is actually the case. The cause of the outliers and the increased spread is also not clear.

Discussion

Before the final conclusions on this experiment a few aspects of the study have to be adressed. Firstly, only a limited number of magnets were used in this experiment. Moreover, although attempted, it was not possible to number the individual magnets and track the effect of the thermal cycles on each magnet. Secondly, it proved difficult to maintain a constant temperature around 135 degrees Celsius with a kitchen oven. One reason for this is that opening the oven door to place the magnets resulted in an unknown temperature drop. Another reason is the slow response time of the oven thermometer which made it difficult to finely control the temperature. As a result the thermal cycle will have been slightly different each time. It is also likely that the temperature was not exactly the same throughout the oven. This could be the cause of the larger spread in measured field densities after the first thermal cycle. Finally, the room in which the measurements were taken was not finely temperature controlled. Therefore, the ambient temperature was probably not equal for each dataset because of the considerable time in between measurements.

Conclusion

A number of conclusions can be drawn from the results of this experiment presented in Figure B.1. The first and most important conclusion is that the Neodymium magnets do seem to follow the behaviour predicted in the literature. After the first thermal cycle a marked drop of almost 10% in magnetic field density is observed. This drop in field density however does not increase significantly after subsequent thermal cycles. Therefore, in final sensor production, it could be useful to stabilize the thermal behaviour of the magnets by preheating them to a little above the sterilization temperature. This would ensure no significant change in performance would occur after later sterilization cycles.

A second relevant conclusion is the fact that there is a significant spread in initial magnetic field density between the different magnets. The initial spread can only partly be explained by the repeatability of the measurements. Moreover, it appears that the spread increases after the first thermal cycles. The increase in spread suggests that not all magnets undergo the same change of field density during the thermal cycle. This could be partly caused by an unequal temperature distribution during the thermal cycle. The difference in field density could lead to a non-zero output of the differential measurement at zero displacement which will have to be accounted for.

C
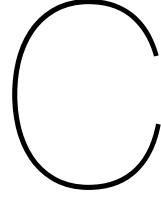
Finite Element Analysis of Rubber

The stiffness of the sensor has a large influence on the measured displacement and final resolution. However, because of the geometry and use of rubber it is not feasible to calculate the stiffness analytically. Therefore, this appendix describes a COMSOL Multiphysics study used to find the force-displacement behaviour of the sensor. This information will be used to determine the displacement that is to be expected in the prototype. Before describing the model and results the modelling of rubber materials is discussed.

Modeling hyperelastic materials

Materials that are able to undergo large deformations without internal losses, due to plastic deformation for example, are classified as hyperelastic materials. Materials in this category include elastomers, polymers, rubber and biological tissues [51]. These materials do not fit the conventional linear-elastic material model based on Hooke's law. Instead, these materials are modeled using the Helmholtz free-energy function Ψ. The free-energy function depends on thermal energy and mechanical strain energy. The thermal part is generally left out in mechanical analyses. Consequently, the free-energy function is only a function of the deformation gradient ($\mathbf{F}$) and indicated with W for strain energy [51, 52].

A material model often used for isotropic rubber-like materials is called the Ogden model [52, 53]. The Ogden model expresses the strain energy function in terms of the principal stretches. The principal stretches are defined as follows:

$$\lambda_i = (L_F/L_0)_i$$

With L_F the length in the deformed state and L_0 the length in the undeformed state in one of the three principal directions. Using the principal stretches the Ogden material model can be written as:

$$W = \sum_{i=1}^{N} \frac{\mu_i}{\alpha_i}(\lambda_1^{\alpha_i} + \lambda_2^{\alpha_i} + \lambda_3^{\alpha_i} - 3) + \sum_{j=1}^{M} \frac{1}{D_j}(J - 1)^{2j}$$

Here $J = det(\mathbf{F})$ and N, μ_i, α_i, M and D_i are empirically found material constants used to describe the non-linear behaviour [51, 52, 53]. The two parts of the model describe the isochoric and volumetric elastic response respectively. When the material is assumed to be incompressible the determinant J equals 1 (by definition) resulting in zero contribution from the volumetric part of the model [51, 52]. For this study a compressible Ogden material model was used. Steck et al. showed that for soft silicone rubber the compressible Ogden material model is better at predicting the stress-strain relation compared to other models [53]. In their work they also present the material parameters needed for FE modelling of Shore 00-30 Ecoflex silicone rubber. Since the sensor will be made out of a similarly soft Ecoflex rubber this is considered a reasonable proxy since no Ogden parameters were found for Shore 00-20 rubber.

73

Model implementation

The mechanical analysis is performed in COMSOL Multiphysics as a non-linear stationary study in 3D using the *Solid mechanics* physics. It is necessary to perform the study with a 3D model because of the large out-of-plane deformation of the compliant layer when it is loaded. The inner and outer ring are included in this model because they also have a limited stiffness. The rings are modelled using the *Linear Elastic Material* domain. The compliant material is modelled using the *Hyperelastic Material* domain. As already introduced in the previous Section the material model used is a Ogden model for compressible materials. Next, there is a *Fixed Constraint* applied to the outer face of the solid outer ring. A *Boundary Load* is applied to the inner face of the inner ring. Finally, all initial displacements and velocities are zero. The main interest of this analysis is the displacement obtained when a specific load is applied. To this end the displacement data is obtained with a point evaluation located at the inner edge of the inner ring in the direction the load is applied.

Three different materials are used for this analysis. The first is ABS, a common structural plastic also used in medical devices, from the COMSOL material library. This material is applied to both the inner and outer ring. Next there is the Shore 00-30 Ecoflex silicone rubber included in the analysis to represent the compliant material. The Ogden model parameters are provided by Table III in Steck et al. [53] and other parameters are based on the Ecoflex datasheet. The bulk modulus used for the nearly incompressible model is predefined in COMSOL. The material parameters used by the COMSOL model are listed in Table C.1.

Table C.1: Physical parameters used in the mechanical analysis of the sensor.

Parameter	Value	Description
ρ_s	$1050 \, \text{kg m}^{-3}$	Density of ABS
E_s	1.8 GPa	Youngs modulus of ABS
ν_s	0.36	Poissons ratio of ABS
ρ_r	$1070 \, \text{kg m}^{-3}$	Density of silicone rubber
μ_1	16.9 kPa	Ogden shear parameter for Shore 00-30 Silicone
μ_2	0.08 kPa	Ogden shear parameter for Shore 00-30 Silicone
μ_3	1.0 kPa	Ogden shear parameter for Shore 00-30 Silicone
α_1	1.3	Ogden alpha parameter for Shore 00-30 Silicone
α_2	5.0	Ogden alpha parameter for Shore 00-30 Silicone
α_3	-2.0	Ogden alpha parameter for Shore 00-30 Silicone

The model used in this analysis consists of a simple 3D geometry. Figure C.1 provides an image of the sensor as it is modelled in COMSOL. The model consists of a solid inner ring that interfaces with the instrument and the compliant layer. On the outside another solid ring is placed. The geometry also includes simplified substitutes for the magnets and sensors. These were included because they result in a varying thickness of the silicone layer which has a large influence on the stiffness The instrument is not included in this model. Table C.2 lists the geometric parameters used to construct the model.

Table C.2: Geometric parameters used in the mechanical analysis of the proposed sensor

Parameter	Value	Description
D_I	3 [mm]	Diameter of the instrument
D_{ir}	9.4 [mm]	Outer diameter of the inner ring
t	3 [mm]	Sensor-magnet distance
D_{oc}	18 [mm]	Outer diameter of the compliant layer
D_o	25 [mm]	Outer diameter of the model
H	5 [mm]	Height of the sensor
Dm	4 [mm]	Magnet diameter
Lm	7 [mm]	Min distance between magnets
Ws	2.5 [mm]	Sensor width
ts	1.57 [mm]	Sensor thickness
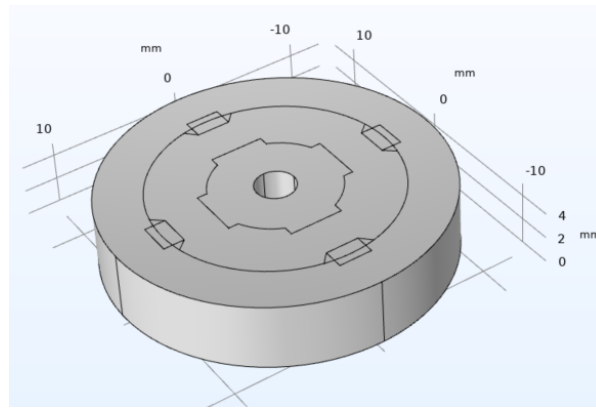
Figure C.1: Geometry used for the mechanical analysis of the sensor

A mesh convergence study was performed to determine the appropriate mesh size. To perform the mesh convergence study the model was put under a 1 N load and the displacement of the inner ring is evaluated. Figure A.1 shows the results obtained by decreasing the mesh size step by step and evaluating the displacement. The figure shows that after $200,000$ Degrees of Freedom the measured displacement does not change significantly. Therefore this mesh setting will be used for all subsequent simulations.
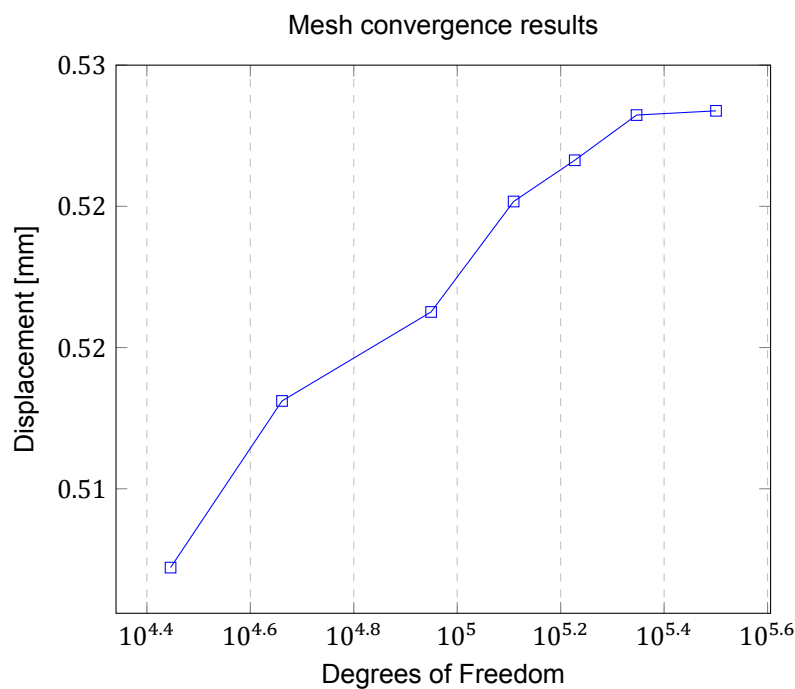
Figure C.2: Results from the mesh convergence study of the mechanical model.

Results

To determine the stiffness of the sensor the load was increased from 0 to 1.5 N in steps of 0.15N using the parametric sweep in COMSOL. The sweep was performed with the compressible Ogden model. The resulting force-displacement results are plotted together with a linear regression of the data in Figure C.3

The large out of plane displacements as well as the non-linear material model suggest a strong
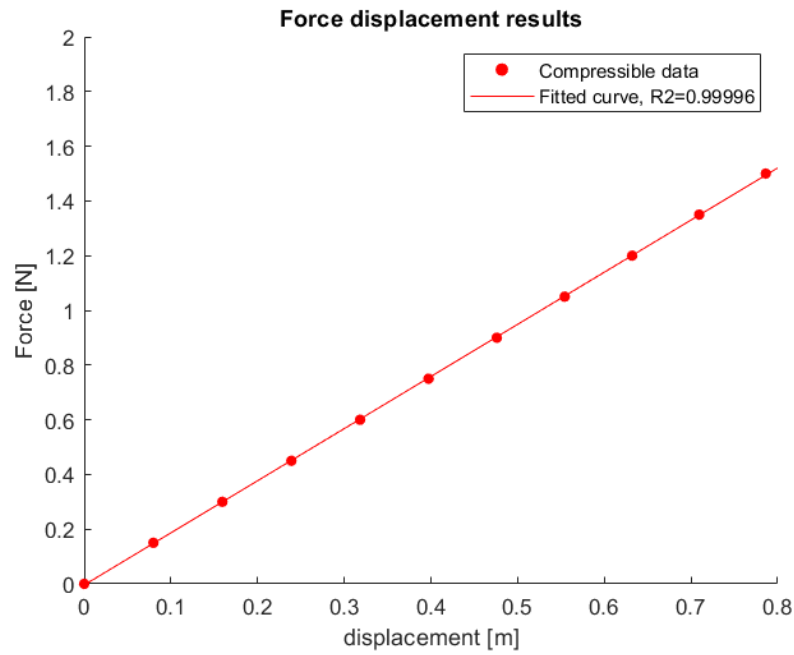
Figure C.3: Force-displacement results from the COMSOL analysis with linear regression for compressible material model

non-linear force-displacement relation. However, the data from the force sweep fit a linear regression with R-squared values of 0.99996 indicating an almost perfectly linear response. The stiffness found in this linear fit equals 1.9 N/mm. A possible explanation for this highly linear response is the fact that the material is shaped in a cylinder. Consequently when one side is compressed the other side is stretched which results in an effect similar to differential measurements.

D
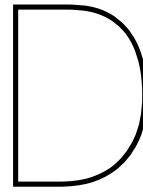
Arduino and Matlab code

This appendix presents the code used in Arduino and Matlab for data acquisition, live plotting and data processing.

D.1. Arduino code for data acquisition

The code listed below is used to acquire samples from the analog pins when a signal is sent from the PC (Matlab) to do so.

```
1  char Incoming_command;
2  int dat;
3  void setup() {
4    // start serial connection
5    Serial.begin(115200);
6    Serial.setTimeout(3);
7  }
8
9  void loop() {
10   // Check for input from Matlab
11     if (Serial.available()>0){
12       Incoming_command = Serial.read();
13
14   //If input is correct send measurements from the four sensors
15       if (Incoming_command == 'm'){
16         Serial.print(analogRead(2));
17         Serial.print("\t");
18         Serial.print(analogRead(3));
19         Serial.print("\t");
20         Serial.print(analogRead(4));
21         Serial.print("\t");
22         Serial.println(analogRead(1));
23       }
24     }
25   }
```

D.2. Matlab code for data acquisition

The code listed below is used to tell the Arduino to send data and plot and store the uncalibrated data that Matlab receives.

77

```matlab
1  arduino = serial('COM5','BaudRate',115200);
2  fopen(arduino);
3
4  %% Correction
5  % determine offset correction factors
6
7  Ncor = 500;        %Number of samples for offset correction for live plotting
8
9  dc = zeros(Ncor,4);
10
11 for ii=1:Ncor                              %Acquire Ncor samples
12     fprintf(arduino,"m");
13     dc(ii,1:4)= fscanf(arduino,'%d');
14 end
15
16 dc_mean = mean(dc,1);
17 Corr = [dc_mean(1)-dc_mean(3),dc_mean(2)-dc_mean(4)];  %Find deviation
       from zero in differential measurements
18
19 %% Live plotting with moving average and correction (diff)
20
21 MA = 5 ;                     %Number of points averaged
22 jj = MA;
23 kk = 1;
24 Tmax = 25;                   %Sampling time
25 data = zeros(Tmax*120,5);    %Oversized data matrix
26 tic
27 Tstart = toc;
28
29 mafdata = zeros(Tmax*150,4);       %Empty matrix for filtered data for
       live plot
30
31 p1 = animatedline('Color', 'r');   %Pre-define animated line
32 p2 = animatedline('Color', 'b');
33
34 while (toc< Tstart+Tmax)
35    jj = jj+1;
36    fprintf(arduino,"m");
37    data(jj,1:4)= fscanf(arduino,'%d');
38    data(jj,5) = toc;
39    mafdata(jj,1:4) = mean(data(jj-MA:jj,1:4),1);   %% apply moving average
40    if (jj == kk*12)
41      addpoints(p1,jj,mafdata(jj,1)-mafdata(jj,3)-Corr(1))
42      addpoints(p2,jj,mafdata(jj,2)-mafdata(jj,4)-Corr(2))
43      drawnow;
44      kk = kk+1;
45    end
46 end
47 data = data(1:jj,:);
48 BW = jj/data(end,5);
49
50 fclose(arduino);
51 clear arduino
52
53
54 %% Store data
```

```matlab
55
56  writematrix(data, '220805_S3_100_4mm_s100.csv');
```

D.3. Matlab code for data processing and calibration

The code listed below is used to process the data from prototype and reference sensor. Afterwards, the code finds the calibration constant, error and noise for the dataset. Finally, the found calibration constants are used to fit the calibration curves for the X- and Y-axis. A very similar code is used to evaluate the error on the validation datasets.

```matlab
1   %% Calibration
2   % Calibration of designed sensor using reference sensor output
3   % Reference sensor: Futek LSB200 2Lbs
4   clear all
5   close all
6
7   %Parameters used on motion stage
8   d = 4.5;                    %Distance traveled [mm]     ADJUST!!!
9   s = 0.5;                    %Speed [mm/s]
10  m = d/s;                    %Movement time [s]
11  p = 5;                      %Pause time [s]
12  tot = 4*p+2*m;              %total measurement time [s]
13
14  %Read sensor data
15  PS = readmatrix('220713_S1_45mm.csv');      %Read data from sensor NOTE:
        data not filtered or corrected
16  PS_diff1 = PS(:,1)-PS(:,3);                 %Construct diff. output 1
17  PS_diff2 = PS(:,2)-PS(:,4);                 %Construct diff. output 2
18  S1_t = PS(:,5);
19
20  %Read reference sensor data
21  RS = load('S1_45mm_20180701-111149.mat');   %Load data strucutre from
        reference sensor
22  RS_V_f = RS.capture.sensor_V;               %Sensor voltage data NOTE: data
         not filtered
23  RS_t = RS.capture.time_s;                   %Time vector
24
25
26  %% Plot to find start time
27  %Find short peak at beginning of measurement and note the time
28  figure
29  subplot(2,1,1)
30  plot(S1_t,PS_diff1); hold on;
31  plot(S1_t,PS_diff2);
32  ylabel('Sensor output','Fontsize',14)
33  title('Uncorrected output from prototype sensor')
34  subplot(2,1,2)
35  plot(RS_t,RS_V_f);
36  title('Uncorrected output from reference sensor')
37  ylabel('Sensor output [V]','Fontsize',14)
38  xlabel('Time [s]')
39
40
41  %% Note the time of the short peak
42  peak_PS = 4.244;            %Peak in prototype sensor data
43  peak_RS = 1.87;             %Peak in reference sensor data
```

```matlab
44
45 % Correct time vectors and filter data
46 [~,tcorr_RS] = min(abs((peak_RS+0.5)-RS_t));        %Find index of
       registred time with margin
47 [~,tcorr_PS] = min(abs((peak_PS+0.5)-S1_t));
48
49 RS_V_f = RS_V_f(tcorr_RS:end);                %Discard data before index
50 RS_t = RS_t(tcorr_RS:end)-(peak_RS+0.5);
51
52
53 PS_diff1 = PS_diff1(tcorr_PS+1:end);
54 PS_diff2 = PS_diff2(tcorr_PS+1:end);
55 S1_t = S1_t(tcorr_PS+1:end)-(peak_PS+0.5);
56
57 %trim end to achieve same length of time
58 if S1_t(end)>= RS_t(end)
59     [~,endcorr_PS] = min(abs(RS_t(end)-S1_t));
60     S1_t = S1_t(1:endcorr_PS);
61     PS_diff1 = PS_diff1(1:endcorr_PS);
62     PS_diff2 = PS_diff2(1:endcorr_PS);
63 else
64     [~,endcorr_RS] = min(abs(S1_t(end)-RS_t));
65     RS_t = RS_t(1:endcorr_RS);
66     RS_V_f = RS_V_f(1:endcorr_RS);
67 end
68
69 % Filter and correct prototype sensor data
70 fw = 20;      %number of points in moving average filter
71
72 S1_diff1_f = movmean(PS_diff1,fw);        %Central moving average filter
73 S1_diff2_f = movmean(PS_diff2,fw);
74
75 S1_offset1 = mean(S1_diff1_f(1:100));
76 S1_diff1_f = S1_diff1_f-S1_offset1;       %Correct offset
77
78 S1_offset2 = mean(S1_diff2_f(1:100));
79 S1_diff2_f = S1_diff2_f-S1_offset2;
80
81 %Filter, offset correct and Calibrate reference sensor data
82 CP = 1.0179;                              %Futek calibration parameter
83 RS_V_filt = movmean(RS_V_f,fw);
84 RS_offset = mean(RS_V_filt(1:300));
85 RS_F = (RS_V_filt-RS_offset).*CP;
86
87 %times of inflection points
88 IP1 = p-peak_RS-0.4;
89 IP2 = IP1+m;
90 IP3 = IP2+2*p;
91 IP4 = IP3+m;
92
93 % Calibrate prototype sensor to match reference sensor
94 mar = m/8;  %Margin on inflection points (to exclude accelerations)
95
96 % Find slopes of reference sensor (using linear regression)
97 [~,ssd] = min(abs((IP1+mar)-RS_t));       % Start slope down
98 [~,esd] = min(abs((IP2-mar)-RS_t));       % End slope down
```

```matlab
 99
100  [~,ssu] = min(abs((IP3+mar)-RS_t));       % Start slope up
101  [~,esu] = min(abs((IP4-mar)-RS_t));       % End slope up
102
103  RS_downslope = RS_F(ssd:esd);
104  RS_upslope = RS_F(ssu:esu);
105
106  [FIT_RS_down,GOF_RS_down] = fit(RS_t(ssd:esd)',RS_downslope','poly1'); %
         Linear fit of slope down
107  [FIT_RS_up,GOF_RS_up] = fit(RS_t(ssu:esu)',RS_upslope','poly1');        %
         Linear fit of slope up
108
109  % Find slopes of prototype sensor (using linear regression)
110  [~,ssd] = min(abs((IP1+mar)-S1_t));       % Find index of start slope down
111  [~,esd] = min(abs((IP2-mar)-S1_t));       % Find index of end slope down
112
113  [~,ssu] = min(abs((IP3+mar)-S1_t));       % Find index of start slope up
114  [~,esu] = min(abs((IP4-mar)-S1_t));       % Find index of end slope up
115
116  PS_downslope1 = S1_diff1_f(ssd:esd);
117  PS_upslope1 = S1_diff1_f(ssu:esu);
118
119  PS_downslope2 = S1_diff2_f(ssd:esd);
120  PS_upslope2 = S1_diff2_f(ssu:esu);
121
122  [FIT_PS_down1,GOF_PS_down1] = fit(S1_t(ssd:esd),PS_downslope1,'poly1'); %
         Linear fit of slope down
123  [FIT_PS_up1,GOF_PS_up1] = fit(S1_t(ssu:esu),PS_upslope1,'poly1');        %
         Linear fit of slope up
124
125  %Compute calibration parameter
126  RS_down_coef = coeffvalues(FIT_RS_down);          %Retrieve coefficient
         values from cfit structure
127  PS_down_coef = coeffvalues(FIT_PS_down1);
128  RS_up_coef = coeffvalues(FIT_RS_up);
129  PS_up_coef = coeffvalues(FIT_PS_up1);
130
131  calpar_down = RS_down_coef(1)/PS_down_coef(1);   %Calibration parameter
         based on downward slope
132  calpar_up   = RS_up_coef(1)/PS_up_coef(1);
133
134  calpar_avg = (calpar_down+calpar_up)/2;           %Average of down and up
135
136  S1_diff1_F = S1_diff1_f*calpar_avg;
137
138  %Resample reference sensor output to match prototype sensor output
139  N1=length(RS_F);
140  N2=length(S1_diff1_F);
141  [P,Q]=rat(N2/N1);
142  RS_F_rs=resample(RS_F,P,Q);
143  RS_t_rs=resample(RS_t,P,Q);
144
145  [~,em] = min(abs(IP4+2-S1_t));          % index to avoid resample noise at end
146
147  % Evaluate error
```

```
148  NRMSEP=100*sqrt(mean((RS_F_rs(1:em)'-S1_diff1_F(1:em)).^2))/max(abs(
         RS_F_rs)); %Normalized RMSE percentage
149
150  [MRE,i_MRE] = max(100*(abs(RS_F_rs(1000:3000)'-S1_diff1_F(1000:3000))./
         RS_F_rs(1000:3000)'));      % MRE and index (excluding zones close to 0)
151
152  % Evaluate noise
153  [~,sp] = min(abs((IP2+mar)-S1_t));      % Start max
154  [~,ep] = min(abs((IP3-mar)-S1_t));      % End max
155
156  noise_band = max(S1_diff1_F(sp:ep))-min(S1_diff1_F(sp:ep));
157  noise_std = std(S1_diff1_F(sp:ep));
158
159  figure
160  plot(RS_t_rs(1:em),RS_F_rs(1:em)); hold on
161  plot(S1_t(1:em),S1_diff1_F(1:em))
162  xlabel('Time [s]')
163  ylabel('Force [N]')
164  legend('Reference sensor','Prototype sensor')
165  title(strcat('Calibrated output of prototype sensor (NRMSE =',num2str(
         NRMSEP),'%)'))
166
167
168  %% Fit insertion depth behaviour
169  Ins_depth = (50:10:140)';
170
171  Calpar_NS = [0.0336, 0.0276, 0.0242, 0.0215, 0.0198, 0.0179, 0.0173,
         0.0164, 0.0159, 0.0156]'; %X-axis calibration constants
172  Calpar_EW = [0.0384, 0.0324, 0.0278, 0.0251, 0.0232, 0.0215, 0.0209,
         0.0202, 0.0197, 0.0204]'; %Y-axis calibration constants
173
174  [FIT_pol4,GOF_pol4] = fit(Ins_depth,Calpar_NS,'poly4');
175  [FIT_pol4_R,GOF_pol4_R] = fit(Ins_depth,Calpar_EW,'poly4');
176
177  figure
178  scatter(Ins_depth,Calpar_NS); hold on
179  plot(FIT_pol4,'b')
180  scatter(Ins_depth,Calpar_EW)
181  plot(FIT_pol4_R,'r')
182  xlabel('Insertion Depth [mm]')
183  ylabel('Linear Calibration parameter')
184  legend('X-axis measured',strcat('X-axis fit (R2 =',num2str(GOF_pol4.
         rsquare),')'),'Y-axis measured',strcat('Y-axis fit (R2 =',num2str(
         GOF_pol4_R.rsquare),')'))
185  title('Fitted fourth-order polynomial for X- and Y-axis')
```

D.4. Matlab code for live plotting of calibrated data

The code listed below is used to acquire and plot live data after the sensor has been calibrated. The output is plotted with a 10 point moving average filter. The output is a single force value without direction. A direction vector could however be easily calculated if needed.

```
1  % Connect to arduino through serial connection
2  arduino = serial('COM5','BaudRate',115200);
3  fopen(arduino);
4
```

```matlab
5  %% Offset correction
6  % determine offset correction factors
7
8  Ncor = 500;        %Number of samples for offset correction for live plotting
9
10 dc = zeros(Ncor,4);
11
12 for ii=1:Ncor                              %Acquire Ncor samples
13     fprintf(arduino,"m");
14     dc(ii,1:4)= fscanf(arduino,'%d');
15 end
16
17 dc_mean = mean(dc,1);
18 Corr = [dc_mean(1)-dc_mean(3),dc_mean(2)-dc_mean(4)];  %Find deviation
         from zero in differential measurements
19
20 %%  Live plotting of calibrated data
21
22 % Provide insertion depth
23 ISD = 100; % insertion depth measured from instrument to trocar tip [mm]
24
25 % Load calibration curves
26 NS_cal_curve = load('220722_NS_cal.mat');
27 EW_cal_curve = load('220722_EW_cal.mat');
28
29 % Determine calibration parameters for given insertion depth
30 Calpar_EW = feval(EW_cal_curve.FIT_a_pol4_R,ISD);
31 Calpar_NS = feval(NS_cal_curve.FIT_t_pol4,ISD);
32
33 MA = 10;                      %Number of points averaged for live plot
34 jj = MA;
35 kk = 1;
36 Tmax = 45;                    %Sampling time
37 data = zeros(Tmax*120,5);    %Oversized data matrix 4
38 tic
39
40 mafdata = zeros(Tmax*120,4);
41 p1 = animatedline('Color', 'r');
42
43 Tstart = toc;
44 while (toc< Tstart+Tmax)
45     jj = jj+1;
46     fprintf(arduino,"m");
47     data(jj,1:4)= fscanf(arduino,'%d');          %Retrieve data from arduino
48     data(jj,5) = toc;
49     mafdata(jj,1:4) = mean(data(jj-MA:jj,1:4),1);    % apply moving average
50     NS_diff = (mafdata(jj,2)-mafdata(jj,4)-Corr(2))*Calpar_NS;      %
         Corrected and calibrated X-axis output
51     EW_diff = (mafdata(jj,1)-mafdata(jj,3)-Corr(1))*Calpar_EW;      %
         Corrected and calibrated Y-axis output
52
53     if (jj == kk*11)                     % Plot every Xth point
54         addpoints(p1,jj,sqrt(EW_diff^2+NS_diff^2))   % Vector sum output of
             total force
55         drawnow;
56         kk = kk+1;
```

```
57     end
58 end
59
60 data = data(1:jj,:);
61 BW = jj/data(end,5);
62
63 fclose(arduino);
64 clear arduino
65
66 %% Store data
67
68 writematrix(data,'220725_S3_SE_100_6mm.csv');
```

Bibliography

[1]   P Puangmali et al. "State-of-the-art in force and tactile sensing for minimally invasive surgery". In: *IEEE Sensors Journal* 8.4 (2008), pp. 371–381.

[2]   AL Trejos, RV Patel, and MD Naish. "Force sensing and its application in minimally invasive surgery and therapy: a survey". In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 224.7 (2010), pp. 1435–1454.

[3]   C Freschi et al. "Technical review of the da Vinci surgical telemanipulator". In: *The International Journal of Medical Robotics and Computer Assisted Surgery* 9.4 (2013), pp. 396–406.

[4]   SF Hardon et al. "A new modular mechanism that allows full detachability and cleaning of steerable laparoscopic instruments". In: *Surgical endoscopy* 33.10 (2019), pp. 3484–3493.

[5]   G Picod et al. "What can the operator actually feel when performing a laparoscopy?" In: *Surgical Endoscopy and Other Interventional Techniques* 19.1 (2005), pp. 95–100.

[6]   A Alassi et al. "Development and kinematic analysis of a redundant, modular and backdrivable laparoscopic surgery robot". In: *2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE. 2018, pp. 213–219.

[7]   F Richter, RK Orosco, and MC Yip. "Open-sourced reinforcement learning environments for surgical robotics". In: *arXiv preprint arXiv:1903.02090* (2019).

[8]   L Toledo et al. "Study of sustained forces and the working space of endoscopic surgery instruments". In: *Annales de chirurgie*. Vol. 53. 1999, p. 587.

[9]   D Tran et al. "Abdominal wall muscle elasticity and abdomen local stiffness on healthy volunteers during various physiological activities". In: *Journal of the mechanical behavior of biomedical materials* 60 (2016), pp. 451–459.

[10]   R.C. Hibbeler. *Mechanics of Materials*. 9th. Pearson, 2013.

[11]   Anton Van Beek. *Advanced engineering design: lifetime performance and reliability*. 7th. Delft University of Technology, 2019.

[12]   WA Rutula, DJ Weber, and HICPAC. "Guideline for Disinfection and Sterilization in Healthcare Facilities". In: (2008).

[13]   Joost van Doornmalen and Klaas Kopinga. "Review of surface steam sterilization for validation purposes". In: *American journal of infection control* 36.2 (2008), pp. 86–92.

[14]   M Suurveld. "Force Sensing Strategies for Robot-Assisted Minimally Invasive Surgery: A Review". *Unpublished*, Available from the author upon request. 2021.

[15]   RS Figliola and DE Beasley. *Theory and design for mechanical measurements*. 6th. John Wiley & Sons, 2016, pp. 490–514.

[16]   PPL Regtien and E Dertien. *Sensors for mechatronics*. Elsevier, 2018.

[17]   R Munnig Schmidt, G Schitter, and A Rankers. *The Design of High Performance Mechatronics-: High-Tech Functionality by Multidisciplinary System Integration*. 2nd. Ios Press, 2014.

[18]   MC Yip, SG Yuen, and RD Howe. "A robust uniaxial force sensor for minimally invasive surgery". In: *IEEE transactions on biomedical engineering* 57.5 (2010), pp. 1008–1011.

[19]   A Cusano, A Cutolo, and M Giordano. "Fiber Bragg Gratings Evanescent Wave Sensors: A View Back and Recent Advancements". In: *Sensors: Advancements in Modeling, Design Issues, Fabrication and Practical Applications*. Springer Berlin Heidelberg, 2008, pp. 113–152.

[20]   V Mishra et al. "Fiber grating sensors in medicine: Current and emerging applications". In: *Sensors and Actuators A: Physical* 167.2 (2011), pp. 279–290.

[21] J Fraden. *Handbook of modern sensors: physics, designs, and applications*. 5th. Springer Science & Business Media, 2016.

[22] Q Liang et al. "Multi-component FBG-based force sensing systems by comparison with other sensing technologies: a review". In: *IEEE Sensors Journal* 18.18 (2018), pp. 7345–7357.

[23] DM Stefanescu. *Handbook of force transducers: principles and components*. Springer Science & Business Media, 2011.

[24] D van der Pol. "Assessment of soft tissue tension during total hip arthroplasty". MA thesis. Delft University of Technology, 2017.

[25] E Noohi, S Parastegari, and M Žefran. "Using monocular images to estimate interaction forces during minimally invasive surgery". In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2014, pp. 4297–4302.

[26] MI Tiwana, SJ Redmond, and NH Lovell. "A review of tactile sensing technologies with applications in biomedical engineering". In: *Sensors and Actuators A: physical* 179 (2012), pp. 17–31.

[27] US Seibold. "An advanced force feedback tool design for minimally invasive robotic surgery". PhD thesis. Technische Universität München, 2013.

[28] D Jones et al. "A soft multi-axial force sensor to assess tissue properties in realtime". In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2017, pp. 5738–5743.

[29] B Darvish et al. "A novel tactile force probe for tissue stiffness classification". In: *arXiv preprint arXiv:1808.02862* (2018).

[30] J Peirs et al. "A micro optical force sensor for force feedback during minimally invasive robotic surgery". In: *Sensors and Actuators A: Physical* 115.2-3 (2004), pp. 447–455.

[31] Y Noh et al. "A Contact Force Sensor Based on S-Shaped Beams and Optoelectronic Sensors for Flexible Manipulators for Minimally Invasive Surgery (MIS)". In: *IEEE Sensors Journal* 20.7 (2019), pp. 3487–3495.

[32] AL Trejos et al. "Design and evaluation of a sterilizable force sensing instrument for minimally invasive surgery". In: *IEEE Sensors Journal* 17.13 (2017), pp. 3983–3993.

[33] M Fakhry, F Bello, GB Hanna, et al. "A real-time compliance mapping system using standard endoscopic surgical forceps". In: *IEEE Transactions on Biomedical Engineering* 56.4 (2009), pp. 1245–1253.

[34] JD Westwood et al. "Quantifying surgeon grasping mechanics in laparoscopy using the blue DRAGON system". In: *Medicine Meets Virtual Reality 12: Building a Better You: the Next Tools for Medical Education, Diagnosis, and Care* 98 (2004), p. 34.

[35] TL Brooks. "Telerobotic response requirements". In: *1990 IEEE International Conference on Systems, Man, and Cybernetics Conference Proceedings*. 1990, pp. 113–120.

[36] GA Fontanelli et al. "An External Force Sensing System for Minimally Invasive Robotic Surgery". In: *IEEE/ASME Transactions on Mechatronics* 25.3 (2020), pp. 1543–1554.

[37] N Zemiti et al. "Mechatronic design of a new robot for force control in minimally invasive surgery". In: *IEEE/ASME Transactions On Mechatronics* 12.2 (2007), pp. 143–153.

[38] S Belsley. *Laparoscopic trocars*. Accessed: 9-6-2021, from: `www.laparoscopic.md/surgery/instruments/trocar`.

[39] TJ Teo, G Yang, and I Chen. "Compliant manipulators". In: *Handbook of Manufacturing Engineering and Technology; Springer: London, UK* (2014), pp. 1–63.

[40] Thierry Rey et al. "Influence of the temperature on the mechanical behaviour of filled and unfilled silicone rubbers". In: *Polymer Testing* 32.3 (2013), pp. 492–501.

[41] Honeywell. *Hall Effect Sensing and Aplication*. Accessed: 10-8-2021, from: `https://sensing.honeywell.com/honeywell-sensing-sensors-magnetoresistive-hall-effect-applications-005715-2-en2.pdf`.

[42] Supermagnete. *Neodymium magneten: Die met de superkrachten*. Accessed: 30-8-2021, from: `https://www.supermagnete.nl/supermagneten-sterke-neodymium-magneten`.

[43] C.R. McMillin. "An assessment of elastomers for biomedical applications". In: *High Performance Biomaterials: A Complete Guide to Medical and Pharmaceutical Applications* (2017), p. 37.

[44] Ryuichiro Yoda. "Elastomers for biomedical applications". In: *Journal of Biomaterials Science, Polymer Edition* 9.6 (1998), pp. 561–626.

[45] Brian J Millar, Sanjukta Deb, et al. "Effect of autoclave sterilisation on the dimensional stability and tear strength of three silicone impression materials". In: *Open Journal of Stomatology* 4.12 (2014), p. 518.

[46] H Terheyden et al. "Sterilization of elastic ligatures for intraoperative mandibulomaxillary immobilization". In: *British Journal of Oral and Maxillofacial Surgery* 38.4 (2000), pp. 299–304.

[47] Agnieszka Szczęsna et al. "Inertial motion capture costume design study". In: *Sensors* 17.3 (2017), p. 612.

[48] Julie Diani, Bruno Fayolle, and Pierre Gilormini. "A review on the Mullins effect". In: *European Polymer Journal* 45.3 (2009), pp. 601–612.

[49] SR Trout. "Material selection of permanent magnets, considering thermal properties correctly". In: *Proceedings: Electrical Insulation Conference and Electrical Manufacturing and Coil Winding Conference (Cat. No. 01CH37264)*. IEEE. 2001, pp. 365–370.

[50] AG Clegg et al. "The temperature stability of NdFeB and NdFeBCo magnets". In: *IEEE Transactions on Magnetics* 26.5 (1990), pp. 1942–1944.

[51] EWV Chaves. *Notes on continuum mechanics*. Springer Science & Business Media, 2013.

[52] GA Holzapfel. "Nonlinear solid mechanics,(2000)". In: *John Wiley&Sons., Ltd* ().

[53] Daniel Steck et al. "Mechanical responses of Ecoflex silicone rubber: Compressible and incompressible behaviors". In: *Journal of Applied Polymer Science* 136.5 (2019), p. 47025.