TU Delft

Delft University of Technology

# Shape fitting

## Application to the estimation of wafer chuck deformation

Vogel, Johan

**DOI**
10.4233/uuid:8f23c2ff-1589-4a8c-8868-03c2f42b4d73

**Publication date**
2016

**Document Version**
Final published version

**Citation (APA)**
Vogel, J. (2016). *Shape fitting: Application to the estimation of wafer chuck deformation*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:8f23c2ff-1589-4a8c-8868-03c2f42b4d73

**Important note**
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

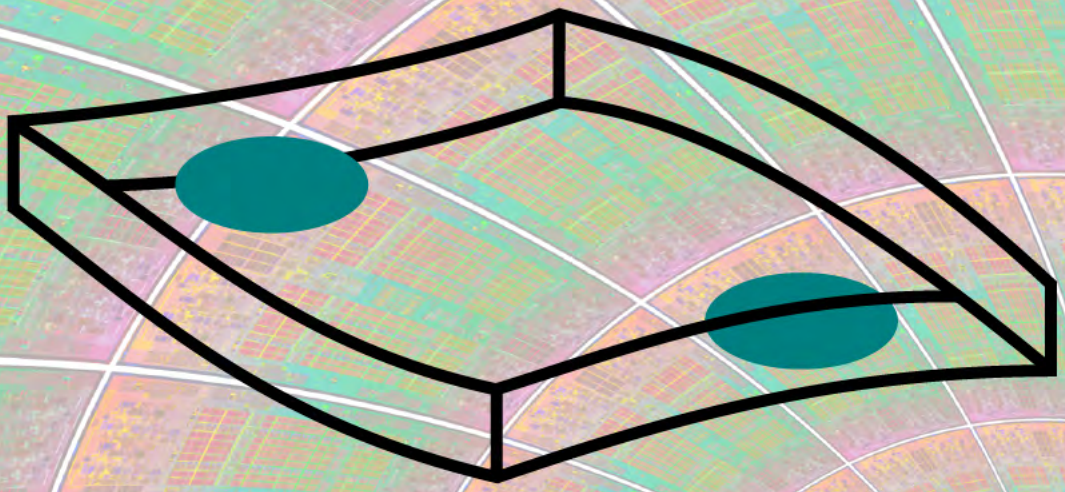
**Copyright**
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

*This work is downloaded from Delft University of Technology.*
*For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.*

# Shape fitting

## Application to the estimation of wafer chuck deformation



Johan Vogel

# SHAPE FITTING

## APPLICATION TO THE ESTIMATION OF WAFER CHUCK DEFORMATION

# SHAPE FITTING

## APPLICATION TO THE ESTIMATION OF WAFER CHUCK DEFORMATION

**Proefschrift**

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.ir. K. C. A. M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op maandag 10 oktober 2016 om 12:30 uur

door

**Johan Gregorius VOGEL**

werktuigkundig ingenieur
geboren te De Bilt, Nederland.

Dit proefschrift is goedgekeurd door de promotor:

Prof. ir. R.H. Munnig Schmidt

Copromotor:

Ir. J.W. Spronck

Samenstelling promotiecommissie:

| | |
|---|---|
| Rector Magnificus, | voorzitter |
| Prof. ir. R.H. Munnig Schmidt, | Technische Universiteit Delft, promotor |
| Ir. J.W. Spronck, | Technische Universiteit Delft, copromotor |
| | |
| Onafhankelijke leden: | |
| Prof. dr. ir. H. Butler, | Technische Universiteit Eindhoven, ASML N.V. |
| Prof. dr. ir. J. van Eijk, | Mice B.V. |
| Prof. dr. ir. F. van Keulen, | Technische Universiteit Delft |
| Dr. A. Tejada Ruiz, | TNO – innovation for life, Helmond |
| Prof. dr. ir. M. Verhaegen, | Technische Universiteit Delft |
| Prof. dr. ir. J.L. Herder, | Technische Universiteit Delft, reservelid |

TU Delft Delft University of Technology

STW Enabling new technology

This research was supported by the Dutch funding agency Stichting Technologie en Wetenschap (STW).

*Keywords:* Shape estimation, shape fitting, wafer chuck deformation, least squares optimisation, position sensing

*Printed by:* Ipskamp Printing, Enschede

*Front:* The foreground shows – with a twist inspired by a famous 20th century Dutch artist – a plate that mimicks a deformed wafer chuck. The background shows a spiral-shaped pattern of chip pictures, representing the continuous persuit of smaller chip dimensions (source of the chip picture: Intel).

*Back:* Picture of the experimetal setup with the position controlled plate.

Copyright © 2016 by J.G. Vogel

ISBN 978-94-028-0338-9

This dissertation is available electronically at http://repository.tudelft.nl/.

*Aan mijn ouders*

# Summary

The performance of computer chips has dramatically increased in the last 60 years.[1] This has, amongst others, been made possible by increasing the accuracy of the lithography machines, the machines that project patterns of light onto the chip substrate and thereby define the small details that form the electronics. An increased accuracy also leads, however, to increased machine costs, so that the machines need to produce more chips per unit of time to limit the costs per chip. → Ch. 1

One of the ways to increase the production rate is to increase the size of the wafer, the chip substrate, so that more chips fit onto a single wafer. In the past, the wafer diameter increased a number of times, up to today's 300 mm. It is almost unavoidable that the wafer size will be increased further to 450 mm.

An increased wafer size automatically leads to increased dimensions of the lithography machine's parts. This has important consequences, especially for the part that supports the wafer, the so-called wafer chuck. The current wafer chucks are relatively thick and stiff, so that they can be considered as a rigid body. The out-of-plane displacement of the point on the wafer that is being exposed, the point of interest, can then be found directly from the four position sensors at the chuck's corners.

Due to mass limitations, it is difficult, however, to maintain the wafer chuck's required stiffness. For example, if the wafer chuck is scaled in order to follow the transition from 300 mm to 450 mm and its mass has to be kept constant, its stiffness decreases by a factor of 25. A less stiff chuck experiences higher deformation when subjected to disturbance forces. In that case the position of the point of interest can not be estimated with sufficient accuracy anymore using the rigid body approach. Therefore, the displacement at the point of interest, including the deformation, needs to be estimated based on measurements of additional sensors.

A commonly employed approach to estimation is *shape fitting*[2]. Shape fitting estimates the shape of the displacement field by fitting a set of shapes to the measurements. Shape fitting does not use the history of the measurement signals; it only uses the current measurement values and recombines them into the estimates. Shape fitting was not applied to the wafer chuck before in literature.

The goal of this thesis is to develop an estimation methodology for wafer

[1] For a Dutch translation of this summary, the reader is referred to p. xi.

[2] This name is introduced in this thesis, as the shape fitting method has no generally accepted name in literature, although it is used extensively.

chuck deformation, based on shape fitting principles, considering the aim for low estimation error and the specific requirements related to the wafer chuck application. In the thesis the selection of the sensor type, the sensor placement and the algorithm for finding an adequate estimator are addressed. Furthermore, the shape fitting method is demonstrated using an experimental setup.

→ Ch. 2

**Sensor type** A displacement field can be estimated from local position measurements with respect to an external reference or from the spatial or temporal derivatives of displacement, for example strain and acceleration.

Strain sensors measure relative displacements in the chuck, which are a function of spatial derivates of the displacement field. The signal of an accelerometer has to be integrated twice to obtain absolute position. This integration increases the noise level at low frequency, so that only the higher-frequency part of the signal is useful. Both accelerometers and strain sensors can be relatively small and do not need an external reference, so that they can be placed relatively unrestricted over the chuck's surface.

The measurement of the position of a point on the chuck requires an external reference. This is only available above the chuck around the lens column. To be as much as possible in close sight of the reference, the position sensors need to be placed at the sides of the wafer chuck. A configuration with four additional position sensors at the centres of the sides, complementing the existing ones at the chuck's corners, is proposed.

→ Ch. 3

**Algorithm** The disturbances that work on the wafer chuck are most dominant at the frequencies well-below the wafer chuck's lowest resonance frequency. Thus, the wafer chuck's response is more or less independent of frequency or *quasi-static*. For this reason shape fitting is an adequate, yet simple, approach for estimation.

In literature, different methods to construct the shape fitting estimator have been presented. Often dynamic modeshapes are used, but for a quasi-static system they are of less value, as none of the eigenmodes is excited at its eigenfrequency. Alternatively, a set of displacement shapes that capture the system's response to its disturbances can be constructed, forming the so-called snapshot-matrix. Methods based on snapshots are appealing because of the straightforward way in which foreknowledge on the system's dynamics and typical disturbances is included. The snapshot-matrix can be partly constructed from measurements, but must include model data if the position of the point of interest can not be measured.

→ Ch. 4

Two snapshot-based techniques from literature are the Proper Orthogonal Mode (POM) and the Least Squares (LS) estimation techniques. The performance of the techniques was compared on the basis of their methodical estimation error and their error due to sensor noise. There is an inherent trade-off between the two error sources and either of the techniques facilitates to make this trade-off. The LS technique, if used in combination with adequate regularisation, leads – by definition – to a lower total estimation error than the POM technique, as it directly minimises that error. Next to that, the LS technique has a higher computational

efficiency and yields results that, unlike the POM technique, are not influenced by the choice of the total set of locations to be estimated for.

As the wafer chuck experiences rigid body motion, its displacement field contains rigid body contributions. These contributions are not quasi-static but dynamic and should, thus, be removed before the shape fitting estimator can be found. It is in practice only possible to *estimate* the rigid body position, but with such an estimate, the so-called flexible residue can be obtained, which is not a function of the rigid body position. Furthermore, if the number of position measurements exceeds the number of rigid body modes, the resulting redundancy in the flexible residues at the sensors need to be removed before calculating the shape fitting estimator. → Ch. 5

To keep the rigid body motion limited, the wafer chuck's position is in practice actively controlled. This leads to dynamic behaviour of the flexible residue around the controller bandwidth. The snapshot matrix must then be build up from sets of snapshots for a grid of frequency samples. The information on relative phase between the signals can be captured in complex numbers. The LS technique was adapted to deal with such a complex snapshot matrix.

In a numerical case study with eight position sensors and disturbances distributed over the wafer-lens interface, the estimation error with shape fitting was shown to reduce by a factor of 70 as compared to the rigid body estimate. The lower the bandwidth for which the shape fitting estimator is optimised, the better the estimation performance. A hybrid estimator was proposed that fuses the estimate of a low-frequency shape fitting estimator band with a higher-frequency position signal that is obtained from double integration of acceleration.

**Sensor placement** Effective placement of the sensors is essential for obtaining good estimation performance. Automated selection is favourable above hand-picking, as sensor placement is not always intuitive. Requirements for a placement algorithm include that the algorithm should fit into the framework of shape fitting and that the algorithm should be effective and fast. → Ch. 6

Four placement algorithms from literature were compared in different test-cases. In a case in which the first eight dynamic modeshapes of a plate had to be distinguished and identified using position sensors, the placement methods yielded comparable results that were often equal to the theoretically optimal placement. For placement of strain sensors, the performance of the algorithms varied. Here, also the influence of the number of candidates in the initial candidate set was tested. In general, the algorithms do not yield better results for larger candidate sets. It was shown that applying the algorithms to multiple relatively small candidate sets leads to better sensor configurations and drastically improves the chance of convergence.

It is proposed to perform sensor placement in the framework of shape fitting by a procedure that distils a set of the most prominent shapes from a snapshot matrix, amongst others by using the proper orthogonal decomposition. Then, the algorithms from literature can be used to optimise for distinguishability of this set of shapes. Still, the actual placement goal in shape fitting can be rather complex and distant from what the algorithms in literature optimise for. The

sensor placement procedure was applied to the numerical case and was shown to reduce the maximum estimation error of the shape fitting estimator by 40 % as compared to the previous manually selected sensor configuration.

→ Ch. 7

**Experimental verification** To validate the system model and demonstrate shape fitting in a physical setup an experimental setup was developed. The setup mimics the wafer chuck system and consists of an in the out-of-plane direction free-floating plate, displacement sensors, a metrology frame, force actuators and a three degrees of freedom control system.

The transfer functions from the actuators to the rigid body displacement and local deformation were measured and compared to the results obtained with the model. As they show close resemblance, it was concluded that the system model is valid.

Shape fitting was experimentally demonstrated by estimating the response to one disturbance actuator based on six flexible residuals in different bandwidths. For an estimation bandwidth of 80 Hz, which is between the control bandwidth and the plate's first eigenfrequency, the resulting estimation error was a factor of 22 smaller than if using the rigid body method. For a larger bandwidth of 150 Hz, which is above the first eigenfrequency, the resulting error of the shape fitting estimator increased by a factor of 10, but the error of the rigid body estimator increased even more, so that the estimation error of shape fitting became a factor of 40 lower than the error of the rigid body estimator.

An accelerometer was mounted to the setup's chuck to verify displacement sensing using the hybrid shape fitting estimator. The error of the hybrid estimator was shown to be close to the noise level of the reference sensor.

→ Ch. 8

**Conclusion** It can be concluded that the shape of the wafer chuck can be estimated effectively based on a limited number of additional sensors using shape fitting. Shape fitting is a relative simple approach, yet it is effective, as the chuck's deformation is mainly quasi-static. The snapshot matrix-based LS technique in combination with regularisation based on the knowledge of the sensors' noise floor leads to a static estimator with minimum total error. The technique was adapted, so that it can deal with rigid body dynamics and dynamic effects of the flexible residue. Effective sensor configurations can be found by incorporating sensor placement methods from literature in the shape fitting framework.

Simulations and practical experiments show that the shape fitting estimator yields a considerably improved position estimation of the point of interest as compared to the current rigid body approach. As such, shape fitting can be considered as a step forward to a lighter chuck. Necessary further steps include developing a method for generating snapshot matrices with both model and measurement data and incorporating the the actual disturbance conditions of the wafer chuck in the shape fitting method.

# Samenvatting

De prestaties van computerchips zijn in de laatste 60 jaar zeer sterk vooruit gegaan. Dit is onder meer te danken aan de verhoging van de nauwkeurigheid van de lithografiemachines, de machines die lichtpatronen op het chipsubstraat projecteren en daarmee de fijne electronische details definiëren. Een hogere nauwkeurigheid leidt echter ook tot verhoogde machinekosten, waardoor de machines per tijdseenheid meer chips moeten produceren om de kosten per chip te beperken. → Ch. 1

Eén van de manieren om de productiesnelheid te verhogen is het vergroten van de *wafer*, het substraat van de chips, zodat er meer chips op een wafer passen. In het verleden is de waferdiameter een aantal keer vergroot, tot de huidige 300 mm. Het is bijna onvermijdelijk dat de wafergrootte in de toekomst verder wordt vergroot naar 450 mm.

Een grotere wafer vraagt automatisch ook om vergroting van andere onderdelen van de lithografiemachine. Dit heeft belangrijke consequenties, vooral op het onderdeel dat de wafer ondersteunt, het zogenaamde *wafer chuck*. De huidige wafer chucks zijn relatief dik en stijf en kunnen daarom beschouwd worden als een star lichaam. De positie in de uit-het-vlak-richting van het punt op de wafer dat belicht wordt, het *point-of-interest*, kan dan direct bepaald worden op basis van de vier positiesensors op de hoeken van het chuck.

Vanwege beperkingen ten aanzien van de massa van het wafer chuck is het echter moeilijk om het chuck voldoende stijf te houden. Als de diameter van het wafer chuck bijvoorbeeld vergroot wordt van 300 mm naar 450 mm, terwijl zijn massa constant moet blijven, dan neemt zijn stijfheid een factor 25 af. Een minder stijf wafer chuck vervormt sterker onder invloed van verstoringskrachten. In dat geval kan de positie van het point-of-interest niet voldoende nauwkeurig meer geschat worden op basis van de huidige star-lichaamsmethode. Daarom moet de verplaatsing van het point-of-interest, waarvan een deel afkomstig is van de vervorming, geschat worden op basis van metingen van extra sensoren.

Een vaak gebruikte schatmethode is *shape fitting*[3]. Shape fitting schat de vorm van het verplaatsingsveld door het interpoleren van de metingen met behulp van een verzameling van vormen. Shape fitting maakt geen gebruik van de geschiedenis van de meetsignalen; het maakt alleen gebruik van de huidige meetwaarden en hercombineert deze tot de schattingen. Shape fitting is nog niet eerder in literatuur toegepast op het wafer chuck.

[3]Deze naam wordt door dit proefschrift geïntroduceerd, omdat de shape fittingmethode – hoewel extensief gebruikt – geen algmeen gangbare naam heeft in de literatuur.

Het doel van dit proefschrift is het ontwikkelen van een schattingmethodologie van de vervorming van het wafer chuck, gebaseerd op de beginselen van shape fitting, waarbij rekening wordt gehouden met de wens van een lage schatfout en de specifieke vereisten in de wafer chucktoepassing. In deze thesis wordt aandacht besteed aan de keuze van het sensortype, aan de selectie van de sensorlocaties en aan het algoritme voor het vinden van een adequate schatter. Verder wordt de shape fittingmethode gedemonstreerd aan de hand van een proefopstelling.

→ Ch. 2

**Sensortype** Een verplaatsingsveld kan geschat worden op basis van lokale positiemetingen ten opzichte van een externe referentie of op basis van de spatiële of temporele afgeleiden van verplaatsing, zoals rek en acceleratie.

Reksensors meten relatieve verplaatsingen op het chuck, die een functie zijn van de spatiële afgeleiden van het verplaatsingsveld. Het signaal van een accelerometer dient tweemaal te worden geïntegreerd om absolute verplaatsing te verkrijgen. Deze integratie verhoogt het ruisniveau op lage frequentie, waardoor slechts het hoogfrequente deel van het signaal bruikbaar is. Zowel accelerometers en reksensors kunnen relatief klein zijn en hebben geen externe referentie nodig, waardoor de plaatsing over het oppervlak van het chuck relatief zonder restricties is.

Voor het meten van de positie van een punt op het chuck is een externe referentie nodig. Deze is alleen aanwezig boven het chuck rondom de lenskolom. Om zoveel mogelijk in het zicht van de referentie te zijn moeten de positiesensors aan de zijkanten van het wafer chuck geplaatst worden. Er wordt een configuratie voorgesteld met, naast de huidige positiesensors op de hoeken van het chuck, vier extra sensors op de middens van de zijkanten.

→ Ch. 3

**Algoritme** De verstoringen die aangrijpen op het wafer chuck zijn het meest dominant bij frequenties die voldoende onder de eigenfrequenties van het wafer chuck liggen. Om die reden is de responsie van het wafer chuck min of meer frequentieonafhankelijk, ofwel *quasistatisch*. Shape fitting is daarom naast een simpele ook een adequate schattingsmethode.

Er zijn verscheidene methodes voor shape fitting gepresenteerd in literatuur. Vaak worden dynamische modale vormen gebruikt, maar voor een quasistatisch systeem zijn deze van minder belang, omdat geen van de eigenmodes wordt geëxciteerd rond zijn eigenfrequentie. In plaats daarvan kan een verzameling worden geconstrueerd uit de verplaatsingsvormen die corresponderen met de responsies van het systeem op zijn verstoringen, waarmee manier de zogenaamde snapshotmatrix gevormd kan worden. Methodes gebaseerd op snapshots zijn aantrekkelijk, omdat ze op een ongecompliceerde manier van de voorkennis omtrent de dynamica van het systeem en de typische verstoringen op een systeem gebruik maken. De snapshotmatrix kan deels worden opgebouwd uit metingen, maar bevat in elk geval de gegevens uit een model indien het point-of-interest niet bemeten kan worden.

→ Ch. 4

Twee snapshot-gebaseerde technieken uit de literatuur zijn de Proper Orthogonal Mode (POM) en de kleinste kwadraten (LS) schattechnieken. De prestaties van deze technieken zijn vergeleken op basis van hun methodische

schattingsfout en de fout ten gevolge van sensorruis. Er moet een inherente afweging gemaakt worden tussen deze twee foutbronnen en beide technieken bieden de mogelijkheid om deze afweging te maken. De LS-techniek leidt, indien gebruikt in combinatie met adequate regularisatie, per definitie tot een lagere totale schatfout dan de POM-techniek, omdat de LS-techniek direct voor deze fout minimaliseert. Daarnaast is de LS-techniek rekentechnisch efficiënter en worden zijn resultaten, in tegenstelling tot die van de POM-techniek, niet beïnvloed door de keuze van de totale verzameling van locaties waarvoor geschat moet worden.

Omdat een wafer chuck star-lichaamsbewegingen ondergaat bevat zijn verplaatsingsveld bijdragen van star-lichaamsmodes. Deze bijdragen zijn niet quasistatisch maar dynamisch en moeten daarom verwijderd worden alvorens de shape fitting-schatter kan worden gevonden. In de praktijk is het alleen mogelijk om de star-lichaamspositie te *schatten*, maar gebruik makend van een dergelijke schatting kan het zogenaamde flexibele residu worden bepaald, welke niet een functie is van de star-lichaamspositie. Verder dient, wanneer het aantal positiemetingen groter is dan het aantal star-lichaamsvrijheidgraden, de resulterende redundantie te worden verwijderd uit de flexibele residuen voordat de shape fitting-schatter bepaald wordt. → Ch. 5

Om de star-lichaamsbeweging te beperken wordt de positie van het wafer chuck in de praktijk actief geregeld. Hierdoor vertoont het flexibele residu dynamisch gedrag rond de regelbandbreedte. De snapshotmatrix moet in dat geval opgebouwd worden uit verzamelingen van snapshots voor een reeks van frequentiemonsters. De informatie omtrent de relatieve fase van de signalen kan worden beschreven met behulp van complexe getallen. De LS-techniek is aangepast om te kunnen werken met een dergelijke complexe snapshotmatrix.

In een numerieke voorbeeldstudie met acht positiesensors en over het vlak tussen de wafer en de lens verpreide verstoringen, leidde het gebruik van de shape fitting-schatter tot een reductie van schattingsfout met een factor 70 ten opzichte van de star-lichaamsschatter. Hoe lager de bandbreedte waarvoor de shape fitting-schatter wordt geoptimaliseerd, hoe beter schatter presteert. Er wordt een hybride schatter voorgesteld die de schatting van een laagfrequente shape fitting-schatter samenvoegt met een hoogfrequent positiesignaal verkregen uit tweevoudige integratie van acceleratie.

**Sensorplaatsing** Voor het behalen van een hoge schatterprestatie is effectieve plaatsing van de sensors essentieel. Geautomatiseerde selectie is wenselijker dan handselectie, omdat sensorplaatsing niet altijd intuïtief is. Vereisten voor een sensorplaatsingsalgoritme zijn onder meer dat het algoritme in te passen is binnen het shape fitting-kader en dat het effectief en snel is. → Ch. 6

Vier algoritmen voor sensorplaatsing uit de literatuur zijn vergeleken in verschillende tests. Bij een test waarin de eerste acht dynamische modale vormen van een plaat aan de hand van positiesensors onderscheiden en geïdentificeerd moesten worden behaalden de algoritmen resultaten vergelijkbaar met de theoretisch optimale plaatsing. Bij plaatsing van reksensors verschilden de prestaties van de algoritmes. Hierbij werd ook de invloed van het aantal sensorkandidaten in de initiële kandidaatverzameling getest. In het algemeen leverden de algoritmen geen betere resultaten bij het gebruik van

grotere kandidaatverzamelingen. De toepassing van de plaatsingsalgoritmen op verscheidene relatief kleine kandidaatverzamelingen leidde echter wél tot betere sensorconfiguraties en een aanzienlijk grotere kans op convergentie.

Voor het uitvoeren van sensorplaatsing in het shape fitting-raamwerk is een procedure geïntroduceerd die een verzameling van de meest prominente vormen extraheert uit een snapshotmatrix, onder meer gebruikmakend van de *proper orthogonal decomposition*. Vervolgens kunnen de algoritmes uit de literatuur worden gebruikt voor het optimaliseren van de onderscheidbaarheid binnen de verzameling van vormen. Desondanks kan het precieze optimalisatiedoel voor shape fitting tamelijk complex en verschillend zijn van het doel waarvoor de algoritmes optimaliseren. De procedure voor sensorplaatsing is toegepast op de numerieke voorbeeldstudie waarbij de maximale schatfout van de shape fitting-schatter met 40 % afnam ten opzicht van de eerder handgekozen sensorconfiguratie.

→ Ch. 7

**Experimentele verificatie** Om het systeemmodel te valideren en shape fitting in een fysieke opstelling te demonstreren is er een experimentele opstelling ontwikkeld. De opstelling bootst het wafer chuck-systeem na en bestaat uit een in uit-het-vlak-richting vrij zwevende plaat, verplaatsingssensors, een metrologieconstructie, krachtactuators en een drie-graden-van-vrijheid regelsysteem.

De overdrachtsfuncties van de actuators naar de star-lichaamsverplaatsingen en de lokale vervormingen zijn gemeten en vergeleken met uit het model verkregen resultaten. Omdat beide sterke overeenkomst vertoonden kon geconcludeerd worden dat het systeemmodel valide is.

Shape fitting is experimenteel gedemonstreerd door de responsie ten gevolge van een verstoringsactuator te schatten op basis van zes flexibele residuen. Bij een schatterbandbreedte van 80 Hz, welke ligt tussen de regelbandbreedte en de laagste mechanische eigenfrequentie, was de resulterende schattingsfout een factor 22 kleiner dan bij gebruik van de star-lichaamsmethode. Bij een hogere schattingsbandbreedte van 150 Hz, welke boven de laagste eigenfrequentie ligt, nam de schattingsfout van de shape fitting-schatter met een factor 10 toe, maar de fout van de star-lichaamsschatter nam nog sterker toe, waardoor de schattingsfout van de shape fitting-schatter een factor 40 lager werd dan die van de star-lichaamsschatter.

Een accelerometer werd bevestigd aan het chuck in de opstelling om de hybride shape fitting-schatter te verifiëren. De fout van de hybride schatter lag dichtbij het ruisniveau van de referentie-verplaatsingssensor.

→ Ch. 8

**Conclusie** Er kan geconcludeerd worden dat de vorm van het wafer chuck effectief geschat kan worden met behulp shape fitting op basis van een beperkt aantal toegevoegde sensoren. Shape fitting is een relatief eenvoudige, maar wel effectieve methode, omdat de vervorming van het chuck voornamelijk quasistatisch is. De snapshotmatrix-gebaseerde LS-techniek in combinatie met regularisatie gebaseerd op de voorkennis omtrent het ruisniveau van de sensors leidt tot een statische schatter met minimale totale fout. De techniek is aangepast om te kunnen werken met star-lichaamsdynamica en dynamische effecten van het flexibele residu. Effectieve sensorconfiguraties kunnen worden verkregen

door het invoegen van algoritmes voor sensorplaatsing in het shape fitting-kader.

Simulaties en praktijkexperimenten laten zien dat het gebruik van de shape fitting-schatter leidt tot een aanzienlijke verbetering van de positieschatting van het point-of-interest in vergelijking met de huidige star-lichaamsmethode. Om die reden kan shape fitting gezien worden als een stap voorwaarts naar een lichter chuck. Er zijn nog verschillende verdere stappen nodig, waaronder de ontwikkeling van een methode voor het genereren van snapshot matrices op basis van zowel gegevens uit het model en uit metingen en het rekening houden met de werkelijke verstoringscondities van het wafer chuck.

# Contents


**1 Introduction** **1**
- 1.1 A 'self-fulfilling' prophecy . . . 1
- 1.2 Chip production using wafer scanners . . . 1
- 1.3 Smaller details and a higher throughput . . . 2
- 1.4 Towards a lighter wafer chuck . . . 5
- 1.5 Estimating deformation using shape fitting . . . 6
- 1.6 Shape fitting estimation methodology wafer chuck deformation . . . 7
- 1.7 Method and structure of this thesis . . . 7

**2 System properties of the wafer chuck** **9**
- 2.1 Introduction to the system . . . 9
  - 2.1.1 Exposure path . . . 9
  - 2.1.2 The position and displacement of the wafer chuck . . . 11
  - 2.1.3 Motion control of the wafer chuck . . . 12
  - 2.1.4 Immersion layer . . . 13
  - 2.1.5 Foreknowledge on the disturbances . . . 14
- 2.2 Mechanical model and dynamic eigenmodes . . . 15
  - 2.2.1 Equations of motion in matrix form . . . 15
  - 2.2.2 Definition modeshapes and eigenfrequencies . . . 15
  - 2.2.3 Modal coordinates . . . 16
- 2.3 Finite element modelling . . . 18
  - 2.3.1 Plate finite element . . . 18
  - 2.3.2 Experimental validation . . . 20
- 2.4 Feasible sensor configurations for estimation . . . 21
  - 2.4.1 Measurands and sensor types . . . 21
  - 2.4.2 Obtaining absolute position from inertial measurements . . . 27
  - 2.4.3 Sensor placement candidates . . . 32
- 2.5 Conclusions . . . 33

**3 Estimation using shape fitting** **35**
- 3.1 Definition of the estimation problem . . . 35
- 3.2 Quasi-static system response and estimation . . . 36
- 3.3 Estimation methods in literature . . . 37
  - 3.3.1 Dynamic estimators . . . 38
  - 3.3.2 Static estimators . . . 38
- 3.4 The shape fitting method . . . 40

3.4.1 Procedure . . . 41
3.4.2 Estimation matrix $\boldsymbol{B}$ . . . 43
3.5 Shape fitting using a snapshot matrix . . . 43
3.5.1 Definition of the snapshot and its mathematical model . . . 43
3.5.2 Definition of the snapshot matrix . . . 44
3.5.3 Obtaining a snapshot matrix . . . 45
3.5.4 Integration of the snapshot matrix into shape fitting . . . 45
3.5.5 Estimation error . . . 47
3.6 Conclusions . . . 47

**4 Snapshot matrix based techniques for shape fitting 49**
4.1 Estimation using the Proper Orthogonal Modes (POM technique) . 50
4.1.1 Method . . . 50
4.1.2 Comparison between proper orthogonal modes and dynamic modeshapes for shape fitting . . . 52
4.2 Estimation using the full snapshot matrix (LS technique) . . . 54
4.3 Comparison between the POM and the LS technique . . . 54
4.3.1 Influence of the choice of target points . . . 54
4.3.2 Methodical error . . . 56
4.3.3 Error due to sensor noise . . . 56
4.3.4 Least squares solution for least total error . . . 57
4.3.5 Time complexity . . . 57
4.4 Numerical study of the snapshot techniques in the wafer chuck . . 59
4.4.1 Wafer chuck model . . . 59
4.4.2 Load cases . . . 59
4.4.3 Snapshot matrix construction and estimator calculation . . 61
4.4.4 Estimation error . . . 62
4.4.5 Estimation results and discussion . . . 62
4.5 Conclusions . . . 66

**5 Shape fitting in presence of rigid body dynamics 67**
5.1 Shape fitting in presence of position control . . . 67
5.1.1 Black box approach . . . 70
5.1.2 Split approach . . . 70
5.2 Decoupling rigid body and flexible modes . . . 72
5.2.1 The rigid body estimate . . . 72
5.2.2 The flexible residue . . . 73
5.2.3 Shape fitting using the flexible residue . . . 74
5.3 Shape fitting in a dynamic system . . . 75
5.3.1 Magnitude and relative phase of dynamic signals . . . 75
5.3.2 A complex snapshot matrix . . . 77
5.3.3 Estimator calculation using the complex snapshot matrix . 77
5.4 Numerical study of shape fitting in a dynamic wafer chuck . . . 79
5.4.1 Method . . . 79
5.4.2 Results and discussion . . . 81
5.5 Hybrid shape fitting estimation . . . 83
5.6 Conclusions . . . 83

**6 Sensor placement** **85**
6.1 Selection of sensor placement algorithms for identifying shapes . . 85
6.1.1 The goal of sensor placement . . . . . 86
6.1.2 Evaluation time and convergence . . . . . 86
6.1.3 Placement algorithms in literature . . . . . 87
6.1.4 Selected shape-based algorithms . . . . . 87
6.2 Comparison of the sensor placement algorithms . . . . . 91
6.2.1 Performance metrics . . . . . 93
6.2.2 Placement for random shapes . . . . . 94
6.2.3 $(x, y)$-placement of position sensors to identify modeshapes 95
6.2.4 $(x, y, \theta)$-placement of strain sensors to identify modeshapes 98
6.3 Integration of sensor placement in shape fitting . . . . . 106
6.4 Sensor placement applied to the numerical study of the wafer chuck 108
6.5 Conclusions . . . . . 109

**7 Experimental verification of shape fitting** **113**
7.1 Setup design . . . . . 113
7.1.1 Purpose of the setup . . . . . 113
7.1.2 Requirements . . . . . 114
7.1.3 Suspension concepts . . . . . 114
7.1.4 Detailed design and component selection . . . . . 116
7.2 Setup validation . . . . . 122
7.2.1 Suspension stiffness . . . . . 122
7.2.2 Noise performance . . . . . 123
7.2.3 Parasitic effects of the actuator forces and suspension . . . 124
7.3 Verification of the system model . . . . . 125
7.4 Demonstration of the shape fitting method . . . . . 127
7.4.1 Comparison estimators in the time domain . . . . . 127
7.4.2 Comparison estimators in the frequency domain . . . . . 129
7.5 Hybrid estimation using shape fitting and an accelerometer . . . . 130
7.5.1 Method . . . . . 135
7.5.2 Results . . . . . 136
7.6 Conclusions . . . . . 138

**8 Conclusions and outlook** **141**
8.1 Conclusions . . . . . 141
8.2 Final conclusion and future outlook . . . . . 146

**A Wafer chuck scaling properties** **149**

**B Out-of-plane force due to immersion film dynamics** **151**
B.1 Assumptions . . . . . 151
B.2 Force due to viscous damping . . . . . 151
B.2.1 Force due to inertia . . . . . 153
B.3 Comparison to literature . . . . . 153

**C Double integration of an acceleration signal** **155**

**D Proper orthogonal decomposition** **157**

**E Comparison of commercial accelerometers** **161**
E.1 Accelerometers . . . . . 161
E.2 Floor vibration criteria and measurements . . . . . 162

**F Kalman filtering** **165**

**G Selection and assessment of the Minus-K vibration isolator** **167**
G.1 Comparison of vibration isolators . . . . . 167
G.2 Experimental evaluation of the Minus-K vibration isolator . . . . . 167

**H Corrected plate finite element matrices** **171**
H.1 Matrix $\boldsymbol{R}$ . . . . . 171
H.2 Matrix $\boldsymbol{Q}$ . . . . . 171
H.3 Matrix $\boldsymbol{A^{-1}}$ . . . . . 172

**Bibliography** **173**

**Mathematical notation** **181**

**Parameters and variables** **182**

**Acronyms and abbreviations** **185**

**Acknowledgements** **187**

**Curriculum vitae** **189**

# Chapter 1

# Introduction

## 1.1 A 'self-fulfilling' prophecy

Last year was the 60th anniversary of Moore's law [1], which is not so much a law, but rather an extrapolation based on his observation. In 1965, Gordon Moore observed that the complexity of computer chips, in terms of the number of transistors per chip area, had doubled every year. He predicted that this tendency would continue for at least ten years. In 1975, when it turned out that he had been right, he predicted that the complexity increase would fall off, but still double every two years [2]. Again, his prediction came true. Moreover, chip manufacturers and their machine suppliers adopted his forecast as their target. In that sense Moore's observation may be regarded as a self-fulfilling prophecy, but the word 'self-fulfilling' does not fully do justice to reality. Through the years, engineers have constantly been working at the limits of what is possible, constantly extending the edges. In the 60 years, they have been able to increase the number of transistors per chip area by a factor of $10^{10}$. As a result, today's electronics is able to fulfil wishes people could before only dream of, such as wireless communication and high-resolution non-invasive medical imaging. Institutes, companies and consumers benefit from fast, powerful, and efficient computer chips.

## 1.2 Chip production using wafer scanners

Chips are produced in batch on wafers, i.e. round thin silicon substrates of up to 300 mm diameter, as shown in Figure 1.2a. The electronic circuits are built up upon this substrate with a process called photolithography. Layer by layer, different materials are applied onto these wafers. Parts of each layer are etched away, leaving a three-dimensional structure that forms the electronic circuitry. To etch away only the unwanted parts of a material layer, a photoresistive layer is applied that is partly exposed by light during the exposure step. This changes the local properties of the photoresist, allowing the selective removal of the photoresist during an etching step. In this way, the exposure step defines the

1975 2016

Figure 1.1: An illustration of the performance improvement of computer chips. Despite its significantly smaller volume, a current iPhone 6 smartphone contains about a 1000 times more transistors than supercomputer Cray-1 around 1975. Cray-1 could perform $10^8$, the iPhone 6 $10^{11}$ floating point calculations per second [3].

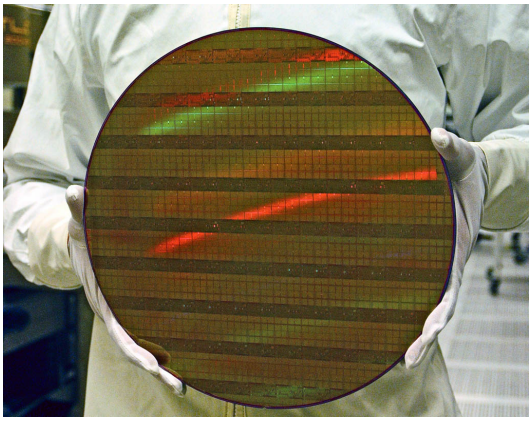
geometry of the layer's electronic patterns. It is a crucial step, as the smaller the electronic details can be made, the more features fit on the chip's area and the more powerful the chip becomes.

The lithography machine that performs the exposure is called the 'wafer scanner' (Figure 1.2b). As the machine performs a difficult task, it is large and highly complex. Figure 1.3 schematically represents the section of the machine where the actual exposure of the wafer takes place. Light from a lens column hits the wafer, which lies on a wafer chuck. The wafer chuck has several functions. It keeps the wafer mechanically and thermally stable and it moves the wafer underneath the lens column, as the lens column has a fixed position and the whole wafer needs to be exposed. The point on the wafer that is exposed at a certain instance in time is called the Point of Interest (PoI). A thin film of immersion fluid is located between the lens column and the exposure area on the wafer. As a glass-fluid transition has a smaller refraction angle than a glass-air transition, immersion systems are able to project smaller details.

Four three Degrees of Freedom (DoFs) optical encoders at the corners of the wafer chuck measure the chuck's position. They measure with respect to a metrology frame that provides a stable displacement reference. From the four sensors at the corners of the chuck the distance between the wafer and the lens column is calculated as well. To do this, a Rigid Body (RB) approach is used, assuming that the wafer chuck does not deform.

## 1.3 Smaller details and a higher throughput

Current wafer scanners use light with a wavelength of 193 nm for the exposure and are able to produce details in the order of 20 nm, which is 10 times smaller than the diffraction limit of the light. To be able to manufacture even smaller features in future, a new machine is currently being developed that uses extreme ultraviolet light for the exposure, effectively decreasing the wavelength of the light with a factor of 14 to 13.5 nm (Figure 1.4a).



(a) A 300 mm diameter wafer. 300 mm is currently the largest substrate size used for manufacturing electronic chips. Source: Intel corporation.

(b) Example of a current wafer scanner. The wafer scanner is performs the exposure of the wafer, thereby defining the geometry of the electronic patterns. Source: ASML.

Figure 1.2: A 300 mm wafer and a wafer scanner.

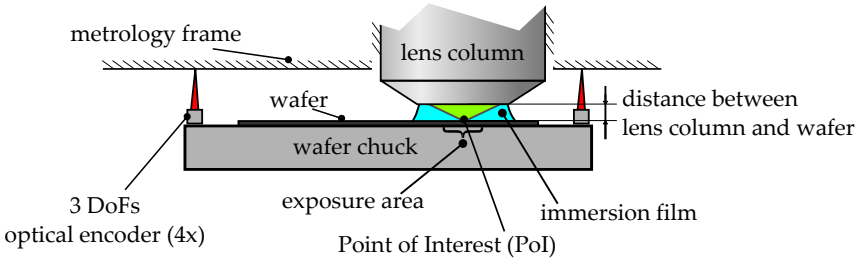



Figure 1.3: The wafer chuck supports the wafer and moves it underneath the lens column. The position of the wafer chuck and the distance between the wafer and lens column are found from the displacement measurements of the optical encoders at the corners of the wafer chuck.

A very important property of a wafer scanner is its throughput, the number of wafers that can be exposed per hour. Due to its extreme specifications a wafer scanner is expensive, but if more chips can be manufactured in the same amount of time, the price of the chips decreases. The throughput of a wafer scanner can, for example, be increased by using larger wafers. Current wafers have a diameter of 300 mm, in future the diameter will increase by a factor of 1.5 to 450 mm (Figure 1.4b). This transition, initiated some years ago in lithography industry [4,5], is currently somewhat simmering, but will at some point in future be a necessary step in the process of improving lithography.

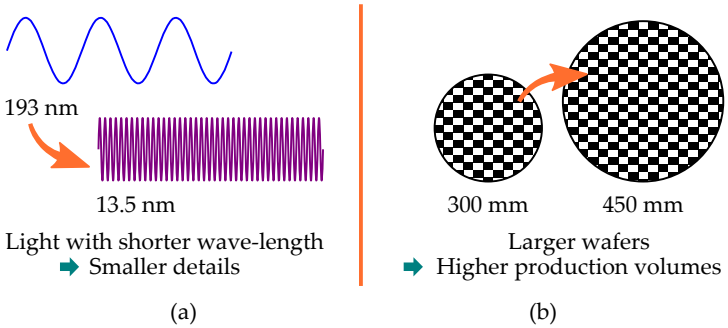



Figure 1.4: Current trends in wafer scanners. New machines are developed for working with extreme ultraviolet light, which makes it possible to project smaller details and thus to build more electronics on a chip's area. On the other hand, to keep the lithography process affordable the wafer size needs at some point to be increased, allowing to manufacture more chips per hour and thus leading to a higher throughput of the wafer scanner.

Still, increasing the wafer diameter is not a trivial step. Not only the wafer but also other parts of the machine need to become larger. The length and width of the wafer chuck, for example, also need to scale up with a factor of 1.5. As for the wafer chuck's thickness, there will be an inherent trade-off between the increase of mass and the increase of deformation.

During the exposure, the wafer chuck deforms due to different disturbance

sources, for example the forces resulting from acoustic effects and the vibrations in the immersion film. The deformations lead to an error between the actual distance of the wafer with respect to the lens column and the value found from the measurements at the corners. Current wafer chucks are relatively thick and stiff to prevent significant deformation. If the stiffness, however, were to be kept constant, the wafer chuck's thickness should be increased [6], leading to larger actuation and reaction forces needed to maintain the current acceleration levels. Those larger actuator forces in turn lead to higher heat loads in the system. Clearly, both the increased heat loads and the increased actuation forces are undesirable in a precision system. Therefore, it would be useful to make the wafer chuck thinner and to accept its lower stiffness as a consequence [7].

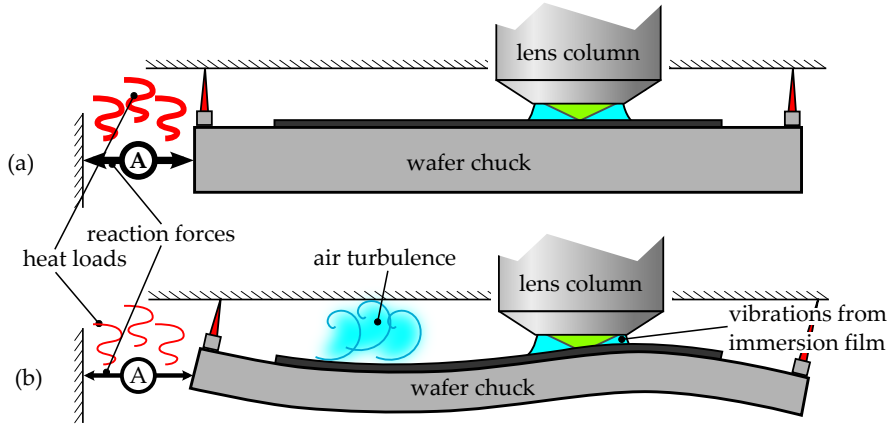



Figure 1.5: If the wafer size is increased to 450 mm, also the size of the wafer chuck should increase. If the stiffness needs to stay equal, the chuck should become thicker, leading to a significantly larger mass and thus larger heat loads and reaction forces (a). Therefore, a thinner chuck would be favourable. But a thinner wafer chuck will also experience larger deformations due to disturbance forces caused by, for example, air turbulence or vibrations in the immersion film (b).

## 1.4 Towards a lighter wafer chuck

If the wafer chuck is made thinner, its deformation increases. The effect is especially large in the out-of-plane direction, where the deformation is primarily caused by bending. As the deformation in out-of-plane direction increases more rapidly than in in-plane-direction for decreasing a decreasing wafer chuck thickness, estimation will be treated using consistently the deformation in the out-of-plane direction.

Currently, an out-of-plane deformation in the order of 5 nm is allowed. In the worst case in which the wafer chuck's geometry is a simple plate and its thickness is chosen such that the chuck's mass stays constant, its first mechanical eigenfrequency decreases by approximately a factor of 8 and its stiffness decreases by a factor of approximately 25. That means that if the current

wafer chuck deforms up to 5 nm, the deformation of the larger one would be up to 120 nm.

If the deformations are such large, the current RB method for finding the distance between the wafer and the lens column at the PoI is not good enough any more. On the other hand, a direct measurement of the distance at the PoI is also difficult because of space limitations in the lens area, the presence of the wafer surface and the non-existence of a stable reference for position measurement close to the lens area.

The solution proposed and explored in this thesis is to place additional sensors – either displacement sensors at other locations than the PoI or sensors measuring other measurands – and *estimate* the deformation at the PoI based on their readings.

## 1.5 Estimating deformation using shape fitting

The physical processes that cause correlations between the measurements and the variables to be estimated make estimation possible. More specifically, for the wafer chuck, the mass and stiffness distribution allow the estimation of static and dynamic deformation. Although estimators could be dynamic, for example based on a Kalman filter, often static estimators are chosen, e.g. [8, 9]. Static estimators make use of the momentary correlation between the measurement quantities at the sensors and the variables to estimate, not taking into account information from the previous measurements.

Static estimators are often chosen intuitively and applied without explicit reasoning. Still, there are legitimate grounds to use them. First of all, if a system behaves statically or quasi-statically, a well-chosen static estimator can lead to equally good results as a well-chosen dynamic estimator. In the wafer chuck case, for example, the low-frequency disturbances are, according to the manufacturer of the systems, predominant; those low-frequency disturbances are well below the resonance frequencies of the chuck, so that the problem may be considered quasi-static or stiffness governed.

Many static estimators in literature make use of a set of basis functions that describe the correlation between the values at the measurement points and the target points. The estimate is then found using a fitting procedure which selects a combination of basis functions that best describes the sensor values. Although this approach is widely used, it does not have a generally accepted name. In the rest of this work, this approach will be referred to as *shape fitting* and the basis functions *fitting shapes*.

Estimation using shape fitting has been studied in many application areas in which a full field measurement of a physical quantity can not be performed. The missing values in the physical field then need to be estimated using a limited number of measurements at other locations or of other measurands. Examples are found in various engineering areas and domains, like estimation of the shape or deformation of bridges [10, 11], pipe lines [12] or antennas [13] or estimation of the thermal profiles of parts in precision machines [14]. Estimation can also be used to generate redundancy for monitoring the condition of the sensors [15].

An estimation problem consists typically of two mutually dependent parts: the selection of the *sensor configuration* and the choice of the *estimation algorithm*. The selection of the sensor configuration consists of all steps of choosing the number of sensors, their locations (i.e. sensor placement) and the type of sensors. For example, not only displacement sensors may be used for measuring a displacement field but also strain sensors, as strain is a spatial derivative of the displacement field. The choice of estimation algorithm consists of all steps of deriving the estimated value from the sensor signals.

Another reason for using shape fitting – next to its suitability, being a static estimation method, to quasi-static problems – is the simplicity of the method. This simplicity has the trivial advantage of being easier to grasp, so that adaptation of the method to specific boundary conditions of the application might be easier, also for non-experts. Another advantage is that intrinsically complicated problems, like the one of sensor placement and the integration with control schemes, might be easier to solve when building upon a relatively simple basis method.

## 1.6 Shape fitting estimation methodology wafer chuck deformation

Shape fitting has often been employed in literature, in different ways and with different levels of refinement. There are, for example, different techniques for obtaining the fitting shapes. Although these techniques have significant differences, they were, to the best of our knowledge, never compared in literature before.

Furthermore, although shape fitting has been applied in high-precision mechatronics before, for example by Koevoets et al. [16], it has not been used in the wafer chuck application yet. The wafer chuck has some specific properties that are important for the way shape fitting might be used, like its RB motion and the controller action that positions the wafer chuck. The properties and the boundary conditions of the wafer chuck also lead to specific requirements to the sensor placement method which need to be addressed.

The goal of this thesis is, therefore, defined as follows:

> The goal of this thesis is to develop an estimation methodology for wafer chuck deformation, based on shape fitting principles, considering the aim for low estimation error and the specific requirements related to the wafer chuck application.

## 1.7 Method and structure of this thesis

In this thesis, several shape fitting methods and sensor placement methods from different application fields and research fields are compared with as a main objective attaining a low estimation error. The methods are adapted further

developed to develop an estimation methodology that is suitable for the wafer chuck application. The adapted methods are implemented on and assessed with an experimental setup to show their applicability to the wafer chuck problem.

Chapter 2 studies the properties and boundary conditions of the wafer chuck that are relevant to the requirements of the estimation and sensor placement algorithms. Based upon the study, a finite element model is selected for further use throughout the thesis and feasible configurations for the sensors that are used for estimation are selected.

Given the properties of the wafer chuck, Chapter 3 selects shape fitting as the estimation method and studies how the method is used in the literature. The so-called 'snapshot matrix' is introduced, which contains the foreknowledge of the mechanics and the typical disturbances that work on the system.

In literature, two methods can be found that can be used to generate a shape fitting estimator. Chapter 4 compares these two method, both in a qualitative and a quantitative way. The numerical results are obtained from wafer chuck case study. The case study was simplified, as the existing shape fitting method cannot directly handle the specific boundary conditions of the wafer chuck related to its RB motion.

Chapter 5 adapts shape fitting method such that it is able to handle the RB motion and the dynamic behaviour of a system. The adapted method is applied to a more realistic numeric model of the wafer chuck. This chapter also proposes a hybrid method that uses shape fitting in combination with acceleration sensing.

Chapter 6 studies sensor placement, the automated choice of the sensor locations. To obtain good estimation results it is essential that the sensors locations are adequate. Several automatic placement algorithms from literature are selected and a procedure is developed for integrating them into the shape fitting framework. Numeric results show a significant improved estimation when utilising the sensor configurations found by this procedure.

To experimentally validate the used numeric methods and to verify the shape fitting method, an experimental setup that mimics the wafer chuck system with its sensors, actuators and controller is developed. Chapter 7 presents the considerations regarding the design of the setup and the estimation results obtained using this setup.

Finally, Chapter 8 presents the conclusions recommendations based on the research in this thesis.

# Chapter 2

# System properties of the wafer chuck

To be able to select an effective method for estimation of the wafer chuck's deformation, it is necessary to be acquainted with the wafer chuck's system properties. The previous chapter briefly introduced the wafer chuck system. This chapter further elaborates on the details of the exposure process and the system's dynamics that are relevant for estimation. Based on this knowledge, the requirements imposed on the estimator are derived.

Section 2.1 provides the details of the exposure process and the motion control system that are relevant to the estimation problem. Section 2.2 introduces the equations of motion that govern the wafer chuck's dynamics and that form the physical model of the system. Based on these equations of motion the concept of dynamic eigenmodes is presented. This concept is useful for understanding the so-called 'quasi-static' behaviour of the wafer chuck system, which proves to be crucial to the development of the estimator. Section 2.3 introduces and experimentally validates the finite element model that is used in the numerical studies in this thesis. Although geometrically simplified, the model shares the same physical principles as the real wafer chuck. Section 2.4 studies different sensor types that can be used for the shape estimation and presents some feasible sensor configurations that fit into the boundary conditions of the wafer chuck. Finally, Section 2.5 concludes the chapter, listing the requirements and boundary conditions derived from the wafer chuck that are imposed on the estimator.

## 2.1 Introduction to the system

### 2.1.1 Exposure path

In lithography, the pattern to be projected onto the wafer is defined by the *reticle*. Thus, the reticle has a function comparable to the slide in a slide projector. The reticle contains a single image, which is projected repeatedly onto the wafer,

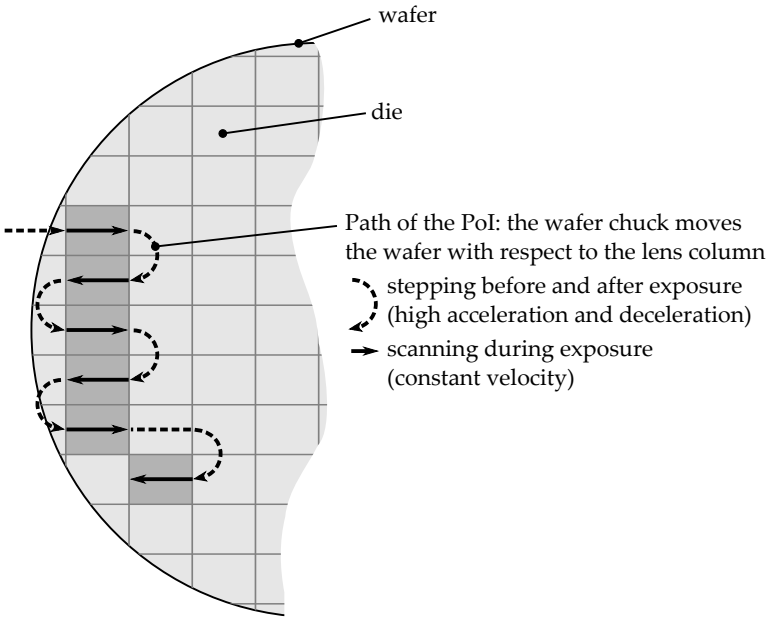



Figure 2.1: The typical path of the PoI on the wafer as a function of time. The wafer chuck moves with respect to the stationary lens column such that this path is formed. One-by-one the dies are exposed during the so-called 'scanning' in which the velocity is kept constant in one direction. In between the exposure of two dies, the wafer is moved as fast as possible to the next die and experiences high acceleration and deceleration levels.

resulting in a grid of this pattern on the wafer, the *dies* (Figure 2.1).

Wafer scanners expose the dies one after the other. This is done in a scanning fashion: The wafer chuck moves the wafer with respect to the fixed lens-column. After a die has been scanned, the wafer steps to the next die and begins with the next scan. In practice, the path of the PoI looks like the one shown in Figure 2.1. The wafer needs to follow an alternating path, such that the reticle, which needs to follow the motion of the image, can slide back and forth.



The stepping part is optimised to take little time, leading to high wafer chuck accelerations. The corresponding high actuation forces lead to deformations and vibrations of the wafer chuck. Right before the starting the scanning motion there is some time for the wafer chuck's vibrations to dampen out. During the exposure, the velocity in the scanning direction is kept constant, whereas the velocities in the other direction are kept zero. In practice, there is a small motion in the out-of-plane direction, such as to follow the height map of the wafer that is obtained before the exposure process [17]. Still, the resulting out-of-plane velocities are relatively small.

### 2.1.2 The position and displacement of the wafer chuck

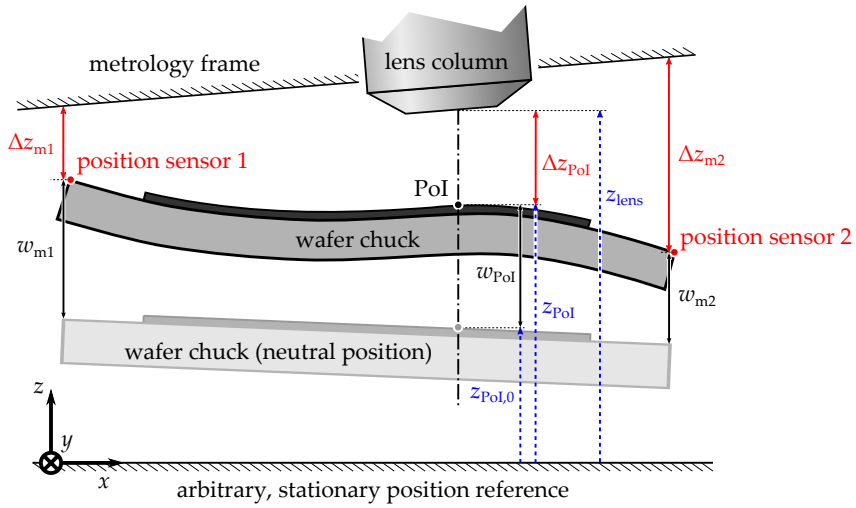
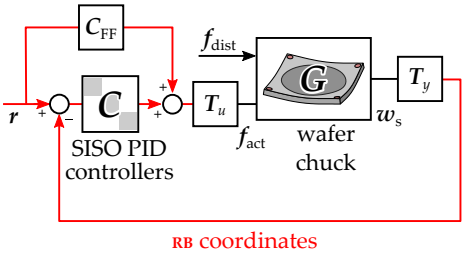
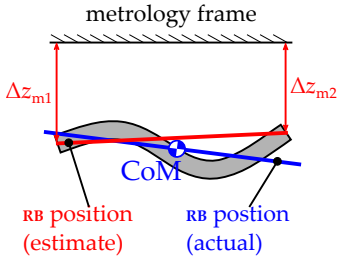



Figure 2.2: The definitions of some absolute and relative positions and displacements of points at the wafer chuck, lens column and metrology frame. Note that the rotation angles, the wafer chuck deformation and the distance between the lens column and the PoI are exaggerated for reasons of clarity.

In immersion lithography machines, the wafer chuck's position is measured in three Degrees of Freedom (DoF) by four optical encoders located at the corners of the wafer chuck. The position of a point can only be measured with respect the position of another point. The encoders measure the position of the wafer chuck's corners with respect to the so-called metrology frame, $\Delta z_{m1}$ to $\Delta z_{m4}$. The metrology frame has a low-stiffness coupling to the world to isolate it from




(a) Basic structure of the wafer chuck's out-of-plane control system.




(b) The RB position is a function the RB coordinates. Those coordinates are estimated from the position measurements at the corners of the chuck with respect to the metrology frame.

Figure 2.3: The wafer chuck's position is controlled using the RB coordinates found from the position sensors at the corners of the chuck.

ground vibrations and has a high internal stiffness to limit its deformation. Next to the wafer chuck's position, also the positions of the parts that need to be stable during the exposure are controlled with respect to the metrology frame. The lens column, for example, is actively kept in position with respect to the metrology frame. It will be assumed that the lens column has no relative motion with respect to the metrology frame.

Figure 2.2 provides an overview of the definitions regarding the out-of-plane positions and displacements of some relevant points at the wafer chuck and its surroundings. The wafer chuck is drawn in an initial, neutral state and in a translated and deformed state. The position of all points on the wafer chuck, in both its states, and on the metrology frame and the lens column can be seen as referenced to the an arbitrary located, stationary reference. Three points on the wafer chuck, corresponding to respectively position sensors 1 and 2 and the PoI, are marked in the figure. As an example, the positions of the PoI points with respect to the stationary reference are indicated.

The displacement field $w$ describes the displacement of the points on the translated and deformed wafer chuck with respect to their neutral position. $w_{m1}$, $w_{m2}$ and $w_{PoI}$ are the respective displacements of the three marked points. $\Delta z_{m1}$ to $\Delta z_{m2}$ are the out-of-plane measurements of two indicated position sensors. The PoI's position with respect to the lens, the quantity that needs to be estimated, is indicated by $\Delta z_{PoI}$.

### 2.1.3 Motion control of the wafer chuck

A control system is used for letting the wafer chuck follow the desired motion in its six DoFs [18]. Figure 2.3a shows the basic structure of the control loop of the out-of-plane direction. The control system consists of a feed-forward action, $\boldsymbol{C}_{FF}$, and a feed-back loop with a controller, $C$. The control system aims at following the position reference $\boldsymbol{r}$ and rejecting the influence of disturbances that act on the wafer chuck, $\boldsymbol{f}_{dist}$.

The chuck's motion is controlled in terms of its Rigid Body (RB) coordinates. In the figure, this is indicated using red lines. Controller $C$ controls the three RB coordinates in a single input/single output (SISO) fashion. Matrix $\boldsymbol{T}_u$ translates the forces in RB coordinates into the actuator forces, $\boldsymbol{f}_{\text{act}}$, that are applied to the wafer chuck, denoted with $G$.

The out-of-plane position, $z_{\text{m}}$, of the wafer chuck's corners with respect to the metrology frame is measured using the sensors at the chuck's corners and converted by matrix $\boldsymbol{T}_y$ into RB coordinates. If the wafer chuck were rigid, the RB coordinates would be exact, i.e. according to the position and rotation of the Centre of Mass (CoM). However, if the wafer chuck is compliant, the RB coordinates should be considered *estimates*, as illustrated in Figure 2.3b.

For the estimation of the wafer chuck's out-of-plane deformation, the feed-forward action is not of interest, as the forces in the out-of-plane direction are kept practically constant during exposure.

### 2.1.4 Immersion layer

In immersion lithography a film of immersion fluid is applied between the lens column and the wafer. This immersion film improves the performance of the optical system by increasing the so-called 'numerical aperture' of the lens [19, 20]. As the light leaving the lens does not enter the air but enters a liquid, it experiences a lower angle of refraction. This in turn means that the higher refractions orders can be imaged without increasing the diameter of the lens, thus improving the imaging resolution.

Next to the optical improvement, the use of an immersion film also brings challenges. It is, for example, necessary to keep the immersion fluid in place during the scanning motion. This is taken into account for by the immersion hood, a large ring around the lens. The immersion hood contains channels for active supply and drain off of the immersion fluid. The immersion film is kept in place by the so-called 'air-knife', a gas is blown out around the immersion film [21]. The fluid and air flow causes disturbance forces to act between the lens column and wafer chuck.

Furthermore, the fluid film adds boundary conditions to the free floating wafer chuck. From Appendix B it follows that the immersion fluid acts as a squeeze film, causing a force between the lens column and the wafer chuck that is composed of parts proportional to their relative velocity, acceleration and velocity squared, i.e.

$$F_z = -\pi R^4 \left( \frac{3\mu \dot{h}}{2h_0^3} + \frac{3\rho \ddot{h}}{20 h_0} - \frac{15\rho \dot{h}^2}{56 h_0^2} \right), \tag{2.1}$$

where $h(t)$ is the height of the squeeze film, equal to the distance between the lens surface and the wafer surface. The formula assumes that $h(t)$ stays close to the nominal film height, $h_0$. With an assumed lens radius of $R = 0.050$ m and a nominal nominal film height of $h_0 = 0.3 \cdot 10^{-3}$ m and assuming water as immersion fluid with a dynamic viscosity that equals $\mu = 10^{-3}$ Pa $\cdot$ s at 20 °C

and a density of $\rho = 1 \cdot 10^3$ kg/m$^3$, the formula can be filled:

$$F_z = -\left(1.1 \cdot 10^3\right)\dot{h} - 10\ddot{h} + \left(58 \cdot 10^3\right)\dot{h}^2. \tag{2.2}$$



If the lens column is assumed to be at stand-still, the first term adds damping to the wafer chuck at the PoI, with a damping constant $c = 1.1 \cdot 10^3$ Ns/m. The second then adds inertia with a mass $m = 10$ kg. The third term is a non-linear term but is not significant if the amplitude of $h(t)$ is in the order of hundreds of nanometres or lower.

Damping is important, as it leads to coupling of the eigenmodes of a system, depending on how large the damping forces are compared to the inertia and the stiffness forces. For the modelling of the wafer chuck in the rest of this thesis it was assumed that the damping is negligible. The contribution of the inertia is significant compared to the wafer chuck's mass. For simplification, this effect was not taken into account in the models of this thesis, but integrating it into more refined models of the wafer chuck is relatively simple.

### 2.1.5 Foreknowledge on the disturbances

As the response of a system is determined by the loads that work on the system, knowledge of these loads is important for estimating the system's response. Part of the loads are known as function of time so that the resulting response may be predicted if a good model is available. The actuation forces calculated by the feed-forward and the feed-back controller are known and, as such, their resulting deformations can be calculated.

The disturbance forces are, on the other hand, not known as function of time. Still, in many cases foreknowledge about the typical location or area where they work, their typical magnitude and their typical frequency content is available. Such foreknowledge makes estimation possible.

Examples of disturbances that work on the wafer chuck in immersion machines include the forces due to the cables that connect to the wafer chuck, the forces due to noise on the actuator signals and the forces induced by the flow in the immersion film. The disturbance sources are extensively studied by wafer scanner manufacturers in the context of error budgeting. This information includes typical location, magnitude and frequency region of the disturbance. The relevant foreknowledge can be used when constructing an estimation algorithm.

The major disturbance source, causing the largest deformations of the wafer chuck are the forces induced by the immersion film [22]. The immersion film disturbances are mainly low-frequency, typically well-below the lowest resonance frequency of the wafer chuck. Being the major disturbance source, the immersion film forces will be used in the numerical analyses throughout this thesis.

The wafer chuck's deformation is not only caused by mechanical disturbances, i.e. forces and moments, but also by thermal disturbances. Heat loads lead to thermal gradients in the wafer chuck, which in turn lead to deformation. As the processes in the thermal domain are typically slow in comparison to

the processes in the mechanical domain, the thermal domain will not excite resonances in the mechanical domain. Therefore, a thermal distribution contributes one-to-one to the deformation in the mechanical domain. Estimation of the deformation due to thermal loads can, therefore, be based on a combination of thermal and mechanical sensors.



## 2.2 Mechanical model and dynamic eigenmodes

This section studies the wafer chuck's mechanics. The concept of the equations of motion and their dynamic eigenmodes and modeshapes is introduced. The concept of eigenmodes proves important for understanding the quasi-static behaviour, which is relevant for the estimator selection in Chapter 3. Furthermore, the equations of motion provide the physical framework of the model that is developed in the next section.

### 2.2.1 Equations of motion in matrix form

The equations of motion for a general linear solid mechanical system with damping can be written in matrix-form as

$$\boldsymbol{M}\ddot{\boldsymbol{w}} + \boldsymbol{C}\dot{\boldsymbol{w}} + \boldsymbol{K}\boldsymbol{w} = \boldsymbol{f}, \tag{2.3}$$

where $\boldsymbol{M}$, $\boldsymbol{C}$ and $\boldsymbol{K}$ are respectively the mass, damping and stiffness matrix, $\boldsymbol{w}$ is a vector describing the generalised displacement field in terms of the nodal coordinates and vector $\boldsymbol{f}$ describes the generalised external forces applied to the system. This equation can be rewritten in the well-known state space form using a set of first order differential equations as

$$\frac{d}{dt}\begin{Bmatrix} \boldsymbol{w} \\ \dot{\boldsymbol{w}} \end{Bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\boldsymbol{M}^{-1}\boldsymbol{K} & -\boldsymbol{M}^{-1}\boldsymbol{C} \end{bmatrix}}_{\boldsymbol{A}} \begin{Bmatrix} \boldsymbol{w} \\ \dot{\boldsymbol{w}} \end{Bmatrix} + \underbrace{\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{M}^{-1} \end{bmatrix}}_{\boldsymbol{B}} \boldsymbol{f}. \tag{2.4}$$

The equivalent to Eq. (2.3) in the frequency domain is

$$-\boldsymbol{M}\boldsymbol{w}'\Omega^2 + \boldsymbol{C}\boldsymbol{w}' j\Omega + \boldsymbol{K}\boldsymbol{w}' = \boldsymbol{f}', \tag{2.5}$$

where the frequency variable $\Omega$ is a capital to better distinguish it from the displacement $w$. $\boldsymbol{w}'$ and $\boldsymbol{f}'$ are respectively the displacements and forces in the frequency domain.

### 2.2.2 Definition modeshapes and eigenfrequencies

At its resonance frequencies, a system's internal forces from inertia, damping and stiffness are in perfect balance at all points of the system. If enough energy is added via external forces and the system's damping is low, relatively large displacement amplitudes can be reached. The displacement field $\boldsymbol{w} = \breve{\boldsymbol{\varphi}}_i$ for which resonance occurs is called a modeshape, the corresponding frequency the

eigenfrequency $\Omega = \Omega_i$. Filling in these definitions in Eq. (2.5) and taking $\boldsymbol{f} = 0$ for internal equilibrium the following equations are obtained

$$-\Omega_i^2 \boldsymbol{M}\breve{\boldsymbol{\varphi}}_i + j\Omega_i \boldsymbol{C}\breve{\boldsymbol{\varphi}}_i + \boldsymbol{K}\breve{\boldsymbol{\varphi}}_i = 0, \tag{2.6}$$

which can be rewritten in the form of an eigenvalue problem

$$\left(\boldsymbol{K} + j\Omega_i \boldsymbol{C} - \Omega_i^2 \boldsymbol{M}\right)\breve{\boldsymbol{\varphi}}_i = \boldsymbol{0}. \tag{2.7}$$

Solving the eigenvalue problem yields the modeshapes and the eigenfrequencies. The scaling of the modeshapes is arbitrary. In a system without damping, the modeshapes and the eigenfrequencies are real, in a system with damping they are complex [23].

As shown in Section 2.1.4, the immersion film adds damping to the wafer chuck in the order of $1.1 \cdot 10^3$ Ns/m, depending on the height of the immersion film. Whether this can be considered significant or not depends on the mass and stiffness of the wafer chuck.

### 2.2.3 Modal coordinates

If the damping in the system is not significant the modeshapes do not transfer energy to each other. This leads to the concept of modal decoupling and modal coordinates, which is not only helpful for improved insight into the dynamics of the system, but also useful in the context of modal order reduction.

It is known that any shape of the displacement field can be described as a linear combination of the system's modeshapes [24]:

$$\boldsymbol{w} = \sum_{n=1}^{N} \breve{\boldsymbol{\varphi}}_n q_n = \breve{\boldsymbol{\Phi}}\boldsymbol{q}, \tag{2.8}$$

where $\breve{\boldsymbol{\Phi}}$ is the matrix with the modeshapes, i.e. $\breve{\boldsymbol{\Phi}} = \begin{bmatrix} \breve{\boldsymbol{\varphi}}_1 & \cdots & \breve{\boldsymbol{\varphi}}_N \end{bmatrix}$, and $\boldsymbol{q}$ is a weighing vector with the so-called *modal coordinates*. By introducing this expression into the system's equations of motion (Eq. 2.3) without the damping term, the equations of motion can be expressed in the modal coordinates,

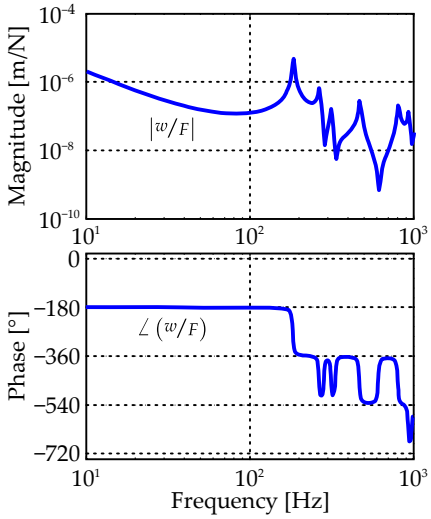
$$\boldsymbol{M}\breve{\boldsymbol{\Phi}}\ddot{\boldsymbol{q}} + \boldsymbol{K}\breve{\boldsymbol{\Phi}}\boldsymbol{q} = \boldsymbol{f}. \tag{2.9}$$

Premultiplying with $\breve{\boldsymbol{\Phi}}^{\mathsf{T}}$ results in

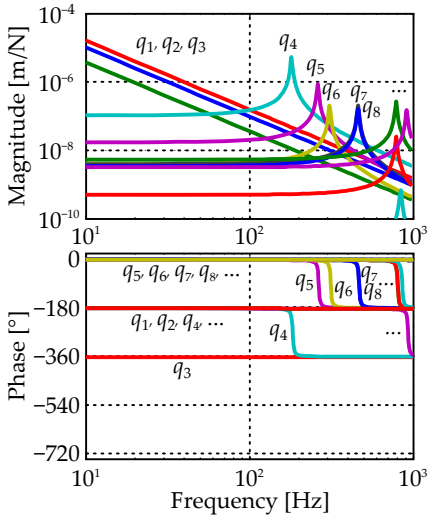
$$\breve{\boldsymbol{\Phi}}^{\mathsf{T}}\boldsymbol{M}\breve{\boldsymbol{\Phi}}\ddot{\boldsymbol{q}} + \breve{\boldsymbol{\Phi}}^{\mathsf{T}}\boldsymbol{K}\breve{\boldsymbol{\Phi}}\boldsymbol{q} = \breve{\boldsymbol{\Phi}}^{\mathsf{T}}\boldsymbol{f}. \tag{2.10}$$

Now the properties of $\breve{\boldsymbol{\Phi}}^{\mathsf{T}}\boldsymbol{M}\breve{\boldsymbol{\Phi}}$ and $\breve{\boldsymbol{\Phi}}^{\mathsf{T}}\boldsymbol{K}\breve{\boldsymbol{\Phi}}$ for a system without significant damping can be used. Modeshapes are known to be orthogonal with respect to the mass and stiffness matrix, so that $\breve{\boldsymbol{\varphi}}_i^{\mathsf{T}}\boldsymbol{M}\breve{\boldsymbol{\varphi}}_j = 0$ and $\breve{\boldsymbol{\varphi}}_i^{\mathsf{T}}\boldsymbol{K}\breve{\boldsymbol{\varphi}}_j = 0$ for any $i \neq j$ [24]. Only $\breve{\boldsymbol{\varphi}}_i^{\mathsf{T}}\boldsymbol{M}\breve{\boldsymbol{\varphi}}_i$ and $\breve{\boldsymbol{\varphi}}_i^{\mathsf{T}}\boldsymbol{K}\breve{\boldsymbol{\varphi}}_i$ are non-zero, so that

$$\boldsymbol{\mathcal{M}} = \breve{\boldsymbol{\Phi}}^{\mathsf{T}}\boldsymbol{M}\breve{\boldsymbol{\Phi}} = \begin{bmatrix} \mathcal{M}_1 & & 0 \\ & \ddots & \\ 0 & & \mathcal{M}_N \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mathcal{K}} = \breve{\boldsymbol{\Phi}}^{\mathsf{T}}\boldsymbol{K}\breve{\boldsymbol{\Phi}} = \begin{bmatrix} \mathcal{K}_1 & & 0 \\ & \ddots & \\ 0 & & \mathcal{K}_N \end{bmatrix}. \tag{2.11}$$




(a) Bode plot of the transfer function from a force $F$ on a point to a displacement $w$ on another point on a plate.




(b) Bode plot of the transfer function in terms of the contributions of the system's eigenmodes.

Figure 2.4: The transfer function of a mechanical system with low damping is build up from the contributions of its eigenmodes, which behave as independent second order systems.

Note that both matrices are diagonal matrices with respectively the so called *modal mass* of the eigenmodes ($\mathcal{M}_i = \breve{\boldsymbol{\varphi}}_i^\mathsf{T} \boldsymbol{M} \breve{\boldsymbol{\varphi}}_i$) and *modal stiffness* ($\mathcal{K}_i = \breve{\boldsymbol{\varphi}}_i^\mathsf{T} \boldsymbol{K} \breve{\boldsymbol{\varphi}}_i$) at its diagonals. The modal mass and modal stiffness depend on the scaling of the modeshapes. Equation (2.10) can thus be rewritten as

$$\mathcal{M}_i \ddot{q}_i + \mathcal{K}_i q_i = \breve{\boldsymbol{\varphi}}_i^\mathsf{T} \boldsymbol{f} \quad i = 1, \ldots, N \tag{2.12}$$

This shows that equations of motion (Eq. 2.3) can be decoupled into a set of independent secondary differential equations using the modal coordinates. Thus, the response of the total mechanical system with insignificant damping, can be simply considered as a combination of separate single mass-spring-systems. In state-space form this can be written as

$$\frac{d}{dt}\begin{Bmatrix} \boldsymbol{q} \\ \dot{\boldsymbol{q}} \end{Bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\boldsymbol{\Omega}^2 & \boldsymbol{0} \end{bmatrix}}_{\boldsymbol{A}_q} \begin{Bmatrix} \boldsymbol{q} \\ \dot{\boldsymbol{q}} \end{Bmatrix} + \underbrace{\begin{bmatrix} 0 \\ \boldsymbol{\mathcal{M}}^{-1} \breve{\boldsymbol{\Phi}}^\mathsf{T} \end{bmatrix}}_{\boldsymbol{B}_q} \boldsymbol{f} \tag{2.13}$$

$$\boldsymbol{w} = \underbrace{\begin{bmatrix} \breve{\boldsymbol{\Phi}} & \boldsymbol{0} \end{bmatrix}}_{\boldsymbol{C}_q} \begin{Bmatrix} \boldsymbol{q} \\ \dot{\boldsymbol{q}} \end{Bmatrix}, \tag{2.14}$$

where $\boldsymbol{\Omega}^2 = \boldsymbol{\mathcal{M}}^{-1} \boldsymbol{\mathcal{K}}$. Figure 2.4 shows the total transfer of a mechanical system and the second order system responses it is build up from.

In the frequency domain, Eqs. (2.8) and (2.12) rewrite to

$$\boldsymbol{w}'(\Omega) = \sum_{n=1}^{N} \frac{\breve{\boldsymbol{\varphi}}_n}{(\Omega_n^2 - \Omega^2)\mathcal{M}_n} \breve{\boldsymbol{\varphi}}_n^{\mathsf{T}} \boldsymbol{f}'(\Omega). \tag{2.15}$$



**Modal truncation** Modal decoupling plays a role in the context of model order reduction. It can be beneficial to reduce the order of a large model, for example to reduce computation time. A method for model order reduction is modal truncation. In this method the equations of motion are translated into their modal form. Then a selection of most important modes is made that is considered enough to accurately describe the system's behaviour, the other eigenmodes are removed [25]. Normally, the eigenmodes corresponding to the higher eigenfrequencies are removed, as they contribute in general less to the system's response. In this case Eq. (2.15) reduces to

$$\boldsymbol{W}'(\Omega) \simeq \sum_{n=1}^{N_{\mathrm{T}}} \frac{\breve{\boldsymbol{\varphi}}_n}{(\Omega_n^2 - \Omega^2)\mathcal{M}_n} \breve{\boldsymbol{\varphi}}_n^{\mathsf{T}} \boldsymbol{f}'(\Omega). \tag{2.16}$$

with $N_{\mathrm{T}} < N$.

## 2.3 Finite element modelling

This section introduces the Finite Element (FE) model that will be used throughout this thesis.

### 2.3.1 Plate finite element

The geometry of a wafer chucks is highly complex. For example, it contains features for clamping the wafer in a highly predictable and stable way and it contains cooling channels for circulating cooling liquid. Furthermore, the geometry has been optimised for good control properties, like low mass and high stiffness, for example for the suspension of the position sensors.

To avoid unnecessary complexity of the model, the wafer chuck is modelled as a plate. As the wafer adds relatively low mass and stiffness to the wafer chuck it is not considered in the dynamic model of the chuck. These simplifications, however, do not change the essence of the problem's physical principles, as the governing equations (Eq. 2.3) are equal. Only the actual geometry and mass and stiffness distribution, captured by the mass matrix, $\boldsymbol{M}$, and the stiffness matrix, $\boldsymbol{K}$, are different.

Ideally, one would want an analytic model for numerical analysis and experiments. With such a model problems like the choice of the spatial discretisation can be avoided and the calculations might be more computationally efficient. However, an analytic model for a plate is less trivial than for, for example, a beam. Hurlebaus and Gaul [26] state that it is impossible to find a good, closed-form analytic model for the plate with all its four sides free (FFFF). They

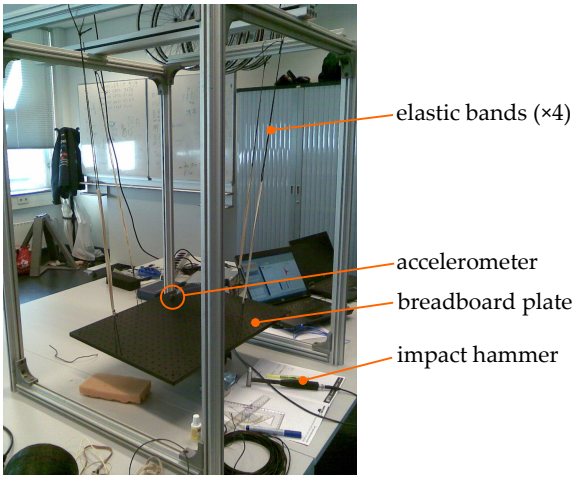



Figure 2.5: A picture of the measurement setup used to validate the FE model. A breadboard plate is suspended by four elastic bands. Using an impact hammer, the FRFs from the impacts in a grid of points to an accelerometer were obtained.

presented an exact solution for the plate's eigenfrequencies, but in the form of a series solution. Therefore, it was chosen to make use of a plate FE model.

Different plate finite elements can be found in literature, for example the one of De Abreu et al. [27] and the one of Charbonneau and Lakis [28] and Kerboua et al. [29]. In the numerical studies of this thesis the latter was utilised. This element allows modelling of the mechanics of rectangular plates, both out-of-plane and in-plane. The finite elements are rectangular and have four nodes with each six degrees of freedom (DoFs), $u$, $v$, $w$, $\partial w/\partial x$, $\partial w/\partial y$ and $\partial^2 w/\partial x \partial y$. The out-of-plane displacement field $w(x, y, t)$ consists of the polynomial terms in $x$ and $y$ up to $x^3 y^3$. These terms are weighed according to linear combinations of the values of the out-of-plane nodal DoFs. The terms of the in-plane displacement field are only weighed according to the in-plane nodal DoFs. In other words, there is no coupling between the in-plane and out-of-plane coordinates of the element, so that the in-plane part of the element can be simply omitted.

Using the definition of the displacement field as function of the nodal DoFs, the mass and stiffness matrices of the plate finite element can be found. This process was fully carried out for this thesis, as the relevant matrices of the intermediate substeps in [29] were found to contain errors. The corrections to these matrices are provided in Appendix H. The element and stiffness matrices were then used to construct the full FE model. Converting the plate model to a modal representation according to Eqs. (2.13) and (2.14) proved to be advantageous, as this allows for modal truncation but also arbitrary placement of actuators.

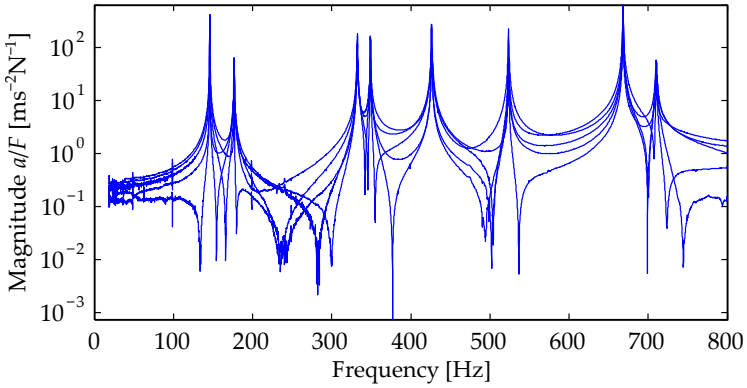



Figure 2.6: Magnitude plot of the FRFs of the plate's response when excited at different locations with an impact hammer and measured with an accelerometer at a single locations. For clarity, only 10 out of the 39 FRFs are shown. Based on the FRFs, the modeshapes and eigenfrequencies are obtained.

### 2.3.2 Experimental validation

To validate the implementation of the FE model, a set of eigenmodes corresponding to the lowest eigenfrequencies were compared to results from an experiment. The experimental setup (Figure 2.5) consisted of a 0.60 m $\times$ 0.45 m $\times$ 12.7 mm breadboard plate (Thorlabs), hanging horizontally and supported by four elastic bands. The elastic bands had a low total stiffness $3 \cdot 10^4$ N/m and were connected at the centres of the plate's sides for minimising the influence of the added stiffness on the low stiffness eigenmodes, i.e. the ones corresponding to the low eigenfrequencies. The plate was hit at a grid of $8 \times 5$ points using an impact hammer. The response of the plate was measured at a single point using an accelerometer. The force signals of the impact hammer and the acceleration signals of the accelerometer were recorded and used for obtaining a set of Frequency Response Functions (FRFs) from impact points to the measurement point (Figure 2.6).

The measured FRFs were used for obtaining the eigenfrequencies and the modeshapes of the plate.[1] The eigenfrequencies were obtained from the measured FRFs by selecting the frequencies at which the sum of the magnitudes of all 39 FRFs attains a maximum. For each in this way obtained eigenfrequency, the values of the FRFs at the eigenfrequency were selected. These values together form a shape which should strictly speaking be called a 'operational deflection shape' [32], as it contains not only contributions of the excited modeshapes itself, but also those of other modeshapes. As the eigenfrequencies of the modes are not closely spaced and the damping is expected to be low, however, the operation deflection shapes at the the resonance frequencies are most probably almost

[1] Initially, the frequency domain parameter estimation method presented by Van der Auweraer et al. [30] and Verboven et al. [31] was used for this purpose. This method, however, involved the use of a weighing function. The choice of the weighing function is not trivial and might affect the results of the comparison. For this reason, the results are not included although they turned out to be very similar to the ones shown in Table 2.1.

equal to the individual modeshapes.

Next to the experimental ones, also a set of eigenfrequencies and modeshapes was found using the FE model. The plate's relevant geometric and material parameters were adopted from the plate's datasheet. As the plate finite element does not support the modelling of holes, the grid of threaded M6 holes were not modelled geometrically, but the Young's modulus $E$ and the density $\rho$ of the plate were uniformly reduced proportional to the volume of the holes.



Table 2.1 compares the first eight modeshapes and eigenfrequencies as they were obtained with the experiment and the FE model. The similarity of the modeshapes of the experiment and the FE model are compared in terms of the Modal Assurance Criterion (MAC) [33]. The MAC value is defined as

$$MAC_{i,j} = \frac{\left|\breve{\boldsymbol{\varphi}}_i^{\mathsf{T}} \bar{\breve{\boldsymbol{\varphi}}}_j\right|^2}{\breve{\boldsymbol{\varphi}}_i^{\mathsf{T}} \bar{\breve{\boldsymbol{\varphi}}}_i \breve{\boldsymbol{\varphi}}_j^{\mathsf{T}} \bar{\breve{\boldsymbol{\varphi}}}_j}, \tag{2.17}$$

where $\breve{\boldsymbol{\varphi}}_i$ and $\breve{\boldsymbol{\varphi}}_j$ are the modeshapes of respectively the experiment and the model. All MAC values are close to 1, which means that the experimental and FE model modeshapes correspond well. The eigenfrequencies, as fitted from the measurements, are close to eigenfrequencies of the model, showing deviations as low as 2 %.

## 2.4 Feasible sensor configurations for estimation

This section studies the possibilities to place additional sensors for use in the estimation of deformation. The study includes both the sensor type but also the locations where sensor can be placed. The section concludes with a feasible sensor configuration based upon position sensors. Optimisation of the sensor locations is covered in Chapter 6.

### 2.4.1 Measurands and sensor types

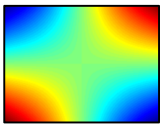
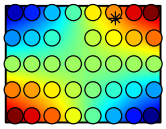
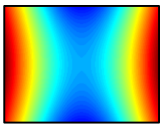
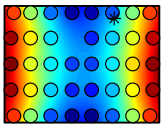
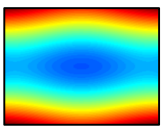
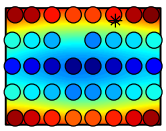
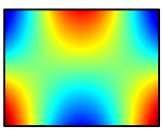
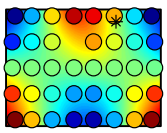
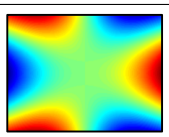
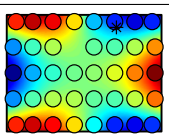
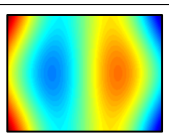
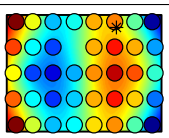
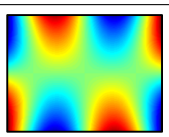
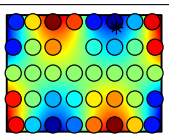
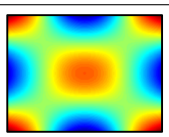
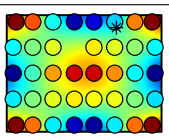
**Position with respect to an external reference**

The position of a point can only be measured with respect to the position of another point. That other point, the measurement reference, may be either located at the same object (an internal reference) or at another object (an external reference).

In current immersion lithography machines, the position of the four corners of the wafer chuck is measured with respect to encoder plates mounted to the metrology frame [34]. This allows for four position measurements in out-of-plane direction. From those the three estimate out-of-plane RB coordinates can be calculated and one deformation DoF. When one of the four position sensors is covered by the immersion hood only the estimate RB coordinates can be found.

Next to the four existing ones extra position sensors could be added, each adding one deformation DoF after decoupling the estimate RB position. Depending on the requirements of the deformation measurement, the resolution

Table 2.1: The modeshapes and eigenfrequencies as obtained with the FE model and the experiment. The experimental modeshapes were determined at the impact points, indicated by the coloured circles ('○'). In between the circles, the experimental modeshapes were interpolated for visual reasons. '∗' indicates the accelerometer point. Having MAC values close to 1, the modeshapes of the model and experiment have almost equal shapes. The eigenfrequencies match closely, deviating less than 2 %.

| Modeshapes | | | Eigenfrequency | | |
|---|---|---|---|---|---|
| from FE model | from experiment | MAC value | from experiment [Hz] | from FE model [Hz] | percentage deviation |
|  |  | 0.99 | 147.1 | 147.8 | −0.47 |
|  |  | 1.00 | 177.8 | 179.0 | −0.67 |
|  |  | 0.99 | 333.3 | 337.9 | −1.4 |
|  |  | 0.99 | 349.8 | 352.7 | −0.83 |
|  |  | 0.99 | 427.0 | 430.5 | −0.82 |
|  |  | 0.99 | 524.0 | 532.6 | −1.6 |
|  |  | 0.98 | 668.4 | 678.1 | −1.4 |
|  |  | 0.99 | 710.4 | 722.4 | −1.7 |

of the additional sensors need not necessarily equal the one of the existing position sensors. Their standoff and range need, of course, to be comparable to the existing ones. The additional sensor can not be placed in the wafer area. Moreover, as the exposure lens and the immersion hood block the view of the metrology frame, the position sensors for the out-of-plane direction must be placed as far to the sides of the wafer chuck as possible.



**Position with respect to an internal reference**

**Strain state** Strain is the ratio between the elongation of an infinitesimal small length in the material due to a load and its unloaded length. The out-of-plane displacement of the points on a solid with respect to their neutral position in the local coordinate frame $x$,$y$,$z$ is described by the displacement field $w(x,y)$. From the displacement field, one-to-one the strain distribution due to bending can be found. Therefore, strain measurements can be used for estimation of the deformation field of the wafer chuck, i.e. the part of the displacement field that is not caused by RB motion.

The in-plane strain of a plate is proportional to the second order spatial derivatives of the displacement field, $w$, and can be described as follows [6]:

$$\varepsilon_x(x,y,z) = -z\frac{\partial^2 w}{\partial x^2}, \quad \varepsilon_y(x,y,z) = -z\frac{\partial^2 w}{\partial y^2}, \quad \varepsilon_{xy}(x,y,z) = -z\frac{\partial^2 w}{\partial x \partial y}, \tag{2.18}$$

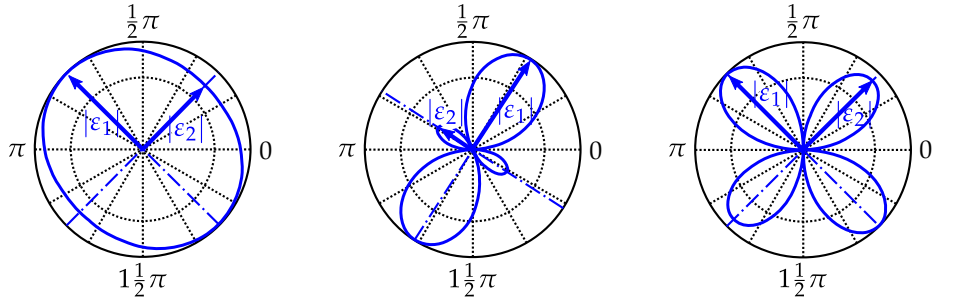
where $x$, $y$, and $z$ are in the plate's local coordinate system. $\varepsilon_x$ and $\varepsilon_y$ are the pure strains in $x$ and $y$-direction and $\varepsilon_{xy}$ is the shear strain, which relates to the engineering strain, $\gamma_{xy}$, as $\varepsilon_{xy} = \frac{1}{2}\gamma_{xy}$ [35]. The strain is proportional to the $z$-coordinate, meaning that the highest in-plane strains occur at the top and bottom surface of a plate, with $z = \pm h/2$. From the strains of Eq. (2.18) the surface strain under an angle $\theta$ with respect to the $x$-axis can be found as follows [36]:

$$\varepsilon(x,y,\theta) = \varepsilon_x \cos^2\theta + \varepsilon_y \sin^2\theta + 2\varepsilon_{xy}\sin\theta\cos\theta. \tag{2.19}$$

This formula shows that the strain is not only a function of the location $(x,y,z)$ of a sensor, but also of its direction $\theta$.

**Direction dependency of strain sensors** Different types of strain sensors include piezoelectric strain sensors, resistive strain gauges and Fibre Bragg Grating (FBG) strain sensors. Depending on the sensor's measurement principle and geometry, a strain sensor has to a greater or lesser extent a preference direction.

A piezoelectric strain sensor in general consist of a thin piece of piezoelectric material with two electrodes at its flat sides. The sensor is fixed with one of the flat surfaces to the measurement target. If the poling of the piezo material is in $z$-direction, elongation in either direction causes an increased voltage over the electrodes [37]. Thus, the shape of the sensor determines the direction dependency. Rectangular sensors, for example, are more direction-dependent than circular sensors.




(a) Example of $\varepsilon(\theta)$ at an umbrella point.

(b) Example of $\varepsilon(\theta)$ at a saddle point.

(c) Example of $\varepsilon(\theta)$ at a saddle point.

Figure 2.7: Three examples of the surface strain, $\varepsilon(\theta)$, as function of the measurement direction. Indicated are the directions corresponding to $\varepsilon_1$ and $\varepsilon_2$. Depending on the relative sign of $\varepsilon_1$ and $\varepsilon_2$, the point with these strains is called a saddle point or an umbrella point. The strain scale is not indicated in the plots.

Resistive strain gauges measure the increase in resistance of wires that elongate due to the strain. As the resistance wires are normally placed in parallel, mainly the strain in the direction of the wires is measured.

A Fibre Bragg Grating (FBG) sensor also measures strain directionally. A FBG is a grating that is manufactured in a glass fibre cable. The grating reflects part of the light that is coupled into the fibre, while transmitting the remainder. The frequency band that is reflected depends on the period of the grating. If the grating is elongated, its pitch increases and this effect can be measured [24, 38]. FBG sensors have often been used for deformation estimation using strain, e.g. [8, 39, 40]. FBGs are short structures, which limits the possibility to integrate strain over longer distances. Alternatively, Fibre Fabry-Pérot (FFP) interferometers may be used. FFP interferometers consist of two FBGs in the fibre, spaced apart at some distance. As the FBGs act as semi-transparent mirrors, light of a certain frequency resonates between the FBGs. As the distance between the FBGs changes when the fibre is elongated, also the resonance frequency changes, which can be measured. Habel et al. show a FFP interferometer with a length of 12 mm [41].

**Strain magnitude** To find the order of magnitude of the strains that need to be measured, the maximum strain of each of the modeshapes is calculated. It is assumed that the sensor is directional, so that the maximum strain not only depends on the position, $(x, y)$, but also on the direction, $\theta$.

For any position $(x, y)$, $\varepsilon(\theta)$ has two optima, $\varepsilon_1$ and $\varepsilon_2$, on $0 \leq \theta < \pi$ – except for the trivial case in which $\partial\varepsilon/\partial\theta$ is constant. The signs of $\varepsilon_1$ and $\varepsilon_2$ may be different or equal. Figure 2.7 provides three examples of the strain as function of the direction, $\theta$. Figure 2.7a corresponds to an 'umbrella point', for which $\varepsilon_1$ and $\varepsilon_2$ have an equal sign. Figures 2.7b and 2.7c correspond to 'saddle points', for which $\varepsilon_1$ and $\varepsilon_2$ have opposite signs.

Figure 2.8 shows at the left the displacement fields, $w(x, y)$, of a plate's first six eigenmodes. At the right the corresponding maximum absolute value of the

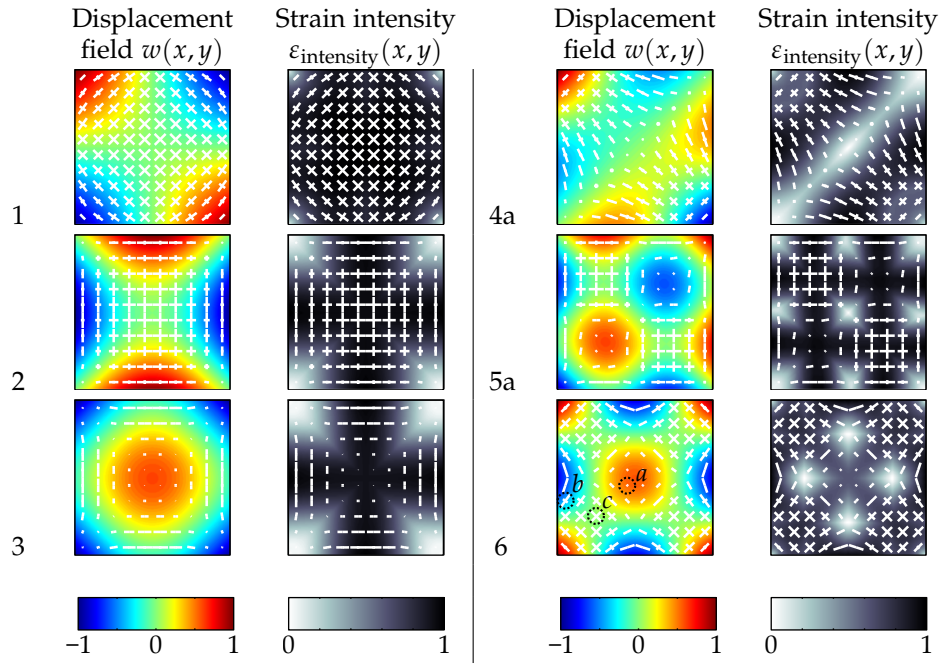



Figure 2.8: The displacement field and corresponding maximum absolute value of the strain intensity $\varepsilon_{\text{intensity}}(x,y)$ at the top surface of a plate for its first six flexible eigenmodes. The stripes point into the directions of maximum strain; the length of the stripes is proportional to the range of the strain in that points, $\varepsilon_{\text{range}}(x,y)$, indicating the sensitivity to the measurement direction. The letters $a$, $b$ and $c$ in the sixth mode correspond to the strain plots of Figure 2.7.

Table 2.2: Maximum surface strain in the modeshapes of plate for a normalised out-of-plane displacement of 1 nm. The plate is made of aluminium and has dimensions of 0.55 m × 0.55 m × 0.032 m.

| Eigenmode | Frequency [Hz] | Displacement [$10^{-9}$ m] | Strain [$10^{-9}$ m/m] |
|---|---|---|---|
| 1 | 508 | 1 | 0.34 |
| 2 | 735 | 1 | 0.70 |
| 3 | 871 | 1 | 0.63 |
| 4a | 1299 | 1 | 0.89 |
| 5a | 2228 | 1 | 1.5 |
| 6 | 2371 | 1 | 1.9 |

strain intensity at the top surface of the plate,

$$\varepsilon_{\text{intensity}}(x,y) = \max(|\varepsilon_1(x,y)|, |\varepsilon_2(x,y)|), \tag{2.20}$$

is shown. The grid of stripes that is superimposed in the plots indicates the direction of the largest strain of the respective modeshape; a strain sensor should be oriented according to these stripes in order to have the highest sensitivity at a certain location. The length of the stripes is proportional to range of the strain at that point, i.e. the difference between maximum and minimum strain:

$$\varepsilon_{\text{range}}(x,y) = |\varepsilon_1(x,y) - \varepsilon_2(x,y)|.$$

This indicates the strain sensor's sensitivity to the mounting direction. At umbrella points, with $\varepsilon_1$ and $\varepsilon_2$ relatively close to each other, the measurement direction is less critical than in saddle points, where $\varepsilon_1$ and $\varepsilon_2$ have different signs.

Table 2.2 lists the maximum strain that occurs in each of the first eight modeshapes, based on the results of Figure 2.8. The displacement fields are scaled such that their maximum displacement is 1 nm. Thus, assuming that displacements are in the order of 1 nm and that the displacement field is well-described by the first eight modeshapes, the typical strains are in the order of $10^{-9}$ m/m.

Note that the surface strain highly depends on the wafer chuck's geometry. Here, a plate is made out of aluminium with dimensions of 0.55 m × 0.55 m × 0.032 m was chosen. A different thickness, for example, would lead to different strain. From Eqs. (2.18) and (2.19) it follows that the surface strain of a plate is proportional to both the thickness of the plate and the scaling of the displacement field, i.e.

$$\varepsilon \propto hw. \tag{2.21}$$

The scaling of the displacement field is proportional to the stiffness of the plate, which is in turn, according to Eq. (A.1), proportional to the thickness to the

power three, so that $F/w = k \propto h^3$. Thus, the surface strain scales according to

$$\varepsilon \propto \frac{F}{h^2}. \tag{2.22}$$

A halve as thick plate would experience a four times higher strain.

The measurement of strains in the order of $10^{-9}$ m/m in a bandwidth up to 100 Hz would require strain sensors with a noise level lower than $10^{-8}$ $1/\sqrt{\text{Hz}}$. Lee (2003) presents an overview of the resolution of FBG strain sensors in literature [38]. The resolutions range from $10^{-8}$ $1/\sqrt{\text{Hz}}$ to $10^{-14}$ $1/\sqrt{\text{Hz}}$. Note that these sensor systems are not commercially available.



### 2.4.2 Obtaining absolute position from inertial measurements

With an inertial sensor, it is possible to measure absolute position of a point directly or to obtain it from its time derivates, namely absolute velocity and acceleration. A practical advantage of measuring absolute position is that the sensor needs no stationary reference in its line of sight, allowing position measurement at points close to the lens column.

This section describes how absolute position can be obtained from inertial measurements with a single mass-spring-damper system and elaborates on to what extend the absolute position can be used for estimation of the wafer chuck's deformation.

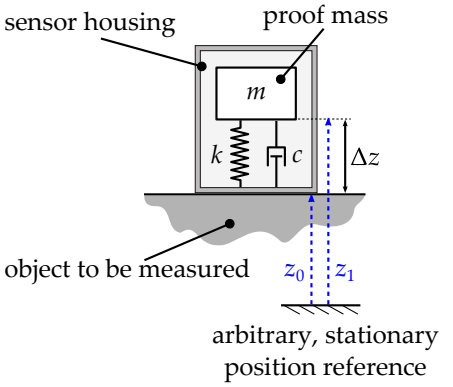
#### Single mass-spring-damper system

Most accelerometers are in essence single mass-spring systems, which can be modelled like in Figure 2.9a. If the so-called proof mass accelerates it experiences inertia forces which causes the mass's suspension to deform. Their relative position, $\Delta z$, is a measure for the acceleration of the body the sensor is connected to, $\ddot{z}_0$. Velocity sensors make use of various physical principles. The geophone, which is frequently used in seismic applications, is based on a mass-spring system like the accelerometer. Here, the velocity of the proof mass, $\Delta\dot{z}$, is translated into an electric signal using a coil and a magnetic field.

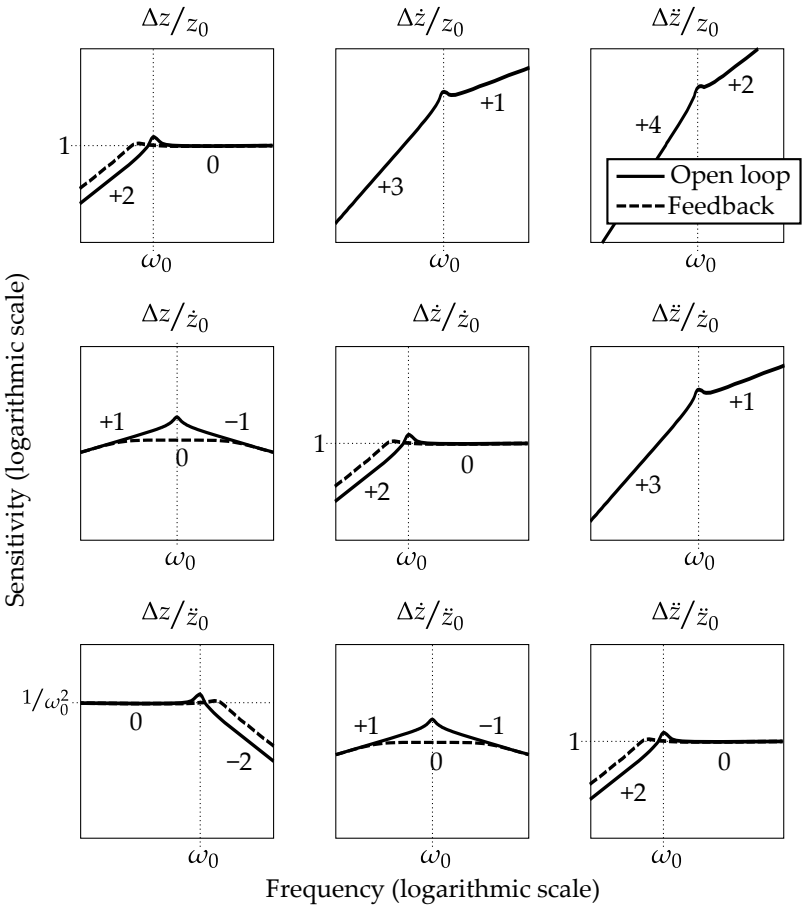
Figure 2.9b presents an overview of the sensitivity of different single mass-spring-damper-based sensors. Either the relative position, $\Delta z$; relative velocity, $\Delta\dot{z}$, or relative acceleration, $\Delta\ddot{z}$, is measured.[2] This measurement is, in turn, a measure of position, $z_0$; velocity, $\dot{z}_0$, or acceleration, $\ddot{z}_0$, of the object of interest. The sensitivity of the sensor is the transfer from the desired output of the sensor to the actually measured quantity and is plotted for the different combinations of position, velocity and acceleration.

The dashed lines In Figure 2.9b indicate how the transfer function can be changed with active feedback. Proportional/Derivate (PD) controllers are used that aim at keeping the either $z_0$ constant or $\dot{z}_0$ or $\ddot{z}_0$ zero. In this way, in the low-eigenfrequency systems (transfer functions $\Delta z/z_0$, $\Delta\dot{z}/\dot{z}_0$ and $\Delta\ddot{z}/\ddot{z}_0$) the eigenfrequency of the mass spring system can be influenced. In the transfer functions $\Delta\dot{z}/\dot{z}_0$ and $\Delta\ddot{z}/\ddot{z}_0$, a damping action can be added, resulting in a more

[2] Here, $\Delta\dot{z}$ is meant to be the time derivative of $\Delta z$, i.e. $d/dt(\Delta z)$.




(a) A sensor based on a single mass-spring-damper system.




(b) Sensitivity plots for the different combinations of the relative motion of the proof mass ($\Delta z$, $\Delta \dot{z}$ and $\Delta \ddot{z}$) and the object to be measured ($z_0$, $\dot{z}_0$ and $\ddot{z}_0$). The numbers in the plot denote the slope of the corresponding line in dB per decade.

Figure 2.9: Accelerometers and geophones can be modelled as mass-spring-damper-systems. In the frequency regions where the sensitivity is flat, the relative movement of the proof mass with respect to the sensor's base is a direct measure of the movement of the object the sensor is connected to.

or less flat sensitivity around the sensor's eigenfrequency. Note that in all cases the measured quantity is unchanged, unlike, for example, a force balance accelerometer.

A sensor should respond equally to signals of different frequencies in the frequency band it is designed for. Of interest are, therefore, the frequency regions for which the sensitivity of the sensor is flat is. The transfer functions $\Delta z/z_0$, $\Delta \dot{z}/\dot{z}_0$ and $\Delta z/\ddot{z}_0$ have such a flat sensitivity. Those transfer functions correspond, respectively, to a modified and a standard geophone and an accelerometer.

The transfer function $\Delta \dot{z}/\dot{z}_0$ can be recognised as the one of a geophone. Colette et al. [42] adapted a standard geophone by measuring the position with respect to the housing instead of the velocity of the proof mass, obtaining the transfer function $\Delta z/z_0$. In that way, they were able to measure absolute position above the sensor's eigenfrequency. As its relevant frequency band is above its eigenfrequency, this sensor, and also the standard geophone, should have a low eigenfrequency, $\Omega_0$. As the eigenfrequency is determined by the ratio between stiffness and mass, low stiffness and high mass is desired. A low-stiffness suspension of the proof mass is typically fragile, which is a problem in the wafer chuck with its high acceleration. A high mass, on the other hand, leads to a large sensor volume, which is not desired either. Thus, geophones are of limited interest for the wafer chuck application.

Accelerometers have the transfer function $\Delta z/\ddot{z}_0$. The frequency band of interest is below the sensors eigenfrequency and, thus, accelerometers can be made relatively compact and robust. As such, accelerometers are a possible candidate for use in the wafer chuck application.

**Double integration of acceleration**

Acceleration needs to be double integrated to find the position of a point. Acceleration is an internally referenced measurand and, therefore, has the fundamental benefit that it does not need an external reference. On the other hand, double integration amplifies the low-frequency noise [43], so that, typically, only the high-frequency part of the signal is useful. In the following, the uncertainty of the obtained position signal as function of time is obtained.

Assume an accelerometer signal consists of the actual acceleration plus a noise contribution and a constant due to sensor bias

$$\tilde{a}(k) = a(k) + \varepsilon(k) + c, \tag{2.23}$$

where $k$ denotes the discrete time index. After double integration (Appendix C) the position is found as follows

$$\tilde{z}(k) = T^2 \sum_{j=1}^{k} \sum_{i=1}^{j} a(i) + \eta(k) + \frac{k^2 + k}{2} T^2 c + kTv(0) + z(0), \tag{2.24}$$

with $\eta(k)$ the noise sequence of $\tilde{z}$ due to the accelerometer's noise. In case that the sensor noise $\varepsilon$ is Gaussian distributed with standard deviation $\sigma_\varepsilon$ and if it

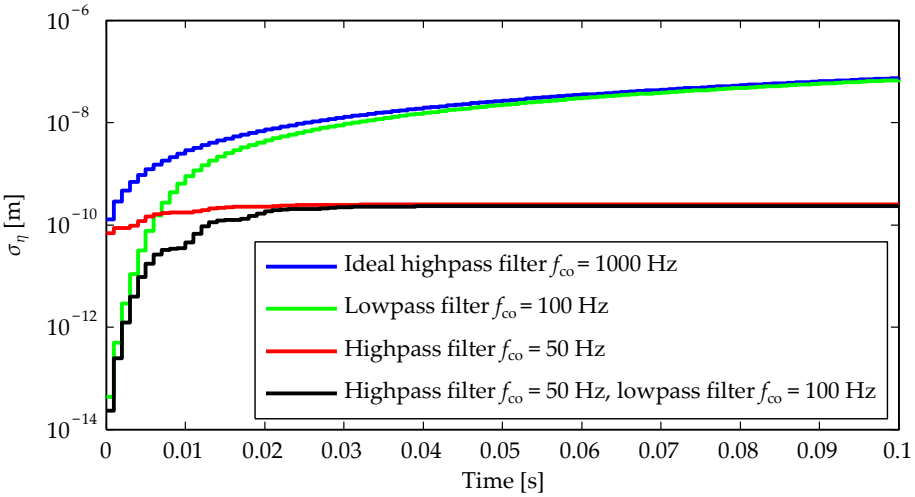



Figure 2.10: The standard deviation of the error when obtaining position from an acceleration signal. Here, only the error due to accelerometer noise, $\varepsilon$, is shown, but also the error on the bias, $c$, and the initial position and velocity, $z(0)$ and $v(0)$ are important.

is white so that there is no correlation between the noise samples, the standard deviation of the position error is, according to Appendix C,

$$\sigma_\eta(k) = T^2\sigma_\varepsilon\sqrt{\frac{k^3}{3}+\frac{k^2}{2}+\frac{k}{6}}, \tag{2.25}$$

in which $k^3/3 + k^2/2 + k/6$ can be recognised as the 'square pyramidal number' [44].

The relatively light MEMS force balance accelerometer Kistler 8330B3 (Appendix E) has a, supposedly, flat noise power density of $P_f = 1.6 \cdot 10^{-11}$ m$^2$s$^{-4}$/Hz, so that the variance of the noise for an ideal lowpass filter with a cutoff frequency of 1000 Hz is

$$\sigma_\varepsilon = \sqrt{(1.6 \cdot 10^{-11})1000} = 1.3 \cdot 10^{-4}\ \frac{\text{m}}{\text{s}^2}. \tag{2.26}$$

Figure 2.10 shows in blue the resulting $\sigma_\eta$ as function of time for a sample frequency of $f_s = 1000$ Hz. The uncertainty of the calculated position signal increases rapidly, caused by the fact that $\sigma_\eta$ is proportional to $\sqrt{k^3}$ for larger $k$. After 5 ms the standard deviation of the error is already over 1 nm, To improve the accuracy of the position signal, techniques like frequent recalibration of the position or filtering must be employed. Note that also the acceleration bias, $c$ and the error on the initial position, $z(0)$, and the initial velocity, $v(0)$, and should be accounted for by calibration or adequate filtering.

The assumption that the noise is white is in practice rather artificial, as not all accelerometers have a flat noise spectrum, e.g. piezoelectric accelerometers (Appendix E). Moreover, even if the noise spectrum is flat up to a certain frequency, it will become non-flat after filtering and sampling. When a signal is sampled, any noise contributions above the Nyquist frequency (half the sample frequency $f_s$) will fold back to the frequency band up to the Nyquist frequency.

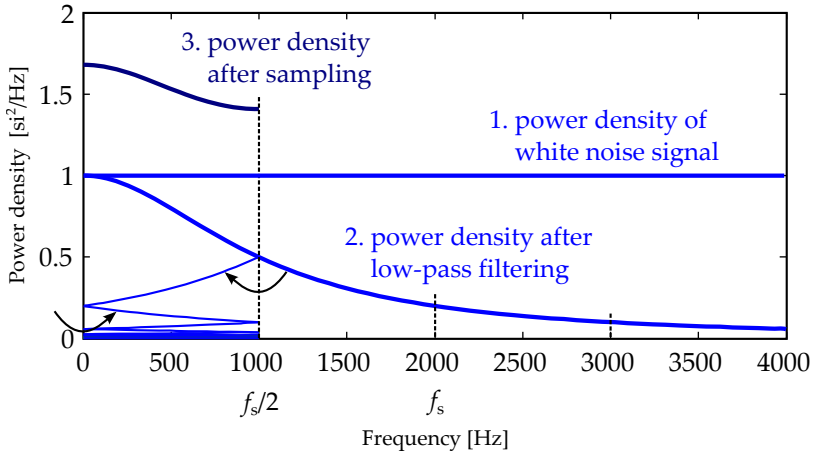


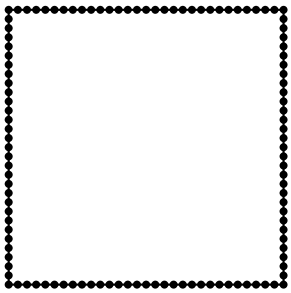

Figure 2.11: The effect of frequency folding due to aliasing when sampling a signal of uncorrelated noise with a sampling frequency $f_s$. All power density content folds back to the frequency band between 0 and the Nyquist frequency ($f_s/2$). The power density of the resulting sampled signal is the sum of all this folded content.

The frequency folding of sensor noise is illustrated in Figure 2.11. Assume a white noise signal, having a flat power density as function of frequency, as depicted by 1. After filtering this signal with a first order lowpass filter with a cutoff frequency equal to the Nyquist frequency its power density is decreased to the level indicated with 2. Now, if the the signal is sampled, its noise power does not change, but relocates by folding back into the frequency band up to the Nyquist frequency. As the noise contributions in the frequency domain are assumed to be uncorrelated, the power density of the sampled signal is the sum of all folded power contributions, denoted in the figure by 3. The resulting noise spectrum is not flat, so that the the assumption that the noise samples are not correlated and the resulting Eq. (2.25) are not valid anymore. Besides, this example also illustrates the noise components from above the Nyquist frequency should be adequately filtered before sampling in order to avoid them to aliase.

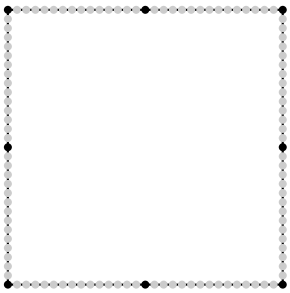
The green line in Figure 2.10 shows the $\sigma_\eta$ as function of time when the signal is lowpass filtered with a 6th order Butterworth filter with a cutoff frequency of 100 Hz. The uncertainty rises, initially, less rapidly. The red line shows $\sigma_\eta$ when only a 6th order Butterworth highpass filter with a cutoff frequency of 50 Hz is applied to the acceleration signal. As expected, the highpass filter removes the low-frequency part of the signal, so that $\sigma_\eta$ converges after some time to a constant value. The black line corresponds to applying both filters.

**Absolute position for estimation of the PoI's position**

The position of the wafer chuck is measured with respect to the metrology frame. Due to the metrology frame's high internal stiffness, its deformation may be neglected. The low-stiffness suspension to the ambient, however, causes a lowpass filter characteristic regarding the floor vibrations. At frequencies



(a) The set of position sensor candidates consist of the locations at the sides of the wafer chuck. Here, a relatively dense set is shown, consisting of 120 candidates.



(b) As a preliminary choice, a configuration consisting of eight sensors at the corners and centres of the sides is adopted. As the sensors are placed far from each other.

Figure 2.12: The chosen preliminary sensor configuration for the position measurement in $z$-direction at eight points on the wafer chuck.

below the metrology frame's suspension eigenfrequency, the metrology frame moves along with the floor vibrations. Thus, below the eigenfrequency, there is a discrepancy between the absolute position measurement at the wafer chuck and the position measurement referenced to the metrology frame. Above the eigenfrequency, however, the metrology frame can be considered at standstill, i.e. having a constant absolute position.

Both inertia-based sensor concepts posses a highpass filter characteristic, i.e. above a certain frequency they yield a useful position signal. This is either due to the physical principle (in case of a position measurement with a proof-mass) or the required highpass filtering (in case of a double integrated acceleration signal).

Thus, above the suspension eigenfrequency of the metrology frame and the cutoff frequency of the inertia-based sensor, the measurement of the absolute position with the inertia-based sensor can be directly related to the position of the metrology frame. Still, if a discrepancy would remain, the absolute position of the metrology frame could be measured with a similar inertia-based sensor.

### 2.4.3 Sensor placement candidates

Figure 2.12a shows sensor candidates for using externally referenced position sensors. Only sensor candidates at the sides of the wafer chuck are taken into account, as these are least covered by the lens column and the immersion hood during the scanning phase. Out of these candidates, a configuration of eight sensors (Figure 2.12b) is selected for further use in the next chapters. Chapter 6 elaborates on the optimisation of the sensor configuration by automated sensor placement.

## 2.5 Conclusions

The most critical part of the lithography process is the exposure. The wafer chuck scans during the exposure beneath the lens column with a highly constant forward velocity, while it is kept still in its other DoFs. The wafer chuck can thus be considered in steady state during exposure.



The deformation of the wafer chuck is partly caused by known and partly by unknown forces. The deformation due to unknown forces, i.e. the disturbances, must be estimated. It is important to make use of the available foreknowledge on the disturbances and the dynamical system when performing estimation. The disturbances are known to be low-frequency, consisting of frequencies typically well-below the resonance frequencies of the wafer chuck.

The dynamics of a solid mechanical system with a low damping can be described by a linear combination of its dynamic modeshapes, according to the corresponding modal coordinates. These modal coordinates have behave as uncoupled second order systems. The form of the mechanical problem is equal for a wafer chuck and a plate, so that for simplification in the further analysis in this thesis a mechanical model based on plate finite elements can be used. The modeshapes and eigenfrequencies of the FE model were found to match closely experimental results.

As an input for a position estimation algorithm, externally referenced position measurements can be used. Those may either be internally or externally referenced. Strain can be seen as a form of internally referenced position. For displacements of 1 nm, typically maximum surface strains in the order of $10^{-9}$ $\mathrm{m/m}$ are encountered. From double integration of acceleration, absolute position can be obtained. This amplifies the low-frequency noise, making only the high-frequency part of the position signal useful. Strain sensors and accelerometers are compact and do not require an external reference, so that the placement and the number of those sensors is relatively unrestricted.

Position sensors that measure with respect to the stable metrology frame are externally referenced sensors. The existing four position sensors may be extended with additional sensors. Those sensors can only be placed outside the wafer area and there where they are in sight of the reference. This restricts the placement to the sides of the wafer chuck. Therefore, a sensor configuration with four additional position sensors at the centres of the sides is proposed.

# Chapter 3

# Estimation using shape fitting

The previous chapter derived requirements for the estimation method, based on the properties of the wafer chuck system. The method should be able to incorporate the foreknowledge on both the system's mechanics and its disturbances. The system is assumed to behave primarily quasi-static, as the disturbances are relatively low-frequency. This chapter selects and further introduces a promising method, which will be called *Shape Fitting* (SF), that utilises this quasi-static behaviour. Furthermore, the *snapshot matrix*, which makes it possible to include information on the system's mechanics and its disturbance distribution, is introduced.

Section 3.1 first formulates the estimation problem. Section 3.2 introduces the quasi-static behaviour of a system without rigid body modes when excited by disturbances well-below the system's lowest dynamic eigenfrequency. Then, in Section 3.3 a literature survey on estimation methods is presented. The basic principles of the SF method, as described in literature, are further elaborated in Section 3.4. Section 3.5 describes the use of snapshot-matrices to build the SF estimator.

## 3.1 Definition of the estimation problem

The previous chapter showed that the estimation of the relative position between the PoI and the lens column boils down to the estimation of the deformation field of the wafer chuck. In such an estimation problem the values at a set of target points need to be estimated from the values at a set of measurement points. Suppose that a physical field is discretised in a grid of $N_P$ points. Let the values at the measurement and target points be stacked as respectively vectors $\boldsymbol{w}_m$ and $\boldsymbol{w}_t$ and let a vector $\boldsymbol{w}$ describe the momentaneous shape at the $N_P$ points:

$$\boldsymbol{w} = \left\{ \begin{array}{c} \boldsymbol{w}_m \\ \boldsymbol{w}_t \end{array} \right\}. \tag{3.1}$$

Here, the symbol $\boldsymbol{w}$ was chosen to refer to displacement or deformations. In the wafer chuck case, the elements of $\boldsymbol{w}$ may correspond to a grid of points, so that

$\boldsymbol{w}$ can indeed be interpreted as a geometric displacement shape, of which some points are measured and the others have to be estimated. In other application fields, $\boldsymbol{w}$ might denote other physical quantities. In the thermal domain for example, it would contain the temperatures of the considered points.

The measurement and the target values could also be combinations of different quantities, e.g. strains and displacements. Moreover, the target values need not necessarily describe the full surface of an object. In the wafer chuck, for example, only the points directly under the lens column are of interest. In such cases, the name 'shape' must not be interpreted geometrically but more generally as a snapshot of values at the same time instance.



The measurements $\tilde{\boldsymbol{w}}_\mathrm{m}$ of the quantities $\boldsymbol{w}_\mathrm{m}$ are assumed to be corrupted by sensor noise $\boldsymbol{\varepsilon}_\mathrm{m}$, so that

$$\tilde{\boldsymbol{w}}_\mathrm{m} = \boldsymbol{w}_\mathrm{m} + \boldsymbol{\varepsilon}_\mathrm{m}. \tag{3.2}$$

The sensor noise is assumed to be Gaussian distributed. This is a common assumption and is from a mathematical perspective convenient, as linear combinations of Gaussian variables are Gaussian themselves. A steady-state bias may be removed through calibration of the sensor so that a zero mean can be assumed. The noise of the $i$th sensor is then characterised by its standard deviation $\sigma_{\varepsilon,i}$. Furthermore, it is assumed that the noise of the different sensors is uncorrelated. Thus, the covariance matrix of the sensor noise, $\boldsymbol{\Sigma}_\varepsilon$, is diagonal, with the variances $\sigma^2_{\varepsilon,i}$ on its diagonal.

From the measurements, $\tilde{\boldsymbol{w}}_\mathrm{m}$, the values of the target points, $\boldsymbol{w}_\mathrm{t}$, need to be estimated.

## 3.2 Quasi-static system response and estimation

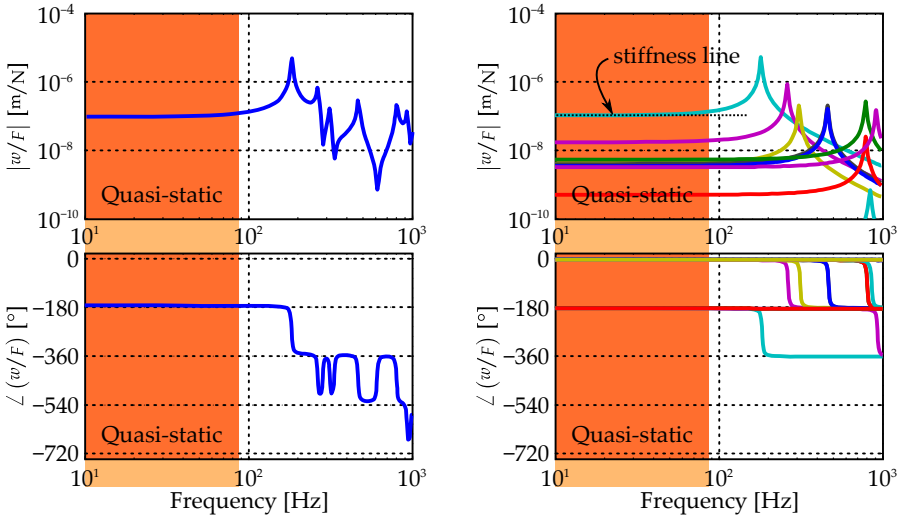
Estimation is possible due to the correlation between the measurement values $\boldsymbol{w}_\mathrm{m}$ and the target values $\boldsymbol{w}_\mathrm{t}$. The correlation is influenced by the mechanics of the object and the disturbances that work on the system. Section 2.1.5 described the available foreknowledge on the disturbances. The main disturbances are induced close to the PoI via the immersion layer below the lens column. Furthermore, the bandwidth of the disturbances is limited to well-below the lowest eigenfrequency of the wafer chuck. This has the important consequence that the wafer chuck can be considered to deform quasi-statically.[1]

In the quasi-static frequency band, inertia and damping forces are small compared to the stiffness forces, so that Eq. (2.3) becomes

$$\boldsymbol{K}\boldsymbol{w} \simeq \boldsymbol{f}. \tag{3.3}$$

A system without Rigid Body (RB) modeshapes has an invertible stiffness matrix $\boldsymbol{K}$ [45], so that its response, $\boldsymbol{w}$, is directly prescribed by its inputs, $\boldsymbol{f}$. Consequently, the past values of the measurement signal's do not contain more information about the system's response than the current measurement

[1] This is true for a wafer chuck that is suspended by a controller with infinite bandwidth. In practice the controller bandwidth is limited, however. Chapter 5 shows that the deformation behaves quasi-statically below the controller bandwidth.




(a) The total transfer function.

(b) The transfer function in terms of the individual eigenmodes.

Figure 3.1: The transfer functions from a force $F$ to a displacement $w$ of a plate without RB motion. For frequencies well-below the first eigenfrequency, the system behaves quasi-statically – its response does not depend significantly on the frequency. In the quasi-static frequency band the individual eigenmodes behave according to their stiffness line; none of the eigenmodes resonates.

values. The estimation method, therefore, does not need to use the history of the measurement signals.

Section 2.2.3 showed that the response of a linear mechanical system without damping can be regarded as the sum of the responses of individual second order systems (Figure 3.1a and b). This means that, if no RB modes are present in a system, the system's response at frequencies lower than the eigenfrequencies is primarily described by the stiffness lines of the system's eigenmodes. Thus, the system's response almost equals the static response, the limit frequency for which the inertia and damping forces are fully zero. Therefore, the system is said to behave quasi-statically in the frequency band well-below its eigenfrequencies. The quasi-static frequency band is indicated in Figure 3.1.

In this and the next chapter, RB motion of the wafer chuck is excluded. Estimation methods that can be applied to systems without RB motion are expected to also work efficiently on systems that have RB motion in combination with a position controller. Chapter 5 will extent the estimation method to the situation in which RB modes are present.

## 3.3 Estimation methods in literature

Many literature sources have been devoted to the topic of estimation of physical fields. The estimators can be classified as either static or dynamic. Static

estimators are estimators that do not take into account the history of the measurement signals, the estimate is fully based on the present sensor readings. Dynamic estimators, by contrast, make use of past measurements for generating the estimate.

### 3.3.1 Dynamic estimators

As dynamic estimators were not expected to yield significant better results for a quasi-static system, no thorough study on dynamic estimators was carried out. An often employed dynamic estimator is the Kalman filter of Kalman and Bucy [46]. The Kalman filter can be regarded as an optimal observer [47]. Knowing the statistical properties of the disturbances working on a system and the sensor noise, the Kalman filter optimally estimates the states of a system, for example the displacements of points on an object. It must be noted that, the basic principles of static estimation via shape fitting and dynamic estimation with a Kalman filter turn out to be similar. This is further elaborated in Appendix F. Chierichetti and Ruzzene presented a method for reconstruction of the dynamic displacement field of a plate using a limited set of measurement [48]. This method, however, assumes periodic signals and must be applied offline. This method can, as such, not be used in the wafer chuck application, which aims at real-time estimation.

### 3.3.2 Static estimators

Most of the considered static estimation methods were formulated in the mechanical domain, a few in the thermal domain. The majority of the methods, however, can equally well be applied in each of the domains. The methods in the mechanical domain often aim at estimating displacement using strain measurements, making use of the relation between strain and deformation. Many papers are written in the context of showing the use of Fibre Bragg Grating (FBG) strain sensors to find the shape of an object. In general, these papers are relatively sensor oriented and less elaborate regarding the estimation method. Roughly speaking, the static estimation methods can be subdivided into inverse FE methods and fitting methods that make use of shapes, which will be called 'SF methods'.

#### Inverse Finite Element methods

Tessler and Spangler developed an inverse FE method to estimate the deformation in plates and shells. The method makes use of the geometric properties of the object to estimate deformation from discrete strain measurements [49]. It was later adapted to other structures, like shear deformable plates, beams and three dimensional frames [50–52]. Gherlone et al. state that "*Because only strain-displacement relations are used in the formulation, both static and dynamic responses can be reconstructed without any a priori knowledge of loading, material, inertial or damping structural properties.*" [52] Thus, their method does not take into account the available foreknowledge but has a certain implicit assumption

on the disturbance distribution. The inverse FE method of Mainçon, in contrast, does allow including this foreknowledge, along with the foreknowledge on the sensor noise [12]. The information is included via variance and covariance data.

### Shape fitting methods

Shape fitting methods use a number of basis vectors or shapes to describe the deformation or displacement field. The amplitudes of the shapes are fitted to the measurements. Three major groups of shapes can be distinguished, i.e. polynomial shapes, dynamic modeshapes and static deformation shapes.



**Polynomials** Jones and his co-workers, Sekouri, Kim and Cho, Wang et al. and Lee and Park made use of polynomial shapes [10, 53–57]. None of them shows, however, how foreknowledge on the disturbances could be included. Kirby et al. show, as an example, also the use of polynomials functions that satisfy the object's boundary conditions [8]. They also mention the possible use of dynamic modeshapes.

**Dynamic modeshapes** Often, the dynamic modeshapes of the system in question are used as fitting shapes [10,11,13,39,58–64]. The dynamic modeshapes are an obvious choice, as they form a linear basis that can describe any displacement field, as presented in Eq. (2.8). In most cases the number of modeshapes is reduced via truncation, keeping only a set of modeshapes corresponding to the lower eigenfrequencies. In case of dynamic disturbances, this seems a reasonable choice. The disturbances are normally bound to a certain maximum frequency. If the system has eigenfrequencies below that frequency, the corresponding modeshapes might be relatively prominent in case of resonance.

Also in case of disturbances that excite the object quasi-statically, the modeshapes corresponding to the lower eigenfrequencies are expected to be relatively prominent, an effect which is demonstrated by Bert [65]. Note that this later effect is the case for the modeshapes in terms of displacement. In terms of the surface strain fields corresponding to the modeshapes, the higher modeshapes of, for example, a plate or a beam stay relatively important. This has to do with the fact that surface strain is the second spatial derivative of displacement and that the spatial frequency of the modeshapes increases with higher eigenfrequency. Thus, when performing strain measurements, relatively many modeshapes need to be taken into account. This effect is illustrated in Examples 3.1 and 3.2. Example 3.1 introduces the displacement modeshapes of a beam and their corresponding strain profiles, which are then used in Example 3.2 to fit the displacement shape and strain profile resulting from a certain load case.

Most papers do not substantiate the truncation, i.e. selection of the number of modeshapes, otherwise than in relation with the number of sensors. Bogert et al., however, base their truncation on the modal strain energy [66]. Koevoets et al. propose weighing factors to prioritise and modeshapes, although they do not clarify how these factors should be chosen [16]. In an example, they use the weighing factors in effect for a truncation. In any case, it is still unclear how

foreknowledge on quasi-static disturbances can be included to yield an effective SF estimator based on modeshapes.

**Static displacement shapes** Another type of fitting shapes are the static displacement shapes. These seem an obvious choice in case of quasi-static disturbances. Weeks uses Green's functions as fitting shapes, which are the static deformation shapes due to a point-force [67, 68]. Her method does attribute weighing factors to the sensor locations but not to the disturbance locations. As such, it does not include foreknowledge on the disturbances, except for the implicit assumption that the disturbances are quasi-static and can work at any included point on the structure. Davis and his co-workers use static deformation shapes that are the result of arbitrarily hand-picked static load combinations [69,70].

The latter method can be considered as a snapshot-based method. A snapshot captures the instantaneous shape of the field of interest. A series of snapshots is combined to a snapshot matrix. The series may consist, for example, of successive snapshots in time, describing the systems transient response, or of separate quasi-static responses. As such, it is possible to capture the typical responses of a system to its disturbances in a snapshot matrix and include the available foreknowledge regarding the disturbances. Koevoets et al. and Ranieri et al. make use of snapshot matrices. They use the Proper Orthogonal Decomposition (POD) to find the most prominent shapes in the snapshot matrix, which are then used for shape fitting [9, 14]. Hakim and Fuchs make use of a snapshot matrix of typical static responses of the system [71]. They directly use the snapshots as fitting shapes.

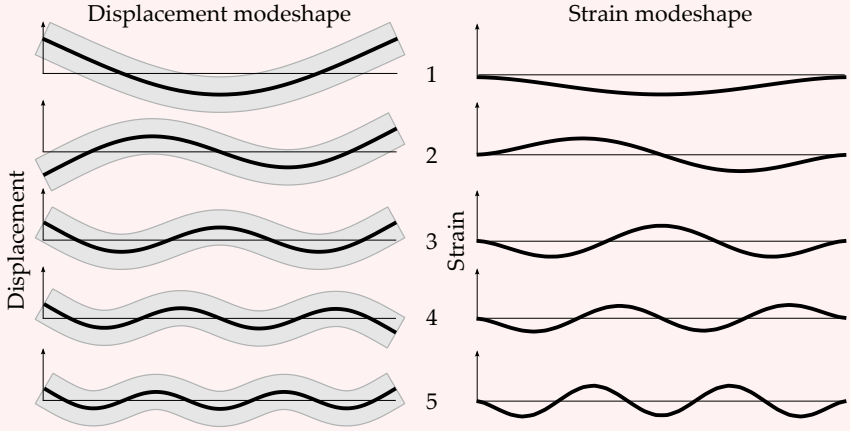
**Number of sensors versus number of shapes** Finally, it is noticeable that the papers have different assumptions regarding the number of sensors, $N_\mathrm{m}$, with respect to the number of fitting shapes, $N_\mathrm{f}$. Some of the papers assume an equal number of sensors and fitting shapes ($N_\mathrm{m} = N_\mathrm{f}$) [13, 60, 61, 69, 70]. They then use an inverse to find the shape amplitudes from the sensor values. Some other papers consider the case that the number of sensors is smaller than the number of fitting shapes ($N_\mathrm{m} \leq N_\mathrm{f}$) [14, 67, 68, 71]. In this case the problem is underconstrained. The majority of the papers only considers the case that the number of sensors is larger than the number of fitting shapes ($N_\mathrm{m} \geq N_\mathrm{f}$) [9–11, 14, 39, 40, 53–59, 62–64, 66, 72], yielding an overconstrained problem. All considered SF methods use a least squares optimisation to solve overconstrained and underconstrained problems. One exception is the method of Weeks, which is, however, closely related to least squares [67, 68].

## 3.4 The shape fitting method

This section introduces the basic principles and the formulas of the SF method.

**Example 3.1: Displacement modeshapes and strain modeshapes of a beam**

The figure shows the first five flexible modeshapes of a free beam, along with the corresponding surface strain profiles, which are called "strain modeshapes". All shapes are arbitrarily scaled.






### 3.4.1 Procedure

The SF method assumes that the momentaneous shape of the field can be well approximated by a linear combination of *fitting shapes*. From the momentaneous values at the measurement points, the amplitudes of the fitting shapes are calculated. The shape estimate is then the sum of the fitting shapes weighted by these amplitudes.

Suppose that the number of fitting shapes is $N_\mathrm{F}$. Each fitting shape is formulated as a vector $\boldsymbol{\varphi}_i$ that describes the values at all $N_\mathrm{P}$ points of the field. The fitting shapes can be combined in a $N_\mathrm{P} \times N_\mathrm{F}$-matrix $\boldsymbol{\Phi}$. The momentaneous shape of the field, $\boldsymbol{w}$, is then approximated by a linear combination of a limited number of fitting shapes:

$$\hat{\boldsymbol{w}} = \boldsymbol{\Phi}\hat{\boldsymbol{q}}, \tag{3.4}$$

where the elements of $\hat{\boldsymbol{q}}$ are the weight values corresponding to the fitting shapes. Those weighing values are found using

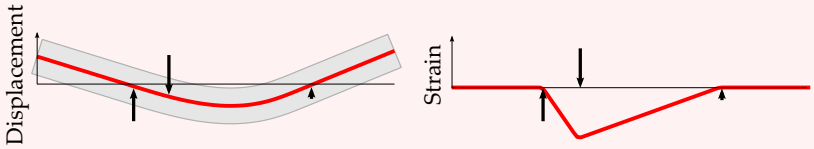
$$\hat{\boldsymbol{q}} = \boldsymbol{\Phi}_\mathrm{m}^{+}\tilde{\boldsymbol{w}}_\mathrm{m}, \tag{3.5}$$

where $N_\mathrm{M} \times N_\mathrm{F}$-matrix $\boldsymbol{\Phi}_\mathrm{m}$ consists of the subset of rows of $\boldsymbol{\Phi}$ that correspond to the measurement points.

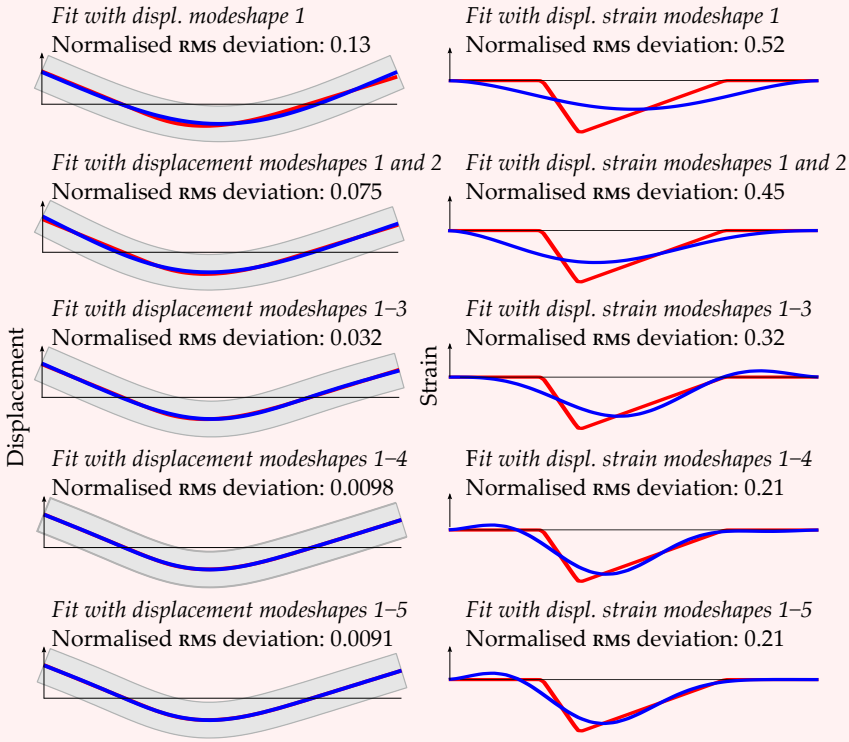
$\boldsymbol{\Phi}_\mathrm{m}^{+}$ is the pseudoinverse of $\boldsymbol{\Phi}_\mathrm{m}$, whose formulation depends on the number of sensors with respect to the number of fitting shapes:

**Example 3.2: Fitting with modeshapes and strain modeshapes**

The following figure plots a static deformation shape due to a disturbance force and two actuator forces (left). Also the corresponding strain profile is drawn (right).




These shapes are in the following figure fitted by sets of respectively the modeshapes and the strain modeshapes (Example 3.1). To evaluate how close a fit approaches the actual shape, the RMS value of the deviation between the fit and the actual shape is calculated and then normalised by the RMS value of the actual shape. For describing the static deformation shape with modeshapes, relatively less modeshapes are needed than for describing the strain shape in terms of the strain modeshape.




Note that the fitting procedure shown in this example is not the same as shape fitting (Section 3.4). In this example the best fit at all points was calculated, whereas shape fitting fits the fitting shapes only at the sensor points and then uses the fitting shapes to extrapolate this fit to the other points of the structure.

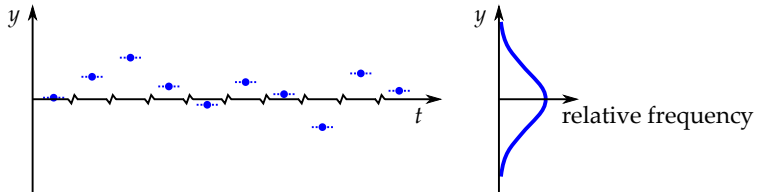



Figure 3.2: Mathematical model of the system's signals. Each of the signals is assumed to assume slowly over time. If the intervals between sampling are sufficiently large the samples are almost uncorrelated (left). The value of each sample can, therefore, be regarded as a realisation of a random variable. The distribution of the random variable is assumed to be, in first approximation, Gaussian with a zero mean (right).



$$\boldsymbol{\Phi}_{\mathrm{m}}^{+} = \begin{cases} \left(\boldsymbol{\Phi}_{\mathrm{m}}^{\mathsf{T}}\boldsymbol{\Phi}_{\mathrm{m}}\right)^{-1}\boldsymbol{\Phi}_{\mathrm{m}}^{\mathsf{T}} & N_{\mathrm{M}} > N_{\mathrm{F}} \\ \boldsymbol{\Phi}_{\mathrm{m}}^{-1} & N_{\mathrm{M}} = N_{\mathrm{F}} \\ \boldsymbol{\Phi}_{\mathrm{m}}^{\mathsf{T}}\left(\boldsymbol{\Phi}_{\mathrm{m}}\boldsymbol{\Phi}_{\mathrm{m}}^{\mathsf{T}}\right)^{-1} & N_{\mathrm{M}} < N_{\mathrm{F}}. \end{cases} \tag{3.6}$$

### 3.4.2 Estimation matrix $\boldsymbol{B}$

Define the $N_{\mathrm{P}} \times N_{\mathrm{M}}$ estimation matrix $\boldsymbol{B}$ as

$$\boldsymbol{B} = \boldsymbol{\Phi}\boldsymbol{\Phi}_{\mathrm{m}}^{+}, \tag{3.7}$$

so that the estimate becomes

$$\hat{\boldsymbol{w}} = \boldsymbol{B}\tilde{\boldsymbol{w}}_{\mathrm{m}}. \tag{3.8}$$

This illustrates that the SF estimates are simply linear combinations of the sensor values, weighed according to the elements of $\boldsymbol{B}$. Alternatively, the estimated shape can be considered as a linear combination of the columns of $\boldsymbol{B}$, weighed by the sensor readings. As such, the columns of $\boldsymbol{B}$ can be regarded as the shapes corresponding to each of the sensors. If the estimate, $\hat{\boldsymbol{w}}$, has the geometrical interpretation as shape, also the columns of $\boldsymbol{B}$ have that. The estimation matrix turns out to be important in the characterisation of the estimation error (Section 3.5.5).

## 3.5 Shape fitting using a snapshot matrix

This section introduces the mathematical interpretation of a snapshot in a quasi-static system and the snapshot matrix that is used for shape fitting. Furthermore, the way the snapshot matrix can be obtained in a practical system and its application for assessing the estimator's performance in terms of the estimation error is discussed.

### 3.5.1 Definition of the snapshot and its mathematical model

As described in Section 3.2, the bandwidth of the disturbances that work at the wafer chuck is limited to well-below the lowest eigenfrequency of the wafer

chuck, so that the wafer chuck's deformation an be considered quasi-static. Thus, the disturbances and the deformation response vary slowly over time and the instantaneous response is a direct result of the instantaneous disturbances.

The deformation response at a certain time instant is called a 'snapshot'. Two snapshots that were recorded with only a short time difference would only differ slightly, due to the slow variation over time. A snapshot that is, however, taken at a time instant sufficiently later, is almost uncorrelated with the first one. Thus, a set of snapshots is more uncorrelated if the snapshots are recorded with a larger time interval (Figure 3.2, left).



Based on probability theory, the following mathematical model of the system's signals is adopted. The system's signals are constant during the running of an experiment. At every initiation of an experiment, the input and outputs attain new values, which are independent from the previous values. Thus, the system's inputs and outputs can be regarded as random variables and their values in each experiment as realisations of those variables.

An important assumption that is made here is that the inputs, i.e. the disturbances, are all Gaussian distributed with a zero mean and a certain standard deviation (Figure 3.2, right). The disturbances may or may not be correlated. Many processes in nature that are the consequence of a number of random effects are, in first approximation, assumed to be Gaussian. This can be substantiated based on the central limited theorem, which states that the sum of a large number of independent random variables is approximately normally distributed, as long as the random variables are identically distributed and have a finite variance [73]. As the disturbances that work in the immersion film are based on a combination of many random effects, this assumption can in first approximation be justified. Still an open question is to what extent this process model does reflect reality and to what extent this influences the final estimation results.

An additional reason for making use of a Gaussian distribution for the disturbance amplitudes is the resulting mathematical simplicity. From a mathematics point of view, the Gaussian distribution is advantageous: the summation of Gaussian distributed variables simply yields a new Gaussian distributed variable. In practice, a simple formulation is desirable, as it makes mathematics more comprehensible. Furthermore, mathematics that is used later on, like sensor placement algorithms (Chapter 6) is on itself computationally intensive and benefits from a relatively simple process model.

### 3.5.2 Definition of the snapshot matrix

Assume that the instantaneous shape of the field, $\boldsymbol{w}$, can be described perfectly by linear combinations of $N_\mathrm{D}$ snapshots that form the columns of a $N_\mathrm{P} \times N_\mathrm{D}$ snapshot matrix $\boldsymbol{W}$. Furthermore, assume that those shapes are weighed according to the elements of a vector $\boldsymbol{\alpha}$, which are independent Gaussian variables with a zero mean and variances given by the elements of vector $\boldsymbol{\sigma}_\alpha$, so that

$$\boldsymbol{w} = \boldsymbol{W}\boldsymbol{\alpha}. \tag{3.9}$$

The elements of $w$ and the rows of $W$ can be partitioned according to whether they belong to a measurement or a target point:

$$\left\{ \begin{array}{c} w_{\mathrm{m}} \\ w_{\mathrm{t}} \end{array} \right\} = \left[ \begin{array}{c} W_{\mathrm{m}} \\ W_{\mathrm{t}} \end{array} \right] \alpha. \tag{3.10}$$

The covariance matrix of the weights $\alpha$ is $\Sigma_{\alpha}$. As the disturbances are assumed to be independent, this is a diagonal matrix.

The snapshot matrix can have different forms, depending on the way the snapshots are obtained. The snapshot matrix can be formed from the snapshots corresponding to one-by-one excitation of the disturbances, as illustrated in Example 3.3, but can also be obtained from random excitation by the disturbances. The snapshots of the second type of snapshot matrix can be seen as linear combinations of the snapshots of the first one. As long as the snapshot matrix of the first type contains enough snapshots, both snapshot matrices carry, in principle, the same information.



### 3.5.3 Obtaining a snapshot matrix

Snapshot matrix $W$ may be acquired from measurements from the actual system or from a model, but neither of these options is trivial. Acquisition of $W$ from measurements, however, is impossible if no measurement can be carried out at the target points. In the wafer chuck application it is not possible to measure the target points in the exposure area during normal operation. It is also not possible to perform the snapshot measurements off-line, when the position of the target points would be measurable, as the typical disturbances are not available then.

If, on the other hand, a model is used, this model should closely resemble the actual system. In addition, also accurate information on the typical disturbances should be available. Alternatively, a combination of measurements of the actual system and the model may be used to generate the snapshot matrix.

In the following, snapshot matrix $W$ and the corresponding weighing vector $\alpha$ are assumed to be perfect, in the sense that they fully and accurately describe the possible shapes of the system and their distribution. Such a snapshot can be used to calculate the SF estimator, but also the expected estimation error.

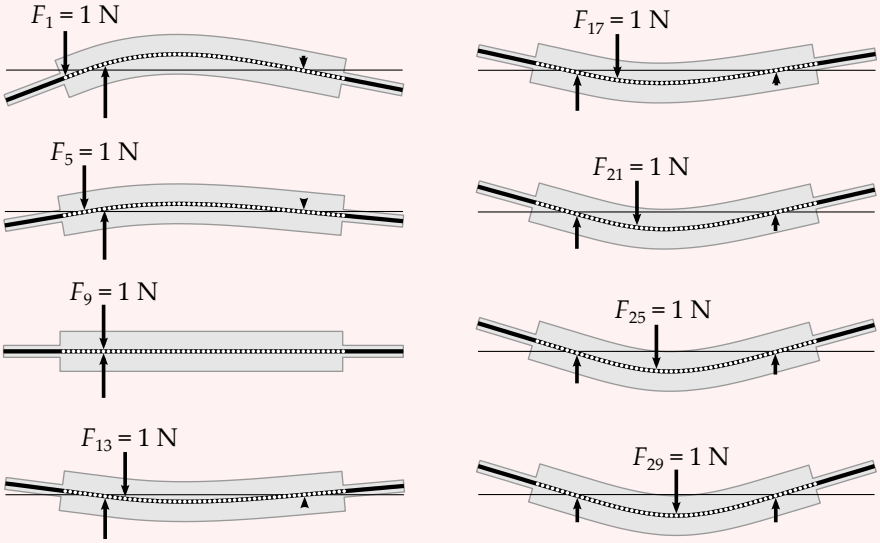
### 3.5.4 Integration of the snapshot matrix into shape fitting

As discussed in the literature review of Section 3.3, two techniques for calculating the SF estimator from a snapshot matrix were encountered in literature. The first technique directly uses the snapshot matrix for shape fitting by simply substituting $\Phi$ by $W$ and $\Phi_{\mathrm{m}}$ by $W_{\mathrm{m}}$ in the Eqs. (3.4–3.8). This will be called the Least Squares (LS) technique. The second technique first calculates the Proper Orthogonal Decomposition (POD) and then forms the matrices $\Phi$ and $\Phi_{\mathrm{m}}$ from a small number ($N_{\mathrm{F}} \leq N_{\mathrm{M}}$) of the most prominent Proper Orthogonal Modes (POMs). This technique will be referred to as the POM technique. Chapter 4 compares the two techniques.

**Example 3.3: Snapshot matrix with static deformation shapes**

Consider the beam of the figure. It is assumed that the quasi-static disturbance forces can only work in vertical direction on the middle part and that the standard deviation of their magnitude is uniform over the surface. The disturbance forces are assumed to be in static equilibrium with two actuator forces. The tips are made slender, as they do not carry loads.

The snapshot matrix is generated as follows. First a grid of possible disturbance locations is defined, in this case 57 points. One by one a disturbance is placed on such a point and the corresponding actuator forces are calculated. Then, the resulting static displacement shapes are calculated, some of which are shown in the figure.



### 3.5.5 Estimation error

Although possibly not the only criterion, the estimation error is a fundamental criterion in characterising an estimation algorithm's performance. The estimation error for the $j$th target point is defined as $\varepsilon_j = \hat{w}_j - w_j$. This can be rewritten using Eqs. (3.2) and (3.8) to (3.10) as a function of the weighing factors $\boldsymbol{\alpha}$ and sensor noise $\boldsymbol{\varepsilon}_\mathrm{m}$:

$$\varepsilon_j = \Big(\boldsymbol{\beta}_j \boldsymbol{W}_\mathrm{m} - \boldsymbol{w}_j\Big)\boldsymbol{\alpha} + \boldsymbol{\beta}_j \boldsymbol{\varepsilon}_\mathrm{m}, \tag{3.11}$$

where $\boldsymbol{\beta}_j$ and $\boldsymbol{w}_j$ are row vectors from the rows of respectively $\boldsymbol{B}$ and $\boldsymbol{W}$ corresponding to target point $j$. The first term can be recognised as the methodical error,

$$\varepsilon_{j,\mathrm{meth}} = \Big(\boldsymbol{\beta}_j \boldsymbol{W}_\mathrm{m} - \boldsymbol{w}_j\Big)\boldsymbol{\alpha},$$

the second as the error due measurement noise,

$$\varepsilon_{j,\mathrm{noise}} = \boldsymbol{\beta}_j \boldsymbol{\varepsilon}_\mathrm{m}.$$

In literature estimation algorithms are often purely assessed on their methodical error, i.e. the estimator performance in presence of ideal sensor signals. In real systems, however, sensors exhibit sensor noise. Depending on the noise level of the sensors this noise may be an important part of the total estimation error. The expected value of the total estimation error squared is then

$$E\left(\varepsilon_j^2\right) = \boldsymbol{w}_j^\mathsf{T} \boldsymbol{\Sigma}_\alpha \boldsymbol{w}_j - 2\boldsymbol{w}_j^\mathsf{T} \boldsymbol{\Sigma}_\alpha \boldsymbol{W}_\mathrm{m}^\mathsf{T} \boldsymbol{\beta}_j + \boldsymbol{\beta}_j^\mathsf{T} \boldsymbol{W}_\mathrm{m} \boldsymbol{\Sigma}_\alpha \boldsymbol{W}_\mathrm{m}^\mathsf{T} \boldsymbol{\beta}_j + \boldsymbol{\beta}_j^\mathsf{T} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\beta}_j. \tag{3.12}$$

The term $\boldsymbol{\beta}_j^\mathsf{T} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\beta}_j$ corresponds to the expectation value of the error due to measurement noise. Clearly, the larger the elements in $\boldsymbol{\beta}_j$, the larger the influence of the sensor noise.

Finally, it is often useful to express the errors of all target points as a single value. Depending on the implications of an estimation error at a single point, different norms could be chosen, like the 1- (the average), 2- (the root mean square) or $\infty$-norm (the maximum).

## 3.6 Conclusions

This chapter compared different methods for estimation of the deformation field of the wafer chuck. The method is required to use the available foreknowledge regarding the system's mechanics and typical disturbance conditions. The disturbances have a frequency content typically well-below the first dynamic eigenfrequency of the wafer chuck, exciting it primarily quasi-statically.

Because of the quasi-static behaviour of the wafer chuck, the use of signal values of the past does not supply more information than the signal's current values. This justifies the use of a static estimator instead of a more complex dynamic estimator. Shape fitting is an often used static method. In this method, the estimates are linear combinations of the sensor values, according to an estimation matrix $\boldsymbol{B}$.

A basic assumption of shape fitting is that the deformation field can be described well by a limited number of shapes; the amplitude of these shapes is calculated from the sensor values. The number of shapes may be smaller, equal or larger than the number of sensors. To attain a low estimation error it is important to choose appropriate fitting shapes.

In literature different types of shapes are used, like polynomials, dynamic modeshapes and static deformation shapes. The often used dynamic modeshapes contain information on the mass and stiffness distribution, but for describing quasi-static behaviour only information about the stiffness distribution is relevant. However, including the available foreknowledge on the disturbance distribution is not trivial for polynomials and modeshapes.



The use of the shapes in a snapshot matrix allows the inclusion of the available foreknowledge on the location and magnitude distribution of the disturbances in the estimator. Assuming point disturbances with Gaussian distributed magnitudes, the quasi-static deformation shapes resulting from each disturbance can be used to build up the snapshot matrix. The SF estimator can be directly based on the full snapshot matrix (LS technique) or on the most prominent pattern in the snapshot matrix (POM technique). If the matrix captures the possible deformations well, the snapshot matrix can also be used to calculate the expected error of the estimator.

# Chapter 4

# Snapshot matrix based techniques for shape fitting

In the previous chapter, the Shape Fitting (SF) method was selected for estimation of wafer chuck's deformation. It was chosen to use shape fitting based on a snapshot matrix. The snapshot-based approach is promising, as the snapshot matrix allows including the relevant foreknowledge on the system and its typical disturbances. Two snapshot matrix-based methods from literature were introduced: a method that first computes the Proper Orthogonal Modes (POMs) based upon the snapshot matrix (the POM technique) and a method that directly uses the snapshot matrix in combination with Least Squares (LS) fitting (the LS technique). These methods were, to the best of our knowledge, not compared in literature before. This chapter compares the two, both based on their theoretical formulation and based on numerical analysis, in order to find which method is most suitable for deriving the estimator. The numerical experiments focus on the deformation of the wafer chuck due to its main disturbance source, the immersion film.

First, in respectively Section 4.1 and 4.2, estimation using the POM technique and the LS technique is introduced. The two methods are then compared in Section 4.3 based on their theoretic formulation. One of the comparison criteria is the estimation error, which consists of methodical error and error due to sensor noise. It is shown that both methods can be adapted in different ways such that a trade-off can be made between those error sources. In Section 4.4 the techniques are numerically assessed and compared. In a simulation, the SF method is applied to a wafer chuck model with highly simplified boundary conditions. Finally, Section 4.5 concludes that for several reasons the LS technique is more suitable than the POM technique.

## 4.1 Estimation using the Proper Orthogonal Modes (POM technique)

### 4.1.1 Method

Each column of snapshot matrix $\boldsymbol{W}$, representing a snapshot of the system, can be imagined as a point in a $N_\mathrm{P}$-dimensional space. The full snapshot matrix can then be thought of as a point cloud with $N_\mathrm{D}$ points. Intuitively, one would want fitting shapes that capture the most prominent patterns in the snapshot matrix. This can be achieved using the Proper Orthogonal Decomposition (POD). For the given snapshot matrix it finds a set of orthogonal vectors and corresponding singular values. The vector with the highest singular value spans the direction in the $N_\mathrm{P}$ space that minimises – in a least squares sense – its total distance to the points in the point cloud, i.e.



$$\arg\min_{\breve{\boldsymbol{\varphi}}_1} \sum_{i=1}^{N_\mathrm{D}} m_i^2 = \sum_{i=1}^{N_\mathrm{D}} \left[ |\boldsymbol{w}_i|^2 - \left( \frac{\boldsymbol{w}_i^\mathsf{T} \breve{\boldsymbol{\varphi}}_1}{|\breve{\boldsymbol{\varphi}}_1|} \right)^2 \right], \tag{4.1}$$

where $m_i$ is the distance between the $i$th point of the snapshot matrix and the direction span by $\breve{\boldsymbol{\varphi}}_1$. The minimisation problem of Eq. (4.1) is equivalent to the maximisation problem

$$\arg\max_{\breve{\boldsymbol{\varphi}}_1} \sum_{i=1}^{N_\mathrm{D}} l_i^2 = \arg\max_{\breve{\boldsymbol{\varphi}}_1} \sum_{i=1}^{N_\mathrm{D}} \left( \frac{\boldsymbol{w}_i^\mathsf{T} \breve{\boldsymbol{\varphi}}_1}{|\breve{\boldsymbol{\varphi}}_1|} \right)^2, \tag{4.2}$$

where $l_i$ is the length of the orthogonal projection of $\boldsymbol{w}_i$ onto $\breve{\boldsymbol{\varphi}}_1$.

Defining $\lambda_1$ as $\sum_{i=1}^{N_\mathrm{D}} l_i^2$, the solution of this optimisation problem (the full derivation is included in Appendix D) is found using the eigenvalue problem

$$\lambda_n \breve{\boldsymbol{\varphi}}_n = \boldsymbol{W}\boldsymbol{W}^\mathsf{T} \breve{\boldsymbol{\varphi}}_n, \tag{4.3}$$

where $\boldsymbol{W}\boldsymbol{W}^\mathsf{T}$ can be recognised as a multiple of the sample covariance matrix of $\boldsymbol{W}$. The eigenvectors $\breve{\boldsymbol{\varphi}}_n$ and eigenvalues $\lambda_n$ are called, respectively, the Proper Orthogonal Modes (POMs) and the singular values of $\boldsymbol{W}$. The POM corresponding to the largest singular value, $\lambda_1$, is the one that minimises Eq. (4.1). It can be reasoned that the other POMs also minimise distance Eq. (4.1), with the additional constraint that it is orthogonal to the vectors with higher singular values. Eventually, $N_\mathrm{P}$ POMs and singular values can be obtained.

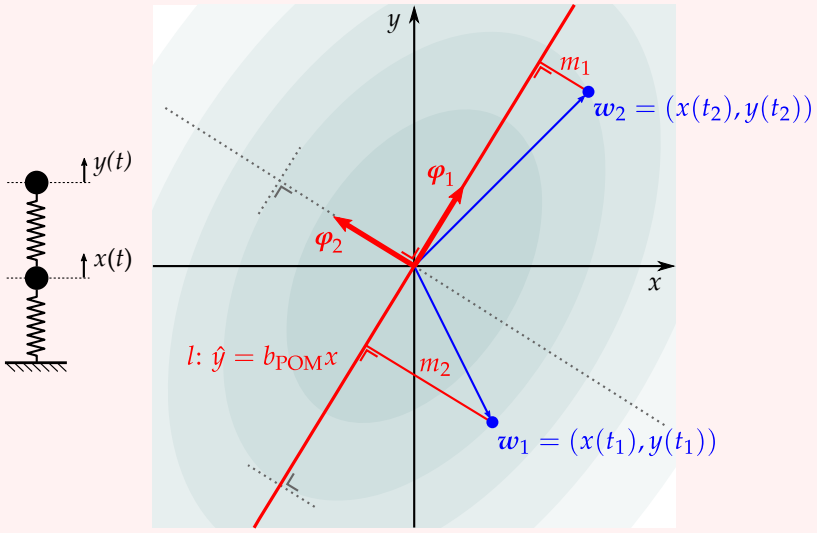
The set of fitting shapes is formed from the POMs with the highest singular values, i.e. the ones that carry the most information about $\boldsymbol{W}$.

$$\breve{\boldsymbol{\Phi}} = \begin{bmatrix} \breve{\boldsymbol{\varphi}}_1 & \cdots & \breve{\boldsymbol{\varphi}}_{N_\mathrm{F}} \end{bmatrix}. \tag{4.4}$$

The number of fitting shapes, $N_\mathrm{F}$, should be chosen such that the estimation error is as low as possible.

Example 4.1 illustrates estimation with the POM technique for a simple two DoFs system.

**Example 4.1: 2 DoFs system (POM technique)**




The figure shows the construction of the fitting shape using POMs for a simple system with two Degrees of Freedom (DoFs). $x$ and $y$ denote the displacements of the two points (DoFs); $x$ belongs to the measurement point and $y$ to the target point, whose displacement must be estimated. The response of the system is assumed to be well-described by two snapshots $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$, together forming snapshot matrix $\boldsymbol{W}$. Thus, it is assumed that the output of the system is a linear combination of shapes $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ with Gaussian distributed weighing factors with equal standard deviation. For this two DoFs system, the SF matrix is simply scalar, $b$, so that $\hat{y} = bx$.

The POMs $\check{\boldsymbol{\varphi}}_1$ and $\check{\boldsymbol{\varphi}}_2$ corresponding to $\boldsymbol{W}$ are shown in the figure. The first POM, $\check{\boldsymbol{\varphi}}_1$, spans the line that minimises $m_1^2 + m_2^2$, the most prominent direction in the snapshot matrix. In this two DoFs case, the direction of POM $\check{\boldsymbol{\varphi}}_2$ is trivial, spanning the least prominent direction in the snapshot matrix. The ellipses show the equidensity contours of the probability density function of the combinations $\boldsymbol{w} = \begin{bmatrix} \boldsymbol{w}_1 & \boldsymbol{w}_2 \end{bmatrix} \begin{Bmatrix} \alpha_1 & \alpha_2 \end{Bmatrix}^\intercal$, where $\alpha_1$ and $\alpha_2$ are standard Gaussian distributed. Note that their semi-axes are spanned by the POMs.

$b_{\text{POM}}$ is found from the direction of most prominent POM, $\check{\boldsymbol{\varphi}}_1$. $b_{\text{POM}}$ can be understood as the slope of the line $\hat{y} = b_{\text{POM}} x$, which is indicated in red.

### 4.1.2 Comparison between proper orthogonal modes and dynamic modeshapes for shape fitting

In many cases, shape fitting is performed using an equal or smaller number of fitting shapes than the number of sensors ($N_M \geq N_F$). In that sense POMs and dynamic modeshapes are used in a highly comparable way, namely to find a set of fitting shapes that leads to an effective estimator.

Feeny and Kapagantu show that for a special class of systems the POMs of a snapshot matrix equal the modeshapes of the system, namely for systems that have a mass matrix that is a multiple of the identity matrix [74]. An example of such a system is a multiple mass-spring system with equal masses. They show that for such a matrix

$$\boldsymbol{W}\boldsymbol{W}^\mathsf{T}\breve{\boldsymbol{\varphi}}_j = \boldsymbol{q}_j^\mathsf{T}\boldsymbol{q}_1\breve{\boldsymbol{\varphi}}_1 + \ldots + \boldsymbol{q}_j^\mathsf{T}\boldsymbol{q}_N\breve{\boldsymbol{\varphi}}_N, \tag{4.5}$$

with $\boldsymbol{q}_i$ a vector with the modal amplitudes of the $i$th mode at time instants $t = t_1, t_2, \ldots, t_{N_D}$. They make use of the fact that for free vibration, the eigenmodes will resonate at different frequencies, so that $\boldsymbol{q}_j^\mathsf{T}\boldsymbol{q}_j \gg \boldsymbol{q}_j^\mathsf{T}\boldsymbol{q}_i$ with $i \neq j$ for a large enough number of snapshots $N_D$. From this and the fact that for such a system the eigenvectors are orthogonal it follows that the eigenvectors and thus the POMs of $\boldsymbol{W}$ are equal to the system's modeshapes.

In a similar fashion it can be shown that the POMs found using quasi-static response – of a system with a unitary mass matrix – to standard Gaussian distributed disturbances at all its inputs equal the system's modeshapes. The quasi-static response ($\Omega = 0$) of the $i$th modal coordinate can using Eq. (2.15) be expressed as

$$\boldsymbol{q}_n^\mathsf{T} = \frac{1}{\Omega_n^2\mu_n}\breve{\boldsymbol{\varphi}}_n^\mathsf{T}\begin{bmatrix} \boldsymbol{f}_1^\mathsf{T} \\ \vdots \\ \boldsymbol{f}_N^\mathsf{T} \end{bmatrix}, \tag{4.6}$$

with $\boldsymbol{f}_i$ a column vector of the forces at the $i$th input point at the at time instants $t = t_1, t_2, \ldots, t_{N_D}$. Thus, the expected value of $\boldsymbol{q}_j^\mathsf{T}\boldsymbol{q}_i$ can be expressed as

$$E\left(\boldsymbol{q}_j^\mathsf{T}\boldsymbol{q}_i\right) = \frac{1}{\Omega_j^2\mu_j}\breve{\boldsymbol{\varphi}}_j^\mathsf{T}E\left(\begin{bmatrix} \boldsymbol{f}_1^\mathsf{T} \\ \vdots \\ \boldsymbol{f}_N^\mathsf{T} \end{bmatrix}\begin{bmatrix} \boldsymbol{f}_1 & \cdots & \boldsymbol{f}_N \end{bmatrix}\right)\breve{\boldsymbol{\varphi}}_i\frac{1}{\Omega_i^2\mu_i}. \tag{4.7}$$

If the disturbances of the different points are indeed standard Gaussian distributed and uncorrelated, $E\left(\boldsymbol{q}_j^\mathsf{T}\boldsymbol{q}_i\right)$ is proportional with $\breve{\boldsymbol{\varphi}}_j^\mathsf{T}\breve{\boldsymbol{\varphi}}_i$ and thus $E\left(\boldsymbol{q}_j^\mathsf{T}\boldsymbol{q}_i\right) = 0$ for $i \neq j$.

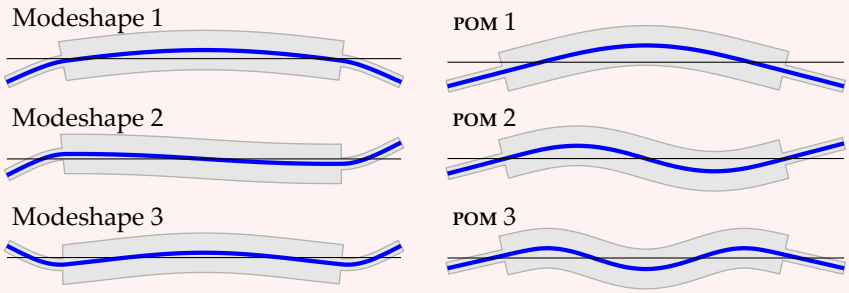
Most systems, however, do not posses a unitary mass matrix, so that the modeshapes are not orthogonal with respect to each other and do not equal the POMs. Furthermore, the choice of equally distributed quasi-static disturbances at all input points does not make sense, as choosing different amplitudes at different points makes it possible to add the foreknowledge on the spatial distribution of the disturbances.

Example 4.2 compares the effectiveness of modeshapes and POMs for fitting beam deformations.

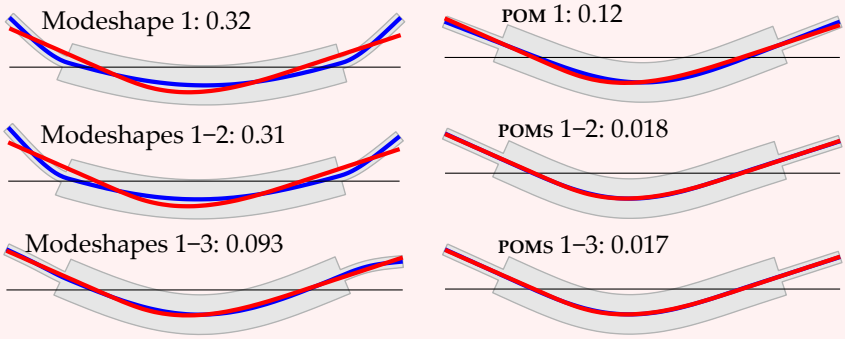
**Example 4.2: Shape fitting with POMs and and dynamic modeshapes**

This example is based on the beam and the snapshot matrix of Example 3.3. The snapshot matrix was build up from the deformation shapes corresponding to disturbances at the middle part of the beam.

The following figure shows the first three modeshapes of the tapered beam. Also the first five POMs, which were calculated using the snapshot matrix, are shown.




The deformation shape corresponding to load case $n = 17$, as shown in Example 3.3, was fitted using different sets of either modeshapes or POMs. The results are plotted in the figure below, along with the normalised RMS residues between the fitted and actual shape. The results illustrate that for accurately fitting this deformation shape with POMs, fewer shapes are needed than when fitting with modeshapes.



## 4.2 Estimation using the full snapshot matrix (LS technique)

The least squares method finds the estimation matrix $\boldsymbol{B}$ that directly minimises the methodical estimation error. This method calculates for each degree of freedom the corresponding optimal coefficients that result in the smallest estimation error, in the least square sense, for all deformation shapes in snapshot matrix $\boldsymbol{W}$. The optimisation for the $j$th degree of freedom reads:

$$\boldsymbol{\beta}_j = \arg\min_{\boldsymbol{\beta}_j} E\left(\varepsilon^2_{j,\mathrm{meth}}\right) = \arg\min_{\boldsymbol{\beta}_j} \left[\left(\boldsymbol{\beta}_j \boldsymbol{W}_\mathrm{m} - \boldsymbol{w}_j\right)\boldsymbol{\alpha}\right]^2, \tag{4.8}$$

where row vectors $\boldsymbol{\beta}_j$, $\boldsymbol{\varepsilon}_j$ and $\boldsymbol{w}_j$ are respectively the $j$th row of the matrices $\boldsymbol{B}$, $\boldsymbol{E}$ and $\boldsymbol{W}$. This problem is clearly a weighted least squares problem, which is known to have the following solution

$$\boldsymbol{\beta}_j = \boldsymbol{w}_j \boldsymbol{\Sigma}_\alpha \boldsymbol{W}_\mathrm{m}^\mathsf{T} \left[\boldsymbol{W}_\mathrm{m} \boldsymbol{\Sigma}_\alpha \boldsymbol{W}_\mathrm{m}^\mathsf{T}\right]^{-1}, \tag{4.9}$$

so that the full matrix $\boldsymbol{B}_\mathrm{LS}$ is given by

$$\begin{aligned} \boldsymbol{B}_\mathrm{LS} &= \begin{bmatrix} \boldsymbol{\beta}_1^\mathsf{T} & \cdots & \boldsymbol{\beta}_{N_\mathrm{P}}^\mathsf{T} \end{bmatrix}^\mathsf{T} \\ &= \boldsymbol{W}\boldsymbol{\Sigma}_\alpha \boldsymbol{W}_\mathrm{m}^\mathsf{T} \left(\boldsymbol{W}_\mathrm{m} \boldsymbol{\Sigma}_\alpha \boldsymbol{W}_\mathrm{m}^\mathsf{T}\right)^{-1}. \end{aligned} \tag{4.10}$$

Note the similarity between this result and Eqs. (3.6) and (3.7). It follows that the LS technique is a special case of shape fitting, in which the set of fitting shapes is formed by snapshot matrix $\boldsymbol{W}$.

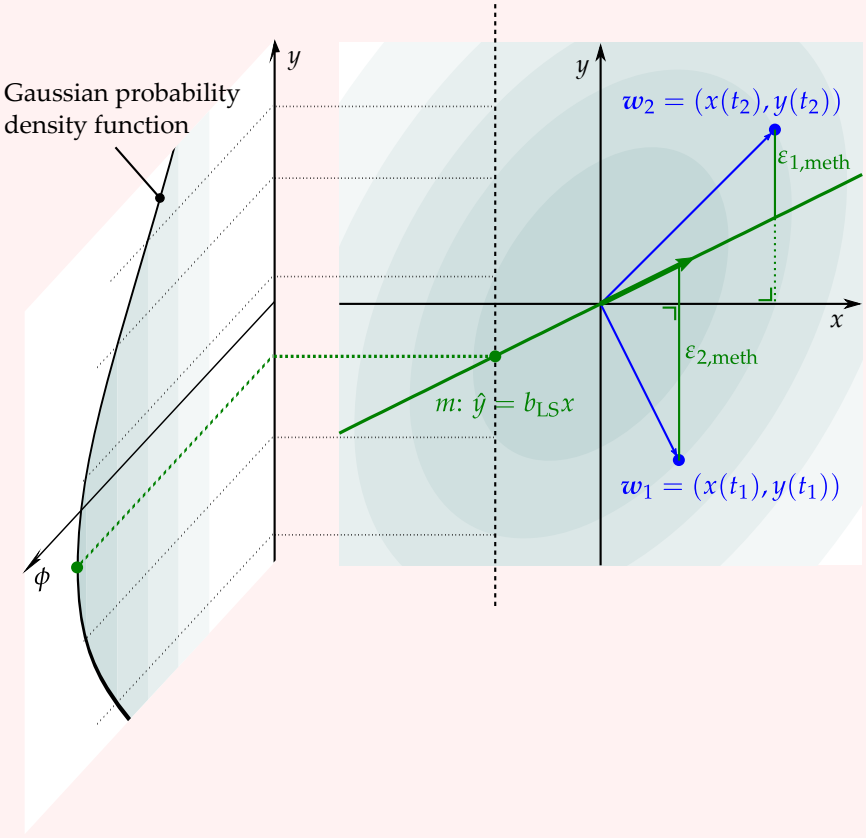
Example 4.3 illustrates estimation with the LS technique for a simple two DoFs system.

## 4.3 Comparison between the POM and the LS technique

### 4.3.1 Influence of the choice of target points

The POM and the LS technique differ fundamentally in the sense that the estimation performance of the POM technique depends on the chosen subset of target points. Typically, the number of target points depends on the application and could be, for example, all points (full field) or just a single point. In any case, one would expect that the estimation of a point does not depend on whether it was calculated as part of the full field or just as a single point. Clearly, the LS technique optimises for each point separately (cf Eq. 4.9). The POM technique, however, makes use of a set of POMs in matrix $\boldsymbol{\Phi}_\mathrm{POM}$ and their orthogonality depends on all chosen points. The question arises how the choice of the target point subset influences the results of the POM technique. Therefore, different sets of target coordinates will be chosen in the numerical study.

**Example 4.3: 2 DoFs system (LS technique)**




For the two DoFs system of Example 3.3 with the same snapshot matrix $\boldsymbol{W} = \begin{bmatrix} w_1 & w_2 \end{bmatrix}$, the optimal estimation parameter $b$ can be calculated with the LS technique. The above figure shows the construction of the line $m$ with slope $b_{\text{LS}}$ that leads to optimal estimation. This line directly minimises estimation error $\varepsilon_1^2 + \varepsilon_2^2$. Unlike the line found with the POM technique, line $m$ does not span the longest semi-axis of the equidensity contour ellipses, but intersects them in their left- and rightmost points. Thus for any given measurement $x$ the LS technique selects the estimate $\hat{y}$ that has the highest probability for the given measurement $x$, as illustrated in the left inset.

### 4.3.2 Methodical error

The POM technique uses POMs to extract a set of the most prominent shapes from snapshot matrix $\boldsymbol{W}$. This set is then used for shape fitting; the less prominent shapes are disregarded. When those less prominent shapes are present in the shape of the field, they will not be recognised as such, but wrongly recognised as contributions from the set of most prominent shapes. This so-called spill-over effect causes the methodical error of the POM technique. The LS technique, however, directly optimises for the smallest methodical error.

### 4.3.3 Error due to sensor noise

Both the POM and the LS technique use inversions that are typically close to singular, meaning that it is difficult to distinguish the fitting shapes from their values at the sensors. An increased effort needed for distinguishing the fitting shapes manifests itself in larger values in matrix $\boldsymbol{B}$. Those large values will, according to Eq. (3.12), amplify the sensor noise so that the estimation error due to sensor noise increases. The POM technique can deal in a natural way with this problem. By reducing the number of selected POMs, $N_\mathrm{F}$, those can be distinguished more easily, so that the estimation error due to sensor noise decreases. This is at the cost of an increased methodical error, because if less modes are taken into account, the methodical error increases. This means that there is an inherent trade-off between the error due to sensor noise and the methodical error. According to [75], the optimal truncation can be found by testing all $N_\mathrm{F} \in \{1 \dots N_\mathrm{P}\}$ and selecting the one leading to the smallest total estimation error.

The choice of a lower truncation number, $N_\mathrm{F}$, can be regarded as a way to regularise the inversion problem. The LS technique, as formulated by [71], does originally not allow for regularisation, but may be regularised in a comparable way as the POM technique. Instead of inverting $\boldsymbol{W}_\mathrm{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\mathrm{m}^\mathsf{T}$ itself, a lower rank approximation of the inverse is calculated that has reduced information about $\boldsymbol{W}$ [76]. If $\boldsymbol{W}_\mathrm{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\mathrm{m}^\mathsf{T}$ is expressed as a singular value decomposition:

$$\boldsymbol{W}_\mathrm{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\mathrm{m}^\mathsf{T} = \boldsymbol{U}\,\mathrm{diag}\{\ \sigma_1 \quad \cdots \quad \sigma_{N_\mathrm{M}}\ \}\,\boldsymbol{V}^\mathsf{T}, \tag{4.11}$$

its inverse can be expressed as

$$\left(\boldsymbol{W}_\mathrm{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\mathrm{m}^\mathsf{T}\right)^{-1} = \boldsymbol{V}\,\mathrm{diag}\left\{\ \sigma_1^{-1} \quad \cdots \quad \sigma_{N_\mathrm{M}}^{-1}\ \right\}\boldsymbol{U}^\mathsf{T}. \tag{4.12}$$

For the lower rank approximation of the inverse, the smallest singular values are replaced by zero:

$$\left(\boldsymbol{W}_\mathrm{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\mathrm{m}^\mathsf{T}\right)^{-1} \simeq \boldsymbol{V}\,\mathrm{diag}\left\{\ \sigma_1^{-1} \quad \cdots \quad \sigma_{N_\mathrm{T}}^{-1} \quad \mathbf{0}\ \right\}\boldsymbol{U}^\mathsf{T}. \tag{4.13}$$

The amount of regularisation depends on the choice $N_\mathrm{T}$. The optimal truncation can be found by testing all $N_\mathrm{T} \in \{1 \dots N_\mathrm{M}\}$ and selecting the one resulting in the smallest total error.

### 4.3.4 Least squares solution for least total error

The previous section showed that both the POM and LS technique allow for a regularisation such that the total estimation error is a balance between methodical error and noise error. This in itself is, however, no guarantee for optimality. Therefore, it is proposed to directly optimise for a minimum total estimation error (Eq. 3.12):

$$\begin{aligned}\nabla_{\boldsymbol{\beta}_j} E\left(\varepsilon^2_{j,\text{total}}\right) &= \nabla_{\boldsymbol{\beta}_j} E\Big(\Big(\boldsymbol{w}_j^\intercal\boldsymbol{\alpha} - \boldsymbol{\beta}\boldsymbol{W}_\text{m}\boldsymbol{\alpha} - \boldsymbol{\beta}_j\boldsymbol{\varepsilon}_\text{m}\Big)\Big(\boldsymbol{\alpha}^\intercal\boldsymbol{w}_j - \boldsymbol{\alpha}^\intercal\boldsymbol{W}_\text{m}^\intercal\boldsymbol{\beta}_j^\intercal - \boldsymbol{\varepsilon}_\text{m}^\intercal\boldsymbol{\beta}_j^\intercal\Big)\Big)\\ &= \nabla_{\boldsymbol{\beta}_j}\Big(\boldsymbol{w}_j^\intercal\boldsymbol{\Sigma}_\alpha\boldsymbol{w}_j - 2\boldsymbol{w}_j^\intercal\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\text{m}^\intercal\boldsymbol{\beta}_j^\intercal + \boldsymbol{\beta}_j\boldsymbol{W}_\text{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\text{m}^\intercal\boldsymbol{\beta}_j^\intercal + \boldsymbol{\beta}^\intercal\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\beta}_j^\intercal\Big)\\ &= -2\boldsymbol{w}_j\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\text{m}^\intercal + 2\boldsymbol{W}_\text{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\text{m}^\intercal\boldsymbol{\beta}_j^\intercal + 2\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\beta}_j^\intercal = \boldsymbol{0}, \qquad (4.14)\end{aligned}$$

so that

$$\boldsymbol{\beta}_j = \boldsymbol{w}_j\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\text{m}^\intercal\big(\boldsymbol{W}_\text{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\text{m}^\intercal + \boldsymbol{\Sigma}_\varepsilon\big)^{-1}. \qquad (4.15)$$

The full matrix $\boldsymbol{B}_\text{LS}$ is then defined as follows

$$\begin{aligned}\boldsymbol{B}_\text{LS} &= \begin{bmatrix}\boldsymbol{\beta}_1^\intercal & \cdots & \boldsymbol{\beta}_{N_\text{P}}^\intercal\end{bmatrix}^\intercal\\ &= \boldsymbol{W}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\text{m}^\intercal\big(\boldsymbol{W}_\text{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\text{m}^\intercal + \boldsymbol{\Sigma}_\varepsilon\big)^{-1}. \qquad (4.16)\end{aligned}$$

Note that compared to Eq. (4.10) a term $\boldsymbol{\Sigma}_\varepsilon$ is added to the inverse, which acts as a regularisation by bringing the matrix to be inverted closer to a diagonal matrix. The solution can be recognised as a form of generalised Tikhonov regularisation [77].

Although Mainçon has a different mathematical formulation and does not use a snapshot matrix [12], his method is, in basis, comparable to the here presented method, as both make use of the foreknowledge on the disturbances and sensor noise in the form of covariance matrices. Indeed, for a system that has no Rigid Body (RB) modes the methods yield equal results.

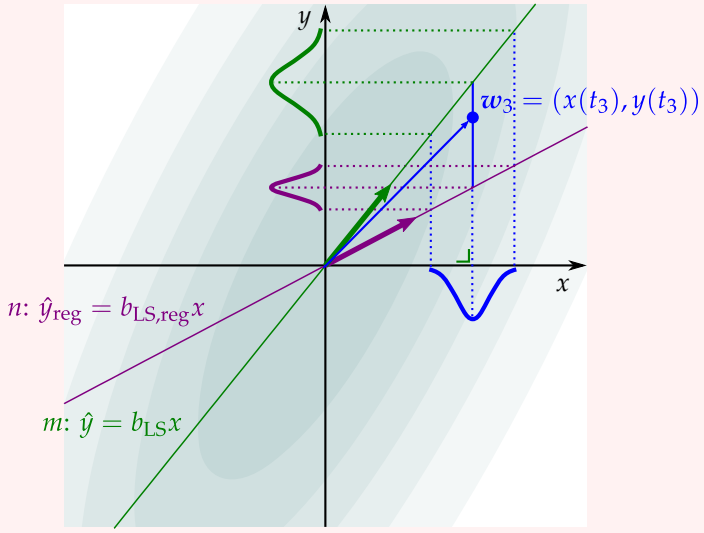
### 4.3.5 Time complexity

Van der Sanden and Philips [78], aiming at extracting the most relevant shapes from a snapshot matrix of a system's temperature responses for use in sensor placement, stated that "*Performing the decomposition using the POD algorithm is very hard if not practically impossible in terms of computation time and memory requirements if the temperature identification matrix is large, as in our case* $\approx$ *50000* $\times$ *3500.*" They used a snapshot matrix with $N_\text{P} = 50000$ target points and $N_\text{D} = 3500$ snapshots to finally estimate using $N_\text{M} = 6$ sensors, probably using shape fitting.

The authors are, presumably, referring to the time complexity of calculating the full POD. The POD is in practice calculated using the Singular Value Decomposition (SVD). Algorithms for obtaining the full SVD have typically a time complexity of $O\big(k_1 N_\text{P}^2 N_\text{D} + k_2 13 N_\text{D}^3\big)$ [79], which grows fast for increasing dimensions of the snapshot matrix. However, as only a small subset of the most prominent POMs is needed for estimation, faster algorithms can be utilised.

**Example 4.4: Regularisation of the snapshot-based LS estimator**

This example is based on the two DoFs system of Example 4.3. The slope $b_{\text{LS}}$ is calculated with the LS technique and is optimal for the given snapshot matrix $\boldsymbol{W} = \begin{bmatrix} \boldsymbol{w}_1 & \boldsymbol{w}_2 \end{bmatrix}$. If the measurement of $x$ contains, however, sensor noise $\boldsymbol{\varepsilon}$, the sensor noise multiplied by $b$ adds up to the estimate $\hat{y}$ (figure below). Thus, the standard deviation of the error on $\hat{y}$ due to noise is $\sigma_{2,\text{noise}} = b_{\text{LS}}\sigma_{\varepsilon}$. To decrease the influence of the sensor noise on the estimate, the slope $b_{\text{LS}}$ should be deceased. This leads to a biased estimator with a higher methodical error. The regularised least squares technique calculates the slope $b_{\text{LS,reg}}$ that leads to the optimal trade-off between methodical and sensor noise.



For example, Matlab's iterative algorithm svds calculates the 15 most prominent POMs from a 50000×3500 snapshot matrix with standard Gaussian distributed entries in 1400 seconds on a laptop with a clock speed of 1.6 GHz.

Both the POM technique and the LS technique calculate the estimator using least squares. The LS technique has typically a large number of fitting shapes, as it fits all snapshots in the snapshot matrix. But also for the LS technique the least squares calculation is relatively fast. Calculation of $\boldsymbol{B} = \boldsymbol{\Phi}\boldsymbol{\Phi}_{\mathrm{m}}^{\mathsf{T}}\left(\boldsymbol{\Phi}_{\mathrm{m}}\boldsymbol{\Phi}_{\mathrm{m}}^{\mathsf{T}}\right)^{-1}$ with $N_{\mathrm{P}} = 50000$, $N_{\mathrm{D}} = 3500$ and $N_{\mathrm{M}} = 6$ takes 0.4 seconds on a laptop with a clock speed of 1.6 GHz. Thus the time complexity of the POM technique, including the calculation of the relevant POMs, is much larger than the one of the LS technique, although by no means prohibitive for off-line calculation of the estimator.

## 4.4 Numerical study of the snapshot techniques in the wafer chuck



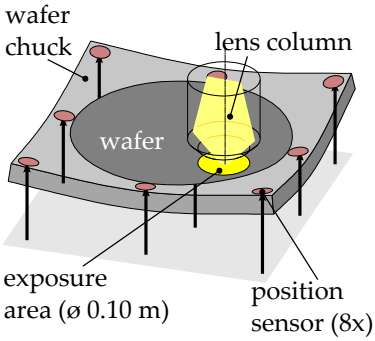
In the following, a numerical study will be presented based on the estimation problem of a wafer chuck's deformation. The geometries, boundary conditions and load cases of an actual wafer chuck are highly complex. These were simplified to keep the study tractable and to not distract from the results regarding estimator performance.

### 4.4.1 Wafer chuck model

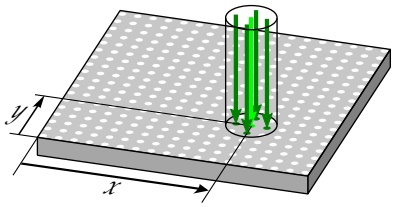
The core of the model is a finite element model of the wafer chuck. The wafer chuck and the wafer on top of it are modelled as a single entity, a square zerodur plate of 0.55 m × 0.55 m × 0.014 m made out of plate finite elements with free boundary conditions (Figure 4.1a). A wafer chucks is kept in place by a control system with sensors and actuators. This was modelled by selecting four actuator positions and choosing actuation force magnitudes such that the net forces and moments, including any applied external loads, on the plate stay zero, see Figure 4.1c. This can be regarded as an ideal control action that fully prevents RB displacement. The hand-picked sensor configuration of Section 2.4.3 with eight position sensors placed at the corners and the centres of the edges was used (Figure 4.1a). This sensor configuration is, presumably, sub-optimal. Approaches to sensor placement for the estimation problem are presented in [75,80]. Automated sensor placement is studied in Chapter 6.

### 4.4.2 Load cases

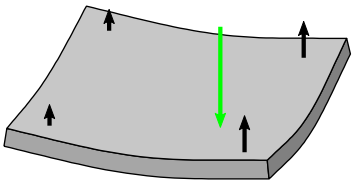
In immersion lithography a fluid layer is applied between the lens and the wafer to the numerical aperture of the lens column. Turbulence in this fluid layer lead to disturbance forces that act onto the wafer chuck. Actually, those forces account for the major part of the wafer chuck's low frequency disturbances, so that the disturbance forces may be assumed to act solely in the exposure area, a circle with a diameter of 0.10 m.




(a) The wafer chuck including the wafer were modelled as a single plate with free boundary conditions. The eight red circles indicate the chosen sensor positions.




(b) The wafer chuck's top surface was discretised in a grid of 51×51 points. Disturbances were assumed to only act at the points in the exposure area. For clarity a smaller discretisation grid is shown, with only 5 disturbance locations in the exposure area; in the numerical experiments 68 disturbance points per lens position $(x, y)$ were used.



(c) Example of a load case corresponding to one disturbance position. A load case was composed from a disturbance force (the green arrow) and four actuator forces (the black arrows) that mimic an ideal controller action by balancing the external forces.

Figure 4.1: Modelling steps for the numerical study of the wafer chuck.

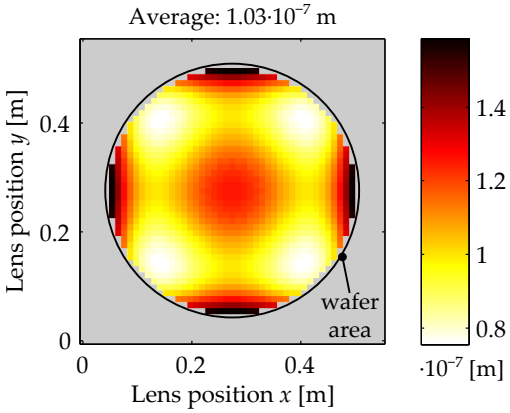



Figure 4.2: Average deformation of the wafer chuck over the exposure area as function of the lens position $(x, y)$.



The wafer chuck's top surface was descretised in a grid of 51×51 points (Figure 4.1b). This discretisation was used both to describe the possible different lens positions $(x, y)$ and the possible positions of disturbance forces. In total 1313 lens positions were studied, i.e. all positions in the 450 mm diameter wafer area. Per lens location the 68 points in the 100 mm diameter exposure area were used as the disturbance locations.

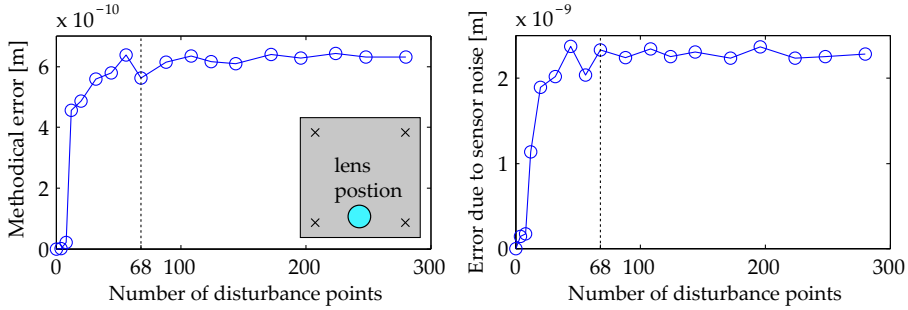
A simple disturbance model was adopted based on the assumption that the total surface integral of the disturbance pressure over the exposure area is 1 N RMS. It was assumed that the disturbances are not spatially correlated, so that the forces at the disturbance positions are normally distributed with a zero mean and a variance of $1/68 \cdot 1\ \mathrm{N}^2$. All disturbance forces combined with the corresponding actuator forces lead to a total of 68 load cases per lens position.

### 4.4.3 Snapshot matrix construction and estimator calculation

The deformed shapes corresponding to the load cases were calculated using the finite element model. As the disturbance forces are assumed to be typically low frequency and well below the lowest eigenfrequency of the wafer chuck, static analysis sufficed. Per lens position the $N_F = 68$ deformation shapes corresponding to the relevant load cases were combined into a $51^2 \times 68$ snapshot matrix $\boldsymbol{W}$. Figure 4.2 shows the average deformation over the exposure area as function of the lens's position $(x, y)$. The weighing of the load cases was uniform, so that $\sigma_q^2 = 1/68$. A RMS noise level of $0.1 \cdot 10^{-9}$ m was assumed for all sensors, so that $\sigma_\varepsilon^2 = 1.0 \cdot 10^{-20}\ \mathrm{m}^2$. Using these choices, matrix $\boldsymbol{B}$ was calculated for any of the following scenarios based on the different estimation technique.

1. POM *(full field):* The full field of target points were included in the calculation of the POD.

2. POM *(exposure area):* Only the target points corresponding to the exposure area were included in the calculation of the POD.




(a) Methodical error.

(b) Error due to sensor noise.

Figure 4.3: The methodical error and the error due to sensor noise as function of the number of disturbance points used in the analysis. The results converge for a large enough number of disturbance points. For the grid of 68 disturbance points, as used in the rest of work, the errors have nearly converged.



3. LS *(standard):* The LS technique was applied according to Eq. (4.10).
4. LS *(reduced basis):* The inversion of Eq. (4.10) was calculated using the singular value decomposition with a limited number of singular values $N_T$, according to Eq. (4.13), such that the estimation error for the respective lens position was minimised.
5. LS *(regularised):* The regularised LS technique was applied.

### 4.4.4 Estimation error

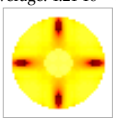
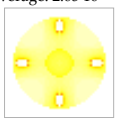
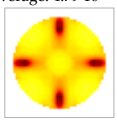
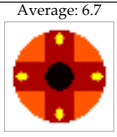
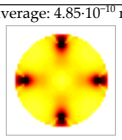
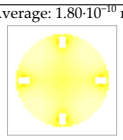
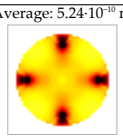
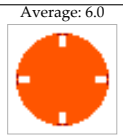
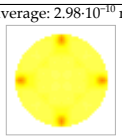
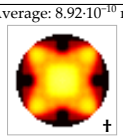
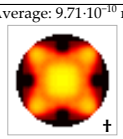
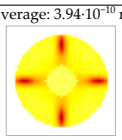
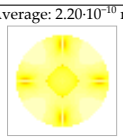
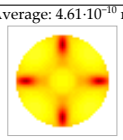
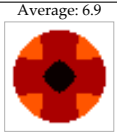
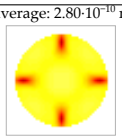
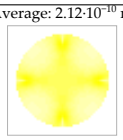
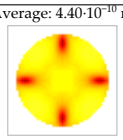
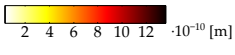
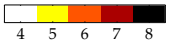
After applying the different estimation techniques, the standard deviation of the estimation error at each target point in the exposure area was calculated. To obtain a single indicator of the error for each lens position, the standard deviations of the target points were averaged. Those averaged were stacked according to their lens positions $(x, y)$ in an error matrix $\boldsymbol{E}_{\text{total}}$. In a comparable way $\boldsymbol{E}_{\text{meth}}$ (methodical estimation error) and $\boldsymbol{E}_{\text{noise}}$ (estimation error due to noise) were found.

To validate the convergence of the analysis, for a single, arbitrary lens location the simulation was carried out multiple times. The number of disturbance and target points below the lens was varied between 1 and 280. Figure shows the results in terms of the methodical error and the error due to sensor noise as a function of the number of points in the lens area. The dotted lines indicate 68 points, as was used in the simulations in this chapter. The figures show that both errors stabilise for a large enough number of points. For 68 points, the method is nearly stabilised.

### 4.4.5 Estimation results and discussion

Table 4.1 shows the estimator performance as function of the lens position for the different scenarios. The left three columns show the error contributions and

Table 4.1: Estimator performance as function of the lens position for different POM and LS based scenarios. † Note that $\boldsymbol{E}_{\text{noise}}$and $\boldsymbol{E}_{\text{total}}$ for the standard LS technique contain relatively high values; those saturate on the given colour scale.

| Scenario | Methodical estimation error ($\boldsymbol{E}_{\text{meth}}$) | Estimation error due to noise ($\boldsymbol{E}_{\text{noise}}$) | Total estimation error ($\boldsymbol{E}_{\text{total}}$) | Truncation number ($N_{\text{POM}}$ or $N_{\text{T}}$) |
|---|---|---|---|---|
| POM (full field) | Average: $4.21 \cdot 10^{-10}$ m  | Average: $2.05 \cdot 10^{-10}$ m  | Average: $4.79 \cdot 10^{-10}$ m  | Average: 6.7  |
| POM (exposure area) | Average: $4.85 \cdot 10^{-10}$ m  | Average: $1.80 \cdot 10^{-10}$ m  | Average: $5.24 \cdot 10^{-10}$ m  | Average: 6.0  |
| LS (standard) | Average: $2.98 \cdot 10^{-10}$ m  | Average: $8.92 \cdot 10^{-10}$ m  † | Average: $9.71 \cdot 10^{-10}$ m  † | *not applicable* |
| LS (truncated) | Average: $3.94 \cdot 10^{-10}$ m  | Average: $2.20 \cdot 10^{-10}$ m  | Average: $4.61 \cdot 10^{-10}$ m  | Average: 6.9  |
| LS (regularised) | Average: $2.80 \cdot 10^{-10}$ m  | Average: $2.12 \cdot 10^{-10}$ m  | Average: $4.40 \cdot 10^{-10}$ m  | *not applicable* |
| |   | | |   |

the total error. The fourth column shows for the POM scenarios the number of POMs that led to the lowest estimation error and for the truncated LS scenario the number of singular values that led to the lowest estimation error.

From a comparison between the plots of the truncation number and the plots of the estimation error the effect of the truncation can be identified. Passing a border to a lower truncation number, the methodical error decreases, whereas the error due to noise increases. This shows the regularising effects of using a smaller set of POMs or singular vectors.

Then, we compare the results of the scenarios. First of all, the POM scenarios show different results, which is expected, as the estimation depends on the choice of the target point subset. The POM for the subset of points at the exposure area performs less well than the POM for the full field. This is remarkable, as the smaller subset seems to be more tailored to the actual exposure area of interest. This may be explained by the fact that the full field contains more information of the sensor points because of the neighbouring points that are also considered. For the POM of the exposure area solely the actual 8 measurement points are considered, so that less sensor information is preserved in the calculation of the POMs.

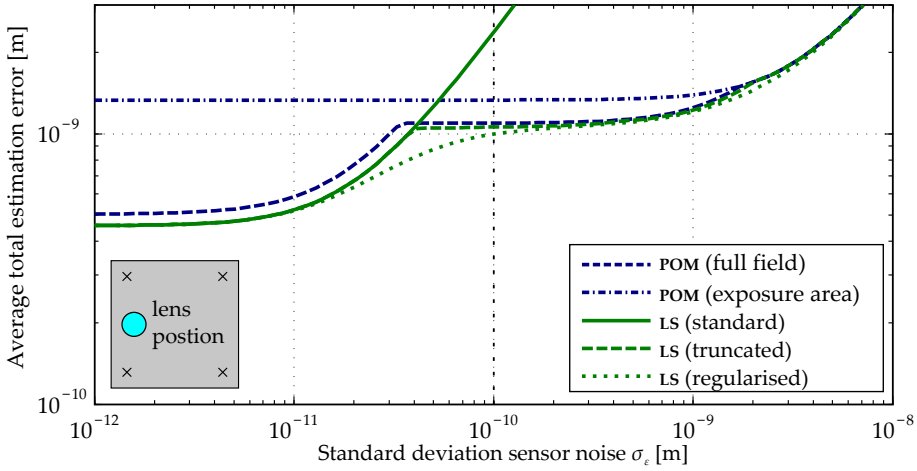


The standard LS scenario leads, as expected, to the lowest methodical estimation error. Noise, on the other hand, leads to large errors, so that the total estimation error is relatively large. This is caused by the lack of regularisation of the estimator. The truncated LS scenario is better regularised; it leads to somewhat larger methodical errors but a considerably smaller noise influence and thus a smaller total estimation error. The total estimation error is also smaller than for the POM scenarios. Finally, it is observed that the regularised LS scenario, as expected from theory, leads to the smallest total estimation error by adequately balancing between methodical error and error due to noise.

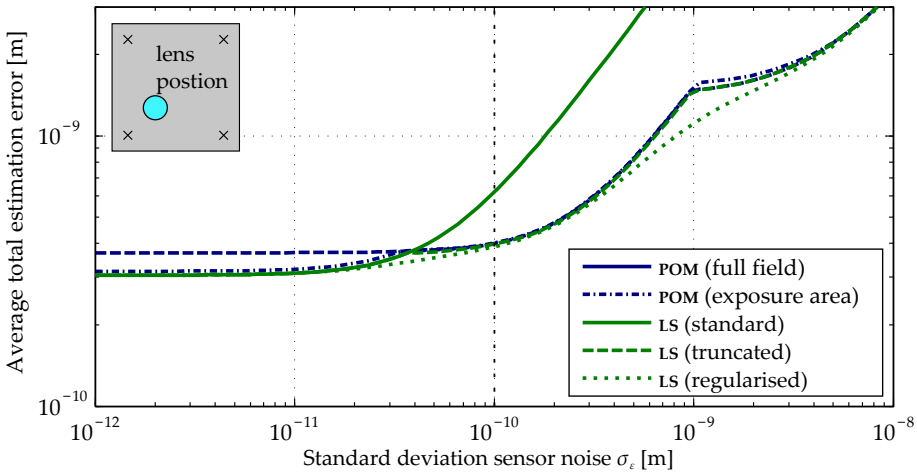
The performance of the different estimation scenarios depends on the noise level. Figure 4.4 shows for two lens positions the average error in the exposure area as function of the sensor noise. For low noise, the total error is dominated by methodical error and stays constant; for high noise the total error grows proportionally to the sensor noise, which is then dominating. The full field POM and truncated LS show clearly the effect of the regularisation via truncation. At certain points the influence of sensor noise is such large that one step further truncation leads to better results. The methodical error is then again limiting. The regularised LS scenario follows the other LS scenarios smoothly, always with the lowest estimation error.

In the preceding, the snapshot matrix $\boldsymbol{W}$ was perfect, in the sense that it did not contain noise and fully reflected the behaviour of the actual system. Indeed, if the snapshot matrix is generated using measurements, noise might influence the estimation's performance. This effect must be reduced by acquiring enough samples to average out the effect of the sensor noise. If the sensors that are used for acquiring the snapshot-matrix are the same as the ones that are used for estimation, or if they have at least the same noise level, this would directly lead to the desired regularisation, as

$$E((\boldsymbol{W}_\mathrm{m}\boldsymbol{\alpha}+\boldsymbol{\varepsilon})(\boldsymbol{W}_\mathrm{m}\boldsymbol{\alpha}+\boldsymbol{\varepsilon})^\mathsf{T}) = E\big(\boldsymbol{W}_\mathrm{m}\boldsymbol{\alpha}\boldsymbol{\alpha}^\mathsf{T}\boldsymbol{W}_\mathrm{m}^\mathsf{T}+\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\mathsf{T}\big) = \boldsymbol{W}_\mathrm{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\mathrm{m}^\mathsf{T}+\boldsymbol{\Sigma}_\varepsilon. \quad (4.17)$$




(a) Lens position 1.




(b) Lens position 2.

Figure 4.4: Estimation error as function of the noise level of the sensors for the five estimation techniques. The error is the average over the exposure area for the lens position indicated by the blue circle in the inset. The noise level of $0.1 \cdot 10^{-9}$ m, as used for the results in Table 4.1, is indicated in the plots.

If the snapshot matrix is obtained using a model, like in the numerical study, it does not suffer from measurement noise, but modelling inaccuracies may be a problem. It is, therefore, important that the model accurately reflects the actual system, which may be achieved through grey box identification methods.

## 4.5 Conclusions

This chapter compared the POM and the LS technique, which are used for the calculation of the SF estimator. A fundamental difference between the POM and LS technique is that the latter optimises for each target point separately, whereas the actual estimation performance of the POM technique depends on the other included target points. Therefore, the selection of output points in the POM is not trivial, as was illustrated by the numerical study.



Both POM and LS allow including foreknowledge on the problem's physics and typical disturbances, which is essential for optimising estimation performance. This foreknowledge is included via a snapshot matrix of the system's output response to typical, Gaussian disturbances. The POM technique extracts the most prominent shapes from the snapshot matrix and uses those for performing shape fitting. The less prominent patterns may, however, cause spillover, leading to the methodical estimation error. The LS technique uses the full snapshot matrix for shape fitting and optimises, by definition, for minimum methodical error.

Not only the methodical error but also sensor noise may lead to a significant estimation error. Both the POM and the LS techniques allow for regularisation in order to reduce the influence of sensor noise. This can in the POM technique be achieved by truncation of the number of prominent POMS and in the LS technique by truncation of the number of singular vectors. Truncation, however, leads to a loss of information about the original snapshot matrix resulting in a higher methodical error. There is a trade-off between methodical error and noise error in order to attain minimum total estimation error. It is proposed to use a regularised LS technique, which directly optimises for lowest total estimation error.

# Chapter 5

# Shape fitting in presence of rigid body dynamics

The formulations of Shape Fitting (SF) in literature do not account for the Rigid Body (RB) motion of a system. For this reason, the SF method was in the previous chapter applied to a wafer chuck model with highly simplified boundary conditions. An ideal control action was assumed, so that the wafer chuck did not experience RB motion.

In reality, the wafer chuck experiences RB motion, as the control action is not ideal. The controller typically has a bandwidth below the first resonance frequency of the mechanical structure, leading to dynamic behaviour in the frequency region that was previously considered quasi-static. The SF method needs to be adapted to deal with RB motion. This chapter further develops the SF method to make it applicable to dynamic systems with RB motion.

Section 5.1 shows to what extent a system with RB motion and a position controller behaves quasi-statically. Two approaches for shape fitting are proposed and one of them is selected for further study. Section 5.2 introduces how the RB and the flexible modes can be decoupled, yielding a flexible residue that can be used for shape fitting. Section 5.3 describes how the snapshot matrix-based Least Squares (LS) technique can be adapted, to also function optimally in the presence of dynamic effects. Section 5.4 presents numerical results of a wafer chuck model with a RB controller. The results show shape fitting is still effective for systems that do not behave solely quasi-static. Section 5.5 introduces a hybrid estimation method based on a both SF estimation and double integration of acceleration.

## 5.1 Shape fitting in presence of position control

Shape fitting is normally applied to systems that purely undergo deformation and no RB motion, due to their stiff connection to the fixed world. The wafer chuck, however, is not connected to the outside world in a passive way, but actively using a position controller. Section 2.1.3 introduced the position control of the wafer chuck. The position of the wafer chuck with respect to the

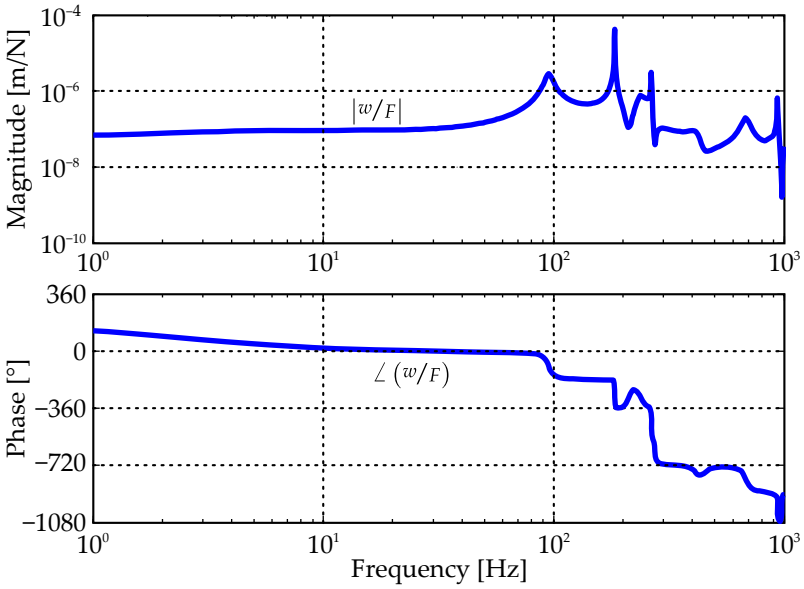


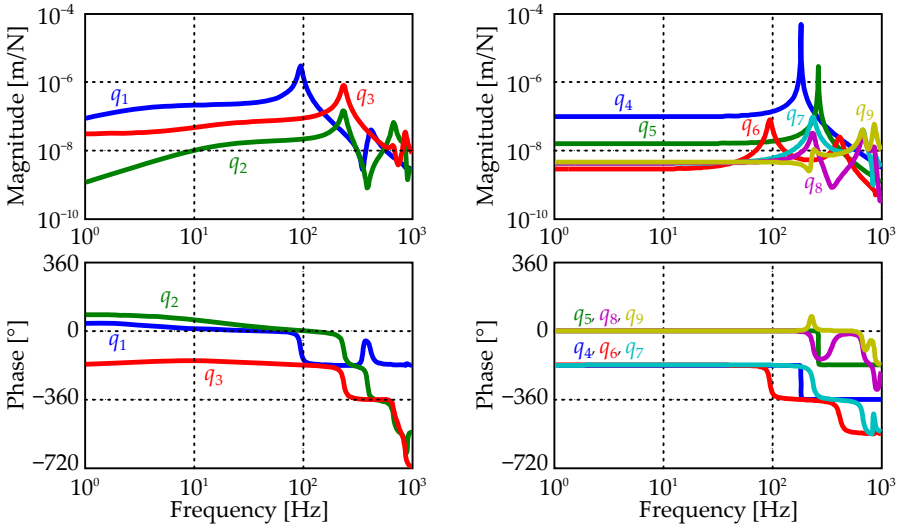

Figure 5.1: The closed-loop transfer function from force to displacement of a fully free plate with RB control. The controller bandwidth is about 90 Hz. The system does not behave quasi-statically below its first eigenfrequency, as the magnitude and phase of its response are not constant.

metrology frame is controlled by controller $\boldsymbol{C}(s)$. An ideal controller with infinite bandwidth would respond instantaneously to the disturbances, such that the net forces and moments on the wafer chuck are fully in equilibrium with the disturbances. Thus, the wafer chuck would not experience RB motion. Chapters 3 and 4 assumed such an ideal controller.

The controller has in reality, however, a limited bandwidth, leading to RB motion and dynamic effects in the frequency band below the wafer chuck's lowest eigenfrequency, which was in the previous presumed quasi-static. Figure 5.1 illustrates this effect by the transfer function of a free-floating plate that is suspended using a RB controller. The controller bandwidth is around 90 Hz, leading to a non-constant magnitude and phase below the first eigenfrequency.

To be able to apply the traditional shape fitting to systems that have RB motion, different approaches are possible. In the following, two approaches are introduced for shape fitting of the actively suspended wafer chuck.

The first approach makes use the knowledge of the actuation forces for the estimation and splits the response due to the known actuation and the unknown disturbance forces. In the other approach the wafer chuck and its controller are treated as a black box system, of which only the outputs are used. The SF estimator is then only fed with these outputs, i.e. the measurements of the displacement of the wafer chuck or related measurands.




(a) Contributions of the RB eigenmodes (translation in $z$-direction ($q_1$), rotation around $x$- ($q_2$) and $y$-axis ($q_3$)).

(b) Contributions of the first six flexible eigenmodes.

Figure 5.2: The contribution of the first nine eigenmodes to the closed-loop transfer function of the RB controlled plate (Figure 5.1). In contrast to the RB eigenmodes, the flexible eigenmodes behave approximately quasi-static for frequencies up to the controller bandwidth (90 Hz).

### 5.1.1 Black box approach

The wafer chuck and its controller can be regarded as a single system with disturbance inputs and displacement outputs. The controller forces are in that case not used, as they are considered internal forces, hence the name 'black box' approach.

Figure 5.2 shows the modal response of a plate with closed-loop RB control. The RB coordinates ($z$, $\theta_x$ and $\theta_y$, Figure 5.2a) behave dynamically, as they do not show a flat magnitude or a flat phase response. The flexible eigenmodes (Figure 5.2b), however, do show approximately quasi-static behaviour up to the controller's bandwidth around 90 Hz. Thus, the response of the system up to the controller bandwidth is a combined effect of RB *dynamics* and *flexible quasi-statics* [81]. Around and above the controller bandwidth, the response of the flexible eigenmodes behaves dynamically.

### 5.1.2 Split approach



The split approach divides the closed-loop response of the wafer chuck to its disturbances into two parts, namely, the part that is directly caused by the disturbances and the part caused by the actuator forces, which are known. The following shows that it is indeed possible to divide the response into those parts.

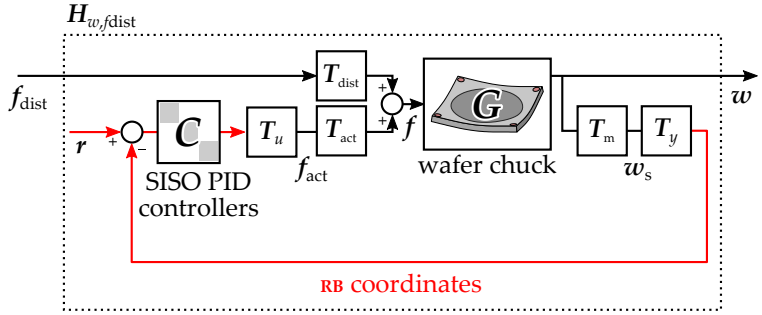
Figure 5.3a shows a block diagram of the system's closed-loop transfer function from the disturbance forces, $\boldsymbol{f}_{\text{dist}}$, to the system's displacements, $\boldsymbol{w}$. This figure is similar to Figure 2.3a, with $\boldsymbol{G}(s)$ the matrix of the wafer chuck's transfer functions from the external forces to the displacements, $\boldsymbol{w}$. The feed-forward path has been omitted to simplify the figure. The matrices $\boldsymbol{T}_{\text{act}}$, $\boldsymbol{T}_{\text{dist}}$ and $\boldsymbol{T}_{\text{m}}$ that select the relevant inputs and outputs of $\boldsymbol{G}$, are shown explicitly. Matrices $\boldsymbol{T}_{\text{act}}$ and $\boldsymbol{T}_{\text{dist}}$ are defined such that $\boldsymbol{G}\boldsymbol{T}_{\text{act}}$ and $\boldsymbol{G}\boldsymbol{T}_{\text{dist}}$ describe the transfer functions of respectively the actuator and the disturbance forces to $\boldsymbol{w}$. In a similar way, matrix $\boldsymbol{T}_{\text{m}}$ selects the outputs of $\boldsymbol{G}$ corresponding to the sensors. Like in Figure 2.3a, Matrix $\boldsymbol{T}_u$ translates the forces in RB coordinates into the actuator forces, $\boldsymbol{T}_y$ translates the displacement at the sensors into RB displacements. The closed-loop transfer function from $\boldsymbol{f}_{\text{dist}}$ to $\boldsymbol{w}$ can be written as:

$$\boldsymbol{H}_{w,f\text{dist}}(s) = \boldsymbol{G}\left(\boldsymbol{I} + \boldsymbol{T}_{\text{act}}\boldsymbol{T}_u\boldsymbol{C}\boldsymbol{T}_y\boldsymbol{T}_{\text{m}}\boldsymbol{G}\right)^{-1}, \tag{5.1}$$

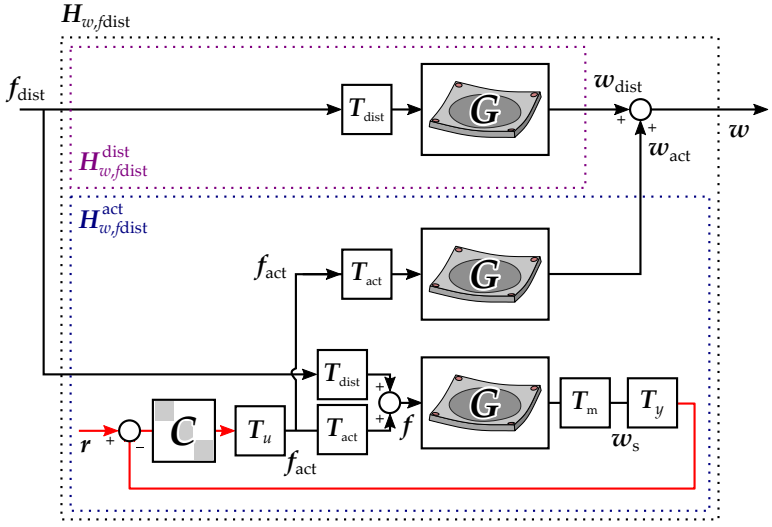
The block diagram that is presented in Figure 5.3b is equivalent to the one of Figure 5.3a. The displacements of the wafer chuck, $\boldsymbol{w}$, is a result of the response the disturbance forces, $\boldsymbol{f}_{\text{dist}}$, the actuation forces, $\boldsymbol{f}_{\text{act}}$, and the initial conditions of the wafer chuck, $\boldsymbol{G}$. As the system is assumed to be linear, the displacement, $\boldsymbol{w}$, can be decomposed into two parts, the displacements $\boldsymbol{w}_{\text{dist}}$ caused by the disturbances forces and the displacements $\boldsymbol{w}_{\text{act}}$ caused by the actuators.

This decomposition can be expressed in terms of transfer functions. The transfer function $\boldsymbol{H}_{w,f\text{dist}}(s)$ is split into the following two parts:

$$\boldsymbol{H}_{w,f\text{dist}} = \boldsymbol{H}^{\text{dist}}_{w,f\text{dist}} + \boldsymbol{H}^{\text{act}}_{w,f\text{dist}}, \tag{5.2}$$




(a) Similar depiction of the system's transfer function $\boldsymbol{H}_{w,f\text{dist}}$ as in Figure 2.3a. Here the matrices $\boldsymbol{T}_{\text{act}}$, $\boldsymbol{T}_{\text{dist}}$ and $\boldsymbol{T}_{\text{m}}$ have been added and the feedforward path has been removed for clarity.




(b) The system's transfer function $\boldsymbol{H}_{w,f\text{dist}}$ can be expressed as the sum of the open loop response to the disturbances ($\boldsymbol{H}^{\text{dist}}_{w,f\text{dist}}$) and the wafer chuck's response to only the actuator forces that are in closed-loop generated in reaction of the disturbances ($\boldsymbol{H}^{\text{act}}_{w,f\text{dist}}$).

Figure 5.3: Two equivalent ways of depicting the closed-loop transfer function $\boldsymbol{H}_{w,f\text{dist}}$ of the out-of-plane motion control system, illustrating the definition of $\boldsymbol{H}^{\text{dist}}_{w,f\text{dist}}$ and $\boldsymbol{H}^{\text{act}}_{w,f\text{dist}}$.

where $\boldsymbol{H}^{\text{dist}}_{w,f\text{dist}}$ denotes the direct open-loop transfer functions from the disturbances to the displacements,

$$\boldsymbol{H}^{\text{dist}}_{w,f\text{dist}} = \boldsymbol{G}\boldsymbol{T}_{\text{dist}}, \tag{5.3}$$

and $\boldsymbol{H}^{\text{act}}_{w,f\text{dist}}$ the part of the response of $\boldsymbol{G}$ that is solely caused by the actuator forces, which in turn result from the closed-loop response to the disturbances, $\boldsymbol{H}_{f\text{act},f\text{dist}}(s)$, so that

$$\boldsymbol{H}^{\text{act}}_{w,f\text{dist}} = \boldsymbol{G}\boldsymbol{T}_{\text{dist}} = \boldsymbol{G}\boldsymbol{T}_{\text{act}}\boldsymbol{H}_{f\text{act},f\text{dist}}, \tag{5.4}$$

with

$$\boldsymbol{H}_{f\text{act},f\text{dist}} = -\boldsymbol{T}_u\boldsymbol{C}\boldsymbol{T}_y\boldsymbol{T}_{\text{m}}\boldsymbol{G}\big(\boldsymbol{I} + \boldsymbol{T}_{\text{act}}\boldsymbol{T}_u\boldsymbol{C}\boldsymbol{T}_y\boldsymbol{T}_{\text{m}}\boldsymbol{G}\big)^{-1}\boldsymbol{T}_{\text{dist}}. \tag{5.5}$$

As the actuation forces $\boldsymbol{f}_{\text{act}}$ are known and assuming that it is possible to obtain $\boldsymbol{w}_{\text{act}}$ based upon them, the contributions $\boldsymbol{w}_{\text{dist}}$ of the disturbance forces to the displacements $\boldsymbol{w}$ can be found at the sensor locations and are then be used for shape fitting. The open-loop part due to disturbances forces (transfer function $\boldsymbol{H}^{\text{dist}}_{w,f\text{dist}}$) only involves the wafer chuck and not the control loop. Thus, the quasi-static frequency band extends up to the wafer chuck's first eigenfrequency, regardless of the controller's bandwidth.

A difficulty of the split approach is, however, that the response to the actuation forces (transfer function $\boldsymbol{H}^{\text{act}}_{w,f\text{dist}}$) needs to be known with sufficient accuracy. This is, however, not trivial, as the system's initial states are not known and that the model that is used for finding the response might not fully equal the real system. For the black box approach, on the other hand, a system model is not necessarily required, as long as the output behaviour of the system is known in the form of a snapshot matrix.

In this thesis it is chosen to develop further with the black box approach, as it extends in a straightforward way the standard SF method that is used for systems without RB motion and a controller. The split approach is left as an alternative that needs further study regarding theoretical feasibility and practical performance.

## 5.2 Decoupling rigid body and flexible modes

In a system that exhibits RB motion it is necessary to decouple the position contributions of the RB modes and flexible modes, as the RB modes typically behave dynamically. This section shows how the RB positions of the points of the object can be found, or rather estimated, and how the estimate contribution of the flexible modes, the flexible residues, can be obtained. Finally, it is shown how the flexible residues can be applied for shape fitting.

### 5.2.1 The rigid body estimate

For decoupling the RB modes, the RB coordinates, which were introduced in Section 2.1.3, are used. The positions in the out-of-plane direction $\boldsymbol{\Delta z}$

can be divided into the contribution of the $N_{\text{rb}}$ RB modeshapes, $\Delta \boldsymbol{z}_{\text{rb}}$, and a displacement contribution due to the flexible modeshapes, $\boldsymbol{w}_{\text{fl}}$:

$$\Delta \boldsymbol{z} = \Delta \boldsymbol{z}_{\text{rb}} + \boldsymbol{w}_{\text{fl}} = \breve{\boldsymbol{\Phi}}_{\text{rb}} \boldsymbol{q}_{\text{rb}} + \boldsymbol{w}_{\text{fl}}, \tag{5.6}$$

where $\breve{\boldsymbol{\Phi}}_{\text{rb}}$ is a $N_{\text{P}} \times N_{\text{rb}}$-matrix consisting of the relevant RB modeshapes and $\boldsymbol{q}_{\text{rb}}$ is a vector with the modal amplitudes of the RB modes. The positions only corresponding to the measurement points read

$$\Delta \boldsymbol{z}_{\text{m}} = \Delta \boldsymbol{z}_{\text{m,rb}} + \boldsymbol{w}_{\text{m,fl}} = \breve{\boldsymbol{\Phi}}_{\text{m,rb}} \boldsymbol{q}_{\text{rb}} + \boldsymbol{w}_{\text{m,fl}}. \tag{5.7}$$

For obtaining the flexible residues, the RB contributions have to be subtracted from the measurement signals. First of all, it is important that all RB modes can be observed at the sensor locations. In order to be able to measure independent information of all RB modeshapes, the placement of the sensors should be such that $\breve{\boldsymbol{\Phi}}_{\text{m,rb}}$ is full column rank. Still, it is not possible to obtain the exact contribution of the RB modes based on the position measurements, as the measurements also contain the flexible contribution. Thus, the RB coordinates, $\boldsymbol{q}_{\text{rb}}$, can not be found exactly, but must be estimated from the displacement measurements

$$\overline{\boldsymbol{q}}_{\text{rb}} = \boldsymbol{T} \Delta \boldsymbol{z}_{\text{m}} = \boldsymbol{T} \breve{\boldsymbol{\Phi}}_{\text{m,rb}} \boldsymbol{q}_{\text{rb}} + \boldsymbol{T} \boldsymbol{w}_{\text{m,fl}}, \tag{5.8}$$

where $\boldsymbol{T}$ is a decoupling matrix of size $N_{\text{rb}} \times N_{\text{M}}$ that is defined in Eq. (5.11). The estimate RB positions are

$$\Delta \overline{\boldsymbol{z}}_{\text{rb}} = \breve{\boldsymbol{\Phi}}_{\text{rb}} \overline{\boldsymbol{q}}_{\text{rb}} = \breve{\boldsymbol{\Phi}}_{\text{rb}} \boldsymbol{T} \Delta \boldsymbol{z}_{\text{m}} = \breve{\boldsymbol{\Phi}}_{\text{rb}} \boldsymbol{T} \Delta \boldsymbol{z}_{\text{m,rb}} + \breve{\boldsymbol{\Phi}}_{\text{rb}} \boldsymbol{T} \boldsymbol{w}_{\text{m,fl}}. \tag{5.9}$$

### 5.2.2 The flexible residue

The flexible residues at the point on the object are defined as the difference between the actual positions and the estimate RB displacement field:

$$\begin{aligned} \overline{\boldsymbol{w}}_{\text{fl}} &= \Delta \boldsymbol{z} - \Delta \overline{\boldsymbol{z}}_{\text{rb}} = \breve{\boldsymbol{\Phi}}_{\text{rb}} \boldsymbol{q}_{\text{rb}} + \boldsymbol{w}_{\text{fl}} - \breve{\boldsymbol{\Phi}}_{\text{rb}} \boldsymbol{T} \breve{\boldsymbol{\Phi}}_{\text{m,rb}} \boldsymbol{q}_{\text{rb}} - \breve{\boldsymbol{\Phi}}_{\text{rb}} \boldsymbol{T} \boldsymbol{w}_{\text{m,fl}} \\ &= \breve{\boldsymbol{\Phi}}_{\text{rb}} \big( \boldsymbol{I} - \boldsymbol{T} \breve{\boldsymbol{\Phi}}_{\text{m,rb}} \big) \boldsymbol{q}_{\text{rb}} + \boldsymbol{w}_{\text{fl}} - \breve{\boldsymbol{\Phi}}_{\text{rb}} \boldsymbol{T} \boldsymbol{w}_{\text{m,fl}}. \end{aligned} \tag{5.10}$$

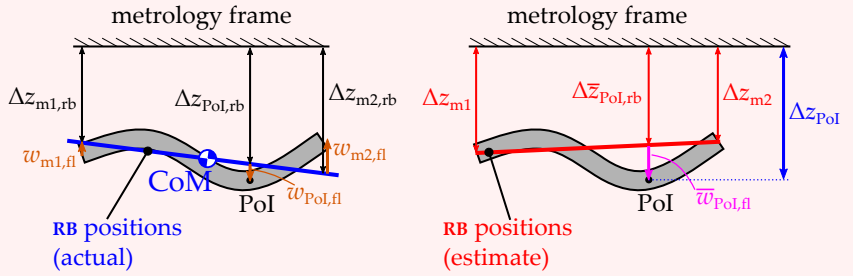
Example 5.1 illustrates the difference between the actual and the estimate RB coordinates and shows the definition of the flexible residue for a beam with two measurement points and a single target point.

For use in shape fitting, the flexible residue should be fully independent of the RB coordinates, $\boldsymbol{q}_{\text{rb}}$. This is the case if decoupling matrix $\boldsymbol{T}$ is chosen such that $\boldsymbol{T} \breve{\boldsymbol{\Phi}}_{\text{m,rb}}$ equals $\boldsymbol{I}$. This is, amongst others, the case if $\boldsymbol{T}$ is the left pseudo-inverse of a matrix $\boldsymbol{S}_{\text{mrb}} \breve{\boldsymbol{\Phi}}_{\text{m,rb}}$, i.e.

$$\boldsymbol{T} = \breve{\boldsymbol{\Phi}}^{+}_{\text{m,rb}} = \left( \breve{\boldsymbol{\Phi}}^{\mathsf{T}}_{\text{m,rb}} \boldsymbol{S}_{\text{mrb}} \breve{\boldsymbol{\Phi}}_{\text{m,rb}} \right)^{-1} \breve{\boldsymbol{\Phi}}^{\mathsf{T}}_{\text{m,rb}} \boldsymbol{S}_{\text{mrb}}, \tag{5.11}$$

where $\boldsymbol{S}_{\text{mrb}}$ is a diagonal $N_{\text{rb}} \times N_{\text{rb}}$-matrix with zeros and ones on its diagonal, which chooses which position sensors are used for estimating the RB coordinates. Of course, $\boldsymbol{S}_{\text{mrb}} \breve{\boldsymbol{\Phi}}_{\text{m,rb}}$ needs to remain full column rank in order to be able to collect independent information of all relevant RB modes.

**Example 5.1: Rigid-body decoupling**




The RB positions of the points on a beam are defined by the position and orientation of the beam's centre of mass (left figure). Eq. (5.6) can for this system be written out as

$$\left\{\begin{array}{c}\Delta z_{m1}\\ \Delta z_{m2}\\ \hline \Delta z_{PoI}\end{array}\right\} = \left\{\begin{array}{c}\Delta z_{m1,rb}\\ \Delta z_{m2,rb}\\ \hline \Delta z_{PoI,rb}\end{array}\right\} + \left\{\begin{array}{c}w_{m1,fl}\\ w_{m2,fl}\\ \hline w_{PoI,fl}\end{array}\right\}, \tag{5.14}$$

$$= \left[\begin{array}{cc}1 & x_{m1}\\ 1 & x_{m2}\\ \hline 1 & x_{PoI}\end{array}\right] \left\{\begin{array}{c}q_{rb1}\\ q_{rb2}\end{array}\right\} + \left\{\begin{array}{c}w_{m1,fl}\\ w_{m2,fl}\\ \hline w_{PoI,fl}\end{array}\right\}, \tag{5.15}$$

where $x_i$ is the lateral position of $i$th point with respect to the beam's centre. The first two lines in Eqs. (5.14) and (5.15) correspond to the measurement point, the last one to the target point, i.e. the PoI.

From position measurements $\Delta z_{m1}$ and $\Delta z_{m1}$ only *estimate* RB positions can be found (right figure). The difference between the PoI's position $\Delta z_{PoI}$ and the RB estimate $\Delta \overline{z}_{PoI,rb}$, the flexible residue $\overline{w}_{PoI,fl}$, may contain RB contributions but is not a function of the RB modal amplitudes.

The flexible residuals at the sensors equal

$$\overline{\boldsymbol{w}}_{m,fl} = \boldsymbol{\Delta z}_m - \boldsymbol{\Delta}\overline{\boldsymbol{z}}_{m,rb} = \left(\boldsymbol{I} - \breve{\boldsymbol{\Phi}}_{m,rb}\boldsymbol{T}\right)\boldsymbol{\Delta z}_m \tag{5.12}$$

Note that if decoupling matrix $\boldsymbol{T}$ is indeed chosen such that $\boldsymbol{T}\breve{\boldsymbol{\Phi}}_{m,rb}$ equals $\boldsymbol{I}$, the flexible residue (Eq. 5.10) rewrites to

$$\overline{\boldsymbol{w}}_{fl} = \boldsymbol{w}_{fl} - \breve{\boldsymbol{\Phi}}_{rb}\boldsymbol{T}\boldsymbol{w}_{m,fl}. \tag{5.13}$$

This formula shows that, although the flexible residue might contain contributions of RB modeshapes, it is *not a function of the* RB *coordinates*, $\boldsymbol{q}_{rb}$. Thus, the flexible residue does not contain the dynamic behaviour of the RB modes.

### 5.2.3 Shape fitting using the flexible residue

The flexible residues are used for shape fitting. To apply shape fitting, the sensor signals need to be independent. This is not the case any more, however, if part of

the signal is used to find the estimate RB coordinates. Indeed, $\left(\boldsymbol{I} - \breve{\boldsymbol{\Phi}}_{\mathrm{m,rb}}\boldsymbol{T}\right)$ is not full rank but has a rank of $N_S - N_{\mathrm{rb}}$. For this reason, the flexible contributions need to be remixed by a $(N_S - N_{\mathrm{rb}}) \times N_S$-transformation matrix $\boldsymbol{R}^\mathsf{T}$, so that $\boldsymbol{R}^\mathsf{T}\left(\boldsymbol{I} - \breve{\boldsymbol{\Phi}}_{\mathrm{m,rb}}\boldsymbol{T}\right)$ is full rank. This matrix has to span the null-space of $\boldsymbol{T}$ ($\boldsymbol{TR} = \boldsymbol{0}$). After transformation, the recombined flexible residuals become

$$\overline{\boldsymbol{w}}_{\mathrm{R,fl}} = \boldsymbol{R}^\mathsf{T}\overline{\boldsymbol{w}}_{\mathrm{m,fl}} = \boldsymbol{R}^\mathsf{T}\left(\boldsymbol{I} - \breve{\boldsymbol{\Phi}}_{\mathrm{m,rb}}\boldsymbol{T}\right)\boldsymbol{\Delta z}_{\mathrm{m}}. \tag{5.16}$$

These recombined residuals can then be used as measurement signals for the SF process in the way described in the previous chapter, so that the estimates of the flexible residue are

$$\hat{\boldsymbol{w}}_{\mathrm{fl}} = \boldsymbol{B}\overline{\boldsymbol{w}}_{\mathrm{R,fl}} = \boldsymbol{B}\boldsymbol{R}^\mathsf{T}\left(\boldsymbol{I} - \breve{\boldsymbol{\Phi}}_{\mathrm{m,rb}}\boldsymbol{T}\right)\boldsymbol{\Delta z}_{\mathrm{m}}. \tag{5.17}$$

$\boldsymbol{B}$ is the SF matrix, which is found in the standard way with Eq. (4.16), substituting snapshot matrix $\boldsymbol{W}$ for a matrix $\overline{\boldsymbol{W}}_{\mathrm{fl}}$ and $\boldsymbol{W}_{\mathrm{m}}$ for a matrix $\overline{\boldsymbol{W}}_{\mathrm{R,fl}}$, so that

$$\boldsymbol{B} = \overline{\boldsymbol{W}}_{\mathrm{fl}}\boldsymbol{\Sigma}_\alpha\overline{\boldsymbol{W}}_{\mathrm{R,fl}}^\mathsf{T}\left(\overline{\boldsymbol{W}}_{\mathrm{R,fl}}\boldsymbol{\Sigma}_\alpha\overline{\boldsymbol{W}}_{\mathrm{R,fl}}^\mathsf{T} + \boldsymbol{R}^\mathsf{T}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{R}\right)^{-1}, \tag{5.18}$$

where $\boldsymbol{R}^\mathsf{T}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{R}$ is the covariance matrix of the sensor noise in the recombined signals. Matrices $\overline{\boldsymbol{W}}_{\mathrm{fl}}$ and $\overline{\boldsymbol{W}}_{\mathrm{R,fl}}$ consist of sets of vectors $\overline{\boldsymbol{w}}_{\mathrm{fl}}$ and $\overline{\boldsymbol{w}}_{\mathrm{R,fl}}$, respectively, which are calculated from a set of snapshot vectors $\boldsymbol{\Delta z}$ by treating them according to Eqs. (5.9), (5.10) and (5.16).

Finally, the position estimates are

$$\boldsymbol{\Delta}\hat{\boldsymbol{z}} = \boldsymbol{\Delta}\overline{\boldsymbol{z}}_{\mathrm{rb}} + \hat{\boldsymbol{w}}_{\mathrm{fl}} = \breve{\boldsymbol{\Phi}}_{\mathrm{rb}}\boldsymbol{T}\boldsymbol{\Delta z}_{\mathrm{m}} + \boldsymbol{B}\boldsymbol{R}^\mathsf{T}\left(\boldsymbol{I} - \breve{\boldsymbol{\Phi}}_{\mathrm{m,rb}}\boldsymbol{T}\right)\boldsymbol{\Delta z}_{\mathrm{m}} \tag{5.19}$$

and the methodic errors with respect to the actual displacements.

$$\boldsymbol{\varepsilon} = \boldsymbol{\Delta}\hat{\boldsymbol{z}} - \boldsymbol{\Delta z} = \boldsymbol{\Delta}\overline{\boldsymbol{z}}_{\mathrm{rb}} + \hat{\boldsymbol{w}}_{\mathrm{fl}} - \boldsymbol{\Delta z} = \breve{\boldsymbol{\Phi}}_{\mathrm{rb}}\boldsymbol{T}\boldsymbol{\Delta z}_{\mathrm{m}} + \boldsymbol{B}\boldsymbol{R}^\mathsf{T}\left(\boldsymbol{I} - \breve{\boldsymbol{\Phi}}_{\mathrm{m,rb}}\boldsymbol{T}\right)\boldsymbol{\Delta z}_{\mathrm{m}} - \boldsymbol{\Delta z}. \tag{5.20}$$

From diverse numeric experiments it was observed that the methodic error of the SF method was unaffected by the actual choice of decoupling matrix $\boldsymbol{T}$. We do not show a proof but hypothesise that this is true in general.

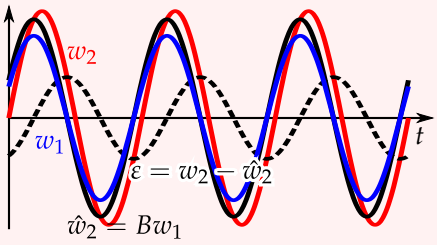
## 5.3 Shape fitting in a dynamic system

Shape fitting can equally well be applied to both quasi-static and dynamic systems. As the magnitude and phase of a dynamic system's signals depend on frequency, the estimation error is expected to be higher. This section illustrates the effect of phase lag on the estimation error and generalises the snapshot matrix such that it can handle signals with a relative phase effectively.

### 5.3.1 Magnitude and relative phase of dynamic signals

In a quasi-static system, the different disturbance sources lead to different signal magnitudes. In a dynamic system, however, the disturbance frequency may lead to different signal magnitudes. Next to that, in a quasi-static system, all signals

**Example 5.2: 2-DoFs estimation with out-of-phase signals**




Consider two sinusoidal signals $w_1(t)$ and $w_2(t)$ with equal frequencies, but different amplitudes $A_1$ and $A_2$ and phases $\varphi_1$ and $\varphi_2$, so that

$$w_1(t) = A_1 \sin(\omega t + \varphi_1), \quad w_2(t) = A_2 \sin(\omega t + \varphi_2). \tag{5.21}$$

Let $w_1(t)$ be the measured signal, which is used to generate the estimate $\hat{w}_2(t)$ of $w_2(t)$ by multiplication with a constant $B$ so that

$$\varepsilon(t) = w_2 - Bw_1 = A_2 \sin(\omega t + \varphi_2) - BA_1 \sin(\omega t + \varphi_1), \tag{5.22}$$

which has a RMS value of

$$\varepsilon_{\text{RMS}} = \sqrt{\frac{\omega}{2\pi} \int_0^{\frac{2\pi}{\omega}} \varepsilon^2(t) dt} = \frac{1}{4} \sqrt{A_1^2 B^2 - A_1 A_2 B \cos(\varphi_1 - \varphi_2) + \frac{1}{2} A_2^2}. \tag{5.23}$$

The RMS error attains a minimum for

$$B = \arg\min_B |\varepsilon|^2 = \frac{A_2}{A_1} \cos(\varphi_1 - \varphi_2), \tag{5.24}$$

with, according to Eq. (5.23), a value of

$$\varepsilon_{\text{RMS}} = \frac{1}{4} A_2 \sqrt{1 - \cos^2(\varphi_1 - \varphi_2)}. \tag{5.25}$$

This result shows that if the signals are in phase or counter-phase ($\varphi_1 - \varphi_2 = k\pi$, with $k$ an integer constant), the error is minimum, with $B = A_2/A_1$. If the signals have, on the other hand, a relative phase of $\pi/2$, the best estimate is 0 (with $B = 0$), leading to the highest possible error for a fixed signal amplitude $A_2$.

are in phase, whereas in a dynamic system signals may have different phase, depending on the frequencies of the disturbances.

Example 5.2 shows a SF estimator for a 2-DoFs system for which a sinusoidal displacement signal has to be estimated from another sinusoidal displacement signal with the same frequency but a different phase and amplitude. If both signals are in phase, the estimation has zero error, but error increases for increasing phase difference.

The SF method can be applied on dynamic systems in the normal way. To attain highest possible performance the snapshot matrix should include not only the snapshots representative to the different disturbance locations and magnitudes, but also to the frequency content of disturbances.

### 5.3.2 A complex snapshot matrix

Section 3.5 introduced the snapshot matrix and how it should be constructed to introduce the knowledge about the system's mechanics and typical disturbances into the SF estimator. A snapshot matrix is representative for the behaviour of a linear system with Gaussian distributed disturbances if the system's response can be linearly composed from the snapshots and if the standard deviation of each snapshot's magnitude is known.



As described in Section 3.5.2, the snapshot matrix can have different forms. In the first form, the snapshots are obtained from random excitation of the disturbances. Such a snapshot matrix naturally includes the magnitude and phase effects the system.

A snapshot matrix of the second form is obtained from one-by-one excitation of the disturbances. For quasi-static systems, such a matrix is obtained using 0 Hz or low-frequency disturbances at the individual disturbance locations. If this form is extended to dynamic systems, the snapshot matrix should include snapshots at a grid of frequencies in the relevant frequency range. Thus, the snapshot matrix does not only contain snapshots that correspond to the disturbance locations, but also to the disturbance frequencies.

To include the relative phase of the signals, multiple snapshots per frequency point are necessary, at least two per frequency point, obtained at a $\pi$ rad phase difference. Alternatively, the snapshots can be constructed using complex values. The major benefit of a complex snapshot matrix is that just one snapshot is needed per frequency point to adequately describe relative phase of the signals. The amplitudes of the sinusoid signals are described by the absolute values of the complex numbers and the relative phases by the arguments. This is similar to a Frequency Response Function (FRF) and makes it possible to directly use the FRF data for constructing the snapshot matrix.

### 5.3.3 Estimator calculation using the complex snapshot matrix

Recall the expression of the estimation error (Eq. 3.11), i.e.

$$\varepsilon_j(t) = \Big(\boldsymbol{\beta}_j \boldsymbol{W}_\mathrm{m} - \boldsymbol{w}_j\Big)\boldsymbol{\alpha} + \boldsymbol{\beta}_j \boldsymbol{\varepsilon}_\mathrm{m}, \tag{5.26}$$

where $\varepsilon_j$, $\boldsymbol{\alpha}$ and $\boldsymbol{\varepsilon}_\mathrm{m}$ were functions of time with Gaussian distributions elements. $\boldsymbol{W}_\mathrm{m}$ and $\boldsymbol{w}_j$ were real numbers.

The equation can be adapted to allow for complex values in the snapshot matrix $\boldsymbol{W}$ and its subsets $\boldsymbol{W}_\mathrm{m}$ and $\boldsymbol{w}_j$. The weighing factors in vector $\boldsymbol{\alpha}$ need then to be changed to not only describe the amplitude of the snapshot, but also the phase of the snapshot. The elements $\alpha_i$ of $\boldsymbol{\alpha}$, therefore, need to be distributed according to $\Re(\alpha_i) \sim \mathcal{N}(0, \sigma^2_{\alpha,i})$ and $\angle\alpha_i \sim \mathcal{U}(0, 2\pi)$. This means that the values $\alpha_i$ are random numbers from independent circularly-symmetric complex random distributions, i.e. $\alpha_i \sim \mathcal{CN}(0, 2\sigma^2_{\alpha,i})$. The resulting $(\boldsymbol{\beta}_j\boldsymbol{W}_\mathrm{m} - \boldsymbol{w}_j\boldsymbol{\alpha})$ is then a complex number, whose real value represents the estimation error for the given weighing factors $\boldsymbol{\alpha}$, so that

$$\varepsilon_j(t) = \Re\Big[\Big(\boldsymbol{\beta}_j\boldsymbol{W}_\mathrm{m} - \boldsymbol{w}_j\Big)\boldsymbol{\alpha}\Big] + \boldsymbol{\beta}_j\boldsymbol{\varepsilon}_\mathrm{m}. \tag{5.27}$$

To be able to calculate the expected value of the total error, Eq. (5.27) is rewritten as

$$\varepsilon_j(t) = \frac{1}{2}\Big[\Big(\boldsymbol{\beta}_j\boldsymbol{W}_\mathrm{m} - \boldsymbol{w}_j\Big)\boldsymbol{\alpha} + \Big(\boldsymbol{\beta}_j\bar{\boldsymbol{W}}_\mathrm{m} - \bar{\boldsymbol{w}}_j\Big)\bar{\boldsymbol{\alpha}}\Big] + \boldsymbol{\beta}_j\boldsymbol{\varepsilon}_\mathrm{m}, \tag{5.28}$$

where $\bar{\boldsymbol{W}}_\mathrm{m}$ and $\bar{\boldsymbol{w}}_j$ are the complex conjugates of $\boldsymbol{W}_\mathrm{m}$ and $\boldsymbol{w}_j$. The expected value of the error squared is

$$\begin{aligned} E\Big(\varepsilon_j^2\Big) =& E\Bigg(\frac{1}{4}\Big[\Big(\boldsymbol{\beta}_j\boldsymbol{W}_\mathrm{m} - \boldsymbol{w}_j\Big)\boldsymbol{\alpha} + \Big(\boldsymbol{\beta}_j\bar{\boldsymbol{W}}_\mathrm{m} - \bar{\boldsymbol{w}}_j\Big)\bar{\boldsymbol{\alpha}} + \boldsymbol{\beta}_j\boldsymbol{\varepsilon}_\mathrm{m}\Big]\times \\ & \Big[\boldsymbol{\alpha}^\mathsf{T}\Big(\boldsymbol{W}_\mathrm{m}^\mathsf{T}\boldsymbol{\beta}_j^\mathsf{T} - \boldsymbol{w}_j^\mathsf{T}\Big) + \boldsymbol{\alpha}^*\Big(\boldsymbol{W}_\mathrm{m}^*\boldsymbol{\beta}_j^\mathsf{T} - \boldsymbol{w}_j^*\Big) + \boldsymbol{\varepsilon}_\mathrm{m}^\mathsf{T}\boldsymbol{\beta}_j^\mathsf{T}\Big]\Bigg), \end{aligned} \tag{5.29}$$

with $\boldsymbol{\alpha}^*$, $\boldsymbol{W}_\mathrm{m}^*$, and $\boldsymbol{w}_j^*$ the transpose complex conjugates of $\boldsymbol{\alpha}$, $\boldsymbol{W}_\mathrm{m}$, and $\boldsymbol{w}_j$. The formula can be simplified by using that $E\big(\boldsymbol{\alpha}\boldsymbol{\varepsilon}_\mathrm{m}^\mathsf{T}\big) = E\big(\bar{\boldsymbol{\alpha}}\boldsymbol{\varepsilon}_\mathrm{m}^\mathsf{T}\big) = E\big(\boldsymbol{\varepsilon}_\mathrm{m}^\mathsf{T}\boldsymbol{\alpha}^\mathsf{T}\big) = E\big(\boldsymbol{\varepsilon}_\mathrm{m}^\mathsf{T}\boldsymbol{\alpha}^*\big) = 0$. Furthermore, for a circularly-symmetric complex random value $a$ with independent real and imaginary parts $a_\Re, a_\Im \sim \mathcal{N}\big(0, \sigma^2\big)$ the following hold

$$E((a_\Re + a_\Im i)(a_\Re + a_\Im i)) = E\Big(a_\Re^2 - a_\Im^2 + 2a_\Re a_\Im i\Big) = \sigma^2 - \sigma^2 = \quad 0 \tag{5.30}$$

$$E((a_\Re + a_\Im i)(a_\Re - a_\Im i)) = \qquad E\Big(a_\Re^2 + a_\Im^2\Big) = \sigma^2 + \sigma^2 = 2\sigma^2 \tag{5.31}$$

$$E((a_\Re - a_\Im i)(a_\Re - a_\Im i)) = E\Big(a_\Re^2 - a_\Im^2 + 2a_\Re a_\Im i\Big) = \sigma^2 - \sigma^2 = \quad 0, \tag{5.32}$$

so that $E(\boldsymbol{\alpha}\boldsymbol{\alpha}^\mathsf{T}) = E(\bar{\boldsymbol{\alpha}}\boldsymbol{\alpha}^*) = 0$ and $E(\boldsymbol{\alpha}\boldsymbol{\alpha}^*) = 2\boldsymbol{\Sigma}_\alpha$, with $\boldsymbol{\Sigma}_\alpha$ the diagonal covariance matrix of $\boldsymbol{\alpha}$. Thus,

$$E\Big(\varepsilon_j^2\Big) = \boldsymbol{\beta}_j\boldsymbol{W}_\mathrm{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\mathrm{m}^*\boldsymbol{\beta}_j^\mathsf{T} - \boldsymbol{\beta}_j\boldsymbol{W}_\mathrm{m}\boldsymbol{\Sigma}_\alpha\boldsymbol{w}_j^* - \boldsymbol{w}_j\boldsymbol{\Sigma}_\alpha\boldsymbol{W}_\mathrm{m}^*\boldsymbol{\beta}_j^\mathsf{T} + \boldsymbol{w}_j\boldsymbol{\Sigma}_\alpha\boldsymbol{w}_j^* + \boldsymbol{\beta}_j\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\beta}_j^\mathsf{T}. \tag{5.33}$$

The optimal $\boldsymbol{\beta}_j$ is the one that leads to a minimum expected value of the estimation error, $E(\varepsilon_j^2)$. $\boldsymbol{\beta}_j$ is found by setting the gradient of $E(\varepsilon_j^2)$ with respect to $\boldsymbol{\beta}_j$ zero. Now the gradient of the error's expected value can be taken, using

the fact that $\nabla_{\boldsymbol{b}} \boldsymbol{b}^{\mathsf{T}} \boldsymbol{A} \boldsymbol{b} = (\boldsymbol{A} + \boldsymbol{A}^{\mathsf{T}})\boldsymbol{b}$:

$$\begin{aligned} \nabla_{\boldsymbol{\beta}} E\left(\varepsilon_j^2\right) &= \left(2\boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{W}_{\mathrm{m}}^{*} \boldsymbol{\beta}_j^{\mathsf{T}} + 2\bar{\boldsymbol{W}}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{W}_{\mathrm{m}}^{\mathsf{T}} \boldsymbol{\beta}_j^{\mathsf{T}}\right) - \\ &\quad \boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{w}_j^{*} - \bar{\boldsymbol{W}}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{w}_j^{\mathsf{T}} + 2\boldsymbol{\Sigma}_{\varepsilon} \boldsymbol{\beta}_j^{\mathsf{T}} \qquad (5.34) \\ &= \left[\boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{W}_{\mathrm{m}}^{*} + \bar{\boldsymbol{W}}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{W}_{\mathrm{m}}^{\mathsf{T}} + \boldsymbol{\Sigma}_{\varepsilon}\right] \boldsymbol{\beta}_j^{\mathsf{T}} - \\ &\quad \left[\boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{w}_j^{*} + \bar{\boldsymbol{W}}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{w}_j^{\mathsf{T}}\right] = 0, \qquad (5.35) \end{aligned}$$

so that

$$\begin{aligned} \boldsymbol{\beta}_j^{\mathsf{T}} &= \left[\boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{W}_{\mathrm{m}}^{*} + \bar{\boldsymbol{W}}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{W}_{\mathrm{m}}^{\mathsf{T}} + \boldsymbol{\Sigma}_{\varepsilon}\right]^{-1} \left[\boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{w}_j^{*} + \bar{\boldsymbol{W}}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{w}_j^{\mathsf{T}}\right] \\ &= \left[\Re(\boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{W}_{\mathrm{m}}^{*}) + \boldsymbol{\Sigma}_{\varepsilon}\right]^{-1} \Re\left(\boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{w}_j^{*}\right) \qquad (5.36) \end{aligned}$$

and thus

$$\boldsymbol{\beta}_j = \Re\left(\boldsymbol{w}_j^{*} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{W}_{\mathrm{m}}\right) \left[\Re(\boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_{\alpha} \boldsymbol{W}_{\mathrm{m}}^{*}) + \boldsymbol{\Sigma}_{\varepsilon}\right]^{-1}. \qquad (5.37)$$

Example 5.3 illustrates this formula for a 2-DoFs system.

5

## 5.4 Numerical study of shape fitting in a dynamic wafer chuck

The numerical study of the last chapter used a highly simplified model of the wafer chuck system. The system did not include a realistic controller to keep the wafer chuck in position, did not undergo RB motion and did not make use of RB decoupling. The numerical experiments in this section follow on the previous study. This time a more realistic wafer chuck model that includes RB motion and a control system is adopted. RB decoupling and shape fitting with a complex frequency domain snapshot-matrix to include the amplitude and phase of the signals are included in the analysis.

### 5.4.1 Method

The method of this numerical study is for the greater part equal to the one of Section 4.4 and their methods are to a large extent comparable. The current study only uses the regularised least squares SF technique, as this technique was shown to yield superior results over the others. As before, a disturbance level of 1 N RMS over the exposure area and a sensor noise level of $0.1 \cdot 10^{-9}$ m were adopted.

Four different scenarios were studied, building up from the scenario with an ideal controller, no RB decoupling and quasi-static disturbances to the realistic scenario with a normal controller, RB decoupling and disturbances in a certain bandwidth.

1. *Ideal controller, no* RB *decoupling, quasi-static:* This scenario replicates the results of Section 4.4.

**Example 5.3: 2-DoFs estimation with out-of-phase signals (2)**

Let the two sinusoid signals of Example 5.2 be described by complex numbers as follows

$$x = a + bi, \quad y = c + di, \tag{5.38}$$

where the parameters $a$, $b$, $c$ and $d$ are real. Choose $A \sim \mathcal{CN}(0, 2\sigma_A^2)$, so that $w_1 = \Re((a+bi)A_1)$ and $w_2 = \Re((c+di)A_2)$.

If the signal $y$ needs to be linearly estimated based on signal $x$, the SF estimate is

$$\hat{y} = B(a + bi), \tag{5.39}$$

where $B$ is a real number that must be chosen such that the estimation error is minimal. $B$ is calculated using Eq. (5.37):

$$B = \Re(xy^*)[\Re(xy^*)]^{-1} = \frac{ac+bd}{a^2+b^2}. \tag{5.40}$$

This result corresponds to Eq. (5.24) in Example 5.2. This follows from the definition of the dot product between $x$ and $y$, when regarded as vectors in the complex plane, $\sqrt{a^2+b^2}\sqrt{c^2+d^2}\cos(\varphi_1 - \varphi_2) = ac + bd$ and the fact that $A_1 = \sqrt{a^2+b^2}$ and $A_2 = \sqrt{c^2+d^2}$.

The expected value of the error squared is, according to Eq. (5.37),

$$\begin{aligned} E\left(\varepsilon_j^2\right) &= (Bx - y)(\bar{x}B - \bar{y}) \\ &= \left((a+bi)\frac{ac+bd}{a^2+b^2} - (c+di)\right)\left((a-bi)\frac{ac+bd}{a^2+b^2} - (c-di)\right) \\ &= \frac{(bc-ad)^2}{a^2+b^2}. \end{aligned} \tag{5.41}$$

Table 5.1: The PID parameters of the RB controllers as used in the numerical study of the wafer chuck.

| | RB coordinate $z$ | RB coordinate $\theta_x$ | RB coordinate $\theta_y$ |
|---|---|---|---|
| $k_\mathrm{p}$ [N/m] | $6 \cdot 10^6$ | $1 \cdot 10^6$ | $1 \cdot 10^6$ |
| $k_\mathrm{d}$ [Ns/m] | $1.6 \cdot 10^3$ | $0.26 \cdot 10^3$ | $0.26 \cdot 10^3$ |
| $k_\mathrm{dN}$ [Ns/m] | $0.1 \cdot 2\pi f_\mathrm{s}$ | $0.1 \cdot 2\pi f_\mathrm{s}$ | $0.1 \cdot 2\pi f_\mathrm{s}$ |
| $k_\mathrm{i}$ [N/ms] | $1 \cdot 10^8$ | $1 \cdot 10^8$ | $1 \cdot 10^8$ |

2. *Normal controller, quasi-static:* In this scenario, a normal controller was applied instead of an ideal controller. This led to RB motion, which needed to be decoupled from the signals of the position sensors. RB decoupling was performed according to the method in Section 5.2. All eight position sensors were used for the decoupling. The disturbance forces were quasi-static ($f_{\text{dist}} = 0$ Hz). The three RB controllers were PID controllers of the form

$$C(s) = k_{\text{p}} + \frac{k_{\text{d}} s}{\frac{1}{k_{\text{dN}}} s + 1} + \frac{k_{\text{i}}}{s}, \tag{5.42}$$

with the controller parameters as listed in Table 5.1.

3. *Normal controller, 0–100 Hz:* This scenario was comparable to the previous scenario, except that the disturbances were dynamic with a bandwidth of 0 to 100 Hz. The snapshot matrix was extended to include a grid of 40 frequencies, evenly distributed within the range of 0 to 100 Hz, per disturbance location. The snapshot matrix was complex, as described in Section 5.3.

4. *Normal controller, 0–100 Hz, RB decoupling corner sensors only:* This scenario was equal to the previous one. The RB decoupling was this time not based on all eight position sensors, but only on the four sensors at the wafer chuck's corners.



### 5.4.2 Results and discussion

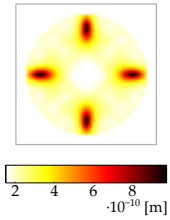
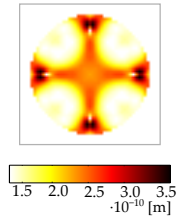
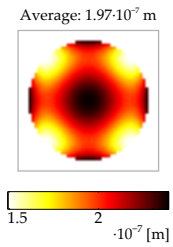
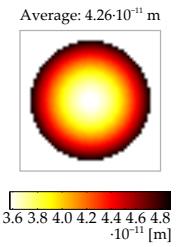
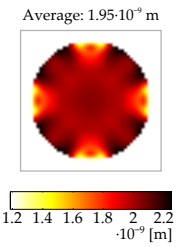
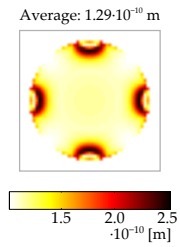
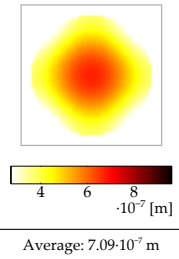
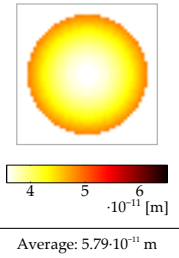
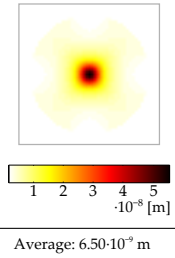
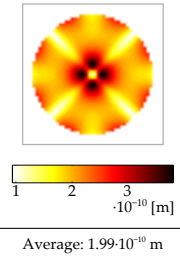
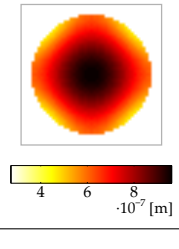
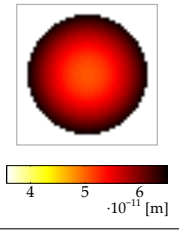
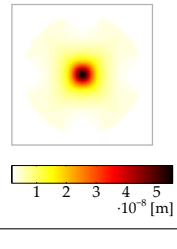
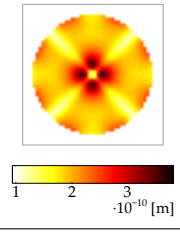
Table 5.2 compares the performance of the RB and the SF estimators in the four scenarios. The performance is presented in a similar way as in Table 4.1; the coordinates $(x, y)$ indicate the position of the lens column; the colours indicate the average of the error's standard deviation over the lens area. The table's left two columns show the estimation errors of the RB estimation, the right two the estimation error of the SF estimation.

The upper row replicates the main results of the numeric experiments of Section 4.4.

The table's second row shows the results for the scenario with the normal position controller and RB decoupling. The methodical error of the RB estimate show a somewhat similar distribution as the chuck's deformation (Figure 4.2). In fact, the error of the RB estimate would equal the chuck's deformation if the RB estimate would be 'ideal', i.e. if the RB estimate would equal the actual RB position. Both the spatial distribution and the magnitude of the methodical error of the SF estimation differ significantly from the scenario with the ideal controller: the magnitude increased by a factor of 10, roughly. This can be explained from the fact that after decoupling the three RB modeshapes only five flexible residue signals are available, decreasing the amount of available information about the deformation.

The third row presents the estimation results for disturbances in the bandwidth from 0 to 100 Hz. 100 Hz is well below the wafer chuck's first resonance frequency, but above the control bandwidth of the RB controller. On average

Table 5.2: The performance of the RB and SF estimation as function of the lens position for different scenarios.

| Scenario | RB estimation | | SF estimation | |
|---|---|---|---|---|
| | Methodical estimation error ($E_{\text{meth,RB}}$) | Estimation error due to noise ($E_{\text{noise,RB}}$) | Methodical estimation error ($E_{\text{meth,SF}}$) | Estimation error due to noise ($E_{\text{noise,SF}}$) |
| Ideal controller, no RB decoupling, quasi-static (Section 4.4) | *not applicable* | *not applicable* | Average: $2.73 \cdot 10^{-10}$ m   | Average: $1.89 \cdot 10^{-10}$ m   |
| Normal controller, quasi-static | Average: $1.97 \cdot 10^{-7}$ m   | Average: $4.26 \cdot 10^{-11}$ m   | Average: $1.95 \cdot 10^{-9}$ m   | Average: $1.29 \cdot 10^{-10}$ m   |
| Normal controller, 0–100 Hz | Average: $4.63 \cdot 10^{-7}$ m   | Average: $4.26 \cdot 10^{-11}$ m   | Average: $6.50 \cdot 10^{-9}$ m   | Average: $2.08 \cdot 10^{-10}$ m   |
| Normal controller, 0–100 Hz, RB decoupling corner sensors only | Average: $7.09 \cdot 10^{-7}$ m   | Average: $5.79 \cdot 10^{-11}$ m   | Average: $6.50 \cdot 10^{-9}$ m   | Average: $1.99 \cdot 10^{-10}$ m   |

the methodical error increases due to the chuck's larger deformation. Also the methodical error of the SF estimator increases, on average by a factor of 3, at maximum by roughly a factor of 25. The distribution of the error changes: the highest error concentrates in the centre of the wafer chuck.

The results show that it is not enough to only study the quasi-static behaviour of a controlled system that is dynamically excited, even if the disturbances are well below the lowest mechanical resonance frequency. It was not studied to what extent the results depend on the choice of controller parameters (Eq. 5.42).

The last row shows the results for the scenario that decoupling is only performed using the four displacement sensors at the chuck's corners. This leads to a higher methodical error of the RB estimation, as could be expected. The methodical error of the SF estimation is equal, however, which is in line with the formulated hypothesis that the standard deviation of this error is not influenced by the specific choice of the RB decoupling. The SF error due to noise decreases slightly.

The estimation error due to sensor noise is relatively small compared to the methodical error, at least a factor of 10 smaller for the given disturbance and sensor noise levels. This means that sensors with a, say, 12 times higher noise level might be used without significantly compromising (less than 10 %) the total estimation error.



## 5.5 Hybrid shape fitting estimation

Shape fitting estimators perform best if optimised for a small, preferably quasi-static, bandwidth. Inertia-based absolute position sensors, on the other hand, have best performance at higher frequencies. Section 2.4.2 introduced absolute position sensors and showed that in the wafer chuck application absolute position should be obtained from acceleration. It is proposed to construct a hybrid estimator that fuses the signal of a SF estimator optimised for lower frequencies with a double integrated accelerometer signal.

The fusion of accelerometer and position signals is well-known in literature. Shaw and Srinivasan [82], for example, showed a method that makes use of complementary filtering; the double integrated acceleration signal is highpass filtered and merged with the lowpass filtered position signal. A more advanced way of filtering and merging signals in order to result a low noise floor is using Kalman filtering, for example [83, 84].

The fusion an accelerometer signal with a SF estimate was, to the best of our knowledge, not shown in literature before. Section 7.5 implements a hybrid SF estimator that uses complementary filtering to the experimental setup.

## 5.6 Conclusions

This chapter studied how shape fitting can be used in a system with RB motion and position controller. The wafer chuck has free boundary conditions and is kept in place using a control system. For shape fitting, a black box approach is adopted, which considers the actuator forces as internal forces that are not used

for estimation. The RB modes of the closed-loop system behave dynamically and the flexible modes quasi-statically, at least up to the controller bandwidth.

In order to remove the contributions of the RB modes from the position measurements, the signals are decoupled into estimate RB positions and flexible residues. These residues can be used as inputs for shape fitting, together with internally referenced measurands. The total position estimates are the sum of the RB estimates and the SF estimates. Numerical results suggested that the total methodical error is independent of the way in which the RB modes are decoupled.

Shape fitting can be applied to dynamic systems in the same way as to quasi-static systems. The snapshot matrix from one-by-one excitation of the disturbances may be generalised using a grid of frequency points. The snapshot matrix that is used for estimating the SF estimator can be formulated compactly using complex numbers. Those complex numbers describe the magnitude and relative phase of a response.

The SF method and the complex snapshot matrix were used in a numerical study on the wafer chuck with dynamical signals. The model of the wafer chuck included a position controller. The estimation error of the SF estimator optimised for the dynamic signals was significantly larger than the error for the quasi-static SF estimation, differing by a factor of 3 on average, locally up to a factor of 25. Still, applying the SF estimator decreases the methodical error by at least a factor of 30 as compared to the RB estimator. The analysis furthermore showed that the required noise level of the sensors can be relieved to 1.2 nm RMS without significantly compromising the total estimation error.

# Chapter 6

# Sensor placement

The previous chapters presented the Shape Fitting (SF) method and adapted it to make it applicable in the wafer chuck problem. It was shown that shape fitting in combination with the least squares technique minimises the estimation error for a given set of sensor locations. It is expected, however, that the estimation results can be further improved by not only optimising the estimation algorithm but also the sensor locations.

As the choice of the location of a sensor also influences the performance of the other sensors, selecting the optimal sensor locations by hand-picking is not trivial. Therefore, an automated procedure for selecting the sensor locations is desired. Fields like wireless sensor networks and experimental dynamics study the automated choice of sensor locations, normally referred to as 'sensor placement', and in these fields many algorithms have been proposed. Still, it is not known how the algorithms should be integrated into shape fitting and which of the algorithms perform best. This chapter compares a number of promising algorithms from literature and presents a procedure that integrates sensor placement into the SF framework.

Section 6.1 presents a brief literature survey on sensor placement. First an overview of some performance metrics and some sensor placement algorithms and their respective optimisation objectives is given. Four promising algorithms from literature are then chosen for further investigation; they are treated in more detail and assessed using displacement and strain shapes in Section 6.2. Section 6.3 presents a method for integrating placement algorithms into the SF framework. Then, Section 6.4 applies the integration procedure to the wafer chuck case of last chapter. The thus optimised sensor configuration indeed leads to a significant decrease of the SF estimation error.

## 6.1 Selection of sensor placement algorithms for identifying shapes

This section first defines the goal of sensor placement. Based on this goal, promising sensor placement algorithms are selected, which are then introduced

in more detail.

### 6.1.1 The goal of sensor placement

The performance of the SF estimator is measured in terms of the estimation error at the $j$th point to be estimated, defined in Eq. (3.11) as

$$\varepsilon_j(t) = \Big(\boldsymbol{\beta}_j \boldsymbol{W}_\mathrm{m} - \boldsymbol{w}_j\Big)\boldsymbol{\alpha} + \boldsymbol{\beta}_j \boldsymbol{\varepsilon}_\mathrm{m},$$

where $\boldsymbol{\beta}_j$ denotes the $j$th row of estimation matrix $\boldsymbol{B}$. For a given sensor configuration, the optimal $\boldsymbol{\beta}_j$ can be found using the LS technique, presented in Eq. (4.15) as

$$\boldsymbol{\beta}_j = \boldsymbol{w}_j \boldsymbol{\Sigma}_\alpha \boldsymbol{W}_\mathrm{m}^\mathsf{T} \big(\boldsymbol{W}_\mathrm{m} \boldsymbol{\Sigma}_\alpha \boldsymbol{W}_\mathrm{m}^\mathsf{T} + \boldsymbol{\Sigma}_\varepsilon\big)^{-1}.$$

These equations show that $\varepsilon_j$ depends on the chosen sensor locations, as $\boldsymbol{W}_\mathrm{m}$ is a subset of the rows of $\boldsymbol{W}$ that correspond to the sensor locations. The rows of $\boldsymbol{W}$ may either correspond to a point that needs to be estimated (a target point), to a location where a sensor can be placed (a sensor candidate point) or to both. The goal of sensor placement is to select the subset of sensor locations that minimises $\varepsilon_j$ at the relevant target points. A corresponding performance metric $J$ could be defined as, for example,

$$J = \sum_j \varepsilon_j^2. \tag{6.1}$$



### 6.1.2 Evaluation time and convergence

Evaluating the objective function for any sensor configuration, the so-called 'brute force' sensor placement, is in practice not feasible. The number of possible sensor configurations is

$$N_\mathrm{PC} = \binom{N_\mathrm{C}}{N_\mathrm{M}} = \frac{N_\mathrm{C}!}{(N_\mathrm{C} - N_\mathrm{M})! N_\mathrm{M}!}. \tag{6.2}$$

This formula shows that when the number of sensor candidates, $N_\mathrm{C}$, is increased, the number of possible sensor configurations, $N_\mathrm{PC}$, typically grows very fast – even if the number of sensors to be placed, $N_\mathrm{M}$, is small.

For example, if 8 sensors were to be placed in a grid of $51 \times 51$ candidates, the number of placement configurations is approximately $5.1 \cdot 10^{22}$. For today's computers, it is unfeasible to evaluate the performance of all these configurations. Thus, the sensor placement algorithms should have a strategy to reach or closely reach the global optimum in limited time.

The time constraint is, however, highly relative. If the sensor placement is incorporated into a larger optimisation, for example, into a topology optimisation, the placement algorithm should be as fast as possible. If, on the other hand, the set of the input parameters is fixed and sensor placement only needs to be carried out once, then the algorithm may take considerable time, for example, a week. In that case it should, however, be guaranteed that the algorithm convergences to a well-performing sensor configuration.

### 6.1.3 Placement algorithms in literature

Large number of articles in literature have been devoted to the topic of sensor placement. Fields in which sensor placement is carried out range from wireless sensor networks [85] to structural damage detection [86]. The optimisation goals and accompanying objective functions differ from field to field. Of special interest for shape fitting are the placement algorithms that aim at distinguishing and identifying a set of shapes at the sensor locations.

In the field of experimental dynamics several relevant placement algorithms have been developed. Papadopoulos and Garcia (1998) provided an overview of sensor placement algorithms for dynamic testing and compared the algorithms [87]. Sensor placement algorithms can be typically divided into two major types; the ones that use the inertia and stiffness distribution of a system and the ones that use a set of modeshapes.

The first type includes algorithms based on Guyan reduction, like the one proposed by Penny et al. (1994) [88]. The latter algorithm aims at eliminating the nodes in a FE model for which the inertia forces are small compared to the stiffness forces. Algorithms that make use of the inertia distribution of the system are, however, considered of limited interest, as the inertia distribution in a quasi-statics is not relevant. Especially the second type of algorithms, is of interest, as they might allow for using fitting shapes instead of modeshapes.

Kammer (1990) introduced the Effective Independence (EFI) algorithm that aims at placing the sensors such that a set of target modeshapes is identified as independently as possible [89]. The algorithm makes only use of the modeshapes at the candidate sensor locations. The number of sensors needs to be at least as big as the number of shapes to be identified. Tasker and Liu (1995) presented a placement algorithm with a comparable aim as the one of Kammer [90]. This algorithm is in the following referred to as 'TA'.

Ranieri et al. (2012) and Ranieri et al. (2014) presented two algorithms that were specifically developed for shape fitting, the latter was called 'FrameSense' [75,91]. These algorithms will be referred to as 'RA12' and 'FS'.

As the EFI, TA, RA12 and FS algorithms are considered promising in the context of shape fitting, they are selected for further study in this chapter.

### 6.1.4 Selected shape-based algorithms

This section first describes the assumptions regarding the estimation principles that the selected sensor placement algorithms have in common. Then the objective functions and main working principles of the selected algorithms are introduced.

**Assumptions regarding the estimator** All studied shape-based placement algorithms assume that the field to be estimated is build up from a limited number of shapes and that the number of sensors to be selected equals or exceeds the number of shapes. In that case, the shape amplitudes can be found free from error from the sensors values, as long as the shapes are observable and distinguishable at the sensor locations and the measurements are noiseless.

In practice, the sensor values do contain noise, so that the amplitudes need to be estimated. All algorithms assume that the noise of the different sensors is uncorrelated and identically distributed.

The algorithms aim at finding a sensor configuration with a low estimation error of the amplitudes. In that case, it would be beneficial to add bias to the estimator via regularisation, leading to methodical error but a decreased error due to sensor noise. None of the placement algorithms, however, assumes a biased estimator.

Under the above assumptions, the estimator and estimation error are found as follows. Let the total shape at the sensors, $\boldsymbol{w}_\mathrm{m}$, be described by a limited number of shapes, $\boldsymbol{\Phi}_\mathrm{m}$, that are weighed by shape amplitudes, $\boldsymbol{q}$, i.e.

$$\boldsymbol{w}_\mathrm{m} = \boldsymbol{\Phi}_\mathrm{m}\boldsymbol{q}. \tag{6.3}$$

If the sensor configuration is such that all shapes are observable and identifiable, the shapes' amplitudes can be estimated free of methodical error, using (amongst others) the pseudo-inverse

$$\tilde{\boldsymbol{q}} = \left(\boldsymbol{\Phi}_\mathrm{m}^\mathsf{T}\boldsymbol{\Phi}_\mathrm{m}\right)^{-1}\boldsymbol{\Phi}_\mathrm{m}^\mathsf{T}\tilde{\boldsymbol{w}}_\mathrm{m} = \boldsymbol{\Phi}_\mathrm{m}^{+}\tilde{\boldsymbol{w}}_\mathrm{m}, \tag{6.4}$$

where $\tilde{\boldsymbol{w}}_\mathrm{m} = \boldsymbol{w}_\mathrm{m} + \boldsymbol{\varepsilon}_\mathrm{m}$ is the noise corrupted measurement. The sensor noise, $\boldsymbol{\varepsilon}_\mathrm{m}$, is assumed to be Gaussian distributed with $\mathcal{N}(0, \sigma_\varepsilon)$.

There is, by definition, no methodical error in this case, so that the error of the estimate shape amplitudes,

$$\boldsymbol{\varepsilon}_q = \tilde{\boldsymbol{q}} - \boldsymbol{q} = \boldsymbol{\Phi}_\mathrm{m}^{+}(\boldsymbol{w}_\mathrm{m} + \boldsymbol{\varepsilon}_\mathrm{m}) - \boldsymbol{q} = \boldsymbol{\Phi}_\mathrm{m}^{+}\boldsymbol{\varepsilon}_\mathrm{m}, \tag{6.5}$$

is solely caused by sensor noise, under the assumption that the number of sensors is equal to or higher than number of shapes. The placement algorithms aim to find sensor configurations that minimise this error.

**Effective independence algorithm (EFI, 1990)** Kammer and his co-workers developed the Effective Independence (EFI) algorithm [89, 92], which aims at finding a sensor configuration that best distinguishes a set of shapes. The algorithm was presented in the framework of modal identification, but the algorithm works equally well for other sets of shapes, such as proper orthogonal modes. The number of shapes must be equal or smaller than the number of sensors to be placed. The algorithm optimises by, starting with the full set of candidate sensor locations, one-by-one eliminating the least favourable sensor candidates until the desired number of sensor is reached.

Kammer [89] reasoned that a good estimator minimises the covariance of the estimate error [77],

$$\boldsymbol{Q} = E((\tilde{\boldsymbol{q}} - \boldsymbol{q})(\tilde{\boldsymbol{q}} - \boldsymbol{q})^\mathsf{T}) = \sigma_\varepsilon^2\left(\boldsymbol{\Phi}_\mathrm{m}^\mathsf{T}\boldsymbol{\Phi}_\mathrm{m}\right)^{-1}, \tag{6.6}$$

and thus maximises the Fisher Information Matrix (FIM), which is defined as

$$\frac{1}{\sigma_\varepsilon^2}\boldsymbol{\Phi}_\mathrm{m}^\mathsf{T}\boldsymbol{\Phi}_\mathrm{m}. \tag{6.7}$$

This expression shows that maximising the FIM also maximises the spatial independence of the shapes and the signal strength at the sensors [92].

The sensor placement algorithm aims at maximising the FIM in terms of its determinant. The algorithm starts with the full set of sensor candidates and then successively removes the sensor candidates that contribute least to the rank of the FIM. The authors showed that the contribution of each of the candidates to the rank can be simply found from the diagonal of the matrix

$$\boldsymbol{E} = \boldsymbol{\Phi}_\mathrm{m}\left(\boldsymbol{\Phi}_\mathrm{m}^\mathsf{T}\boldsymbol{\Phi}_\mathrm{m}\right)^{-1}\boldsymbol{\Phi}_\mathrm{m}^\mathsf{T}, \tag{6.8}$$

where the number of rows of $\boldsymbol{\Phi}_\mathrm{m}$ decreases, as in each elimination the row corresponding to the eliminated sensor is removed. Li et al. showed a fast way for calculating the diagonal of $\boldsymbol{E}$, which uses a QR-decomposition to avoid the time intensive inversion in Eq. (6.8) [93].

**Singular value-based algorithm by Tasker and Liu (TA, 1995)**

Tasker and Liu (1995) developed a sensor placement algorithm that, like EFI, starts from the covariance of the estimate error (Eq. 6.6) and the FIM (Eq. 6.7) [90]. Tasker and Liu aimed at minimising the trace of $\boldsymbol{Q}$, i.e. the sum of the diagonal's entries, which equals

$$\mathrm{tr}(\boldsymbol{Q}) = E((\tilde{\boldsymbol{q}} - \boldsymbol{q})^\mathsf{T}(\tilde{\boldsymbol{q}} - \boldsymbol{q})), \tag{6.9}$$

which is equal to the sum of the variances of the shape amplitude errors. The authors showed that the trace of $\boldsymbol{Q}$ decreases by an amount

$$\Delta_i\sigma^2 = \frac{\varphi_i^\mathsf{T}\boldsymbol{Q}\boldsymbol{Q}\varphi_i}{1 - \varphi_i^\mathsf{T}\boldsymbol{Q}\varphi_i} \tag{6.10}$$

when eliminating the $i$th sensor candidate. $\varphi_i$ denotes the $i$th column of $\boldsymbol{\Phi}_\mathrm{m}$. The authors proposed to start with the full set of sensor candidates and to then eliminate one-by-one the sensor candidate with lowest $\Delta\sigma^2$. Using a singular value decomposition, $\Delta_i\sigma^2$ can be efficiently calculated for all $i$ out of the remaining sensor candidates.

**Ranieri (RA12, 2012)**

Ranieri et al. (2012) introduced a sensor placement algorithm that aims at best identifying a set of shapes for the estimation of temperature fields [91]. According to the authors, the optimal set of sensor locations is the one that yields the lowest condition number of matrix $\boldsymbol{\Phi}_\mathrm{m}$

$$\mathrm{cond}(\boldsymbol{\Phi}_\mathrm{m}) = \sigma_\mathrm{max}/\sigma_\mathrm{min},$$

where $\sigma_\mathrm{max}$ and $\sigma_\mathrm{min}$ are the largest and smallest singular value of $\boldsymbol{\Phi}_\mathrm{m}$, respectively. They argued that the condition number can be minimised by removing one-by-one the sensor candidates that show largest correlation. The algorithm

first normalises the rows of the initial $\boldsymbol{\Phi}_\mathrm{m}$ and then calculates a matrix $\boldsymbol{G}$ with the dot products of each of the rows of $\boldsymbol{\Phi}_\mathrm{m}$ with each of the other rows,

$$\boldsymbol{G} = \boldsymbol{\Phi}_\mathrm{m}\boldsymbol{\Phi}_\mathrm{m}^\mathsf{T} - \boldsymbol{I}.$$

Then one-by-one, the $n$th sensor candidate corresponding to the lowest dot product $G_{n,j}$ is eliminated, where $j$ comprises the set of rows corresponding to the still available, other sensor candidates. Note that $\boldsymbol{G}$ is not recalculated during the elimination process.

The authors suggested to find $\boldsymbol{\Phi}_\mathrm{m}$ from the proper orthogonal modes using the POD of the snapshot matrix. As proper orthogonal modes are normal, we will first normalise the columns of $\boldsymbol{\Phi}_\mathrm{m}$.

**FrameSense (FS, 2014)**

Ranieri and his co-workers [9, 75] proposed FrameSense (FS), an algorithm for sensor placement that aims to best identify a set of shapes. They applied the algorithm to the estimation of temperature fields. Like the EFI algorithm, the number of shapes must be equal or smaller than the number of sensor to be placed. Also this algorithm starts with a set of candidate sensor locations and one-by-one eliminates the less favourable candidates until the desired number of sensors is reached.

The FS algorithm aims at finding a placement that leads to a minimum error on the estimate of the shape amplitudes. This corresponds to minimising the expected value of the amplitude estimate error, which the authors called the Mean Square Error (MSE)

$$MSE = E((\tilde{\boldsymbol{q}} - \boldsymbol{q})^\mathsf{T}(\tilde{\boldsymbol{q}} - \boldsymbol{q})) = E\Big(\big(\boldsymbol{\Phi}_\mathrm{m}^{+}\boldsymbol{\varepsilon}_\mathrm{m}\big)^\mathsf{T}\boldsymbol{\Phi}_\mathrm{m}^{+}\boldsymbol{\varepsilon}_\mathrm{m}\Big) = \sigma_\varepsilon^2\big\|\boldsymbol{\Phi}_\mathrm{m}^{+}\big\|_2^2, \tag{6.11}$$

According to the authors, directly optimising for low MSE is difficult. Therefore, they suggest a closely related metric, the Frame Potential (FP):

$$FP(\boldsymbol{\Phi}_\mathrm{m}) = \big\|\boldsymbol{\Phi}_\mathrm{m}\boldsymbol{\Phi}_\mathrm{m}^\mathsf{T}\big\|_2^2. \tag{6.12}$$

Their estimation algorithm aims at minimising the frame potential of $\boldsymbol{\Phi}_\mathrm{m}$ by starting with all sensor candidates and eliminating one-by-one the sensor candidates that, after being eliminated, yield minimum frame potential. Although not given as part of the Algorithm 1 in [75], the rows of $\boldsymbol{\Phi}_\mathrm{m}$ need to be normalised before starting the algorithm, which is mentioned later in the paper. Indeed, we observed that this is essential for the algorithm to yield proper results.

**Maximisation versus minimisation** We argue that the FS algorithm as presented in [9, 75] contains an error, leading to extremely bad sensor configurations. Ranieri et al. [75] described the elimination procedure of FS as "*It is a greedy "worst-out" algorithm: at each iteration it removes the row of* $\boldsymbol{\Psi}$ *that maximally increases the FP. In other words, we define a set of locations* $\mathcal{S}$ *that are not suitable*

*for sensing and at each iteration we add to $\mathcal{S}$ the row that maximizes the following cost function:*" The authors then defined the cost function

$$F(\mathcal{S}) = FP(\boldsymbol{\Psi}) - FP\Big(\boldsymbol{\Psi}_{\mathcal{N}\setminus\mathcal{S}}\Big), \tag{6.13}$$

where $\boldsymbol{\Psi}$ is the $\boldsymbol{\Phi}_{\mathrm{m}}$ corresponding to the initial sensor candidate set and $\mathcal{N} = \{1, \ldots, N\}$ is the set of initial candidates, so that $\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\mathcal{N}}$. However, the then presented FS algorithm prescribes to eliminate in each step the $i$th location that *minimises* the $F(\mathcal{S} \cup i)$. This is in contradiction with the authors' previous description of the procedure. The FS algorithm was also presented in [9, 94], without a description of the optimisation goal, but only the algorithm, again with the minimisation.

The author of this thesis argues that $F(\mathcal{S} \cup i)$ should be *maximised*. $F(\mathcal{S} \cup i)$ can be understood as the decrease of the FP with respect to the initial FP when removing the $i$th sensor location along with the already eliminated locations $\mathcal{S}$. It can be split up into two parts, as follows,

$$F(\mathcal{S} \cup i) = FP(\boldsymbol{\Psi}) - FP\Big(\boldsymbol{\Psi}_{\mathcal{N}\setminus(\mathcal{S}\cup i)}\Big) = a + b,$$

where

$$a = FP(\boldsymbol{\Psi}) - FP\Big(\boldsymbol{\Psi}_{\mathcal{N}\setminus\mathcal{S}}\Big)$$

is the decrease of the FP after eliminating the locations $\mathcal{S}$ and

$$b = FP\Big(\boldsymbol{\Psi}_{\mathcal{N}\setminus\mathcal{S}}\Big) - FP\Big(\boldsymbol{\Psi}_{(\mathcal{N}\setminus\mathcal{S})\setminus i}\Big)$$

the further decrease when eliminating an additional location $i$. As $a$ does not depend on $i$, maximising $a + b$ is equal to maximising $b$. By eliminating the $i^*$th location that maximises the decrease of the FP, the remaining FP is minimised, which is the aim of the authors. Thus, $F(\mathcal{S} \cup i)$ should be *maximised* by eliminating the $i^*$th location

$$i^* = \arg\max_i F(\mathcal{S} \cup i). \tag{6.14}$$

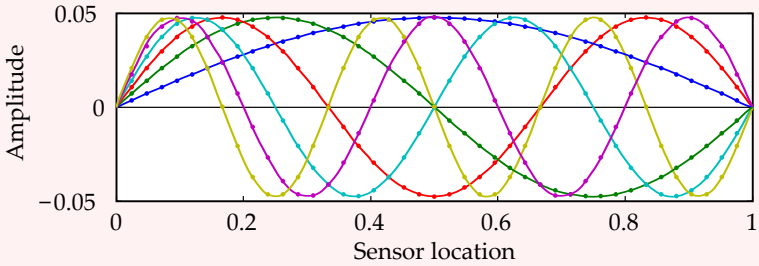
The scaling of the columns of the initial $\boldsymbol{\Psi}$ remains unclear in the papers, although it can significantly influence the sensor placement. [9] proposes a way for obtaining $\boldsymbol{\Psi}$ via the POD. As proper orthogonal modes are normal, we will first normalise the columns of $\boldsymbol{\Psi}$ in studies in this thesis.

Example 6.1 applies FS to a beam, using either a minimisation or a maximisation and with or without the normalisation of the rows of $\boldsymbol{\Psi}$. The results support the statement that a maximisation should be used and that the columns of $\boldsymbol{\Psi}$ should be normalised prior to sensor placement.

## 6.2 Comparison of the sensor placement algorithms

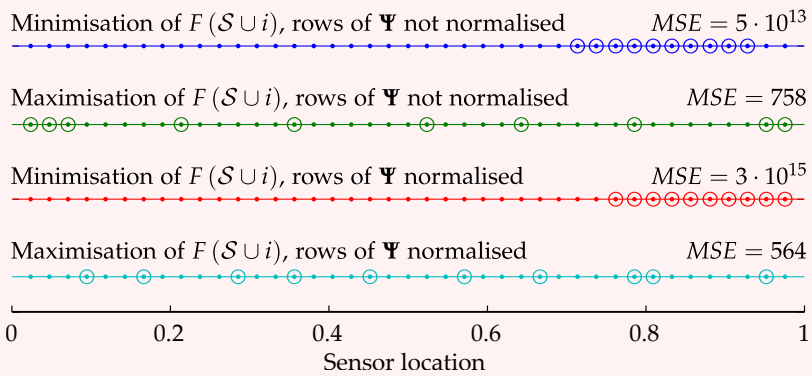
This section compares the performance of the sensor placement algorithms by using numerical experiments. To make a fair comparison, the conditions of the

**Example 6.1: FrameSense – optimisation objective and normalisation**




A shape matrix $\mathbf{\Psi}$ was composed of $K = 6$ shapes, $\psi_i$, with $N = 41$ sensor candidates, as shown in the above figure. The shapes are sinusoidal and thus, by definition, orthogonal. It was chosen to normalise them, so that $\psi_i^T \psi_i = 1$.

Two versions of $\mathbf{\Psi}$ were calculated, one with normalised and the other without normalised rows. Then, the FS algorithm, both in its original and its modified version with the maximisation, was used to find optimised sensor configurations with $L = 10$ sensors. To measure the relative performance of the sensor configurations, their MSEs (Eq. 6.11) were calculated.




The above figure presents the sensor configurations along with their respective MSEs. The configurations found by the FS algorithm with the minimisation have their sensors placed closely together. Intuitively, these configurations seem very weak, as the sensor measure relatively dependent signals. Indeed, in terms of the MSEs of the configurations found with the maximisation perform with 758 and 564 much better than the ones found with the minimisation, which have large MSEs of $5 \cdot 10^{13}$ and $3 \cdot 10^{15}$. Furthermore, the results of the maximisation suggest the normalisation of the rows of $\mathbf{\Psi}$ to be beneficial, as the resulting sensor configuration has a 25 % lower MSE.

experiments will be according to the assumptions of the algorithms in Section 6.1.4; i.e. the number of shapes will be equal or smaller than the number of sensors and the noise of the sensors is uncorrelated and from equal Gaussian distributions.

As the relative performance of placement algorithms may depend on the types of shapes, it is important to test the algorithms with different types of shapes. The algorithms are first applied to shapes with random entries and then more specifically to the modeshapes of a plate both in terms of displacement and strain.

### 6.2.1 Performance metrics

To asses the performance of the algorithms for quantitative comparison, an appropriate performance metric needs to be selected. The metric of Eq. (6.1) is important for shape fitting. However, it is based on a the snapshot matrix, which normally has more shapes than the number of sensors. Thus, another metric that fits the assumptions in Section 6.1.4 should be adopted for the comparison.

Different metrics and objective functions have been proposed in literature. The EFI algorithm [89] optimises for a high determinant of the FIM, $\det\left(1/\sigma_\varepsilon^2 \cdot \boldsymbol{\Phi}_\mathrm{m}^\mathsf{T}\boldsymbol{\Phi}_\mathrm{m}\right)$. Penny [88] proposed assessment of the sensor configuration based on the ratio between the maximum and minimum singular value of $\boldsymbol{\Phi}_\mathrm{m}$, which equals the condition number. Also Ranieri et al. [91] proposed the condition number of $\boldsymbol{\Phi}_\mathrm{m}$, $\mathrm{cond}(\boldsymbol{\Phi}_\mathrm{m}) = \sigma_\mathrm{max}/\sigma_\mathrm{min}$. Tasker and Liu [90] optimised for minimum trace, $\mathrm{tr}\left(\sigma_\varepsilon^2\left(\boldsymbol{\Phi}_\mathrm{m}^\mathsf{T}\boldsymbol{\Phi}_\mathrm{m}\right)^{-1}\right)$. Ranieri et al. [75] used the MSE, $\sigma_\varepsilon^2\left\|\boldsymbol{\Phi}_\mathrm{m}^+\right\|_2^2$.

The MSE is of interest, as it has a direct interpretation in terms of the estimation error of the shape amplitudes. More specifically, it is proportional to the expected value of the Root Mean Square (RMS) error of the shape amplitudes estimates. The RMS error of the amplitude estimates is defined as

$$\varepsilon_{q,\mathrm{rms}} = \sqrt{\frac{1}{N_\mathrm{q}}(\tilde{\boldsymbol{q}} - \boldsymbol{q})^\mathsf{T}(\tilde{\boldsymbol{q}} - \boldsymbol{q})} \tag{6.15}$$

and, using Eq. (6.5), its expected value can be written as

$$\sigma_q^2 = E\left(\varepsilon_{q,\mathrm{rms}}^2\right) = E\left(\frac{1}{N_q}\left(\boldsymbol{\Phi}_\mathrm{m}^+\boldsymbol{\varepsilon}_\mathrm{m}\right)^\mathsf{T}\boldsymbol{\Phi}_\mathrm{m}^+\boldsymbol{\varepsilon}_\mathrm{m}\right) = \frac{1}{N_q}\sigma_\varepsilon^2\left\|\boldsymbol{\Phi}_\mathrm{m}^+\right\|_2^2, \tag{6.16}$$

with $N_q$ the number of shapes. Indeed this result equals the MSE, apart from the constant $1/N_q$.

In a comparable fashion, an objective function in terms of the error of the estimates can be defined, which is of more interest in the context of shape fitting. The error of the estimates is

$$\boldsymbol{\varepsilon}_w = \tilde{\boldsymbol{w}}_\mathrm{t} - \boldsymbol{w}_\mathrm{t} = \boldsymbol{\Phi}_\mathrm{t}\boldsymbol{\Phi}_\mathrm{m}^+\boldsymbol{\varepsilon}_\mathrm{m}, \tag{6.17}$$

where the subscript 't' denotes that only the subset of rows of $\boldsymbol{w}_\mathrm{t}$ and $\boldsymbol{\Phi}_\mathrm{t}$ of target points is selected, the points that are of direct interest for the estimation.

The RMS value of $\boldsymbol{\varepsilon}_w$ is

$$\varepsilon_{w,\text{rms}} = \sqrt{\frac{1}{N_w}\boldsymbol{\varepsilon}_w^\mathsf{T}\boldsymbol{\varepsilon}_w}, \tag{6.18}$$

with $N_w$ the number of target points. The expected value of $\varepsilon_{w,\text{rms}}$ is

$$\sigma_w^2 = E\left(\varepsilon_{w,\text{rms}}^2\right) = E\left(\frac{1}{N_w}\left(\boldsymbol{\Phi}_\text{t}\boldsymbol{\Phi}_\text{m}^+\boldsymbol{\varepsilon}_\text{m}\right)^\mathsf{T}\boldsymbol{\Phi}_\text{t}\boldsymbol{\Phi}_\text{m}^+\boldsymbol{\varepsilon}_\text{m}\right) = \frac{1}{N_w}\sigma_\varepsilon^2\left\|\boldsymbol{\Phi}_\text{t}\boldsymbol{\Phi}_\text{m}^+\right\|_2^2. \tag{6.19}$$

From this, we define a performance metric called 'Sensor Noise Amplification' (SNA):

$$SNA = \frac{\sigma_w}{\sigma_\varepsilon} = \sqrt{\frac{1}{N_w}\left\|\boldsymbol{\Phi}_\text{t}\boldsymbol{\Phi}_\text{m}^+\right\|_2^2}. \tag{6.20}$$

This metric can be understood as the amplification factor with which the sensor noise level is multiplied when obtaining the estimate from the sensor readings. The unit of SNA depends on the units of $\sigma_w$ and $\sigma_\varepsilon$. For example, if $\sigma_w$ and $\sigma_\varepsilon$ are they are both in terms of displacement, the SNA is dimensionless; if they are in terms of displacement and strain, respectively, the unit of SNA is metre. The SNA does not depend on the scaling of the columns of $\boldsymbol{\Phi}$, as opposed to the metric $\sigma_q$ and the MSE. This is an advantage, as the SNA is, as such, a more objective metric than the MSE.

The SNA is tailored to the assumptions of the sensor placement algorithms under review. When, however, used outside these assumptions, it should be noted that the SNA does not consider the methodical error due to possible spill-over effects if the number of shapes to be identified exceeds the number of sensors. Furthermore, as the SNA is a measure of the estimate's noise level caused by the sensor noise, it should be compared to the quantity to be estimated for evaluating whether or not the noise level is acceptable.

### 6.2.2 Placement for random shapes

Ranieri et al. [75] compared the placement performance of their FS algorithm with the performance of three other sensor placement algorithms and random placement. The performance was tested with different types of a shape matrices with random entries. The advantage of random shapes is that the placement results can be easily reproduced and thus allow for fast comparison.

#### Method

Four types of random matrices were used: matrices with a Gaussian distribution $\mathcal{N}(0,1)$, matrices with a Gaussian distribution $\mathcal{N}(0,1)$ whose rows were normalised, matrices with a Gaussian distribution $\mathcal{N}(0,1)$ whose columns were orthonormalised and matrices with a Bernouilli distribution $1/30\mathcal{B}(1,1/2)$ (i.e. having matrix entries with equal probability of 0 or $1/30$). The number shapes was 30, the number of sensor candidates was in all cases 100 and the number of sensors to be placed was varied between 35 en 60.

**Results**

Figure 6.1 shows the performance of the different algorithms applied to the $100\times30$ matrices with the Gaussian distribution $\mathcal{N}(0,1)$. The results of Figure 6.1a are copied from Ranieri et al. [75]. They investigated four algorithms: FS, 'Determinant', 'Mutual information', 'MSE' and random sensor placement. Figure 6.1b is from our own experiments with the three other algorithms we selected, i.e. EFI, TA and RA12. Each experiment was carried out 100 times; the standard deviation is indicated with dashed lines. In a similar way, the results for the other types of matrices are presented in Figures 6.2, 6.3 and 6.4.

The figures show that the RA12 algorithm performs relatively poor for all four matrix types. Although their FS algorithm in most cases outperforms the other algorithms the authors compared it to, our results show that EFI and TA perform at least as good and in most cases better than FS.

TA performs systematically better than the EFI algorithm. Interestingly enough, TA and EFI show also good results for the Bernoulli matrices, unlike the other methods, that not considerably exceed the performance of the random sensor placement.

### 6.2.3 $(x,y)$-placement of position sensors to identify mode-shapes

This section compares the performance of the sensor placement algorithms for identifying a set of modeshapes of a square plate.



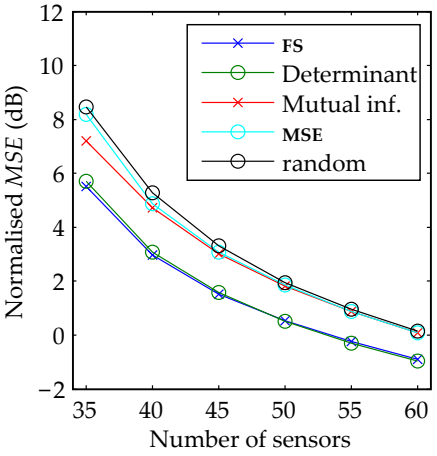
**Method**

Figure 6.5 shows the out-of-plane modeshapes of a plate that correspond to the plate's lowest eigenfrequencies. Generally speaking, those are the modeshapes that are most important for describing the plate's displacement field. It was chosen to distinguish eight modeshapes, of which three Rigid Body (RB) modeshapes. The selected sensor placement algorithms had to select 8 position sensors from a grid of $51\times51$ sensor candidates.

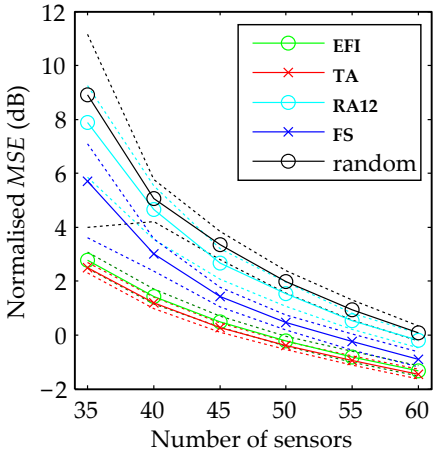
**Results**

Figure 6.6 shows the sensor configurations found by the four algorithms. The EFI and TA algorithms yield the same placement configuration as the hand-picked one in Chapter 2, namely sensors at the chuck's corners and centres of the sides. With SNAs of 1.6 and 1.2, respectively, the performance of the sensor configurations of RA12 and FS are weaker than the configuration of the others ($SNA = 0.86$).

Figure 6.7 plots, as function of position, the order in which the sensor candidates were eliminated by each of the three algorithms. The EFI and TA algorithm show a relatively smooth elimination order. The elimination orders of RA12 and FS shows relatively arbitrary patterns, with early and late removed sensor candidates closely spaced together. This suggests that the objective

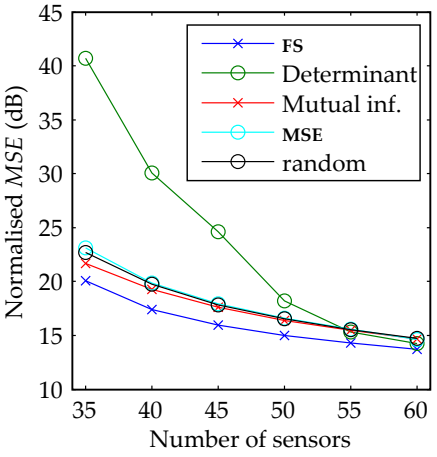



(a) Results of four placement algorithms and random sensor placement, all adopted from [75].

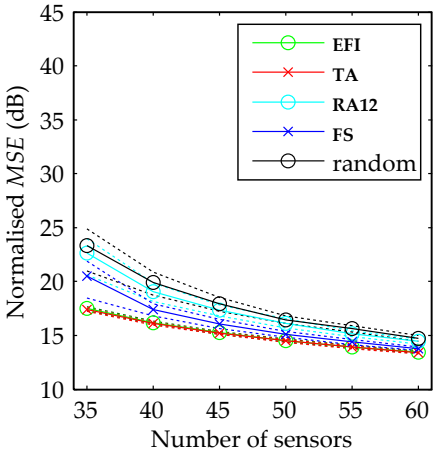



(b) Experimental results of this thesis with FS and three additional placement algorithms. The dashed lines show the standard deviation over 100 experiments.

Figure 6.1: The performance of the different sensor placement algorithms and random sensor placement when applied to matrices with a Gaussian distribution $\mathcal{N}(0,1)$. The performance is expressed in terms of the MSE of the optimised configuration.

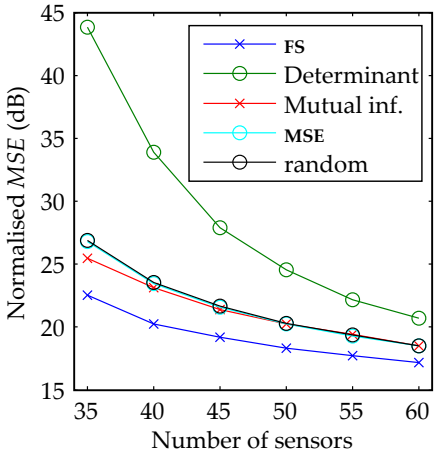



(a) Results adopted from [75].

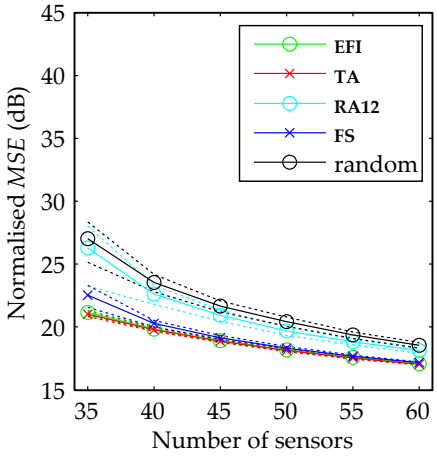



(b) Experimental result of this thesis.

Figure 6.2: Results comparable to Figure 6.1, but using matrices with a Gaussian distribution $\mathcal{N}(0,1)$ whose rows were normalised.



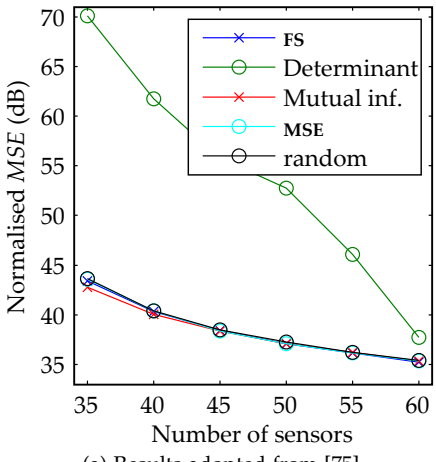

(a) Results adopted from [75].



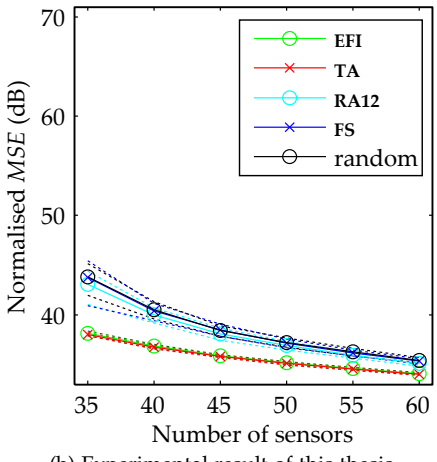

(b) Experimental result of this thesis.

Figure 6.3: Results comparable to Figure 6.1, but using matrices with a Gaussian distribution $\mathcal{N}(0,1)$ whose columns were orthonormalised.






(a) Results adopted from [75].




(b) Experimental result of this thesis.

Figure 6.4: Results comparable to Figure 6.1, but using matrices with a Bernouilli distribution $1/30\mathcal{B}(1,1/2)$.

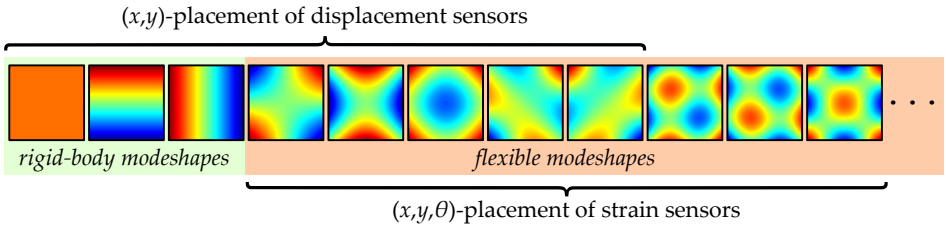
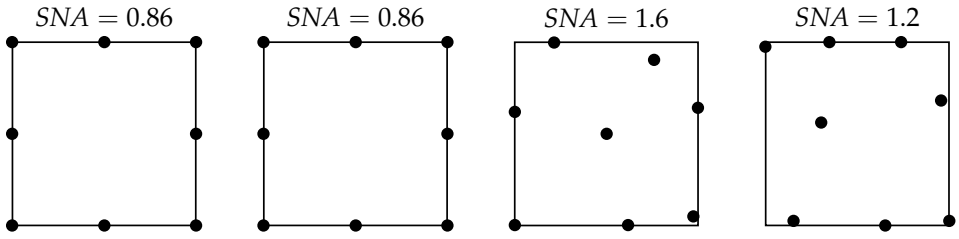



Figure 6.5: The placement of the position sensors was optimised for distinguishing three RB modeshapes and five flexible modeshapes of a square plate. For the placement of the strain sensors, eight flexible modeshapes had to be distinguished. All mode-shapes were out-of-plane.




(a) EFI algorithm. (b) TA algorithm. (c) RA12algorithm. (d) FS algorithm.

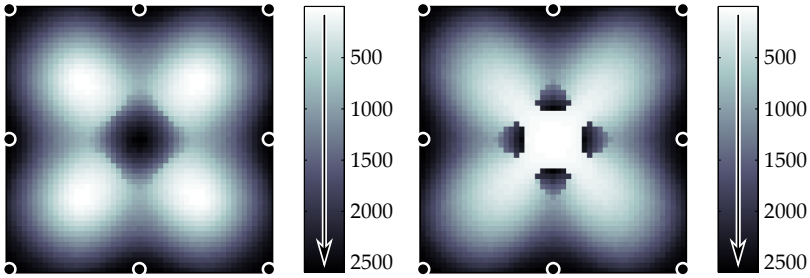
Figure 6.6: The sensor configurations obtained with the four placement algorithms from literature. The sensor candidate set consisted of a 51 × 51 grid. To performance metric SNA of each configuration, here in terms of displacement over displacement, is indicated.

function of RA12 and FS is less adequate than those of the other algorithms, which could be a reason for the weaker sensor configurations of RA12 and FS.

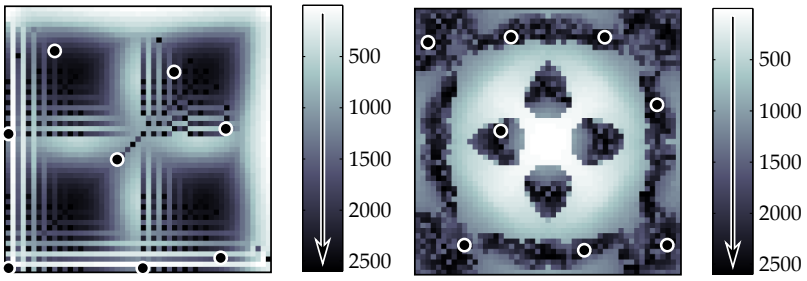
### 6.2.4 $(x, y, \theta)$-placement of strain sensors to identify modeshapes

In the last section, the sensor placement algorithms were compared in a case in which eight position sensors had to distinguish eight modeshapes of a plate. In this section, a similar comparison is made, this time based on strain sensors. As strain sensors can not observe RB modes, eight flexible modes were selected to be distinguished (Figure 6.5).

According to Section 2.4.3, the strain sensors can be arbitrarily placed over the wafer chuck's surface. Unlike displacement measurement, strain measurement is direction dependent. Therefore, the set of sensor candidates in $x$- and $y$-direction has to be extended with the $\theta$-direction, which denotes the strain sensor's orientation.




(a) EFI algorithm. (b) TA algorithm.




(c) RA12algorithm. (d) FS algorithm.

Figure 6.7: The elimination order of the sensor candidates as function of their position $(x, y)$ on the plate. The finally obtained sensor locations are indicated with circles. The shape set for which the placement was performed comprised eight modeshapes of a plate, including the RB modeshapes, as shown in Figure 6.6.

As it unknown how the candidate set's size influences the performance of different algorithms, candidate sets of different sizes were investigated. Also the speed of the estimation algorithms is investigated, as candidate sets with three directions get easily large.

### Method

The number of candidates in the plane ($x/y$-direction) was varied from $1^2$ to $31^2$ and the number of sensor orientation ($\theta$-direction) candidates was varied from 1 to 21. Given a candidate set, a matrix $\boldsymbol{\Phi}_{\mathrm{m,full}}$ consisting of the strain amplitudes of the sensor candidates was calculated using Eq. (2.19). Then, the placement algorithms were applied to $\boldsymbol{\Phi}_{\mathrm{m,full}}$ to obtain the optimised sensor configurations, with both 8 and 24 sensors. Finally, the SNA of each sensor configuration was calculated.

### Results

As the performance of the sensor configurations differed highly, the are presented in two ways. Firstly, the SNAs are presented as function of the number of sensor candidates on a linear, saturated scale, allowing a comparison of the algorithms and the size of the candidate set in terms of the best performing sensor configurations. Secondly, the results are presented on a logarithmic scale, allowing comparison of the algorithms in terms of their spread in performance and outliers.



Table 6.1 shows the performance of the obtained configuration as function of the size of candidate set. The placement results depend on the number of sensors that are kept after the candidate elimination process. The first column of the table corresponds to 8 sensors (equal to the number of shapes to identify), the second column corresponds to 24 sensors.

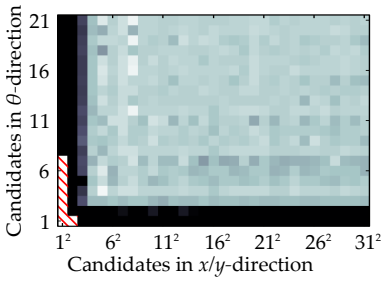
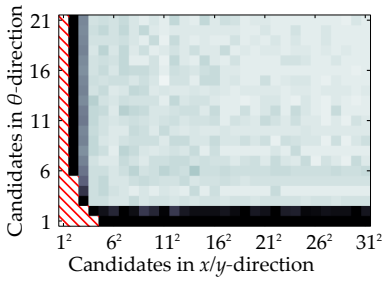
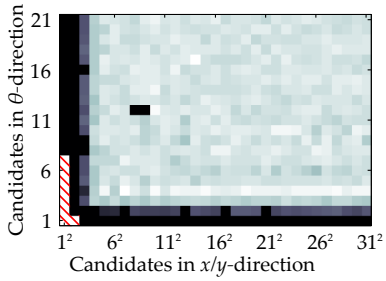
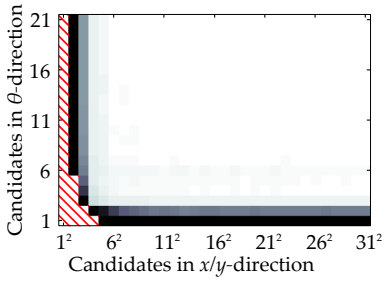
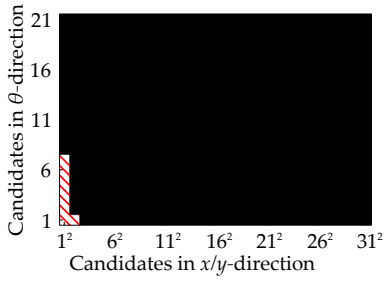
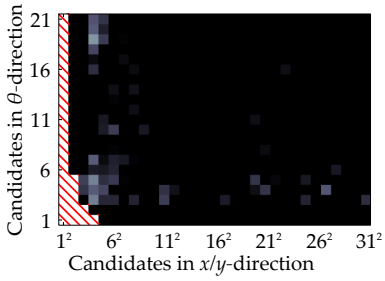
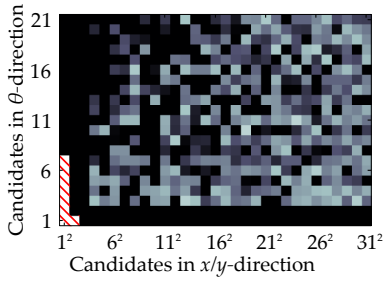
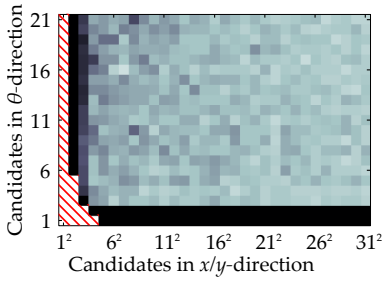
The performance is given in terms of the SNA, according to Eq. (6.20). As the SNAs are in this case in terms of displacement over strain, their unit is metre. Note that the colour scales are saturated to be able to present high detail, so that outliers can not be identified in the plots.

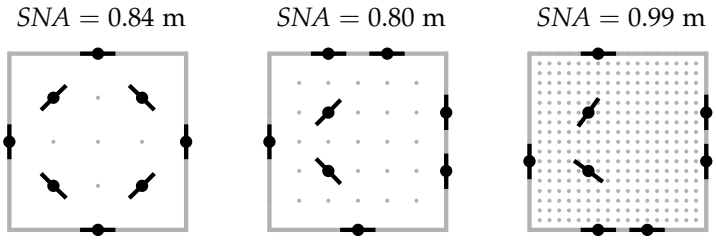
Overall, the TA algorithm leads to better placement results than the EFI algorithm, which in turn performs better than the FS and the RA12 algorithm. To expect good results, the number of candidates in $x$- and $y$-direction and in $\theta$-direction should be larger than 3. The RA12 algorithm performs better for a larger numbers of $x/y$-candidates. Surprisingly, the TA and EFI algorithm yield better results when the number of candidates is smaller than when the candidate set is larger.

Figure 6.8 shows for each algorithm the best configuration of eight strain sensors that was obtained for the different candidate sets. The configurations show clear similarities, some sensors are placed at and in parallel to the sides of the plate, the others are placed under an angle of around 45° at the plate's diagonals.

Figure 6.9 shows cumulative histograms of the performance of the sensor configurations, making it possible to study the outliers. The configurations

Table 6.1: The performance of the strain sensor placement algorithms as function of the size of the set of sensor candidates. The sets were build up from a grid of candidates in $x/y$-direction and a number of orientations of the strain sensors ($\theta$-direction). The table's columns correspond to a final sensor configuration of, respectively, 8 and 24 sensors. The SNA is in terms of displacement over strain and has as unit *metre*.

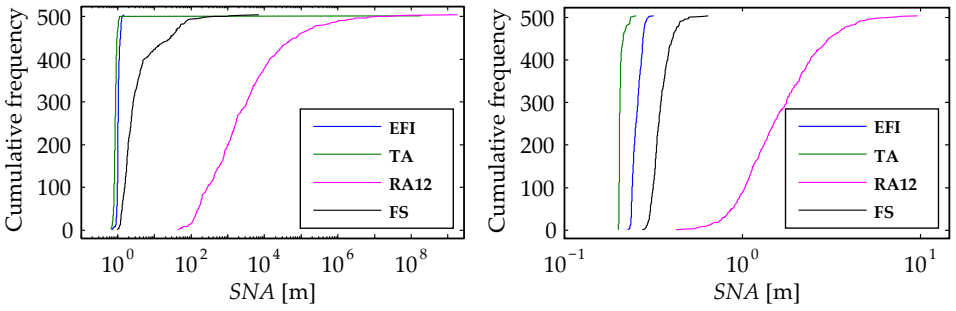
| Placement algorithm | Sensor configuration performance in terms of the SNA using 8 sensors | using 24 sensors |
|---|---|---|
| | 0.8 1 1.2 1.4 1.6 1.8 [m] | 0.5 0.6 0.7 0.8 0.9 [m] |
| **EFI** |  |  |
| **TA** |  |  |
| **RA12** |  |  |
| **FS** |  |  |




(a) EFI algorithm. (b) TA algorithm. (c) FS algorithm.

Figure 6.8: Optimised configuration of eight strain sensors for identifying eight mode-shapes of a plate as found by different placement algorithms. The optimised configurations correspond to the best ones of Table 6.1.






(a) Placement of 8 sensors. (b) Placement of 24 sensors.

Figure 6.9: Cumulative histograms of the performance in terms of the SNA of the strain sensor configurations found by the different algorithms. In total 504 configurations are taken into account. The configurations less than 3 candidates in $\theta$-direction and less than $3 \times 3$ in $x/y$-direction are omitted in the plots.

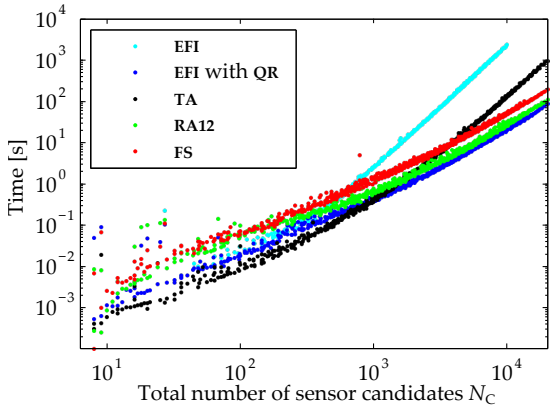



Figure 6.10: The execution time of sensor placement algorithms as function of the size of the candidate sets. The candidate sets correspond to the ones of Table 6.1. The EFI algorithm was implemented both using its standard formulation and a faster implementation using a QR decomposition. The algorithms were executed on a laptop with a clock speed of 1.6 GHz.

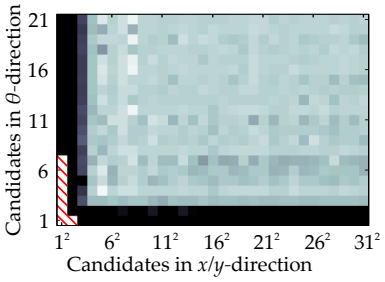
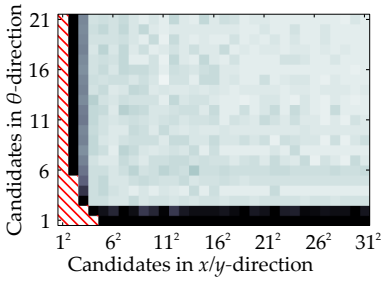
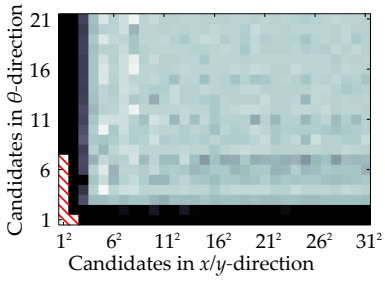
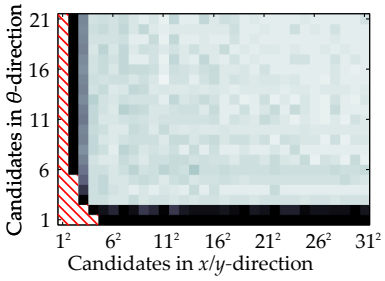
with less than 3 candidates in $\theta$-direction and less than $3 \times 3$ in $x/y$-direction are omitted, as these correspond to relatively small candidate sets and perform systematically weak.

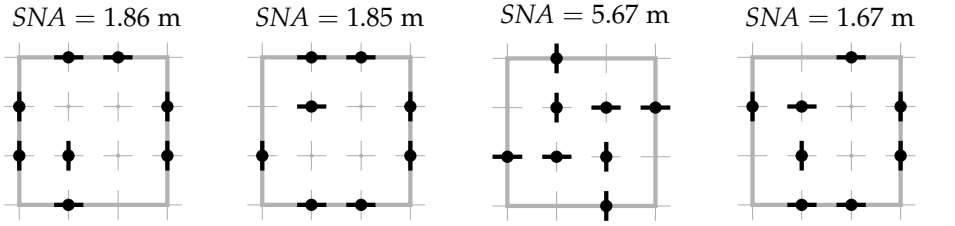
The EFI and the TA algorithm yield configurations with relatively close SNAs. The TA algorithm yielded, however, some outliers with high SNA. When placing 8 sensors, the FS algorithm produces relatively many outliers ($SNA$>10). The figures show that the RA12 algorithm often produces very weak configurations ($SNA$>$10^3$). As the range of the SNAs of the RA12 and FS algorithm are relatively wide, these algorithms can be said to be less robust in terms of converging to the optimum than EFI and TA.

Based on these observations, a placement strategy is proposed that applies the TA algorithm on several different smaller candidate sets instead of a single large set. This increases the chance of good converging and of finding a high-performing sensor configuration, especially when taking into account that placement with small sensor candidate sets is much faster than with large sets. For example, placement with the TA algorithm for a single large set with $31 \times 31 \times 21$ candidates took 220 s on a laptop with a clock speed of 1.6 GHz. Placement for all 113 candidate sets with up to $11 \times 11 \times 11$ candidates took only 142 s.

**Execution time** The time needed to execute the placement algorithms on a standard laptop a laptop with a clock speed of 1.6 GHz was monitored. Figure 6.10 shows the execution time as function of the number of candidates. As the EFI algorithm involves inversion of a matrix, it is relatively slow if the number of candidates exceeds 1000. An implementation using a QR-decomposition, as described by Li et al. [93], reduces the execution time significantly, so that EFI becomes the fastest algorithm of the three.

Table 6.2: The performance of the strain sensor placement for as function of the size of the set of sensor candidates. The sets were build up from a grid of candidates in $x/y$-direction and a number of orientations of the strain sensors ($\theta$-direction).

| Placement algorithm | Sensor configuration performance in terms of the SNA<br>using 8 sensors<br>0.8 1 1.2 1.4 1.6 1.8 [m] | using 24 sensors<br>0.5 0.6 0.7 0.8 0.9 [m] |
|---|---|---|
| **EFI** (1 elimination per iteration) |  |  |
| **EFI** (1 % elimination per iteration) |  |  |




(a) EFI algorithm. (b) TA algorithm. (c) FS algorithm. (d) Brute force optimum.

Figure 6.11: Comparison of the optimised sensor configurations found by the placement algorithms and the brute force optimum. The sensor candidate sets consisted of 4 candidates in $x$- and $y$- and 2 in $\theta$-direction.

A speed improvement can be obtained by eliminating more than one sensor per iteration step. For example, when eliminating 1 % of the sensor candidates per step with the EFI algorithm, the placement performance did not alter significantly (Table 6.2).

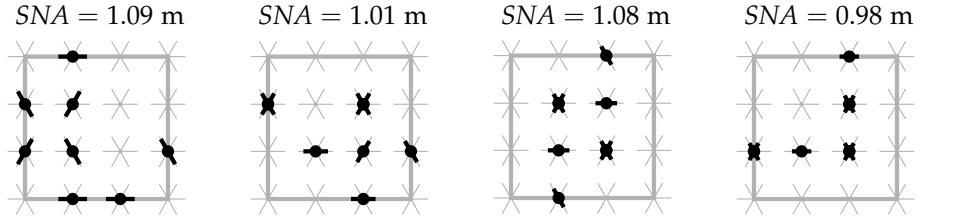
### Closeness to the global optimum

The performance of a sensor configuration can be rated using the SNA, allowing to compare the configurations. However, as the globally optimum sensor configuration is not known, the closeness of an optimised configuration to this global optimum is also unknown. To get an idea how close the placement algorithms approach the performance of the best sensor configuration for a given sensor candidate set, brute force placement was performed on two different sensor candidate sets.

The first candidate set consisted of $4 \times 4$ candidates in $x/y$- and 2 in $\theta$-direction, leading to a total of $4^2 \cdot 2 = 32$ candidates. Eight sensors had to be selected out of the candidate sets. For all $10 \cdot 10^6$ possible sensor configurations the SNA was calculated. The same was procedure was carried out with the second – larger – candidate set, that consisted of 3 instead of 2 candidates in $\theta$-direction, leading to $4^2 \cdot 3 = 48$ candidates and $377 \cdot 10^6$ possible sensor configurations.

Figure 6.11 shows the sensor configurations found by the placement algorithms and the sensor configuration corresponding to the globally optimal SNA for the first candidate set and Figure 6.12 the ones for the second candidate set. Although the configurations differ significantly, they all have a SNA within 12 % of the global optimum, except for the SNA of FS with the first candidate set. The configuration found by the TA algorithm approaches the global optimum SNAs the closest, to within 11 % for the first and 3 % for the second candidate set.

For this sensor candidate set, the methods indeed approach the global optimum closely. Thus, one may hypothesise that this is also the case for larger candidate sets.




(a) EFI algorithm. (b) TA algorithm. (c) FS algorithm. (d) Brute force optimum.

Figure 6.12: Comparison of the optimised sensor configurations found by the placement algorithms and the brute force optimum. The sensor candidate sets consisted of 4 candidates in $x$- and $y$- and 3 in $\theta$-direction.

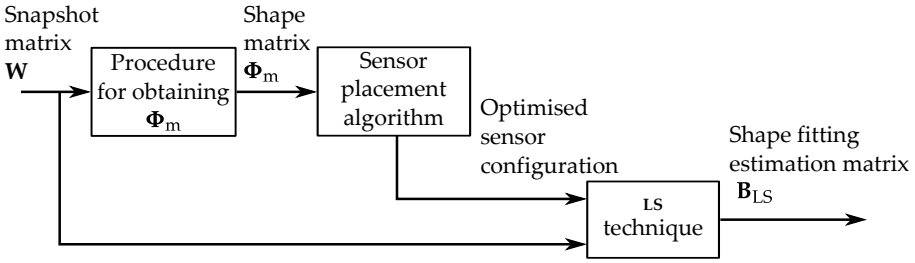



Figure 6.13: The steps needed to obtain the SF estimation matrix, $\boldsymbol{B}_{\text{LS}}$. The LS technique makes use of the optimised sensor locations that are found using sensor placement algorithms from literature. A procedure extracts from snapshot matrix $\boldsymbol{W}$ the patterns at the sensor candidates that correspond to the most prominent deformation patterns at the target points. The resulting shape matrix $\boldsymbol{\Phi}_{\text{m}}$ is then used for sensor placement.

## 6.3 Integration of sensor placement in shape fitting

In the previous sections sensor placement algorithms were selected and assessed numerically. This section shows how the placement algorithms can be integrated into the SF framework.

Figure 6.13 shows the proposed process for obtaining the SF estimation matrix $\boldsymbol{B}_{\text{LS}}$ from snapshot matrix $\boldsymbol{W}$. The SF estimator is based on both the snapshot matrix and the optimised sensor configuration. The sensor configuration is found using a sensor placement algorithm, which also makes use of the snapshot matrix. However, the snapshot matrix should first be made suitable for applying the selected sensor placement algorithms. For this purpose a procedure is developed for obtaining, based upon the snapshot matrix, a shape matrix $\boldsymbol{\Phi}_{\text{m}}$ which can directly be used as input for the sensor placement algorithms. This procedure consist of four steps, i.e. 1) the separation of the flexible residues from the snapshot matrix; 2) conversion to a real snapshot matrix, in case the snapshot matrix is complex; 3) the extraction of most prominent shapes from the snapshot matrix; and 4) the addition of the flexible modeshapes.

**Procedure to obtain shape matrix $\boldsymbol{\Phi}_{\mathrm{m}}$**

**Separation of the flexible residues** The displacements in snapshot matrix $\boldsymbol{W}$ are build up from RB contributions $\boldsymbol{W}_{\mathrm{rb}}$ and flexible contributions $\boldsymbol{W}_{\mathrm{fl}}$. These contributions need to be separated, such that the flexible part of the snapshot matrix, $\boldsymbol{W}_{\mathrm{fl}}$, is left.

**Conversion to a real snapshot matrix** If $\boldsymbol{W}_{\mathrm{fl}}$ is complex, here denoted with $\boldsymbol{W}_{\mathrm{fl}}^{\star}$, the complex values capture both the relative magnitudes and phase of the signals. The sensor placement algorithms, however, may not be able to deal with complex numbers, so that the phase relations need to be captured with real numbers only. This can be achieved by defining a new matrix consisting of real valued snapshots that are recombinations of the snapshot shapes with random magnitude and phase. $\boldsymbol{W}_{\mathrm{fl}} = \Re(\boldsymbol{W}_{\mathrm{fl}}^{\star}\boldsymbol{A})$, with $\boldsymbol{A}$ a $N_{\mathrm{D}}^{\star} \times N_{\mathrm{D}}$-matrix consisting of circularly-symmetric complex random elements $a_{ij} \sim \mathcal{CN}(0,1)$. $N_{\mathrm{D}}^{\star}$ is the original number of shapes in the snapshot matrix, $\boldsymbol{W}_{\mathrm{fl}}^{\star}$, and $N_{\mathrm{D}}^{\star}$ is the desired number of shapes in the real valued snapshot matrix $\boldsymbol{W}_{\mathrm{fl}}^{\star}$. $N_{\mathrm{D}}$ should be sufficiently higher than $N_{\mathrm{D}}^{\star}$ to capture magnitude and phase distribution of the original shapes well.

**Extraction of most prominent shapes** The sensor placement algorithms that were selected in Section 6.1 are typically tailored to distinguish a number of shapes that is equal or smaller than the number of sensors ($N_{\mathrm{F}} \leq N_{\mathrm{M}}$). However, the number of snapshots in the snapshot matrix easily exceeds the number of sensors and in that case the size of matrix $\boldsymbol{W}_{\mathrm{fl}}$ need to be reduced. This can be achieved using the Proper Orthogonal Decomposition (POD) [9, 78], which extracts the most prominent shapes from a matrix (Appendix D).

Obtaining the most prominent shapes is, however, not a trivial step, as the subset of target points to be estimated is not necessarily equal to the subset of measurement points. The part of the shape of which the most prominent patterns need to be identified consists of the target points. Those points need to be included in the POD. The measurement points, however, are the points for which sensor placement has to be carried out. Thus, a procedure is needed to find a matrix $\boldsymbol{\Phi}_{\mathrm{m,fl}}$ that describes the shapes at the sensor points corresponding to the most prominent shapes at the target points based on snapshot matrix $\boldsymbol{W}$.

The POD of matrix $\boldsymbol{W}_{\mathrm{t,fl}}$ (the subset of rows of $\boldsymbol{W}_{\mathrm{fl}}$ corresponding to the target points) is calculated, yielding the POMs $\breve{\boldsymbol{\varphi}}_i$. The most dominant POMs are collected in a matrix as follows,

$$\breve{\boldsymbol{\Phi}}_{\mathrm{t,fl}} = \begin{bmatrix} \breve{\boldsymbol{\varphi}}_1 & \cdots & \breve{\boldsymbol{\varphi}}_{N_{\mathrm{F}}} \end{bmatrix}, \tag{6.21}$$

where $N_{\mathrm{F}}$ needs to be equal to or smaller than the desired number of sensors to be placed. The rows of matrix $\breve{\boldsymbol{\Phi}}_{\mathrm{t,fl}}$ correspond to the target points, but the sensor placement algorithm needs only information of the sensor candidate points. This can be solved by determining how the columns of $\breve{\boldsymbol{\Phi}}_{\mathrm{t,fl}}$ can be formed from linear combinations of $\boldsymbol{W}_{\mathrm{t,fl}}$ and then recombining the columns of $\boldsymbol{W}_{\mathrm{m,fl}}$ according to the thus found weights. The weights are found from the pseudo-inverse of $\boldsymbol{W}_{\mathrm{m,fl}}$

multiplied by $\check{\boldsymbol{\Phi}}_{t,fl}$:

$$\check{\boldsymbol{\Phi}}_{m,fl} = \boldsymbol{W}_{m,fl}\boldsymbol{W}_{t,fl}^{T}\left(\boldsymbol{W}_{t,fl}\boldsymbol{W}_{t,fl}^{T}\right)^{-1}\check{\boldsymbol{\Phi}}_{t,fl}. \tag{6.22}$$

**Addition of the flexible modeshapes** Matrix $\check{\boldsymbol{\Phi}}_{m,fl}$ consists only of the flexible contributions to the displacements at the sensor locations. As the RB coordinates also have to be estimated from the sensor signals, the RB shapes at the sensor locations, $\check{\boldsymbol{\Phi}}_{m,rb}$, need to be added to the set of flexible shapes,

$$\check{\boldsymbol{\Phi}}_{m} = \begin{bmatrix} \check{\boldsymbol{\Phi}}_{m,rb} & \check{\boldsymbol{\Phi}}_{m,fl} \end{bmatrix}. \tag{6.23}$$

Shape matrix $\check{\boldsymbol{\Phi}}_{m}$ can then be used as input for the sensor placement algorithm.

**In time changing disturbances**

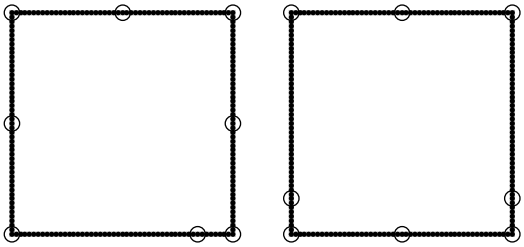
Still, an important issue concerning the integration of sensor placement is that the placement algorithms only optimise for a single set of shapes. However, in the case in which the disturbances are caused by the immersion film, the disturbance conditions change as function of time. Thus, also the snapshot matrix and the optimal sensor configuration change. Therefore, a sensor configurations needs to be found that, although possibly being suboptimal for each of the individual disturbance conditions, over time leads to an – on average – minimum estimation error.

As a solution, the snapshot matrix could be formed based on all disturbances conditions, neglecting their time dependency, or based on a limited number of conditions that are expected to lead to the largest estimation error. In both cases, however, some foreknowledge on the disturbances is lost.

## 6.4 Sensor placement applied to the numerical study of the wafer chuck

In Chapters 4 and 5 a case study was presented, in which shape fitting was applied to the wafer chuck and its main disturbance source, the disturbances introduced by the immersion layer. The case study used the hand-picked sensor configuration of Section 2.4.3, which consists of eight position sensors at the corners and the side centres. This section applies sensor placement to the case study, using the procedure of Section 6.3 and shows that the estimation error indeed significantly reduces.

**Method** Shape matrix $\check{\boldsymbol{\Phi}}_{m}$ was calculated from snapshot matrix $\boldsymbol{W}^{\star}$ using the procedure of last section. As the lens column and its disturbances travel with respect to the wafer chuck, the snapshot matrix changes as function of PoI's location. The first row of Table 6.3 reproduces the estimation results of Chapter 5 that were obtained with the hand-picked sensor configuration. The highest methodical estimation error of shape fitting is mainly concentrated at the centre



(a) EFI algorithm. (b) TA algorithm.

Figure 6.14: Two new position sensor configurations found with the EFI and TA algorithms. The configurations are optimised for disturbances at the centre of the wafer chuck.

of the plot, corresponding to the PoI being at the wafer chuck's centre. For this reason, it was chosen to avoid the problem of multiple snapshots matrix, by taking only into account the snapshots corresponding to the centre PoI location. Thus, $\boldsymbol{W}^\star$ consisted of in total 6900 shapes (69 spatial points times 100 frequency points). The sensor candidates were restricted to the sides of the wafer chuck.

From the complex valued $\boldsymbol{W}^\star$ a real valued snapshot matrix $\boldsymbol{W}$ was generated consisting of $10^4$ snapshots. After obtaining the shape matrix $\check{\boldsymbol{\Phi}}_\mathrm{m}$ the EFI and TA sensor placement algorithms were applied. The algorithms were modified to not eliminate the four sensor candidates at the chuck's corners. Then in the same manner as in Chapter 5 the estimation errors were obtained for the two new sensor configurations.



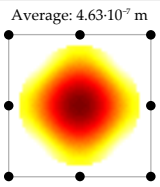
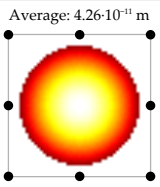
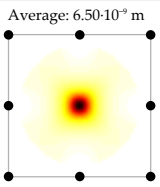
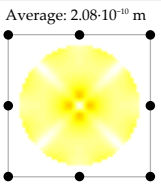
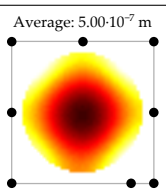
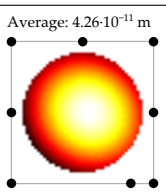
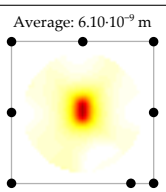
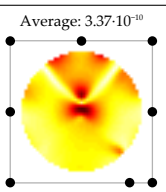
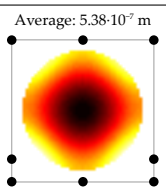
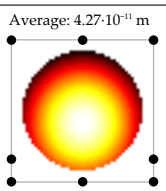
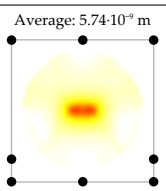
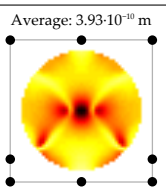
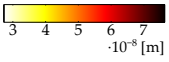
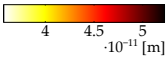
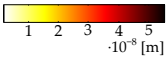
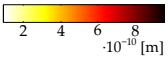
**Results** Figure 6.14 shows the two optimised sensor configurations. As compared to the hand-picked configuration, the EFI algorithm displaces *one* sensor from its location at the side centre, the TA algorithm *two*.

The rows of Table 6.3 show the estimation for the different sensor configurations. Although the two new sensor configurations lead to a higher methodical error of the RB estimator, the methodical error of the SF estimator decreases. Especially in terms of the maximum error this decrease is significant; the decrease is are about 20 % and 40 % for the EFI and TA algorithm, respectively. On the other hand, the estimation error of shape fitting due to noise increases, for the TA algorithm by more than a factor of 2.

## 6.5 Conclusions

This chapter studied the implementation of automatic sensor placement in the framework of shape fitting. The algorithms in literature have various optimisation goals. Four algorithms, EFI, TA, RA12 and FS, were selected for further study. These algorithms place sensors based on a set of shapes, aiming at finding the amplitudes of the shapes with least error. This goal is related with, but different from our estimation goal, i.e. finding shape estimates with

Table 6.3: A comparison between the estimation errors of the RB and the SF estimator when using different sensor configurations. The error is shown as function of the location of the lens column. The results are obtained using the standard controller, 8 sensors for estimation and a disturbance bandwidth of 0–100 Hz. The first row is for clarity duplicated from Table 5.2.

| Scenario | RB estimation | | SF estimation | |
|---|---|---|---|---|
| | Methodical estimation error ($E_{\text{meth,RB}}$) | Estimation error due to noise ($E_{\text{noise,RB}}$) | Methodical estimation error ($E_{\text{meth,SF}}$) | Estimation error due to noise ($E_{\text{noise,SF}}$) |
| Initial sensor configuration |   |   |   |   |
| Sensor configuration EFI |   |   |   |   |
| Sensor configuration TA |   |   |   |   |
| |   |   |   |   |

least error. All four algorithms start with a set of sensor candidates, of which the less important sensors are removed one-by-one. As such, the algorithms are fast but none of them can guarantee global optimality.

The placement algorithms were assessed in different ways. They were, amongst others, applied to sets of randomly distributed shapes. The EFI and TA algorithms performed systematically better than the other selected algorithms, also the algorithms tested in similar experiments in literature before.

Furthermore, the algorithms were applied to a set of eight modeshapes of a plate, which had to be distinguished either using position sensors or strain sensors. In both cases the TA algorithm obtained superior results over the other algorithms. In general, sensor placement with a smaller candidate sets has an equal or even higher chance to yield a well-performing sensor configuration than placement with a larger candidate set. Convergence to a well-performing sensor configuration was observed to be not guaranteed, especially with RA12. Based on these observations, a placement strategy is proposed in which several small candidate sets are tested, instead of a single larger one.

For two small sensor candidate sets, a brute force evaluation of all sensor possible sensor configurations was carried out to find the optimal sensor configuration. The TA algorithm closely approached the performance of the optimal sensor configuration, to within 11 % and 3 % in terms of the SNA.

Finally, a procedure was developed for incorporating the placement algorithms into the SF framework. To be able to use a snapshot matrix as input for the sensor placement algorithms, the number of shapes need to be reduced and the RB contributions need to be removed. The procedure starts with calculating the flexible residues corresponding to the snapshot matrix and extracts from them the most prominent patterns at the target points using the POD. Then the most likely patters at measurement points are calculated, which are the shapes to perform sensor placement with. After sensor placement, the SF estimator is calculated using the LS technique. The procedure was in combination with the EFI and TA sensor placement algorithms applied in the numerical case of the wafer chuck. Although the placement procedure does not take into account a travelling disturbance source, the procedure significantly reduced the SF estimation error.

# Chapter 7

# Experimental verification of shape fitting

In the previous chapters, the Shape Fitting (SF) method was introduced and further developed for use in the wafer chuck application. The method was assessed using numerical models and simulations. It is important to experimentally verify the numerical models, as they may not include all relevant physical effects and boundary conditions and are also prone to programming errors. This chapter describes the design of an experimental setup and the results obtained with it.

Section 7.1 first defines the purpose and the requirements of the setup. Then, based on these requirements, the design choices regarding the setup and its components are presented. Section 7.2 describes the results of the validation experiments that were performed to prove that the setup operates as desired. The setup is then utilised in Section 7.3 to verify the numeric model of the wafer chuck and its controller. Section 7.4 assesses the SF method experimentally. Section 7.5 shows experimental results of a hybrid estimator that combines shape fitting with absolute position measurement based on an accelerometer signal.

## 7.1 Setup design

### 7.1.1 Purpose of the setup

An experimental setup may be used for many different reasons. Three major purposes of a setup are 1) the exploration of new physical phenomena; 2) the verification of model-based expectations and 3) the demonstration in practice of a method or principle.

The physical principles upon which this thesis builds are considered well-understood and need no further exploration. In fact, the setup is designed such that phenomena that are not well-understood are kept non-significant. Verification of the models is important because a model may have inherent shortcomings in describing all relevant physical effects and boundary conditions. Significant discrepancies might be discovered using experiments with a physical setup.

Furthermore, even if a model is able to describe the physical phenomena, the model or a derived simulation could contain programming errors, which might be revealed by experimental verification. Therefore, the main purposes of this setup are the verification of the models and the demonstration of shape fitting on a physical setup.

### 7.1.2 Requirements

The real wafer chuck system is highly complex, but for the setup it is not necessary to mimic all its aspects. The setup should only contain the physical principles that are relevant to the research in this thesis. Therefore, the setup need to have a similar topological architecture, but its parts can be highly simplified as compared to the actual wafer chuck system.

The wafer chuck is mimicked by a simple plate. Only the out-of-plane motion of the setup is relevant, so that the plate should be constrained in the in-plane directions and should be free-floating in the out-of-plane direction. The plate is kept in position in out-of-plane direction with position sensors, force actuators and a Rigid Body (RB) control system as described in Section 2.1.3. The setup should be sufficiently linear.

The purpose of the setup is to verify the models and the SF method, for which it is important that both RB and flexible behaviour can be measured. This means that the noise floor of the measurements should be low.

To make the setup usable for different types of experiments the sensors and the actuators should be easily removed and placed at other locations. The setup should therefore be modifiable, but in such a way that the measurements remain repeatable.

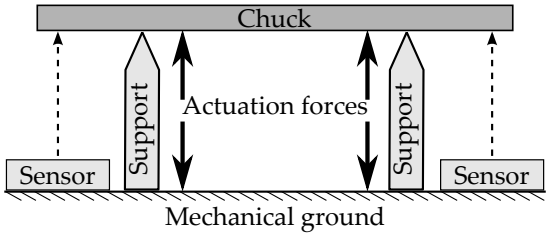


### 7.1.3 Suspension concepts

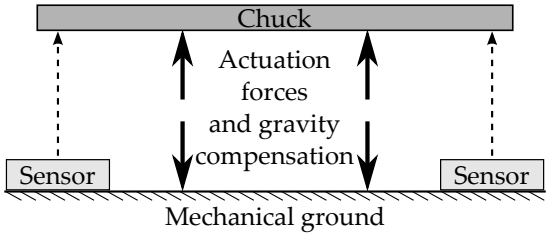
During design of the setup, the decisions on conceptual and detailed level were highly interwoven. This section presents the fundamental conceptual choices regarding the suspension of the chuck and the sensors. The next section elaborates on the more detailed design decisions.

A reason for choosing a free-floating plate is that the system is then topologically similar to the wafer chuck system. A free-floating plate setup is, however, much more complicated than one with a directly supported plate. Yet, the free-floating plate concept also brings the advantage of a reduced noise floor.

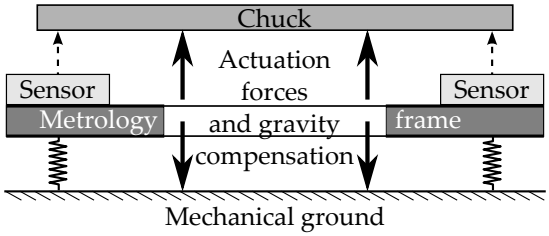
In order to find how large the reduction of the noise floor is, a noise analysis was performed on three different suspension concepts, as depicted in Figure 7.1. In the first concept the plate is directly supported by the ground and the sensors and actuators are mounted to the ground. In the second concept the chuck is free-floating and levitated by a position control system. The sensors are mounted on the ground. In the third concept, the sensors are mounted on a metrology frame, a vibration isolation system with a low-stiffness connection to the ground. To levitate the chuck in the second and third concept, the actuators apply additional forces for gravity compensation.




(a) *Non-levitated*. The chuck is directly supported by the ground; the sensors are directly mounted to the ground.




(b) *Levitated*. The chuck is free-floating using RB control; the sensors are directly mounted to the ground.




(c) *Levitated/metrology frame*. The chuck is free-floating using RB control; the sensors are connected with low stiffness to ground.

Figure 7.1: Three possible suspension concepts for the setup. The third concept, the levitated chuck with the metrology frame, leads to the lowest noise and was chosen for the setup.

Table 7.1: Comparison between the noise levels of the three suspension concepts. The values per noise source are translated into their deformation equivalents and are given as the peak-to-peak ($8\sigma$) values up to 1 kHz.

| Noise source | Suspension concept | | |
|---|---|---|---|
| | Non-levitated [nm] | Levitated [nm] | Levitated/metrology frame [nm] |
| Floor vibrations | 27 | 20 | $< 0.20$ |
| Reaction forces | 19 | 5.0 | $< 0.050$ |
| Current amplifiers | 5.6 | 5.7 | 5.7 |
| Sensor noise | 27 | 27 | 27 |
| *Total* | *43* | *34* | *28* |

Four noise sources were identified, namely floor vibrations, reaction forces from the actuators, noise of the current amplifiers that feed the actuators and sensor noise. Their typical signal levels were measured or estimated and translated into deformation equivalent noise, i.e. the apparent deformation of the plate due to the noise source. Note that the reaction forces are correlated with the actuation signals and can not be lowered by averaging multiple experiments, making this from a different type of noise source than the others.

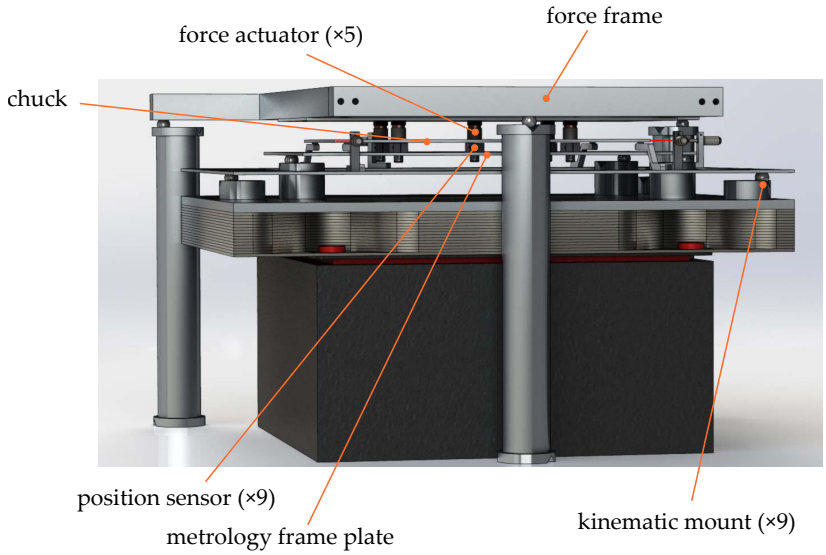
The floor vibrations, known to be significant in the laboratory in which the setup was to be build, were measured using geophones (Appendix E). The noise of the current amplifiers was measured using a current probe, the supplied current had almost no influence on the noise level. The sensor noise was measured stably fixated reference. The amplitude of the reaction forces was estimated to be 1 N.

Table 7.1 lists the the deformation equivalent noise levels for the three suspension concepts. Note the large reduction of the influence of the floor vibrations and reaction forces for the levitated concept with a metrology frame. In this concept the relative motion between the plate and metrology frame is highly reduced due to the low-stiffness connection between both the plate and the ground and the metrology frame and the ground.

Clearly, the levitated concept with metrology frame has the lowest total noise. For this reason and the fact that this concept has a similar topology as the actual system, it is selected for further development.

### 7.1.4 Detailed design and component selection

Figure 7.2 shows photo-renders of the setup that was designed and build. The design of its most important components, which are indicated in the figure, are further described in this section.



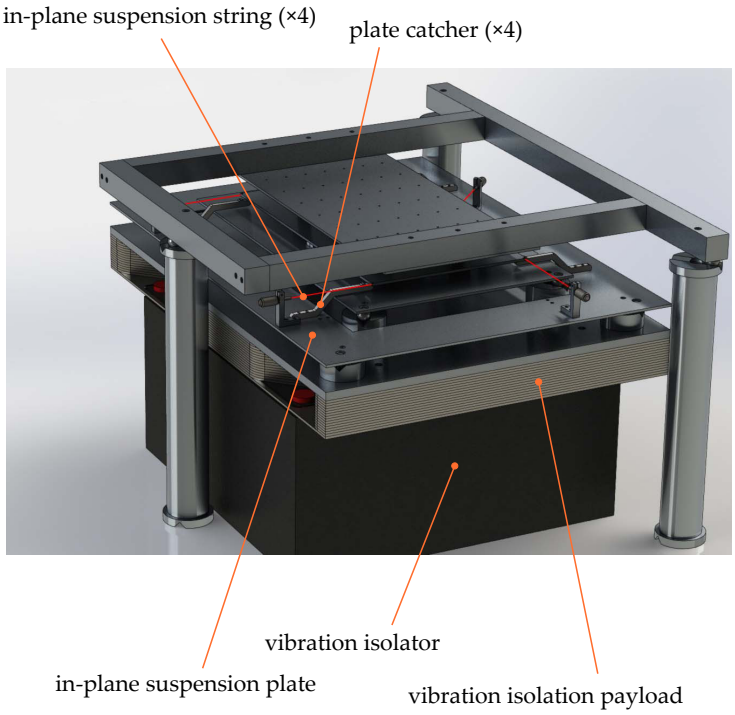




Figure 7.2: Photo renders of the setup from different viewing points. Indicated are the most important components, whose design and selection is described in Section 7.1.4.

### Metrology frame

The metrology frame consists of a vibration isolation payload and a metrology plate that holds the sensors The metrology frame is connected via a vibration isolator to the ground. The vibration isolator provides a low-eigenfrequency suspension which lowpass filters the ground vibrations and the vibrations caused by the reaction forces. In a small market survey the commercial Minus-K BM-4 isolator was identified as a well-performing and cost-effective solution (Appendix G.1). This fully passive system contains a spring with a high positive and a second with a negative stiffness that in parallel result in a low positive stiffness in vertical direction. The system allows for eigenfrequencies as low as 0.5 Hz. To not only reduce the floor noise but also the disturbances that directly work on the metrology frame and decrease the influence of sensor cable stiffness, the isolator variant with highest allowed payload was selected for best performance. Together with a stack of steel plates, a total payload of about 90 kg was installed.

### In-plane suspension

The out-of-plane motion need to be kept free, whereas the in-plane direction must be stiffly suspended. Different concepts were studied, of which a concept with strings was selected. The three in-plane DoFs are suspended by four strings with a pretension of 0.1 N. This results in a relatively low out-of-plane stiffness between about $\pm 2.5$ N/m. The in-plane suspension strings are held by an in-plane-suspension plate, which also comprises four plate catchers, which support the plate during disassembly.



The in-plane suspension could either be mounted to the ground or to the metrology frame. The latter option was chosen, as this reduces the influence of in-plane ground vibration. The influence from the plate's vibrations on the metrology frame, on the other hand, is expected to be low, due to the high mass of the metrology frame.

### Plate design

To apply the SF method the chuck's RB motion and its deformation should both be measurable with the same sensors. The deformation is typically small compared to the RB motion. This puts demands on the sensors' specifications, as sensors with a large range in combination with small resolution steps are expensive. Therefore, the ratio between the two needs to be minimised.

To achieve large deformations, the plate's stiffness should be low. But from a RB point of view, the plate should be stiff, because this increases its eigenfrequencies. Higher eigenfrequencies allow for a higher controller stiffness and, thus, to smaller RB motions. In addition, a higher stiffness controller with a higher bandwidth is desired, because it increases the frequency band that is useful for applying shape fitting. After all, performing low-frequency experiments is difficult, due to thermal drift and long required measurement time.

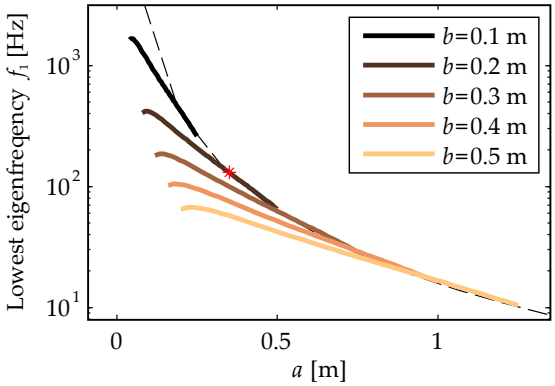



Figure 7.3: The lowest eigenfrequency, $f_1$, of a plate as function of the plate's dimensions $a$ and $b$, according to Eqs. (7.2) and (7.3). The plate has material parameters of aluminium and a thickness of 3 mm. The dashed line indicates the first eigenfrequency of a beam (i.e. $b \ll a$) according to [95]. The red asterisk indicates a size of 350 mm $\times$ 200 mm, the plate size that was chosen for the setup.

When selecting the chuck's size, clearly a trade-off has to be made between the height of the lowest eigenfrequency and the stiffness of the plate. This was done in a parametric analysis with simple analytic formulas.

The controller stiffness $K_\mathrm{p}$ that can be attained is proportional to the plate's lowest eigenfrequency, $f_1$, squared:

$$K_\mathrm{p} \propto f_1^2. \tag{7.1}$$

According to Leissa [96], the lowest eigenfrequency of a plate can be expressed as function of the plate's geometry and material properties as

$$f_1 = \lambda_1 \frac{1}{2\pi\sqrt{12}} \sqrt{\frac{E}{\rho(1-\nu^2)}} \frac{h}{a^2}, \tag{7.2}$$

where $a$ denotes the length, $b$ the width and $h$ the thickness of the plate. $\lambda_1$ is a function of the ratio $a/b$. Leissa supplies $\lambda_1$ for only four values of $a/b$. To obtain a continuous function, the values were fitted with a second order polynomial, taking into account the fact that $b^2\lambda_1(a/b) = a^2\lambda_1(b/a)$, yielding

$$\lambda_1 = -6.06\left(\frac{a}{b}\right)^2 + 26.2\left(\frac{a}{b}\right) - 6.06. \tag{7.3}$$

with $0.4 \leq a/b \leq 2.5$. Figure 7.3 shows eigenfrequency $f_1$ as function of the plate's dimensions $a$ and $b$.

To describe the stiffness of the plate, the analytic formula of a beam is used. A beam with thickness $h$, length $a$ and width $b$ that is supported with hinges at its ends has a stiffness

$$k_\mathrm{beam} = 4\frac{Eh^3 b}{a^3}. \tag{7.4}$$

If the plate's deformation is modelled as two beams in series, one bending in

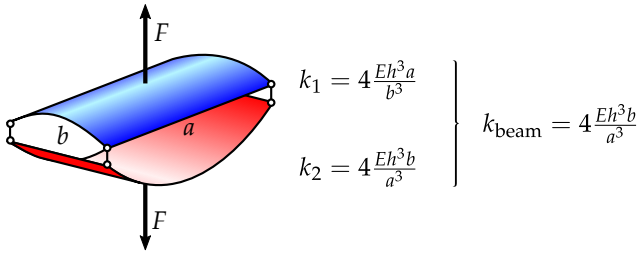



Figure 7.4: The bending stiffness of a plate, $k_{\text{plate}}$, is approximated by the stiffness of two bending beams in series, one bending in the direction of its width the other in the direction of its length.

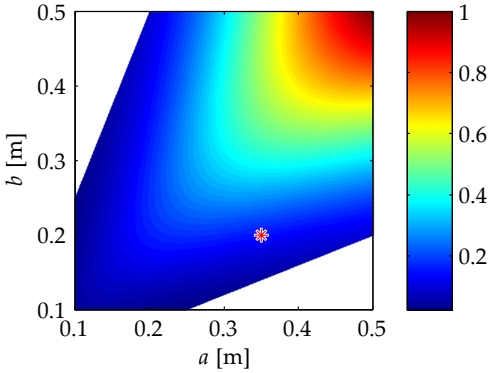



Figure 7.5: The normalised ratio between the actuation stiffness and the bending stiffness, $r_k$, of a plate as function of its length, $a$, and width, $b$. The red asterisk indicates a size of 350 mm $\times$ 200 mm, the plate size that was chosen for the setup.



the direction of the width $b$, the other in the direction of length $a$, as illustrated in Figure 7.4, the total stiffness becomes

$$k_{\text{plate}} \approx \frac{4Eh^3ab}{a^4 + b^4}. \tag{7.5}$$

Combining Eqs. (7.1–7.5), the ratio between the bending stiffness and the controller stiffness, can be expressed as

$$r_k(a,b) \approx \frac{k_{\text{plate}}}{K_{\text{p}}} \propto \frac{\rho hab}{\left(-\frac{6.06}{b^2} + \frac{26.2}{ab} - \frac{6.06}{a^2}\right)^2 (a^4 + b^4)},$$

which needs to be minimised. Figure 7.5 shows a normalised $r_k$ as function of the plate's dimensions. Smaller values of $r_k$ are attained for smaller plate dimensions and higher aspect ratios.

A plate geometry of 350 mm $\times$ 200 mm $\times$ 3 mm was chosen (indicated with the red asterisk in Figure 7.5), also based on considerations like the availability of enough space for the sensors and actuators. A rectangular plate geometry was chosen instead of a square geometry to remove symmetry from the system,

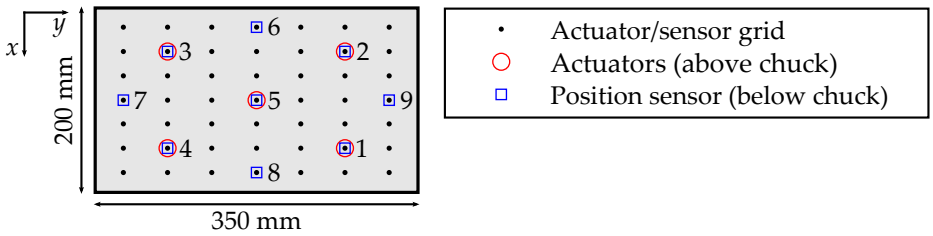



Figure 7.6: Initial configuration of the setup's actuators and sensors. The numbers indicate the actuator and sensor numbering convention.

thereby avoiding multiplicity of eigenmodes. As the plate needs to be electrically conductive for measurement reasons aluminium was selected as material.

### Control system

A RB coordinate controller (Section 5.2.1) was implemented for levitation of the chuck. This controller is run on a real-time xPC system with analogue/digital and digital/analogue converters.

### Sensor selection

Different types of positions sensors were compared, mainly based on the ratio between their range and resolution and their price. The required ratio was determined using the FE model in combination with the designed controller. The required sensor range was found from the maximum translation amplitude of the plate at a frequency below the bandwidth of the system and the required resolution from the quasi-static deformation caused by a force at the centre of the plate and the corresponding actuator forces. To be able to measure the amplitude with high enough resolution, the required ratio, $r_s$, was multiplied by a factor of 100, leading to

$$r_s = \frac{0.01\ \mathrm{m/N}}{5 \cdot 10^{-6}\ \mathrm{m/N}} \cdot 100 = 2 \cdot 10^5.$$

Based on a market survey, Micro-Epsilon CapaNCDT6200 capacitive sensors with CS1 heads and DL6230 signal conditioners were selected. These allowed for measuring over a range of 1 mm with a resolution determined by the sensor's noise level of 12 nm ($1\sigma$, bandwidth 1 kHz).

In total 9 sensors were placed on the locations indicated in Figure 7.6. Sensors 1 to 4 are used for the RB control.

### Actuator selection and gravity compensation

Voice coil actuators were selected as actuators because their motor constant (generated force per applied current) is relatively constant over their stroke. To minimise the mass that is added to the chuck the compact Moticont lvcm-016-013-01 actuators were selected, whose magnet housings have a mass of 12 g. Based upon the information on their datasheet, their stiffness at nominal current

is at maximum $\pm 300$ N/m at the extremes of their range (0 to 6 mm). In practice, a range of 1 mm is used, so that, if operated around their zero stiffness point, the maximum stiffness at nominal current is lower, e.g. $\pm 25$ N/m. The actuators will provide the forces needed for gravity compensation. This requires a current of 3.5 times the nominal current of a single actuator. The actuator stiffness should be multiplied by this value to obtain the maximum stiffness in $z$-direction.

To avoid wires connected to the chuck and heat loads, the coils were connected to the force frame and the magnet housings to the chuck. The housings were connected to the chuck with simple disk magnets that were glued to the chuck. This construction also allows for in-plane alignment of the magnet housings with respect to their coils. In total five actuators were placed, four symmetrically and one in the centre (Figure 7.6).

The sensors and actuators were mounted at opposite sides of the plate to allow for collocation. The actuators were mounted above the plate to decrease heat introduction to the plate, which causes thermal deformation. Moreover, the coils are conductively cooled by there direct connection to the force frame.

#### Kinematic mounts

To facilitate simple modification and repeatable replacement of the setup's parts, kinematic mounts were designed for the connections between the in-plane suspension, the metrology frame, the force frame and their respective neighbours (Figure 7.2).



## 7.2 Setup validation

### 7.2.1 Suspension stiffness

The stiffness of the plate's suspension should be low compared to the controller stiffness. Based upon the setup's design it is expected that the suspension stiffness is dominated by the actuator stiffness, which is circa $\pm 100$ N/m. The stiffness caused by the in-plane suspension is expected to be negligible.

An experiment was performed to determine the stiffness in $z$-direction. First, the distance between the force frame and the metrology frame with the plate in a fixed position was manually adapted by changing the height of the vibration isolator's platform in order to approach closely the zero-stiffness point of the actuators. Then the $z$-setpoint of the controller was modified in a number of steps from 0.050 mm to 0.950 mm. For each step the controller forces were recorded. From these measurements, the suspension stiffness in $z$-direction was obtained. A maximum stiffness of about $\pm 100$ N/m at the extremes of the range was found, which can be explained well by the actuator stiffness. As the chosen controller stiffness is typically $10^4$ N/m, the additional stiffness can be considered low.

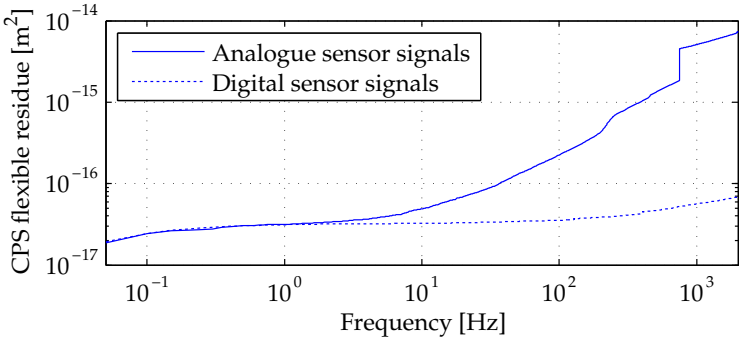



Figure 7.7: Cumulative power spectrum of the flexible residue at the centre sensor ($\overline{w}_{5,\text{fl}}$). The flexible residue was calculated based on both the analogue and digital sensor signals.

### 7.2.2 Noise performance

**Vibration isolation** A geophone was used to examine the vertical transmissibility of the Minus-K vibration isolation system. Although the eigenfrequency of the vertical motion was not tuned to be as low as possible, the measured transmissibility at the centre of the Minus-K (Appendix G.2) resembles the curve from the datasheet well. It was observed, however, that the Minus-K has a relatively high stiffness in the out-of-plane rotation direction, which is caused by the rubber tilting pad that makes the out-of-plane rotations possible. This stiffness leads to an almost full transmission of vibrations around 8 Hz at the corners of the isolator.

**Deformation equivalent noise** The noise performance of the system as a result of all unwanted disturbances was validated. As the setup is developed for evaluating the estimation of deformation, it is important to study the deformation that is equivalent to the noise. The deformation is expressed in terms of the flexible residue at the centre sensor, $\overline{w}_{5,\text{fl}}$. The noise floor has to be low, typically 50 nm, which is a factor of 100 below the expected desired deformation amplitudes.

To find the deformation noise level, the signals of the position sensors were recorded for 20 s while the plate was in closed-loop control and no disturbances were applied via the actuators. From the recorded signals the flexible residue at the centre sensor ($\overline{w}_{5,\text{fl}}$) was calculated. The flexible residue was obtained from both the analogue signals and the digital signals (EtherCAT) of the sensor signal conditioners. The analogue signals were digitised using a NI6220 data acquisition card.

Figure 7.7 shows the Cumulative Power Spectrum (CPS) of the flexible residue. The noise level of the residue obtained with the analogue signals starts increasing faster than the one obtained with the digital signals. According to the sensor manufacturer, the digital sensor signals are lowpass filtered with a cutoff frequency of 400 Hz, but this filter can not explain the discrepancy between the two spectra. The difference between the spectra below 200 Hz might be

explained by a white noise with a level of $1.3 \cdot 10^{-9}$ m/$\sqrt{\text{Hz}}$. Indeed the data acquisition card produces random noise, but $1.3 \cdot 10^{-9}$ m/$\sqrt{\text{Hz}}$ is still a factor of 2.1 higher than the expected contribution of the noise level of the data acquisition card to the flexible residue.

The CPS of the analogue signals shows a jump around 750 Hz. As the jump has a width of only one frequency sample it does most probably not stem from the mechanical but from the electronic domain. The jump was also present when the current amplifiers of the actuators were turned off and when a single sensor was read out with an external data acquisition box (NI6211). As the jump's frequency changes when the sampling frequency is changed, the noise component most probably aliases in from a higher frequency. Furthermore, it was observed that the height of the jump is proportional to the distance between the sensor and the plate. Thus, it is hypnotised that the peak is caused by the sensor signal conditioners.

The use of the digital sensor signals leads to a noise level of 7.4 nm ($1\sigma$, 1 kHz bandwidth), which is about a factor of 2 higher than the predicted noise level of Table 7.1. The analogue readout leads to a noise level of 71 nm, which is significantly higher than the prediction. This is still a factor of 70 lower than the typical deformation amplitudes and is, therefore, considered acceptable.

**Delay noise** The digital sensor outputs have an internal sampling frequency that is not synchronised with the sampling of the xPC system. This causes delay and at regular time intervals repetition of the position samples, in that way adding noise proportional to the slope of the signals. As the effect is not synchronised amongst all 9 sensor channels, the use of the digital signals leads to considerable error in the estimation experiments. For this reason the analogue signals are used in the rest of this chapter.



### 7.2.3 Parasitic effects of the actuator forces and suspension

For performing valid deformation measurements not only the setup's noise level should be low, but also its parasitic deformation. The reaction forces of the actuators cause vibration of the ground, which is partly transmitted via the vibration isolator to the metrology frame. Furthermore, the actuator stiffness couples the motion of the force frame to the chuck and the string suspensions introduce stiffness between the chuck and the metrology plate. All these effects lead to a parasitic relative motion between the chuck and the metrology frame.

The control system interprets the parasitic relative motion as a displacement of the wafer chuck and responds by exerting forces to the wafer chuck. Those forces lead to parasitic deformation of the wafer chuck.

The parasitic deformations are, as opposed to the deformations due to the random floor vibrations, correlated with the desired response to the actuator signals. Thus, the parasitic deformation is easily misinterpreted as part of the intended response of the wafer chuck. Moreover, the parasitic deformation can not be reduced by averaging techniques as it is not random.

A validation experiment was carried out to find the level of the parasitic deformation and to compare it to the level of the desired deformation. The

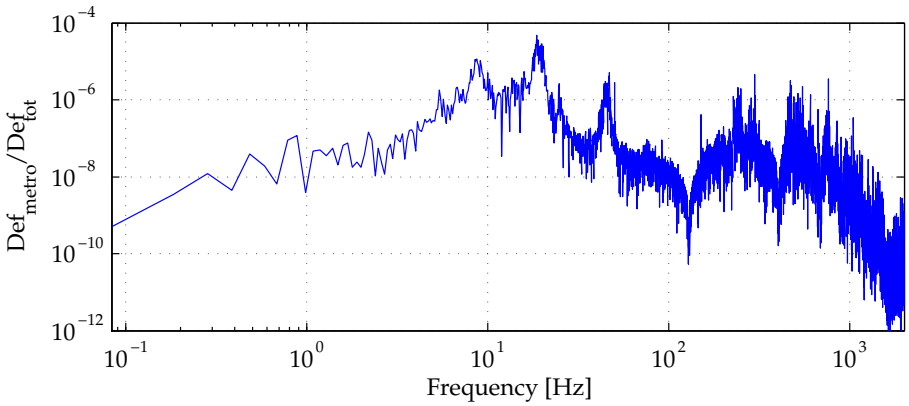



Figure 7.8: The ratio between the parasitic deformation and the total deformation of the wafer chuck due to the centre actuator. The parasitic deformation is insignificant as it is a factor of $10^{-4}$ or smaller for all frequencies.

transfer function of the centre actuator to the position of the metrology frame was obtained using a geophone. The closed loop transfer function from an output disturbance on the sensors to the deformation of the plate was obtained using the system's model. The two transfer functions were multiplied to find the full transfer function of the parasitic deformation to the actuator force. This transfer function was compared to the transfer function of the deformation due to the actuation forces.

Figure 7.8 compares the transfer functions by showing the ratio between the parasitic deformation and the total deformation. This ratio stays below $10^{-4}$ for all frequencies. Thus, for a deformation amplitude of $5 \cdot 10^{-6}$ m the expected deformation error is well below $5 \cdot 10^{-10}$ m, which is lower than the noise level of the experimental setup.

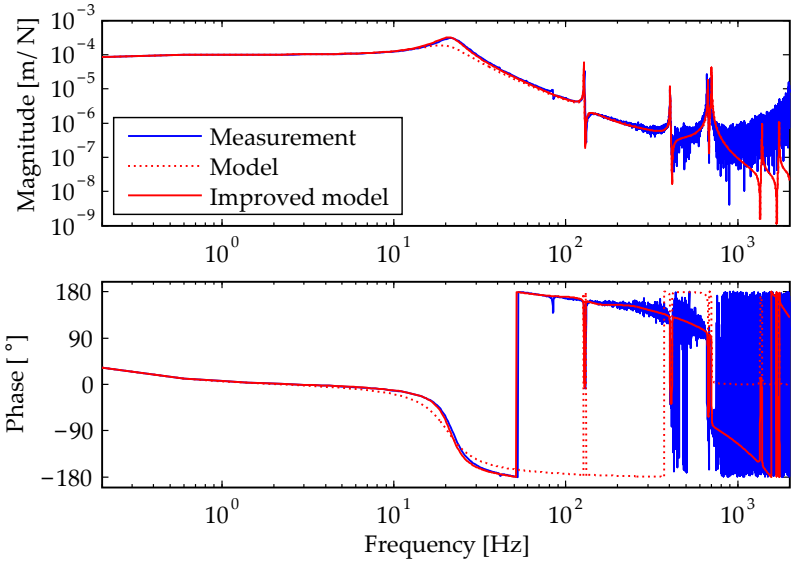


## 7.3 Verification of the system model

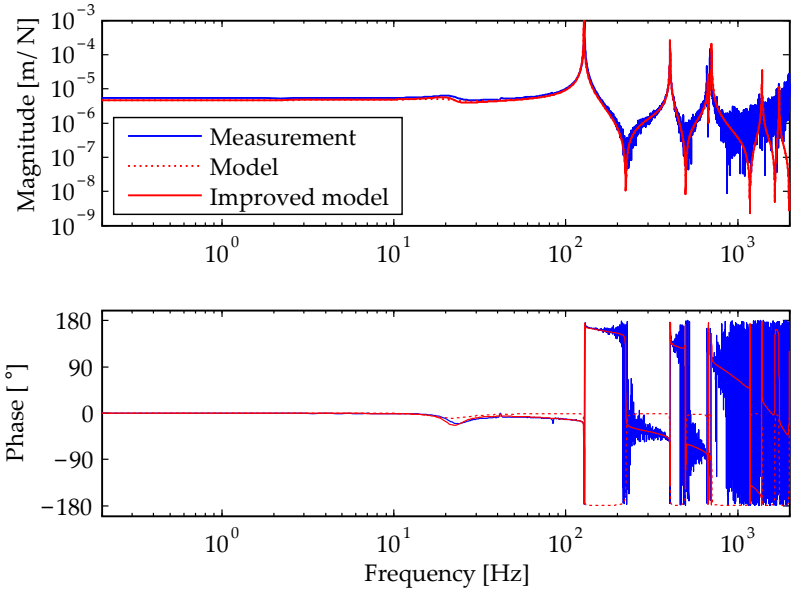
To verify the model of the total system, i.e. the finite element model and the controller in combination with the decoupling matrices (Section 5.2.1), the most important transfer functions were compared to experimental results. To obtain the transfer functions of the setup, a block signal with a period of 5 seconds was applied to the centre actuator. The sensor response was recorded for 120 seconds. From this data, the transfer functions were calculated.

Figure 7.9a shows a Bode plot of the closed-loop transfer from a force at the disturbance actuator to the RB estimate of the centre PoI. Both the centre disturbance actuator and PoI sensor were placed at location 5, the centre of the plate. Figure 7.9b shows a Bode plot of the flexible residue for the same actuator and sensor pair.

Two parameters were adapted to obtain better agreement between the model and setup. Firstly, the motor constants were multiplied by about 1.6, leading




(a) Transfer function from the centre actuator to the RB estimate at the centre sensor.




(b) Transfer function from the centre actuator to the flexible residue at the centre sensor.

Figure 7.9: Bode plots of the transfer functions of the centre actuator to the RB estimate and the flexible residue at the centre position sensor. The improved model contains a lowpass filter and a delay and closely resembles the measurement.

to equal RB estimate at low frequency, but a 13 % larger deformation than modelled. Secondly, the modal damping constant of the model was fitted to the measurements. Figures 7.9a and 7.9b show the transfer functions obtained from the model in dotted red lines.

For low-frequencies, the most important differences between the experimental results and the model are found around the controller bandwidth, which is at 20 Hz for the translation and at 10 Hz for the rotations. These differences are primarily caused by a Butterworth lowpass filter with a cutoff frequency of 200 Hz. This filter was used to reduce the noise component at 750 Hz (Section 7.2.2), which led to instability of the motion controller. At higher frequency, delay causes the setup to show more phase lag than the model. Figures 7.9a and 7.9b show the transfer functions of the improved model with the added lowpass filter and a delay in solid red lines. The transfer functions of the improved model and the setup show close resemblance. It can, therefore, be concluded that the model of the FE model in combination with the control system describes the RB controlled plate system well.

## 7.4 Demonstration of the shape fitting method

To verify the SF method and to demonstrate its practical applicability it was implemented on the experimental setup. The setup has five actuators that can be used to apply disturbance forces and eight sensors for estimation. Thus, the number of disturbance locations is relatively limited. For this reason it was chosen to not use more disturbance locations than the number of estimation sensors to obtain more snapshots than estimation sensors, but instead to use grids of frequencies to obtain the snapshots. The results of different frequency bands are compared. This section starts by assessing shape fitting in the time domain and then studies the shape fitting performance in the frequency domain.



### 7.4.1 Comparison estimators in the time domain

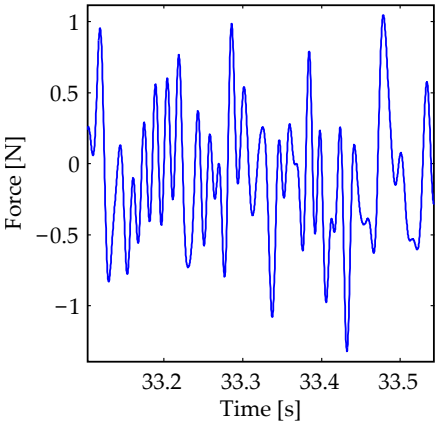
#### Method

In the experiments a single actuator was used as a disturbance source, namely the centre actuator (position 5, Figure 7.6). The centre position sensor (position 5) was selected as the target point (the PoI). The displacement of the target point was estimated both with the RB estimator and the SF estimator. All sensors, except the PoI sensor, – in total 8 sensors – were used for estimation.

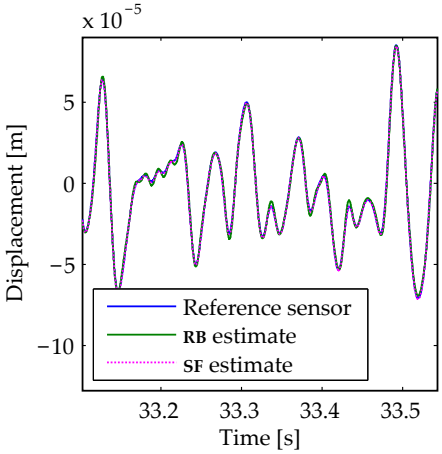
Estimation was performed for noise signals with two frequency bands, one from 0.1 to 80 Hz and the other from 0.1 to 150 Hz. The band 0.1 – 150 Hz includes the first resonance frequency of the plate.

The experimental procedure consisted of two steps, first a calibration step, then an estimation step.

**Calibration step** A square wave signal with a period of 5 seconds was applied to the disturbance actuator. The sensor response was recorded for 120 seconds. The data of all sensors, including the PoI sensor, was used to generate a complex




(a) Detail (0.44 s) of the disturbance force of the centre actuator.




(b) Detail (0.44 s) of the displacement measurement of the reference sensor and its estimates.

Figure 7.10: Results of the time domain experiment with a disturbance signal in the frequency band 0.1 – 80 Hz.

frequency-domain snapshot matrix (cf. Section 5.3), according to the frequency band for which the SF estimator was optimised. For each band the corresponding sensor influence matrix $\boldsymbol{B}$ was calculated according to Eq. (5.37).

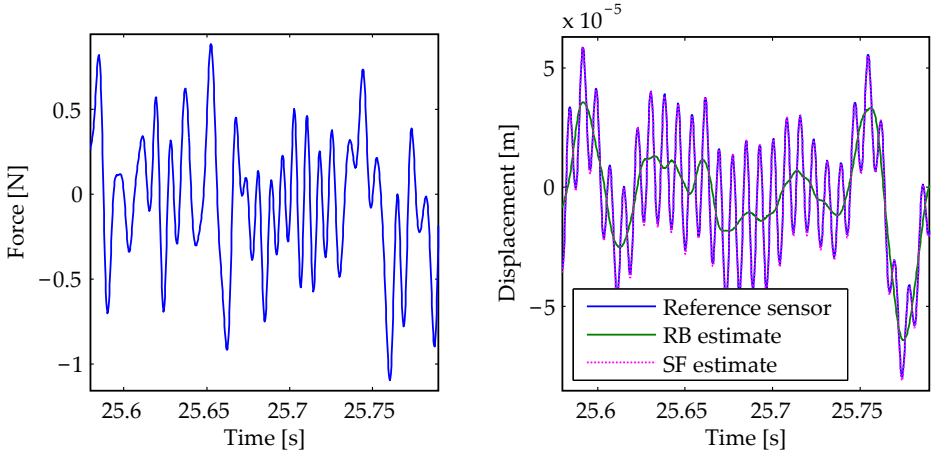
**Estimation step** Two random signals with frequency content corresponding to the band for which the estimator was optimised were constructed. The signals consisted of a high number of frequency components linearly distributed through the frequency band, each component having a Gaussian distributed amplitude and an uniformly distributed phase. Figures 7.10a and 7.11a show details of the disturbance signals. The signals were applied with the disturbance actuator and the responses of the position sensors were recorded.

The signals were highpass filtered using a cutoff frequency of 0.1 Hz to remove the sensor offset and the drift of the plate's shape due to thermal effects. Then shape fitting was used performed with the signals to estimate the PoI's position as function of time.

**Results**

Figures 7.10b and 7.11b show details of the displacement measurement of the reference sensor and the RB and SF estimates for the frequency bands from 0.1 to 80 Hz and 0.1 to 150 Hz, respectively. Especially at the disturbance bandwidth 0.1 – 150 Hz the RB estimator performs

Table 7.2 lists the RMS values of the reference sensor and the estimation errors. The level of the displacement measured at the reference sensor is equal in the two experiments. The error of the RB estimation equals – by definition – the flexible residue. As expected, the SF method performs better than the estimation




(a) Detail (0.21 s) of the disturbance force of the centre actuator.

(b) Detail (0.21 s) of the displacement measurement of the reference sensor and its estimates.

Figure 7.11: Results of the time domain experiment with a disturbance signal in the frequency band 0.1 – 150 Hz.

error of the RB method. The level of RB estimation error is for the disturbance band 0.1 – 150 Hz more than ten times larger than for the band 0.1 – 80 Hz, most likely due to the modeshape corresponding to the plate's lowest resonance frequency. The level of the SF estimation error increases only by about a factor of 1.4, indicating that the SF estimator can deal well with the modeshape.



### 7.4.2 Comparison estimators in the frequency domain

The previous section studied the estimation performance in the time domain. The results showed that the estimator performance depends on the frequency band of the disturbances. This section studies the performance of the estimators further in the frequency domain.

Table 7.2: A comparison of shape fitting in the time domain in terms of the RMS values of the estimation error.

| | Disturbance bandwidth | |
|---|---|---|
| | 0.1 – 80 Hz | 0.1 – 150 Hz |
| | RMS value [m] | RMS value [m] |
| Reference position sensor | $28 \cdot 10^{-6}$ | $28 \cdot 10^{-6}$ |
| Flexible residue | $1.3 \cdot 10^{-6}$ | $19 \cdot 10^{-6}$ |
| Error RB estimation | $1.3 \cdot 10^{-6}$ | $19 \cdot 10^{-6}$ |
| Error SF estimation | $0.36 \cdot 10^{-6}$ | $0.49 \cdot 10^{-6}$ |

### Method

In the previous section, the estimator performance was studied for two frequency bands. In the experiments of this section four frequency bands were selected, all starting at 0.1 Hz and extending to, respectively, 1, 10, 80 and 150 Hz. In the bands 0.1 – 1 Hz and 0.1 – 10 Hz the deformation of the system is expected to be mostly quasi-static. In the band 0.1 – 80 Hz also dynamic effects due to the system's bandwidth are included. In the band 1 – 150 Hz is next to that also the first resonance frequency of the plate included.

As in the previous section, the experimental procedure consisted of a calibration step followed by an estimation step. The calibration was performed in the same manner and with the same measurement data as in Section 7.4.1. For each of the four frequency bands a sensor influence matrix $\boldsymbol{B}$ was optimised.

In the estimation step, the same measurement data was used as in the calibration step.

### Results and discussion

Figures 7.12 up to 7.15 show the behaviour of the RB and SF estimators for the different frequency bands and compare them to the actual displacement. The figures have linear instead of logarithmic frequency axes, which allows easier interpretation of the results if the disturbances can be assumed uniformly distributed over the frequency band.

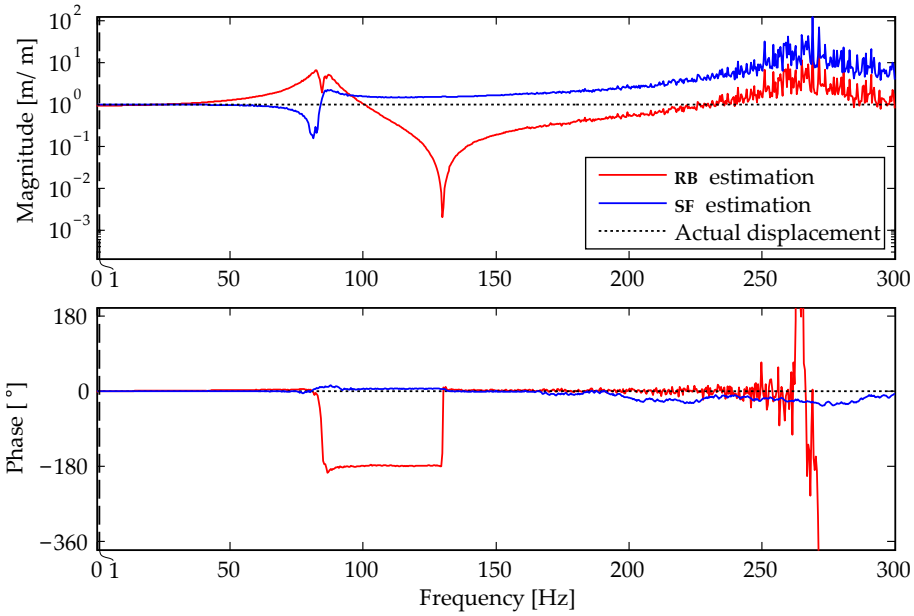
The upper figures show Bode plots of the transfer functions from the actual to the estimate displacement. The transfer function from actual displacement to the RB estimate is equal for all figures, as the RB estimator does not use foreknowledge on the typical disturbances – in this case the frequency content of the disturbances. An ideal estimator has a magnitude equal to 1 and a phase of 0° for all frequencies. All transfer functions start relatively good with magnitudes close to 1 and phases close to 0°. The transfer function of the SF estimator seems to have in general both a better magnitude and phase behaviour.

The lower figures show the transfer from actuation force to estimation error. For most frequencies in the frequency band the SF estimator performs significantly better then the RB estimator. This is often even the case outside the frequency band that was optimised for.

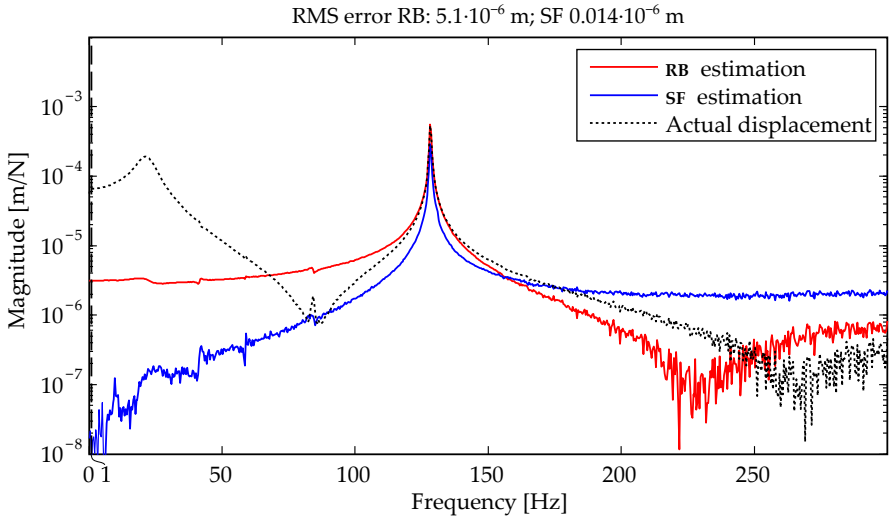
The RMS values of the estimation errors are indicated above the lower figures. The values correspond to disturbance signals that are white within the respective frequency band and have a RMS value of 1 N. Note that the frequency components of the error contributions need to be squared when calculating the total estimation error caused by the noise frequency band.

## 7.5 Hybrid estimation using shape fitting and an accelerometer

Section 5.5 proposed a hybrid method for estimating position using a combination of shape fitting and an absolute position measurement. This section

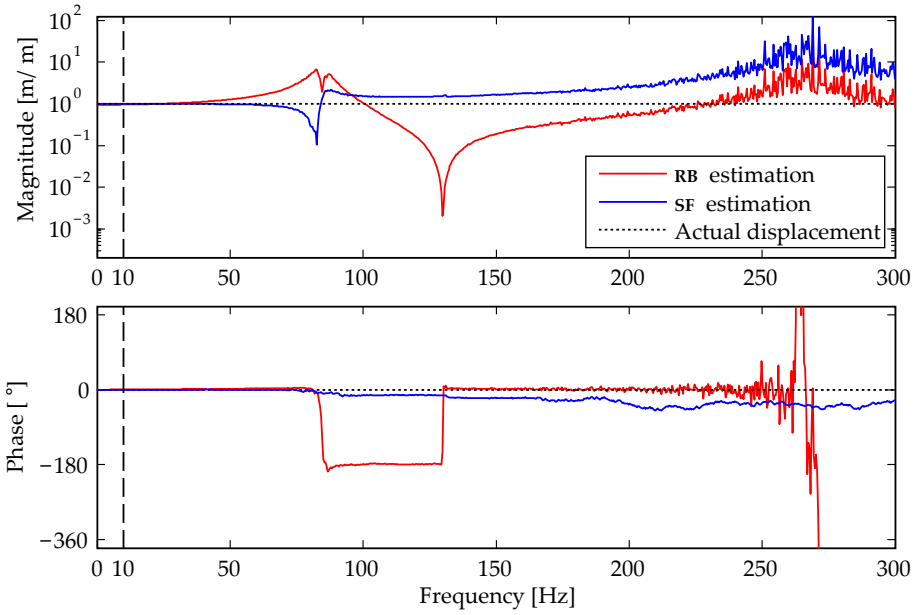



(a) Bode plot of the transfer function from actual displacement to the RB and SF estimates.

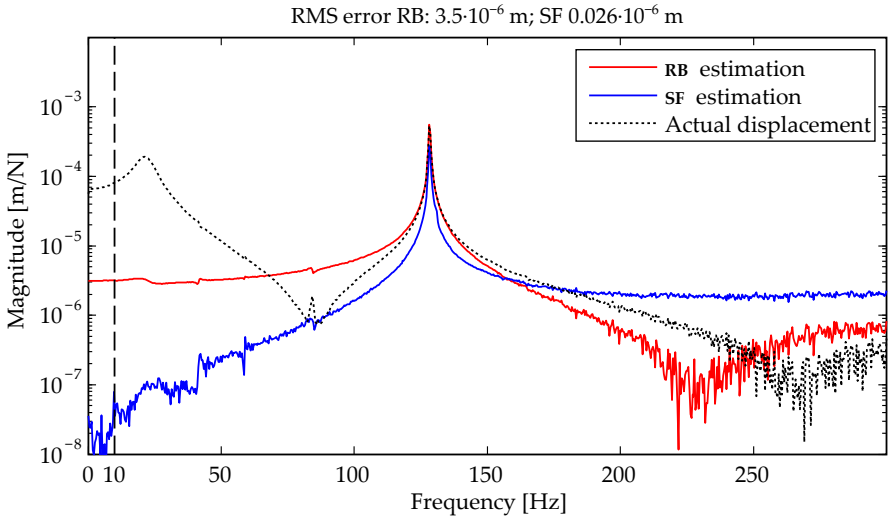



(b) Transfer function from the disturbance force to the error of the RB and SF estimates and the actual displacement.

Figure 7.12: Optimisation for disturbances in the frequency band 0.1 – 1 Hz.

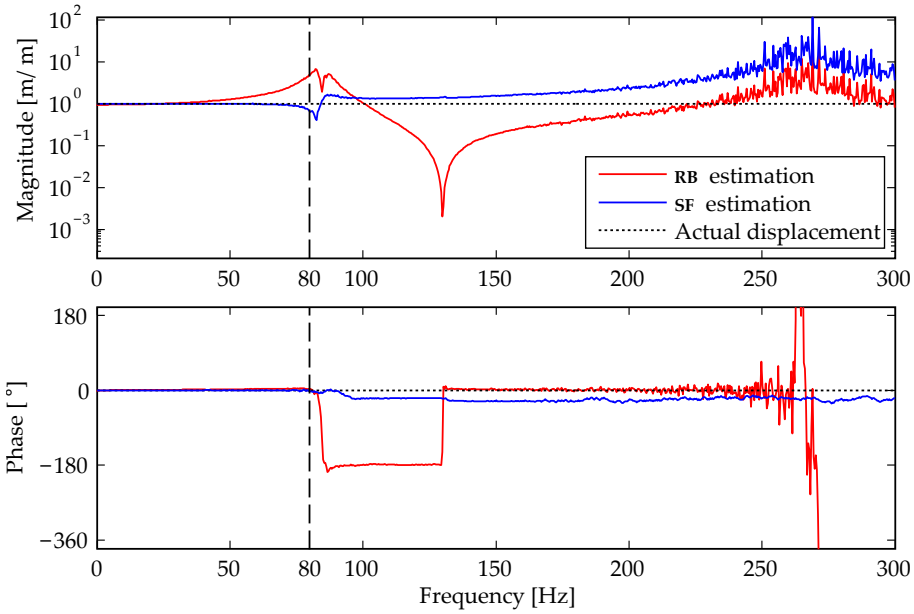



(a) Bode plot of the transfer function from actual displacement to the RB and SF estimates.

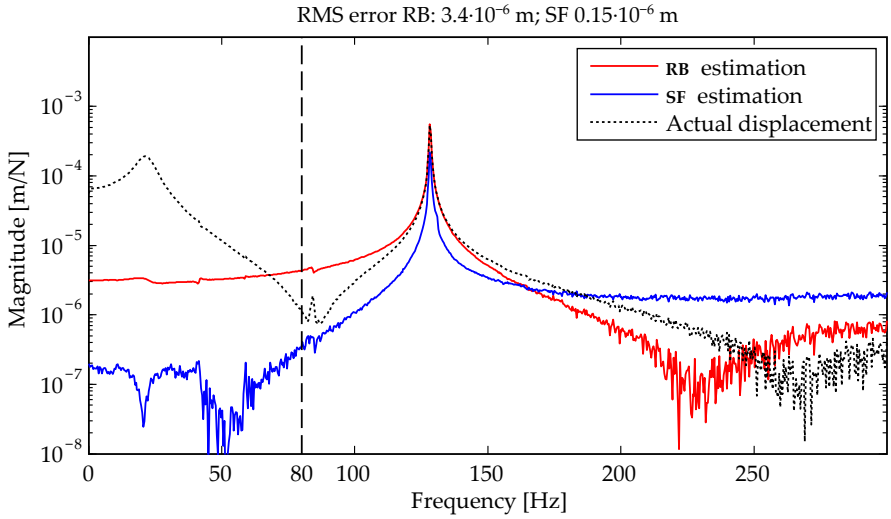



(b) Transfer function from the disturbance force to the error of the RB and SF estimates and the actual displacement.

Figure 7.13: Optimisation for disturbances in the frequency band 0.1 – 10 Hz.

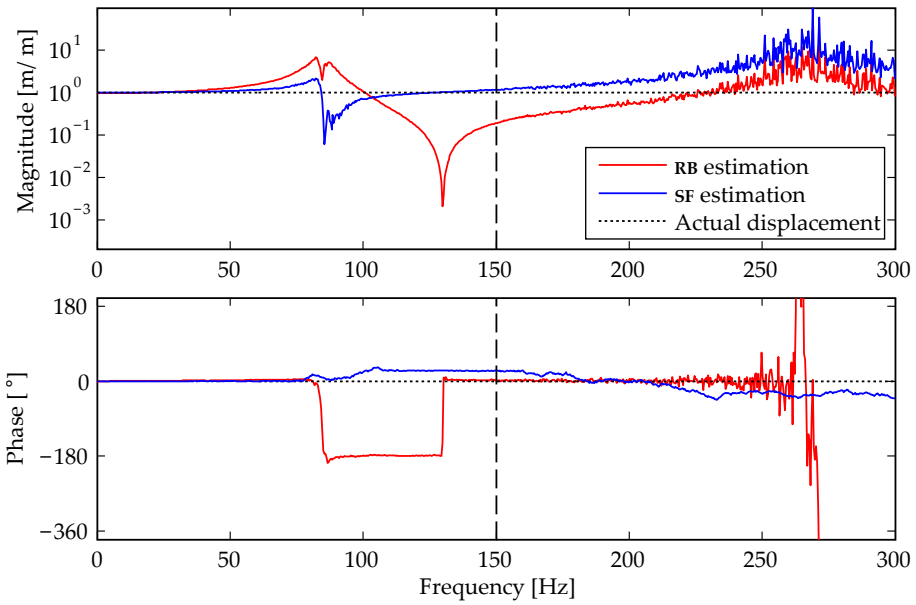



(a) Bode plot of the transfer function from actual displacement to the RB and SF estimates.

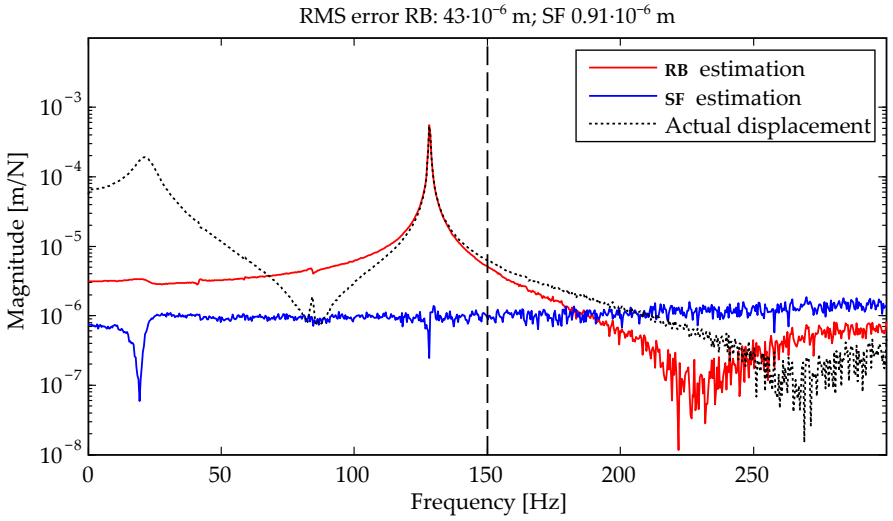



(b) Transfer function from the disturbance force to the error of the RB and SF estimates and the actual displacement.

Figure 7.14: Optimisation for disturbances in the frequency band 0.1 – 80 Hz.




(a) Bode plot of the transfer function from actual displacement to the RB and SF estimates.




(b) Transfer function from the disturbance force to the error of the RB and SF estimates and the actual displacement.

Figure 7.15: Optimisation for disturbances in the frequency band 0.1 – 150 Hz.

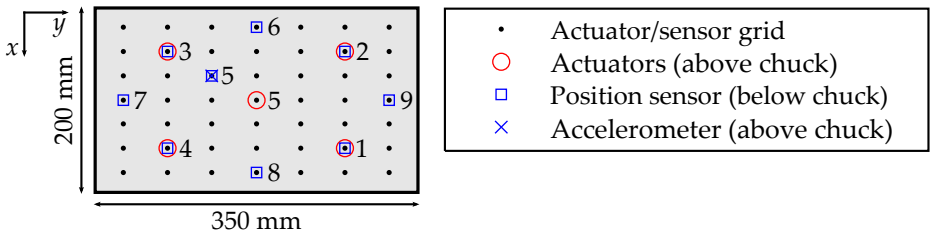



Figure 7.16: Configuration of the actuators and sensors used for the double integration of acceleration experiments. As compared to the initial sensor configuration (Figure 7.6), the centre position sensor was moved so that the accelerometer could be placed above it.

implements the hybrid method on the experimental setup. The absolute position is obtained from an accelerometer. First, the method is presented and then experimental results are shown.

### 7.5.1 Method

An accelerometer was added to the setup. The accelerometer had to be low-noise and light-weight and was selected based on an overview of high-performance commercial accelerometers (Appendix E). It was observed that the noise level of the accelerometers is inversely proportional with their mass. Based on a comparison, the capacitive MEMS accelerometer 2220-2 of SiliconDesigns (similar to Dytran 7500A3) was selected. This sensor has a mass of 10 g, a flat noise level of $80 \cdot 10^{-6}$ $\mathrm{ms}^{-2}/\sqrt{\mathrm{Hz}}$ and a measurement range from 0 to 400 Hz.

The accelerometer was used to obtain a position signal by double integration. The accelerometer was placed in line with the centre position sensor, which was used as a reference sensor. The sensors were placed according to Figure 7.16, they were moved away from the centre of the chuck to avoid interference of the accelerometer with the actuator.

**Construction of the hybrid estimate** The hybrid signal is constructed from the double integrated accelerometer signal and the SF estimate. After double integration, the accelerometer signal was filtered using a third-order Butterworth filter. This filter order is sufficiently high as to remove a double integrated constant acceleration signal. The position signals were lowpass filtered using the filter complementary to this third-order Butterworth filter. They were then used for shape fitting. Finally, the SF estimate was combined with the double integrated accelerometer signal to yield the hybrid estimate. As the two filters by definition add up to a magnitude of 1 and a phase of 0, this hybrid estimate is exact if the input signals are also exact (Figure 7.17).

**Calibration step** First a calibration signal and then the actual signals for estimation were applied to the setup using the centre actuator. The signals of the position sensor and the accelerometer were recorded for 20 s. The signals were highpass filtered with a 0.1 Hz second-order Butterworth filter to remove

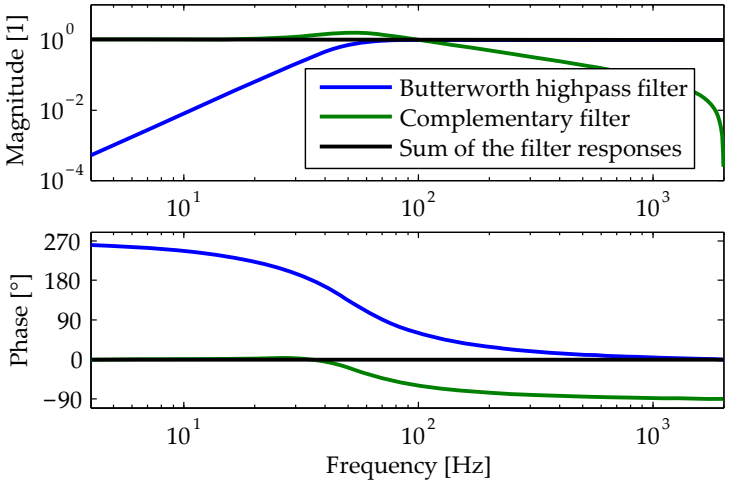



Figure 7.17: Responses of the filters used for the hybrid estimator. The double integrated acceleration signal was highpass filtered using a third order Butterworth filter (blue). The SF estimate is filtered with the complement of the Butterworth filter (green). By definition, the two filters add up to a magnitude of 1 and a phase of 0 (black).

calibration errors and low-frequency drift. All signals were lowpass filtered with a 1 kHz lowpass Butterworth filter to reduce the noise level of the signals.

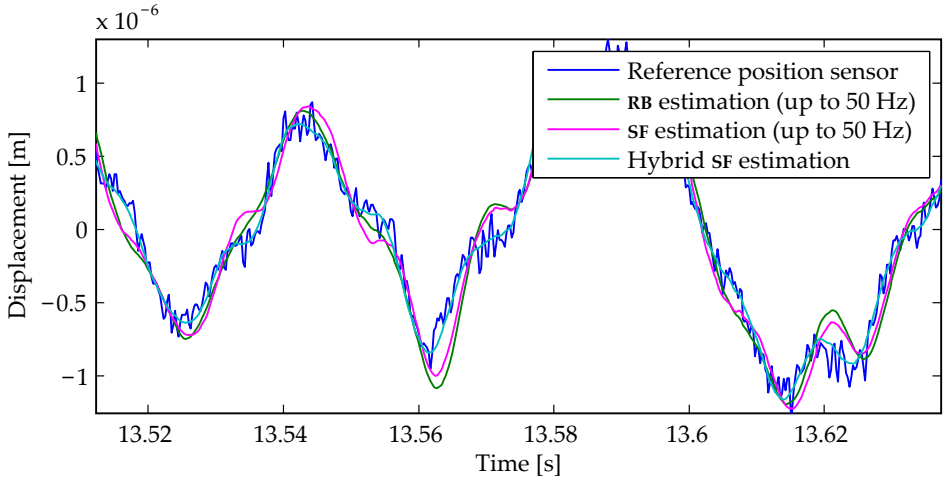
The calibration signal was a pseudo-random binary signal. The corresponding sensor signals were recorded. Then, estimation matrix $\boldsymbol{B}$ was calculated for standard shape fitting. The position signals were lowpass filtered using the complementary lowpass filter to form a time-domain snapshot matrix representative for the bandwidth 1 – 50 Hz.



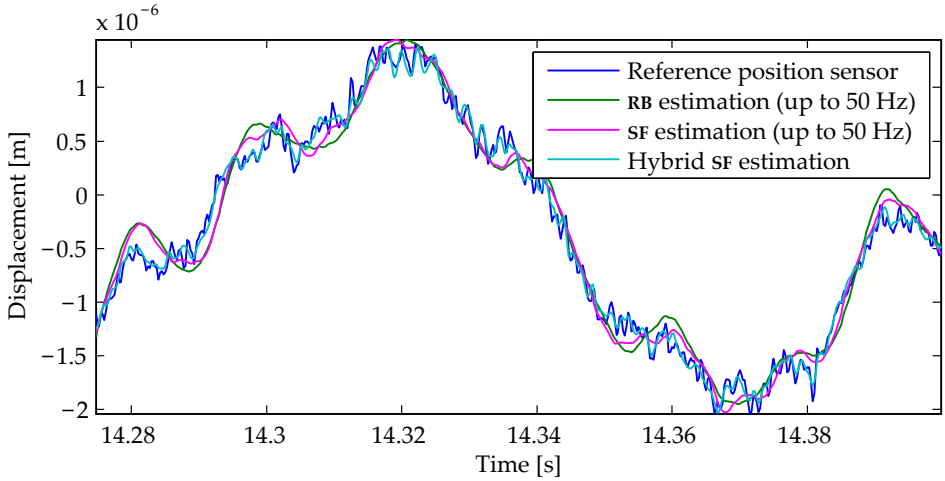
**Estimation step** Estimation was performed using two measurements, one with a white noise disturbance signal that was filtered with a sixth-order lowpass filter with a cutoff frequency of 80 Hz, the other with an unfiltered white noise disturbance signal. Then the hybrid estimates were calculated.

### 7.5.2 Results

Figure 7.18 shows the actual PoI displacement and the estimates in a time interval of 0.125 s. The upper figure shows the results for the lowpass filtered disturbance signal, the lower figure the results for the fully white noise signal. The RB and SF estimates contain only low-frequency signal content and deviate relatively much from the actual PoI displacement. The hybrid signal estimates better the actual displacement. The actual displacement shows comparable spikes in both plots, suggesting that these spikes must be attributed to sensor noise rather than high-frequency displacement. The hybrid estimate for the lowpass filtered disturbance (upper figure) is, however, smoother than the one for the white noise disturbance (lower figure), suggesting that the hybrid estimate contains less noise than the reference sensor.




(a) White noise disturbance at centre actuator, lowpass filtered at 80 Hz.




(b) White noise disturbance at centre actuator, not filtered.

Figure 7.18: Comparison between the displacement estimate of the hybrid SF estimator and the displacement measurement of the reference sensor. The hybrid estimator makes use of the estimate of a SF estimator that was optimised for the low-frequency part of the response up to 50 Hz. The higher frequencies are obtained from double integration of the acceleration signal. For illustration, the RB estimate and the SF estimate up to 50 Hz are included in the plots.

Table 7.3: A comparison of the hybrid SF estimation with the standard SF estimate and the RB estimate. The comparison is in terms of the RMS values of the estimation error.

| | White noise disturbance | |
|---|---|---|
| | lowpass filtered $f_{co} = 80$ Hz | not filtered |
| | RMS value [m] | RMS value [m] |
| Reference position sensor | $1.403 \cdot 10^{-6}$ | $0.864 \cdot 10^{-6}$ |
| Flexible residue | $0.098 \cdot 10^{-6}$ | $0.158 \cdot 10^{-6}$ |
| Error RB estimation | $0.098 \cdot 10^{-6}$ | $0.158 \cdot 10^{-6}$ |
| Error SF estimation | $0.084 \cdot 10^{-6}$ | $0.113 \cdot 10^{-6}$ |
| Error RB estimation (up to 50 Hz) | $0.188 \cdot 10^{-6}$ | $0.202 \cdot 10^{-6}$ |
| Error SF estimation (up to 50 Hz) | $0.173 \cdot 10^{-6}$ | $0.180 \cdot 10^{-6}$ |
| Error hybrid SF estimation | $0.083 \cdot 10^{-6}$ | $0.075 \cdot 10^{-6}$ |

Table 7.3 shows RMS values of the signals. The results of the one experiment can not be compared directly to the other, as the disturbance power of the first experiment was higher, which can be noted from its higher RMS PoI displacement. By definition, the error of the RB estimate equals the flexible residue. The error of the RB and SF estimates after filtering with the complementary filter with a cutoff frequency of $f_{co} = 50$ Hz are higher than the RB and SF estimates as they are compared to the reference PoI displacement, which is not filtered with that filter. In the first experiment, with disturbances up to 80 Hz, the hybrid SF estimate and the standard SF estimate show a comparable error level around $0.084 \cdot 10^{-6}$ m. This is, however, close to the noise level of the reference sensor. It should be noted that the RMS noise level of the sensors at a bandwidth of 1 kHz is around $0.071 \cdot 10^{-6}$ m, so that smaller estimation error can not be detected without averaging techniques. In the second experiment with the unfiltered white noise disturbance, the hybrid SF estimate is significantly improved with respect to the standard SF estimate.



## 7.6 Conclusions

An experimental setup, mimicking the wafer chuck system, was designed, build and validated. The purpose of the setup was to verify the system model and to demonstrate shape fitting in a physical setup.

The setup consisted of a plate, fixed in the in-plane directions and position controlled in the out-of plane direction using position sensors, force actuators and a RB controller. To be able to measure both the RB motion and the relatively small deformation of the plate, the position sensors were selected to have a 1 mm range and a high resolution. Furthermore, a metrology frame with a high mass and a low eigenfrequency was implemented to minimise the influence of floor noise and actuator reaction forces on the sensor signals.

The setup validation showed that the setup was suitable for estimation. The stiffness in out-of-plane direction, originating from the Lorenz actuators and in-plane suspension, was found to be small, i.e. at least 100 times smaller than the controller stiffness. The noise floor in terms of deformation equivalent noise was 71 nm ($1\sigma$, bandwidth 1000 Hz), which was much higher than expected. This was primarily caused by a higher sensor noise. Also, the transfer from actuator reaction force to motion of the metrology frame was examined. The resulting unwanted parasitic signals proved to be low compared to the real deformations caused by the actuator forces, at least a factor of $10^4$ smaller.

The setup was used to verify the system model by comparing its transfer function to the one of the model. The Bode plots of the model's transfer from the centre disturbance actuator to the RB displacement and deformation were shown to closely resemble the ones of the setup.

Finally, the setup was used for the demonstration of shape fitting. This was done for a case with one disturbance actuator and five flexible residuals. First, a calibration experiment was performed to be able to calculate the estimation matrix and the expected estimation error. Then, the actual estimation was carried out using this estimation matrix. To be able to reach the expected small estimation errors it proved to be important to remove the static and low-frequency deformation content. For a bandwidth from 0.1 to 150 Hz, a SF error of $0.49 \cdot 10^{-6}$ m ($1\sigma$) was obtained, which is almost a factor of 40 smaller than the error of the RB method. The results for different disturbance bands show that the SF estimation error increases for an increasing upper frequency of the band.

A hybrid estimator combines the low-frequency signal content of a SF estimate with the high-frequency content of a double integrated acceleration signal. This allows to optimise the SF estimator for a relatively narrow frequency band and on the other hand to use the double integrated acceleration with a lower noise level, leading to a low total noise level. For the response to a white noise disturbance, the hybrid estimator yielded an error of $0.075 \cdot 10^{-6}$ m ($1\sigma$, bandwidth 1000 Hz), which is close to the noise level of the reference sensor, showing the effectiveness of the hybrid estimator.

# Chapter 8

# Conclusions and outlook

The goal of this thesis is to develop an estimation methodology for wafer chuck deformation, based on shape fitting principles, considering the aim for low estimation error and the specific requirements related to the wafer chuck application. The thesis shows that the estimation can be performed effectively using the relatively simple shape fitting method. It is a promising method for the estimation of wafer chuck deformation and can be one of the steps forward in the transition to larger wafers and, ultimately, cheaper and more powerful electronic chips.

Shape fitting estimates the positions of the target points using the measurement signals of the position or other measurands at the measurement points. The Shape Fitting (SF) estimator makes use of the correlation between the measurements and the quantity to be estimated. The estimate is found from a linear combination of the measurements. The method does not take into account the measurement signals' history and is, as such, of particular interest for quasi-static or stiffness-governed problems.

To meet the goal of the thesis several research steps were carried out. Existing SF methods in literature were collected, reviewed and compared. The requirements and boundary conditions of the wafer chuck application were studied and the SF method was developed further to make it applicable in the application. The SF method was applied in a numerical study with a simplified wafer chuck model and to an experimental setup.

This chapter presents the conclusions that are drawn from the results in this thesis. Furthermore, the recommendations based on the work are presented.

## 8.1 Conclusions

**Constructing the shape fitting estimator using a snapshot matrix**

- To obtain high estimation performance it is important to include the relevant foreknowledge on the wafer chuck's mechanics and its typical disturbances when constructing the estimator. In the wafer chuck system it is known that the majority of the disturbances in immersion lithography

→ Sct. 2.1

machines is concentrated in the exposure area and that frequencies of the disturbance are typically well below the lowest mechanical eigenfrequency of the wafer chuck.

→ Sct. 3.3

- Different types of shapes can be used for shape fitting. Traditionally, shape fitting is used in combination with modeshapes. This is a sensible choice if no specific information about the disturbances is known. In that case it may be assumed that the eigenmodes corresponding to the lowest eigenfrequencies are excited around their eigenfrequencies and form, as such, the major contribution to the deformation.
  However, in a system that is excited with frequencies well below the lowest mechanical eigenfrequency relatively many modeshapes contribute to the deformation. If strain is chosen as a measurand, modeshapes are especially inefficient, as a relatively many modeshapes are needed to describe the strain profile of a static deformation shape. It is, therefore, better to use the deformation shapes that correspond to the disturbances that typically work at the system.

→ Sct. 3.5

- The typical deformation shapes form together the so-called snapshot matrix. From this matrix the SF estimator is calculated, thus including the foreknowledge on the system's mechanics and its typical disturbances. A good snapshot matrix describes the typical deformations of the system well, such that the system's response can be described as a linear combination of the snapshots.
  The snapshot-matrix can be obtained from measurements or from a model, but neither possibility is trivial. The model of the system's mechanics and disturbances might not fully match the actual system, whereas measurement of the system's full response, including the PoI's response in case of the wafer chuck, is not available under the actual disturbance conditions.

**Techniques for the calculation of the shape fitting estimator**

→ Sct. 3.5

- The SF estimator can be calculated based on the snapshot matrix using the POM and the LS technique. The POM technique is known in the field of temperature estimation in high-precision machines and CPUS. The LS technique is often used in the field of shape estimation based on strain measurements.

→ Sct. 4.3

  A fundamental difference between the Proper Orthogonal Mode (POM) and Least Squares (LS) technique is that the latter optimises for each target point separately. The estimation performance of the POM technique, however, depends also on the other included target points, as was illustrated in the numerical study. For this reason the selection of target points in the POM is not trivial.

→ Sct. 4.3

- The POM technique extracts the most prominent shapes from the snapshot matrix and uses them to perform shape fitting. The less prominent patterns may, however, cause spillover, leading to the methodical estimation error. The LS technique uses the full snapshot matrix for shape fitting, by definition leading to a minimum methodical error.

- Not only the methodical error but also sensor noise may lead to significant estimation error. An estimation matrix $\boldsymbol{B}$ was defined, whose columns correspond to the applied sensors and whose rows correspond to the target degrees of freedom. This matrix is helpful in assessing the estimation error due to sensor noise. → Sct. 3.5
  Both the POM and the LS techniques allow for regularisation in order to reduce the influence of sensor noise, the POM technique by truncating the number of prominent shapes, the LS technique by truncating the number of singular vectors. The truncation, however, leads to a loss of information on the original snapshot matrix, resulting in a higher methodical error. To attain minimum total estimation error, a trade-off has to be made between methodical error and noise error. It is proposed to use a regularised LS technique, which optimises directly for lowest total estimation error. → Sct. 4.3

**Measurands and feasible sensor configurations**

- The position of the wafer chuck in out-of-plane direction is in immersion lithography systems found from four position sensors at the chuck's corners. The deformation on top of the Rigid Body (RB) position can be estimated using shape fitting. To do this, additional position sensors or sensors that measure measurands proportional to deformation, such as strain and curvature, can be placed. Strain sensors may be placed at will at the wafer chuck's surface. → Sct. 2.4

- Position measurements need to be performed with respect to a stable reference, i.e. the metrology frame in lithography machines. The metrology frame is located above the wafer chuck, around the lens column. Position sensors can not be mounted at the wafer area of the chuck and preferably not too close to the wafer area. Thus, the possibility for position measurements is limited to the sides of the wafer chuck. → Sct. 2.4
  Absolute position can be measured using inertia-based sensors. Examples include the position measurement with respect to a decoupling mass, like a geophone, or the double integration of an accelerometer signal. The inertia-based sensor have the advantage of not needing an external reference but suffer from relatively large noise at lower frequency. → Sct. 2.4

**Shape fitting in a system with a position controller**

- The wafer chuck is virtually connected to the metrology frame by a position control system. The controller introduces forces to the system, as a response to the wafer chuck's displacement and deformation. Shape fitting using the black box approach generalises the shape fitting in a simple way to such a position controlled system. It performs estimation directly based on the sensor signals, without using the applied actuation signals. → Sct. 5.1

- The positions of the measurement and target point can be expressed as a combination of RB modeshapes weighed by the RB coordinates and the flexible modeshapes weighed by the flexible coordinates. The RB → Sct. 5.2

coordinates of the closed-loop controlled wafer chuck behave dynamically, also in the frequency band up to the wafer chuck's lowest mechanical eigenfrequency. The flexible coordinates, on the other hand, show only limited dynamic behaviour, although they do not behave fully quasi-statically – especially near the controller bandwidth – in the frequency band up to the lowest mechanical eigenfrequency.

→ Sct. 5.2

- The RB positions at the measurement and target points can be estimated from at least three position measurements with respect to the stable metrology frame. The RB modes are decoupled by subtracting the estimate RB position from the actual position, yielding the *flexible residue*. Although the flexible residue is both composed of RB modeshapes and flexible modeshapes, it is not a function of the RB coordinates, but solely of the flexible coordinates. The flexible residue shows, therefore, only limited dynamic behaviour, at least up to the controller bandwidth, allowing to apply shape fitting in an effective way.

→ Sct. 5.3

- If the flexible residue does not behave fully quasi-statically, the response at the different frequencies needs to be taken into account. If the snapshot matrix is obtained from a model, this is possible by sampling a grid of frequencies in the frequency band that must be optimised for.
  A snapshot matrix should, in case of dynamic effects, not only account for the relative magnitudes of the signals, but also for their relative phase. It is possible to achieve this by including more snapshots at different phase angles. We introduced the complex snapshot matrix that includes the phase behaviour in an efficient way. The regularised LS technique was adapted to facilitate the use of a complex snapshot matrix.

→ Sct. 5.3

- The performance of the SF estimator depends on the bandwidth of the disturbances. The SF estimator should be optimised for the bandwidth of the disturbances. A larger bandwidth, however, leads typically to a larger

→ Sct. 7.4

  estimation error. Results from the experimental setup showed that the SF estimation can also be performed if the disturbances band includes a resonance frequency. Although the SF estimation error was higher for such a large disturbance bandwidth, it was still a factor of 40 smaller than the RB estimate error.

→ Sct. 5.5

- A hybrid estimator was developed that fuses the signal of a SF estimator with a double integrated acceleration signal. In this way, the SF estimator can be optimised for a relatively small bandwidth for which it yields a relatively low error, whereas only the higher frequency part of the double integrated acceleration signal needs to be used, which has a relatively low

→ Sct. 7.4

  noise level. The hybrid estimator was implemented in the experimental setup and showed to be effective, yielding an estimation error close to the noise level of the reference sensor.

→ Sct. 5.4

- Numerical experiments were carried out with a plate and a controller mimicking the wafer chuck. The sensor configuration with eight position sensors was used and the disturbances were distributed over the exposure

area with a RMS magnitude of 1 N. The average methodical error of the SF estimator ranged between 2.0 nm (quasi-static disturbances) and 6.5 nm (disturbances up to controller bandwidth). The error due to the sensor noise (0.1 nm RMS) ranged between 0.13 nm and 0.21 nm, so that sensor noise requirements could be relieved by approximately a factor of 12 without a significant increase of the total estimation error.

**Sensor placement**

- To be able to perform estimation effectively it is important to select an adequate sensor configuration. Choosing such a configuration is in general not trivial, especially when placing strain sensors. Therefore, automated sensor placement algorithms from literature were examined. Especially the placement algorithms that were developed in the field of experimental dynamics are of interest in the context of shape fitting, as they aim at identifying and distinguishing shapes. → Sct. 6.1
  The methods of Kammer (EFI) and Tasker and Liu (TA) were selected for further study, along with the methods of Ranieri and his co-workers (RA12 and FS), which were developed directly for shape fitting. All four methods start with an initial sensor candidate set, of which the least promising sensors are eliminated one-by-one. The methods aim at finding the sensor configuration that minimises the estimation error of the amplitudes of a limited number of shapes. This optimisation goal is not equal, but closely related to the aim of shape fitting. → Sct. 6.1

- The authors of the FS algorithm showed that FS leads to better placements than a few other placement methods. We replicated the results of FS for placement with different types of random shapes and showed that FS is in most cases outperformed by EFI and TA. Next to this, it was shown that the formulation of the FS algorithm contains an error, which, if not corrected, leads to extremely poor sensor placements. → Sct. 6.1

- A new performance metric, the Sensor Noise Amplification (SNA), was proposed to evaluate the sensor configurations. The methods were compared in different numerical experiments in which a limited number of modeshapes had to be distinguished either using position sensors or strain sensors. When placing eight position sensors, the EFI and TA algorithms yielded the same result, namely a placement with four sensors at the corners and the other sensors at the centres of the sides, which has a lower SNA than the configurations of RA12 and FS algorithm. → Sct. 6.2

- Also placement of strain sensors was performed. Identifying the modeshapes using strain sensors involves not only the determination of the sensor location but also their orientation. The algorithms were, amongst others, compared based on their performance in terms of the resulting SNAs and the convergence to well performing sensor configurations. The TA algorithm yielded in general superior results over the other algorithms, but did not converge for all sensor candidate sets. It was, furthermore, → Sct. 6.2

observed that a small candidate set can equally well lead to a high performance sensor configuration as a large candidate set. Thus, given the fact that calculation time is restricted, sensor placement should be performed on many small sensor candidate sets rather than a single large one. This practically eliminates the chance of no convergence and improves the chance of obtaining a high-performance sensor configuration.

→ Sct. 6.3

- A procedure was developed to incorporate the placement algorithms in the SF framework. This procedure generates the input shapes for the placement algorithms by extracting the most prominent shapes from the snapshot matrix using the Proper Orthogonal Decomposition (POD), while accounting for the fact that the POD does not distinguish measurement and target points. Still, it is not trivial how sensor placement should be carried out for the wafer chuck, as the disturbance source travels in time, so that the sensor configuration should optimise for some average performance.

→ Sct. 6.4

  In a numerical experiment, the sensor configuration was optimised for situation in which the lens column is above the centre of the wafer chuck. Using the TA algorithm, a sensor configuration with eight position sensors was found that reduced the maximum estimation error by 40 % as compared to the manually selected sensor configuration.

## 8.2 Final conclusion and future outlook

- The shape fitting method is a suitable candidate for estimation of the wafer chuck's deformation. Traditionally, shape fitting is used in combination with modeshapes. In combination with the snapshot matrix, however, it allows for including foreknowledge on the mechanics and the disturbances. The snapshot matrix also allows to incorporate the uncertainty of the relative magnitude and position of the disturbances. Shape fitting is, compared to dynamic methods, a relatively simple method and allows the integration of existing sensor placement methods. The contributions of methodical error and sensor noise can be compared in a straightforward manner, making it possible to derive sensor specifications and to decide whether or not the number of sensors involved in the estimation should be increased or deceased.

- The experiments with the model and the setup demonstrated the effectiveness of SF estimation. SF estimation for average deformations of 200 nm, caused by quasi-static disturbances in the exposure area, yielded an average methodic estimation error 2 nm. This result was obtained with four additional displacement sensors at the chuck's side centres and a relatively thin plate mimicking the wafer chuck.
  Of course, the performance of a SF estimator depends to a large extent on the mechanics of the system and the disturbances that work on the system and their uncertainties. To obtain a better idea of the performance in a wafer scanner, it is recommended to evaluate the method using a more realistic model of the wafer chuck and its controller, together with more

accurate information regarding the disturbance distribution and the noise levels of the sensors. Of these three, especially the disturbance distribution is likely to have most influence on the estimation performance.
The numeric wafer chuck cases in this thesis focussed on estimation in immersion wafer scanners. The new generation of lithography machines which is currently under development makes use of Extreme Ultraviolet (euv) light. The euv machines are developed for 300 mm wafers, but also in those machines a transition to larger wafer seems ultimately unavoidable. The disturbances work in euv machines at different locations of the wafer chuck than in immersion machines and may, for example, be related to actuator noise and cooling system disturbances. Still, there is no reason why shape fitting could not be applied in these machines.

- The main building blocks for performing shape fitting were presented in this thesis. The method needs to be further refined to meet some of the specific properties of the wafer chuck.
  For example, the inherent problems of obtaining a snapshot matrix need to be addressed further. In case the matrix is obtained from a model, uncertainties in the mechanical model and the disturbance model might lead to an increased estimation error. The snapshot matrix can, however, not be fully obtained from measurements, as a direct measurement at the target points is not possible – at least not in presence of the realistic disturbances. This can possibly be solved by matching the finite element model to the real system using measurements, so that the improved model can be used to generate the snapshot matrix.

- Next to estimation with shape fitting, the wafer chuck's deformation could also be estimated using dynamic estimation methods. An example of such a technique is the unknown input Receding Horizon Input Estimator (rhie), which is topic of research by a colleague in the same project at Delft University of Technology. This method was adapted for use in wafer chuck and, not long ago in a collaborative effort, tests with this method on the experimental setup (Chapter 7) were initiated. One of the goals of this research collaboration is the comparison of the performance of the rhie and the sf estimator, a goal which is of interest for both industry and academia.

# Appendix A

# Wafer chuck scaling properties

This appendix corresponds to the Sections 1.3, 1.4 and 2.4. It illustrates how the main parameters and properties of the wafer chuck scale when the wafer chuck's size is increased, while its stiffness, lowest mechanical eigenfrequency or mass is kept constant.

If the wafer diameter is increased from 300 mm to 450 mm, the wafer chuck's length, $l$, and width, $w$, need to scale by a factor of 1.5 ($S_l = S_w = 1.5$). The thickness, $h$, of the chuck is studied, along with the chuck's mass, $m$; stiffness, $k$; lowest eigenfrequency, $f_1$; required actuator forces, $F$, and the generated actuator heat in terms of power, $P$. Table A.1 summarises the scale factors for the three scenarios in which the chuck's stiffness, lowest mechanical eigenfrequency or mass is kept constant.

Table A.1: Scale factors $S$ of some of the wafer chuck's parameters and properties in three scenarios.

| Scenario | $S_l = S_w$ | $S_h$ | $S_k$ | $S_f$ | $S_m = S_F$ | $S_P$ |
|---|---|---|---|---|---|---|
| Constant stiffness | 1.5 | 1.3 | 1 | 0.67 | 2.9 | 8.7 |
| Constant lowest eigenfrequency | 1.5 | 1.7 | 2.3 | 1 | 3.9 | 15 |
| Constant mass | 1.5 | 0.44 | 0.039 | 0.13 | 1 | 1 |

**Constant stiffness scenario** If the static deformation of the chuck needs to stay equal, the chuck's thickness has to increase. The bending stiffness of a plate is assumed to be comparable to that of a beam (cf Eq. 7.5),

$$k = \frac{F}{\delta} = \frac{48E\frac{1}{12}bh^3}{l^3} = \frac{4Eh^3}{l^2}, \tag{A.1}$$

For a constant stiffness it holds that $h \propto \sqrt[3]{l^2}$, so that the chuck's thickness scales with a factor $S_h = S_l^{2/3} = 1.3$.

According to Leissa [96] (Eq. 7.2), the lowest eigenfrequency of a square plate with all sides free depends in the following way on its dimensions:

$$f_1 \propto \frac{\sqrt{h^3}}{l^2}.$$

Thus, the lowest eigenfrequency changes a factor $S_f = S_h^{1,5} S_l^{-2} = 0.67$.

The chuck's mass increases by factor $S_m = S_l^2 S_h = 2.9$. If the chuck's acceleration $a$ remains unchanged, the actuation forces and the reaction forces increase, according to $F = ma$, as $S_F = S_m = 2.9$.

The current through a Lorenz actuator is proportional to the applied forces. The Joule heating of the actuators, is in turn a function of the actuator current, with $P = I^2 R$, so that $S_P = S_F^2 = 8.7$.

**Constant lowest eigenfrequency scenario** If the lowest eigenfrequency is kept constant, the chuck's thickness scales, according to $1 = S_f = S_h^{3/2} S_l^{-2}$, a factor of $S_h = S_l^{4/3} = 1.7$. The chuck's stiffness scales as $S_k = S_h^3 S_l^{-2} = S_l^2 = 2.3$. The chucks mass scales as $S_m = S_l^{10/3} = 3.9$ and the heat produced by the actuators as $S_P = S_l^{20/3} = 15$.

**Constant mass scenario** In case the chuck's mass should remain equal, the chuck's thickness must decrease as $1 = S_m = S_l S_w S_h$, so that $S_h = S_l^{-2} = 0.44$. The chuck's stiffness scales as $S_k = S_h^3 S_l^{-2} = S_l^{-8} = 0.039$ and its lowest eigenfrequency as $S_f = S_h^{3/2} S_l^{-2} = S_l^{-5} = 0.13$. The actuator forces and the produced heat stay equal for constant mass.

# Appendix B

# Out-of-plane force due to immersion film dynamics

This appendix corresponds to Section 2.1.4 and derives the forces that work at the interface of the wafer chuck and the immersion film due damping and inertia in the film if there is relative motion between the wafer chuck and the lens column. The analysis makes use of squeeze film assumptions and compares the resulting equations to an expression from literature.

## B.1 Assumptions

Consider the axis-symmetrical immersion film as depicted in Figure B.1. The layer has a height $h(t)$ and radius $R$. Coordinates $y$ ($-\frac{h}{2} \leq y \leq \frac{h}{2}$), $r$ and $\theta$ describe the position in the fluid film. The height $h(t)$ of the immersion film is orders of magnitude smaller than the radius if the immersion layer. It is therefore acceptable to use squeeze-film assumptions.

Conservation of volume in a cylindrical control volume is described as follows

$$h(t)\pi(r + \hat{r}(r,t))^2 = h_0 \pi r^2, \tag{B.1}$$

in which $\hat{r}$ denotes the bulk displacement at position $r$ and $h_0$ the nominal height, for which there is just no bulk displacement. From this equation the bulk displacement $\hat{r}(r,t)$ can be found:

$$\hat{r}(r,t) = \frac{-h(t) + \sqrt{h(t)h_0}}{h(t)} r = \left( \sqrt{\frac{h_0}{h(t)}} - 1 \right) r. \tag{B.2}$$

## B.2 Force due to viscous damping

The bulk velocity is defined as the time derivative of bulk displacement

$$v_{\text{bulk}}(r,t) = \frac{d}{dt}\hat{r}(r,t) = -\frac{\sqrt{h_0}\dot{h}(t)}{2\sqrt{h^3(t)}} r. \tag{B.3}$$

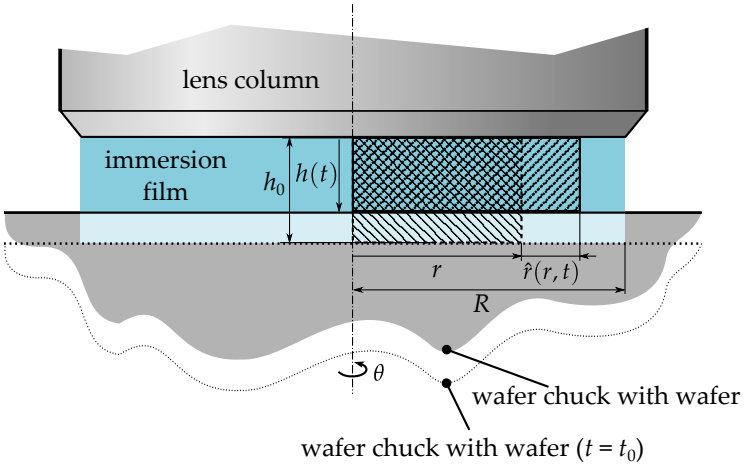



Figure B.1: Model of the bulk displacement of the fluid in the immersion film. The volume of the fluid stays constant of the wafer chuck moves from the initial position $h_0$ with respect to the lens column to its new position $h(t)$. Thus, the radius of a cylindrical control volume changes by $\hat{r}$ with respect to its initial radius $r$.

The flow is assumed to be laminar and to have a parabolic velocity profile according to

$$\frac{\partial u}{\partial y} = 12\frac{y}{h^2}v_{\text{bulk}}.$$

Due to the viscosity of the immersion fluid power is dissipated, causing damping. The power dissipated in a hollow cylinder with an infinitesimal wall thickness $dr$ is

$$\begin{aligned} dP_{\text{damping}}(r,t) &= 2\pi r\left[\int_{-\frac{h}{2}}^{\frac{h}{2}} \mu\left(\frac{\partial u}{\partial y}\right)^2 dy\right]dr \\ &= 24\pi\mu\frac{v_{\text{bulk}}^2}{h}rdr. \end{aligned} \tag{B.4}$$

The force difference between inner surface and outer surface of the hollow cylinder is

$$dF_{\text{damping}}(r,t) = \frac{24\pi\mu\frac{v_{\text{bulk}}^2(r,t)}{h(t)}rdr}{v_{\text{bulk}}(r,t)} = 24\pi\mu\frac{v_{\text{bulk}}(r,t)}{h(t)}rdr, \tag{B.5}$$

leading to a pressure difference

$$dp_{\text{damping}}(r,t) = \frac{24\pi\mu\frac{v_{\text{bulk}}(r,t)}{h(t)}rdr}{2\pi r\cdot h(t)} = 12\mu\frac{v_{\text{bulk}}(r,t)}{h(t)^2}dr. \tag{B.6}$$

This result is integrated to obtain the pressure with respect to the ambient ($r = R$), assuming that $R$ is not a function of time:

$$\begin{aligned} p_{\text{damping}}(r,t) &= \int_r^R dp(r,t) \\ &= -3\mu\frac{\sqrt{h_0}\dot{h}(t)}{\sqrt{h^7(t)}}\Big(R^2 - r^2\Big). \end{aligned} \tag{B.7}$$

Finally, the total force is the surface integral of the pressure:

$$F_{\text{damping}}(t) = \int_0^R \int_0^{2\pi} p(r,t) r d\varphi dr = -\frac{3}{2}\pi\mu\frac{\sqrt{h_0}\dot{h}(t)}{\sqrt{h^7(t)}}R^4. \tag{B.8}$$

### B.2.1 Force due to inertia

The bulk acceleration $a_{\text{bulk}}$ is found using Eq. (B.3):

$$a_{\text{bulk}} = \frac{d^2\hat{r}(r,t)}{dt^2} = \sqrt{h_0}\left[\frac{3\big(\dot{h}(t)\big)^2}{4\sqrt{h^5(t)}} - \frac{\ddot{h}(t)}{2\sqrt{h^3(t)}}\right] r. \tag{B.9}$$

The force (per unit of length) needed to accelerate a hollow cylinder with infinitesimal width $dr$ is:

$$F_{dr}(r,t) = \rho(h(t)\cdot 2\pi r\cdot dr)a_{\text{bulk}} \tag{B.10}$$

and the corresponding power (per unit of length):

$$\begin{aligned} P_{dr,\text{inerta}}(r,t) &= \rho(h(t)\cdot 2\pi r\cdot dr)a_{\text{bulk}}v_{\text{bulk}} \\ &= -\rho\pi h_0\left[\frac{3\big(\dot{h}(t)\big)^2}{4h^3(t)} - \frac{\ddot{h}(t)}{2h^2(t)}\right]\dot{h}(t)r^3 dr. \end{aligned} \tag{B.11}$$

Integrating this yields the total power:

$$P_{\text{inerta}}(t) = -\rho\pi h_0\left[\frac{3\big(\dot{h}(t)\big)^2}{16h^3(t)} - \frac{\ddot{h}(t)}{8h^2(t)}\right]\dot{h}(t)R^4, \tag{B.12}$$

so that, after division with $\dot{h}(t)$, the force exerted on the wafer chuck due the inertia is found:

$$F_{\text{inerta}}(t) = -\rho\pi h_0\left[\frac{\ddot{h}(t)}{8h^2(t)} - \frac{3\big(\dot{h}(t)\big)^2}{16h^3(t)}\right]R^4. \tag{B.13}$$

## B.3 Comparison to literature

The total force that is exerted on the wafer chuck is the sum of Eqs. (B.8) and (B.13):

$$F_{\text{total}}(t) = F_{\text{damping}}(t) + F_{\text{inerta}}(t) = -\frac{3}{2}\pi\mu\frac{\sqrt{h_0}\dot{h}(t)}{\sqrt{h^7(t)}}R^4.$$

Assume that the displacement is a small perturbation around $h(t)$ of the form $h(t) = h_0 + A\sin\omega t$, with $A \ll h_0$. In that case the total force can be rewritten as:

$$F_{\text{total}}(t) = F_{\text{damping}} + F_{\text{inerta}} = -\pi R^4 \left[ \frac{3\mu\dot{h}(t)}{2h_0^3} + \frac{\ddot{h}(t)}{8h_0} - \frac{3\left(\dot{h}(t)\right)^2}{16h_0^2} \right]. \tag{B.14}$$

When comparing this result to the equation derived by [97],

$$F = -\pi R^4 \left[ \frac{3\mu\dot{h}}{2h_0^3} + \frac{3\rho\ddot{h}}{20h_0} - \frac{15\rho\left(\dot{h}(t)\right)^2}{56h_0^2} \right], \tag{B.15}$$

it is noticed that the first term (the damping term) of both equations is equal and that the second and the third term have the same form, but a different prefactor.

The prefactors of the second term and the third term of Eq. (B.14) differ by respectively 17 % and 30 % compared to the prefactors of Eq. (B.15), most likely due to a different modelling assumption. At the same time, this result shows that the model developed in this appendix is useful for acquainting qualitative insight of the damping and inertia effects in the immersion layer. Section 2.1.4 applies the the formula of Kuzma to the wafer chuck application and shows that the third term is non-significant for the typical parameters in the wafer chuck application.

# Appendix C

# Double integration of an acceleration signal

This appendix corresponds to Section 2.4.2 and derives the variance of the position signal that is obtained from double integration of an accelerometer signal. The derivation is formulated in the discrete time domain; the noise of the acceleration samples is assumed to be uncorrelated.

Assume an accelerometer signal consists of the actual acceleration plus a noise contribution and a constant due to sensor bias

$$\tilde{a}(k) = a(k) + \varepsilon(k) + c, \tag{C.1}$$

where $k$ denotes the discrete time. Velocity and displacement can be found from the acceleration measurements $\tilde{a}$ as follows,

$$\tilde{v}(k) = T \sum_{i=1}^{k} (a(i) + \varepsilon(i)) + kTc + v(0) \tag{C.2}$$

and

$$\begin{aligned} \tilde{z}(k) &= T \sum_{j=1}^{k} \left( T \sum_{i=1}^{j} (a(i) + \varepsilon(i)) + jTc + v(0) \right) + z(0) \\ &= T^2 \sum_{j=1}^{k} \sum_{i=1}^{j} a(i) + \eta(k) + \frac{k^2 + k}{2} T^2 c + kTv(0) + z(0) \end{aligned} \tag{C.3}$$

with $\eta(k)$ the noise sequence of $\tilde{z}$ due to the accelerometer's noise, being

$$\begin{aligned} \eta(k) &= T^2([\varepsilon(1)] + [\varepsilon(1) + \varepsilon(2)] + \ldots + [\varepsilon(1) + \ldots + \varepsilon(k)]) \\ & \quad T^2(k\varepsilon(1) + (k-1)\varepsilon(2) + \ldots + \varepsilon(k)). \end{aligned} \tag{C.4}$$

with a variance as function of time

$$\begin{aligned}\sigma_\eta^2(k) &= E\Big(\eta(k)^2\Big)\\ &= E\Big(T^4(k\varepsilon(1)+(k-1)\varepsilon(2)+\ldots+\varepsilon(k))\Big)\\ &= E\Big(T^4\Big(k^2\varepsilon^2(1)k+2k(k-1)\varepsilon(1)\varepsilon(2)+(k-1)^2\varepsilon^2(2)+\ldots+\varepsilon^2(k)\Big)\Big).\end{aligned} \quad \text{(C.5)}$$

If the sensor noise $\varepsilon$ is Gaussian distributed with standard deviation $\sigma_\varepsilon$ and if it is white so that there is no correlation between the noise samples, Eq. (C.5) becomes

$$\sigma_\eta^2(k) = T^4\Big(k^2\sigma_\varepsilon^2+(k-1)^2\sigma_\varepsilon^2+\ldots+\sigma_\varepsilon^2\Big) = T^4\sigma_\varepsilon^2\left(\frac{k^3}{3}+\frac{k^2}{2}+\frac{k}{6}\right), \quad \text{(C.6)}$$

in which $k^3/3+k^2/2+k/6$ can be recognised as the 'square pyramidal number' [44].

# Appendix D

# Proper orthogonal decomposition

This appendix introduces the Proper Orthogonal Decomposition (POD). The POD is used in the Sections 3.4, 4.1 and 6.3. The POD of a matrix consists a set of vectors, the Proper Orthogonal Modes (POMs), and a set of scalars, the Proper Orthogonal Values (POVs). The POMs corresponding to the highest POVs describe the most prominent patterns in the columns of the initial matrix.

Assume a $N_\mathrm{P}$-DoFs system of which snapshots have been obtained. The snapshot at the $i$th time-instant is stored in a vector $\boldsymbol{w}_i$. If we perform multiple measurements, a matrix $\boldsymbol{W}$ can be combined from the snapshots:

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{w}_1 & \cdots & \boldsymbol{w}_{N_\mathrm{D}} \end{bmatrix} = \left[ \begin{Bmatrix} w_{1,1} \\ \vdots \\ w_{1,N_\mathrm{P}} \end{Bmatrix} \cdots \begin{Bmatrix} w_{N_\mathrm{D},1} \\ \vdots \\ w_{N_\mathrm{D},N_\mathrm{P}} \end{Bmatrix} \right]. \tag{D.1}$$

Figure D.1 shows a point cloud of snapshots of a system with two DoFs, $w_1$ and $w_2$.

Each column of snapshot matrix $\boldsymbol{W}$ can be imagined as a point in a $N_\mathrm{P}$-dimensional space. The full snapshot matrix can then be thought of as a point cloud consisting of $N_\mathrm{D}$ points. The POD of a snapshot matrix consists of a set of orthogonal vectors and their corresponding singular values. The POD can be obtained in the following way.

The vector with the highest singular value spans the direction in the $N_\mathrm{P}$ space that minimises – in a least squares sense – its total distance to the points in the point cloud:

$$\arg\min_{\breve{\boldsymbol{\varphi}}_1} \sum_{i=1}^{N_\mathrm{D}} m_i^2 = \sum_{i=1}^{N_\mathrm{D}} \left[ |\boldsymbol{w}_i|^2 - \left( \frac{\boldsymbol{w}_i^\mathsf{T} \breve{\boldsymbol{\varphi}}_1}{|\breve{\boldsymbol{\varphi}}_1|} \right)^2 \right], \tag{D.2}$$

where $m_i$ is the distance between the $i$th point of the snapshot matrix and the direction span by $\breve{\boldsymbol{\varphi}}_1$. Defining $l_i$ as the length of the orthogonal projection of $\boldsymbol{w}_i$ onto $\breve{\boldsymbol{\varphi}}_1$, the minimisation problem of Eq. (D.2) is equivalent to the maximisation

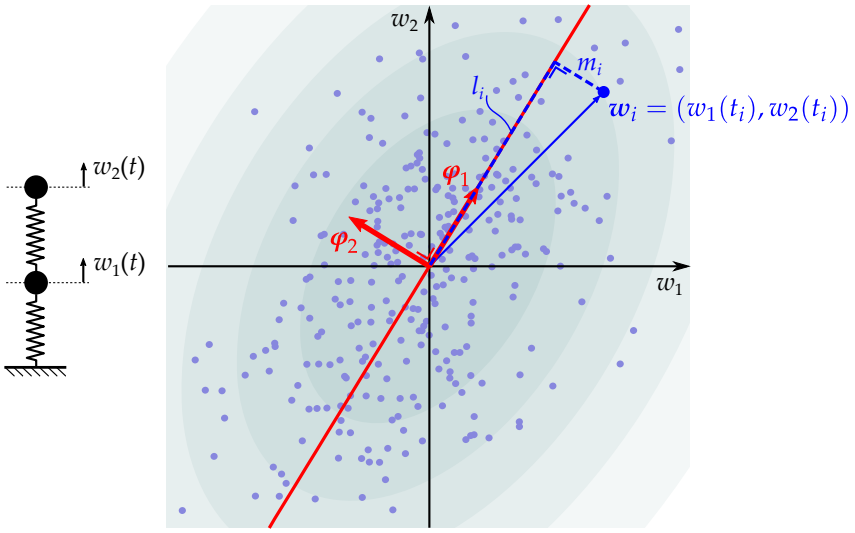



Figure D.1: Point cloud of snapshots of a two DoFs system ($w_1, w_2$). The distance between a snapshot $\boldsymbol{w}_i$ and vector $\breve{\boldsymbol{\varphi}}_1$ is called $m_i$ and the length of the projection of $\boldsymbol{w}_i$ onto direction $\breve{\boldsymbol{\varphi}}_1$ is called $l_i$. The first POM, $\breve{\boldsymbol{\varphi}}_1$, minimises the sum of all $m_i$ in a squared sense, or, equivalently, maximises the sum of all $l_i$ in a squared sense. Thus, $\breve{\boldsymbol{\varphi}}_1$ spans the most prominent direction of the point cloud. The second POM, $\breve{\boldsymbol{\varphi}}_2$, optimises for the same criterion, under the condition that it is orthogonal to the first POM.

problem

$$\arg\max_{\breve{\boldsymbol{\varphi}}_1} \sum_{i=1}^{N_\mathrm{D}} l_i^2 = \arg\max_{\breve{\boldsymbol{\varphi}}_1} \sum_{i=1}^{N_\mathrm{D}} \left( \frac{\boldsymbol{w}_i^\mathsf{T} \breve{\boldsymbol{\varphi}}_1}{|\breve{\boldsymbol{\varphi}}_1|} \right)^2, \tag{D.3}$$

as $l_i^2 + m_i^2 = |\boldsymbol{w}_i|^2$ and thus $\sum_i l_i^2 + \sum_i m_i^2 = \sum_i |\boldsymbol{w}_i|^2$.

Length $l_i$ can be expressed as

$$l_i^2 = \frac{\left(\boldsymbol{w}_i^\mathsf{T} \breve{\boldsymbol{\varphi}}_1\right)^2}{|\breve{\boldsymbol{\varphi}}_1|^2} = \frac{\breve{\boldsymbol{\varphi}}_1^\mathsf{T} \boldsymbol{w}_i \boldsymbol{w}_i^\mathsf{T} \breve{\boldsymbol{\varphi}}_1}{\breve{\boldsymbol{\varphi}}_1^\mathsf{T} \breve{\boldsymbol{\varphi}}_1}. \tag{D.4}$$

Let $\lambda_1 = \sum_{i=1}^{N_\mathrm{D}} l_i^2$ be the sum of the squared lengths, so that

$$\lambda_1 = \sum_{i=1}^{N_\mathrm{D}} \frac{\breve{\boldsymbol{\varphi}}_1^\mathsf{T} \boldsymbol{w}_i \boldsymbol{w}_i^\mathsf{T} \breve{\boldsymbol{\varphi}}_1}{\breve{\boldsymbol{\varphi}}_1^\mathsf{T} \breve{\boldsymbol{\varphi}}_1}, \tag{D.5}$$

which can be rewritten as follows

$$\lambda_1 \breve{\boldsymbol{\varphi}}_1^\mathsf{T} \breve{\boldsymbol{\varphi}}_1 = \sum_{i=1}^{N_\mathrm{D}} \breve{\boldsymbol{\varphi}}_1^\mathsf{T} \boldsymbol{w}_i \boldsymbol{w}_i^\mathsf{T} \breve{\boldsymbol{\varphi}}_1 \tag{D.6}$$

$$= \breve{\boldsymbol{\varphi}}_1^\mathsf{T} \sum_{i=1}^{N_\mathrm{D}} \boldsymbol{w}_i \boldsymbol{w}_i^\mathsf{T} \breve{\boldsymbol{\varphi}}_1, \tag{D.7}$$

$$0 = \check{\boldsymbol{\varphi}}_1^{\mathsf{T}} \left( \sum_{i=1}^{N_\mathrm{D}} \boldsymbol{w}_i \boldsymbol{w}_i^{\mathsf{T}} - \lambda_1 I \right) \check{\boldsymbol{\varphi}}_1. \tag{D.8}$$

For the nontrivial case that $\check{\boldsymbol{\varphi}}_1 \neq \mathbf{0}$

$$\left( \sum_{i=1}^{N_\mathrm{D}} \boldsymbol{w}_i \boldsymbol{w}_i^{\mathsf{T}} - \lambda_1 I \right) \check{\boldsymbol{\varphi}}_1 = \mathbf{0}. \tag{D.9}$$

Note that this result is equivalent to $\left( \sum_{i=1}^{N_\mathrm{D}} \boldsymbol{w}_i \boldsymbol{w}_i^{\mathsf{T}} - \lambda_1 I \right) \check{\boldsymbol{\varphi}}_1 \perp \check{\boldsymbol{\varphi}}_1$.

Eq. (D.9) can be rewritten as

$$\lambda_1 \check{\boldsymbol{\varphi}}_1 = \sum_{i=1}^{N_\mathrm{D}} \boldsymbol{w}_i \boldsymbol{w}_i^{\mathsf{T}} \check{\boldsymbol{\varphi}}_1 \tag{D.10}$$

$$= \boldsymbol{W}\boldsymbol{W}^{\mathsf{T}} \check{\boldsymbol{\varphi}}_1, \tag{D.11}$$

where $\boldsymbol{W}\boldsymbol{W}^{\mathsf{T}}$ can be recognised as a matrix proportional to the empirical sample correlation matrix of snapshot matrix $\boldsymbol{W}$. This equation has the form of a standard eigenvalue problem. As the aim is to maximise $\lambda_1$, the first POM $\check{\boldsymbol{\varphi}}_1$ is selected as the eigenvector of $\boldsymbol{W}\boldsymbol{W}^{\mathsf{T}}$ with the largest corresponding eigenvalue $\lambda_1$.

As $\boldsymbol{W}\boldsymbol{W}^{\mathsf{T}}$ is per definition a Hermitian matrix (i.e. equal to its conjugate transpose), the eigenvectors will be orthogonal. Only in case of degenerate eigenvalues, eigenvectors may not be orthogonal, but in that case it is always possible to construct an orthogonal set of eigenvectors from a linear combination of the degenerate eigenvectors.

When POM $\check{\boldsymbol{\varphi}}_1$ is found, the second POM $\check{\boldsymbol{\varphi}}_2$, *orthogonal to* $\check{\boldsymbol{\varphi}}_1$ needs to be found. As the the eigenvectors of $\boldsymbol{W}\boldsymbol{W}^{\mathsf{T}}$ are orthogonal $\check{\boldsymbol{\varphi}}_2$ is the eigenvector of $\boldsymbol{W}\boldsymbol{W}^{\mathsf{T}}$ corresponding to the *one-but-largest* eigenvalue $\lambda_2$. In the same way the full set of POMs is found from the eigenvectors and eigenvalues of $\boldsymbol{W}\boldsymbol{W}^{\mathsf{T}}$.

# Appendix E

# Comparison of commercial accelerometers

This appendix corresponds to the Sections 2.4.2 and 7.5. In the following a comparison of high-performance accelerometers is presented, based on their noise floor, bandwidth and mass, as supplied by the manufacturers.

The fields in which the accelerometers are applicable range from dynamics and motion control to floor vibrations and seismic measurements. To put the data in perspective also a few velocity sensors are included, along with floor vibration criteria and floor vibration measurements at a granite table in the relatively noisy Mechatronics laboratory at Delft University of Technology.

## E.1 Accelerometers

Figure E.1 compares 16 accelerometers based on their noise level. The figure plots the amplitude density of the noise for the sensor's measurement bandwidth. The mass of the accelerometers is indicated behind the names of the sensors.

The accelerometers in this comparison include capacitive MEMS, force feedback and piezoelectric accelerometers. The RSensors MTSS-1041A accelerometer makes use of molecular electronic transfer: a gas moves, as a result of the inertia forces, through channels in the sensor, thereby transporting charge between electrodes [98].

All data is obtained from the datasheets supplied by the sensors' manufacturers. For capacitive MEMS and force feedback sensors, the noise level is in general specified by a single value. A flat noise as function of frequency was assumed for these sensors. The noise level of piezoelectric accelerometers is in most cases specified at a few frequencies. The noise amplitude density decreases typically by a factor of 10 if the frequency increases by a factor of 100; if only a single noise specification was supplied for a piezoelectric accelerometer, a noise level according to this trend was assumed.

In most cases, multiple frequency ranges are specified for a piezoelectric

accelerometer. These correspond to the amount of tolerated deviation from the sensor's nominal sensitivity. The bandwidth corresponding to 5 % – 6 % (0.5 dB) deviation is indicated with squares connected by a solid line, 10 % – 11 % (1 dB) deviation with a cross and a solid line, and 30 % (3 dB) deviation by a dotted line without markings.

In the plot the general tendency can be observed that accelerometers with a lower noise level have higher mass.

Two velocity sensors are included in the plot, the Trillium compact 20 s and the low-cost geophone GS-11D. Their noise levels were converted to equivalents in terms of acceleration. The noise level of the GS-11D (which was most likely limited by the noise level of the NI6211 data acquisition box) was obtained from measurements.

The noise level of the geophone is between 0.2 Hz and 9 Hz better than the relatively expensive accelerometer EpiSensor ES-U2, showing that the geophone is a cost-effective solution for floor vibration measurements.

## E.2 Floor vibration criteria and measurements

Figure E.1 also provides measurement results regarding the vertical floor vibrations at a granite table in the Mechatronics laboratory (measurement 'Granite table (a)' and 'Granite table (b)') and the noise level at the top of the Minus-K vibration isolator[1]. The figure furthermore includes the ISO guidelines and Vibration Criteria [99, 100].

The measurements at the granite table and the Minus-K vibration isolator were obtained using a geophone. The measurement can be trusted up to about 150 Hz, from that frequency, the noise level of the geophone becomes dominant (measurement 'Granite table (a)').

The floor noise is highest around 8 Hz. At that frequency, the noise level is as high as the ISO specification for office floor noise. At the top plate of the Minus-K vibration isolator, however, the noise level is reduced to below level VC-D of the vibration criteria.

[1] Appendix G.2 provides additional information on the Minus-K vibration isolator.

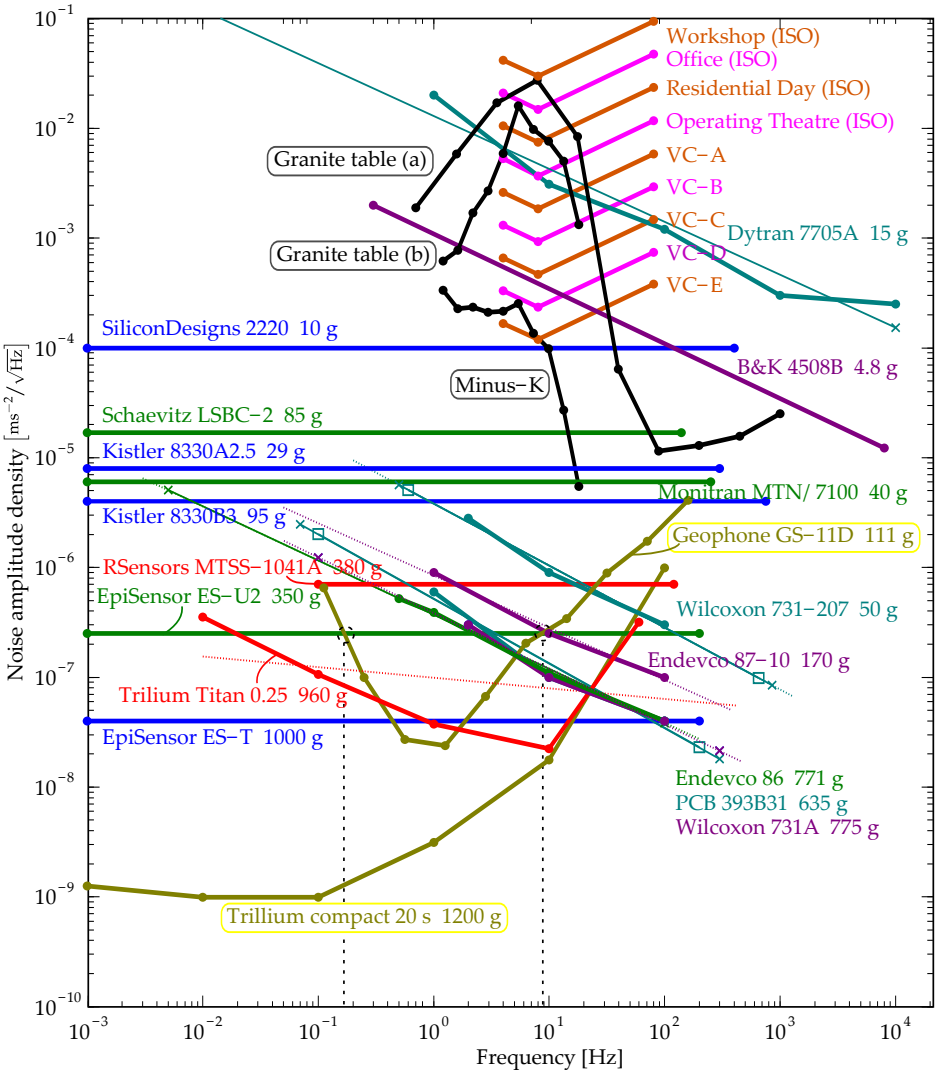



Figure E.1: The noise level in terms of noise amplitude density for some high-performance commercial accelerometers. For creating this overview, the specifications supplied by the manufacturers were used. If the noise-level was specification by a single value, the noise level was assumed to be comparable to other sensors of the same working principle. For comparison, the noise level in terms of acceleration of two velocity sensors were included in the overview (the Trillium compact 20 s and the GS-11D geophone, the latter from measurement). As a reference, the floor noise level at a granite table in the Mechatronics laboratory at Delft University of Technology, together with floor noise requirements for general and laboratory environments according to the ISO guidelines and Vibration Criteria [99,100] were included.

# Appendix F

# Kalman filtering

This appendix corresponds to Section 3.3. The Kalman filter is introduced and compared to the shape fitting estimator.

Consider a mechanical system that is described in state-space system as follows:

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{u}(k) + \boldsymbol{E}\boldsymbol{b}(k) + \boldsymbol{w}(k) \\ \boldsymbol{y}(k) &= \boldsymbol{C}\boldsymbol{x}(k) + \boldsymbol{D}\boldsymbol{u}(k) + \boldsymbol{G}\boldsymbol{b}(k) + \boldsymbol{v}(k). \end{aligned} \tag{F.1}$$

This system is represented by the block diagram that is highlighted by the blue rectangle in Figure F.1. $\boldsymbol{w}$ are the force disturbances and $\boldsymbol{v}$ is the measurement noise. $\boldsymbol{u}$ are the unknown input forces, i.e. the disturbances. $\boldsymbol{b}$ are the known input forces, which could be the feed-forward forces, but not the feedback forces, as those are internal and captured in system matrix $\boldsymbol{A}$.

A Kalman filter [46] can be constructed for this system as follows:

$$\begin{aligned} \tilde{\boldsymbol{x}}(k+1) &= \boldsymbol{A}\tilde{\boldsymbol{x}}(k) + \boldsymbol{E}\boldsymbol{b}(k) + \boldsymbol{K}\boldsymbol{e}(k) \\ \tilde{\boldsymbol{y}}(k) &= \boldsymbol{C}\tilde{\boldsymbol{x}}(k) + \boldsymbol{G}\boldsymbol{b}(k). \end{aligned} \tag{F.2}$$

The Kalman filter is highlighted by the red rectangle in Figure F.1. Error $\boldsymbol{e}$ is the difference between the actual output of the system, $\boldsymbol{y}$, and the Kalman estimate, $\tilde{\boldsymbol{y}}$, and is caused by errors $\boldsymbol{w}$ and $\boldsymbol{v}$ and the possible difference between the initial state of the actual system, $\boldsymbol{x}(0)$ and the initial state of the Kalman filter, $\tilde{\boldsymbol{x}}(0)$. Next to that, it may also contain error due to discrepancy between the system matrices of the actual system and the Kalman filter.

The Kalman filter aims at estimating the system's states $\boldsymbol{x}$ optimally – in a least squares sense –, given the system's known inputs $\boldsymbol{b}$ and measured outputs $\boldsymbol{y}$ [101]. This is achieved by choosing a Kalman gain $\boldsymbol{K}$ that takes into account the uncertainty of the system's states $\boldsymbol{w}$ and output $\boldsymbol{y}$. Often, it is assumed that the level of these uncertainties have reached steady-state. In that case a constant Kalman gain can be used. If there is no correlation between the disturbances $\boldsymbol{w}$ and the measurement noise $\boldsymbol{v}$, the Kalman gain is calculated from

$$\boldsymbol{K} = \boldsymbol{A}\boldsymbol{\Sigma}_x\boldsymbol{C}^\mathsf{T}(\boldsymbol{C}\boldsymbol{\Sigma}_x\boldsymbol{C}^\mathsf{T} + \boldsymbol{\Sigma}_v)^{-1}, \tag{F.3}$$

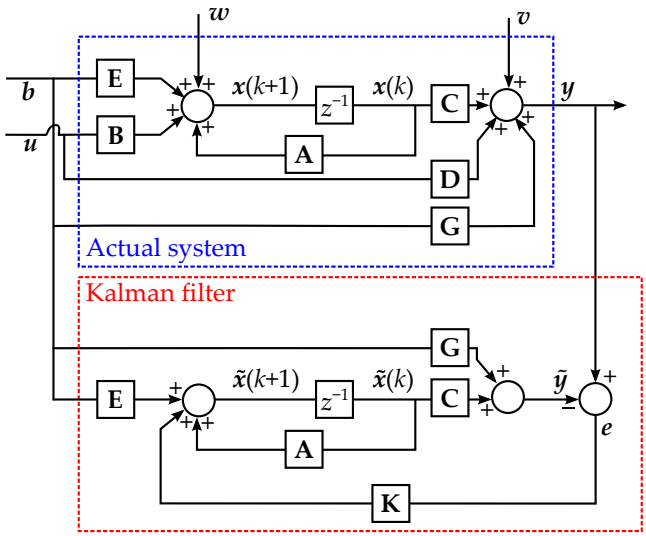



Figure F.1: Block diagram of a system and the corresponding Kalman filter. The Kalman filter estimates the system's state vector, $\boldsymbol{x}$, using the known system inputs, $\boldsymbol{b}$, and measured outputs, $\boldsymbol{y}$.

in which $\boldsymbol{\Sigma}_v$ is the covariance matrix of the measurement noise and $\boldsymbol{\Sigma}_x$ the covariance matrix of the state vector, which is found by solving the discrete time Ricatti equation.

The Kalman gain of Eq. (F.3) shows close resemblance to the SF estimation matrix $\boldsymbol{B}_{\mathrm{LS}}$ as obtained via the regularised LS technique (Eq. 4.16). From the measurements $\tilde{\boldsymbol{w}}_{\mathrm{m}}$, the SF estimator that uses the LS technique yields the following shape amplitudes:

$$\hat{\boldsymbol{\alpha}} = \boldsymbol{\Sigma}_\alpha \boldsymbol{W}_{\mathrm{m}}^{\mathsf{T}} \left( \boldsymbol{W}_{\mathrm{m}} \boldsymbol{\Sigma}_\alpha \boldsymbol{W}_{\mathrm{m}}^{\mathsf{T}} + \boldsymbol{\Sigma}_\varepsilon \right)^{-1} \tilde{\boldsymbol{w}}_{\mathrm{m}}. \tag{F.4}$$

Indeed, the Kalman filter is, under certain conditions, equivalent to the SF estimator. This was previously described by, for example, Kailath [101] and, more specifically in the application field of shape fitting by Andersson et al. [11].

# Appendix G

# Selection and assessment of the Minus-K vibration isolator

To obtain a low noise level in the experimental setup, it was chosen to suspend the metrology frame using a vibration isolation system (Section 7.1.3). This appendix compares some commercially available vibration isolators and evaluates the transmissibility of the selected Minus-K vibration isolator.

## G.1 Comparison of vibration isolators

Table G.1 provides an overview of some commercially available table-top vibration isolation systems. Both active systems, i.e. using sensors and actuators, and passive systems are listed. The resonance frequency $f_1$ and the transmissibility are obtained from the specifications supplied by the manufacturers.

Based on this overview, the Minus-K BM-4 vibration isolator was selected for use in the experimental setup.[1] This isolator is fully passive and thus avoids possible vibrations induced by sensor noise in an active system. Of the reviewed systems, the Minus-K obtains the lowest eigenfrequency and lowest transmissibility in vertical direction. It is also a cost-effective solution compared to the active systems and does not need supply of air or gas.

## G.2 Experimental evaluation of the Minus-K vibration isolator

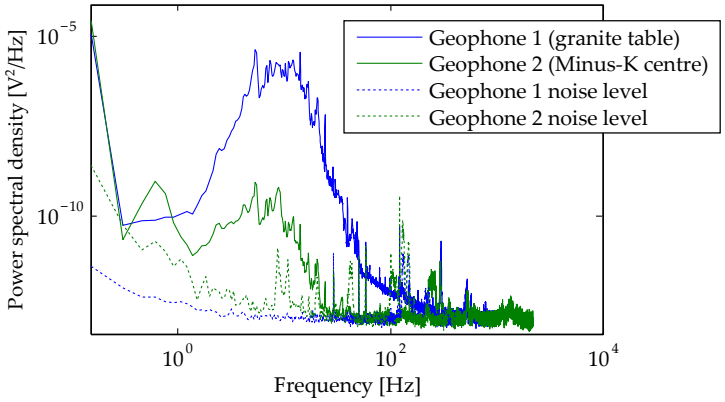
The vertical transmissibility of the Minus-K vibration isolation system was evaluated experimentally. The measurement results showed that the vibration isolation at the centre of the isolator is according to the manufacturer's specification, but also that the vibrations at the corner of the top plate are suppressed

[1] The choice for the model (Minus-K 200BM-4), which determines the payload range, is further substantiated in Section 7.1.4.

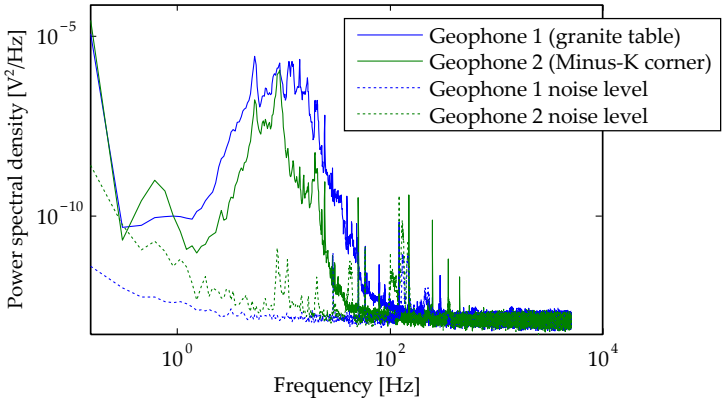
Table G.1: Specifications of some commercially available table-top vibration isolators.

| Model | Passive/ active* | $f_1^\dagger$ [Hz] | Transmissibility† [1] | | |
|---|---|---|---|---|---|
| | | | at $f_1$ | at 1 Hz | at 10 Hz |
| Table Stable 150 | Active<br>(piezo) | 1 Hz<br>1 Hz? | 1<br>1 | 1<br>1–1.4 | 0.01<br>0.02 |
| TMC TableTop PZT | Active<br>(piezo) | n/a<br>n/a | < 1<br>< 1 | 0.3<br>0.3 | 0.03<br>0.03 |
| Minus-K BM-4 | Passive | 0.5 Hz<br>0.5 Hz‡ | 2<br>3 | 0.2<br>1 | 0.003<br>0.06 |
| TMC 66 Series TableTop CSP | Passive<br>(air supply) | 3 Hz<br>2.3 Hz | 10<br>3 | 1<br>1 | 0.1<br>0.1 |
| Thorlabs PTT600600 | Passive | 3.5 Hz<br>n/a | 16<br>3 | 1<br>1 | 0.15<br>0.15 |
| Newport BenchTop Compact | Passive (air supply) | 3.2 Hz<br>3.6 Hz | 3.1<br>3 | 1<br>1 | 0.09<br>0.09 |

*Air supply: Pressurised air or gas supply is required; Piezo: Piezoelectrically actuated.
† The first line indicates the resonance frequency ($f_1$) or transmissibility for the vertical motion direction, the second line the one for the horizontal motion direction. n/a: specification was not available.
‡ This resonance frequency is payload-dependent.




(a) First experiment, geophone 2 is placed at the centre of the isolator's top plate.




(b) Second experiment, geophone 2 is placed at the front-left corner of the isolator's top plate.

Figure G.1: The power spectral density of the recorded geophone signals in response to floor vibrations. Up to 30 Hz, the signals are well above the geophones' noise level.

much less effectively. The latter can be explained by the mechanical construction of the isolator.

**Method** A geophone GS-11D was place at the granite table, at about 10 mm distance from the front-left side of the Minus-K vibration isolation system. A second geophone was placed at the centre of the isolator's top plate. No external excitation was applied, except for the always present ground vibration. The voltage output of the geophones were recorded for 120 s using a NI611 data acquisition box. The procedure was repeated after placing the geophone to the front-left corner of the isolator's top plate.

**Results** Figure G.1 compares the Power Spectral Density (PSD) of the recorded geophone signals to the measurements of the noise level of the geophones. The

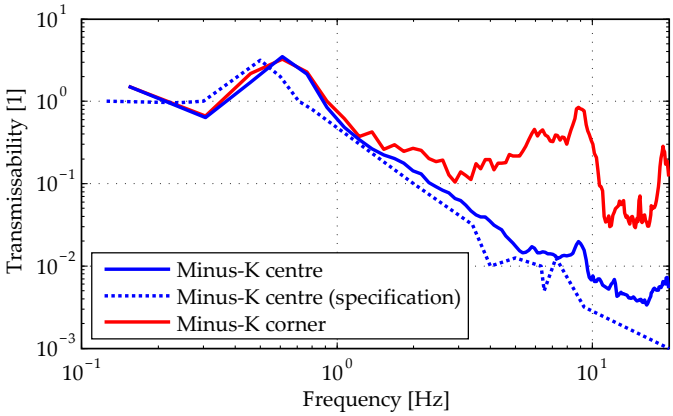



Figure G.2: The transmissibility of the Minus-K vibration isolator. The measured transmissibility from the base of isolator to the centre of its top plate is closely comparable to the specifications of the manufacturer. The transmissibility to the corner of the top plate is much higher than specified, however.

transmissibility of the Minus-K isolator can be estimated by dividing the PSD of the geophone placed on the Minus-K by the PSD of the geophone at the granite table next to the base of the Minus-K and taking the square root of the result. Both should – in terms of power – be at least a factor of five above the noise floor of the geophones to yield a valid result. For the measurements at the centre this is the case between 1.2 Hz and 18 Hz; for the measurements at the corner between 0.5 Hz and 30 Hz. At the lowest frequency point the PSD is relatively high, the reason for this drift is unknown.

Figure G.2 shows the measured transmissibility from the isolator's base to the centre and corner of its top plate. The transmissibility of the centre shows good resemblance with the manufacturer's specifications. The figure, however, also shows that the Minus-K suppresses the floor vibrations not as well at the corner as at the centre.

The lower performance at the corner of the isolator's table can be explained by the isolator's mechanical construction. The table top is connected to the isolator by a rubber pad, to isolate the vibrations in out-of-plane rotation direction. The resulting eigenfrequency is still relatively high – around 8 Hz. The eigenfrequency could be decreased by making the size of the pads smaller, thereby reducing the stiffness. The vibration isolator would, however, become more sensitive to the position of the mass at its table top. Alternatively, the rotational moment of inertia of the table could be increased by redistributing the payload mass more outward from the centre.

# Appendix H

# Corrected plate finite element matrices

The finite element models used for static and dynamic mechanical analyses throughout this thesis make use of the plate finite element of Kerboua et al. [29]. This finite element is introduced in Section 2.3 of this thesis. Kerboua et al. provided the matrices that are needed to implement the finite element, but those were found to contain errors. This appendix provides the correct replacements for the erroneous terms. The terms were derived using symbolic computation, following the approach described by the original authors.

## H.1 Matrix $R$

Matrix $\boldsymbol{R}$ is correct.

## H.2 Matrix $Q$

The following lists only the erroneous entries of the original matrix $\boldsymbol{Q}$, each followed by its correct replacement.

$Q_{4,15} = \frac{x}{A^3}$; must be $-\frac{x}{A^3}$.
$Q_{5,22} = -\frac{x^3}{2A^3B^2}$; $-\frac{x^3}{6A^3B^2}$.
$Q_{5,23} = -\frac{x^2y}{6A^2B^3}$; $-\frac{x^2y}{2A^2B^3}$.
$Q_{5,24} = -\frac{x^3y}{6A^{2B^3}B^2}$ *(sic)*; $-\frac{x^3y}{6A^3B^3}$.
$Q_{6,17} = -\frac{2y}{A^2B}$; $-\frac{2y}{AB^2}$.

## H.3 Matrix $A^{-1}$

The following lists only the erroneous entries of the original matrix $\boldsymbol{A}^{-1}$, each followed by its correct replacement.

$(A^{-1})_{2,7} = -\frac{A}{x_e}$; must be $\frac{A}{x_e}$.

$(A^{-1})_{4,1} = -\frac{AB}{x_e y_e}$; $\frac{AB}{x_e y_e}$.

$(A^{-1})_{4,7} = \frac{AB}{x_e y_e}$; $-\frac{AB}{x_e y_e}$.

$(A^{-1})_{4,13} = -\frac{AB}{x_e y_e}$; $\frac{AB}{x_e y_e}$.

$(A^{-1})_{4,19} = \frac{AB}{x_e y_e}$; $-\frac{AB}{x_e y_e}$.

$(A^{-1})_{7,20} = -\frac{AB}{x_e y_e}$; $\frac{B}{y_e}$.

$(A^{-1})_{8,2} = \frac{AB}{x_e}$*; $\frac{AB}{x_e y_e}$.

$(A^{-1})_{8,8} = -\frac{AB}{x_e}$*; $-\frac{AB}{x_e y_e}$.

$(A^{-1})_{8,14} = \frac{AB}{x_e}$*; $\frac{AB}{x_e y_e}$.

$(A^{-1})_{8,20} = -\frac{AB}{x_e}$*; $-\frac{AB}{x_e y_e}$.

$(A^{-1})_{14,3} = -\frac{12A^3}{x_e^2}$*; $-\frac{6B^2}{y_e^2}$.

$(A^{-1})_{14,3} = \frac{2B^2}{y_e}$* (*sic*, defined a second time differently); $-\frac{6B^2}{y_e^2}$.

$(A^{-1})_{14,5} = \frac{2B^2}{y_e}$; $-\frac{4B^2}{y_e}$.

$(A^{-1})_{14,21} = \frac{12A^3}{x_e^2}$*; $\frac{6B^2}{y_e^2}$.

$(A^{-1})_{14,23} = \frac{B^2}{y_e}$; $-\frac{2B^2}{y_e}$.

$(A^{-1})_{16,6} = -\frac{2A^2B}{x_e}$; $-\frac{4A^2B}{x_e}$.

$(A^{-1})_{17,24} = \frac{4AB^2}{y_e}$; $-\frac{2AB^2}{y_e}$.

$(A^{-1})_{20,3} = 0$; $\frac{36A^2B^2}{x_e^2 y_e^2}$.

$(A^{-1})_{20,9} = -\frac{36A^2B^2}{x_e^2 y_e}$*; $-\frac{36A^2B^2}{x_e^2 y_e^2}$.

$(A^{-1})_{20,15} = 0$; $\frac{36A^2B^2}{x_e^2 y_e^2}$.

$(A^{-1})_{20,16} = -\frac{24A^2B^2}{x_e y_e^2}$; $-\frac{12A^2B^2}{x_e y_e^2}$.

$(A^{-1})_{20,21} = 0$; $-\frac{36A^2B^2}{x_e^2 y_e^2}$.

$(A^{-1})_{20,24} = \frac{24A^2B^2}{x_e y_e}$; $\frac{8A^2B^2}{x_e y_e}$.

$(A^{-1})_{22,3} = \frac{72AB^2}{x_e^3 y_e^2}$* $-\frac{72A^3B^2}{x_e^3 y_e^2}$.

$(A^{-1})_{22,4} = -\frac{36A^3B^3}{x_e^2 y_e^2}$**; $-\frac{36A^3B^2}{x_e^2 y_e^2}$.

$(A^{-1})_{22,9} = -\frac{72AB^2}{x_e^3 y_e^2}$*; $\frac{72A^3B^2}{x_e^3 y_e^2}$.

$(A^{-1})_{22,10} = -\frac{36A^3B^3}{x_e^2 y_e^2}$**; $-\frac{36A^3B^2}{x_e^2 y_e^2}$.

$(A^{-1})_{22,15} = \frac{72AB^2}{x_e^3 y_e^2}$*; $-\frac{72A^3B^2}{x_e^3 y_e^2}$.

$(A^{-1})_{22,16} = \frac{36A^3B^3}{x_e^2 y_e^2}$**; $\frac{36A^3B^2}{x_e^2 y_e^2}$.

$(A^{-1})_{22,21} = -\frac{72AB^2}{x_e^3 y_e^2}$*; $\frac{72A^3B^2}{x_e^3 y_e^2}$.

$(A^{-1})_{22,22} = \frac{36A^3B^3}{x_e^2 y_e^2}$**; $\frac{36A^3B^2}{x_e^2 y_e^2}$.

$(A^{-1})_{23,3} = \frac{72A^2B^3}{x_e^2 y_e^2}$*; $-\frac{72A^2B^3}{x_e^2 y_e^3}$.

$(A^{-1})_{23,9} = -\frac{72A^2B^3}{x_e^2 y_e^2}$*; $\frac{72A^2B^3}{x_e^2 y_e^3}$.

$(A^{-1})_{23,15} = \frac{72A^2B^3}{x_e^2 y_e^2}$*; $-\frac{72A^2B^3}{x_e^2 y_e^3}$.

$(A^{-1})_{23,21} = \frac{72A^2B^3}{x_e^2 y_e^2}$*; $\frac{72A^2B^3}{x_e^2 y_e^3}$.

$(A^{-1})_{24,5} = -\frac{72A^3B^3}{x_e^2 y_e^2}$**; $\frac{72A^3B^3}{x_e^3 y_e^2}$.

$(A^{-1})_{24,11} = \frac{72A^3B^3}{x_e^2 y_e^2}$**; $-\frac{72A^3B^3}{x_e^3 y_e^2}$.

$(A^{-1})_{24,17} = \frac{72A^3B^3}{x_e^2 y_e^2}$**; $-\frac{72A^3B^3}{x_e^3 y_e^2}$.

$(A^{-1})_{24,23} = -\frac{72A^3B^3}{x_e^2 y_e^2}$**; $\frac{72A^3B^3}{x_e^3 y_e^2}$.

* (should be dimensionless)

** (should have dimension L)

# Bibliography


[1] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, no. 8, 1965.

[2] G. E. Moore *et al.*, "Progress in digital integrated electronics," *IEDM Tech. Digest*, pp. 11 – 13, 1975.

[3] Cray Research, Inc., *The Cray-1 Computer System*, 1977.

[4] J. Pettinato and D. Pillai, "Technology decisions to minimize 450-mm wafer size transition risk," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 18, pp. 501 – 9, Nov. 2005.

[5] C.-F. Chien, J. Wang, T.-C. Chang, and W.-C. Wu, "Economic analysis of 450mm wafer migration," pp. 1 – 4, Oct. 2007.

[6] G. Kirchhoff, "Über das gleichgewicht und die bewegung einer elastischen scheibe," *Journal für die reine und angewandte Mathematik*, vol. 40, pp. 51 – 88, 1850.

[7] D. Laro, R. Boshuizen, D. Oom, L. Sanders, and J. van Eijk, "Lightweight 450 mm wafer stages enabled by over-actuation," in *Proceedings of the 10th international conference of the EUSPEN*, vol. 1, pp. 433 – 6, 2010.

[8] G. C. Kirby III, D. K. Lindner, M. A. Davis, and A. D. Kersey, "Optimal sensor layout for shape estimation from strain sensors," in *Proceedings of SPIE*, vol. 2444, pp. 367 – 76, SPIE, 1995.

[9] J. Ranieri, A. Vincenzi, A. Chebira, D. Atienza, and M. Vetterli, "Near-optimal thermal monitoring framework for many-core systems-on-chip," *IEEE Transactions on Computers*, vol. 64, pp. 3197 – 209, Nov. 2015.

[10] H. Lee and H. Park, "Estimation of deformed shapes of beam structures using 3d coordinate information from terrestrial laser scanning," *Computer Modeling in Engineering and Sciences*, vol. 29, no. 1, pp. 29 – 44, 2008.

[11] M. S. Andersson and E. F. Crawley, "Structural shape estimation using shaped sensors," in *Proceedings of the 36th Structures, Structural Dynamics and Materials Conference*, vol. 5, (New Orleans, LA, USA), pp. 3368 – 78, 1995.

[12] P. Mainçon, "Inverse fem i: Load and response estimates from measurements," in *Progress in Structural Engineering, Mechanics and Computation*, pp. 967 – 71, 2004.

[13] G. Lesueur, H. Gilles, S. Girard, T. Merlet, and M. Queguiner, "Optical sensor for real-time reconstruction of distortions on electronically steered antenna," *Photonics Technology Letters, IEEE*, vol. 20, no. 21, pp. 1763 – 5, 2008.

[14] A. Koevoets, J. van der Sanden, and T. Ruijl, "Thermal-elastic compensation models for position control," in *Proceedings of the ASPE Annual Meeting 2009*, 2009.

[15] J. Kullaa, "Sensor validation using minimum mean square error estimation," *Mechanical Systems and Signal Processing*, vol. 24, no. 5, pp. 1444 – 57, 2010. Special Issue: Operational Modal Analysis.

[16] A. Koevoets, H. Eggink, J. van der Sanden, J. Dekkers, and T. Ruijl, "Optimal sensor configuring techniques for the compensation of thermo-elastic deformations in high-precision systems," in *13th International Workshop on Thermal Investigation of ICs and Systems*, pp. 208 – 13, Sept. 2007.

[17] B. M. La Fontaine, J. Hauschild, M. V. Dusa, A. Acheta, E. M. Apelgren, M. Boonman, J. Krist, A. Khathuria, H. J. Levinson, A. Fumar-Pici, *et al.*, "Study of the influence of substrate topography on the focusing performance of advanced lithography scanners," in *SPIE Microlithography 2003*, pp. 570 – 81, SPIE, 2003.

[18] H. Butler, "Position control in lithographic equipment [applications of control]," *Control Systems, IEEE*, vol. 31, no. 5, pp. 28 – 47, 2011.

[19] R. H. Munnig Schmidt, "Ultra-precision engineering in lithographic exposure equipment for the semiconductor industry," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1973, pp. 3950 – 72, 2012.

[20] J. Mulkens, B. Streefkerk, M. Hoogendorp, R. Moerman, M. Leenders, F. de Jong, M. Stavenga, and H. Boom, "Immersion lithography exposure systems: today's capabilities and tomorrow's expectations," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 5754, pp. 710 – 24, 2005.

[21] X. Yang, "The squeezed film with normal oscillation," Master's thesis, Delft University of Technology, 2014.

[22] "Error budgetting of immersion lithography machines," tech. rep., ASML, 2013.

[23] J. Woodhouse, "Linear damping models for structural vibration," *Journal of Sound and Vibration*, vol. 215, no. 3, pp. 547 – 69, 1998.

[24] Y.-J. Rao, "In-fibre bragg grating sensors," *Measurement Science and Technology*, vol. 8, no. 4, pp. 355 – 75, 1997.

[25] E. Davison, "A method for simplifying linear dynamic systems," *Automatic Control, IEEE Transactions on*, vol. 11, pp. 93 – 101, Jan. 1966.

[26] S. Hurlebaus, L. Gaul, and J.-S. Wang, "An exact series solution for calculating the eigenfrequencies of orthotropic plates with completely free boundary," *Journal of Sound and Vibration*, vol. 244, no. 5, pp. 747 – 59, 2001.

[27] G. L. C. M. de Abreu, J. F. Ribeiro, and V. Steffen Jr, "Finite element modeling of a plate with localized piezoelectric sensors and actuators," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 26, no. 2, pp. 117 – 28, 2004.

[28] E. Charbonneau and A. Lakis, "Semi-analytical shape functions in the finite element analysis of rectangular plates," *Journal of Sound and Vibration*, vol. 242, no. 3, pp. 427 – 43, 2001.

[29] Y. Kerboua, A. Lakis, M. Thomas, and L. Marcouiller, "Hybrid method for vibration analysis of rectangular plates," *Nuclear Engineering and Design*, vol. 237, no. 8, pp. 791 – 801, 2007.

[30] H. van der Auweraer, P. Guillaume, P. Verboven, and S. Vanlanduit, "Application of a fast-stabilizing frequency domain parameter estimation method," *Journal of Dynamic Systems, Measurement, and Control*, vol. 123, no. 4, pp. 651 – 8, 2001.

[31] P. Verboven, P. Guillaume, B. Cauberghe, S. Vanlanduit, and E. Parloo, "A comparison of frequency-domain transfer function model estimator formulations for structural dynamics modelling," *Journal of Sound and Vibration*, vol. 279, pp. 775 – 98, 2005.

[32] M. H. Richardson, "Is it a mode shape, or an operating deflection shape?," *SV Sound and vibration*, vol. 31, no. 1, pp. 54 – 61, 1997.

[33] R. J. Allemang, "The modal assurance criterion: Twenty years of use and abuse," *SV Sound and vibration*, vol. 37, pp. 14 – 23, 2003.

[34] T. Castenmiller, F. Van de Mast, T. De Kort, C. Van de Vin, M. De Wit, R. Stegen, and S. Van Cleef, "Towards ultimate optical lithography with nxt: 1950i dual stage immersion platform," in *Optical Microlithography XXIII*, vol. 7640, SPIE, 2010.

[35] K. Chandrashekhara, *Theory of plates*. Universities press, 2001.

[36] R. Hibbeler, *Mechanics of Materials*. Prentice Hall, Upper Saddle River, 2005.

[37] J. Sirohi and I. Chopra, "Fundamental understanding of piezoelectric strain sensors," *Journal of Intelligent Material Systems and Structures*, vol. 11, no. 4, pp. 246 – 57, 2000.

[38] B. Lee, "Review of the present status of optical fiber sensors," *Optical Fiber Technology*, vol. 9, no. 2, pp. 57 – 79, 2003.

[39] L.-H. Kang, D.-K. Kim, and J.-H. Han, "Estimation of dynamic structural displacements using fiber bragg grating strain sensors," *Journal of Sound and Vibration*, vol. 305, no. 3, pp. 534 – 42, 2007.

[40] S. Rapp, L.-H. Kang, J.-H. Han, U. C. Mueller, and H. Baier, "Displacement field estimation for a two-dimensional structure using fiber bragg grating sensors," *Smart Materials and Structures*, vol. 18, no. 2, pp. 1 – 12, 2009.

[41] W. Habel, D. Hofmann, and B. Hillemeier, "Deformation measurements of mortars at early ages and of large concrete components on site by means of embedded fiber-optic microstrain sensors," *Cement and Concrete Composites*, vol. 19, no. 1, pp. 81 – 102, 1997.

[42] C. Collette, L. Fueyo-Roza, and M. Horodinca, "Prototype of a small low noise absolute displacement sensor," *IEEE Sensors Journal*, vol. 14, no. 1, pp. 91 – 5, 2014.

[43] Y. K. Thong, M. S. Woolfson, J. A. Crowe, B. R. Hayes-Gill, and R. E. Challis, "Dependence of inertial measurements of distance on accelerometer noise," *Measurement Science and Technology*, vol. 13, no. 8, pp. 1163 – 72, 2002.

[44] J. H. Conway and R. Guy, *The Book of Numbers*. Copernicus, New York, 1996.

[45] C. Felippa, K. Park, and M. Justino Filho, "The construction of free–free flexibility matrices as generalized stiffness inverses," *Computers & Structures*, vol. 68, no. 4, pp. 411 – 8, 1998.

[46] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Fluids Engineering*, vol. 83, no. 1, pp. 95 – 108, 1961.

[47] B. Friedland, *Control system design: an introduction to state-space methods*. Courier Corporation, 1986.

[48] M. Chierichetti and M. Ruzzene, "Dynamic displacement field reconstruction through a limited set of measurements: Application to plates," *Journal of Sound and Vibration*, vol. 331, no. 21, pp. 4713 – 28, 2012.

[49] A. Tessler and J. Spangler, "A variational principle for reconstruction of elastic deformations in shear deformable plates and shells," tech. rep., NASA, 2003.

[50] A. Tessler and J. Spangler, "Inverse fem for full-field reconstruction of elastic deformations in shear deformable plates and shells," in *Proceedings of Second European Workshop on Structural Health Monitoring*, pp. 83 – 90, 2004.

[51] M. Gherlone, P. Cerracchio, M. Mattone, M. D. Sciuva, and A. Tessler, "Shape sensing of 3d frame structures using an inverse finite element method," *International Journal of Solids and Structures*, vol. 49, no. 22, pp. 3100 – 12, 2012.

[52] M. Gherlone, P. Cerracchio, M. Mattone, M. D. Sciuva, and A. Tessler, "An inverse finite element method for beam shape sensing: theoretical framework and experimental validation," *Smart Materials and Structures*, vol. 23, no. 4, pp. 1 – 13, 2014.

[53] R. Jones, T. A. Berkoff, D. G. Bellemore, D. Early, J. S. Sirkis, M. A. Putnam, E. J. Friebele, and A. D. Kersey, "Cantilever plate deformation monitoring using wavelength division multiplexed fiber bragg grating sensors," in *Proceedings of SPIE* (H. D. R. Murphy Kent A., ed.), vol. 2718, (San Diego, CA, USA), pp. 258 – 68, 1996.

[54] R. T. Jones, D. G. Bellemore, T. A. Berkoff, J. S. Sirkis, M. A. Davis, M. A. Putnam, E. J. Friebele, and A. D. Kersey, "Determination of cantilever plate shapes using wavelength division multiplexed fiber bragg grating sensors and a least-squares strain-fitting algorithm," *Smart Materials and Structures*, vol. 7, no. 2, pp. 178 – 88, 1998.

[55] M. Sekouri, *Strain-based shape estimation for plate structures*. PhD thesis, L'École de technologie supérieure de Montréal, 2004.

[56] N.-S. Kim and N.-S. Cho, "Estimating deflection of a simple beam model using fiber optic bragg-grating sensors," *Experimental Mechanics*, vol. 44, no. 4, pp. 433 – 9, 2004.

[57] Y. Wang, N. Chen, B. Yun, and Y. Cui, "Use of fiber bragg grating sensors for determination of a simply supported rectangular plane plate deformation," *Photonics Technology Letters, IEEE*, vol. 19, pp. 1242 – 4, Aug. 2007.

[58] A. C. Pisoni, C. Santolini, D. E. Hauf, and S. Dubowsky, "Displacements in a vibrating body by strain gage measurements," in *Proceedings of the 13th International Conference on Modal Analysis*, pp. 119 – 25, SPIE, 1995.

[59] G. Foss and E. Haugse, "Using modal test results to develop strain to displacement transformations," in *1995 IMAC XIII – 13th International Modal Analysis Conference*, pp. 112 – 8, SPIE, 1995.

[60] C.-J. Li and A. G. Ulsoy, "High-precision measurement of tool-tip displacement using strain gauges in precision flexible line boring," *Mechanical Systems and Signal Processing*, vol. 13, no. 4, pp. 531 – 46, 1999.

[61] G. Lesueur, A. Marciniak, H. Gilles, S. Girard, M. Laroche, T. Merlet, and M. Queguiner, "Polarization optical sensor for dynamical deformation measurements on 1-d or 2-d mechanical structures," *Photonics Technology Letters, IEEE*, vol. 21, pp. 1311 – 3, Sept. 2009.

[62] G. Kessenich, S. Bhola, B. Pletner, W. Horth, and A. Hosoi, "Sub-nanometer broadband measurement of elastic displacements in optical metrology frames and other critical elements," in *Metrology, Inspection, and Process Control for Microlithography XXIII*, vol. 7272, SPIE, 2009.

[63] H.-I. Kim, L.-H. Kang, and J.-H. Han, "Shape estimation with distributed fiber bragg grating sensors for rotating structures," *Smart Materials and Structures*, vol. 20, no. 3, pp. 1 – 11, 2011.

[64] A. Derkevorkian, J. Alvarenga, S. F. Masri, H. Ryaciotaki-Boussalis, and W. L. Richards, "Computational studies of a strain-based deformation shape prediction algorithm for control and monitoring applications," in *Industrial and Commercial Applications of Smart Structures Technologies 2012*, vol. 8343, SPIE, Mar. 2012.

[65] C. Bert, "Relationship between fundamental natural frequency and maximum static deflection for various linear vibratory systems," *Journal of sound and vibration*, vol. 162, no. 3, pp. 547 – 57, 1993.

[66] P. B. Bogert, E. Haugse, and R. E. Gehrki, "Structural shape identification from experimental strains using a modal transformation technique," in *Proceedings of 44th AIAA/ASME/ASCE/AHS Structures. Structural Dynamics, and Materials Conference, AIAA*, vol. 1626, 2003.

[67] Weeks, "Shape determination and control for large space structures," tech. rep., NASA, 1981.

[68] C. J. Weeks, "Static shape determination and control for large space structures: I. the flexible beam," *Journal of Dynamic Systems, Measurement, and Control*, vol. 106, no. 4, pp. 261 – 6, 1984.

[69] M. A. Davis, A. D. Kersey, J. S. Sirkis, and E. J. Friebele, "Fiber optic bragg grating array for shape and vibration mode shape sensing," in *Smart Structures and Materials 1994: Smart Sensing, Processing, and Instrumentation*, vol. 2191, pp. 94 – 102, 1994.

[70] M. A. Davis, A. D. Kersey, J. Sirkis, and E. J. Friebele, "Shape and vibration mode sensing using a fiber optic bragg grating array," *Smart Materials and Structures*, vol. 5, no. 6, p. 759, 1996.

[71] S. Hakim and M. Fuchs, "Shape estimation of distorted flexible structures," *Structural Optimization*, vol. 12, no. 4, pp. 237 – 43, 1996.

[72] G.-M. Lee, "Prediction of strain responses from the measurements of displacement responses," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 1143 – 52, 2007.

[73] F. Dekking, C. Kraaikamp, H. Lopuhaä, and L. Meester, *A Modern Introduction to Probability and Statistics - Understanding Why and How*. Springer Verlag London Limited, 2005.

[74] B. Feeny and R. Kappagantu, "On the physical interpretation of proper orthogonal modes in vibrations," *Journal of Sound and Vibration*, vol. 211, no. 4, pp. 607 – 16, 1998.

[75] J. Ranieri, A. Chebira, and M. Vetterli, "Near-optimal sensor placement for linear inverse problems," *Signal Processing, IEEE Transactions on*, vol. 62, pp. 1135 – 46, March 2014.

[76] P. C. Hansen, "The truncated svd as a method for regularization," *BIT Numerical Mathematics*, vol. 27, no. 4, pp. 534 – 53, 1987.

[77] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter estimation and inverse problems*. Academic Press, 2013.

[78] J. van der Sanden and P. Philips, "Fem model based pod reduction to obtain optimal sensor locations for thermo-elastic error compensation," in *DSPE conference on precision mechatronics*, pp. 65 – 70, Sept. 2012.

[79] G. H. Golub and C. F. V. Loan, *Matrix Computations*. John Hopkins University Press, 3rd ed., 1996.

[80] J. G. Vogel, A. Tejada, J. W. Spronck, and R. H. Munnig Schmidt, "Estimation of the deformation of a plate using a limited number of sensors," in *Proceedings of the ASPE Spring Topical Meeting on Precision Control for Advanced Manufacturing Systems*, pp. 70–75, 2013.

[81] J. G. Vogel, J. W. Spronck, and R. H. Munnig Schmidt, "Estimating deformation of a free-floating wafer chuck," in *Proceedings of the 14th international conference of the EUSPEN*, 2014.

[82] F. Shaw and K. Srinivasan, "Bandwidth enhancement of position measurements using measured acceleration," *Mechanical Systems and Signal Processing*, vol. 4, no. 1, pp. 23 – 38, 1990.

[83] H. Liu and G. Pang, "Accelerometer for mobile robot positioning," *IEEE Transactions on Industry Applications*, vol. 37, pp. 812 – 19, May 2001.

[84] J. Kim, K. Kim, and H. Sohn, "Autonomous dynamic displacement estimation from data fusion of acceleration and intermittent displacement measurements," *Mechanical Systems and Signal Processing*, vol. 42, no. 1, pp. 194 – 205, 2014.

[85] M. Younis and K. Akkaya, "Strategies and techniques for node placement in wireless sensor networks: A survey," *Ad Hoc Networks*, vol. 6, no. 4, pp. 621 – 55, 2008.

[86] K. Worden and A. Burrows, "Optimal sensor placement for fault detection," *Engineering Structures*, vol. 23, no. 8, pp. 885 – 901, 2001.

[87] M. Papadopoulos and E. Garcia, "Sensor placement methodologies for dynamic testing," *AIAA Journal*, vol. 36, pp. 256 – 63, Feb. 1998.

[88] J. Penny, M. Friswell, and S. Garvey, "Automatic choice of measurement locations for dynamic testing," *AIAA journal*, vol. 32, no. 2, pp. 407 – 14, 1994.

[89] D. Kammer, "Sensor placement for on-orbit modal identification and correlation of large space structures," in *American Control Conference, 1990*, pp. 2984 – 90, IEEE, 1990.

[90] F. A. Tasker and C. Liu, "Variance-based sensor placement for modal identification of structures," *Journal of Guidance, Control, and Dynamics*, vol. 18, no. 3, pp. 627 – 30, 1995.

[91] J. Ranieri, A. Vincenzi, A. Chebira, D. Atienza, and M. Vetterli, "Eigenmaps: Algorithms for optimal thermal maps extraction and sensor placement on multicore processors," in *Design Automation Conference (DAC), 2012 49th ACM/EDAC/IEEE*, pp. 636 – 41, June 2012.

[92] D. Kammer and L. Yao, "Enhancement of on-orbit modal identification of large space structures through sensor placement," *Journal of Sound and Vibration*, vol. 171, no. 1, pp. 119 – 39, 1994.

[93] D.-S. Li, H.-N. Li, and C.-P. Fritzen, "A note on fast computation of effective independence through qr downdating for sensor placement," *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 1160 – 8, 2009.

[94] D. El Badawy, J. Ranieri, and M. Vetterli, "Near-optimal sensor placement for signals lying in a union of subspaces," in *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)*, pp. 880 – 4, Sept. 2014.

[95] S. S. Rao, *Mechanical vibrations. 2004.* Pearson Prentice Hall, Inc. NJ.

[96] A. Leissa, "The free vibration of rectangular plates," *Journal of Sound and Vibration*, vol. 31, no. 3, pp. 257 – 93, 1973.

[97] D. C. Kuzma, "Fluid inertia effects in squeeze films," *Applied Scientific Research*, vol. 18, no. 1, pp. 15 – 20, 1968.

[98] H. Huang, V. Agafonov, and H. Yu, "Molecular electric transducers as motion sensors: A review," *Sensors*, vol. 13, no. 4, pp. 4581 – 97, 2013.

[99] C. G. Gordon, "Generic criteria for vibration-sensitive equipment," in *Part of the SPIE Conference on Current Developments in Vibration Control for Optomechanical Systems*, pp. 22 – 33, SPIE, 1992.

[100] B. Bessason, C. Madshus, H. A. Frøystein, and H. Kolbjørnsen, "Vibration criteria for metrology laboratories," *Measurement Science and Technology*, vol. 10, no. 11, pp. 1009 – 14, 1999.

[101] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*. Prentice Hall Upper Saddle River, NJ, 2000.

# Mathematical notation

| | |
|---|---|
| $\Re(a)$ | Real part of $a$. |
| $\Im(a)$ | Imaginary part of $a$. |
| $\boldsymbol{A}$, $\boldsymbol{a}$, $a$ | Matrix $\boldsymbol{A}$, vector $\boldsymbol{a}$, scalar $a$. |
| $[\vdots\vdots\vdots]$ | Matrix composed of multiple matrices, vectors or scalars. |
| $\{\vdots\}$ | Column vector composed of multiple column vectors or scalars. |
| $a_{i,j}$ | The element in the $i$th row and the $j$th column of $\boldsymbol{A}$. |
| $\boldsymbol{A}^\mathsf{T}$, $\boldsymbol{a}^\mathsf{T}$ | Transpose of $\boldsymbol{A}$, (idem) $\boldsymbol{a}$. |
| $\bar{\boldsymbol{A}}$, $\bar{\boldsymbol{a}}$, $\bar{a}$ | Element-by-element complex conjugate of $\boldsymbol{A}$, (idem) $\boldsymbol{a}$, complex conjugate of $a$. |
| $\boldsymbol{A}^*$, $\boldsymbol{a}^*$ | Complex conjugate matrix of $\boldsymbol{A}$ , (idem) $\boldsymbol{a}$. (Thus, $\boldsymbol{A}^* = \bar{\boldsymbol{A}}^\mathsf{T}$, $\boldsymbol{a}^* = \bar{\boldsymbol{a}}^\mathsf{T}$) . |
| $\|\boldsymbol{A}\|_2$ | The vector 2-norm of $m \times n$ matrix $\boldsymbol{A}$, so that $\|\boldsymbol{A}\|_2 = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\|a_{ij}\|^2}$. |
| $\boldsymbol{I}$ | Identity matrix of an appropriate size. |
| $E(X)$ | Expected value of random variable $X$. |
| $\mathcal{B}(1,p)$ | Bernoulli distribution with a probability of $p$ for success, and $1-p$ for failure. |
| $\mathcal{N}(\mu,\sigma^2)$ | Gaussian distribution with mean $\mu$ and standard deviation $\sigma$; $\mathcal{N}(0,1)$ denotes the standard Gaussian distribution. |
| $\mathcal{CN}(0,2\sigma^2)$ | Circularly-symmetric complex Gaussian distribution with zero mean and standard deviation $\sigma$. |
| $\boldsymbol{A}^+$ | Pseudo-inverse of matrix $\boldsymbol{A}$, as defined in Eq. (3.6). |
| $\nabla_{\boldsymbol{a}} f(\boldsymbol{a})$ | Gradient of scalar function $f(\boldsymbol{a})$. |

# Parameters and variables

The following lists a selection of the variables and parameters, which are used at multiple places in the text. If applicable, the number of the equation in which the variable or parameter is first introduced is indicated.

| | | |
|---|---|---|
| → (3.9) | $\boldsymbol{\alpha}$ | Snapshot weights |
| → (3.7) | $\boldsymbol{B}$ | Estimation matrix ($N_\mathrm{P} \times N_\mathrm{M}$) |
| → (3.11) | $\boldsymbol{\beta}_j$ | Row vector from the row of $\boldsymbol{B}$ corresponding to the $j$th target point |
| | $\boldsymbol{C}(s)$ | Diagonal matrix of RB position controllers $C(s)$ |
| → (5.42) | $C(s)$ | RB position controllers |
| → (3.2) | $\boldsymbol{\varepsilon}_\mathrm{m}$ | Vector describing the instantaneous sensor noise |
| → (3.11) | $\varepsilon_j$, $\varepsilon_{j,\mathrm{meth}}$, $\varepsilon_{j,\mathrm{noise}}$ | The estimation error for the $j$th target point, the part due to the methodical and due to the noise error, respectively |
| → (3.3) | $\boldsymbol{f}$ | Vector of input forces on a system |
| | $f_\mathrm{co}$ | Cutoff frequency of a lowpass or highpass filter |
| | $f_\mathrm{s}$ | Sample frequency |
| | $\theta$ | Orientation angle of a strain sensor |
| → (3.3) | $\boldsymbol{K}$ | Stiffness matrix of a system |
| | $\boldsymbol{M}$ | Mass matrix of a system |
| → (6.11) | $MSE$ | Mean square error, i.e. expected value of the amplitude estimate error |
| → (6.2) | $N_\mathrm{C}$ | Number of sensor candidates |
| | $N_\mathrm{D}$ | Number of shapes in a snapshot matrix |
| | $N_\mathrm{F}$ | Number of fitting shapes after truncation |
| | $N_\mathrm{M}$ | Number of sensors |
| | $N_\mathrm{P}$ | Number of measurement and target points |

| | | |
|---|---|---|
| $N_{\mathrm{PC}}$ | Number of possible sensor placement configurations | → (6.2) |
| $N_{\mathrm{rb}}$ | The number of rigid body modeshapes in the out-of-plane direction | |
| $N_w$ | The number of target points | → (6.18) |
| $\hat{\boldsymbol{q}}$ | Estimate weights of the fitting shapes in $\boldsymbol{\Phi}$ | → (3.2) |
| $\tilde{\boldsymbol{q}}$ | Estimate weights of the shapes in $\boldsymbol{\Phi}_{\mathrm{m}}$, only corrupted by sensor noise | → (6.4) |
| $\boldsymbol{q}_{\mathrm{rb}}$ | Rigid body coordinates, i.e. modal amplitudes of the rigid body modes | → (5.6) |
| $\overline{\boldsymbol{q}}_{\mathrm{rb}}$ | Estimate of the rigid body coordinates | → (5.8) |
| $\boldsymbol{R}^{\intercal}$ | Transformation matrix for recombining the flexible residues ($(N_S-3)\times N_S$) | → (5.16) |
| $\boldsymbol{S}_{\mathrm{mrb}}$ | Diagonal matrix weighing the use of the position sensors for estimating the rigid body coordinates ($N_{\mathrm{rb}}\times N_{\mathrm{rb}}$) | → (5.11) |
| $SNA$ | Sensor noise amplification, i.e. expected value of the shape estimate error | → (6.20) |
| $\boldsymbol{\Sigma}_{\alpha}$ | Covariance matrix of the snapshots weights, $\boldsymbol{\alpha}$ | → (3.10) |
| $\boldsymbol{\Sigma}_{\varepsilon}$ | Covariance matrix of the sensor noise, having the sensor noise variances $\sigma^2_{\varepsilon,i}$ on its diagonal | → (3.2) |
| $\sigma_{\varepsilon}$ | Standard deviation of the the noise of each of the sensors | → (6.7) |
| $\sigma_{\varepsilon,i}$ | Standard deviation of the the noise of the $i$th sensor | → (3.2) |
| $\boldsymbol{T}$ | Decoupling matrix (defined in Eq. 5.11) ($N_{\mathrm{rb}}\times N_{\mathrm{M}}$ ) | → (5.8) |
| $t$ | Time variable | |
| $u$, $v$, $w$ | Displacement field in respectively $x$, $y$ and $z$-direction. | |
| $\boldsymbol{W}$ | Snapshot matrix ($N_{\mathrm{P}}\times N_{\mathrm{D}}$) | → (3.9) |
| $\boldsymbol{w}$ | Vector describing the momentaneous displacement field of an object | → (3.9) |
| $\hat{\boldsymbol{w}}$ | Estimate of $\boldsymbol{w}$ | |
| $\boldsymbol{w}_j$ | Row vector from the row of $\boldsymbol{W}$ corresponding to the $j$th target point | → (3.11) |
| $\boldsymbol{w}_{\mathrm{fl}}$ | The contributions of the flexible modeshapes to $\boldsymbol{\Delta z}$ | → (5.6) |
| $\overline{\boldsymbol{W}}_{\mathrm{fl}}$ | Matrix with the flexible residues of a position snapshot matrix | → (5.18) |
| $\overline{\boldsymbol{w}}_{\mathrm{fl}}$ | The flexible residue, i.e. the difference between the actual displacement field and the estimate rigid body displacement field | → (5.10) |
| $\boldsymbol{W}_{\mathrm{m}}$ | Subset of the rows of $\boldsymbol{W}$ corresponding to the sensor locations ($N_{\mathrm{M}}\times N_{\mathrm{D}}$) | → (3.10) |

| | | |
|---|---|---|
| → (3.1) | $\boldsymbol{w}_\mathrm{m}$ | Subset of the elements of $\boldsymbol{w}$ corresponding to the sensor locations |
| | $\boldsymbol{w}_\mathrm{m,fl}$ | Subset of the elements of $\boldsymbol{w}_\mathrm{fl}$ corresponding to the sensor locations |
| | $\overline{\boldsymbol{w}}_\mathrm{m,fl}$ | Subset of the elements of $\overline{\boldsymbol{w}}_\mathrm{fl}$ corresponding to the sensor locations |
| → (3.2) | $\tilde{\boldsymbol{w}}_\mathrm{m}$ | The measurement of $\boldsymbol{w}_\mathrm{m}$, i.e. corrupted with sensor noise $\boldsymbol{\varepsilon}_\mathrm{m}$ |
| → (5.17) | $\hat{\boldsymbol{w}}_\mathrm{fl}$ | Estimate of $\boldsymbol{w}_\mathrm{fl}$ |
| → (5.18) | $\overline{\boldsymbol{W}}_\mathrm{R,fl}$ | Matrix with the recombined flexible residues of a position snapshot matrix |
| → (5.16) | $\overline{\boldsymbol{w}}_\mathrm{R,fl}$ | The recombined flexible residues |
| → (3.10) | $\boldsymbol{W}_\mathrm{t}$ | Subset of the rows of $\boldsymbol{W}$ corresponding to the target locations |
| → (3.1) | $\boldsymbol{w}_\mathrm{t}$ | Subset of the elements of $\boldsymbol{w}$ corresponding to the target locations |
| | $\boldsymbol{\Phi}$ | Matrix consisting of the fitting shapes ($N_\mathrm{P} \times N_\mathrm{F}$) |
| → (2.8) | $\breve{\boldsymbol{\Phi}}$ | Matrix consisting of modeshapes |
| → (4.4) | $\breve{\boldsymbol{\Phi}}$ | Matrix consisting of proper orthogonal modes ($N_\mathrm{P} \times N_\mathrm{F}$) |
| → (2.6) | $\breve{\boldsymbol{\varphi}}$ | Modeshape |
| → (4.1) | $\breve{\boldsymbol{\varphi}}$ | Proper orthogonal mode ($N_\mathrm{P} \times 1$) |
| | $\boldsymbol{\Phi}_\mathrm{m}$ | Subset of the rows of $\boldsymbol{\Phi}$ corresponding to the sensor locations ($N_\mathrm{M} \times N_\mathrm{F}$) |
| | $\breve{\boldsymbol{\Phi}}_\mathrm{m,rb}$ | Subset of the rows of $\breve{\boldsymbol{\Phi}}_\mathrm{rb}$ corresponding to the sensor locations ($N_\mathrm{M} \times N_\mathrm{rb}$) |
| | $\breve{\boldsymbol{\Phi}}_\mathrm{rb}$ | Matrix consisting of rigid body modeshapes ($N_\mathrm{P} \times N_\mathrm{rb}$) |
| | $x$, $y$, $z$ | Cartesian coordinates, of which $z$ is in out-of-plane direction |
| → (5.6) | $\boldsymbol{\Delta z}$ | The positions in out-of-plane direction of the points on an object with respect to a certain reference |
| → (5.19) | $\boldsymbol{\Delta}\hat{\boldsymbol{z}}$ | Estimate of $\boldsymbol{\Delta z}$ |
| | $\boldsymbol{\Delta z}_\mathrm{m,rb}$ | Subset of the elements of $\boldsymbol{\Delta z}_\mathrm{rb}$ corresponding to the sensor locations |
| | $\boldsymbol{\Delta}\overline{\boldsymbol{z}}_\mathrm{m,rb}$ | Subset of the elements of $\boldsymbol{\Delta}\overline{\boldsymbol{z}}_\mathrm{rb}$ corresponding to the sensor locations |
| → (5.6) | $\boldsymbol{\Delta z}_\mathrm{rb}$ | The contribution of the rigid body modeshapes to $\boldsymbol{\Delta z}$ |
| → (5.9) | $\boldsymbol{\Delta}\overline{\boldsymbol{z}}_\mathrm{rb}$ | The estimate rigid body displacement field of the points on an object |

# Acronyms and abbreviations

| | |
|---|---|
| CoM | Centre of Mass |
| CPS | Cumulative Power Spectrum |
| DoF(s) | Degree(s) of Freedom |
| EFI | Effective Independence |
| EUV | Extreme Ultraviolet |
| FBG | Fibre Bragg Grating |
| FE | Finite Element |
| FFP | Fibre Fabry–Pérot |
| FP | Frame Potential |
| FRF | Frequency Response Function |
| FS | Frame Sense |
| LS | Least Squares |
| MAC | Modal Assurance Criterion |
| MEMS | Microelectromechanical system |
| MSE | Mean Square Error |
| PID | Proportional, Integral and Derivative |
| POD | Proper Orthogonal Decomposition |
| PoI | Point of Interest |
| POM | Proper Orthogonal Mode |
| POV | Proper Orthogonal Value |
| PSD | Power Spectral Density |
| RA12 | Sensor placement algorithm of Ranieri et al. [91] |

| | |
|---|---|
| RB | Rigid Body |
| RHIE | Receding Horizon Input Estimator |
| RMS | Root Mean Square |
| SF | Shape Fitting |
| SISO | Single Input/Single Output |
| SNA | Sensor Noise Amplification |
| SVD | Singular Value Decomposition |
| TA | Sensor placement algorithm of Tasker and Liu [90] |

# Acknowledgements

This thesis is the result of four and a half years of research work. It is not the only result, however. During these years I had the chance to learn new skills and to further develop others. I went on learning how to perform research, widened my knowledge of mechatronics and had some experience with teaching tasks. I would like to thank those many people that in some way or another contributed to this.

First of all, I want to thank my promotor, Rob Munnig Schmidt. Owing his lectures I chose to go further into Mechatronics. Rob, I enjoyed working on your book and on your course and I have vivid memories of our discussions, like on the value of matrices and the charge in piezo-electrics. I would also like to thank my daily supervisor Jo Spronck for all discussions and support. Jo, it was amazing that, although you were always busy, you had always *'even*[1]' time. I remember you lessons about 'what do we *really* want?', 'if you want it, it is possible' and on keeping critical throughout the process.

Furthermore, I want to thank my temporal supervisors Arturo Tejada Ruiz and Paolo Tiso for their help in brainstorming on this subject. Arturo, I remember, for example, the efforts you put in helping me with my first oral presentation at a conference (at MIT). I would also like to thank Jan van Eijk, with whom I had several discussions, which were of crucial importance for the directions I took in this research.

I would like to thank Ron van Ostayen for organising the MSD-meetings, where we learned each week about the topics of a Ph.D. or a graduate student and got a flavour of the various topics in our field.

Furthermore, I would like to thank the people of ASML with whom I had contact, amongst others Marc van de Wal, Chris Valentin, Jeroen de Best and George Leenknegt. Thank you for the discussions and your help during the period that I worked at ASML.

I would also like to thank my colleague researcher in the IMWACOL project, Ruxandra Mustata, for the nice discussions and the collaboration in coaching one of the master's students. I enjoyed to learn a bit of the DCSC-way of approaching similar problems.

During the four and a half years I shared a room and many ideas with my fellow Ph.D.s and postdocs, Guido, Jeroen, Chris, Patrice, Pablo, Rudolf, Arjan, Jan, Teun, David, Alejandro, Takeshi, Phuc (Foppe), Ruijun, Oscar, Stefan

[1] The Dutch word.

and, although not of MSD, Evert, Kostiantyn and Gijs. Thank you for the nice and comradely atmosphere and all fun we had. I remember the office bike, which was used for easy transportation within the office, especially when close to stressful conference deadlines. We survived the intermittently recurring floodings of our office due to the, seemingly, insoluble leakage. A special thanks to my good friend Oscar, who was always willing to help and think along, also about obscure mathematical concepts that were possibly not always of the highest interest to him.

Also the students of our mechatronics lab contributed to the good atmosphere, thank you for this! We worked in the lab, had coffee breaks and lunch together and learned about each other's projects. We lifted each other's spirits and laughed about salespeople that persistently call the RMS noise of their sensors peak-to-peak.

A special word of thanks to the students whom I coached during their graduation: Stefan van der Kleij and Floris Klein Horsman. You helped me realising the setup of the position controlled plate. You helped designing, manufacturing, building and qualifying the setup and doing experiments with it. Excellent work, which partly found a place in Section 2.4.2 and Chapter 7. It was a pleasure to work with you and to contribute to your transformation from a student to an engineer.

I would like to thank Rob Luttjeboer, for his professional way of keeping the lab workable and safe and his readiness to help with materials and practical advice. I would also like to thank Harry and Patrick for their practical assistance. I would like to thank Dennis de Klerk, Paul van der Valk and Maarten van der Seijs for their coaching during the dynamic tests of the Thorlabs plate.

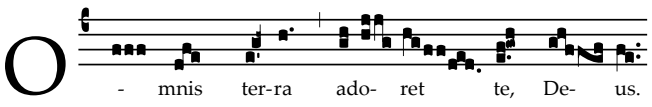
Maaike Beliën thaught two courses related to academic writing. I would like to thank her for her individual advice in setting up the structure of the thesis and reviewing some of the texts. Her help saved months of reinventing wheels, as writing is to some extent comparable to composing music: you need to know the basic rules and to keep sticking to them until you are experienced enough to deviate.

Also a word of thanks to my new colleagues and the students at the Electronic Instrumentation Laboratory. Thanks for the support and the discussions.

Finally, I would like to thank my parents for their continuous support and my family and friends, amongst others of the church and the choirs, for all fun and moments of relaxation next to the research work. I would like to thank Karel de Haan, who introduced me to computers, blocking the games and teaching programming in Basic and Delphi, thereby awaking interests which proved valuable throughout my studies and the Ph.D. time.

Johan Vogel

Delft, September 2016



# Curriculum vitae


| | |
|---|---|
| Name | ir. J.G. Vogel |
| E-mail address | j.g.vogel@tudelft.nl |


## Biography

Johan studied Mechanical Engineering at Delft University of Technology. His bachelor research project (2008) was in the field of biomechanical engineering, where he did research on shape lock mechanisms for guiding coloscopes. Specialising in mechatronic system design, he graduated Cum Laude (2011) on the design of a parallel haptic device for pick-and-place operation, for which he received the Best Graduate Award of the faculty Mechanical, Maritime and Materials Engineering. From 2011 to 2015, he worked as a Ph.D. candidate in the group of Mechatronic System Design of prof.ir. Rob Munnig Schmidt, focussing on the development of methods for measuring and estimating deformation in high-precision mechatronic systems such as lithography machines. As of October 2015, he started as a postdoctoral researcher within the group of Industrial Electronic Instrumentation at Delft University of Technology, developing and improving Eddy-current position sensors to let them meet the requirements of high-precision applications.

Johan has a wide interest in mechanical engineering, especially in topics in which theory and practice go together, which is often the case in mechatronic problems. He is particularly interested in mechanisms, mechatronic design, dynamics and motion control. Furthermore, he is interested in metrology, frequency domain techniques, inverse problems and optimisation.

## Education and Employment

| | |
|---|---|
| 2015 – now | Postdoctoral researcher on *Eddy-Current Position Sensing for High-Precision Applications*, advisor: Dr. M.Sc. Stoyan Nihtianov, Delft University of Technology, department of Microelectronics, Electronic Instrumentation Laboratory. |
| 2011 – now | Ph.D. candidate Mechanical Engineering on *Shape Estimation in Lithography Machines*, advisors: ir. Jo Spronck and |

| | |
|---|---|
| | prof.ir. Rob H. Munnig Schmidt, Delft University of Technology, department of Precision and Microsystems Engineering, group Mechatronic System Design. |
| 2009 – 2011 | M.Sc. Mechanical Engineering on *the Design of a 5-DoF Parallel Haptic Robot* (obtained 17 May 2011, Cum Laude), advisors: P. Lambert, M.Sc. and prof.ir. Rob H. Munnig Schmidt, Delft University of Technology, department of Precision and Microsystems Engineering, group Mechatronic System Design. |
| 2010 – 2011 | Student assistant Mechatronic System Design (course assistant), Delft University of Technology. |
| 2009 | Honours Programme in Mechanical Engineering, Delft University of Technology, 2009. |
| 2005 – 2008 | B.Sc. Mechanical Engineering on *Shape Fixation for Endoscopes* (obtained 24 October 2008, Cum Laude), Delft University of Technology. |
| 2005 – now | Coaching Math (all levels secondary school), Latin, Biology. |
| 1999 – 2005 | VWO at Utrechts Stedelijk Gymnasium (Nature&Health and Nature&Technology) |

## Awards

- Supervisor of graduate student Stefan van de Kleij, who graduated Cum Laude, December 9, 2014, Delft University of Technology.
- Poster Award 1st Prize, Euspen 12th International Conference, June 4 – 8 2012, Stockholm, Sweden.
- Best Graduate of the Faculty Mechanical, Maritime and Materials Engineering (3mE) Award 2010–2011, Delft University of Technology.
- M.Sc. Mechanical Engineering Cum Laude, May 17, 2011, Delft University of Technology.

## Publications

- A. Loeve, O.S. van de Ven, J.G. Vogel, P. Breedveld, J. Dankelman, "*Vacuum packed particles as flexible endoscope guides with controllable rigidity*," Granular Matter vol. 12, pp. 543 – 54, 2010.
- J.G. Vogel, J.W. Spronck, R.H. Munnig Schmidt, "*Deformation measurement of a plate with damping contact*," in: Proceedings of the 12th international conference of the EUSPEN, Vol. 1, pp. 210 – 3, 2012.
- J.G. Vogel, A. Tejada, J.W. Spronck, R.H. Munnig Schmidt, "*Estimation of the deformation of a plate using a limited number of sensors*," in: Proceedings of the ASPE Spring Topical Meeting on Precision Control for Advanced Manufacturing Systems, pp. 70 – 5, 2013.

- J.G. Vogel, J.W. Spronck, R.H. Munnig Schmidt, *"Estimating deformation of a free-floating wafer chuck*," in Proceedings of the 14th international conference of the EUSPEN, 2014.
- J.G. Vogel and S. Nihtianov, *"Modelling the inductance of a novel eddy-current position sensor for high-precision applications*," Sensors Applications Symposium (SAS), IEEE, pp. 1 – 5, 2016.
- J.G. Vogel and S. Nihtianov, *"Tilt sensitivity of an eddy-current position sensor for high-precision applications,"* in Proceedings of the 16th international conference of the EUSPEN, 2016.
- V. Chaturvedi, J.G. Vogel, S. Nihtianov, *"Suppression efficiency of the correlated noise and drift of self-oscillating pseudo-differential eddy current displacement sensor,"* in Proceedings of the 30th EuroSensors conference, 2016.
- J.G. Vogel and S. Nihtianov, *"Study of the self-resonance frequency of a flat coil for an eddy-current position sensor"* in Proceedings of the IEEE Sensors conference, 2016.

## Teaching activities

- Supervision of graduation projects
- WBTP303 Bachelor's Research and Design Project (coaching of student groups)
- Book *The design of High Performance Mechatronics* (Munnig Schmidt et al., 2011, 2nd ed. 2014) (proof reading and contributions)
- WB2427 Mechatronic System Design (development of graded assignments, assignment feedback lecture)
- WB5452-14 Intro lab PME (preparation of setups, instruction)
- Member of exam committees and review of journal papers

## Other activities

- Presenter Open Days TU Delft (boot on haptic systems)
- Presenter for TU Delft and 3TU at exhibitions (Precision Fair and Euspen international conference)
- Website design and development for several commercial and non-profit organisations
- Secretary of the board at the Gooise Koorschool choir

## Personal interests

- Classical singing, piano
- Building and operating 3D printers

In wafer scanners – the machines that define the details of electronic chips – there is a need for highly accurate deformation measurements of the machine components during the chip manufacturing process.

This thesis develops an estimation methodology, based on shape fitting principles, that aims at a low estimation error and addresses the specific requirements related to one of the components of a wafer scanner, the wafer chuck.