Data-Driven Optimal Control
for Adaptive Optics

Cover: Telescopes in a wavefront landscape
Jeroen ten Asbroek, jeroen@danceonair.nl

DATA-DRIVEN OPTIMAL CONTROL FOR ADAPTIVE OPTICS

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op

dinsdag 23 januari 2007 om 12:30 uur

door

Karel Johannes Gerhardus HINNEN

natuurkundig ingenieur
geboren te Enschede

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. M.H.G. Verhaegen

Samenstelling promotiecommisie:

Rector Magnificus,	voorzitter
Prof. dr. ir. M.H.G. Verhaegen,	Technische Universiteit Delft, promotor
Prof. dr. R. Babuška,	Technische Universiteit Delft
Prof. dr. ir. B. De Moor,	Katholieke Universiteit Leuven
Prof. dr. J.S. Gibson,	University of California, Los Angeles
Prof. dr. ir. M. Steinbuch,	Technische Universiteit Eindhoven
Dr. ir. N.J. Doelman,	TNO Industrie en Techniek
Dr. J. -M. Conan,	Office National d'Etudes et de Recherches
Prof. dr. ir. J. Hellendoorn,	Technische Universiteit Delft, reservelid

Dr. ir. N.J. Doelman heeft als begeleider in belangrijke mate aan de totstandkoming van het proefschrift bijgedragen.

This dissertation has been completed in partial fulfillment of the requirements of the Dutch Institute of Systems and Control DISC for graduate study.

This research has been conducted in the framework of the "Knowledge Center for Aperture Synthesis" (KAS). The knowledge center is a long-term co-operation of TNO Science and Industry and Delft University of Technology to develop advanced technologies for optical aperture synthesis.

Printed in the Netherlands.

Acknowledgments

Looking back, the past four years have been a valuable but turbulent experience. It is only because of the support and encouragement of many people – my family, friends and colleagues – that I have finally been able to overcome the distortions and finish my (re)search in peace and quiet. Therefore, I would like to thank all of you who have turned the past four years into a memorable time. Thanks to you, doing research has really been great!

Besides this general word of gratitude, a number of people deserve special mention. Over years there was always the support of my supervisor Michel Verhaegen. Michel, you taught me to find my way through the labyrinth of possibilities that a scientist encounters in doing research. Starting up together a new research direction has really been exciting but not always easy. During the project we had to spend a great deal of our time in getting familiar with existing concepts and techniques. Your solid background in system identification has helped me to get grip on the problem. Thank you for your enthusiasm, your stimulating discussions, your never remitting trust and the freedom you gave me in doing my research. My other adviser has been Niek Doelman from TNO Science and Industry. Niek, thank you for all your interesting ideas and feedback. Thank you also for guarding the link between theory and practical relevance.

I am also very grateful to TNO who provided the financial support and the lab facilities to enable this research. My special thanks go out to Kees van den Berg who has played a crucial role in getting the AO setup started. Without your experience on digital camera's and real-time Linux it would never have been possible to validate the control theory. Thanks go also to Rob Vink, Harry de Man and Amir Vosteen for their help in optically related questions, to Wouter Mulckhuyse for his help in solving problems with real-time Linux and Johan Leijtens for finding the funds for a new deformable mirror. I also want to thank Gleb Vdovin from OKO Technologies for his advise and his efforts in minimizing the delivery time of the mirror. I wish to thank Rudolf Le Poole and Remko Stuik from Leiden Observatory. Rudolf, thank you for your patience in explaining the fundamental limitations in wavefront sensing and your never ending enthusiasm for everything related to astronomy instrumentation (and sailing). Thanks also to Richard Wilson, Matthias Schöck and Christoph Keller for providing open-loop WFS data from the William Herschel, Keck and McMath-Pierce telescope.

Special thanks go to Rufus Fraanje and Stoyan Kanev (Tony) who have been

far more than just good colleagues. Thank you both for the good working atmosphere, for the support, and for all the long and lively discussions also outside office hours. Tony, you have always been willing to help me, read my papers and manuscripts and provide me with indispensable feedback. For this I'm very grateful. Thank you also for being a challenging opponent in several tennis and squash games. Rufus, thank you for all your good ideas and advices and for sharing your positive view on life. The many cultural and social activities have really broadened my mind. I also want to acknowledge the pleasant cooperation with Rogier Ellenbroek over the last two years. Having a colleague working in application field has been inspiring, hopefully for the both of us. Thank you also for the many squash games – I still hope to defeat you once.

Furthermore, I want to thank all members from the Delft Center for Systems and Control and former of Control Systems Engineering group for making me feel at home. In particular, I want to mention Robert Bos, Xavier Bombois, Arjan den Dekker, Rufus Fraanje, Dirk Poot, Stoyan Kanev and Vincent Verdult with whom I had oh so many interesting conversations during lunch and coffee breaks. Also thanks to secretaries Kitty Dukker and Ellen van den Berg who helped me out with dozens of questions and many forms and to the system administrators, Will van Geest, Daan Noteboom and Arjan van Dijke for sharing their knowledge and keeping my computer in good shape. I truly enjoyed the several nice workshop, symposia and conference trips during the project. In this context I want to mention Bas Swinkels, Edgar Vuelban, Freek Stoffelen, Mariëlle van Veggel and Niek Doelman. Your good company during the conferences and subsequent round trips is greatly acknowledged.

And last but not least, I want to thank my parents Hans and Adri and my brother Janwillem for their ongoing interest, encouragement, trust and care over all those years. Your unconditional support has been indispensable in coming this far. As you will understand, words are not adequate to express my gratitude towards you!

Delft, September 2006
Karel Hinnen

Summary

For hundreds of years, atmospheric turbulence has imposed a serious constraint on the angular resolution of ground-based astronomical telescopes. Without any form of compensation, atmospheric turbulence blurs the images and limits the angular resolution in the visible to about 0.5 to $1 \mathrm{arcsec}$. Adaptive optics (AO) is a technique for correcting the optical wavefront distortions introduced in a light beam as it propagates through a turbulent medium. An AO imaging system compensates for the wavefront phase errors by sensing the perturbation with a wavefront sensor (WFS) and adding the estimated conjugated phase by actively adjusting the optical path length differences with a deformable mirror (DM). This thesis focuses on the control aspects of AO.

Most AO systems are based on a simple control law that consists of a static wavefront reconstruction step followed by a series of parallel single-input single-output temporal compensators. Important drawbacks of this approach are that it implicitly assumes a decoupling of the spatial and temporal dynamics and that it is not able to explicitly account for the DM and WFS dynamics and the temporal evolution of the wavefront. Furthermore, when using a minimum-variance wavefront reconstructor, the modified wavefront statistics as a result of closed-loop operation are usually neglected. The temporal error caused by the finite time delay between measurement and correction is know to be one of the main limitations on the performance of an AO system. Also the measurement noise is known to constitute an error source of significant importance, especially when the AO system is operating on faint guide stars. The purpose of the research presented in this thesis is to demonstrate that these errors can be reduced by using a rigorous control strategy that is able to exploit the spatio-temporal correlation in the wavefront and explicitly accounts for DM and WFS dynamics.

The control strategy used to achieve these goals is one of data-driven disturbance and system modeling followed by a minimum-variance or $\mathcal{H}_2$-optimal control design. In this approach the second-order statistics of the atmospheric wavefront distortions are modeled as a regular stochastic process. The problem of finding the spectral factor that accurately describes the relevant turbulence dynamics, is the most complicated and fundamental step. For this reason, a significant part of the work has been devoted to stochastic disturbance modeling. Two strategies for identifying an atmospheric disturbance model have been elaborated. The first strategy is based on approximating the theoretical turbulence spectrum. To this end, a subspace-based algorithm for estimating the

minimum-phase spectral factor from samples of a matrix-valued power spectrum has been developed. The algorithm has been successfully applied to approximate the non-rational Kolmogorov power spectrum, but lacks computational efficiency required to identify a full multi-variable atmospheric disturbance model for the number WFS channels in an AO system. The second approach for determining an atmospheric disturbance model has been proven to be more suitable for this. It consists of a dedicated subspace identification algorithm that estimates the atmospheric disturbance model directly from open-loop WFS. In combination with a re-parametrization of the WFS space, the developed subspace algorithm is sufficiently efficient to identify an atmospheric disturbance model for AO systems with up to a few hundred degrees of freedom. Since the approach does not assume any form of decoupling, the identified disturbance model should be able to capture the spatio-temporal correlation imposed by frozen turbulence satisfying the Taylor hypothesis.

Given the identified atmospheric disturbance model and a model of the AO system dynamics, the AO control problem has been formulated in a $\mathcal{H}_2$-optimal control framework. It has been shown that, as a result of the minimum-property of the disturbance model, the general solution to the $\mathcal{H}_2$-optimal control problem can be simplified so that, instead of two, at most one Riccati equation has to be solved. Moreover, if the model of AO system dynamics is minimum-phase or has a known inner-outer factorization, also the second Riccati equation can be eliminated, giving rise to an analytical way of computing the optimal controller. This observation has been used to derive an analytical expression for the $\mathcal{H}_2$-optimal controller in the case that the AO system can be characterized by a scalar-dynamic transfer function consisting of an integer number of samples delay and a two taps impulse response. By analyzing the dynamic behavior of the WFS camera, this particular model structure has shown to be valid for any AO system with a DM that has a time-constant that is short to the WFS exposure time.

The data-driven optimal control approach obtained by combining the proposed subspace identification algorithm and the analytical expression for the $\mathcal{H}_2$-optimal controller has been experimentally demonstrated on an AO laboratory setup. An extensive validation study has shown that, compared to the common control law consisting of a minimum-variance wavefront reconstructor and a first-order lag filter, optimal control is effective in reducing the temporal error. This implies that the gain in performance is especially large at high Greenwood to sample frequency ratios, where the temporal error becomes dominant. Optimal control may therefore help to improve the performance of current AO systems in heavy turbulence conditions, including high wind speeds and small Fried parameters, as well as under low level light conditions where high sampling frequencies is ruled out because of the measurement noise. Even this has not been considered in the validation experiments, optimal control is also expected to be useful in reducing the error contribution due to measurement noise as it explicitly accounts for the spatio-temporal correlation in the wavefront.

Samenvatting

Atmosferische turbulentie heeft eeuwenlang het oplossend vermogen van telescopen op aarde beperkt. Zonder enige vorm van compensatie, vervaagt het de beeldvorming en beperkt de resolutie tot ongeveer $0.5 - 1$ arcsec voor zichtbaar licht. Adaptieve optica (AO) is een techniek waarmee het mogelijk is om de golffront verstoringen die ontstaan in een turbulent medium actief te compenseren. Gebruik makend van een golffront sensor (WFS) schat het systeem de verstoringen, waarna deze worden gecompenseerd door actief het optische weglengte verschil in het systeem aan te passen met een vervormbare spiegel (DM). Dit proefschrift concentreert zich op de regeltechnische aspecten van AO.

De meeste AO systemen zijn gebaseerd op een eenvoudig regelschema, dat bestaat uit een statische golffront reconstructie stap, gevolgd door een serie onafhankelijke parallelle servo compensatoren. Een nadeel van deze aanpak is dat het impliciet een ontkoppeling van de spatiële en temporele dynamica veronderstelt. Daarnaast laat de structuur van de servo compensatoren het veelal niet toe om expliciet rekening te houden met de dynamica van de DM en WFS en de temporele evolutie van het golffront. Bovendien wordt bij gebruik van een minimum-variantie golffront reconstructor de gewijzigde golffront statistiek als gevolg van het feit dat de verstoringen in gesloten lus worden gemeten, meestal buiten beschouwing gelaten. Het is bekend dat de temporele fout veroorzaakt door de tijdsvertraging tussen meting en correctie, één van de belangrijkste beperkingen is met betrekking tot de prestaties van een AO systeem. Ook de meetruis levert een belangrijke bijdrage aan het totale fouten budget van het AO systeem, met name wanneer er gewerkt wordt met een zwakke hulp ster. In dit onderzoek wordt aangetoond, dat deze fouten kunnen worden beperkt door gebruik te maken van een regelstrategie die rekening houdt met zowel de DM- en WFS dynamica als de spatiële en temporele dynamica van de golffront verstoring.

Om dit doel te bereiken wordt voor het modelleren van zowel de golffront verstoringen als de AO dynamica gebruik gemaakt van data gebaseerde systeem identificatie. De geïdentificeerde modellen worden vervolgens benut om een minimum-variantie of een $\mathcal{H}_2$-optimale regelaar te bepalen. In deze aanpak wordt de tweede-orde statistiek van de golffront verstoringen gemodelleerd als een regulier stochastisch proces. Het vinden van de spectrale factor die de relevante turbulentie dynamica nauwkeurig beschrijft is hierbij de moeilijkste en de meest fundamentele stap. Een belangrijk deel van het onderzoek is daarom aan het modelleren van stochastische turbulentie gewijd. Er zijn twee strategieën ten aanzien van het identificeren van een verstoringsmodel nauwkeurig onderzocht. De eerste

methode is gebaseerd op het benaderen van het theoretische verstoringsspectrum. Hiertoe is een op subspace identificatie gebaseerd algoritme ontwikkeld dat als doel heeft de minimum fase spectrale factor te vinden die de gegeven samples van het powerspectrum zo nauwkeurig mogelijk beschrijft. Het algoritme is met succes gebruikt om het niet-rationele Kolmogorov power spectrum te benaderen, maar ontbeert de efficiëntie om een volledig multi-variabel atmosferisch verstoringsmodel te identificeren voor het aantal WFS kanalen dat in AO gebruikelijk is. De tweede methode blijkt hiervoor meer geschikt en maakt gebruik van een subspace identificatie algoritme, dat het atmosferisch verstoringsmodel direct op basis van open-loop WFS schat. In combinatie met een reparametrizatie van de WFS ruimte, is dit algoritme voldoende efficiënt om een verstoringsmodel voor AO systemen met maximaal een paar honderd vrijheidsgraden te identificeren. Aangezien de methode geen enkele vorm van ontkoppeling veronderstelt, zou het verstoringsmodel de spatiële en temporele correlatie, zoals opgelegd door een als bevroren te beschouwen stroming van turbulentie in overeenstemming met de Taylor hypothese, moeten kunnen beschrijven.

Uitgaande van het geïdentificeerde verstoringsmodel en het model van het AO systeem, kan het AO regelprobleem worden geformuleerd als het bepalen van de $\mathcal{H}_2$-optimale regelaar. Aangezien het verstoringsmodel minimum fase is, kan de algemene oplossing van het $\mathcal{H}_2$-optimale regelprobleem worden vereenvoudigd zodat in plaats van twee hooguit één Riccati vergelijking moet worden opgelost. Een verdere vereenvoudiging kan worden bereikt indien het model van het AO systeem minimum fase is of een bekende *inner-outer* factorizatie heeft. In dit geval kan ook de tweede Riccati vergelijking worden geëlimineerd zodat een analytisch uitdrukking ontstaat. Op deze manier kan er een analytische uitdrukking voor de $\mathcal{H}_2$-optimale regelaar worden afgeleid in het geval dat het AO systeem kan worden gekarakteriseerd door een scalaire overdrachtsfunctie bestaande uit een geheel aantal samples vertraging en een impulsresponsie met twee coëfficiënten. Door het dynamisch gedrag van de WFS camera te analyseren blijkt dat deze modelstructuur voor ieder AO systeem met een DM die een tijds-constante heeft die kort is ten opzichte van de WFS belichtingstijd geldt.

De data gedreven optimale regelstrategie, verkregen door het combineren van het subspace identificatie algoritme en de analytische uitdrukking voor de $\mathcal{H}_2$-optimale regelaar, is experimenteel gevalideerd op een AO laboratorium opstelling. Een validatie studie laat zien dat, in vergelijking tot de algemeen gangbare regelstrategie bestaande uit een minimum variantie golffront reconstructor en een eerste order filter, de voorgestelde optimale regelstrategie een effectief middel is om de temporele fout te reduceren. De winst van optimaal regelen ligt derhalve in het gebied waar de Greenwood frequentie groot is ten opzichte van de sample frequentie en de temporele fout dominant wordt. Optimaal regelen kan daardoor bijdragen aan het verbeteren van de prestaties van bestaande AO systemen, met name tijdens zware turbulentie met hoge wind snelheden en kleine Fried parameters, maar ook in situaties waar de sample frequentie wordt gelimiteerd door meetruis. Voorts is het aannemelijk, dat optimaal regelen een bijdrage kan leveren aan het reduceren van de foutbijdrage ten gevolge van meetruis, aangezien het de spatiële en temporele correlatie in het golffront in rekening brengt.

Contents

1 | CHAPTER

Introduction

A daptive optics is nowadays a well established technique to actively compensate the wavefront distortions introduced in a light beam as it propagates trough a turbulent medium. It has found widespread use in ground-based astronomical telescopes to restore the image quality by counteracting the devastating effect of atmospheric turbulence on the angular resolution.

This thesis focuses on the control aspects of adaptive optics (AO). To provide some additional background information and to better position the work, this introductory chapter will start with a brief overview of the basic principles and developments in the field of AO. Since AO is still a rapidly developing area the overview will be far from exhaustive. For more background information the reader is referred to the standard text books (Tyson 1998; Roddier 1999; Hardy 1998; Tyson 2000).

After this elementary overview, different control strategies in AO are reviewed. This forms the starting point for motivating the followed research strategy. The chapter finishes with an overview of the main contributions and a brief outline of the remaining chapters of the thesis.

1.1 Adaptive optics in astronomy

For thousands of years, astronomical observations play a crucial role in our attempts to reveal the mysteries of the universe. The introduction of the telescope at the beginning of the seventeenth century, resulted in a quantum leap in the angular resolution with respect to the naked-eye. Several important discoveries and breakthroughs in astronomy can be attributed to technical improvements of the telescope. These days, ground-based optical and near-infrared astronomical telescopes (see Figure 1.1) are indispensable tools in astronomy.

1.1.1 The need for adaptive optics

Two important properties of a telescope are its light collecting power and angular resolution. Like any optical imaging system, the angular resolution of a telescope is ultimately limited by diffraction (Hecht 1987; Born and Wolf 1999). Adopting Lord Rayleigh's criterion, the diffraction limited angular resolution of an optical imaging system with a circular aperture of diameter D is given by

$$\sin\theta \approx 1.22\frac{\lambda}{D},\tag{1.1}$$

where λ denotes the observing wavelength and the resolution θ is expressed in units of radians. The amount of detail that can be resolved increases hence with the telescope diameter D. This in combination with the improved light collecting power, forms an important drive to build larger and larger telescopes.
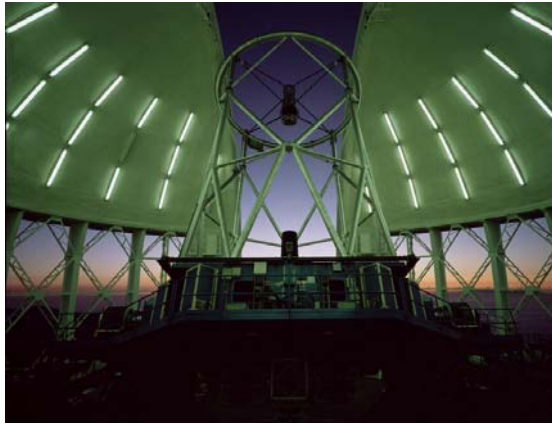
Figure 1.1: Gemini North telescope with open wind vents and observing slit. Gemini North is a 8m telescope located on Mauna Kea, Hawaii. Image Credit: Gemini Observatory / Association of Universities for Research in Astronomy.

The size of modern telescopes, however, has increased to such an extend that atmospheric turbulence has become a limiting factor. Atmospheric turbulence arises from large scale temperature inhomogeneities caused by solar heating of the Earth's atmosphere. By mixing air of different temperatures, it is responsible for random local fluctuations in the refractive index. The light collected by the telescope will hence experience time and space varying random fluctuations in the optical path length. As a result, a perfectly plane wavefront from a far and distant star will be no longer flat at the time it arrives at the telescope aperture. Some parts of the incoming light beam will be delayed with respect to other parts, resulting in a distorted wavefront.

Without taking any counter measures, the atmospheric wavefront distortions limit the achievable angular resolution to about $1\,\mathrm{arcsec}$ for observations in the near infrared. This type of resolutions correspond to the diffraction limited performance of a telescope with a diameter in the order of only 10 to $20\,\mathrm{cm}$. So when

the telescope diameter is increased beyond this size, the improvement in angular resolution will stay behind what can be expected from (1.1). The resolution will be completely determined by the atmosphere and is said to be seeing limited. To further improve the imaging quality of large ground-based telescopes it is necessary to reduce the devastating effect of the atmospheric wavefront distortions on the imaging process. This is the task of an adaptive optics system.

1.1.2 Principle of adaptive optics

To explain the principle of AO, consider the schematic drawing in Figure 1.2. When light from a distant star arrives at the outer layers of the atmosphere, it has a perfectly plane wavefront. However, this plane wavefront will never reach the telescope as the turbulent atmosphere will introduce time and space varying optical path length differences. This gives rise to a turbulence induced phase profile $\phi(\rho, t)$, where $\rho \in \mathbb{R}^2$ specifies the spatial position in the telescope aperture and t denotes time. The AO system tries to cancel out these wavefront distortions by actively introducing optical path length differences of opposite phase.
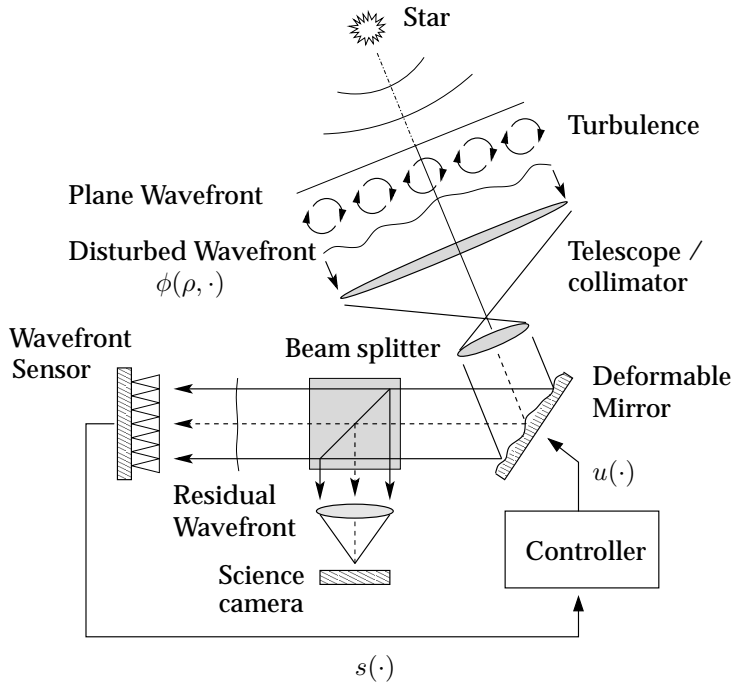
Figure 1.2: Schematic representation of an AO system, and its main components.

An AO system is typically composed of the following components – a wavefront sensor (WFS), an active component to influence the optical path length differences or phase and a feedback controller. In most systems, like the one depicted in Figure 1.2, the active optical component is a deformable mirror (DM). For the ease of discussion we will simply assume that the active component is a DM.

Light entering the AO system is first directed to the DM. By actively changing the mirror shape, the DM is able to apply a phase correction $\phi_m(\rho, t)$. The residual phase error is the difference between the turbulence induced wavefront and the applied correction, i.e. $\epsilon = \phi - \phi_m$. After applying the wavefront correction, a beam splitter divides the reflected light beam in two parts. The first part of the corrected light beam leaves the AO system and is used by the science camera to form an image of the object of interest. The remainder of the light is directed to the WFS, which provides quantitative information about the residual wavefront. Based on the WFS measurements $s(\cdot)$, the controller has to determine the actuator inputs $u(\cdot)$ to the DM. The controller should adapt the input signal in such a way that the DM cancels out most of the distortions.

By counteracting the wavefront distortions, AO is able to reduce the devastating effect of atmospheric turbulence on the imaging process. If the AO system is working properly, the light to the science camera should have an almost flat wavefront, as if there where hardly any distortions. In this way, the corrected image can be recorded without being spread out when using long exposure times. By using AO, large ground-based based telescopes may reach close to diffraction limited performance in the near infrared (Rousset et al. 1990; Beckers 1993). Figure 1.3 provides an example of the gain in angular resolution that can be achieved by AO. It shows an image of a star-burst galaxy obtained with the Canada-France-Hawaii Telescope (CFHT), atop of the Mauna Kea volcano, Hawaii, with and without AO. The AO corrected image has a much higher resolution and shows more details.
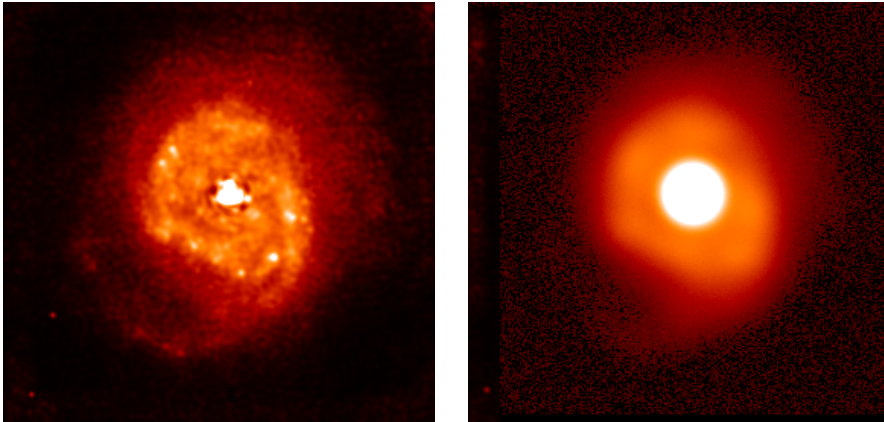
Figure 1.3: NGC7469, a galaxy belonging to the class known as star-burst galaxies. The left image has been obtained with AO compensation and has a resolution of 0.13 arcsec. The right panel shows the image that would have been obtained without AO compensation and has an resolution of 0.7 arcsec. Image Credit: Canada-France-Hawaii Telescope (CFHT).

In astronomy, it is common to make a distinction between adaptive optics and active optics. The latter term is used to refer to the technique of compensating static and low frequency errors in the primary mirror geometry of the telescope itself. Examples of the compensated error sources include mechanical errors in-

troduced by gravitational sag and wind forces at different telescope inclinations. Active optics is also used to align the different segments in telescopes with a segmented primary mirror. In contrast to adaptive optics, active optics operates at a fairly low temporal frequency in the order of $0.05\,\mathrm{Hz}$ or less.

1.1.3   Early developments of adaptive optics

The idea of AO was already known in the fifties. In 1953, Babcock (Babcock 1953) proposed the usage of an Eidophor[1] in feedback configuration to compensate atmospheric wavefront distortions. However, the demanding technical requirements of this idea delayed the realization of AO systems for several decades. In the early seventies, the US Air Force built the first operational AO systems both for improved imaging of satellites and for the projection of high energy laser beams onto missiles. In the late 1980's, the European Southern Observatory (ESO) and the US National Optical Astronomy Observatories (NOAO) decided to start a program for the development of AO for infrared astronomy.

In 1989, ESO unveiled the first non-military AO system for large telescopes. The Come-On system, developed through collaboration between ESO, ONERA and astronomical institutes in France, was the first system to provide diffraction-limited astronomical images in tests on the $1.52\,\mathrm{m}$ telescope (Rousset et al. 1990). The system was later put into service on the ESO $3.6\,\mathrm{m}$ telescope at la Stilla (Chili). Nowadays AO is a well established technique. By the mid-1990s virtual all large telescopes had either been retrofitted or had integrated AO system in their design. So most modern observatories are able to offer their astronomers an AO system.

1.1.4   Improved sky coverage by laser guide stars

An important complication in astronomical imaging is that the amount of light available for wavefront sensing is typically very small. If the science object has a stellar magnitude (Roddier 1999; Hardy 1998) of 13 or more, it is too faint for the WFS to function properly. Under these circumstances, AO can only be used if there is a bright guide star close to the science object that can be used as alternative target for wavefront sensing. The idea behind this is that light from this reference source experiences almost the same turbulence as the light from the faint science object. This implies that the WFS measurements from the guide star can be used to correct the wavefront distortions associated with the science image.

For an reference source to be a suitable guide star, it should be sufficiently close to the science object to ensure that the WFS measures approximately the same distortions. A quantitative measure for the admissible angular separation between guide star and science object is the so called isoplanatic angle. The isoplanatic angle θ_0 is defined as angle between two light beams so that they experience a mean-square non-common path phase difference of $1\,\mathrm{rad}^2$. Realistic values of the

---

[1]An Eidophor is an old-fashioned television projection system that was used to create theater-sized images. Its working is based on electrostatic deformation of an oil surface by electron bombardment.

isoplanatic angle are in the order of a few arc seconds. This means that still only 0.1 to 1% of the stars in the sky are close enough to a bright star to benefit from AO (Rigaut and Genrdon 1992). Unfortunately, many interesting science objects lack therefore a sufficiently bright natural guide star.

Similar limitations existed for many military applications of AO. To deal with this issue and improve the sky coverage, the US Air Force started to develop laser-beacon assisted AO in the 1980's. In laser-beacon assisted AO (see e.g. Fugate et al. 1991), a laser beam is projected into the air to create an artificial guide star or laser guide star (LGS), in the vicinity of the science object. Two strategies are used for creating artificial guide stars. The first concept is based on back-scattering of pulsed laser light by sodium resonance in the high mesosphere at an altitude of approximately $90\,\mathrm{km}$, while the second approach uses Rayleigh diffusion of particles and molecules in the low stratosphere at 10 to $20\,\mathrm{km}$.

An important disadvantage of the use of LGS is that it leads to a significant increase of the complexity and costs of the AO system. Especially the need for expensive high performance and high quality lasers, which produce an enormous amounts of heat, is complicating matters. Also the problem of projecting the laser beam to the sky with minimal distortions is a challenging task. Another complication of laser assisted AO is that LGS cannot be used as an absolute reference to stabilize the image motion. The laser beam projected from the ground is also effected by turbulence, resulting in random displacements of the guide star position. Since the absolute position of the LGS is unknown, it is impossible to measure the global tilt and tilt modes. To compensate these low order modes, an additional faint natural reference guide star is needed. The source brightness required for tilt tracking (typically around 16 on the magnitude scale), however, is much less than that needed for compensating higher order modes.

Also the effect known as conical anisoplanatism is posing a physical limitation on the use of LGSs. Since the artificial guide star is located at a relatively low altitude, the back-scattered light forms a conical beam which samples a different part of the atmosphere than the light coming form a distant star. The error caused by conical anisoplanatism scales with the telescope diameter and the observing wavelength. Consequently, LGS assisted AO becomes unusable for telescopes larger than $8\,\mathrm{m}$ diameter in the visible (Ragazzoni et al. 2000). To overcome this problem, a constellation of multiple guide stars has to be used. At this moment, several large ground-based telescopes have an operational LGS system.

1.1.5 Extension to multi-conjugate adaptive optics

Besides the achievable angular resolution and the sky coverage, also the field of view (FoV) is an important property in astronomical telescopes. Because of anisoplanatism, the AO performance is not uniform over the sky but degrades with the angular distance from the guide star. Since the turbulence is distributed in a volume above the telescope, light waves from different directions experience different distortions so that the wavefront corrections derived from WFS measurements of a guide star in one part of the sky will only compensate turbulence in

a small FoV. To improve the FoV of classical AO systems, a technique known as multi-conjugate adaptive optics (MCAO) has been proposed.

The basic concept of MCAO has already been proposed as early as 1975 (Dicke 1975), and received renewed attention after the papers of Beckers (1989) and Ellerbroek (1994). Whereas classical AO uses a single DM and WFS to compensate the wavefront distortions in a single plane conjugated to the telescope aperture, MCAO tries to compensate the turbulence in the three-dimensional volume above the telescope. This is achieved by using multiple DMs and WFSs. By combining the WFS measurements obtained from guide stars in different directions it is possible to determine the wavefront distortions at different heights. The process of converting the WFS measurements in a three-dimensional turbulence profile is called turbulence tomography. With multiple DMs conjugated at turbulent layers at different height, MCAO is able to correct the wavefront distortions in a three-dimensional fashion. This results in a more uniform correction that is less sensitive to the direction of observation over a much larger FoV.

Ragazzoni et al. (2000) has experimentally demonstrated the principle of atmospheric turbulence tomography by collecting WFS data from a constellation of three off-axis natural guide stars ≈ 15 arcsec from a central star. The measurement data from the off-axis stars has been used to compute the wavefront distortions in the direction of the central star. By comparing the computed distortions with real WFS measurements from the central star, it is shown that tomographic reconstruction reduces the wavefront anisoplanatic error by a factor three compared to the wavefront reconstructed by simply taking the arithmetic average over the neighboring guide stars. Other studies to access the feasibility of MCAO have shown that for an $8\,\text{m}$ telescope with three mirrors and five LGSs, a ten-fold increase in the area of the compensated FoV should be possible (Ellerbroek and Rigaut 2000).

With MCAO being based on multiple DMs, WFSs and LGSs it is far more complex than a classical AO system. MCAO is still in an experimental stage, and there is a lot of research effort in developing the necessary concepts and technical expertise to demonstrate it on sky. A MCAO system for the Gemini South telescope is currently under development and expected to be commissioned in 2007. Meanwhile, ESO is working on the multi-conjugate adaptive optics demonstrator (MAD) to investigate the feasibility of different MCAO techniques for the next generation of very large telescopes (VLTs).

1.1.6   Non-astronomic application of adaptive optics

Now that AO has reached a certain level of maturity, it is starting to have spin-offs beyond the traditional applications in astronomy and the military. The technology has found its way in several new developments in medicine, manufacturing and laser communication. One of the first commercial applications is in the field of ophthalmology. Here AO is used in laser eye surgery and to improve the resolution of images of the human retina for early detection of eye diseases.

Besides correcting for optical distortions, AO can be used to alter the characteristics of an optical system. The objective here is not to improve the image

quality, but rather to adapt the wavefront to achieve a particular effect. Some of the emerging AO applications, such as femto-second pulse shaping and accurate focus control of laser beams, are based on this idea.

Furthermore, AO is finding its way in several three-dimensional in vivo and in vitro imaging applications, systems for vision assessment, in applications of laser communications and material processing, in confocal microscopy, optical data storage and defense applications such as laser beam delivery and surveillance applications. For an extensive overview of medical and industrial applications of AO, as well as an overview of its market prospects, the reader is referred to Greenway and Burnett (2004).

1.2 Imaging through atmospheric turbulence

Both for the design and performance evaluation of AO systems, it is important to have a basic understanding of the physical processes that cause turbulence. A good appreciation of the structure of turbulence is particularly useful as it shows that the wavefront aberrations are not completely random, but can be predicted to a certain extent. Furthermore, a physical description of turbulence provides more insight in the way design parameters, like the WFS resolution, the number of DM actuators and the control bandwidth, influence the overall performance.

There is a wealth of literature on the characterization of atmospheric turbulence and its influence on image formation. It is beyond the scope of this thesis to provide a complete overview. In this section some of the main results that are useful in motivating the chosen control strategy will be summarized. For a more extensive overview, the reader is referred to the standard works of Tatarskii (1971); Goodman (1985); Léna (1997); Hardy (1998); Roddier (1981) and (Roddier 1999). At the end of the section we will relate the image quality of a turbulence degraded image to a convenient performance measure for AO.

1.2.1 Wavefront phase-distortions

Light propagating through the atmosphere is affected by random fluctuations in the refractive index. The physical source of these inhomogeneities is the turbulent mixing of air of different temperatures. This effect is most profound at the interface of different wind-layers, where wind shear forms the driving force. Before entering the atmosphere, light from a distant astronomical object forms essentially plane waves. However, in the atmosphere, light propagating through regions of high refractive index will be delayed with respect to other regions. When the light arrives at the telescope, the wavefront is no longer flat but severely distorted.

To a first approximation, the so called near-field approximation, the optical path differences Δl in the telescope aperture can be expressed as

$$\Delta l(\rho) = \int n(\rho, z) \, \mathrm{d}z, \tag{1.2}$$

where z is a coordinate along the line of sight, $n(\rho, z)$ denotes refractive index and $\rho \in \mathbb{R}^2$ is the spatial position in the aperture. The refractive index can be approximated as the sum of a wavelength dependent part $n(\lambda)$ and a randomly fluctuating part $n_f(\rho, z)$ dependent on the temperature and pressure variations, i.e. $n(\rho, z) = n(\lambda) + n_f(\rho, z)$. Since the fluctuating $n_f(\rho, z)$ part is fairly wavelength independent, also the shape of the DM needed to compensate the optical path length differences of light with different wavelengths is the same. This is particularly important as it implies that the WFS can be operated at a different wavelength than the wavelength used for the observations.

The effect of the wavefront distortions on the image formation process, on the other hand, shows a strong wavelength dependence. The image quality is not determined by the absolute optical path differences but by their relative impact, that is by the optical phase differences. The phase and the optical path length are related as $\phi = k\Delta l$, where $k = {}^{2\pi}/\lambda$ is the wave number and λ denotes the wavelength. Hence, the phase fluctuations are inversely proportional with the wavelength. This explains why atmospheric turbulence has a less detrimental effect on the imaging performance at longer wavelengths.

1.2.2   Spatial structure of atmospheric turbulence

Atmospheric turbulence is a random process that can only described in terms of statistical quantities. The theoretical framework for understanding imaging through turbulence relies heavily on the model of the velocity of motion in a turbulent medium as proposed by Kolmogorov (1960). This model assumes that energy is added to the medium in the form of large-scale inhomogeneities with a characteristic size L_0, the so called outer scale. In the case of atmospheric turbulence, the ultimate energy source for generating these inhomogeneities is solar heating, which leads to kinetic energy in the form of convection and wind shear.

Kolmogrov suggested that these large scale disturbances successively break down, transferring the kinetic energy into smaller and smaller structures. If the characteristic size of the turbulent vortexes reaches the level l_0 for which it is no longer possible to sustain the turbulence, the energy is dissipated as heat by viscous friction. For the process to continue at a stable rate, the energy transferred to each of the disturbances with a characteristic size $l_0 \leq l \leq L_0$ should equal the energy dissipation. Under the assumption that the atmosphere can be considered as locally homogeneous and isotropic, this energy balance can be used to derive a relation between the velocity fluctuations and the characteristic disturbance size. This gives rise to a statistical description of spatial distribution of the velocity fluctuations in a turbulent medium.

In describing the phase fluctuations it is commonly assumed that a passive additive, i.e. a quantity that does neither affect the dynamics nor the composition of the turbulent medium, has the same spatial distribution as the velocity fluctuations. This implies that also the refractive index fluctuations follow a Kolmogorov power law. By performing the integration (1.2) over the refractive index it is now possible to determine the spatial distribution of the phase distortions in the tele-

scope aperture. The spatial distribution of the phase distortions is usually characterized by the structure function, which is the variance of the phase difference between two points separated by a vector r, i.e.

$$D_\phi(r) = \left\langle |\phi(\rho, \cdot) - \phi(\rho + r, \cdot)|^2 \right\rangle, \tag{1.3}$$

where $\langle \rangle$ denotes the ensemble average over different realization of $\phi(\rho, \cdot)$ and $\cdot$ a single time-instant. The structure function was introduced by Kolmogorov to describe non stationary random functions with a slowly varying mean. Because the difference between two nearby phase points is not affected by the slowly varying mean, the increment can be considered stationary whereas the phase itself is not. Structure functions are also useful as one is not interested in the absolute phase, but only in the relative fluctuations over the aperture.

Since the turbulence is assumed to be homogeneous and isotropic, the structure function does not depend on the absolute positions ρ and $\rho + r$, but only on their mutual distance $|r|$. For Kolmogorov turbulence, the phase structure function over the aperture is given by

$$D_\phi(r) = 6.88 \left(\frac{|r|}{r_0} \right)^{5/3}, \tag{1.4}$$

where

$$r_0 = \left[0.423 \left(\frac{2\pi}{\lambda} \right)^2 \sec(\zeta) \int C_n^2(h) \mathrm{d}h \right]^{-3/5}, \tag{1.5}$$

is the Fried parameter, ζ the angular distance of the source from zenith and $C_n(h)$ the refractive index structure coefficient characterizing the turbulence strength at a height h above the ground. An interesting interpretation of r_0 is that it specifies the aperture diameter for which the mean-square wavefront phase error is approximately $1\,\mathrm{rad}^2$. Typical values for r_0 in the visible range from less than $5\,\mathrm{cm}$ in strong daytime turbulence to over $20\,\mathrm{cm}$ at good sites at night.

The phase structure function plays an important role in characterizing the effect of atmospheric turbulence on the image formation. Under the assumption that the wavefront distortions are Gaussian random process, it is related to the coherence function as

$$\Gamma(\kappa) = \mathrm{e}^{-1/2 D_\phi(\lambda f \kappa)}, \tag{1.6}$$

where f is the focal length of the imaging system and κ denotes the spatial frequency in the aperture. For long exposure images, the coherence function can be interpreted as the optical transfer function (OTF) (see e.g. Hecht 1987) of the atmosphere. If the turbulence rather than the telescope is limiting the imaging performance, the Fourier transform of the coherence function defines the seeing limited point spread function (PSF) in the focal plane. The turbulence limited PSF, the so called seeing disk, has a full width of half maximum (FWHM) of $0.98\lambda/r_0$, which corresponds in good approximation to the diffraction limited performance of a telescope with a diameter D of r_0. Consequently, a telescope with $D \gg r_0$ is seeing limited whereas a telescope with $D \ll r_0$ is limited by diffraction.

Furthermore, equation 1.4 allows one to calculate an estimate of the wavefront fitting error, caused by the inability of the DM to assume any arbitrary shape. The mean-square fitting error is given by

$$\sigma_f^2 = a_f \left(\frac{d}{r_0} \right)^{5/3},$$ (1.7)

where d is the characteristic size of the inter-actuator spacing and a_f is a fitting error coefficient depending on the influence function of the corrector. Wavefront phase fluctuations with a variance below a threshold of about 1 radian have little effect on the image quality and therefore typically do not need to be compensated. Using this threshold as a rule of thumb, equation (1.7) can be used to estimate the required d/r_0 and hence the minimal number of mirror actuators. Likewise the size of anisoplanatic error can be estimated from the structure function.

1.2.3 Temporal evolution of the wavefront

The Kolmogorov model for turbulence describes only the spatial distribution of the wavefront distortions. For control design, however, it is also useful to have some insight in the time scale on which the distortions evolve. The common way to model the temporal evolution of the wavefront is by using the Taylor hypothesis of frozen turbulence (Taylor 1938).

The Taylor hypothesis is based on the assumption that the atmospheric turbulence is concentrated in a number of discrete layers which each move with their own wind speed and direction over the telescope aperture. Since the lifetime of the turbulent refractive index inhomogeneities is assumed to be much longer than the time needed for the layers to cross the telescope aperture, the different layers can be considered as frozen phase-screens. With the spatial distribution of the layers being fixed, the temporal evolution of the wavefront is entirely due to the wind transport. Considering a single layer of turbulence moving with a wind velocity $v \in \mathbb{R}^2$, the phase distortion at point the ρ and time $t + \tau$ can be related to the phase at time t as

$$\phi(\rho, t + \tau) = \phi(\rho - v\tau, t).$$

The temporal difference τ is thus transformed into a spatial difference $|r| = v\tau$, which is characterized by (1.4). When the wavefront is affected by several layers, the temporal evolution of a single point in the aperture can be approximated as a phase screen propagating with a turbulence weighted velocity $\bar{v}$, defined as

$$\bar{v} = \left(\frac{\int C_n^2(z)|v(z)|^{5/3}\mathrm{d}z}{\int C_n^2(z)\mathrm{d}z} \right)^{3/5},$$ (1.8)

where the integration is performed over the line of sight. As a result, the temporal phase structure function $D_\phi(\tau)$ is obtained by substituting $|r| = \tau\bar{v}$ in equation (1.4). The characteristic time for the turbulence to move over a distance τ_0, i.e. $\tau_0 = r_0/\bar{v}$ is called the turbulence coherence time. Typical values for the wind speed are in the order of $10 \, \mathrm{m/s}$ with peak values up to $40 - 50 \, \mathrm{m/s}$.

The validity of the frozen flow hypothesis depends on the time scales that are considered. Whereas the model is a good approximation for short time scales, it cannot be valid for long time scales because of the temporal evolution of the turbulence itself. This effect is also referred to as boiling. By analyzing data from the $1.5\,\mathrm{m}$ and $3.5\,\mathrm{m}$ telescopes at the Starfire Optical Range, it has been demonstrated that the frozen flow hypothesis is an accurate description of the temporal development of atmospheric turbulence on time scales shorter than approximately to 10 to $20\,\mathrm{ms}$ (Schöck 1998; Schöck and Spillar 2000).

Besides the structure function also the power spectrum is often used to describe the phase fluctuations. The power spectrum $\Phi(f)$ and the structure function $D_\phi(\tau)$ are related through the Wiener-Khinchin theorem. The (temporal) power spectrum corresponding to the temporal phase structure function $D_\phi(\tau)$ is given by (Conan et al. 1995; Glindemann et al. 2000)

$$\Phi(f) = 0.077 r_0^{-5/3} \frac{1}{\bar{v}} \left(\frac{f}{\bar{v}} \right)^{-8/3}. \tag{1.9}$$

This power spectrum can be used to derive an estimate of the temporal errors introduced by pure time delays and the finite bandwidth of the control system. Greenwood (1977) has shown that the temporal wavefront error as a result of bandwidth limitations for an AO system with a first-order temporal feedback controller or a low-pass controller with an infinitely sharp cut-off, is given by

$$\sigma_t^2 = a_t \left(\frac{f_G}{f_S} \right)^{5/3}, \tag{1.10}$$

where f_S is the bandwidth of the feedback system and $f_G = 0.427 \bar{v}/r_0$ is a characteristic frequency known as the Greenwood frequency. The scaling constant $a_t \in \mathbb{R}$ depends on the type of feedback controller and is equal to 1 for the first-order controller. Finally, the wavefront error caused by a pure time delay $\tau_d \in \mathbb{R}$ in the system is given by $\sigma_d^2 = 28.4(\tau_d f_G)^{5/3}$ (Fried 1990).

1.2.4 Formulation of the AO control objective

The quality of a turbulence degraded image is often expressed in terms of the Strehl ratio. The Strehl ratio is defined as the peak intensity of the image of a point source, normalized to the diffraction limited peak intensity (Tyson 1998; Hardy 1998). This is an useful and sensitive performance measure as any wavefront error is expected to diffract light away from the center of the image, thereby reducing the peak intensity. For an optical system with a (residual) phase distortion $\epsilon(\rho, \cdot)$, the Strehl is given by

$$S = \left\langle \frac{1}{A} \left| \iint e^{ik\epsilon(\rho, \cdot)} \, d\rho \right|^2 \right\rangle, \tag{1.11}$$

where the integration extends over the opening aperture and A denotes the light collecting area. From equation (1.11) it is clear that Strehl ratio for an undistorted wavefront, i.e. $\epsilon(\rho, \cdot) = 0$, is equal to $S = 1$. In the presence of any wavefront

aberration, the Strehl ratio will be less than 1. The general objective of AO can hence be formulated as maximizing the Strehl.

From a practical point of view, the above expression for Strehl is not very convenient. Evaluating the Strehl requires accurate knowledge of the wavefront $\epsilon(\rho, \cdot)$ over the entire aperture, which is generally not available. Because of the random nature of turbulence, at most the statistical properties of the wavefront are known. However, if the wavefront distortions are not too large, it is possible to relate the Strehl to the variance of the phase error over the aperture σ_ϵ^2. By expanding the exponential in (1.11) and preserving the first few terms, the Strehl can be approximated as (Born and Wolf 1999)

$$S \approx 1 - \sigma_\epsilon^2 \approx \exp(-\sigma_\epsilon^2) \tag{1.12}$$

where

$$\sigma_\epsilon^2 = \frac{1}{A} \left\langle \iint \epsilon^2(\rho, \cdot)\mathrm{d}\rho - \left(\iint \epsilon(\rho, \cdot)\mathrm{d}\rho \right)^2 \right\rangle. \tag{1.13}$$

The above expressions provide a reasonable description of S for phase errors with variances up to respectively 0.4 and $4\,\mathrm{rad}^2$. The latter of the two approximation is sometimes called the extended Maréchal approximation.

With the Strehl being a strictly decreasing function of σ_ϵ^2, the approximations in (1.12) suggest that the objective of maximizing the AO imaging quality can be replaced by that of minimizing the residual phase variance. An analysis performed by Herrmann (1992) confirms that even though the above approximations might be rather crude, the minimum-variance wavefront leads indeed to the maximum Strehl. Since the objective of maximizing the Strehl is rather awkward, the AO control objective is usually reformulated as that of minimizing σ_ϵ^2.

1.3 Classical AO control approach

This section provides a concise overview of what we will refer to as the classical AO control approach. The considered control strategy is still the most widespread control approach used in AO systems for ground-based telescopes (see e.g. Brase et al. 1998; van Dam et al. 2004). For this reason, the outlined control procedure will form the baseline for comparing the control strategies developed in this thesis. The shortcomings of the classical AO control approach will be used in Section 1.6 to motivate the followed research strategy. More background information on most of the theory introduced in this section can be found in the standard works Tyson (1998); Roddier (1999); Hardy (1998).

1.3.1 A closer look at the AO control system

In the previous sections we have seen that the general objective of AO is to suppress the wavefront distortions introduced by the turbulent atmosphere. Furthermore, it has been pointed out that optimizing the imaging performance can be

achieved by minimizing the residual phase variance. The AO control problem should therefore be directed to the design of a controller that achieves this goal. Figure 1.4 provides a schematic representation of a standard AO control loop consisting of a WFS, a DM and a controller C, each indicated by a shaded block. The block diagrams within the shaded blocks show the mathematical structure of the models of the components as used in the classical AO control approach.
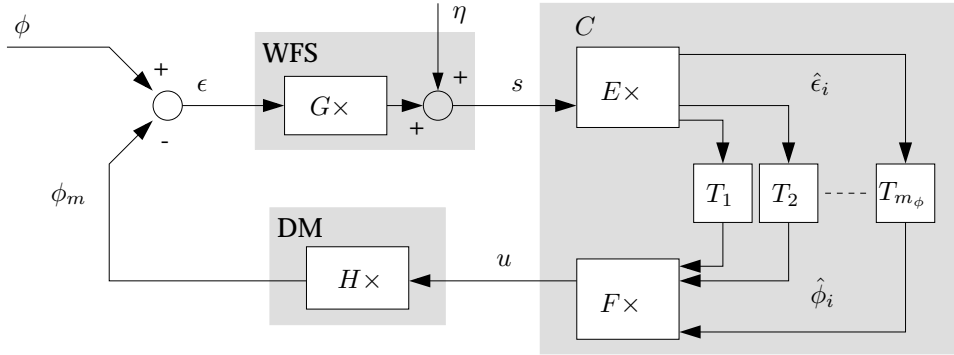
Figure 1.4: Schematic representation of a standard AO control loop consisting of a WFS, a DM and a controller C. The controller has the structure of what we refer to as the classical AO control law.

An important complication in controlling the AO system is that it is not possible to directly measure the residual phase error. In contrast to what might be expected from the name, a WFS typically provides a signal that is a measure of the slope or curvature of the wavefront. Different types of WFS, based on different physical principles, have been developed including the curvature WFS, the pyramid WFS, the shearing interferometer and the Shack-Hartmann sensor. Of these WFSs, the Shack-Hartmann sensor is the most commonly used. Although in this thesis we will restrict our attention to the Shack-Hartmann type of WFS, the other types of WFS can be handled in a similar framework.

The working principle of a Shack-Hartmann WFS is illustrated in Figure 1.5. Its main component is a grid of identical lenses, the so called lenslet array, that segments the telescope aperture into a number of sub-apertures. Each of the sub-apertures focuses the incident light into a spot on a charged-coupled device (CCD). When the incoming wave-front is plane (left-hand side of figure), each spot is located exactly under the geometrical center of its respective sub-aperture, defining a regular grid. Any wavefront aberration (right-hand side), causes the spots to depart from their reference positions. The displacements of each of the spots is proportional to the averaged wavefront slope across the corresponding sub-aperture. By determining the displacement of the spots, the Shack-Hartmann sensor is hence able to measure the wavefront slopes. The current state-of-the-art telescopes have AO systems with in the order of a few hundred Shack-Hartmann sub-apertures and DM actuators (e.g. the Keck telescope has a WFS with 304 sub-apertures and a DM with 349 actuators van Dam et al. (2004)).

The classical AO control approach decomposes the controller in a cascade of
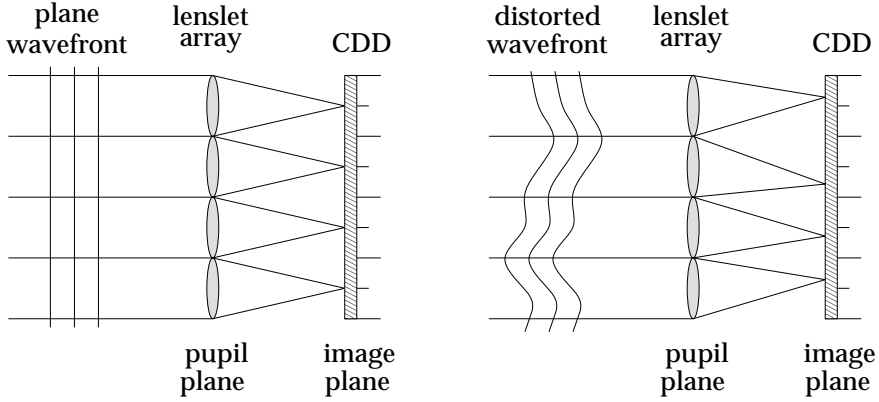
Figure 1.5: Schematic representation of a Shack-Hartmann wavefront sensor for a plane wavefront (left) and a distorted wavefront (right).

static mapping and a series of parallel single-input single-output (SISO) feedback loops. For the ease of discussion we will temporally assume that the parallel feedback loops are in between two of such static mappings. This situation is depicted in Figure 1.4. To deal with the discrepancy between measurement and control objective, the classical AO control approach includes a separate wavefront reconstruction step. The wavefront reconstruction problem is solved in a static sense and corresponds to the first matrix multiplication, denoted by E, in Figure 1.4. Since the AO system is operated in closed loop, the wavefront reconstructed from the WFS measurements s is an estimate of the residual phase error ϵ. This implies that the estimated wavefront is not the wavefront correction that has to be applied by DM but can be seen as an increment to the current correction. It is the task of the dynamic temporal compensator consisting of the parallel feedback loops $T_1, \ldots, T_{m_\phi}$ to estimate the required wavefront correction out of these increments. The temporal compensator is hence responsible for achieving stability and closed loop performance. The second static mapping, denoted by F, projects the estimated phase correction on the actuator space. The wavefront reconstruction and the temporal compensator will be discussed in the subsequent subsections.

1.3.2 Finite-dimensional representation of the wavefront

The atmospheric wavefront distortions $\phi(\rho, \cdot)$, the applied wavefront correction $\phi_m(\rho, \cdot)$ and the residual wavefront $\epsilon(\rho, \cdot)$ are continuous functions of time and space and can therefore be regarded as an infinite-dimensional signals. The WFS on the other hand produces only a finite-dimensional measurement signal $s(k) \in \mathbb{R}^{m_s}$ at discrete-time instants $t = kT$, where T denotes the sampling time. A Shack-Hartmann WFS for instance, probes the incoming wavefront distortion with a finite spatial resolution set by sub-aperture spacing. The dimension of the WFS signal in this case is twice the number of active spots. Given only the WFS signal $s(k)$ it is impossible to reconstruct the infinite-dimensional wavefront distortions.

For this reason it is common practice to represent the wavefront distortion over the aperture by means of a discrete-time finite-dimensional vector signal.

Depending on the way the finite-dimensional vector signal is obtained, we distinguish zonal and modal representation. A zonal representation refers to the spatial sampling of the wavefront by considering the local deformation at specific positions in the aperture. Let the finite-dimensional vector representation of the wavefront $\phi(\rho, k)$ be denoted by $\phi(k)$ and let the sampling locations be denoted by $\rho_i \in \mathbb{R}^2$, $i \in \{1, \ldots, m_\phi\}$, then the ith component of $\phi(k)$ is obtained as $\phi_i(k) = \phi(\rho_i, k)$. Note that to avoid an abundance of notation we use the same symbol for the continuous and sampled wavefront. The two are distinguished from each other by their arguments. The modal representation of the wavefront is on the other hand obtained by expanding the phase on a set of basis functions

$$\phi(\rho, k) = \sum_{i=0}^{m_\phi} \phi_i(k) Z_i(\rho), \tag{1.14}$$

where the coefficients $\phi_i(k)$ form the components of the vector signal $\phi(k)$ and $Z_i(\rho)$, $i \in \{1, \ldots, m_\phi\}$ is a set of suitable basis functions. Frequently used basis functions include the Zernike basis, the Karhunen-Loéve basis or a basis composed of the eigenmodes of the mirror. Whether a zonal or modal representation is used is irrelevant for the remaining discussion. It is however assumed that the 2-norm of the obtained vector signal $\phi(k)$ provides a good approximation of the phase variance over the aperture.

1.3.3   Static reconstruction and the actuator projection step

In the classical AO control approach, the wavefront reconstruction step and the projection of the required wavefront correction onto the actuator space are both considered in a static setting. Using the finite-dimensional vector representation of the wavefront, an incoming wavefront $\phi(k)$ and the corresponding (open-loop) WFS measurements can be related as

$$s(k) = G\,\phi(k) + \eta(k) \tag{1.15}$$

where $G \in \mathbb{R}^{m_s \times m_\phi}$ is the so called geometry matrix and $\eta(k)$ represents the measurement noise, which is assumed to be zero-mean, white and uncorrelated to $\phi(k)$. The precise form of G depends on the modal basis functions used in the finite-dimensional vector representation of $\phi(\rho, \cdot)$, or on the locations of the sampling points ρ_i when using a zonal representation. In fact for a Shack-Hartmann WFS, the positions ρ_i are typically chosen in such a way that the WFS slope measurements can easily approximated in a finite-difference setting. Common configurations for defining the phase reconstruction points include the Hudgin and Fried geometry (see Figure 1.6). Let the static relation between $s(k)$ and the wavefront estimated $\hat{\phi}(k)$ be given by $\hat{\phi}(k) = Es(k)$, then the wavefront reconstruction problem can be formulated as

$$E = \arg\min_{\widetilde{E}} \left\langle \left\| \phi(k) - \widetilde{E}s(k) \right\|_2^2 \right\rangle, \tag{1.16}$$

where $\langle \rangle$ denotes the conditional expectation over different realizations of $\phi(k)$, and $s(k)$ is given by equation (1.15). Note that the objective of minimizing the expected value of the difference between $\phi(k)$ and $\hat{\phi}(k)$ squared is in agreement with the objective of minimizing the variance of the wavefront estimation error.
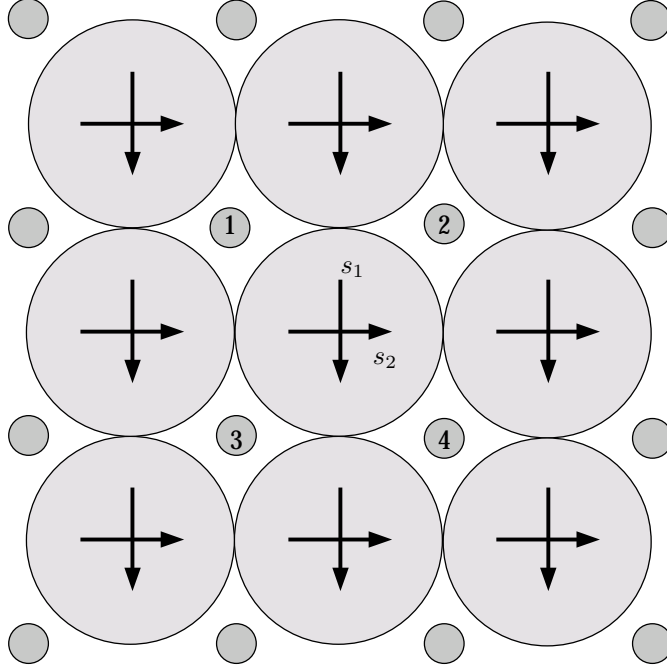
Figure 1.6: Schematic representation of the Fried geometry. The large circles represent the sub-apertures of the Shack-Hartman WFS and the small circles mark the points of reconstructed phase. The arrows and denote averaged slope over the aperture. Using the Fried geometry, the slopes s_1 and s_2 are related to the phase points using the finite-difference approximations $s_1 = (\phi_3 + \phi_4 - \phi_1 - \phi_2)/2l$ and $s_2 = (\phi_2 + \phi_4 - \phi_1 - \phi_3)/2l$, where l is the pitch size and $\phi_i, i \in \{1, 2, 3, 4\}$ denote the phase in corresponding phase points.

The earliest approaches to wavefront reconstruction (Fried 1977; Hudgin 1977; Herrmann 1980) neglected the stochastic nature of the wavefront distortions. This implies that the wavefront distortion $\phi(k)$ in (1.16) is seen as a deterministic signal. Under this simplifying assumption the wavefront reconstructor reduces to

$$E = G^T (GG^T)^{-1}.$$

This least-squares type of wavefront reconstructor has been successfully applied in many AO systems and is still applied today. Even though it has a reasonable performance for high signal to noise rations, it is rather sensitive to measurement noise. To improve the wavefront reconstruction performance under low light level conditions, it is useful to return to the stochastic setting of equation (1.16). This has lead to the following minimum-variance of maximum a posteriori (MAP) es-

timator Wallner (1983); Bakut et al. (1994); Law and Lane (1996)

$$E = C_\phi G^T \left(G C_\phi G^T + C_\eta \right)^{-1},$$ (1.17)

where $C_\phi \doteq \mathcal{E}\{\phi(k)\phi^T(k)\}$ and $C_\eta \doteq \mathcal{E}\{\eta(k)\eta^T(k)\}$ denote the covariance matrices of $\phi(k)$ and $\eta(k)$, respectively. From the above expression it is clear that the minimum-variance reconstructor can be interpreted as a regularized pseudo-inverse where the regularization reduces the sensitivity to poorly sensed modes. Up to now, the wavefront reconstruction problem has been considered in an open-loop setting. However, when used in an AO system, the WFS does not measure the open-loop wavefront distortions $\phi(k)$ but the residue $\epsilon(k)$. This implies that in (1.17) the covariance matrix C_ϕ should be actually replaced by the residual wavefront variance $C_\epsilon \doteq \{\epsilon(k)\epsilon^T(k)\}$. In the classical AO control approach this modification in the wavefront statistics is usually neglected.

The projection of the required wavefront correction on the actuator space is also described by a static matrix multiplication, i.e. $u(k) = F\hat{\phi}(k)$. In estimating the required actuator inputs, it is usually assumed that the DM can be modeled as $\phi_m(k) = Hu(k)$, where H is the so called influence matrix. Considering this DM model, the problem of finding the projection matrix F that minimizes the wavefront fitting error is given by

$$F = \arg \min_{\widetilde{F}} \left\| \hat{\phi}(k) - H\widetilde{F}\hat{\phi}(k) \right\|_2^2.$$ (1.18)

In contrast to the wavefront estimation and reconstruction problem in equation (1.16), this is a deterministic optimization problem and its solution is given by $F = (H^T H)^{-1} H^T$. Furthermore, note that since the temporal feedback $T_1, \ldots, T_{m_\phi}$ loops in the classical AO control approach are all the same, the static multiplications F and G can be combined in a single matrix $R = FE$. This matrix is usually referred to as the control matrix. By combining the matrices, the classical AO control law can be implemented as a single matrix-vector multiply followed by a temporal compensator.

1.3.4 Temporal compensator design

As outlined above, the task of the temporal compensator is to update the estimate of the required wavefront correction $\hat{\phi}(k)$ on the basis of the estimate of residual wavefront error $\hat{\epsilon}(k)$. Furthermore, we have seen that in the wavefront reconstruction step and in the projection on the actuator space typically all dynamics are neglected. On the other hand, in temporal compensator design it is impossible to neglect the AO system dynamics. In any AO system there will be a finite delay between measurement and correction. Pure time delays, for instance, are caused by CCD read out times in the WFS and the time needed for processing the WFS image to produce the measurement signal $s(k)$. Typical values for the time required to read out the CCD detectors are in the order of $0.5 - 1\,\text{ms}$. The time delay between the middle of the exposure period and the availability of the reconstructed wavefront is typically about 2 frames for astronomical AO systems.

In the temporal compensator design one has to pay attention to this type of delays. The temporal compensator design is responsible for achieving stability and closed loop performance in spite of the presence of any system dynamics.

The temporal compensator used in classical AO control consists of a series of parallel feedback loops. This decomposition is based on the implicit assumption that the spatial and temporal dynamics can be decoupled. It is assumed that each component of the required wavefront correction $\hat{\phi}_i(k)$ can be obtained as

$$\widehat{\phi}_i(k) = T_i(z)\widehat{\epsilon}_i(k), \quad i \in \{1, \ldots, m_\phi\},$$

where the sub-index i is used to denote the ith component of the corresponding vector. Furthermore, it is commonly assumed that all feedback loops are the same, i.e. $T_i(z) = T(z)$. The design of the scalar feedback compensator $T(z)$ typically has its roots in classical control design. Commonly applied control structures include first-order lag filters, proportional-integral (PI) controllers and Smith predictors. Given the control structure, the control design problem boils down to the choice of the appropriate control parameters, which determine the effective bandwidth. For a bright guide star, the bandwidth requirements are set by the turbulence. The effect of the bandwidth on the residual wavefront error for a first-order compensator has been extensively studied in literature (Greenwood and Fried 1976; Greenwood 1977; Tyler 1994). To achieve a good performance, the bandwidth should be chosen sufficiently high without loss of stability.

Things get more involved for faint guide stars where measurement noise becomes an important issue. In this case the choice of the control bandwidth is a trade off between disturbance rejection and measurement noise amplification. A large control bandwidth is attractive for reducing temporal latency errors, but also increases the propagation of measurement noise. The optimal control bandwidth is the bandwidth that minimizes the sum of these error contributions and depends on both the signal-to-noise ratio (SNR) of the WFS measurements and the prevalent turbulence conditions, while retaining stability. A simulation study by (Brigantic et al. 1998) shows that under low light levels the choice of the bandwidth is mainly determined by minimizing the measurement noise errors and is rather insensitive to the prevalent seeing conditions.

The classical control approach has already a reasonable performance under moderate conditions. This can be explained by the fact that the turbulence has a power spectrum in which the low frequency components are dominant. A controller with integrating action gives rise to a high loop gain in this frequency range, which implies a good disturbance rejection for slowly varying processes. It has been shown by Looze (2005a, 2006) that if the temporal dynamics of the atmospheric wavefront distortions can be modeled as independent first-order autoregressive processes, i.e.

$$\phi_{k+1} = a_i\phi_k + w(k),$$

with $a_i \in \mathbb{R}$ and $w(k)$ a zero-mean white noise process, the classical AO control structure is optimal under some stringent conditions. More specifically, it is assumed that the only dynamics in the AO system is a one-sample delay and that

the DM influence matrix H is invertible. Under these conditions, the minimum-variance controller is composed of the MAP wavefront reconstructor (1.17) (with C_ϕ replaced by C_ϵ), followed by a first-order lag filter with a time-constant equal to the time constant of the considered atmospheric disturbance model. Here, it is important to note that the resemblance between the atmospheric disturbance model and the optimal controller is no coincidence but a direct consequence of the well know Internal Model Control (IMC) principle (Francis and Wonham 1976).

1.3.5 Separation principle

In Section1.3.3, the problems of wavefront reconstruction and the subsequent projection on the actuator space have been considered independently. Wallner (1983) has analyzed these problems in an unified framework. The analysis accounts for the fact that the wavefront distortions over the aperture are continuous functions and includes the effect that the WFS averages the slopes over a finite area. Also Wiberg et al. (2004, 2005) have analyzed the wavefront reconstruction and fitting problem in the setting of infinite-dimensional vector signals. The analysis shows that the unified approach of wavefront reconstruction and the projection on the actuator space as pursued by Wallner (1983) does not yield any additional performance gain. It is shown that the reconstruction and fitting problem can be solved separately, without any loss in performance. This is in fact a direct consequence of the well-known separation principle in control Wonham (1968); Doyle et al. (1989).

Furthermore, the analysis of Wiberg et al. shows that the residual wavefront error in the static reconstruction and fitting problem can be decomposed into three errors sources; that is, turbulence modes that cannot be controlled by the DM, modes that cannot be observed by the WFS and errors in the estimation process due to measurement noise. This leads to the definition of the concepts of controllability and observability. Fusco et al. (2001) derived a minimum-variance estimator for estimating the phase corrections to be applied by each of the DMs in an MCAO system. The reconstructor is directed to the minimization of the mean-square residual phase error over a desired FoV. It is shown that the optimal estimator consists of a full tomographic reconstruction of the turbulent atmosphere, followed by a projection onto the DMs accounting for the FoV. Also this is a consequence of the separation principle.

1.4 More recent developments in AO control

Besides the classical AO control approach discussed in the previous section, an number of other approaches have been proposed in literature. This section provides a concise overview of the most important developments in this field. However, one should keep in mind that especially the most advanced control schemes are still not common practice. The most important differences with the strategies developed in this thesis will be briefly reviewed at the end of subsection 1.6.3.

1.4.1 Modal gain optimization

Modal gain optimization is basically an extension of the classical AO control approach discussed in Section 1.3. The temporal compensator used in the classical AO control consists of a series of parallel but exactly the same feedback loops, i.e. $T_i(z) = T(z)$. As we have seen, the choice of the control parameters of these feedback loops is a trade off between disturbance rejection, noise amplification and closed-loop stability. To relax this trade off, modal gain optimization (Gendron and Léna 1994, 1995; Ellerbroek et al. 1994) uses a modal representation of the wavefront in which the feedback loop corresponding to each model coefficient, $T_i(z)$ is allowed to be tuned differently. Instead of using a common control bandwidth, the bandwidth is hence optimized on a mode-by-mode basis. Since modes with a low spatial frequencies content have a higher SNR, these modes should be assigned a larger bandwidth in minimizing the residual wavefront error.

Gendron and Léna (1994, 1995) developed and experimentally demonstrated a procedure for optimizing the modal gains of an integrator feedback law. It uses a fixed *a priori* chosen set of basis functions that resembles a Karhunen-Loève decomposition. The optimization of the modal gains is entirely based on open-loop WFS data obtained from the AO system to make it independent from assumptions on the turbulence structure. The modal gain optimization approach has been taken a step further by Ellerbroek et al. (1994), who proposed a procedure for simultaneous optimization of both the basis functions used in the modal expansion and their associated control bandwidth. Using this approach with optimized basis functions, it is shown that an AO control system with only two distinct control bandwidths, of which one is zero, achieves almost the same performance as a full optimization. Such a control system has implementational advantages in making the algorithm adaptive to turbulence changes as it requires to adjust only a single control bandwidth rather than an entire set.

1.4.2 Control and wavefront prediction

Apart from the measurement noise and the finite control bandwidth, the inherent time delay in the feedback loop is known to be one of the main limitations on the AO correction performance. It has been suggested that wavefront prediction can be used to reduce the effect of these latency errors. Several papers, like for instance Aitken and McGaughey (1996); Lloyd-Hart and McGuire (1995), have shown that atmospheric turbulence can be predicted over a short time horizon. McGuire et al. (2000) have compared several prediction strategies, including recursive least squares (RLS), a fixed finite impulse response filter (FIR) referred to as matrix-inversion least squares algorithm and different training rules for neural networks, on simulated open-loop data. To reduce the time delay error, Wild (1996) has developed a predictive optimal estimator. This predictor does not truly predict future wavefront distortions from a set of past WFS measurements, but minimizes the time delay error in an ensemble averaged sense.

Many demonstrations of the predictability of atmospheric turbulence have been based on open-loop simulations scenarios. A predictive controller that oper-

ates in closed-loop has been proposed by (Dessenne et al. 1997, 1998; Madec 1999). The so called modal linear predictive controller has a lot in common with the modal gain approach. Like modal gain optimization, the temporal compensator is assumed to be composed of a series of parallel scalar feedback loops applied to modal coefficients. Instead of optimizing only the gain of an integrator, each temporal compensator is fully parametrized by an autoregressive moving average (ARMA) filter. Under the assumption that the only dynamics in the AO system is a delay of 2 samples, a recursive least-squares algorithm is used to estimate the optimal ARMA coefficients that minimize the mean-square residual phase error. The algorithm determines the parameters on open-loop measurements. The modal linear predictive controller has been successfully demonstrated on a $1.52\,\mathrm{m}$ telescope (Dessenne et al. 1999) at the Observatoire de Haute Provence, France.

1.4.3   Static wavefront reconstruction for large telescopes

As telescope designs move to larger and larger apertures, also the number of degrees of freedom of the AO system (i.e. the number of DM actuators and WFS channels) will increase substantially. Things will even get worse since virtually all large telescope designs will incorporate MCAO to exploit the full potential of the increased aperture size. For the least-squares and minimum-variance wavefront reconstruction techniques discussed in Section 1.3, the complexity of computing the control matrix and applying it to the WFS data in the form of a vector-matrix multiply scales respectively as $\mathcal{O}(n^3)$ and $\mathcal{O}(n^2)$, where n is either the number of DM actuators or the number of WFS sub-apertures. As a result, current reconstruction techniques will become computationally intractable for future extremely large telescopes (ELT) using AO systems with $10^4 - 10^5$ degrees of freedom.

To overcome these problems, there has been recently quite some effort in improving the scalability of the wavefront reconstruction and projection step. This has resulted in a variety of iterative algorithms that exploit the specific structure of the reconstruction problem and avoid the explicit computation of the reconstructor matrix (see e.g. Ellerbroek 2002; Gilles et al. 2002; Poyneer et al. 2002; Gilles et al. 2003; Gilles 2003; MacMartin 2003). Besides the gain in computational efficiency, these methods have a memory advantage in that they do not need to store the entire reconstructor matrix. Ellerbroek (2002) proposed the use of a slight modification to the regularization term appearing in the minimum-variance reconstructor to obtain a representation that can be efficiently solved by sparse matrix techniques. This approximation appears to have a negligible effect on the estimation accuracy, and the sparse matrix techniques reduce the overall complexity of the approximate reconstructor to $\mathcal{O}(n^{3/2})$.

The algorithm described by Gilles et al. (2002) approximates the phase covariance matrix by a block circulant matrix with circulant blocks, which can be efficiently inverted in Fourier domain. Then a multi-grid preconditioned conjugate-gradient (MGCG) method is used to reconstruct the minimum-variance estimate of the wavefront with a total complexity in the order of $\mathcal{O}(n \log n)$. Another reconstruction technique that uses fast Fourier transforms is based on direct inversion of the WFS model in the frequency domain (Poyneer et al. 2002). Also this algorithm

has a computational complexity of $\mathcal{O}(n \log n)$. The use of preconditioned conjugate gradient solvers for MCAO has been elaborated in (Gilles et al. 2003). The two preconditioners considered in this paper require an off-line sparse Cholesky factorization of complexity $\mathcal{O}(n^{3/2})$, but their averaged update rate is typically 4-5 orders of magnitude lower than the sample rate. All on-line computations scale linear in the number of sensors and actuators, i.e. $\mathcal{O}(n)$. Similar preconditioners are considered in Gilles (2003). The use of local reconstructors wherein the actuator commands depend only on sensor information in a neighboring region has been analyzed by MacMartin (2003). To prevent performance degradation from global modes, the local estimators are either augmented by a second hierarchic layer of global estimators, or extended to include past local estimates, resulting in an overall computational complexity of $\mathcal{O}(n^{4/3})$ and $\mathcal{O}(n^{3/2})$, respectively. To arrive at a control strategy that offers full scalability for future generations of large DMs (see e.g. Hamelinck et al. 2004, 2005a,b, 2006), (Ellenbroek et al. 2006) proposed a distributed control framework for AO in which each actuator has a separate processor that can communicate with a few direct neighbors.

1.4.4 Optimal control for adaptive optics

Paschall and Anderson where the first to formulate the AO control problem in a Linear Quadratic Gaussian (LQG) optimal control framework (Paschall 1991; Paschall and Anderson 1993). In contrast to the control approaches discussed up to now, this is a model-based control approach which explicitly tries to minimize the residual phase variance. The control design starts with deriving dynamic models of both the AO system and the turbulent atmosphere. Given these models, LQG control theory can be used to compute the optimal controller that minimizes the mean-square residual phase error. By deriving appropriate and accurate models, the LQG framework is not only able to account for pure time delays in the AO system but also for DM and WFS dynamics. Which of these effects are included in the control design, depends entirely on the detail of modeling. Hence, the quality of the models will make or break the success of LQG.

The centralized LQG control design by Paschall and Anderson is based on an atmospheric disturbance model that describes the wavefront distortions in terms of the first 14 Zernike modes with modal coefficients generated by an independent first-order Markov process (for definition Markov process, see Shanmugan and Breipohl 1988). The atmospheric disturbance model is identified from simulation data obtained from analytical expressions for the auto-correlation kernels of the Zernike coefficients. The mirror dynamics are modeled as a first-order lag filter and a one sample delay is included to account for the time delay in the feedback loop. A modal control algorithm that uses the LQG formalism has been proposed by Looze et al. (1999, 2003). In this approach the controller is decomposed in a set of modes derived by expanding the wavefront distortions on a Karhunen-Loève basis. The individual modes are assumed to be decoupled and the effect of the atmospheric turbulence on each of them is described by a separate ARMA model. Open-loop WFS data are used to identify the model parameters. Under the assumption of modal decoupling, the control design reduces to a bunch of scalar

LQG optimal control problems. Gavel and Wiberg (2003) recently proposed an optimal control approach which is explicitly based on the assumption that the atmosphere can be described by a single layer of frozen Kolmogorov turbulence, satisfying the Taylor hypothesis. To perform the time update the so called near-Markov approximation is introduced. This approximation states that all information needed to predict the wavefront phase at a certain time-instant is contained in the phase distortion over the aperture at the previous time-instant. Using this assumption, only the conditional mean of the wavefront in the aperture is needed to convey all statistical information from the past.

The use of the LQG optimal control formalism has been extended by Le Roux et al. (2004) to make it suitable for MCAO. The basic idea is to discretize the FoV over the area over which the performance has to be optimized. Considering a finite number of discrete turbulence layers, the problem of optimizing the mean-square residual phase error can again be formulated as minimizing the norm of a finite-dimensional vector signal. Under the assumption that the DM dynamics can be neglected, the control law reduces to a Kalman filter for estimating the phase distortions followed by a static state feedback law. Also in this case, the control design is based on a simple atmospheric turbulence model that describes modal coefficients of a Zernike expansion of the wavefront distortions by means of a first-order AR model. The AR coefficient matrix is chosen diagonal and its elements have been adjusted to enforce a correlation time that decreases with the Zernike radial order. The covariance matrix of the zero-mean white noise input is determined in such a way, that the phase covariance matrix satisfies the Kolmogorov spatial distribution. The same atmospheric disturbance model is used in the Kalman based optimal control approaches presented in (Petit et al. 2004, 2006). The first paper shows that the flexibility of optimal control also allows counteracting the effect of telescope vibrations. The latter reports the first experimental results of Kalman based optimal control for active vibration suppression and extending the FoV.

1.4.5   Adaptive and quasi-adaptive control schemes for AO

The control schemes discussed so far, are all time invariant. The most advanced concepts rely upon *a priori* knowledge of the turbulence statistics to improve their performance. The required *a priori* knowledge is obtained by matching physical parameters to the expected turbulence conditions or are acquired in a separate identification experiment. Since atmospheric parameters like the wind velocity and the atmospheric turbulence strength change on time scales of minutes, it becomes necessary to constantly update the controller to the changing atmospheric statistics in order to prevent a loss in performance. To account for the varying turbulence conditions and track the chances in the spatial and temporal correlations, different adaptive control schemes have been proposed (Ellerbroek and Rhoadarmer 1998, 2001; Gibson and Chang 1999; Gibson et al. 1999, 2000, 2001; Chang and Gibson 2000; Liu and Gibson 2004).

The adaptive control scheme proposed by Ellerbroek and Rhoadarmer (1998, 2001), relies on a recursive least squares (RLS) algorithm to adaptively optimize

the wavefront reconstructor of an AO system on the basis of closed-loop WFS measurements. The approach applies also to wavefront reconstructors that include multiple frames of past measurement data to compensate for the system latency. The main idea of this control strategy is to decompose the WFS measurement vector in two subspaces, namely the range space of the DM actuator commands and its orthogonal complement. By introducing a suitably chosen constraint on the reconstructor, it is fully specified on the first subspace but unconstrained on the second. Furthermore, the constraint guarantees that both the estimated residual phase error and the second component of the WFS measurement vector are completely independent form the current DM actuator command so that an open-loop RLS can be used for optimizing the second part of the wavefront reconstructor. This results in a control strategy which has a fixed reconstructor for nominal performance, augmented by an adaptive feedback loop.

Another adaptive control strategy for AO is based on multi-lattice RLS filtering (Gibson and Chang 1999; Gibson et al. 1999, 2000, 2001; Chang and Gibson 2000; Liu and Gibson 2004). Like the previous approach, this control strategy employs augmentation of a LTI feedback loop with an adaptive loop. The design of the controller is posed as an generic adaptive feedforward disturbance rejection or a noise cancellation problem. The common denominator of the different flavors of this approach is that they all reconstruct an estimate of the open-loop WFS signal by subtracting the influence of the DM wavefront corrections from the closed-loop measurements. Using this estimate as a reference input, RLS updates the coefficients of a FIR filter with the objective to minimize the variance of an estimate of the residual wavefront. The adaptive RLS filter is implemented in the form of a multi-channel lattice filter to provide the required numerical stability and speed required for real-time AO. Furthermore, an efficient re-parametrization of the actuator space leads to a control space from which the degrees of freedom in the actuator space, that are most effective for wavefront correction, can be easily inferred. In this way, it is possible to disregard ineffective degrees of freedom which reduces the necessary on-line computations. Finally, it has been recently pointed out by Liu and Gibson (2004) that a quasi adaptive controller which updates the gains periodically may be as effective as a full adaptive control loop.

1.5   Scope of the thesis

This thesis focuses on the control aspects of AO. The requirements on the control strategy are strongly intertwined with the system design. Many choices in the AO system design, like the choice of the DM and WFS, have a large influence on the achievable (control) performance. It is therefore to be expected that the overall performance may benefit from an integrated AO system and control design. Ideally, both the AO system and control design should be tailored to the prevalent atmospheric seeing conditions. In practice, however, the AO system design choices are rather permanent and cannot easily be adapted to the seeing. As a result, the AO system components are typically optimized for averaged seeing conditions. The AO control design, on the other hand, is more flexible and can

hence be adjusted to turbulence conditions. For this reason, it will be assumed that the AO system design is fixed. The effect of important AO system design choices, such as the type of WFS and DM, the optimization of the sensor-actuator layout and the sampling time, are beyond the scope of this thesis. Furthermore, it will be assumed that the design has been optimized to such an extend that there is still room for improvement by control. This means that under the considered seeing conditions the control performance is one of the limiting factors.

The objective of this research is to demonstrate that advanced control strategies are able to improve the performance of operational AO systems by explicitly accounting for both the turbulence and WFS dynamics. In analyzing the control problem it will be assumed that the AO system components, that is the DM and WFS, can be modeled as LTI systems. Later on, this will be narrowed down to the assumption that the DM can be considered to be static and the only dynamics in the AO feedback link is caused by the integrating action of the WFS. This is a reasonable assumption for most AO systems as the time-constant of the DM mirror is typically much smaller than the WFS exposure-time. The proposed control strategy is not restricted to a particular type of WFS. The only requirement however is that the optical transformation from phase to wavefront slopes, curvature or whatever quantity is measured, can be described by a static linear mapping. Furthermore, it will be assumed that second-order statistics of the atmospheric distortions can be modeled as a stationary regular stochastic process, at least over a time-span that is long compared to the sampling rate and the time constant of the fluctuations themselves. This assumption on the turbulence structure will be considered in more detail in Section 1.6.

In this thesis we will not consider the problem of making the proposed algorithms scalable for future generations of large telescope with many hundreds to thousands of sensors and actuators. As we have see in Section 1.4.3, the development of efficient wavefront reconstruction algorithms for this type of telescopes is an important research area in itself. For the proposed control strategy, the off-line identification of the atmospheric disturbance model is the most computational demanding step. This step can be performed within a few minutes on a general purpose PC with a $3\,\mathrm{GHz}$ Intel Pentium IV processor and $512\,\mathrm{Mb}$ of internal memory, for AO systems with up to a few hundred degrees of freedom. This corresponds to the typical size of current AO systems. Finally, we will restrict our attention to a classical AO system where only a single WFS, conjugated to the telescope aperture, is used to sense the wavefront distortions. Even though MCAO will not be considered, there are no fundamental limitations to extend the proposed control strategy to MCAO. More specifically this can be achieved by using the theoretical framework similar to the one proposed by Conan et al. (1995).

1.6   Motivation and research strategy

In Section 1.3 we have seen that most AO systems are still based on a control strategy that decomposes the control problem in a static wavefront reconstruction

step and a temporal compensator design. Even though this classical control approach provides already remarkable results, there is still a need for improvement. As pointed out by Roddier (1998), the compensational efficiency of current AO systems is sometimes unduly low. Rather than focusing only on the development of AO systems with more sensors and actuators it may therefore rewarding to improve the performance of current AO systems. The development of improved AO control strategies may help to achieve this goal, especially when operating under low light level conditions and rough turbulence. Starting from the shortcomings of the classical control approach, this section will explain and motivate the research strategy followed in this thesis.

1.6.1 Limitations of classical AO control approach

From a control engineering point of view, the classical AO control approach leaves a lot of room for improvement. The structure of the control law is mainly based on physical insights. Rather than treating the control problem as an integral disturbance rejection problem, the control design is seen as servo control problem in which the wavefront correction applied by the DM has to follow the distortion reconstructed from the WFS measurements. The main shortcomings of this approach can be summarized as follows:

- The decomposition of the control law in a static wavefront reconstruction step and a temporal compensator consisting of parallel feedback loops, implicitly assumes that the spatial and temporal dynamics are decoupled.

- Since the parallel feedback loops typically have a fixed servo controller structure (e.g. PI, first-order lag filter or Smith Predictor), it usually is not possible to account for the AO system and turbulence dynamics in an optimal way. The temporal controller design is pragmatic and not model based.

- The wavefront reconstruction problem is usually solved assuming open-loop conditions. When using a minimum-variance reconstructor, the modified wavefront statistics as a result of closed-loop operation, are neglected.

To illustrate that the above limitations are indeed rather restrictive, it is useful to reconsider the structure of the atmosphere. As explained in Section 1.2, the temporal evolution of the wavefront is usually described by considering the atmosphere as being composed of different wind blown frozen phase-screens. Assume for the moment that this is true and that there is only one dominant turbulence layer moving with a velocity $v \in \mathbb{R}^2$ (see Figure 1.7). Since the turbulence is assumed to be frozen, the wavefront distortions observed at time t will have only been shifted over a distance $v\tau$ at time $t + \tau$. This implies that the phase observed at the point $x - v\tau$ and time t can be used to predict the distortion at the point $x \in \mathbb{R}^2$ and time $t + \tau$. More generally, wavefront distortions observed at the windward side provide direct information on (the future) development of the turbulence elsewhere in the aperture.
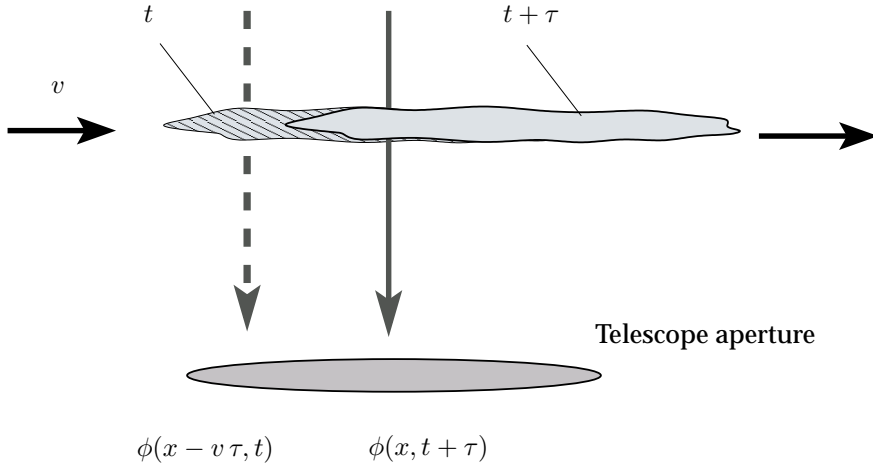
Figure 1.7: Visualization of the Taylor approximation. A layer of frozen turbulence is passing over the telescope aperture.

Even though this example is overly simplistic, it clearly shows that the spatial and temporal dynamics of the wavefront distortions cannot be decoupled. Because of the existing coupling, WFS measurement from neighboring channels can be used to improve the estimate of the wavefront and to predict its future development. The ability to predict future wavefront distortions is particularly useful to reduce the latency errors in the control system. The spatio-temporal correlation in the wavefront may also be exploited to reduce the sensitivity to measurement noise, since photons collected by different WFS channels at different time instants may all contribute to the wavefront estimation process. This is particularly interesting when using a faint guide star. Also in more complicated cases where the turbulence consists of several frozen layers or is partially non-frozen, there is no reason to believe that the spatial and temporal dynamics are decoupled. Any control approach (like the classical AO control approach) that departs from this assumption, will therefore inevitably sacrifice a part of the achievable performance.

1.6.2 An control engineering approach to AO

The above considerations call for an integrated control strategy that departs from a minimal number of prior assumptions on the spatio-temporal structure of the turbulence. In this thesis we will approach the AO control problem from a control engineering point of view. The AO control problem will be formulated as a multi-variable disturbance rejection problem. More specifically, the proposed control design strategy is one of data-driven disturbance modeling followed by a model based $\mathcal{H}_2$-optimal control design (Shu and Chen 1995; Chen and Francis 1995; Zhou et al. 1996; Burl 1999; Hassibi et al. 1999). The use of data-driven identification is attractive as it yields a good match with the prevalent turbulence conditions without the need to accurately estimate all kinds of physical parame-

ters. Furthermore, it provides the necessary flexibility to ensure that the control approach does not depend on restrictive assumption such as Kolmogorov turbulence and the frozen flow assumption. Before elaborating the proposed control approach, it is useful to formulate the general purpose of this research.

- Research objective -

The aim of the research presented in this thesis is to demonstrate that the performance of the current generation of AO systems can be improved by applying advanced control strategies that are able to account for the AO system dynamics, the spatio-temporal correlation in the wavefront and the fact that the AO system is operated in closed-loop.

In the above statement, the notion performance improvement should be interpreted in the broadest sense of the word. Advanced control is expected to realize performance improvements with respect to the following objectives:

- **Imaging performance** - AO is used to restore the imaging quality despite the presence of turbulence. The main objective is therefore to optimize imaging performance; i.e. achieve higher Strehls.

- **Sky coverage** - By reducing the sensitivity to measurement noise, it should be possible to relax the requirements on the limiting magnitude of the guide star. This will improve the sky coverage.

- **Reliability** - The performance of the current AO systems is very sensitive to the particular observing conditions. If this sensitivity can be reduced this will improve the usability and reliability of AO.

As already mentioned, the proposed control design strategy is based on data-driven identification followed by an $\mathcal{H}_2$-optimal control design. The reason why this approach is expected to improve the performance of current AO systems is that it basically removes the major limitations of the classical control approach. If the data-driven identification provides accurate models of the AO system dynamics and the atmospheric disturbances, the $\mathcal{H}_2$-framework will automatically account for the DM and WFS dynamics and the spatio-temporal correlation in the wavefront. Furthermore, the $\mathcal{H}_2$-framework inherently yields a closed-loop control design. In the case that there are no model errors, the $\mathcal{H}_2$-optimal control design will find the controller that is optimal in the sense that it minimizes the residual phase variance. Because of the relation between phase variance and Strehl, the controller also optimizes the imaging quality.

The extend to which the data-driven optimal control design strategy is able to exploit the spatio-temporal correlation in the wavefront depends on the atmospheric disturbance model. In the data-driven identification procedure, the only assumption imposed on the wavefront distortions is that their second-order statistics can be modeled as a regular stochastic process. In other words, it is assumed

that one can think of the wavefront distortions as being generated by a finite-dimensional LTI system with a white noise input. This way of modeling stochastic distortions is quite common in the field of signal processing and systems and control (Åström and Wittenmark 1997; Hayes 1996; Kailath et al. 2000; Shanmugan and Breipohl 1988) and often yields a reasonable description for bandwidth limited processes. An important consequence of this assumption, however, is that the wavefront statistics are assumed to be stationary. Whether or not this is realistic for AO depends on the considered time-scales. The statistical properties of the wavefront distortions are known to change on a time-scale in the order of minutes, while the fluctuations themselves have time-constant in the order of a few milliseconds. This implies that at least on a time-scale of minutes the stationary assumption is quite reasonable. To guarantee a good performance over longer periods however, a regular update of the disturbance model might be required.

The proposed control strategy will use a dedicated subspace identification algorithm to identify the atmospheric turbulence model on the basis of open-loop WFS data. In the end, the goal is to identify a full multi-variable ARMA disturbance model that does not assume any form of decoupling between the channels. Such an atmospheric disturbance model is sufficiently general to capture the spatio-temporal correlation imposed by one or more frozen layers. In this way, the proposed control strategy will be able to benefit from the existence of one or more frozen layers, without being dependent on it. The problem of identifying such a full multi-variable disturbance model is probably the most complicated step, in the entire procedure. Because of the large number of WFS channels, identifying a disturbance model without any form of decoupling is a challenging problem which requires efficient algorithms. For this reason, a significant part of this work is devoted to this problem.

1.6.3   Difference with existing control approaches

The $\mathcal{H}_2$-optimal control strategy presented in this thesis is closely related to the LQG and Kalman filter based minimum-variance approaches discussed in Section 1.4.4. Indeed, the LQG and the $\mathcal{H}_2$-optimal control framework are equivalent in the sense that an LQG problem can be reformulated in an $\mathcal{H}_2$-optimal problem and visa versa (Zhou et al. 1996; Hassibi et al. 1999). The main difference with the proposed control strategies is therefore not in the framework of analysis but in the model structure used to describe the disturbance model.

The existing LQG approaches are based on an atmospheric disturbance model that either assumes modal decoupling, or restricts the spectral factor to the class of first-order autoregressive (AR) processes. Neither of these model structures is able to account for the spatio-temporal correlation imposed by a frozen flow. The first model structure is too restricted because it assumes a decoupling of the spatial and temporal dynamics. In this perspective it is important to note that also a Karhunen-Loève basis (see e.g. Roddier 1999; Hardy 1998; Wang and Markey 1978) does not provide spatio-temporal decoupling as it diagonalizes only the first coefficient matrix of the auto-correlation function. For time-shifts larger than one, the covariance matrices still contain off-diagonal entries so that there remains a

coupling between the individual modes. On the other hand, the impulse response of a first-order AR process is too short for modeling the temporal dynamics.

Like the LQG and Kalman filter based approaches, also the adaptive control strategies discussed in subsection 1.4.5 impose a restriction on the spatio-temporal correlation of the atmospheric turbulence that can be accounted for. The model structure of the adaptive control loop is typically restricted to a finite impulse response (FIR) filter with only a few taps. This implies that the atmospheric turbulence as well as the residual wavefront error after compensation by the additional LTI feedback loop, is implicitly assumed to have this structure. Furthermore, the proposed control approach does not require any tuning of parameters like the step-size, is not prone to convergence problems and has an on-line computational demand that is much lower than that of an adaptive approach. An important advantage of adaptive control, however, is that it is able to respond to gradual changes in the turbulence statistics. The proposed control strategy has to be extended with some kind of model update scheme to cope with such changes.

1.7 Main contributions

Considering the research strategy presented in the previous section, this section provides a brief overview of the main contributions. The different contributions as well as their mutual relation will be discussed in more detail in Section 1.8. The contributions in this thesis can be summarized as follows:

- A new subspace-based identification algorithm for estimating the minimum-phase spectral factor of a (matrix-valued) stochastic process given samples of its power spectrum. One of the key features of the algorithm is the use of conic linear programming to guarantee the preservation of the positive realness property of the estimated power spectrum.

- A stochastic subspace identification algorithm that avoids the need for spectral factorization by directly estimating the state-space matrices of the one-step ahead predictor, from time-domain data. By paying special attention to an efficient implementation, the algorithm is able to identify a full multivariable disturbance model for AO systems with up to a few hundred degrees of freedom using a single general purpose PC with $3\,\mathrm{GHz}$ Intel Pentium IV processor and $512\,\mathrm{Mb}$ internal memory.

- The combination of data-driven subspace identification and $\mathcal{H}_2$-optimal control design in the context of AO. The $\mathcal{H}_2$-optimal control framework provides a convenient way to deal with the discrepancy between the WFS signal and the control objective. The minimum-phase property of the identified disturbance model implies that at most one Riccati equation has to be solved in computing the optimal controller.

- The transfer function from control inputs to WFS outputs has been analyzed for the common case that the DM can be considered static and the only dynamics in the AO system derives from the integrating action of the WFS. A

data-driven identification algorithm has been developed for identifying the model parameters of the derived model structure. For this model structure an analytical expression for the $\mathcal{H}_2$-optimal controller has been derived.

- Experimental validation of the proposed data-driven $\mathcal{H}_2$-optimal control approach on an AO laboratory setup. By analyzing the dominant error sources it has been shown that optimal control is able to reduce the latency errors in the AO system. Furthermore, the experiments show that the same level of performance can be achieved at lower sample rates, enabling longer WFS integration times, reducing the noise. This may improve the sky coverage.

1.8   Organization of the thesis

This section provides a brief overview of the organization of the thesis. This thesis consists of 5 chapters and an appendix. Each of the subsequent chapters, except for Chapter 5, are basically extended versions of papers that have been published or submitted for publication in well-established journals in the field of optics and control. Chapter 5 concludes the dissertation. It summarizes the main conclusions that can be drawn and gives suggestions for future research. Finally, appendix A includes the formal proof of the main theorem in Chapter 3. The results presented in Chapters 2 to 4 are based on the following papers:

1. Hinnen, K., Verhaegen, M., Doelman, N., Oct. 2005. Robust spectral factor approximation of discrete-time frequency domain power spectra. Automatica 41 (10), 1791–1798.

2. Hinnen, K.J.G., Verhaegen, M., Doelman, N. J., Mar. 2006. A data-driven $\mathcal{H}_2$-optimal control approach for adaptive optics, submitted to IEEE Transactions on Control Systems Technology.

3. Hinnen, K., Verhaegen, M., Doelman, N., Jun. 2006. Exploiting the spatio-temporal correlation in adaptive optics using data-driven $\mathcal{H}_2$-optimal control, submitted to Journal of the Optical Society of America A.

Each of these papers forms the backbone of a separate chapter of the thesis. The different papers have been extended to include more background information which makes them more accessible for people from both the optics and control communities. In addition, the link between the first paper and the remaining part of the dissertation has been clarified. The other two papers have been augmented to include some additional simulations and experimental results. Part of these results have previously appeared in conference proceedings. A complete list of publications and papers submitted during this PhD research period is included in the list of publications at the end of thesis.

Using this structure, the three main chapters of the thesis can be regarded as separate entities, which can be read independently. A consequence of this setup, however, is that it also gives rise to a certain extend of overlap between the individual chapters. To maintain the independent nature of Chapters 2 to 4, the

references are included at the end of the corresponding chapter. In order to ensure that the papers form an integral part of the dissertation, all references have also been collected in the bibliography at the end of the thesis. For similar reasons, an attempt has been made to keep the notation consistent throughout the thesis. A concise overview of the content of the main chapters is provided by the paper abstracts at the beginning of each chapter.

1.9 Outline of the thesis

This section provides an outline of the thesis. Its main objectives are to motivate the different choices in elaborating the research strategy and to present an overview of the relation between the main research results. Furthermore, the mutual relation between the different chapters will be explained in more detail. As a result, the reader may use this section as a guideline to find his way through the thesis. The discussion is organized topologically, but shows a rather close resemblance with the chapter division.

1.9.1 Disturbance modeling from discrete-time power spectra

As explained in Section 1.6, the general strategy pursued in this thesis is to capture second-order statistics of the wavefront distortions in a minimum-phase spectral factor, and then use the identified spectral factor to solve the control problem in an optimal control framework. When no measurement data are available, the spectral factor has to be determined on the basis of a theoretical description of the atmosphere. For this purpose, the description of atmospheric turbulence in terms of frozen layers with a Kolmogorov spatial distribution is sufficient. Given the turbulence strength, the wind speed and the direction of the different layers, such a model provides a complete specification of the second-order statistics of the wavefront. So, the problem is to find a minimum-phase spectral factor that provides an accurate description of the second-order statistics imposed by the theoretical model. Considering a single point in the aperture, the theoretical atmospheric disturbance model gives rise to the power spectrum defined in equation (1.9). Likewise, by accounting for the contribution of the different layers, it is possible to derive a matrix valued power spectrum for the wavefront phase vector $\phi(k) \in \mathbb{R}^{m_\phi}$. For this reason it would be desirable to have a way of estimating the minimum-phase spectral factor directly from power spectrum data.

The problem of estimating a minimum-phase spectral factor from a matrix valued power spectrum is not restricted to this particular application. It occurs in any application that requires a filtered white noise model of a stochastic process that is characterized by its power-spectrum only. Using power spectra to describe a stochastic process may, for instance, be attractive when considering large data sequences which are difficult to handle in the time-domain. Indeed, by going to the frequency domain one can condense the relevant information in a smaller amount of data points. Chapter 2 considers the spectral factor approximation problem

in a general setting. It provides a subspace-based algorithm for identifying the state-space quadruple $[A, B, C, D]$ of a minimum-phase spectral factor that accurately approximates the given spectral data. Since the objective is to determine a state-space realization of a spectral factor that provides a good match with the given power spectrum, the algorithm will be referred to as the spectral factor approximation algorithm. The proposed spectral factor approximation algorithm basically consists of a subspace identification algorithm followed by a parametric optimization procedure. The rationale behind this is that the subspace algorithm provides a first reasonable estimate of the spectral factor. This estimate is then used to initialize the parametric optimization procedure to further minimize the weighted 2-norm of the error between the estimated and the given power spectrum. Since the subspace algorithm generally provides good initial estimates of the spectral factor, the drawbacks associated with the highly nonlinear and non-convex nature of the parametric optimization problem are partially relaxed.

The proposed subspace identification algorithm is based on an approximate relation between the inverse-discrete Fourier transform (IDFT) of the power spectrum and the system matrices $[A, C]$ of the corresponding spectral factor. The considered approximation can be interpreted as neglecting the effect of aliasing, an assumption that is quite reasonable whenever the stochastic process is well sampled. In comparison with the exact relation used in the subspace algorithm for the identification of discrete-time frequency domain power spectra, proposed by Van Overschee et al. (1997), the use of this approximation avoids the need to carry out a complicated split, which is very sensitive to errors on the power spectrum.

Another key aspect of the proposed subspace algorithm is the way it deals with the well-known positive realness issue in stochastic identification. Like most stochastic subspace identification algorithms for time series (see e.g. Aoki 1990; Van Overschee and De Moor 1993a; Mari et al. 2000a; Dahlén 2001), the proposed subspace algorithm does not directly estimate the minimum-phase spectral factor. Instead of this, the algorithm first identifies a model that describes the covariances or power spectrum of the stochastic process, which is then factorized to obtain an estimate of the minimum-phase spectral factor. For this approach to be successful, it is important to ensure that the power spectrum model identified in the intermediate step is positive real. If this condition is not satisfied, the identified power spectrum has no physical meaning and the spectral factorization cannot be performed (see e.g. Sayed and Kailath 2001; Kailath et al. 2000). The risk of getting stuck on the spectral factorization step is especially large when the power spectrum is contaminated by measurement noise or when it cannot be attributed to a finite-dimensional linear time invariant system.

To prevent the proposed subspace algorithm from ending up in this situation, the intermediate power spectrum estimate has to preserve the positive realness property. In the proposed subspace algorithm this is achieved by imposing an additional constraint on estimating the pair $[B, D]$. More specifically, for a given pair $[A, C]$, the problem of finding the optimal $[B, D]$ is formulated as a least squares problem with linear matrix inequality constraint. It will be shown that the resulting constrained least squares problem can be efficiently solved via conic linear programming (CLP). This is computationally far more efficient than the

standard approach in which the constrained least squares problem is solved as a large scale semi-definite program, and avoids the problems associated with local minima that may occur when using direct nonlinear least squares optimization over the unknowns $[B, D]$ (Mari et al. 2000a; Van Overschee et al. 1997). In this perspective, it is important to note that the idea of using conic linear programming to deal with the positive realness issue is not restricted to the proposed spectral factor approximation algorithm. Indeed, since the positive real requirement in stochastic subspace algorithms for time domain identification can be formulated analogously, also these algorithms may benefit from this idea.

Furthermore, note that since CLP is convex, the problem of estimating the pair $[B, D]$ can be solved in a globally optimal sense. This property is used to derive a numerical analogue of the separable least-squares principle which is used to facilitate the parametric optimization procedure over the parameters of the minimum-phase spectral factor. Separable least squares reduces the number of parameters in the optimization problem and is known to give rise to a better numerical conditioning and reduction in the required number of iterations (Bruls et al. 1999; Sjöberg and Viberg 1997; Ribarits et al. 2003).

The performance of the proposed spectral factor approximation algorithm has been demonstrated by applying it to a number of challenging simulation examples. The first simulation example is not directly related to AO and is used to investigate the performance in the case the power spectrum is indeed generated by a stochastic process with a spectral factor that belongs to the class of finite-dimensional systems. Furthermore, the example is used to investigate the performance robustness with respect to estimation errors on the power spectrum. The stochastic process considered in the simulations is vector-valued and has a challenging power spectrum that consists of a number of distinct resonance and anti-resonance peaks. The simulation experiments show that in the case the spectral factor approximation algorithm is applied directly to the true power spectrum, it is able to find a minimum-phase spectral factor that approximates the spectrum close to the numerical precision and round-off errors. Also in the case that the power spectrum is estimated by using an averaged periodogram and is hence contaminated by estimation errors, the algorithm is able to provide a good fit to the provided spectral data. In fact, the weighted 2-norm of the difference between true spectrum and power spectrum corresponding to the estimated minimum-phase spectral factor is smaller than the weighted 2-norm of difference between the true spectrum and the averaged periodogram.

The paper in Chapter 2 has been extended with a second simulation example that bears a much closer relation with the purpose for which the spectral factor approximation algorithm has been actually developed. It considers the problem of estimating the spectral factor that describes the temporal evolution of the wave-front distortions at a single point in the aperture. The theoretical power spectrum of the phase at a single point in the aperture is given by equation (1.9). This is a non-rational power law, which gives rise to a spectral factor of infinite-dimension. In contrast to the previous example, the power spectrum can no longer be described by a stochastic process that is in the model set considered by the spectral factor approximation algorithm. The problem at hand is therefore to approxi-

mate the power spectrum of an infinite-dimensional stochastic process by a finite-dimensional minimum-phase spectral factor.

The simulation example shows that the algorithm is able to provide a good match with the given power spectrum over a rather large bandwidth. Only in the low frequency range, the approximation error is relatively large. The relatively large fitting error in this frequency region can be explained by the fact that the Kolmogorov power spectrum has a singularity at $f = 0$, which causes the power spectrum to approach infinity. Since the spectral factor approximation algorithm always requires the specification of the power spectrum at a number of discrete sample points including $f = 0$, where inevitably an error is introduced. Note that this is not really a shortcoming of the proposed algorithm. It is well known that due to the singularity at $f = 0$, the Kolmogorov turbulence is physically not realistic as it corresponds to a process with infinite variance (Hardy 1998; Tatarskii 1971; Goodman 1985). Indeed, the Kolmogorov model is only valid in the inertial sub-range which excludes the low frequency range. To avoid the singularity problem, one should account for the finite outer scale L_0 by using the Von Kármán spectrum. The bottom line of the simulation experiment, however, is that at least in the relevant frequency range the temporal dynamics can be accurately modeled by a regular stochastic process. When the spatial and temporal dynamics are assumed to be decoupled, as in the common control approach, the identified spectral factor can be used to compute the optimal servo controller.

1.9.2   Disturbance modeling from time-domain data

Given the theoretical power spectrum of the wavefront distortions at a single point in the aperture, the proposed spectral factor approximation algorithm can be used to identify the minimum-phase spectral factor that captures the temporal dynamics in the wavefront. The identified spectral factor can then be used to account for the temporal dynamics in the control design. The ultimate goal however, is not only to account for the temporal dynamics, but to exploit the full spatio-temporal correlation in the control design. This requires a multi-variable disturbance model that provides a full description of the auto-correlation of, and the cross-correlation between, all WFS channels or phase points in the aperture. Such a model could, in principle, be obtained by first deriving all the spectra and cross-spectra from the theoretical frozen flow model, and then using the spectral factor approximation algorithm to determine the minimum-phase spectral factor.

Yet, from a practical point of view, this approach seems to be unnecessarily laborious and restrictive. First of all, determining the spectra and cross-spectra on the basis of the frozen flow assumption requires accurate knowledge of the number of frozen layers, their respective turbulence strengths, wind speeds and directions. To obtain accurate knowledge of the current value of these physical parameters, they have to be estimated on the basis of WFS data. Estimating the parameters is not that simple, and are only needed to construct the spectra and cross-spectra. Actually, it is much easier to directly estimate the required power spectra, or the equivalent correlation functions in the time-domain, from the WFS data. This has the additional advantage that the estimated power spectra depend

no longer on the theoretical atmospheric disturbance model. Indeed, estimating the power spectrum from the WFS data does not require any assumption on the structure or underlying physical model of the turbulent atmosphere.

Also, when using open-loop WFS data, a procedure consisting of estimating the power spectra and then using the spectral factor approximation algorithm is unnecessarily involved. The first step in estimating the pair $[A, C]$ is to compute the IDFT, by which we effectively return to the time-domain. Moreover, estimating the power spectrum from time-domain data is a delicate process which involves the choice of one out of many different methods and settings. It is also questionable whether or not the formulation of the identification problem as a weighted 2-norm approximation of the spectrum is very natural in the sense that it is able to guarantee consistency of the covariance parameters. Only in the case that the spectrum is weighted by the true power spectrum, the chosen criterion reduces to the classical prediction criterion. Finally, several simulation experiments show that the spectral factor approximation algorithm is computationally too complex to determine the spectral factor for a realistic number of phase points or WFS channels. Even the use of CLP cannot avoid that the number of independent channels that can be modeled, without making serious concessions to the spectral resolution, remains limited to roughly five to ten channels.

The above arguments are strongly in favor of a time-domain identification approach that estimates the minimum-phase spectral factor directly from open-loop WFS data. Simply choosing to work in the time-domain, however, does not automatically solve all problems. As already mentioned in Section 1.9.1, the positive real issue in stochastic identification is not restricted to the frequency domain. It can be formulated in a completely analogous way in the time domain. Indeed, many algorithms for stochastic time-domain identification, estimate a rational covariance model as an intermediate step. As pointed out by Lindquist and Picci (1996), such a procedure might easily fail if there is no additional constraint to ensure that the estimated covariance model is positive real and has a valid spectral factor. Also in the time-domain, the positive realness requirement can be formulated as a constraint least squares problem. Since solving this problem is the most computationally intensive step in the power spectrum approximation algorithm, it is to be expected that only a little can be gained by considering stochastic identification algorithms that estimate a rational covariance model as intermediate step. Furthermore, it should be pointed out that both the dimension and the number of unknowns in the least squares problem scales quadratically with the number of channels. It is therefore not likely that this type of algorithms can be easily scaled to the dimension of a realistic AO system.

The first part of Chapter 3 presents a dedicated subspace-identification algorithm that is able to avoid the need for spectral factorization. This is achieved by working directly with the raw time domain data, rather than first estimating a covariance model. The proposed subspace algorithm is basically an output only version of a subspace algorithm introduced by Jansson (2003), which has been improved to make it more efficient when dealing with large numbers of channels. By first estimating the coefficients of a high order AR model, more structure is added to the data equations used by the subspace algorithm to estimate a basis for the

state vector. This opens the possibility to directly estimate the state-space matrices A, $\Sigma_1 C$ and K of the *innovation predictor model* corresponding to the minimum-phase spectral factor. In this way the minimum-phase requirement is translated to a stability requirement on the system matrices of the innovation predictor model and the minimum-phase spectral factor, $A - K\Sigma_1 C$ and A, respectively. Another advantage of the proposed subspace algorithm is that it is based on a method originally developed for closed-loop identification. As a result, the algorithm can be easily extended to the case that the WFS data are collected in closed-loop. This is especially important since this may be used to update the disturbance model without interrupting the observations.

To arrive at an efficient implementation of the subspace algorithm, a different weighting scheme is used in reconstructing the state sequence. The original subspace algorithm described by Jansson (2003) uses canonical correlation analysis (CCA) to estimate a state-space basis from the data equation. The proposed subspace algorithm, on the other hand, is based on a weighting scheme similar to the one used in the MOESP class of subspace algorithms (Verhaegen 1994). By considering these weights it is possible to arrive at a square root implementation of the algorithm that avoids the explicit computation of the inverse square-root of the correlation matrix of filtered future measurement data. Apart from the fact that this leads to a more efficient implementation, avoiding the explicit computation of the inverse square-root covariance matrix is attractive from the viewpoint of numerical robustness. In the square-root implementation, a single RQ factorization of the stacked block-Hankel matrix of past and future data can be used both for computing the required AR coefficients and for data compression. Indeed, the entire subspace algorithm can be expressed in terms of the R factor of the block-Hankel matrix. A further efficiency improvement is achieved by exploiting the displacement structure of the stacked block-Hankel matrix. To this end, the proposed identification algorithm uses the fast RQ factorization algorithm for block-Hankel matrices as proposed by Mastronardi et al. (2001). This leads to an efficient implementation in terms of the number of flops and the memory requirements.

Another important issue in making the identification of a full multi-variable atmospheric disturbance model computationally feasible for a realistic number of WFS channels, is the choice of the number of block rows in the proposed subspace algorithm. As in most subspace algorithms, the number of block-rows used in constructing the block-Hankel matrices of past and future data are important user defined parameters that have a large influence on the computational complexity, especially in the multi-variable case. When the stochastic process is generated by a finite-dimensional linear stochastic process with a minimal state-space representation of order n, a sufficient condition to guarantee that the extended observability matrix can be estimated from the data equations is that the number of future and past block rows, $f \in \mathbb{N}$ and $p \in \mathbb{N}$ are chosen larger than n. For this reason, it is common practice to choose the parameters f and p much larger than the expected model order and the order of the identified model. For multi-variable systems, this condition is usually overly conservative. To reduce the computational load we will generally choose f, p much smaller than the order of the identified disturbance model. Finally, in identifying an atmospheric disturbance model it is

advantageous to account for the fact that only a part of the wavefront can be reconstructed from the WFS measurements. By introducing a new set of basis functions it is possible to reduce the dimension of both the WFS signal and the vector representation of the wavefront, without loosing any information. Apart from reducing the dimension of the identification problem, this is also advantageous as it leads to an improved numerical conditioning of the identification problem.

The proposed subspace identification algorithm has been thoroughly tested in numerous simulation experiments. To assess whether or not the algorithm is suitable for atmospheric disturbance modeling, especially the simulation experiments based on open-loop WFS are of interest. The data used in these experiments have been obtained from both an AO laboratory setup at TNO Science and Industry, the Netherlands, and from real turbulence measurements with the JOSE (Saint-Jacques 1998) seeing-monitor at the William Herschel Telescope situated on Las Palma, Canary Islands. The performance of the proposed subspace algorithm has been evaluated by first identifying a disturbance model and then using this model to predict open-loop wavefront distortions one sample in advance. The performance of this one-step ahead predictor is compared with the optimal one-step ahead predictor for a random walk or Wiener process (Shanmugan and Breipohl 1988). For a Wiener process, the optimal one-step ahead predictor is equal to the current value of the sample, that is, the change in the signal is completely unpredictable. Since atmospheric turbulence has basically a low pass spectrum, the random walk predictor usually has already a reasonably good performance.

The close link between wavefront prediction and control has already been pointed out before. Indeed, it can be shown that the performance of the random walk predictor is equal to the performance of the common AO control approach in which the temporal compensator is a pure integrator. The simulation experiments show that the proposed subspace algorithm is able to identify a full multi-variable atmospheric disturbance model for an AO system with in the order of hundred active WFS spots. Furthermore, the one-step ahead predictor based on the identified atmospheric disturbance model is able to achieve a considerable reduction in the mean-square wavefront prediction error when compared to the random walk approach. The identified disturbance model provides an adequate model of the atmosphere, which can be used for controller design. Finally, the simulation experiments on the JOSE data confirm that the atmospheric wavefront distortions can be regarded wide sense stationary; at least on a time scale in the order of a few minutes.

1.9.3 $\mathcal{H}_2$-optimal controller design for AO

As explained in Section 1.6, the identified atmospheric disturbance model is used as one of the basic ingredients in computing the optimal controller. The problem of computing the optimal controller is the central issue in the second part of Chapter 3. Suppose for the moment that the transfer function from control input to the wavefront correction reconstructed from the WFS measurements, $\mathcal{H}(z)$, can be described by a LTI system. Then, by using an atmospheric disturbance model, the

problem of finding the optimal controller that minimizes the mean-square residual wavefront error, can be conveniently expressed in a $\mathcal{H}_2$-optimal control framework. Since the $\mathcal{H}_2$-optimal control framework makes a clear distinction between performance outputs and measured outputs, it is perfectly suitable to account for the fact that the WFS is not able to directly measure the phase.

After having formulated the AO control problem in a $\mathcal{H}_2$-optimal control framework, standard $\mathcal{H}_2$-optimal control theory can be used to compute the closed-loop optimal controller. This generally involves the numerical solution to two Riccati equations. Due to the special structure of the AO control problem it is, however, possible to simplify the computations. In the second part of Chapter 3, it will be shown that, because of the minimum-phase property of the spectral factor, at least one of the Riccati equations can be avoided. Also the second Riccati equation can be avoided if the model of the DM and WFS dynamics, $\mathcal{H}(z)$, is minimum-phase or has a known inner-outer factorization. It will be shown that this is the case if the transfer function is of the form $\mathcal{H}(z) = z^{-d}(1 + \alpha z^{-1})H$. As a result it is possible to derive an analytical expression for the $\mathcal{H}_2$-optimal controller in this case.

Having an analytical expression for the optimal controller is attractive both from a computational and a numerical robustness point of view. Solving the large-scale Riccati equations associated with the standard solution to the $\mathcal{H}_2$-optimal control problem may become impractical, especially since the changing turbulence conditions may require a regular update of the controller (Looze et al. 2003). From a numerical point of view, the analytical expressions are attractive since, as a result of the particular nature of turbulence spectrum, the poles of the disturbance model typically cluster in the vicinity of the point $z = -1$. Standard Riccati solvers may suffer from convergence problems and increased sensitivity when the poles are located too close to the unit circle. Furthermore, the numerical sensitivity of most Riccati solvers typically increases with the order of the system. By using analytical solution these problems can be avoided. For more information on numerical methods for solving Riccati equations the reader is referred to (Arnold and Laub 1984; Benner et al. 1997) and references therein. Finally, note that the control design procedure obtained by combining the proposed subspace identification algorithm and the analytical expressions for the $\mathcal{H}_2$-optimal controller is exclusively based on standard matrix manipulations. It provides a non-iterative procedure to go from open-loop WFS data to optimal control design.

Apart from the computational advances, the analytical expressions are attractive as they provide more insight in the precise structure of the optimal controller. From the derived expressions it is clear that the optimal controller can be decomposed in a part that is concerned with the problem of predicting the uncorrected wavefront distortions over a time-horizon of the pure delay z^{-d} in the transfer function $\mathcal{H}(z)$, and a dynamic filter that is responsible for projecting the estimated wavefront on the actuator space. The dynamic filter is basically a regularized inverse of the remaining part of the transfer function $\mathcal{H}(z)$, i.e. the two FIR taps $(1 + \alpha z^{-1})H$, where the regularization makes a trade off between the objective of finding the best actuator inputs to compensate for the estimated wavefront and the objective of minimizing the contribution of the control effort to the cost function. The above interpretation confirms the close relation between wavefront pre-

diction and optimal control. Indeed, if $\mathcal{H}(z)$ consists of a pure delay the optimal controller reduces to a d-step ahead predictor and a static mapping onto the actuator space. Furthermore, the above interpretation of the optimal controller shows the close resemblance with the common AO control approach that also decomposes in a wavefront estimation step and a projection on the actuator space. The main difference between both approaches, however, is that the static wavefront reconstructor and the temporal compensator in the common AO control approach are replaced by a dynamic wavefront prediction scheme that uses current and past WFS measurements from all WFS channels to exploit the spatio-temporal correlation in the wavefront. Another important difference is that the wavefront projection on the actuator space becomes dynamic, whenever the $\mathcal{H}(z)$ deviates from being a pure delay. At the end of Chapter 3, the proposed data-driven optimal control approach is demonstrated by simulations on open-loop WFS data.

1.9.4 Experimental validation of $\mathcal{H}_2$-optimal control strategy

In order to demonstrate the practical feasibility of the proposed data-driven $\mathcal{H}_2$-optimal control approach, it has been implemented on the previously mentioned AO laboratory setup. Validation experiments with this setup serve as a proof of concept for demonstrating the ability of optimal control to improve the wavefront disturbance rejection by exploiting the spatio-temporal correlation in the wavefront. Furthermore, the experiments are used to gain a better insight in the conditions under which optimal control is able to outperform the common AO control law. To this end, the performance of the optimal controller is compared with the common AO control law. In the discussion so far, the transfer function $\mathcal{H}(z)$ from control input to the reconstructed wavefront, has been assumed to be given. In reality however, the model of the DM and WFS dynamics is not known in advance.

Apart from describing the validation experiments, Chapter 4 is concerned with modeling the transfer function $\mathcal{H}(z)$. Accurate modeling of the wavefront correction link is of utmost importance, as it is a part of the control loop. As a result, model errors on $\mathcal{H}(z)$ have a large influence on the overall performance and may, in the worst case, even destabilize the loop. Moreover, the gain in performance brought about by the predictive capabilities of the atmospheric disturbance model may be completely lost by inaccurate modeling of the wavefront correction link. On the other hand, in order to apply the derived analytical expressions for the $\mathcal{H}_2$-optimal controller, the transfer function $\mathcal{H}(z)$ has to possess the special structure of scalar dynamics consisting of an integer number of samples delay and a two tap FIR impulse response. These requirements may seem very conflicting and restrictive. By analyzing the WFS sampling process, however, it is shown that the assumed model structure is still rather general. Indeed, any AO system can be modeled in this way when it has a wavefront correction device that can be considered static in the sense that its time-constant is short compared to the WFS exposure-time. Since the transfer function is influenced by many factors (like the precise alignment of the WFS) the relevant parameters in the transfer function $\mathcal{H}(z)$ are preferably estimated on the basis of measurement data. Hence, also in modeling the AO system we adhere to a data-driven identification approach. Also in the

common AO control approach the influence matrix H is often estimated on the basis of measurement data collected in identification experiments. The problem of identifying the unknowns α and H in the transfer function $\mathcal{H}(z) = z^{-d}(1+\alpha z^{-1})H$ is formulated as finding the parameters that minimize a prediction-error criterion. Since $\mathcal{H}(z)$ contains a term with the product αH, this gives rise to a non-convex optimization problem.

In Chapter 4, it will be shown that the resulting optimization problem can be efficiently solved by using the concept of separable least squares (Golub and Pereyra 1973). This gives rise to an efficient data-driven identification procedure to identify the relevant parameters in the structured transfer function $\mathcal{H}(z)$. From a system theoretical point of view it is interesting to note that the different parts of the generalized plant model, i.e. the atmospheric disturbance model and the model of the wavefront correction link, are identified separately. Separate identification has the advantage that more structure is added to the generalized plant model, which reduces the overall size of the identification problem. Indeed, both models have their own state without any cross-coupling between them. Another advantage of considering separate identification is that the atmospheric disturbance model may require a regular update to cope with the changing turbulence conditions, while the mirror can be considered constant. Finally, separate identification is expected to give more accurate results as there is no need to distinguish between the contributions of turbulence and the mirror to the WFS measurements.

In validating both control approaches on the AO laboratory setup, a wide variety of performance measures has been used. The considered performance criteria are based on measurements from the WFS as well as on images obtained from the science camera. The performance criteria derived from the WFS measurements include an estimate of the mean-square residual wavefront error and a sample estimate of the cost function. The second class of performance criteria provides a measure of the optical quality of the science image and includes an estimate of the full-width of half maximum (FWHM), the improvement in Strehl ratio and the normalized encircled energy. Each of these performance criteria has been evaluated over a large range of Greenwood to sample frequencies. The different performance measures are all consistent in that the performance improvement at low Greenwood to sample frequency ratios is rather moderate, while a considerable gain in performance is observed at high ratios. It is shown that this behavior can be related to a reduction of the temporal error.

The dominant error source in the AO laboratory setup depends on the actual turbulence conditions. By analyzing the error sources it is shown that the total mean-square residual wavefront error can be decomposed in the wavefront fitting error and the temporal error. The wavefront fitting error depends only on the turbulence strength and the wavefront correction device, and can hence be considered constant. The temporal error, on the other hand, is an exponentially increasing function of the Greenwood to sample frequency ratio and is strongly influenced by the controller design. As a result, at low Greenwood to sample frequency ratios the wavefront fitting error is the dominant error source and little can be achieved by using optimal control. At high Greenwood to sample frequency ratios, however, the temporal error becomes the limiting factor. By exploiting

the spatio-temporal correlation in the wavefront, optimal control is able to reduce this error source. Optimal control is only able to achieve a considerable performance error reduction if one of the dominating error sources is influenced by the controller design. As a result, optimal control will be especially beneficial in conditions like violent turbulence and weak guide stars, which require a low sample frequency.

2 | CHAPTER

Robust Spectral Factor Approximation of Discrete-Time Frequency Domain Power Spectra

This paper presents a subspace-based identification algorithm for estimating the state-space quadruple $[A, B, C, D]$ of a minimum-phase spectral factor from matrix valued power spectrum data. The key step in the algorithm is the preservation of the positive realness (PR) property of the estimated power spectrum derived from the spectral factor. For a given pair $[A, C]$ with A stable, this PR property is guaranteed via the solution of a conic linear programming (CLP) problem. In comparison with the classical LMI-based solution, this results in a more efficient way to minimize the weighted 2-norm of the error between the estimated and given power spectrum.

The property that the CLP problem can be solved in a globally optimal sense, is exploited in the derivation of a separable least squares procedure for the (local) minimization of the above 2-norm with respect to the parameters of a parametrization of the minimum-phase spectral factor.

The advantages of the derived subspace algorithm and the iterative local minimization procedure are illustrated in a brief simulation study. In this study the effect of dealing with short length data sets for computing the power spectrum, on the estimated spectral factor, is illustrated.

Keywords: Spectral factorization; Subspace identification; Power spectra; Stochastic realization; Conic linear programming

2.1   Introduction

Estimating a minimum-phase spectral factor from a matrix valued discrete-time frequency domain power spectrum is a challenging problem. In several applications it is important to have an accurate estimate of the spectral factor of a stochastic process. In active vibration suppression for instance, the spectral factor of the disturbance source plays an important role in designing the optimal feedback controller (Fraanje, Verhaegen, Doelman, and Berkhoff 2004). Furthermore, it has been demonstrated that the spectral factor of the disturbances measured by an accelerometer may be useful in determining the modal parameters of civil engineering structures (Kirkegaard and Andersen 1997). One can think of a number of situations in which only the power spectrum of a stochastic process is known. This is, for instance, the case when considering the optical wavefront fluctuations introduced by atmospheric turbulence, which are theoretically described by a Kolmogorov spectrum. To deal with such stochastic processes it is desirable to have an algorithm that is able to estimate the minimum-phase spectral factor directly on the basis of power spectrum data. A promising approach to tackle this problem is based on subspace identification.

Subspace identification has its origin in stochastic state-space realization theory as developed in the 1960s. Some of the main principles on which subspace identification is based can already be recognized in the classical contribution by Ho and Kalman (1966). Over the years, the concept of subspace identification has been refined and by now a large number of algorithms is available (Larimore 1990; Verhaegen 1994; Van Overschee and De Moor 1996; Viberg 1995). A concise overview of different subspace identification algorithms is provided by Viberg (1995). Also for the identification of stochastic systems, a large variety of subspace algorithms have been proposed (Aoki 1990; Van Overschee and De Moor 1993; Mari, Stoica, and McKelvey 2000). Most of these algorithms however operate on time-domain data. Up to our knowledge, the algorithm described by Van Overschee, De Moor, Dehandschutter, and Swevers (1997) is the only subspace algorithm to identify a stochastic system directly from discrete-time frequency domain power spectrum data.

The use of subspace identification algorithms has a number of important advantages over classical prediction error methods (Ljung and Glad 1994). Prediction error methods are based on the optimization of system parameters over a suitable cost function, which is usually a highly nonlinear optimization problem. Due to the existence of local minima and non-convexity, the performance may be very sensitive to both the starting point in the optimization procedure and the chosen model parametrization. Subspace identification does not suffer from these problems. There is no need for model parametrization and no iterative nonlinear optimization is required. In contrast, they rely on numerically robust tools of linear algebra and their computational complexity is modest. Finally, subspace identification algorithms inherently treat the multi-variable case.

This paper presents an algorithm for estimating a minimum-phase spectral factor that accurately approximates a given power spectrum. Since the objective is to determine a spectral factor that provides a good match to the given spec-

tral data, we will refer to the algorithm as the spectral factor approximation algorithm. The algorithm includes a number of provisions to assure a reasonable performance in the case that the given power spectrum cannot be exactly described by a finite-dimensional LTI spectral factor. This is for instance the case for non-rational and estimated power spectra. The proposed estimation algorithm roughly consists of a subspace identification algorithm, supplemented by a parametric optimization procedure. Here, the subspace identification algorithm is used to provide an accurate initial estimate of the spectral factor, which is then further optimized with respect to the desired cost function. In the latter step, the spectral factor is parametrized in output normal form. Since this parametrization only represents the class of stable systems, the additional optimization over the cost function will not jeopardize the stability of the spectral factor that is explicitly enforced on the obtained initial estimate by using the Schur re-stabilization procedure as proposed in (Mari, Stoica, and McKelvey 2000). The subspace identification algorithm differs in several aspects from one proposed by Van Overschee et al. (1997). The main contributions presented in this paper are:

- It will be show that in relating the inverse discrete Fourier transform (IDFT) of the given power spectrum to the system matrices of the spectral factor, it is advantageous to introduce an approximation which amounts to neglecting the effect of aliasing. In comparison with the algorithm described by Van Overschee et al. (1997), this avoids the need to carry out a numerically very sensitive split. As a result, the introduced approximation results in better estimates of the pair $[A, C]$ and hence of the quadruple $[A, B, C, D,]$. Furthermore, the proposed simplification is not restricted to square Hankel matrices, which is beneficial from a computational point of view.

- A key issue in the identification of stochastic systems is to ensure that the estimated spectrum is positive semi-definite on the unit circle. This property is needed to ensure that the estimated spectrum has a minimum-phase spectral factor. The requirement imposes a constraint on the estimation of the pair $[B, D]$. It will be shown that for a given estimate of the pair $[A, C]$, the problem of finding the optimal values of the pair $[B, D]$ can be formulated as a conic linear program (CLP). This approach is numerically far more efficient than solving formulating the estimation problem as a large semi-definite program or a nonlinear least squares problem (Mari et al. 2000; Van Overschee et al. 1997).

- The ability to solve the CLP in a global optimal sense is exploited to derive a numerical analogue of the separable least squares principle (Golub and Pereyra 1973; Ruhe and Wedin 1980). Separable least squares reduces the number of parameters that have to be optimized in the parametric optimization. It is know that this may result in better numerical conditioning of the optimization problem and less iterations (Bruls et al. 1999; Sjöberg and Viberg 1997; Ribarits et al. 2003).

The remainder of this paper is organized as follows. In Section 2.2, we will first provide an accurate description of the considered discrete-time power spectrum identification problem and introduce the necessary notation. Sections 2.3, 2.4

and 2.5 will subsequently deal with each of the three items discussed above. In Section 2.6, the performance of the proposed algorithm is demonstrated by means of two simulation experiments. In the first simulation experiment the algorithm is applied to a spectrum estimated on the basis of a small amount of measurement data. During this experiment the performance robustness of the proposed algorithm with respect to estimation errors has been demonstrated. The second simulation experiment considers the problem of approximating a non-rational power spectrum, as encountered in turbulence modeling. The paper concludes with a short discussion in Section 2.7.

2.2   Problem formulation and notation

This section provides a description of the stochastic identification problem that is considered in this paper. Before focusing on the actual problem formulation, we will first consider some notational issues. In this paper, X^T is used to denote the transpose of the matrix X and X^* denotes its conjugate transpose. Furthermore, $\|X\|_F$ denotes the Frobenius norm of X and $X^\dagger$ its Moore-Penrose pseudo inverse. $\lambda_i(X)$, $i \in \{1, \ldots, m\}$ are the eigenvalues of the real-valued $m \times m$ matrix X. The Euclidean norm of a vector x is denoted by $\|x\|_2$. $\mathcal{E}\{x\}$ denotes the expected value of a process x, with respect to the underlying probability distribution. $\hat{x}$ is an estimate of x. The Kronecker product between two matrices X and Y is denoted by $X \otimes Y$. Furthermore, we use the notation $X \succcurlyeq 0$ to denote that the matrix X is positive semi-definite. The shorthand $\text{vec}(X)$ is introduced for the operator that stacks the columns of a matrix $X \in \mathbb{R}^{k \times m}$ in a km-dimensional column vector. Each conjugate symmetric matrix X can be split as $X = L + D + L^*$, where the matrix L is strictly lower-triangular and D is diagonal. The operator $\text{vec}_\text{s}(X)$ performs this split and stacks the nonzero elements of $(L + D/2)$ in a $l(l+1)/2$-dimensional vector. The working of the operator is illustrated in the following example:

$$X = \begin{bmatrix} 2 & 3 - j \\ 3 + j & 4 \end{bmatrix} \iff \text{vec}_\text{s}(X) = \begin{bmatrix} 1 \\ 3 + j \\ 2 \end{bmatrix}. \tag{2.1}$$

Consider a stationary stochastic process $y(k) \in \mathbb{R}^l$ with discrete-time index $k \in \mathbb{N}$ and power spectrum $\Phi(z)$. Furthermore, assume that there exists a stable linear time invariant (LTI) system $S(z)$ satisfying the relation $\Phi(z) = \Phi^S(z)$ where

$$\Phi^S(z) \doteq S(z)S^T(z^{-1}). \tag{2.2}$$

The system $S(z)$ provides a model of the stochastic process $y(k)$ in the sense that the signal, obtained by filtering a unit variance white noise sequence $u(k)$, has the same second order statistics. When the system $S(z)$ has a stable inverse (i.e. $|\lambda_i(A)| < 1$ and $|\lambda_i(A - BD^{-1}C)| < 1$, $\forall i \in \{1, \ldots, n\}$), it is called a minimum-phase spectral factor of $\Phi(z)$. Let the minimum-phase spectral factor $S(z)$ have a

state-space realization

$$x(k+1) = Ax(k) + Bu(k) \tag{2.3}$$
$$y(k) = Cx(k) + Du(k), \tag{2.4}$$

where $x(k) \in \mathbb{R}^n$ denotes the state-vector and $u(k) \in \mathbb{R}^l$ and $y(k) \in \mathbb{R}^l$ are the input and output signals, respectively. In the following it will be assumed that the state-space realization of $S(z)$ is minimal and we will briefly refer to the state-space realization as $\{A, B, C, D\}$. The state-space realization $\{A, B, C, D\}$ and the power-spectrum are related through equation (2.2) via the transfer function representation

$$S(z) \doteq C(zI - A)^{-1}B + D. \tag{2.5}$$

For practical reasons, it is convenient to introduce an alternative expression for the power spectrum $\Phi(z)$ in terms of the system matrices $\{A, B, C, D\}$. Let us define the transfer function $L(z)$ and the matrix $Q \in \mathbb{R}^{n+l}$ as follows:

$$L(z) \doteq \left(\begin{array}{c|c} C(zI - A)^{-1} & I \end{array} \right) \quad Q \doteq \begin{pmatrix} B \\ D \end{pmatrix} \begin{pmatrix} B^T & D^T \end{pmatrix}. \tag{2.6}$$

From equation (2.5) and (2.2) one can easily infer that the power-spectrum associated with the input-output system (2.3)-(2.4) can be expressed as

$$\Phi^S(z) = L(z)QL^T(z^{-1}). \tag{2.7}$$

This representation of the power spectrum separates the system matrices into two groups. The transfer function $L(z)$ depends only on the pair $\{A, C\}$, while the matrix Q depends only on the pair $\{B, D\}$. This partitioning is of particular interest in identifying the system matrices. It is important to note that not every real valued matrix Q corresponds to a valid power spectrum. From equation (2.6) it is clear that Q has to be symmetric and positive semi-definite (i.e. $Q \succeq 0$). Only this class of Q matrices yields a valid spectrum that can be factorized to recover the system matrices B and D that describe the spectral factor. This constraint on the class of valid Q matrices corresponds to the positive realness condition, which requires that $\Phi(z) \succeq 0$ for all $|z| = 1$. Any physically meaningful spectrum has to satisfy this condition. Finally, note that besides the above formulation, the power spectrum can also be defined as the z-transform of the covariance function $R_k \doteq \mathcal{E}\{y_{k+m}y_m^T\}$, which is given by (Mari et al. 2000)

$$R_k = \begin{cases} CPC^T + DD^T & \text{for } k = 0 \\ CA^{k-1}G & \text{for } k \geq 1 \end{cases} \tag{2.8}$$

and $R_{-k} = R_k^T$, where $G \doteq APC^T + BD^T$ and $P \doteq \mathcal{E}\{x(k)x(k)^T\}$, the state covariance matrix, can be obtained as the solution to the discrete Lyapunov equation $P = APA^T + BB^T$.

The stochastic identification problem can now be formulated as follows. Given $(N + 1)$ equidistantly distributed samples $\Phi_k \in \mathbb{C}^{l \times l}$ of power-spectrum $\Phi(z)$, i.e.

$$\Phi_k = \Phi(e^{j(2\pi k/2N)}), \ k = 0, \ldots, N \ ,$$

find a state-space realization $\{A, B, C, D\}$ of the unit variance minimum-phase spectral factor $S(z)$ that provides the best fit to samples Φ_k. The accuracy of the fit is quantified by means of the cost function

$$J \doteq \sum_{k=0}^{N} \mathtt{vec_S} \left(\Phi_k - \Phi_k^S\right)^* W_k \, \mathtt{vec_S} \left(\Phi_k - \Phi_k^S\right) \tag{2.9}$$

where $\Phi_k^S \doteq \Phi^S(e^{j(2\pi k/2N)})$ and $W_k \in \mathbb{R}^{l(l+1)/2}, W_k \succcurlyeq 0$ is a suitably chosen frequency dependent weighting function. This cost function is a quadratic form with full weighting at each frequency. The frequency dependent weighting might be useful when the samples Φ_k are estimated empirically. Increasing the penalty in frequency regions where the variance is expected to be small, makes it possible to improve the accuracy with respect to the true underlying power spectrum. When the estimator used to determine the power spectrum is unbiased, has a Gaussian distribution and provides estimates $\hat{\Phi}_k$ that are uncorrelated for different frequencies, we may turn the cost function into a maximum likelihood type of criterion by choosing W_k equal to the inverse of the variance of the estimate. Frequency weighting is also interesting in the case that the samples Φ_k are obtained from a process that does not belong to the class of linear finite-dimensional stochastic systems. Here frequency weighting may be used to specify a region of interest.

In the above discussion it is assumed that the system $S(z)$ is square. There exist, however, situations in which the number of noise inputs required to model the stochastic process is smaller than the number of output channels l. The problem of modeling a stochastic process with the minimal number of white noise inputs is non-trivial and is beyond the scope of this paper. For time-domain state-space identification, this problem has been addressed by Johansson et al. (2001).

2.3   Initial estimate of the A and C matrices

As outlined in the introduction, the proposed spectral factor approximation algorithm uses subspace identification to obtain an initial estimate of the system parameters. This section describes the part of the subspace algorithm that is used to provide an initial estimate of $\{A, C\}$. The approach is related to the algorithm of Van Overschee et al. (1997) and the main differences will be discussed.

The first step in the algorithm is to expand the given samples Φ_k to $2N$ points over the entire unit circle. By exploiting the periodicity and conjugate symmetry of the spectrum, we have $\Phi_k = \Phi_{2N-k}^*$ for $k = N+1, \ldots, 2N$, where * denotes the conjugate transpose. The IDFT of the power spectrum can now obtained as:

$$\phi_m = \frac{1}{2N} \sum_{k=0}^{2N-1} \Phi_k \, e^{j(2\pi mk/2N)} \, . \tag{2.10}$$

The following theorem provides the relation between the IDFT coefficients ϕ_m and the system matrices of $S(z)$, and forms the basis of the algorithm.

Theorem 2.1 (Van Overschee et al. (1997)) *Let G be defined as in equation (2.8) and let $M \doteq (I - A^{2N})^{-1}$, then the IDFT ϕ_k, $m \in 1, \ldots, 2N - 1$ of the extended sequence Φ_k is given by:*

$$\phi_0 = R_0 + CA^{2N-1}MG + G^T(A^T)^{2N-1}M^TC^T \tag{2.11}$$
$$\phi_m = CA^{m-1}MG + G^T(A^T)^{2N-m-1}M^TC^T \tag{2.12}$$

An important observation is that the IDFT coefficients ϕ_m in (2.12) consists of two terms. Since the minimum-phase spectral factor S is stable, the effect of the second term becomes negligible whenever $2N - m - 1$ gets sufficiently large. This gives rise to the result summarized in the following corollary:

Corollary 2.1 *Define the block-Hankel matrix $H_{(q,s)}$ as:*

$$H_{(q,s)} \doteq \begin{bmatrix} \phi_1 & \phi_2 & \ldots & \phi_s \\ \phi_2 & \phi_3 & \ldots & \phi_{s+1} \\ \vdots & & \ddots & \vdots \\ \phi_q & \phi_{q+1} & \ldots & \phi_{q+s-1} \end{bmatrix}, \tag{2.13}$$

where the number of block rows q and block columns is at least the system order n (i.e. $n \leq q, s < \infty$). Then:

$$\lim_{N \to \infty} H_{(q,s)} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix} \begin{bmatrix} G & AG & \cdots & AG^{s-1} \end{bmatrix} = \mathcal{O}\mathcal{C}, \tag{2.14}$$

with $\mathcal{O}$ the extended observability matrix and $\mathcal{C}$ the extended controllability matrix. From equation (2.8) it is clear that in the limit $N \to \infty$ the block-Hankel matrix (2.13) is a matrix of process covariances R_k.

The factorization of $H_{(q,s)}$ as given in Corollary 2.1 holds only in the limit $N \to \infty$. When N is sufficiently large however, it is possible to choose q and s such that the effect of the second term in equation (2.12) is negligible for all $m \leq q + s - 1$. With each block being well approximated by $\phi_m \approx CA^{m-1}MG$, we have the approximate factorization $H_{(q,s)} \approx \mathcal{O}M\mathcal{C}$. This implies that the block-Hankel matrix $H_{(q,s)}$ can be factorized to obtain an estimate the column space of $\mathcal{O}$. By exploiting the shift invariant structure of the extended observability matrix $\mathcal{O}$ (see, e.g. Verhaegen (1994)), the pair $\{A, C\}$ may be determined up to a similarity transformation. To make this more explicit, consider the singular value decomposition (SVD) of $H_{(q,s)}$

$$H_{(q,s)} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}, \tag{2.15}$$

where the diagonal matrix $\Sigma_1 \in \mathbb{R}^{n \times n}$ contains the dominant singular values and U_1 and V_1^T denote matrices composed of the corresponding singular vectors. Since

the spectral factor is assumed to be minimal, the rank of the matrix $H_{(q,s)}$, and hence the number of singular values in Σ_1, provides an estimate of its order n. Furthermore, the columns of $U_1\Sigma_1^{1/2}$ provide a basis for the column space of $H_{(q,s)}$. This, in combination with Corollary 2.1, implies that

$$\widehat{\mathcal{O}} = U_1\Sigma_1^{1/2}T,$$

for some non-singular matrix T. The estimated extended observability matrix $\widehat{\mathcal{O}}$ determines the system matrices A and C up to the similarity transformation T. When T is set to identity, it follows from the definition of $\mathcal{O}$ that an estimate of the matrix C can be obtained by reading off the first l rows of $\mathcal{O}$. An estimate of the matrix A, on the other hand, can be determined from the set of overdetermined equation, $\overline{\mathcal{O}}A = \underline{\mathcal{O}}$, where $\overline{\mathcal{O}}$ and $\underline{\mathcal{O}}$ are the matrices obtained by deleting the l first and last rows of the matrix $\mathcal{O}$, respectively. Even though the described subspace approach often yields a stable A matrix, stability is by no means guaranteed. In order to avoid unstable A matrices, the Schur re-stabilization procedure as described in Mari et al. (2000) is applied if necessary. This rescue is especially important when working on noisy power spectra and power spectra that cannot be attributed to a finite dimensional linear system.

The subspace approach used to obtain an initial estimate of the pair $\{A, C\}$ differs from the power spectrum identification algorithm proposed by Van Overschee et al. (1997). Figure 2.1 provides a schematic representation of the main steps in both algorithms. The steps that correspond to our approach are enclosed in the shaded block.
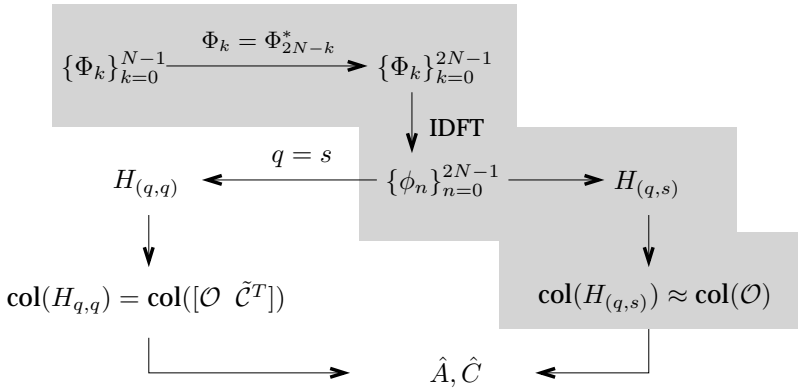
Figure 2.1: Comparison between the applied subspace identification approach and the subspace identification approach described by Van Overschee et al. (1997). The approach described in this paper is enclosed in the shaded area.

From Figure 2.1 it is clear that the first few steps, up to and including the calculation of the inverse DFT ϕ_m, are the same. Furthermore, both algorithms are based on equation (2.12). The main difference however is that the algorithm of VanOverschee et al. explicitly accounts for the contribution due to the second

term. In Van Overschee et al. (1997) it is shown that if $q, s \geq 2n$, the column space of the matrix $H_{(q,s)}$ (2.13) is equal to the column space of $[\mathcal{O} \; \tilde{\mathcal{C}}^T]$, with the reversed extended controllability matrix $\tilde{\mathcal{C}}$ defined as $\tilde{\mathcal{C}} \doteq [A^{q-1}G \; \ldots \; AG \; G]$. This implies that the rank of the block-Hankel matrix $H_{(q,s)}$ provides an estimate of two times the system order and that its column space is an approximation of the space spanned by $[\mathcal{O} \; \tilde{\mathcal{C}}^T]$.

The main difficulty of this approach, is to find a similarity transformation $T \in \mathbb{R}^{2n \times 2n}$ that separates the column space of $H_{(q,s)}$ into a part that corresponds to the matrix $\mathcal{O}$ and a part that corresponds to the matrix $\tilde{\mathcal{C}}^T$. Adopting the same notation for the singular value decomposition of $H_{q,q}$ as in equation 2.15, this implies that one has to find a matrix T such that $U_1 S_1^{1/2} T = [\mathcal{O} \; \tilde{\mathcal{C}}^T]$. Finding this transformation appears to be a very delicate and sensitive process, due to the assumption that one of the intermediate matrices has precisely n stable and n unstable reciprocal eigenvalues. This assumption might easily fail, especially when the spectrum is perturbed by noise. Finally it is important to note that in determining the similarity transformation T, it is implicitly assumed that the matrix $H_{(q,s)}$ is square ($q = s$). This is a disadvantage as it results either in large block Hankel matrices or matrices with a relatively small number of block-rows. By neglecting the contribution due to the second term in equation (2.12), the need for finding a similarity transformation that separates the column space of the block-Hankel matrix in two separate parts is avoided. As a result, the proposed subspace algorithm is numerically more robust. Finally, recall that the second term in equation (2.12) can be neglected whenever N is sufficiently large. This condition can be interpreted as the requirement that the given power spectrum $\Phi(e^{j(2\pi k/2N)})$ is sampled with a sufficient resolution. In this perspective, the contribution of the second term can be seen as the analog of aliasing which occurs when a process is under-sampled in the time-domain. So, whenever the power spectrum is sampled with sufficient resolution it is attractive to introduce the proposed approximation.

2.4   Initial estimate of B and D matrices

In this section we consider the second part of the subspace algorithm, which is concerned with providing an initial estimate of the system matrices B and D. The most direct approach would probably be to estimate the matrix G, from the row space of $H_{(q,s)}$. Once the matrices A, C and G are known, it is possible to compute the pair $\{B, D\}$ by solving the Riccati equation:

$$P = APA^T + (G - APC^T)(R_0 - CPC^T)^{-1}(\cdot)^T \qquad (2.16)$$

where R_0 is obtained from equation (2.11). The pair $\{B, D\}$ of the minimum-phase spectral factor $S(z)$ is now obtained from the positive definite solution of equation (2.16) as:

$$B = (G - APC^T)(R_0 - CPC^T)^{-\frac{1}{2}}, \; D = (R_0 - CPC^T)^{\frac{1}{2}}. \qquad (2.17)$$

The B and D matrices obtained in this way guarantee that the resulting spectral factor (2.3)-(2.4) is minimum-phase.

The main disadvantage of the above approach of estimating B and D is that the Riccati equation (2.16) may fail to have a positive semi-definite solution. The risk of failure is especially high when the samples Φ_k are corrupted by measurement noise. This problem is closely related to the covariance extension problem, which is extensively discussed by Lindquist and Picci (1996); Dahlén (2001). The infinite covariance sequence synthesized by substituting the estimated triplet $\{A, C, G\}$ in equations (2.8) and (2.11) may not be a valid covariance sequence in the sense that its discrete-time Fourier transform (DTFT) $\hat{\Phi}(z)$ is not positive semi-definite on the unit circle, i.e. the estimated spectrum is not positive real. When this is the case, the estimated spectrum $\hat{\Phi}(z)$ is not physically meaningful and does not have a spectral factor $S(z)$. The same problem occurs in Van Overschee et al. (1997), when the system matrices B and D are determined by estimating G from the column space of $H_{(q,q)}$.

One way to deal with the positive real requirement on the spectrum is to formulate the problem of estimating $\{B, D\}$ as an optimization problem over the cost function (2.9). With the alternative representation of the spectrum (2.7) and a fixed pair $\{A, C\}$, this gives rise to the following constraint least squares problem:

$$Q^* = \arg\min_{Q \succcurlyeq 0} \sum_{k=0}^{N} \left\| W_k^{\frac{1}{2}} \, \mathsf{vecs} \left(\Phi_k - L_k Q L_{-k}^T \right) \right\|_2^2 \tag{2.18}$$

where $L_k \doteq L(\mathrm{e}^{\frac{j2\pi k}{2N}})$ and the constraint $Q \succcurlyeq 0$ ensures that the matrix Q^* can be factorized to find B and D. In Van Overschee et al. (1997) two methods are presented to solve the constraint least squares problem. In the first approach the constraint least squares problem is expressed as semi-definite program. This has the advantage that one is guaranteed to find the global optimum of equation (2.18). A drawback, however, is that the resulting large scale LMIs are computationally very demanding. A similar problem occurs in time-domain stochastic subspace identification. In Mari et al. (2000) it is shown that the positive realness requirement in stochastic identification can be imposed by fitting a filtered version of the process covariances. This results in an LMI having the same structure. To reduce the size of the LMIs, Mari et al. (2000) takes into account only a small number of process covariances. Similarly we could restrict the constraint least squares problem (2.18) to a small number of sample points Φ_k. This kind of simplification, is not very attractive since it throws away valuable data.

The second approach uses a nonlinear least squares solver to optimize the cost function (2.9) directly over B and D. Because of the structure of the matrix Q (see equation (2.6)), the constraint $Q \succcurlyeq 0$ is automatically satisfied. An initial guess of the matrices B and D is obtained by perturbing the estimated R_0 to such an extend that the Riccati equation has a positive semi-definite solution. This method is computationally more efficient but it may end in a local minimum.

A more efficient way to solve the constraint least squares problem (2.18) is to express it as a conic linear program (CLP) (Ben-Tal and Nemirovski 2001). The

Matlab toolbox SEDUMI (Sturm 2002) provides an efficient tool to find numerical solutions to this kind of problems. This results in an algorithm that is far more efficient than the LMI approach. Moreover, since the approach is based on convex optimization, it will be able to find the global optimum with respect to the matrices B and D. Before deriving the CLP, we will introduce some definitions.

We are particularly interested in two types of symmetric cones; the cone of positive semi-definite matrices (denoted by $\mathcal{K}_s$) and the Lorentz cone (denoted by $\mathcal{K}_l$). The most natural way the define the cone of positive semi-definite matrices, would be in terms of matrices. For the purpose of this paper, a definition in terms of vectors is more convenient. This issue is resolved by vectorization and results in the following definition:

$$\mathcal{K}_s \doteq \left\{ \text{vec}\,(X) \mid X \in \mathbb{R}^{w \times w},\ X^T = X \succcurlyeq 0 \right\},$$

$$\mathcal{K}_l \doteq \left\{ \begin{pmatrix} \gamma \\ x \end{pmatrix} \in \mathbb{R} \times \mathbb{C}^v \mid \|x\|_2 \le \gamma \right\}$$

Furthermore, we introduce generalized inequalities. We say that $x \succcurlyeq_\mathcal{K} y$ if and only if $x - y \in \mathcal{K}$, with $\mathcal{K}$ a positive cone. The notation $X \succcurlyeq 0$ without subscript, remains in use as a shorthand to denote positive semi-definite matrices. With the above notation and definitions, we are now ready to formulate the constrained least squares problem of equation (2.18) as a CLP:

Theorem 2.2 *Let P be the projection matrix that maps a vector $\text{vec}\,(X) \in \mathbb{C}^{l^2}$ onto the vector $\text{vec}_\text{s}\,(X) \in \mathbb{C}^{l(l+1)/2}$ (i.e. $\text{vec}_\text{s}\,(X) = P\text{vec}\,(X)$) and define $\tilde{\Phi} \in \mathbb{C}^{(N+1)l(l+1)/2}$ and $\tilde{L} \in \mathbb{C}^{(N+1)l(l+1)/2 \times v}$ as follows:*

$$\tilde{\Phi} \doteq \begin{bmatrix} W_0^{1/2} P\,\text{vec}\,(\Phi_0) \\ W_1^{1/2} P\,\text{vec}\,(\Phi_1) \\ \vdots \\ W_N^{1/2} P\,\text{vec}\,(\Phi_N) \end{bmatrix}, \quad \tilde{L} \doteq \begin{bmatrix} W_0^{1/2} P\mathcal{L}_0 \\ W_1^{1/2} P\mathcal{L}_1 \\ \vdots \\ W_N^{1/2} P\mathcal{L}_N \end{bmatrix} \tag{2.19}$$

with $\mathcal{L}_k \doteq L(\mathrm{e}^{-j(2\pi k/2N)}) \otimes L(\mathrm{e}^{j(2\pi k/2N)})$ and $v = (n+l)^2$ the number of entries in the matrix Q. Furthermore, let A, b and c be defined as:

$$A^T \doteq \begin{bmatrix} -1 & 0_{1\times v} \\ 0 & \tilde{L} \\ \hline 0_{v\times 1} & I_{v\times v} \end{bmatrix}, \quad b \doteq \begin{bmatrix} 1 \\ \hline 0_{1\times v} \end{bmatrix}, \quad c \doteq \begin{bmatrix} 0 \\ \tilde{\Phi} \\ \hline 0_{v\times 1} \end{bmatrix}, \tag{2.20}$$

then finding the solution to the constrained least squares optimization problem of equation (2.18), is equivalent to solving the following CLP:

$$\min_z \left\{ b^T z \mid c - A^T z \succcurlyeq_\mathcal{K} 0 \right\} \tag{2.21}$$

where $\mathcal{K} \doteq \mathcal{K}_l \times \mathcal{K}_s$ is the Cartesian product of a Lorentz cone and a cone of positive semi-definite matrices. The optimal vector z^* and the optimal matrix Q^* are related as $z^* = (\gamma \mid \text{vec}\,(Q^*)^T)^T$.

Proof: Let us first apply the projection matrix P as defined in Theorem 2.19 to get rid of the $\mathtt{vec_S}$ operators in equation (2.18). By using the relation $\mathtt{vec}\,(ABC) = (C^T \otimes A)\mathtt{vec}\,(B)$, we obtain the following equivalent optimization problem:

$$\min_{Q \succcurlyeq 0} \sum_{k=0}^{N} \left\| W_k^{1/2} P \big(\mathtt{vec}\,(\Phi_k) - \mathcal{L}_k \,\mathtt{vec}\,(Q)\, \big) \right\|_2^2,$$

with $\mathcal{L}_k$ defined as in the theorem. Stacking the vectors that correspond to each of the samples and introducing the auxiliary variable γ, gives:

$$\min_{Q} \left\{ \gamma \mid Q \succcurlyeq 0, \ \| \tilde{\Phi} - \tilde{L}\mathtt{vec}\,(Q) \|_2 \leq \gamma \right\}$$

where $\tilde{\Phi}$ and $\tilde{L}$ are defined as in (2.19). With the vector $z \doteq (\gamma \mid \mathtt{vec}\,(Q)^T)^T$, the above optimization problem can be formulated as:

$$\min_{z} \left\{ [1 \mid 0]z \ \middle| \ \begin{bmatrix} 0 \\ \tilde{\Phi} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & -\tilde{L} \end{bmatrix} z \succcurlyeq_{\mathcal{K}_l} 0, \ [0 \mid I]z \succcurlyeq_{\mathcal{K}_s} 0 \right\},$$

with $\mathcal{K}_s$ the cone of positive semi-definite matrices and $\mathcal{K}_l$ the Lorentz cone. By combining the two conic constraints and using the definitions in (2.20), we observe that the constrained least squares problem (2.18) is equivalent to the CLP in equation (2.21). $\qquad\square$

From Theorem 2.2 it is clear that the optimal solution to the constrained least squares problem (2.18) can be found by solving the CLP in equation (2.21). The optimal Q^* satisfies the constraint $Q \succcurlyeq 0$ and can be factorized (see equation (2.6)) to determine a B and D. This factorization is performed by first extracting the matrices BB^T, BD^T and DD^T as the $n \times n$ upper left sub-matrix, the $n \times l$ upper right sub-matrix and the $l \times l$ lower right sub-matrix of Q^*. The sub-matrices are then used to determine the matrices R_0 and G as defined in equation (2.8). As in the approach described at the start of this section, the system matrices B and D can now be obtained from the identified quadruple $\{A, G, C, R_0\}$ by solving Riccati equation (2.16). The positive semi-definite solution P finally determines the B and D by means of equation (2.17). Since $\{B, D\}$ is computed through the positive semi-definite solution of (2.16) the spectral factor is guaranteed to be minimum-phase.

2.5   Optimization of the cost function

The initial estimate obtained in Section 2.3 and 2.4 may be refined by means of parametric optimization. This is especially useful for power spectra that are contaminated by noise or do not belong to the class of finite dimensional LTI systems. In this section we will consider both direct optimization over all system parameters and a numerical variant of the separable least squares approach (Golub and Pereyra 1973; Ruhe and Wedin 1980).

In the additional optimization step, the pair (A, C) is parametrized in output normal form (Hanzon and Peeters 2000) with parameter vector $\theta_{AC} \in \mathbb{R}^{nl}$, while the pair (B, D) is described by stacking all entries in a vector $\theta_{BD} \in \mathbb{R}^{(n+l)l}$. This parametrization has the advantage that it only represents the class of stable systems and observable state-space systems for $-1 \leq \theta_{AC} \leq 1$. The estimate of the spectral factor may therefore be refined by using it as an initial estimate in the following optimization problem:

$$(\theta_{AC}^*, \theta_{BD}^*) = \arg \min_{\theta_{AC} \in [-1,1], \theta_{BD}} J(\theta_{AC}, \theta_{BD}) \tag{2.22}$$

where J is the cost function (2.9) expressed in terms of the introduced parametrization. The most straightforward approach to tackle this problem is to optimize the cost function directly over all $(2n + l)l$-elements in θ_{AC} and θ_{BD}, by using nonlinear optimization techniques. In our implementation, we use the Matlab routine `lsqnonlin`, which is based on a interior-reflective Newton method. This routine, is able to take into account the inequality constraint on the entries of θ_{AC}.

Apart from direct optimization over all $(2n + l)l$-parameters, it is possible to exploit the special structure in (2.22). This approach, which is based on the observation that optimization with respect to θ_{BD} is much easier than optimization with respect to θ_{AC}, can be interpreted as the numerical variant of the separable least squares principle. Empirical evidence shows that separable least squares optimization is more robust and requires less iterations than direct optimization over all parameters (Bruls et al. 1999; Ribarits et al. 2003). From Section 2.4 it is clear that for every fixed θ_{AC}, the problem of determining the optimal θ_{BD}^*, i.e.

$$\Psi(\theta_{AC}) \doteq \min_{\theta_{BD}} J(\theta_{AC}, \theta_{BD}), \tag{2.23}$$

is a convex optimization problem. This gives rise to the following relation between (local) minimizers of $\Psi(\theta_{AC})$ and $J(\theta_{AC}, \theta_{BD})$.

Theorem 2.3 *Let $\theta_{BD}^*(\theta_{AC})$ denote the global minimizer of (2.23) for fixed θ_{AC} and let $J(\theta_{AC}, \theta_{BD})$ be the cost function as defined in equation (2.9). With Ω an open subset of $\mathbb{R}^{nl}$, the following statements hold:*

1. *If $\theta_{AC}^* \in \Omega$ is a local minimizer of $\Psi(\theta_{AC})$ and $\theta_{BD}^* = \theta_{BD}^*(\theta_{AC})$, then $(\theta_{AC}^*, \theta_{BD}^*)$ is a local minimizer of $J(\theta_{AC}, \theta_{BD})$ in $\theta_{AC} \in \Omega$. Furthermore the following equality holds $J(\theta_{AC}^*, \theta_{BD}^*) = \Phi(\theta_{AC}^*)$.*

2. *$(\theta_{AC}^*, \theta_{BD}^*)$ is a global minimizer of $J(\theta_{AC}, \theta_{BD})$ in $\theta_{AC} \in \Omega$, if and only if θ_{AC}^* is a global minimizer of $\Phi(\theta_{AC})$ in Ω and $\theta_{BD}^* = \theta_{BD}^*(\theta_{AC})$.*

Proof: Assume that θ_{AC}^* is a local minimizer of $\Psi(\theta_{AC})$ in Ω and let θ_{BD}^* be the corresponding global minimizer of (2.23), i.e. $\theta_{BD}^* = \theta_{BD}^*(\theta_{AC})$. From the definitions of $\Psi(\theta_{AC})$ and $\theta_{BD}^*(\theta_{AC})$, it is immediately clear that $J(\theta_{AC}^*, \theta_{BD}^*) = \Psi(\theta_{AC}^*)$. Let us now assume that $(\theta_{AC}^*, \theta_{BD}^*)$ is not a local minimizer of $J(\theta_{AC}, \theta_{BD})$ in Ω. With $B_\epsilon(\xi) \doteq \{\zeta | \; \|\xi - \zeta\|_2 \; < \epsilon\}$ an ϵ-environment around the point ξ, this implies that for any $\epsilon > 0$ there exists a $(\tilde{\theta}_{AC}, \tilde{\theta}_{BD})$ such that $[\tilde{\theta}_{AC}^T \; \tilde{\theta}_{BD}^T]^T \in$

$B_\epsilon\left([\theta_{AC}^{*T} \ \theta_{BD}^{*T}]^T\right)$, $\tilde{\theta}_{AC} \in \Omega$, and $J(\tilde{\theta}_{AC}, \tilde{\theta}_{BD}) < J(\theta_{AC}^*, \theta_{BD}^*)$. As the function $\theta_{BD}^*(\theta_{AC})$ is defined as the global minimizer of sub-problem (2.23), we observe that $\Psi(\tilde{\theta}_{AC}) = J(\tilde{\theta}_{AC}, \theta_{BD}^*(\tilde{\theta}_{AC})) \leq J(\tilde{\theta}_{AC}, \tilde{\theta}_{BD}) < J(\theta_{AC}^*, \theta_{BD}^*) = \Psi(\theta_{AC}^*)$. From this inequality, we infer that $\Psi(\tilde{\theta}_{AC}) < \Psi(\theta_{AC}^*)$. Furthermore it is easy to see that $\|\theta_{AC}^* - \tilde{\theta}_{AC}\|_2 < \epsilon$ and hence $\tilde{\theta}_{AC} \in B_\epsilon(\theta_{AC}^*)$. Because the above arguments hold for any $\epsilon > 0$, it is not possible to find an ϵ-environment $B_\epsilon(\theta_{AC}^*)$ such that $\Psi(\theta_{AC}^*) \leq \Psi(\theta_{AC})$ for all $\theta_{AC} \in B_\epsilon(\theta_{AC}^*)$. This is in contradiction with the assumption that θ_{AC}^* is a local minimizer of $\Psi(\theta_{AC})$ in Ω. Therefore $(\theta_{AC}^*, \theta_{BD}^*)$ has to be a local minimizer of $J(\theta_{AC}, \theta_{BD})$ in Ω, which finishes the proof of part 1 of Theorem 2.3. The proof of part 2 can be found in (Golub and Pereyra 1973). $\square$

From Theorem 2.3 it is clear that is reasonable to replace the original optimization problem (2.22) by the following alternative problem:

$$\min_{-1 \leq \theta_{AC} \leq 1} \Psi(\theta_{AC}) = \min_{-1 \leq \theta_{AC} \leq 1} J\left(\theta_{AC}, \theta_{BD}^*(\theta_{AC})\right) \tag{2.24}$$

Note that the alternative optimization problem (2.24) has only nl variables, in comparison with the $(2n + l)l$-dimensions of the original problem. The price paid for this reduction in dimensions is that each function evaluation of $\Psi(\theta_{AC})$ is quite time consuming as it needs the solution of the $(n + l)l$-dimensional optimization problem (2.23). In fact the optimization over the parameter vectors θ_{AC} and θ_{BD} is performed separately, which explains the name separable least squares. As in the direct optimization over all parameters, we use the Matlab nonlinear least squares solver `lsqnonlin` to carry out the optimization over the θ_{AC}-vector.

Once we have determined the (local) optimum of the parameter vectors θ_{AC}^* and θ_{BD}^*, either via direct optimization or by using the separable least squares approach, it is straightforward to determine the system matrices A, B, C and D of the corresponding spectral factor. The used parametrization guarantees that the spectral factor is stable, but the quadruple $\{A(\theta_{AC}^*), B(\theta_{BD}^*), C(\theta_{AC}^*), D(\theta_{BD}^*))\}$ might not be minimum-phase. This issue is resolved, by calculating the matrix products BB^T, BD^T and DD^T and proceeding along the steps as described at the end of Section 2.4. In this way we find new B and D such that the spectral factor is stable, non-minimum-phase and still has the same performance.

2.6   Numerical validation

The performance of the spectral factor approximation algorithm has been demonstrated by means of two simulation examples. This section describes the test procedure and summarizes the most important results. The first simulation example is used to demonstrate the performance of the proposed algorithm in the case that the given power spectrum is in the model class of finite-dimensional LTI systems. Also the robustness to estimation errors on the given power spectrum is considered in this example. The second simulation example considers the problem of approximating a non-rational power spectrum over a certain frequency region of interest. In contrast to the first example, the given power spectrum does no longer belong to the class of systems considered by the spectral factorization algorithm.

2.6.1   Example I: Approximation of an estimated power-spectrum

In the first simulation example we consider a finite-dimensional stochastic system with known spectral factor. The considered spectral factor has order $n = 10$ and describes the displacement measured at $l = 2$ positions of a vibrating plate excited by a broadband noise. The true power-spectrum $\Phi^{tr}(z)$, evaluated at $N = 500$ points, is given by the solid line in Figure 2.2 and shows a number of distinct resonance peaks. Apart from the true spectrum, we consider the spectrum estimated
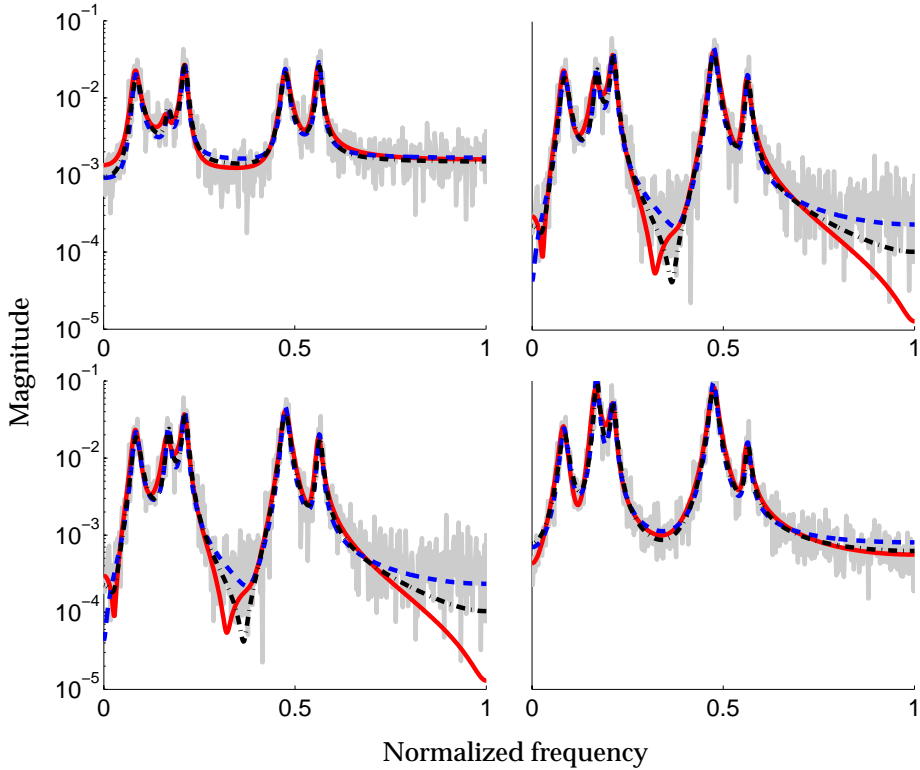
Figure 2.2: True auto and cross power-spectra (dark solid line), Averaged periodogram obtained from simulation experiment (gray noisy line in the background), Power spectrum of initial estimate (dashed line). Power spectrum obtained with the separable least squares variant of the proposed algorithm (dashed-dotted line). The frequency axis has been normalized on the Nyquist frequency.

from 5000 samples $y(k) \in \mathbb{R}^2$ of the stochastic process. The samples are generated by driving the spectral factor with a zero mean unit variance white noise sequence. An estimate of the power spectrum is then obtained by determining the averaged periodogram. To this end, the output sequence $\{y(k)\}_{k=1}^{5000}$ is divided in 5 batches of 1000 samples, i.e. $\{y_i(k)\}_{k=1}^{1000} = \{y_{1000(i-1)+j}\}_{j=1}^{1000}$, where the superscript i is used to distinguish between the different data batches. Using this notation, the

averaged periodogram Φ_k^{av} is calculated as:

$$\Phi_k^{av} = \frac{1}{5} \sum_{i=1}^{5} Y_k^i (Y_k^i)^* \in \mathbb{R}^{2 \times 2} \quad \text{for} \quad k = 0, \ldots, N$$

where Y_k^i is the DFT of $\{y_i(k)\}_{k=1}^{1000}$. This may seem a laborious way of generating noisy spectral data, but for validation purposes it has the advantage that the true spectrum is known. The averaged periodogram Φ_k^{av} used in the simulations is depicted by the gray noisy line in the background of Figure 2.2. It has a large variance, giving rise to a challenging identification problem.

In the tests on the averaged periodogram (i.e. $\Phi_k^{av} = \Phi_k$), the following function has been used as the weight W_k, to accentuate the relative fitting error:

$$W_k^{av} = \begin{bmatrix} \epsilon + |\Phi_{11k}^{av}| & 0 & 0 \\ 0 & \epsilon + |\Phi_{21k}^{av}| & 0 \\ 0 & 0 & \epsilon + |\Phi_{22k}^{av}| \end{bmatrix}^{-2} \tag{2.25}$$

where Φ_{ijk}^{av} denotes the $\{i,j\}^{th}$-element of Φ_k^{av}. The regularization $\epsilon = 10^{-2}$ is included to avoid large weights. In a similar way, the weighting function W_k^{tr} is defined by replacing the averaged periodogram Φ_k^{av} in the above definition with the true power spectrum Φ_k^{tr}. This weighting function will be used in the simulation experiments on the true power spectrum.

Two different performance criteria are used to quantify the performance of the algorithm in the identification experiments. The first performance criterion is basically a scaled version of cost function (2.9), where Φ_k is the power spectrum used as the input in the identification experiments and the weighting function W_k is chosen equal to W_k^{av}. This gives rise to the following definition:

$$P_{av}^w \doteq \frac{2}{N+1} \sum_{k=0}^{N} \text{vec}_{\mathsf{S}} \left(\Phi_k^{av} - \Phi_k^S \right)^* W_k^{av} \, \text{vec}_{\mathsf{S}} \left(\Phi_k^{av} - \Phi_k^S \right), \tag{2.26}$$

where Φ_k^S denotes the power spectrum corresponding to the identified spectral factor. This performance index provides a measure of the fitting error with respect to the averaged periodogram Φ_k^{av}. Furthermore, it is desirable to have a quantitative measure of the fitting error with respect to the true spectrum Φ_k^{tr}. Such a measure is provided by the second performance index P_{tr}^w, which is derived from (2.26) by replacing the functions Φ_k^{av} and W_k^{av} with Φ_k^{tr} and W_k^{tr} respectively.

As an initial test, the proposed algorithm and the power spectrum subspace identification algorithm of Van Overschee et al. (1997) have both been applied to the true spectrum Φ_k^{tr}. The block-Hankel matrix used in the algorithm of Van Overschee et al. (1997) was square with $q = s = 500$, while the proposed algorithm used a rectangular matrix with $s = 20, q = 481$. The results of the simulations are included in second part of Table 2.1. The initial estimate obtained by the proposed algorithm fits the true spectrum to such an extent that there is no need to consider further optimization. Also the computation time on a PC with $2.7\,\text{GHz}$ Intel Pentium® IV processor with 512Mb of internal memory, is included in the

table. Table 2.1 shows that the proposed algorithm performs much better than the algorithm of Van Overschee et al. (1997). The proposed algorithm is faster and gives rise to better results. The relatively large computation time of the algorithm of Van Overschee et al. (1997) is partially caused by the necessity of having a square Hankel matrix (see Section 2.3). This results in large Hankel matrices of which calculating the SVD is computationally intensive.

	P_{av}^w	P_{tr}^w	Time [s]
Properties of Test Data			
True spectrum	7.87×10^{-2}	**0**	–
Averaged periodogram	**0**	7.93×10^{-2}	–
Tests on True Spectrum			
Initial estimate	–	8.58×10^{-17}	44.3
Van Overschee et al.	–	1.74×10^{-1}	80.6
Tests on Averaged Periodogram			
Initial estimate	6.29×10^{-2}	2.22×10^{-2}	**36.4**
All parameters	5.72×10^{-2}	1.79×10^{-2}	**216**
Separable least squares	5.73×10^{-2}	1.80×10^{-2}	1.0×10^4
Van Overschee et al.	–	–	–

Table 2.1: Simulations results. Performance of the proposed spectral factor approximation algorithm and the power spectrum identification algorithm by Van Overschee et al. (1997), with respect to the performance criteria P_{av}^w and P_{tr}^w.

The second set of simulation experiments is concerned with estimating the minimum-phase spectral factor from the noisy periodogram Φ_k^{av}. In these experiments we consider the performance of each of the three possible estimates provided by the proposed spectral factor approximation algorithm, i.e. the initial estimate, the estimate after optimization over all system parameters and the estimate obtained by the separable least squares approach. The results of the experiments are included in Table 2.1. Figure 2.2 shows the power spectra corresponding to the initial estimate (dotted line) and the spectral factor obtained after separable least squares optimization (dash-dotted line). Also the algorithm of Van Overschee et al. (1997) has been tested on the averaged periodogram Φ_k^{av}. The algorithm however failed to produce a sensible result, as it did not succeed in finding a similarity transformation to split the column space of the block-Hankel matrix $H_{(q,s)}$ (see Section 2.3). At a certain point in the procedure used to find the similarity transformation, it is assumed that the block-Hankel matrix has n stable and n unstable eigenvalues. When the spectrum is perturbed by noise, such as the averaged periodogram Φ_k^{av}, this assumption might easily fail. This problem has also been recognized in Van Overschee et al. (1997), where they refer briefly

to an alternative approach based on parametrizing the null space of the extended observability matrix. This approach is however not pursued, as it is numerically less reliable.

From Table 2.1 and Figure 2.2 it is clear that the proposed spectral factor approximation algorithm is indeed able to find a minimum-phase spectral factor that provides an accurate description of the given power spectrum. Furthermore, the algorithm appears to be robust with respect to the estimation errors that are introduced when the power spectrum is estimated on the basis of a finite amount of measurement data. Indeed, the initial estimate obtained by the spectral approximation algorithm provides already a better approximation of the true spectrum than the averaged periodogram that is used as input. The additional parametric optimization step gives rise to a further performance improvement. The performance improvement achieved with direct optimization over all parameters and the numerical variant of the separable least squares approach, is almost identical. The fitting error with respect to P_{av}^w and P_{tr}^w reduces by 9.0% and 18.9%, respectively. These results demonstrate that the numerical variant of the separable least squares principle really works in practice. The separable least squares algorithm appears to be computationally more demanding. The simulation experiments however confirm that separable least squares approach requires less iterations (13 against 25) to reach approximately the same performance. From Bruls et al. (1999); Ribarits et al. (2003) it is to be expected that the separable least squares has advantages in the sense of numerical robustness.

2.6.2 Example II: Approximation of Kolmogorov spectrum

In the second simulation example, we consider the problem of finding the spectral factor that accurately approximates a non-rational power spectrum satisfying a $-8/3$ power law. In contrast to the previous example, the underlying true power spectrum no longer belongs to the class of regular stochastic processes that can be described by a finite-dimensional spectral factor. Indeed, the fractional power in the power law could be expanded in terms of an infinite series, giving rise to a spectral factor of infinite dimension. Since the considered power spectrum is outside the model class considered by the spectral approximation algorithm, an exact match over the entire frequency range is out of the question. The problem is therefore to approximate the given power spectrum over a certain region of interest.

The particular spectrum considered in this simulation study, is the well-known Kolmogorov power spectrum, which usually arises in the theoretical framework of imaging through atmospheric turbulence. Accurate knowledge of the minimum-phase spectral factor that describes the temporal evolution of the optical phase fluctuations, might be useful in adaptive optics (AO) control design. For a single layer of frozen turbulence, the temporal power spectrum of the phase fluctuations at a single point in space in radians squared per Hertz is given by (Hardy 1998; Conan et al. 1995; Glindemann et al. 2000)

$$\Phi(f) = 0.077 r_0^{-5/3} \frac{1}{v} \left(\frac{f}{v} \right)^{-8/3}, \qquad (2.27)$$

where v is the speed of the wind blown frozen layer and r_0 denotes the so-called Fried parameter or turbulence coherence length. In the simulation experiments, we will use the typical values $r_0 = 20 \times 10^{-2}$m and $v = 10$m/s. At this point it is important to note that the Kolmogorov power spectrum has a pole at $f = 0$. Because of this singularity, the integral over the temporal power spectrum and hence the variance of the phase fluctuations, is infinite. This physically non-realistic situation is a well-known property of the Kolmogorov spectrum (Tatarskii 1971). In fact, the Kolmogorov model is known to provide only a reasonable description of the phase fluctuations in the inertial sub-range $v/L_0 \ll f \ll v/l_0$, where L_0 is the outer of turbulence and l_0 is the inner scale at which the viscous dissipation starts. To describe the turbulence outside the inertial sub-range more advanced models, like the Von Kármán phase spectrum, are required.

In this paper, we will restrict our attention to the problem of approximating the Kolmogorov spectrum as given in (2.27). To this end, the spectrum is sampled by evaluating it on a grid of equidistantly distributed frequency points, i.e.

$$\Phi_k = \Phi\left(\frac{k f_s}{2N}\right), \quad k = 1, \ldots, N$$

where f_s denotes the sample frequency of the stochastic process. To avoid the problem of evaluating the power spectrum at $f = 0$, the spectral sample at this frequency, i.e. Φ_0, is chosen equal to the value of Φ_1. This implies that the spectrum that is approximated is chopped off and deviates from the $-8/3$ Kolmogorov asymptote for frequency smaller than $f = f_s/(2N)$. The choice of the sample frequency f_s and the number of samples N, therefore determine the frequency region of interest in approximating the Kolmogorov spectrum. In generating the data for the simulation experiments, the Kolmogorov power spectrum is sampled at $N = 1000$ points and the sample frequency is set to $f_s = 500\,\text{Hz}$. This implies that Φ_k provides undistorted samples of the Kolmogorov power spectrum in the frequency range $f \in [0.25 \quad 250]\,\text{Hz}$. The samples Φ_k, $k = 1, \ldots, N$ used in the identification experiments are included as the crosses in Figure 2.3. To accentuate the relative fitting error and to put more emphasis on the frequencies in the mid-frequency range, the fitting error is weighted by the inverse spectrum and a Gaussian window. This gives rise to the following weighting function

$$W_k = \frac{1}{|\Phi_k|^2} \, e^{-\alpha^2 \left(\frac{k - N/2}{N/2}\right)^2}, \tag{2.28}$$

where the parameter α is set to 2.5. In evaluating the performance, the previously defined cost function P_{tr}^w will be used as a performance criterion.

The proposed algorithm has been used to identify spectral factors of different orders. The number of block rows used to form the block-Hankel matrix is $s = 15$, and the number of block columns is $q = 984$. Figure 2.3 shows the power spectra of the identified spectral factors of order $n = 2$ (dashed line) and $n = 6$ (solid line), obtained using parametric optimization over all parameters. The identified second order model provides a reasonable approximation of the given spectral samples. With a sixth order model it is already possible to obtain an almost perfect
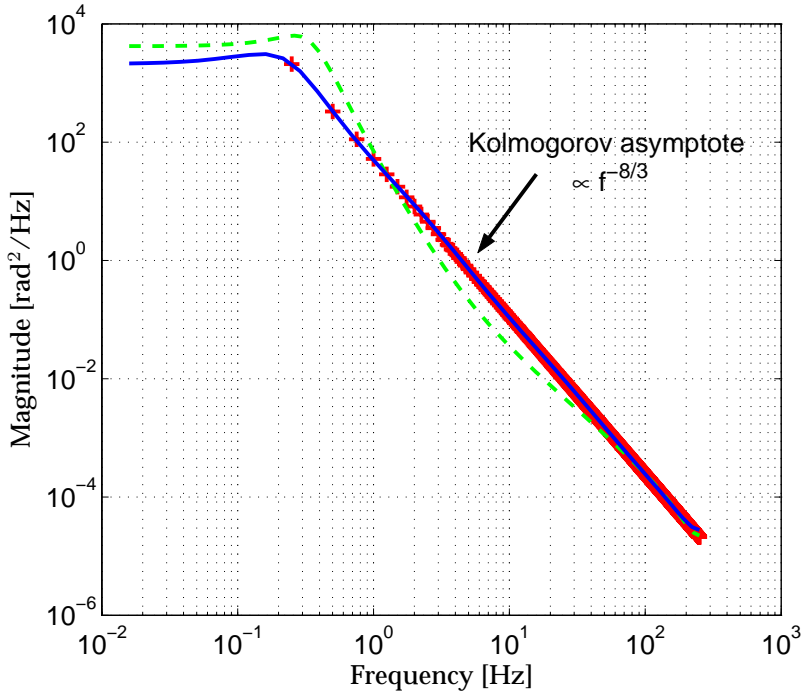
Figure 2.3: Samples of the non-rational Kolmogorov phase spectrum, used as the input Φ_k to the power spectral factor approximation algorithm (crosses). Power spectrum of identified spectral factor of order $n = 2$ (dashed line). Power spectrum of identified spectral factor of order $n = 6$ (solid line).

fit to the given spectral samples, over the entire frequency range. The performance of the identified spectral factor for different model orders, both before and after parametric optimization, has be included in Table 2.2. Apart from the computed performance measure P_{tr}^w, the table includes the total computation time on a PC with AMD Athlon™ 64 X2 Dual Core 4800+ Processor and 4Gb of internal memory. Furthermore, the table indicates whether or not Schur re-stabilization has been used to guarantee stability of the identified minimum-phase spectral factor.

From Table 2.2 it is clear that the subspace identification algorithm generally provides already a reasonable good initial estimate of the minimum-phase spectral factor. Especially for low model orders the additional gain achieved by parametric optimization is rather modest. The parametric optimization step seems to be particularly useful in the case that Schur re-stabilization has been applied to enforce stability. This is to be expected, since the Schur re-stabilization procedure is formulated as a matrix nearness which is concerned with the problem of finding a stable A matrix that is as close as possible to the identified matrix, with respect to some weighted norm. There is no guarantee that this procedure is optimal with respect to the cost function considered in the spectral factor approximation

algorithm. The table shows that Schur re-stabilization is only necessary for model orders larger than $n = 6$. This is in accordance with the general observation that the chance of identifying an unstable system in subspace identification generally increases with the selected model order. Finally, note that the performance after parametric optimization increases monotonically with the model order. Also this is to be expected since the increased model order gives rise to additional degrees of freedom, which can be exploited to reduce the fitting error. The small fitting error, especially when the model order is chosen sufficiently large, confirms that the proposed algorithm is able to accurately approximate the given spectral samples. As a result, the proposed spectral factor approximation algorithm provides a way to obtain a reasonable approximation of the non-rational Kolmogorov power spectrum over a frequency range set by the sampling.

n	P_{tr}^w	Time [s]	P_{tr}^w	Time [s]	Schur
	Initial estimate		*After parametric optimization*		
1	2.0354×10^{-3}	0.75995	2.0354×10^{-3}	1.2048	no
2	1.9815×10^{-3}	0.66981	1.9632×10^{-3}	41.434	no
4	9.3376×10^{-5}	2.5923	4.1842×10^{-5}	74.571	no
5	3.4852×10^{-5}	3.7426	3.4484×10^{-5}	18.942	no
6	3.7943×10^{-5}	7.4214	3.5044×10^{-5}	80.412	no
8	4.3463×10^{-7}	14.925	1.3824×10^{-7}	32.881	yes
10	2.4415×10^{-7}	13.999	9.8324×10^{-9}	30.260	yes
12	4.4671×10^{-7}	15.331	1.9854×10^{-9}	40.976	yes

Table 2.2: Simulation results. Performance of the proposed spectral factor approximation algorithm in modeling a non-rational Kolmogorov phase spectrum.

2.7   Conclusions

We have developed a subspace-based algorithm for estimating a minimum-phase spectral factor that accurately approximates a given discrete-time frequency domain power spectrum. A key issue in solving the stochastic realization problem is to guarantee that the estimated power spectrum satisfies the positive semi-definiteness condition on the unit circle. It has been shown that this issue can be resolved by formulating it as a CLP. The ability to solve the CLP in a global optimal sense, has been used to derive a numerical analogue of the separable least squares principle. Since the procedure inherently attempts to minimize the (weighted) 2-norm of the power spectrum error, the procedure is robust with respect to errors on the given power spectrum.

The performance and robustness have been studied in two simulation examples. In the first simulation example we have studied the case that the given power spectrum is indeed generated by a stochastic system within the class of finite-dimensional LTI systems. The simulation example has also been used to demonstrate the performance robustness with respect to the estimation error on the power spectrum. In the second simulation example we have considered the problem of approximating the non-rational Kolmogorov spectrum over a certain frequency region of interest. It is show that the algorithm is able to provide a good approximation to the selected set of equidistantly distributed spectral samples. As a result, the proposed algorithm can be used to identify a finite dimensional spectral factor that is close to the theoretical power spectra describing the temporal evolution of the optical phase fluctuations due to atmospheric turbulence. Knowledge of this spectral factor may be useful in designing the feedback controller or temporal compensator in an adaptive optics system.

References

Aoki, M., 1990. State Space Modeling of Time Series. Springer.

Ben-Tal, A., Nemirovski, A., 2001. Lectures on modern convex optimization: analysis, algorithms, and engineering applications. SIAM.

Bruls, J., Chou, C. T., Haverkamp, B. R. J., Verhaegen, M., 1999. Linear and nonlinear system identification using separable least-squares. European Journal of Control 5, 116–128.

Conan, J.-M., Rousset, G., Madec, P.-Y., Jul. 1995. Wave-front temporal spectra in high-resolution imaging through turbulence. Journal of the Optical Society of America A 12 (7), 1559–1570.

Dahlén, A., 2001. Identification of stochastic systems: Subspace methods and covariance extension. Ph.D. thesis, Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden.

Fraanje, R., Verhaegen, M., Doelman, N., Berkhoff, A., Aug. 2004. Optimal and robust feedback controller estimation for a vibrating plate. Control Engineering Practice 12 (8), 1017–1027.

Glindemann, A., Hippler, S., Berkefeld, T., Hackenberg, W., Apr. 2000. Adaptive optics on large telescopes. Experimental Astronomy 10 (1), 5–49.

Golub, G. H., Pereyra, V., Apr. 1973. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables are separate. SIAM Journal on Numerical Analysis 10 (2), 413–432.

Hanzon, B., Peeters, R. L. M., 2000. Balanced parameterization of stable SISO all-pass systems in discrete time. Math. Control Signals Systems 13, 240–276.

Hardy, J. W., 1998. Adaptive Optics for Astronomical Telescopes. Oxford series in optical and imaging sciences. Oxford University Press, New York.

Ho, B. L., Kalman, R. E., 1966. Effective construction of linear time state-variable models from input/output functions. Regelungstechnik 14, 545–548.

Johansson, R., Verhaegen, M., Chou, C. T., Robertsson, A., 2001. Residual models and stochastic realization in state-space identification. International Journal of Control 74 (10), 988–995.

Kirkegaard, P. H., Andersen, P., Feb. 1997. State space identification of civil engineering structures from output measurements. In: Fracture and Dynamics. Vol. 3089 of 15th International Modal Analysis Conference. Orlando, Florida, United States, pp. 889–895.

Larimore, W. E., Dec. 1990. Canonical variate analysis in identification, filtering, and adaptive control. In: 29th IEEE Conference on Decision & Control. Honolulu, Hawaii, United States, pp. 569–604.

Lindquist, A., Picci, G., May 1996. Canonical correlation analysis, approximate covariance extension and identification of stationary time series. Automatica 32 (5), 709–733.

Ljung, L., Glad, T., 1994. Modeling of Dynamic Systems. Prentice Hall PRT, Englewood Cliffs.

Mari, J., Stoica, P., McKelvey, T., Jul. 2000. Vector ARMA estimation: A reliable subspace approach. IEEE Transactions on Signal Processing 48 (7), 2092–2104.

Ribarits, T., Deistler, M., Hanzon, B., Aug. 2003. Separable least squares data driven local coordinates. In: 13th IFAC Symposium on Systems Identification. Rotterdam, pp. 1922–1927.

Ruhe, A., Wedin, P. A., Jul. 1980. Algorithms for separable nonlinear least squares problems. SIAM Review 22 (3), 318–337.

Sjöberg, J., Viberg, M., Sep. 1997. Separable non-linear least squares minimization - possible improvements for neural net fitting. In: Proceeding of IEEE Workshop in Neural Networks for Signal Processing. Amelia Island Plantation, Florida, pp. 345–354.

Sturm, J. F., 2002. Implementation of interior point methods for mixed semidefinite and second order cone optimization problems. Optimization Methods and Software 17 (6), 1105–1154.

Tatarskii, V. I., 1971. The Effects of the Turbulent Atmosphere on Wave Propagation. Israel Program for Scientific Translations Ltd. Keter Press, Jarusalem.

Van Overschee, P., De Moor, B., May 1993. Subspace algorithms for the stochastic identification problem. Automatica 29 (3), 649–660.

Van Overschee, P., De Moor, B., 1996. Subspace Identification For Linear Systems, Theory - Implementation - Applications. Kluwer Academic Publishers, Dordrecht.

Van Overschee, P., De Moor, B., Dehandschutter, W., Swevers, J., Dec. 1997. A subspace algorithm for the identification of discrete time frequency domain power spectra. Automatica 33 (12), 2147–2157.

Verhaegen, M., Jan. 1994. Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. Automatica 30 (1), 61–74.

Viberg, M., 1995. Subspace-based methods for the identification of linear time-invariant systems. Automatica 31 (12), 1835–1851.

3 | CHAPTER

Data-Driven $\mathcal{H}_2$-Optimal Control for Adaptive Optics

Adaptive optics is currently used in ground-based astronomical telescopes to improve the resolution by counteracting the effects of atmospheric turbulence. Most AO systems are based on a simple control law that neglects the temporal evolution of the distortions introduced by the atmosphere.

This paper presents a data-driven control design approach that is able to exploit the spatio-temporal correlation in the wavefront, without assuming any form of decoupling between the different wavefront sensor channels and phase reconstruction points. The approach consists of a dedicated subspace-identification algorithm, used to identify an atmospheric disturbance model from open-loop wavefront sensor data, followed by $\mathcal{H}_2$-optimal control design.

It is shown that in the case that the deformable mirror and wavefront sensor dynamics can be represented by a delay and an impulse response of two taps, it is possible to derive an analytical expression for the $\mathcal{H}_2$-optimal controller. Together with the identification algorithm, this provides a non-iterative way to go from open-loop measurement data to closed-loop controller design. Numerical simulation experiments demonstrate a performance improvement with respect to the common AO control approach.

Keywords: Adaptive optics; Data-driven disturbance modeling; Stochastic identification; Optimal control

3.1   Introduction

Adaptive optics is a technique to actively sense, estimate and correct the wavefront distortions that are introduced in a light beam as it propagates through a turbulent medium. Nowadays, adaptive optics has found widespread application in ground-based astronomical imaging, where it is used to compensate the wavefront distortions introduced by the turbulent atmosphere. By using measurements from a wavefront sensor (WFS), AO tries to cancel out most of the wavefront distortions by varying the optical path difference with an active optical element, such as a deformable mirror (DM). Whereas atmospheric turbulence otherwise severely limits the angular resolution, AO may now help to improve the image quality and enable the recording of long-exposure images with resolutions close to the diffraction limit. For an extensive overview of AO, the reader is referred to Hardy (1998); Tyson (1998); Roddier (1999), and the references therein.

   This paper focuses on the control aspects of AO. An important complication in the AO control problem is that the WFS typically provides some measure of the slope (or curvature) of the residual wavefront, while the performance is often evaluated in terms of the mean-square wavefront. The common way to deal with the discrepancy between measurement and control objective is to include a separate wavefront reconstruction step. Given the reconstructed wavefront, the problem of imposing the proper shape on the DM is seen as a servo control problem. As a result, the majority of the AO systems are based on a control law that consists of a separate wavefront reconstruction step, the projection of the estimated wavefront on the DM actuator space and a dynamic servo compensator responsible for stability and closed-loop performance. The wavefront reconstruction and DM fitting problem are usually solved in a static setting (Roddier 1999; van Dam et al. 2004). In the simplest case the wavefront reconstructor and the DM fitting matrix are obtained as the pseudo-inverse of the phase-to-slope mapping and the DM influence matrix. Both maximum likelihood and maximum *a posteriori* techniques are used to improve the accuracy of the wavefront estimate by incorporating prior knowledge on the second order statistics of the spatial distribution of the wavefront (Roddier 1999; Law and Lane 1996). The modified statistics due to closed-loop operation, however, are often neglected. In designing the dynamic servo controller it is typically assumed that the control loop can be decoupled in a series of independent single-input single-output (SISO) feedback loops with a predefined control structure having its roots in classical control theory. Common servo controller structures include the leaky integrator, the proportional-integral (PI) controller and the Smith predictor. The choice of the control parameters is a trade off between disturbance rejection, noise propagation and closed-loop stability. It can be easily shown that the higher the control bandwidth, the better the disturbance rejection but also the higher the noise propagation and risk of closed-loop instabilities. The optimal bandwidth depends both on the signal-to-noise ratio (SNR) and the atmospheric turbulence conditions. In the modal control optimization approach (Ellerbroek et al. 1994; Gendron and Léna 1994), the wavefront is decomposed in a set of modal basis functions and the servo gain is optimized on a mode to mode basis. Since the SNR decreases with the spatial frequency, modes with a high spatial frequency should have a smaller bandwidth. Stability criteria

for first-order closed-loop AO servo systems have been analyzed in Wild (1998).

Even though the above described common AO control strategy does already a remarkable job under favorable conditions, there is still need for performance improvement. As pointed out by Roddier (1998), especially the compensation efficiency of large AO systems is unduly low. Rather than focusing only on the development of AO systems with more sensors and actuators, it may therefore be rewarding to search for ways of improving the performance of current AO systems. It is clear that the common AO control approach does not explicitly account for the temporal evolution of the wavefront disturbance and the dynamics of the AO system components. On the other hand, the finite time delay between measurement and correction is known to be one of the major limitations on the AO performance (Dessenne et al. 1997; van Dam et al. 2004). A promising way to reduce the effect of the temporal error is to exploit the temporal correlation in the wavefront to anticipate future wavefront distortions. This has for instance motivated the development of the predictive optimal estimator (Wild 1996), which minimizes the time delay error in an ensemble averaged, sense rather than actually predicting future wavefronts. In the spirit of modal optimization, a modal linear predictive controller, whose parameters are optimized by recursive least-squares, has been introduced (Dessenne et al. 1997). Since the latter approach assumes decoupling of the different modes however, it is still not able to take advantage of the spatio-temporal correlation in the wavefront.

This paper presents a data-driven control design approach that take full advantage of the spatio-temporal correlation. In contrast to the above approaches, it does not assume any form of decoupling between the spatial and temporal dynamics. The proposed control design approach consists of two major steps. In the first step a dedicated subspace identification algorithm is used to identify a multivariable atmospheric disturbance model on the basis of open-loop WFS data. In the second step the identified atmospheric disturbance model is used to compute the optimal controller by formulating the AO control problem in an $\mathcal{H}_2$-optimal control framework. Apart from the fact that the $\mathcal{H}_2$-optimal control provides a more elegant framework of dealing with the discrepancy between measurement and control objective, it is closely related to Linear Quadratic Gaussian (LQG) control design. The LQG framework has been used by Paschall and Anderson (1993) to design an AO controller under the restrictive assumption that the atmospheric wavefront distortions can be described in terms of the first 14 Zernike modes with coefficients generated by independent first-order Markov processes. Looze et al. (1999, 2003) have used LQG to design a diagonal modal controller based on an atmospheric disturbance model in which each individual mode is described by an autoregressive moving average (ARMA) model identified from open-loop wavefront data. The LQG control approach has proved to be suitable for both classical and multi-conjugated AO systems Le Roux et al. (2004). Each of these LQG approaches use a rather restrictive atmospheric disturbance model in the sense that modal decoupling is assumed. By assuming model decoupling, it is not possible to make optimal use of spatio-temporal correlation that is imposed by the Taylor hypothesis (Hardy 1998; Roddier 1999). From the Taylor hypothesis, which states that the atmospheric turbulence evolves at a time scale much longer than

the time it takes for the inhomogeneities to cross the line of sight, it is clear that the upper wind WFS channels provide direct information on the future development of the turbulence elsewhere in the aperture. Gavel and Wiberg (2003), recently proposed an optimal control approach which is explicitly based on the Taylor hypothesis. Even though this is very elegant from a theoretical point of view, the Taylor hypothesis may only be partially satisfied in practice. The proposed control approach uses a model structure that is sufficiently general to exploit the spatio-temporal correlation imposed by the Taylor hypothesis, but does not depend on it. Furthermore, the data-driven modeling approach has the advantage that it yields a good match with the prevalent turbulence conditions and it does not require accurate estimates of physical parameters like the wind speed and direction of the frozen layers.

Apart from introducing a more general model structure for the disturbance model, this paper shows that the special structure of the AO control problem can be exploited in computing the $\mathcal{H}_2$-optimal controller. In general, computing the $\mathcal{H}_2$-optimal controller requires the numerical solution of two Riccati equations. Due to the special structure of the identified disturbance model, in the worst case only one Riccati equation needs to be solved. Furthermore it will be shown that an analytical expression for the closed-loop optimal controller can be derived when the deformable mirror dynamics can be represented as a delay and a finite impulse response (FIR) model of two taps. This is a realistic assumption when the characteristic time of the deformable mirror is small compared to wavefront sensor integration or exposure time.

The remainder of this paper is organized as follows. Section 3.2 provides an accurate description of the AO control problem and introduces most of the necessary notation. The subspace identification algorithm used to obtain a control-relevant atmospheric disturbance model from open-loop wavefront sensor data will be considered in more detail in Section 3.3. The algorithm is capable to identify an accurate atmospheric disturbance model for small to medium sized AO systems. This will be illustrated in Section 3.4 by means of two simulation examples on the basis of open-loop WFS data obtained from an AO laboratory setup and real-life data from the William Herschel telescope. Given the identified disturbance model, Section 3.5 presents a general strategy to determine the controller that minimizes the mean-square residual phase error. The AO control problem will be formulated in the $\mathcal{H}_2$-optimal control framework for which the standard solution is known. The central topic in Section 3.6 is how to exploit the special structure of the AO problem in computing the $\mathcal{H}_2$-optimal controller. Section 3.7 presents a validation study in which the performance of the proposed control design strategy is compared with the common AO control law. Also these simulations are performed on the basis of open-loop wavefront sensor data obtained from an AO test bench. Finally, the paper concludes with a short discussion in Section 3.8.

3.2   The adaptive optics control problem

Figure 3.1 provides a schematic representation of the functional relation between the main components of a classical AO system. Light with an atmospherically distorted phase profile $\phi(\cdot)$ enters the system and is reflected from a deformable mirror (DM), which introduces a phase correction $\phi_m(\cdot)$. Part of the compensated light, with residual phase error $\epsilon = \phi - \phi_m$, is directed to a wavefront sensor (WFS). The WFS signal $s(\cdot)$ forms the input to the controller, which is responsible for determining the actuator commands $u(\cdot)$. The measurement noise is represented by an additive noise term $\eta(\cdot)$.

   A common objective in AO is to maximize the Strehl ratio, which is defined as the on-axis intensity of a point source relative to that of the diffraction limit. Through the Maréchal approximation (Born and Wolf 1999), this is equivalent to minimizing the mean-square residual phase error. The AO control problem can hence be defined as the problem of finding the closed-loop controller that minimizes the mean-square residual phase error. In this paper it will be assumed
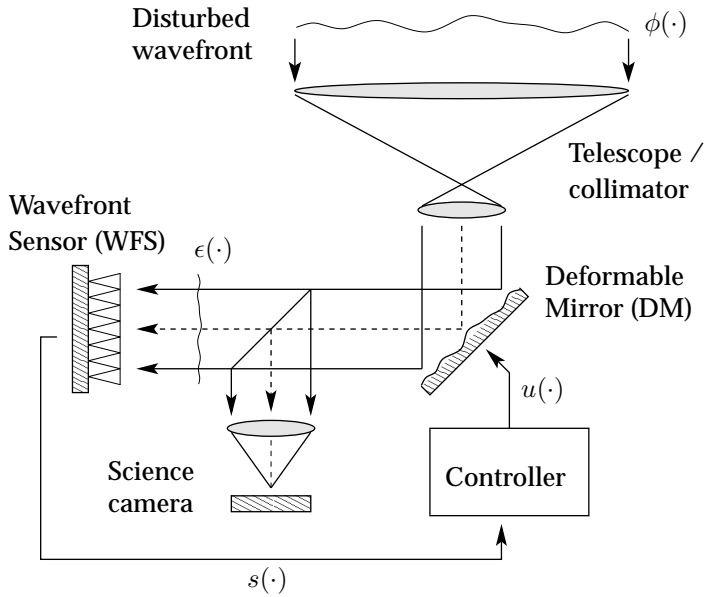
Figure 3.1: Functional relationship between main components of an AO system.

that the phase distortion profile over the telescope aperture can be represented by a finite-dimensional discrete-time vector signal. At each time instant $k \in \mathbb{N}$, the uncorrected wavefront $\phi(\cdot)$, the phase correction introduced by the DM $\phi_m(\cdot)$ and the residual wavefront error $\epsilon(\cdot)$ are described by the vectors $\phi(k) \in \mathbb{R}^{m_\phi}$, $\phi_m(k) \in \mathbb{R}^{m_\phi}$ and $\epsilon(k) \in \mathbb{R}^{m_\phi}$. Whether the vector signal $\phi(k)$ provides a zonal or modal description of the wavefront is irrelevant as long as its mean-square error provides a good approximation of the mean-square error of the unsampled

wavefront over the telescope aperture.

An important issue in the AO control problem is that the wavefront distortion $\phi(k)$ cannot be measured directly. Only the WFS slope measurements $s(k) \in \mathbb{R}^{m_s}$ are available for identification and control. It is generally not possible to reconstruct the entire wavefront from the measurement signal. In order to arrive at a well-posed control and identification problem it is important to exclude the unobservable part of the wavefront from the problem formulation. This is achieved by introducing a signal of lower dimension that describes only the observable part of the wavefront $\phi(k)$. To this end, let the wavefront $\phi(k)$ and the open-loop WFS slope measurements $s(k)$ be related as

$$s(k) = \mathcal{G}(z)\phi(k) + \eta(k), \tag{3.1}$$

where $\mathcal{G}(z) = g(z)G$ is the cascade of a scalar stable linear time invariant (LTI) system $g(z)$, which accounts for the WFS dynamics, and the phase-to-slope geometry matrix $G \in \mathbb{R}^{m_s \times m_\phi}$, which describes the optical transformation from phase to slopes. Here, the assumption of scalar dynamics is not very restrictive since the WFS dynamics are typically determined by the finite CCD exposure time, the readout delays and the data processing delays, which are the same for each of the channels. From equation (3.1) it is clear that only the part of the wavefront that is in the row space of G can be reconstructed from the measurements. The number of unobservable modes is equal to the rank deficiency of the matrix G. A reduced representation of the observable part of the wavefront can be obtained by considering the singular value decomposition (SVD)

$$G = U\Sigma V^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix},$$

where $U \in \mathbb{R}^{m_s \times m_s}$ and $V \in \mathbb{R}^{m_\phi \times m_\phi}$ are orthonormal matrices and the partitioning of $\Sigma \in \mathbb{R}^{m_s \times m_\phi}$ is such that $\Sigma_1 \in \mathbb{R}^{m_y \times m_y}$ contains all nonzero singular values. Substituting the SVD in (3.1) and pre-multiplying both sides with U_1^T, gives rise to the following reduced WFS model

$$y(k) = \Sigma_1 \varphi(k) + \nu(k), \tag{3.2}$$

where $y(k) \doteq U_1^T s(k) \in \mathbb{R}^{m_y}$ and $\varphi(k) \doteq g(z)V_1^T\phi(k) \in \mathbb{R}^{m_y}$. The signal $\varphi(k)$ can be interpreted as a filtered reduced representation of the observable part of the wavefront $\phi(k)$. This can be easily seen by noting that due to the orthogonality of V the wavefront can be decomposed as $\phi(k) = V_1 V_1^T \phi(k) + V_2 V_2^T \phi(k)$. Whereas the first term has a direct influence on the measured output $s(k)$, the second term cannot be observed as it lies in the null-space of G. Furthermore, it is clear that the signal $g(z)\phi(k)$ and $\varphi(k)$ have the same 2-norm. The signal $y(k)$ can be regarded as a representation of the informative part of $s(k)$. Here it is interesting to note that projection U_1^T removes the measuring noise that is not in the range space of G and cannot be caused by the wavefront. Since G is generally tall, the proposed projection results in a considerable reduction of the dimension of the signal that has to be modeled. In accordance with above definitions, the filtered reduced representation of the DM wavefront correction and residual phase error are defined

as $\varphi_m \doteq g(z)V_1^T\phi_m$ and $\varepsilon(k) \doteq \varphi(k) - \varphi_m(k)$. Linearity of the WFS implies that the measured output corresponding to the reduced residual wavefront $\varepsilon(k)$ can be expressed as $r(k) \doteq y(k) - y_m(k)$, where $y(k)$ and $y_m(k)$ denote the contributions due to atmospheric turbulence and the DM.

It will be assumed that the relation between the actuator inputs $u(k) \in \mathbb{R}^{m_u}$ and the DM wavefront correction $\varphi_m(k) \in \mathbb{R}^{m_y}$, can be described by a stable LTI system with state-space realization

$$\mathcal{H}(z) = C_m(zI - A_m)^{-1}B_m. \tag{3.3}$$

The absence of a direct feed-through term in the DM model is not very restrictive as there is always at least one sample delay between measurement and correction. Without loss of generality this delay can be included in the DM model. Furthermore, it should be noted that by the definition of the signal $\varphi(k)$ also the WFS dynamics $g(z)$ are implicitly included in the DM model. Because of this, the reduced WFS model (3.2) is now free of any dynamics. The reason for this particular choice of the reduced signals $\varphi(k)$, $\varphi_m(k)$ and $\varepsilon(k)$ will become more clear when defining the cost function in (3.5).

Optimizing the AO system performance requires accurate knowledge of the statistical properties of the uncorrected wavefront and the corresponding WFS signal. In this paper it is assumed that the observable part of the uncorrected wavefront $\varphi(k)$, and hence also the reduced WFS signal $y(k)$, can be modeled as a wide sense stationary (WSS) regular stochastic process. This is a reasonable assumption since the statistical properties of the wavefront change on a time scale that is long compared to the time scale of the fluctuations themselves. To guarantee a good performance over longer time periods however, it is necessary to update the disturbance model on a regular basis. Under the assumption of wide sense stationarity, the second order statistics of the reduced WFS signal $y(k)$ can be described as the output of an LTI system with a zero-mean white noise input $v(k) \in \mathbb{R}^{m_y}$ and covariance matrix $R_v \doteq \mathcal{E}\{v(k)v^T(k)\}$. This in combination with the reduced WFS model (3.2) motivates the choice of the following model structure

$$\mathcal{S} : \begin{cases} x(k+1) = & A_d x(k) + K_d v(k) \\ y(k) = \Sigma_1 C_d x(k) + & v(k) \\ \varphi(k) = & C_d x(k) + \zeta(k), \end{cases} \tag{3.4}$$

where $A_d - K_d\Sigma_1 C_d \in \mathbb{R}^{n_d \times n_d}$ and $A_d \in \mathbb{R}^{n_d \times n_d}$ are assumed to be stable, and $\zeta(k) = \Sigma_1^{-1}(v(k) - \nu(k))$ is again a zero-mean white noise sequence with covariance matrix $R_\zeta \doteq \mathcal{E}\{\zeta(k)\zeta(k)^T\}$. The cross-covariance between the $v(k)$ and $\zeta(k)$ will be denoted by $R_{v\zeta} \doteq \mathcal{E}\{v(k)\zeta(k)^T\}$. The above description of the atmospheric turbulence is assumed to be minimal in the sense that output signals cannot be described by a model of order less than n_d. An important characteristic is that the model is in innovation form with respect to the WFS output $y(k)$. It provides a minimum-phase spectral factor of the stochastic process $y(k)$, which will appear to be a very useful property in determining the optimal controller.

The problem considered in this paper divided into two sub-problems. Given a batch of $N_i \in \mathbb{N}$ open-loop WFS observations $y(k)$, the first sub-problem is

to determine an estimate of the system matrices $A_d \in \mathbb{R}^{n_d \times n_d}$, $K_d \in \mathbb{R}^{n_d \times m_y}$ and $\Sigma_1 C_d \in \mathbb{R}^{m_y \times n_d}$ such that the estimated atmospheric disturbance model $\mathcal{S}$ matches the second order statistics of the process $y(k)$. This part of the problem is referred to as the stochastic identification problem. Starting from the identified disturbance model $\mathcal{S}$, the second sub-problem is to find the optimal controller $\mathcal{C}(z)$ that minimizes the cost function

$$J = \mathcal{E}\left\{\varepsilon^T(k)\varepsilon(k)\right\} + \mathcal{E}\left\{u^T(k)Qu(k)\right\}, \tag{3.5}$$

where $Q \geq 0$ is a positive regularization matrix which makes a trade-off between the expected mean-square reduced residual wavefront error $\mathcal{E}(\varepsilon(k)\varepsilon^T(k))$ and the expected amount of control effort $\mathcal{E}(u(k)u^T(k))$. By increasing the control effort weighting Q it is possible to reduce the amount of energy dissipated by the DM and make the controller more robust to model uncertainties. The matrix Q will be typically chosen diagonal, allowing for a penalty on the control effort on each of the actuators separately. Furthermore it is important to recall that signal $\varepsilon(k)$ has been defined in such a way that it incorporates the WFS dynamics $g(z)$. Minimizing the first term of (3.5) is therefore equivalent to minimizing the mean-square error of the observable part of the filtered signal $g(z)\epsilon(k)$. Even though it is possible to explicitly account for the WFS dynamics in the definition of cost function, this is usually not sensible because the WFS usually has low-pass characteristics. This implies that the WFS dynamics mainly distorts the high frequency region of ε, while the turbulence is dominant at low-frequencies. Moreover by inverting the WFS dynamics one risks the chance of high-frequency noise amplification.

3.3   Data-driven disturbance modeling

In this section we present a dedicated subspace identification algorithm that is able to deal with the stochastic identification problem considered in previous section. Based on open-loop WFS data, the algorithm provides a full spatio-temporal atmospheric disturbance model $\mathcal{S}$, without assuming any form of decoupling between the channels. A consequence of this rather extensive description is that even for relatively small AO systems a huge identification problem has to be solved. Computational efficiency is therefore an important issue. The proposed algorithm is able to identify a full atmospheric disturbance model for AO systems with up to a few hundred degrees of freedom on a PC with a $3\,\mathrm{GHz}$ Intel Pentium IV processor and 512 Mb of internal memory, within the period of only a few minutes time.

Based on numerically robust matrix operations, subspace algorithms bypass the need for model parametrization and nonlinear optimization. This is an important advantage over the more traditional maximum likelihood and prediction error methods (Ljung and Glad 1994; Ljung 1999), which rely on the optimization of a suitably chosen cost function of structural model parameters. Apart from the increased risk of ending up in a local minimum, the computational complexity of these algorithm grows rapidly with the number of independent parameters. This becomes especially a problem in the multivariate case where the mapping from

each input to each output is often parametrized independently. With subspace identification the multivariate case can be handled within the same framework.

The subspace identification algorithm presented in this paper is basically an efficient output only implementation of the SSARX algorithm proposed by Jansson (2003). In analogy with the name SSARX, the new algorithm will be called SSAR. An important advantage of the proposed algorithm is that it provides a direct estimate of the state-space matrices of the minimum-phase spectral factor. Most subspace identification algorithms for stochastic identification, are based on a two-step procedure where the minimum-phase spectral factor is obtained after the factorization of some intermediate estimate. Apart from the fact that such a two step procedure is computationally more complex, it is important to note that, depending on the type of intermediate estimates, the spectral factorization problem may fail to have a solution. The class of subspace algorithms which first estimate a rational covariance model (Van Overschee and De Moor 1993), for instance, suffers from this problem. As demonstrated by Lindquist and Picci (1996), the spectral factorization problem has only a solution when the rational covariance model is positive real. A possible workaround when the positive realness condition has failed is to modify the system matrices of the covariance estimate in a spirit of a matrix nearness problem (Mari et al. 2000). Even though the resulting covariance matching problem can be efficiently formulated in terms of a semi-definite program, this approach is computationally not feasible for the size of systems considered in this paper. The SSAR algorithm avoids the need for spectral factorization by directly estimating the system matrices of the Kalman predictor model corresponding to the minimum-phase spectral factor. The minimum-phase requirement is translated to a stability requirement on the system matrices of the Kalman filter and the minimum-phase spectral factor. This requirement can be easily checked and if necessary stability can be enforced by using the Schur procedure in (Mari et al. 2000).

The computational demands of the proposed subspace algorithm have been reduced by using an efficient implementation. By applying a different weighting in estimating the state sequence, a single RQ factorization of a stacked block-Hankel matrix of past and future data can be used both for computing the required AR coefficients and for data compression. This leads to an efficient implementation both in terms of the number of flops and the memory requirements. Before providing a detailed description of the new SSAR implementation, the SSARX algorithm as presented by (Jansson 2003) will be briefly reviewed. Since the interest is in stochastic identification, also the SSARX algorithm will be considered in an output only setting. Reviewing the SSARX algorithm will prove to be valuable both for explaining the proposed algorithm and in outlining the difference with a straightforward output only implementation of SSARX.

The SSARX algorithm is based on an alternative representation of the stochastic process $y(k) \in \mathbb{R}^{m_y}$. Consider the stochastic disturbance model introduced in equation (3.4). By using the output equation to eliminate the noise input $v(k) \in \mathbb{R}^{m_y}$ from the state-update equation, the stochastic process $y(k)$ can be represented

in the alternative form

$$\widetilde{\mathcal{S}} : \begin{cases} x(k+1) = \widetilde{A}x(k) + \widetilde{K}y(k) \\ \quad y(k) = \widetilde{C}x(k) + v(k), \end{cases} \tag{3.6}$$

where the matrices $\widetilde{A}$ and $\widetilde{C}$ are defined as $\widetilde{A} \doteq A_d - K_d\widetilde{C}$, $\widetilde{C} \doteq \Sigma_1 C_d$ and where $\widetilde{K}$ is defined as $\widetilde{K} = K_d$ to achieve conformity in notation. The above representation of $y(k)$ can be seen as the Kalman predictor model corresponding to (3.4). Furthermore, it is clear that the minimum-phase requirement on the stochastic disturbance model is equivalent to demanding that $\widetilde{A}$ is asymptotically stable. Let the vectors of stacked past and future outputs and the vector of future innovations be defined as

$$\begin{aligned} y_p(k-p) &\doteq \begin{bmatrix} y^T(k-p) & \dots & y^T(k-2) & y^T(k-1) \end{bmatrix}^T, \\ y_f(k) &\doteq \begin{bmatrix} y^T(k) & y^T(k+1) & \dots & y^T(k+f-1) \end{bmatrix}^T, \\ v_f(k) &\doteq \begin{bmatrix} v^T(k) & v^T(k+1) & \dots & v^T(k+f-1) \end{bmatrix}^T, \end{aligned}$$

with $p, f \in \mathbb{N}$ some user defined parameters, whose selection will be discussed in more detail at the end of the section. Then, by iteratively applying the state-update equation in (3.6), it can be shown (Jansson 2003; Peternell et al. 1996) that for a stable $\widetilde{A}$ and p sufficiently large, the state $x(k)$ can be approximated as a linear combination of the past outputs

$$x(k) \approx \mathcal{K}_p y_p(k-p), \tag{3.7}$$

where the matrix $\mathcal{K}_p \in \mathbb{R}^{n_d \times pm_y}$ is defined as

$$\mathcal{K}_p \doteq \begin{bmatrix} \widetilde{A}^{p-1}\widetilde{K} & \dots & \widetilde{A}\widetilde{K} & \widetilde{K} \end{bmatrix}. \tag{3.8}$$

Here it should be noted that the ordering of the output data in the vector of past observations $y_p(k-p)$ differs form the one used in Jansson (2003). While in the above definition the observations are ordered forward in time, the vector of past observations in the original presentation of SSARX is ordered backward in time. This is by no means a fundamental difference. The only effect of reordering the observations in $y_p(k-p)$ is that also the block columns of $\mathcal{K}_p$ are ordered in a reverse way. The considered ordering of $y_p(k-p)$ will appear useful in achieving an efficient implementation of the proposed algorithm. Furthermore, by using (3.6) it is possible to establish the following relation between past and future outputs

$$y_f(k) = \mathcal{O}_f x(k) + \mathcal{T}_f y_f(k) + v_f(k), \tag{3.9}$$

where the extended observability matrix $\mathcal{O}_f$ and Toeplitz matrix of Markov parameters $\mathcal{T}_f$ are defined as

$$\mathcal{O}_f \doteq \begin{bmatrix} \widetilde{C} \\ \widetilde{C}\widetilde{A} \\ \vdots \\ \widetilde{C}\widetilde{A}^{f-1} \end{bmatrix}, \quad \mathcal{T}_f \doteq \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \widetilde{C}\widetilde{K} & \ddots & & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \widetilde{C}\widetilde{A}^{f-2}\widetilde{K} & \cdots & \widetilde{C}\widetilde{K} & 0 \end{bmatrix}. \tag{3.10}$$

Consider the state estimate in (3.7) and the relation between past and future outputs (3.9), then the SSARX algorithm consists of the following steps. In the first step a high order autoregressive model is estimated from the data to get an unstructured estimate of the Markov parameters $\widetilde{C}\widetilde{A}^i\widetilde{K}$ for $i \in \{0, \ldots, f-2\}$. The estimated Markov parameters are then used to construct an estimate $\widehat{\mathcal{T}}_f$ of the Toeplitz matrix $\mathcal{T}_f$. In the next step the constructed estimate $\widehat{\mathcal{T}}_f$ is used to define the signal $z(k)$ as $z(k) \doteq (I - \widehat{\mathcal{T}}_f)y_f(k)$. Together with the definition of $\mathcal{B} \doteq \mathcal{O}_f\mathcal{K}_p$, the approximation of $z(k)$ obtained by substituting (3.7) in (3.9) can be expressed as

$$z(k) \approx \mathcal{B}\, y_p(k-p) + v_f(k). \tag{3.11}$$

This equation can be viewed as a low rank linear regression problem in $\mathcal{B}$ and is used to obtain an estimate of $\mathcal{K}_p$. The least squares estimate of $\mathcal{B}$, in the sense that it minimizes the conditional expectation of the mean-square error between $z(t)$ and $\mathcal{B}y_p(k-p)$, is given by $\Gamma_{zp}(\Gamma_{pp})^{-1}$, where Γ_{zp} is the cross-correlation matrix between $z(k)$ and $y_p(k-p)$ and Γ_{pp} is the correlation matrix of $y_p(k-p)$. Given only a finite data set, the correlation matrices are approximated by their finite sample estimates $\widehat{\Gamma}_{zp}$ and $\widehat{\Gamma}_{pp}$, with $\widehat{\Gamma}_{zp}$ defined as

$$\widehat{\Gamma}_{zp} = \frac{1}{N} \sum_{k=0}^{N-1} z(k)y_p^T(k-p), \tag{3.12}$$

and $\widehat{\Gamma}_{pp}$ defined accordingly. This gives rise to the following approximation $\widehat{\mathcal{B}} = \widehat{\Gamma}_{zp}\widehat{\Gamma}_{pp}^{-1}$ of $\mathcal{B}$. An estimate of the row space of the matrix $\mathcal{K}_p$ can now be obtained by factorizing $\widehat{\mathcal{B}}$. This is achieved by computing the SVD of

$$M \doteq W_1\widehat{\mathcal{B}}W_2 = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^T, \tag{3.13}$$

where $\widetilde{U}$ and $\widetilde{V}^T$ are orthonormal matrices, $\widetilde{\Sigma}$ is a diagonal matrix composed of the singular values arranged in non-decreasing order, and W_1 and W_2 are non-singular weighting matrices. The estimate of $\mathcal{K}_p$ obtained from the factorization, up to within a similarity transformation T, is then given by $\widehat{\mathcal{K}}_p = V_{n_d}^T W_2^{-1}$, where $V_{n_d}^T$ is the matrix composed of the first n_d columns of $\widetilde{V}^T$. In the SSARX algorithm the row-space of $\mathcal{K}_p$ is estimated by performing a canonical correlation analysis (CCA) on the signals $z(k)$ and $y_f(k-p)$. This is equivalent to choosing the weighting matrices in the SVD as $W_1 = \widehat{\Gamma}_{zz}^{-1/2}$ and $W_2 = \widehat{\Gamma}_{pp}^{1/2}$. These weighting matrices have the nice statistical property that they lead to the maximum likelihood estimate for a Gaussian linear regression problem with a rank constraint on the coefficient matrix (Peternell et al. 1996).

The next step of the SSARX algorithm consists of substituting the estimate of $\mathcal{K}_p$ in equation (3.7) to obtain an estimate of the corresponding state sequence $x(k)$. By replacing the true state with the estimated state $\hat{x}(k) = V_{n_d}^T W_2^{-1} y_p(k-p)$, the system matrices can now be estimated by finite linear regression in the state-space equations (3.6). Here the term finite linear regression refers to the linear regression problem obtained by replacing the conditional expectation by finite sample estimates, as was done to obtain $\widehat{\mathcal{B}}$ from (3.11). In this way, an estimate of

the system matrices $\tilde{A}$ and $\tilde{K}$ is obtained by regressing $\hat{x}(k+1)$ on $\hat{x}(k)$ and $y(k)$. Likewise, the system matrix $\widetilde{C}$ is estimated by regressing $y(k)$ on $\hat{x}(k)$.

To arrive at an efficient implementation of the algorithm SSAR proposed in this paper, the data equations are expressed in terms of block Hankel matrices. Given N samples of the stacked data vector of past or future outputs $y_q(k)$, $q \in \{p, f\}$, the data block Hankel matrix $Y_{i,q,N}$ is defined as

$$Y_{i,q,N} \doteq \begin{bmatrix} y_q(i) & y_q(i+1) & \cdots & y_q(i+N-1) \end{bmatrix} \in \mathbb{R}^{qm_y \times N},$$

where the first entry of the subscript of $Y_{i,q,N}$ refers to the time index of its top left entry, the second refers to the number of block rows and the third refers to the number of columns. Using the same notational convention, the block Hankel matrix constructed from the vector of future innovations $v_f(k)$ will be denoted by $V_{i,f,N}$. By stacking time-shifted versions of (3.9), the equivalent data equation in terms of block Hankel matrices can be expressed as

$$(I - \widehat{\mathcal{T}}_f)Y_{p,f,N} \approx \mathcal{B}Y_{0,p,N} + V_{0,f,N}. \tag{3.14}$$

Recall that the first step in the SSARX algorithm is to identify a high order AR model to get an unstructured estimate of the Markov parameters $\widetilde{C}\widetilde{A}^i\widetilde{K}$ $i \in \{0, \ldots, f-1\}$. An important observation is therefore that equation (3.14) is nothing but the stacked outputs of an AR model. This implies that if the AR model order is not too high, the problem of identifying the Markov parameters can be conveniently expressed in terms of the block Hankel matrices of the above equation. By selecting the right block rows, the block Hankel matrices in (3.14) contain sufficient information for identifying an AR model of order $p + f - 1$. However, since exploratory experiments show that for the considered type of data the choice of the AR model order has no or little influence on the overall performance, it will be chosen as small as possible for reasons of efficiency. The order of the AR model will therefore be chosen $f - 1$, which is the minimal amount needed to construct the estimate $\widehat{\mathcal{T}}_f$. The problem of identifying such an AR model can now be formulated as the following least squares optimization problem (Ljung and Glad 1994; Ljung 1999)

$$\widehat{\widetilde{C}\mathcal{K}}_{f-1} = \arg \min_{\widetilde{C}\mathcal{K}_{f-1}} \|Y_{p+f-1,1,N} - (\widetilde{C}\mathcal{K}_{f-1})Y_{p,f-1,N}\|_F^2, \tag{3.15}$$

where $\| \cdot \|_F$ denotes the Frobenius norm and where, in accordance with definition (3.8), the matrix $\widetilde{C}\mathcal{K}_{f-1}$ is composed of the first $f-1$ Markov parameters. It is well known that the above optimization problem has the following solution

$$\widehat{\widetilde{C}\mathcal{K}}_{f-1} = Y_{p+f-1,1,N}(Y_{p,f-1,N})^\dagger, \tag{3.16}$$

where $(\cdot)^\dagger$ denotes the pseudo-inverse. Since the matrices $Y_{p+f-1,1,N}$ and $Y_{p,f-1,N}$ correspond to respectively the last and first $f-1$ block rows of $Y_{p,f,N}$, the solution to the problem of identifying the Markov parameters necessary for constructing the estimate of $\mathcal{T}_f$, can be completely expressed in terms of the data block Hankel matrices in (3.14).

Also the problem of determining an estimate for the extended controllability $\mathcal{K}_p$ can be conveniently formulated in terms of the data block Hankel matrices. By comparing the definitions of the block Hankel matrix and the finite sample estimate of the correlation matrices Γ_{zp} and Γ_{pp} in (3.12), it is clear that the finite sample estimates can be expressed as

$$C\widehat{\Gamma}_{zp} = \frac{1}{N}(I - \widehat{\mathcal{T}}_f)Y_{p,f,N}Y_{0,p,N}^T$$

$$\widehat{\Gamma}_{pp} = \frac{1}{N}Y_{0,p,N}Y_{0,p,N}^T.$$

With the correlation matrix $\widehat{\Gamma}_{zz}$ expressed in a similar way, the above equations could in principle be used to compute the matrix $M \doteq W_1\widehat{\mathcal{B}}W_2$ for the canonical correlation weights W_1 and W_2 as defined before. Instead of proceeding in this way, the proposed SSAR algorithm uses a different set of weights. More specifically, the weighting matrices are chosen as $W_1 = I$ and $W_2 = (\frac{1}{N}\Gamma_{pp})^{1/2}$, which corresponds to the weighting scheme used by the MOESP class of subspace algorithms (Verhaegen 1994). This choice of weighting matrices is motivated by the desire to reduce the computational complexity. Even though the CCA weighting scheme might be more attractive from a statistical point of view, it involves the additional step of computing the inverse square root of the estimated correlation matrix $\widehat{\Gamma}_{zz}$. For the particular choice of weights, explicit computation of square root matrices can be avoided by using a square root implementation based on a single RQ factorization of the block Hankel matrices. This will be made more clear in Theorem 3.1. At this stage it is important to note that the square root implementation could be extended to include the CCA weighting scheme but this would involve the need for an additional RQ factorization step. Considering the new weights, the matrix M as defined in equation (3.13) becomes equal to $M = \sqrt{N}(\widehat{\Gamma}_{zp})(\widehat{\Gamma}_{pp})^{-1/2}$. By substituting the finite sample estimates (3.17) and (3.17), this gives rise to the following expression in terms of the data block Hankel matrices $Y_{0,p,N}$ and $Y_{p,f,N}$

$$M = (I - \widehat{\mathcal{T}}_f)Y_{p,f,N}Y_{0,p,N}^T \left(Y_{0,p,N}Y_{0,p,N}^T\right)^{-1/2}. \tag{3.17}$$

In the same way, the corresponding estimate of the extended controllability matrix, $\widehat{\mathcal{K}}_p = \sqrt{N}V_{n_d}^T\widehat{\Gamma}_{pp}^{-1/2}$, can be expressed as

$$\widehat{\mathcal{K}}_p = V_{n_d}^T \left(Y_{0,p,N}Y_{0,p,N}^T\right)^{-1/2}. \tag{3.18}$$

Given the estimate of the extended controllability matrix $\mathcal{K}_p$, the remaining steps of proposed subspace identification algorithm are again to reconstruct the approximate state sequence $\hat{x}(k) = V_{n_d}^T W_2^{-1}y_p(k-p)$, from which the system matrices are estimated by finite linear regression in the state-space equations. By selecting the right block-rows, also the finite linear regression step can be conveniently expressed in terms of the data block Hankel matrices $Y_{0,p,N}$ and $Y_{p,f,N}$. Using the same notational convention as for the data block Hankel matrices, the sequence of reconstructed states can be expressed as $\widehat{X}_{p,1,N} = \widehat{\mathcal{K}}_p Y_{0,p,N}$. Likewise, the sequence of time-shifted states is given by $\widehat{X}_{p+1,1,N} = \widehat{\mathcal{K}}_p Y_{1,p+1,N}$. By

considering the state-space equations (3.6), it should be clear that the finite linear regression problem used to estimate the system matrices $\widetilde{A}$ and $\widetilde{K}$ from the reconstructed state sequences $\widehat{X}_{p,1,N}$ and $\widehat{X}_{p+1,1,N}$, can formulated as the following least squares optimization problem

$$\min_{[\widetilde{A}|\widetilde{K}]} \left\| \widehat{X}_{p+1,1,N} - [\widetilde{A}|\widetilde{K}] \begin{bmatrix} \widehat{X}_{p,1,N} \\ Y_{p,1,N} \end{bmatrix} \right\|_F^2, \tag{3.19}$$

In the same way, the estimate of the system matrix $\widetilde{C}$ obtained by regressing $y_p(k-p)$ on $\hat{x}(k)$ can be expressed as the minimizing argument of the least squares optimization problem

$$\min_{\widetilde{C}} \|Y_{p,1,N} - \widetilde{C}\widehat{X}_{p,1,N}\|_F^2. \tag{3.20}$$

The above least squares optimization problems and the expressions for the reconstructed state sequences $\widehat{X}_{p,1,N}$ and $\widehat{X}_{p+1,1,N}$, give rise to the following estimate of the system matrices $\widetilde{A}$, $\widetilde{C}$ and $\widetilde{K}$

$$C[\widehat{\widetilde{A} \mid \widetilde{K}}] = \widehat{\mathcal{K}}_p Y_{1,p+1,N} \begin{bmatrix} \widehat{\mathcal{K}}_p Y_{0,p,N} \\ Y_{p,1,N} \end{bmatrix}^{\dagger} \tag{3.21}$$

$$\widehat{\widetilde{C}} = Y_{p,1,N}(\widehat{\mathcal{K}}_p Y_{0,p,N})^{\dagger}. \tag{3.22}$$

The obtained estimate of the triple $(\widetilde{A}, \widetilde{C}, \widetilde{K})$ in combination with the singular values Σ_1 of the geometry matrix, provide a complete specification of the system matrices of atmospheric disturbance model (3.4). By applying the definitions of $\widetilde{A}$, $\widetilde{C}$ and $\widetilde{K}$ it is possible to explicitly compute A_d, C_d and K_d. As will be demonstrated in Section 3.6, this last step is not always necessary for computing the controller.

   The outcome of each of the steps of the proposed subspace identification algorithm, that is the identification of the AR model, the problem of estimating the extended controllability matrix $\mathcal{K}_p$ and the finite linear regression step used to estimate the system matrices, has been expressed in terms of the data block Hankel matrices $Y_{0,p,N}$ and $Y_{p,1,N}$. Even though these expressions provide a direct way to estimate the system matrices $\widetilde{A}$, $\widetilde{C}$ and $\widetilde{K}$, this is computationally not attractive since the number of columns in the data block Hankel matrices is typically very large. However, due to specific choice of the weighting matrices W_1 and W_2, the solution to each of the sub-problems can be computed from the R factor of the stacked data block Hankel matrix $Y_{0,p+f,N}$. This results in an efficient implementation, both with respect to the number of flops and required memory storage. Furthermore, since the approach avoids the explicit computation of the square root matrices in (3.17) it will help to improve the numerical robustness. The precise relation between each of the steps of the subspace algorithm and the R factor of the data block Hankel matrix is summarized in Theorem 3.1. The use of a RQ factorization to improve the computational efficiently of subspace algorithms is by no means new and has been proposed before in Verhaegen (1994).

Theorem 3.1 (SSAR via RQ factorization) *Given a signal $y(k) \in \mathbb{R}^{m_y}$, consider the economy size RQ factorization of the block Hankel matrix $Y_{0,p+f,N} \in \mathbb{R}^{(p+f)m_y \times N}$*

$$
\begin{bmatrix} Y_{0,1,N} \\ Y_{1,p-1,N} \\ Y_{p,1,N} \\ Y_{p+1,f-2,N} \\ Y_{p+f-1,1,N} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & 0 & 0 \\ R_{21} & R_{22} & 0 & 0 \\ \vdots & \vdots & \ddots & 0 \\ R_{51} & R_{52} & \dots & R_{55} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_5 \end{bmatrix}, \tag{3.23}
$$

where the R and Q factor are partitioned in accordance with the partitioning of the block-rows of $Y_{0,p+f-1,N}$ and $p, f \in \mathbb{N}$ are such that $n_d \leq p\, m_y$, $f m_y \leq N$. Furthermore let the matrices M and $\widehat{\mathcal{K}}_p$ be defined as in (3.17) and (3.18). Then the solutions (3.16), (3.21) and (3.22) to the optimization problems (3.15), (3.19) and (3.20), respectively, can be completely characterized in terms of the R factor only:

1. *The solution to optimization problem (3.15) is given by*

$$
\widehat{\widetilde{C}\mathcal{K}}_{f-1} = \begin{bmatrix} R_{51} & \cdots & R_{54} \end{bmatrix} \begin{bmatrix} R_{41} & \cdots & R_{44} \end{bmatrix}^\dagger.
$$

2. *Let $\widehat{\mathcal{T}}_f$ denote the matrix obtained by replacing $\widetilde{C}\widetilde{A}^i\widetilde{K}$ in (3.10) by the $f-i^{th}$ block column of the estimate $\widetilde{C}\mathcal{K}_{f-1}$, then*

$$
\widehat{M} = (I - \widehat{\mathcal{T}}_f) \begin{bmatrix} R_{31} & R_{32} \\ R_{41} & R_{42} \\ R_{51} & R_{52} \end{bmatrix} \quad \widehat{\mathcal{K}}_p = V_{n_d}^T \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix}^\dagger,
$$

   where $V_{n_d}^T$ contains the right singular vectors corresponding to the n_d largest singular values of M.

3. *Given $\widehat{\mathcal{K}}_p$ and $V_{n_d}^T$, the solutions to the optimization problems (3.19) and (3.20) can be expressed as*

$$
\widehat{[\widetilde{A}|\widetilde{K}]} = \widehat{\mathcal{K}}_p \begin{bmatrix} R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \left[\begin{array}{cc|c} V_{n_d}^T & & 0 \\ R_{31} & R_{32} & R_{33} \end{array} \right]^\dagger
$$

$$
\widehat{\widetilde{C}} = \begin{bmatrix} R_{31} & R_{32} \end{bmatrix} (V_{n_d}^T)^\dagger.
$$

The solution to each of the optimization problems is unique if and only if the matrices in the pseudo-inverse are non-singular.

Proof: In order to prove the first statement, note that $Y_{p+1,f-3,N}$ and $Y_{p+f-2,1,N}$ can be easily expressed in terms of the sub-matrices of (3.23). The partitioning of the RQ factorization is such that these matrices correspond to the fourth and the fifth block row. Substituting the resulting expressions for $Y_{p+1,f-3,N}$ and $Y_{p+f-2,1,N}$ in (3.16) and using the orthonormality of Q gives the desired result.

The second and third statement can be proved by proceeding in a similar way. The matrices $Y_{0,p,N}$ and $Y_{p,p+f,N}$ are expressed in terms of the sub-matrices of RQ factorization (3.23) and the resulting expressions are substituted in (3.17) and (3.18). Using the orthogonality of Q leads to the given expressions for M and $\mathcal{K}_p$. The condition $n_d \leq p\,m_y, f m_y \leq N$ guarantees that M is sufficiently large to have at least n right singular vectors, as needed in constructing $V_{n_d}^T$.

In proving the third statement, $Y_{0,p,N}$, $Y_{1,p+1,N}$ and $Y_{1,p+1,N}$ are expressed in terms of the RQ factorization and the result is substituted in equations (3.21) and (3.22). Again due to the orthogonality the Q-factor drops out the equation. Finally by substituting the obtained expression for $\mathcal{K}_p$ in the pseudo-inverse term of the equations finishes the proof. $\qquad\square$

The above theorem forms the starting point for an efficient implementation of the proposed subspace identification algorithm. It shows that the entire SSAR algorithm can be characterized in terms of the R factor of the block Hankel matrix $Y_{0,p+f,N}$ and that here is no need to actually compute the Q factor. Since the computation of the R factor is the only operation performed directly on the data block Hankel matrices, it has a large influence on the overall computational efficiency of the algorithm. To arrive at an efficient implementation, the R factor is computed by using the fast algorithm described by Mastronardi et al. (2001). Especially for large block Hankel matrices, this algorithm is far more efficient than a standard RQ decomposition based on Householder transformations, as it is able to exploit the displacement structure in the block Hankel matrix $Y_{0,p+f,N}$.

The precise choice of the user defined parameters f and p remains a difficult issue in subspace identification (Bauer 2004). In the SSARX algorithm, it is assumed that the parameters f and p are always chosen strictly larger than the order of the system n_d. With the stochastic disturbance model (3.4) being a minimal representation of the stochastic process, this is a sufficient condition to ensure that the matrix $\mathcal{B} = \mathcal{O}_f \mathcal{K}_p$ has rank n_d and can be factorized to find an estimate of the extended controllability matrix $\mathcal{K}_p$. For multi-variable systems, the requirement $f, p > n_d$ may be overly conservative. As will be demonstrated by the numerical validation experiments in Section 3.7, it may be useful to choose the parameters f and p much smaller than the order of the identified disturbance model. The parameters f and p should however be chosen sufficiently large to guarantee that the rank of the matrix $\mathcal{B}$ is larger than the order of the identified disturbance model. A thorough consistency analysis of both the SSARX and the proposed algorithm is still a topic for future research.

3.4 Numerical validation disturbance modeling

To demonstrate that the proposed subspace identification algorithm can indeed be used to identify an accurate atmospheric disturbance model, numerous of simulation experiments have been performed. This section reports on the results of two such identification experiments, each performed on a different data set. The first

simulation example considers open-loop WFS data obtained from an AO laboratory setup at TNO Science and Industry. This data set is particularly interesting as is will also be used in the closed-loop simulation examples in Section 3.7. The data set is generated by a turbulence simulator that can be considered as truly frozen. The second simulation example considers open-loop WFS data from a real telescope and is important to verify if proposed subspace identification algorithm is also useful in modeling realistic turbulence scenarios.

3.4.1   Simulation procedure and performance measures

Each of the simulation experiments considered in this paper is based on a data set consisting $N_t = 1.0 \times 10^4$ samples of open-loop WFS data. This data batch is divided into two parts. The first $N_i = 8000$ samples are used to identify an atmospheric disturbance model with the proposed SSAR subspace identification algorithm, while the remaining $N_v = 2000$ samples are reserved for validation purposes. In accordance with this functional division, the first part of the data set is referred to as the identification set, while the second part is called the validation set. The performance of the identified atmospheric disturbance model is judged by considering its ability to predict the wavefront distortions, or more specifically, the wavefront at next sample instant. For an atmospheric disturbance model of the from (3.4), the optimal one-step ahead predictor that minimizes the prediction error in a mean-square sense, is given by (see e.g. Kailath et al. 2000)

$$\begin{cases} x(k+1) = \widetilde{A}\,x(k) + \widetilde{K}\,y(k) \\ \hat{\varphi}(k|k-1) = \Sigma_1^{-1}\widetilde{C}\,x(k) \end{cases}, \tag{3.24}$$

where $\hat{\varphi}(k|k-1)$ denotes the estimate of $\varphi(k)$ given all past WFS measurements up to and including time $k-1$. The performance of the optimal one-step ahead predictor derived from the identified atmospheric disturbance model is compared with the optimal predictor for a random walk process. Here, a random walk process is defined as a process of which the increments from sample to sample are zero-mean and white. Hence, for such a process, the optimal wavefront estimate at the next sample is equal the wavefront reconstructed from the current WFS measurement, i.e. $\hat{\varphi}(k|k-1) = \Sigma_1^{-1}y(k-1)$. Given the estimates from the optimal and the random walk predictor, the wavefront prediction error is computed as

$$\epsilon(k) = \varphi(k) - \hat{\varphi}(k|k-1).$$

The wavefront prediction error is evaluated for both predictors on both the identification and the validation set. The reason for considering the random walk predictor is motivated by the fact that it bears a close resemblance with the common AO control law. Indeed, it can be shown that the temporal error in an AO system using static wavefront reconstruction and pure integrators as feedback compensators, gives rise to the same prediction error as the random walk approach.

In quantifying the performance of both the random walk predictor and the optimal one-step ahead predictor derived from the identified atmospheric disturbance model, two different performance measures have been considered. Both of

these performance measures are computed on the basis wavefront prediction error $\varepsilon(k)$. The first performance measure can be interpreted as a normalized version of the mean-square prediction error and is defined as

$$J_1 = \frac{\sum_{k=1}^{N_s} \text{tr}(\varepsilon(k)\varepsilon^T(k))}{\sum_{k=1}^{N_s} \text{tr}(\varphi(k)\varphi^T(k))}, \tag{3.25}$$

where N_s denotes the number of samples in the performance evaluation, which is N_i for the identification set and N_v for the validation set. The performance index J_1 provides a quantitative measure of the total reduction in the mean-square prediction error and is independent of the scaling of the wavefront. To obtain some more insight in the temporal dynamics of the controller it is useful to consider the normalized averaged power spectrum $P(\omega)$ of the simulated residual phase error. The normalized averaged power spectrum $P(\omega)$ is defined as

$$P(\omega) = \frac{\sum_{j=1}^{m_y} \Phi_j(\omega)}{1/(N_s - 1)\sum_{k=1}^{N_s} \text{tr}(\varphi(k)\varphi^T(k))}, \tag{3.26}$$

where $\Phi_j(\omega)$ is the estimated power spectral density of the jth component of the residual wavefront $\varepsilon(k)$, evaluated at the frequency ω. The spectrum $\Phi_j(\omega)$ is computed as the Welch averaged periodogram with a window size of 256. The performance of the optimal one-step ahead predictor has been evaluated for atmospheric disturbance models of different orders.

3.4.2 Open-loop WFS data AO laboratory setup

The open-loop WFS data used in the first set of simulation experiments have been obtained from an AO test bench. The AO test bench uses a turbulence simulator consisting of a circular plan parallel glass plate that is rotated through the optical beam by means of a driving stage. One side of the glass plate has been machined in such a way that the glass plate introduces spatial wavefront distortions characterized by a Kolmogorov spectrum with a $D/r_0 = 5$, where D is the diameter of the simulated telescope aperture and r_0 denotes the Fried parameter (Hardy 1998). This type of turbulence simulator gives rise to a single frozen layer satisfying the Taylor hypothesis. The open-loop WFS data are recorded by using a Shack-Hartmann WFS with an 16×16 orthogonal micro-lens array. After aligning the setup, only the best illuminated micro-lenses and least distorted spots were selected for wavefront sensing. The active spots have been selected in such a way that no micro-lenses within the convex hull of the selected spots are excluded. In other words, there are no holes or missing spots in the set of selected active micro-lenses. This resulted in a total of 53 active spots, which corresponds to a WFS signal $s(k)$ of dimension $m_s = 106$. The geometry matrix G used for reconstructing the wavefront was defined according to the Fried geometry. Considering this geometry matrix, projecting out the unobservable modes leads to a reduced WFS signal $y(k) = U_1^T s(k)$ of dimension $m_y = 69$. The open-loop WFS has been collected at a sampling rate of $f = 25\,\text{Hz}$. The rotational speed of the glass plate results in a Greenwood frequency of $f_G = 0.95\,\text{Hz}$. Since the temporal error scales

as $\sigma_T^2 \propto (f_G/f)^{5/3}$ (Hardy 1998), the AO test bench has the same temporal error as an AO system with a sample frequency of $f = 750\,\mathrm{Hz}$ and a Greenwood frequency of $f_G = 28.5\,\mathrm{Hz}$.
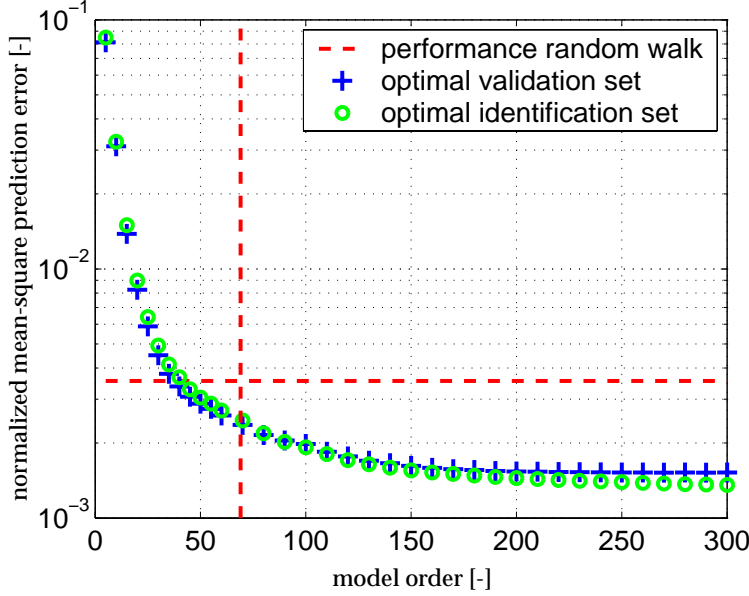

Figure 3.2: Normalized mean-square wavefront prediction error J_1 as a function of the order of the identified disturbance model on AO breadboard data. The center of the cross formed by the dashed lines indicates the performance and the effective model order of the random walk predictor.

Figure 3.2 shows the normalized mean-square prediction error J_1, for the optimal one-step ahead predictor, as a function of the order n_d of the identified disturbance model. The crosses in the figure correspond to the performance estimates computed on the basis of the validation data set, while circles are used to indicate the performance corresponding to the identification data set. Furthermore, the large cross formed by the dashed lines indicates the performance level and the effective model order of the random walk approach. Note that since the random walk approach uses only the last WFS measurements to obtain an estimate of the wavefront, it only needs to remember this measurement, which gives rise to an effective model order equal to the number of WFS channels. The results depicted in Figure 3.2 have been obtained by using $p = 15$ past and $f = 15$ future block-rows in the proposed subspace identification algorithm. Similar results are obtained for $p = f \in \{20, 25, 30, 35\}$, with a slightly lower performance for the higher values of p and q. This weak dependence, simplifies the choice parameters p and q. The slightly improved performance for smaller p and q is probably caused by the fact that the smaller the number of block-rows, the more columns of the block-Hankel matrix can be formed, which in turn has a positive influence on the accuracy of the estimated extended observability. It does not need any explanation that p and

q should be chosen sufficiently large to capture all dynamics. The figure shows that the performance of the identified disturbance model increases monotonically with the model order, and that for model orders $n_d \geq 40$ the identified disturbance model provides already a better estimate than the random walk approach. When the model order is equal to the equivalent order of the random walk predictor, i.e. $n_d = 69$, the normalized mean-square prediction error on the identification and validation set are $J_1 = 2.38 \times 10^{-3}$ and $J_1 = 2.49 \times 10^{-3}$. Compared to the random walk predictor with $J_1 = 3.55 \times 10^{-3}$, this is an improvement of respectively 33.0% and 29.9%. The figure also shows that for increasing model orders the performance of the optimal one-step ahead predictor converges to a more or less constant level. For $n_d = 300$, the performance is $J_1 = 1.35 \times 10^{-3}$ for the identification set and $J_1 = 1.52 \times 10^{-3}$ for the validation set. This corresponds to an performance improvement of respectively 65.7% and 60.5%, with respect to the random walk predictor. So, even though the performance on the validation set is somewhat lower that on the identification set, the simulation example shows that there still plenty of room to improve on the random walk predictor.

Finally, it is interesting to note that the proposed subspace identification algorithm is indeed able to capture (at least part of) the frozen like behavior of the wavefront. To illustrate this, the identified atmospheric disturbance model for $n_d = 300$ has been excited with a zero-mean white noise sequence. Figure 3.3 shows 24 out of the $N_t = 1.0 \times 10^4$ frames of the observable part of the wavefront $V_1 V_1^T \phi(k) = V_1 \varphi(k)$ generated in this way. The depicted sequence of frames clearly shows a wavefront that is propagating over the aperture. Since the input to the disturbance model 3.4 has been zero-mean and white, the observed frozen like behavior can only be due to the disturbance model. In other words, the observed frozen like behavior has to be part of the identified atmospheric disturbance model.

3.4.3   Open-loop WFS data from William Herschel Telescope

The previous subsection has shown that the proposed subspace identification algorithm is able to identify an atmospheric disturbance model that can be used to improve on random walk prediction. It is important to verify whether or not this also holds for real turbulence. For this reason, the second data set consists of real-life open-loop WFS data obtained from the William Herschel Telescope (WHT). With a diameter of $4.2\,\mathrm{m}$, the WHT is the largest of the Isaac Newton Group of telescopes at the Observatorio del Roque de Los Muchachos, La Palma, Spain. The data set has been collected by one of the two seeing monitors, implemented as part of the Joint Observatories Seeing Evaluation (JOSE) project. The JOSE seeing monitor has been designed to resemble the WFS sensor of the WHT AO system as closely as possible, in order to examine both the spatial and temporal properties of turbulence on the appropriate scale. It consists of a Shack-Hartmann WFS with an 8×8 orthogonal micro-lens array, which partitions the telescope aperture in sub-apertures with an effective size of $0.5\,\mathrm{m}$. Each sub-aperture has a field of view of $2.2\,\mathrm{arcsec}$ and is imaged on a 8×8 sub-region of a 64×64 CCD camera, giving rise to a sensitivity of approximately 0.27 arcsec/pixel. The device is operated in
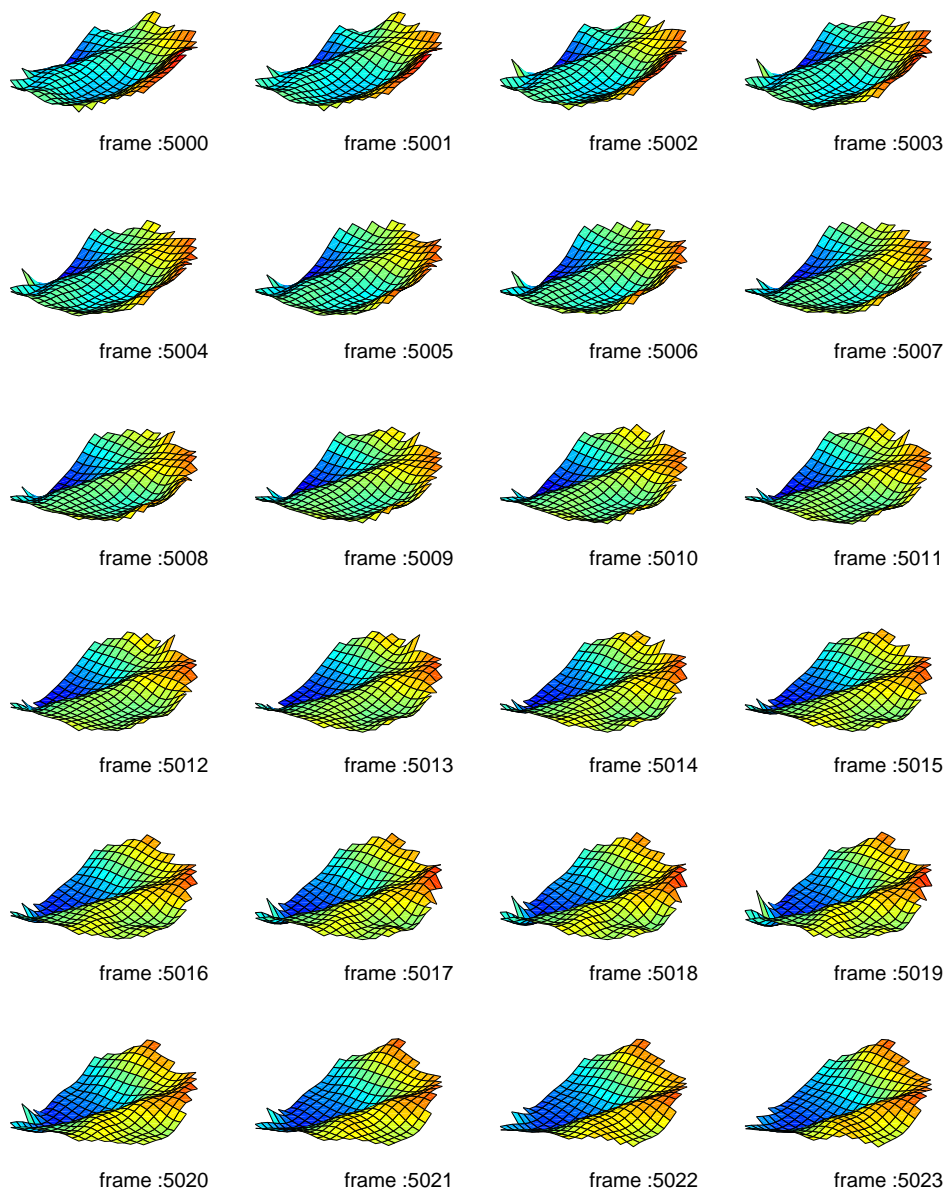
Figure 3.3: Wavefront distortions generated by exciting the identified atmospheric disturbance model ($p = f = 15$ and $n_d = 300$) by zero-mean white noise. The wavefront sequence clearly possess a frozen flow component. This shows that the subspace identification algorithm is able to capture the second-order statistics associated with the frozen WFS data obtained from the AO breadboard.

the visible at a wavelength around $700\,\mathrm{nm}$. For a more detailed description of the JOSE seeing monitor the reader is referred to Saint-Jacques (1998). The data used in the simulations have been recorded on June 5, 1997, under good seeing conditions characterized by a full width of half maximum (FWHM) of $0.5 - 0.7$ arcsec. The sampling rate during the data acquisition was $296\,\mathrm{Hz}$ and the centroid positions have an accuracy of approximately 0.1 pixels. Just as in simulations on the breadboard data, the phase-to-slope matrix G is defined according to a Fried geometry. Considering only the 48 best illuminated spots, this gives rise to a reduced WFS signal of dimension $m_y = 67$, which is close to the number of channels in the breadboard data set. An important difference with the previous data set, however, is that the WFS layout of the JOSE seeing monitor contains a central obscuration, because of the telescope's secondary mirror.
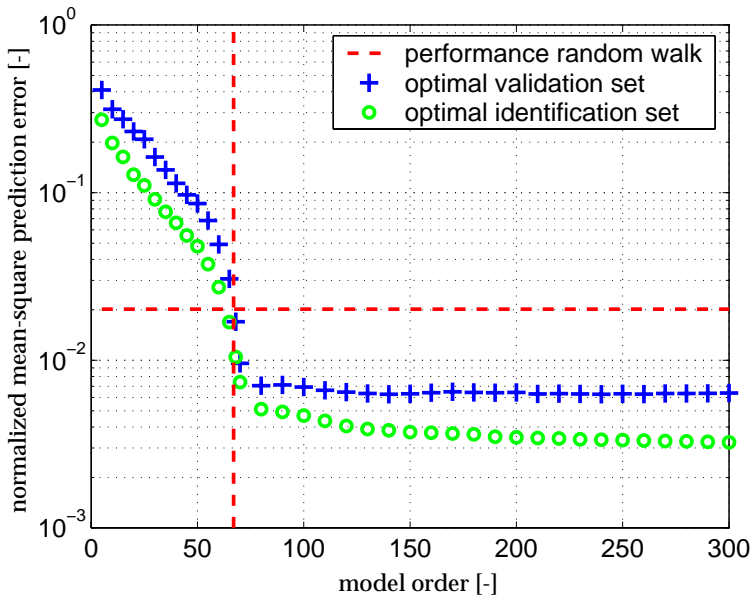

Figure 3.4: Normalized mean-square wavefront prediction error J_1 as a function of the order of the identified disturbance model on William Herschel telescope data. The center of the cross formed by the dashed lines indicates the performance and the effective model order of the random walk predictor.

The outcome of the simulation experiments on the open-loop telescope data is depicted in Figure 3.4. Just as in the previous experiments, the circles denote the normalized mean-square prediction error J_1 computed on the basis of the identification data set, while the crosses correspond to the performance on the validation data. The center of the large cross formed by the dashed lines is again used to denote the performance level and the effective model order of the random walk approach, which in this case are given by $n_d = 67$ and $J_1 = 2.02 \times 10^{-2}$. The simulation results presented in the figure have been obtained by choosing $p = f = 30$. Like the simulations of the breadboard data, the performance of the subspace

identification algorithm has been analyzed for $p = f \in \{15, 20, 25, 30\}$ with almost identical results. In contrast to the previous experiments, however, the performance is slightly better for higher values of p and q. In the trade off between modeling the higher order dynamics and improving the statistical averaging by increasing the number columns in the block-Hankel matrix, the modeling of the higher order dynamics now seems to gain the upper hand. Furthermore, note the simulation experiments on the breadboard data and the real WFS data show a similar trend. In both cases, the wavefront prediction error decreases monotonically with the model order and seem to converge to a constant level. The order at which the optimal one-step ahead predictor starts to improve on the random walk approach however, is much closer to the effective order of the random walk predictor. A possible explanation for this is that the turbulence can no longer be seen as a single frozen layer. For frozen turbulence, many of the WFS channels have common dynamics, which may help to reduce the model order. The normalized mean-square residual prediction error achieved for a model order $n_d = 300$ is $J_1 = 3.24 \times 10^{-3}$ for the identification set and $J_1 = 6.38 \times 10^{-3}$ for the validation set. Compared to the random walk approach, this corresponds to a reduction of respectively 84.0% and 68.4%. The relative improvement on real telescope data is therefore even larger than on the breadboard data. This can be explained by the fact that, when considering the temporal error, the equivalent Greenwood frequency for the AO breadboard data at the sampling rate of the telescope is only $f_G = 11.25\,\mathrm{Hz}$. Since typical values for the Greenwood frequency are around the $20 - 30\,\mathrm{Hz}$, the temporal error is likely to be much smaller on the breadboard data set than on the telescope data set. This implies that there is potentially more to gain by applying the optimal one-step ahead predictor on the real telescope data.

For $n_d = 300$, Figure 3.5 shows the normalized averaged power spectra $P(\omega)$ of the open-loop wavefront distortions (upper dashed-dotted line), of the wavefront prediction error using the random walk approach (solid line), of the wavefront prediction error using the optimal one-step ahead predictor on validation data (dashed line) and of the wavefront prediction error using the optimal one-step ahead predictor on the identification data (lower dashed-dotted line). Also from this figure it is clear that the proposed subspace identification algorithm is able to identify an accurate atmospheric disturbance model that can be used to considerably reduce the wavefront prediction error. Note that the wavefront prediction error corresponding the optimal one-step ahead predictor is almost white. This shows that the identified atmospheric disturbance model provides an accurate description of spatial averaged temporal correlation. The wavefront prediction error on the validation set, on the other hand, is slightly colored. Note that the slightly reduced performance on the validation set is also apparent from Figure 3.4. Whether this is caused by the fact that the lower frequencies are more difficult to model or by slow changes in the turbulence statistics is hard to distinguish. To verify if the assumption of modeling the turbulence as a regular stochastic process is reasonable on a time scale of a few minutes, two data sets each consisting of $N_t = 1.0 \times 10^4$ have been recorded 5 minutes and 2 seconds apart. The first data set has been used to identify an atmospheric disturbance model, while the second set is used for performance evaluation. Considering this validation set, the relative reduction of the mean-square wavefront prediction error compared to
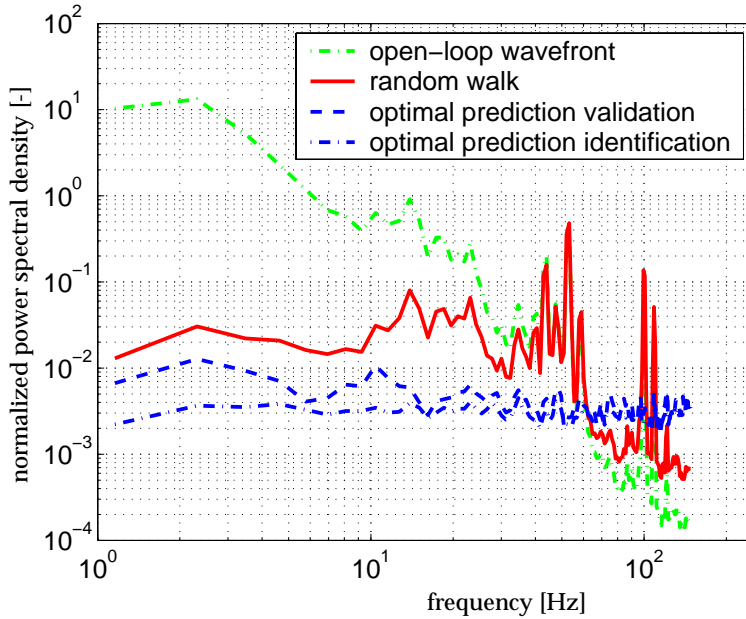
Figure 3.5: Normalized averaged power spectrum $P(\omega)$ of wavefront prediction error on William Herschel telescope data.

the random walk approach is still 58.9%. This shows that the stationarity assumption is still usable over a time horizon of minutes, but that in order to guarantee a good performance over significant longer time scales a regular update of the disturbance model is probably necessary.

3.5 AO in an optimal control framework

Given an atmospheric disturbance model of the form (3.4), this section provides a general recipe to approach the AO control problem as formulated in Section 3.2. An important aspect of the AO control problem is that there is a difference between the objective of minimizing the mean-square residual WFS signal $r(k)$ and the actual cost function (3.5). The $\mathcal{H}_2$-optimal control framework provides an attractive way to deal with this discrepancy between measurement and control objective. It will be shown that the AO control problem can be conveniently expressed as an $\mathcal{H}_2$-optimal control problem. Standard $\mathcal{H}_2$-optimal control theory can then be used to compute the closed-loop optimal controller.

Consider the block-diagram in Figure 3.6, which provides a schematic representation of the closed-loop AO control system. The shaded box in this figure defines the so-called generalized plant $\mathcal{P}(z)$, which forms the starting point in formalizing the $\mathcal{H}_2$-optimal control framework. The generalized plant makes a
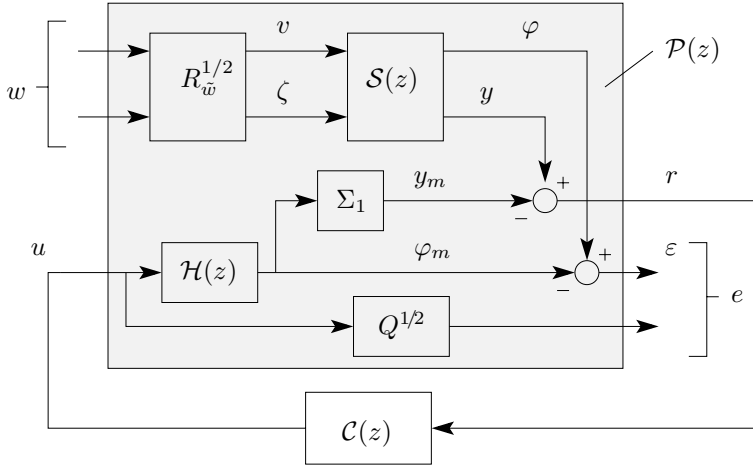
Figure 3.6: Block-diagram of AO system in generalized plant representation

clear distinction between exogenous zero-mean white noise inputs $w(k)$ and control inputs $u(k)$ on the one hand and measurement outputs $r(k)$ and performance outputs $e(k)$ on the other hand. In the following, it will be assumed that the zero-mean white noise input $w(k)$ has unit covariance $\mathcal{E}\{w(k)w^T(k)\} = I$ and that the generalized plant has the following state-space description

$$\begin{bmatrix} \xi(k+1) \\ \hline e(k) \\ r(k) \end{bmatrix} = \begin{bmatrix} A & B_w & B_u \\ \hline C_e & D_{ew} & D_{eu} \\ C_r & D_{rw} & 0 \end{bmatrix} \begin{bmatrix} \xi(k) \\ \hline w(k) \\ u(k) \end{bmatrix}, \qquad (3.27)$$

where $\xi(k)$ denotes the state variable. To facilitate the discussion, it is useful to partition the generalized plant $\mathcal{P}$ in the same way as the input and output signals. The input-output relation is described by the open-loop transfer functions $\mathcal{P}_{ew}(z)$, $\mathcal{P}_{eu}(z)$, $\mathcal{P}_{rw}(z)$ and $\mathcal{P}_{ru}(z)$, where the subscripts refer to the corresponding input and output signals. Using this notation the feedback connection of $\mathcal{P}$ and $\mathcal{C}$ is given by

$$\mathcal{F}(\mathcal{P}, \mathcal{C}) = \mathcal{P}_{ew} + \mathcal{P}_{eu}(I - \mathcal{C}\mathcal{P}_{ru})^{-1}\mathcal{C}\mathcal{P}_{rw}.$$

$\mathcal{H}_2$-optimal control theory deals with the problem of finding the causal controller $\mathcal{C}(z)$ that minimizes the $\mathcal{H}_2$-norm of the closed-loop transfer function $\mathcal{F}(\mathcal{P}, \mathcal{C})$. The following theorem provides a solution to the general $\mathcal{H}_2$-optimal control problem (see Shu and Chen 1995; Chen and Francis 1995; Zhou et al. 1996).

Theorem 3.2 (Discrete-time $\mathcal{H}_2$-optimal control) *Consider the generalized plant $\mathcal{P}(z)$ with state-space realization* (3.27) *and let the following conditions be satisfied*

- (A, B_u) *is stabilizable and* (A, C_r) *is detectable;*

- $R_{eu} \doteq D_{eu}^T D_{eu} > 0$ *and* $R_{rw} \doteq D_{rw}D_{rw}^T > 0$;

- *the matrices*

$$\begin{bmatrix} A - \lambda I & B_u \\ C_e & D_{eu} \end{bmatrix}, \begin{bmatrix} A - \lambda I & B_w \\ C_r & D_{rw} \end{bmatrix}$$

have full rank for all $\lambda \in \mathbb{C}$ such that $|\lambda| = 1$.

Under these conditions, the following Riccati equations have a unique and stabilizing solution $X = X^T \geq 0$ and $Y = Y^T \geq 0$

$$\begin{aligned} X &= AXA^T - (AXC_r^T + S_{rw}^T)(C_r X C_r^T + R_{rw})^{-1}(\cdot)^T + Q_w \\ Y &= A^T Y A - (A^T Y B_u + S_{eu})(B_u^T Y B_u + R_{eu})^{-1}(\cdot)^T + Q_e \end{aligned}$$

where $Q_w \doteq B_w B_w^T$, $Q_e \doteq C_e^T C_e$ $S_{eu} \doteq C_e^T D_{ew}$, and $S_{rw} \doteq \mathcal{B}_w D_{rw}^T$. With X and Y the stabilizing solutions to the above Riccati equations, define the matrices

$$\begin{aligned} F &\doteq (B_u^T Y B_u + R_{eu})^{-1}(B_u^T Y A + S_{eu}^T) \\ F_0 &\doteq (B_u^T Y B_u + R_{eu})^{-1}(B_u^T Y B_w + D_{eu}^T D_{ew}) \\ L &\doteq (AXC_r^T + B_w D_{rw}^T)(C_r X C_r^T + R_{rw})^{-1} \\ L_0 &\doteq (FXC_r^T + F_0 D_{rw}^T)(C_r X C_r^T + R_{rw})^{-1} \end{aligned}$$

Then the proper real-rational controller $\mathcal{C}(z)$ which internally stabilizes $\mathcal{P}(z)$ and is minimizing the $\mathcal{H}_2$-norm of the transfer function $\mathcal{F}(\mathcal{P}, \mathcal{C})$ from e to w, is given by

$$\begin{bmatrix} \hat{\xi}(k+1) \\ u(k) \end{bmatrix} = \left[\begin{array}{c|c} A + B_u L_0 C_r - B_u F - L C_r & B_u L_0 - L \\ \hline F - L_0 C_r & -L_0 \end{array} \right] \begin{bmatrix} \hat{\xi}(k) \\ r(k) \end{bmatrix},$$

with $\hat{\xi}(k)$ the estimate of $\xi(k)$ given $y(i)$, $i \leq k - 1$.

As depicted in Figure 3.6, the AO control system can be extended to fit in the generalized plant framework. By moving the system boundaries and considering the atmospheric disturbance model as a part of the generalized plant, the only exogenous inputs are the zero-mean white noise inputs $v(k)$ and $\zeta(k)$ of the disturbance model. Even though the signals are zero-mean and white, the signal $\tilde{w}(k) = [v^T(k) \; \zeta^T(k)]^T$ obtained by stacking $v(k)$ and $\zeta(k)$ cannot be directly used as the exogenous input $w(k)$ of the generalized plant because it does not have unit covariance. In order to ensure that also this condition is satisfied, the atmospheric disturbance model is augmented with a static matrix multiplication to normalize the input covariance. Let the covariance matrix of $\tilde{w}(k)$ be denoted by $R_{\tilde{w}}$, then the input weight needed to generate a signal $\tilde{w}(k)$ with the desired statistical properties out of the exogenous input signal $w(k)$, is equal to the square root of the covariance matrix $R_{\tilde{w}}$. The input to the atmospheric disturbance model $\tilde{w}(k)$ and the exogenous input to the generalized plant $w(k)$ are therefore related as $\tilde{w}(k) = R_{\tilde{w}}^{1/2} w(k)$. In order to evaluate $R_{\tilde{w}}^{1/2}$, note that the measurement noise $\nu(k)$ shoud be uncorrelated with the wavefront distortions $\varphi(k)$ and the state vector $x(k)$. By multiplying the second output equation in (3.4) from the right by $\nu^T(k)$ and substituting the definition of $\zeta(k)$, we can therefore write

$$\begin{aligned} \mathcal{E}\left\{ \varphi(k)\nu^T(k) \right\} &= \mathcal{E}\left\{ (C_d x(k) - \Sigma_1^{-1}\zeta(k))\nu^T(k) \right\} = 0 \\ &= \Sigma_1^{-1}(R_{v\nu} - R_\nu), \end{aligned}$$

where $R_{v\nu} \doteq \mathcal{E}\{v(k)\nu^T(k)\}$ and $R_\nu \doteq \mathcal{E}\{\nu(k)\nu^T(k)\}$. From the above equation we infer that $R_{v\nu} = R_\nu$, which can be used to express the cross-covariance $R_{v\zeta}$ as

$$R_{v\zeta} = \mathcal{E}\left\{v(k)\left((v^T(k) - \nu^T(k))\Sigma_1^{-1}\right)\right\} = (R_v - R_\nu)\Sigma_1^{-1}.$$

In the same way, the coveriance matrix R_ζ can be expressed as $R_\zeta = \Sigma_1^{-1}(R_v - R_\nu)\Sigma_1^{-1}$. From this it can be easily verified that the weighting matrix $R_{\tilde{w}}^{1/2}$ is given by

$$R_{\tilde{w}}^{1/2} = \begin{bmatrix} R_v & (R_v - R_\nu)\Sigma_1^{-1} \\ \Sigma_1^{-1}(R_v - R_\nu) & \Sigma_1^{-1}(R_v - R_\nu)\Sigma_1^{-1} \end{bmatrix}^{1/2} = \begin{bmatrix} R_v^{1/2} & 0 \\ \Xi_1 & \Xi_2 \end{bmatrix},$$

where the matrices $\Xi_1 \doteq \Sigma^{-1}(R_v^{1/2} - R_\nu R_v^{-1/2})$ and $\Xi_2 \doteq \Sigma_1^{-1}(R_\nu - R_\nu R_v^{-1} R_\nu)^{1/2}$, as we will see, do not effect the controller design. Since the atmospheric disturbance model is assumed to be minimal, $R_v^{1/2}$ is non-singular. This in combination with the non-singularity of the measurement noise covariance matrix R_ν, implies that the chosen input weighting $R_{\tilde{w}}^{1/2}$ has full rank. Because $R_{\tilde{w}}^{1/2}$ is upper triangular, the input weighting preserves the minimum-phase property of the atmospheric disturbance model with respect to $y(k)$.

In order to express the AO control problem in a $\mathcal{H}_2$ optimal control framework, it is still necessary to choose an appropriate performance output $e(k)$. The performance output $e(k)$ should be chosen in such way that it is consistent with the AO control objective. This can be achieved by choosing the performance output as $e(k) = [\varepsilon^T(k) \ u^T(k)Q^{1/2}]^T$. For this particular choice, the MSE of the performance output takes the same value as the cost function (3.5). Furthermore, it follows by the Parseval theorem that the $\mathcal{H}_2$-norm of the closed-loop transfer function $\mathcal{F}(\mathcal{P}, \mathcal{C})$ is equal to the MSE value of $e(k)$, i.e.

$$J = \text{tr}\mathcal{E}\{e(k)e^T(k)\} = \|\mathcal{F}(\mathcal{P}, \mathcal{C})\|_2^2$$

where $\|\cdot\|_2^2$ denotes the $\mathcal{H}_2$-norm. This implies that the AO control problem reduces the standard $\mathcal{H}_2$-optimal control problem of finding the controller $\mathcal{C}(z)$ that minimizes $\mathcal{H}_2$-norm of the closed-loop transfer function $\mathcal{F}(\mathcal{P}, \mathcal{C})$.

The shaded block in Figure 3.6 illustrates the relation between the generalized plant with performance output $e(k)$ and the different components in the AO system like the DM, the WFS and the atmospheric disturbance model. By combining the WFS model (3.2) with the state-space representations of the DM model (3.3), the atmospheric disturbance model (3.4) and the input weighting (??), it is possible to derive the following state-space description for the generalized plant $\mathcal{P}(z)$

$$\left[\begin{array}{c} \xi(k+1) \\ \hline e(k) \\ \hline r(k) \end{array}\right] = \left[\begin{array}{cc|cc|c} A_m & 0 & 0 & 0 & B_m \\ 0 & A_d & K_d R_v^{1/2} & 0 & 0 \\ \hline -C_m & C_d & \Xi_1 & \Xi_2 & 0 \\ 0 & 0 & 0 & 0 & Q^{1/2} \\ \hline -\Sigma_1 C_m & \Sigma_1 C_d & R_v^{1/2} & 0 & 0 \end{array}\right] \left[\begin{array}{c} \xi(k) \\ \hline w(k) \\ \hline \tilde{u}(k) \end{array}\right].$$

Given the above state-space realization of the generalized plant, a general strategy for computing the optimal controller is to apply Theorem 3.2.
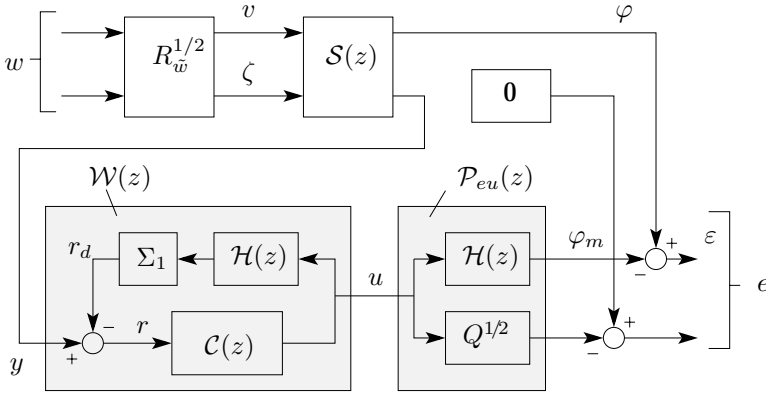
Figure 3.7: Block-scheme of equivalent feedforward AO control system

3.6   Computing the optimal controller

In the previous section a general recipe to approach the AO control problem has been provided. From Theorem 3.2 it is clear that this strategy typically requires the solution of two Riccati equations. In this section it will be shown that since the atmospheric disturbance model is in innovation form, only one Riccati equation needs to be solved to compute the optimal controller. In fact, the Riccati equation which is directly related to the spectral factorization problem can be avoided. Furthermore, it will be shown that in the case that the settling time of the step response of the DM is negligible with respect to the WFS exposure time an analytical expression for the optimal controller can be derived. An analytical solution is attractive from a computational point of view because the poles of the atmospheric disturbance model typically cluster in the neighborhood of the point $z = -1$. When the poles are close to the unit circle, standard Riccati solvers (Arnold and Laub 1984) may suffer from convergence problems and increased numerical sensitivity. An analytical solution provides also more insight in the relation with the common AO control approach than a numerical solution.

Instead of approaching the AO control problem by applying Theorem 3.2, a different strategy is used. In this section the AO feedback control problem is transformed in a feedforward control problem, which is solved by causal Wiener filtering. This strategy is known as the Internal Model Control (IMC) approach (Morari and Zafiriou 1989). The main difficulty in determining the controller $\mathcal{C}(z)$ is that it enters the residual wavefront error $\varepsilon(k)$ in a nonlinear way. Since the WFS and DM mirror models are assumed to be known, the open-loop wavefront signal $y(k)$ can be reconstructed by subtracting the influence of the DM on the measured WFS signal $r(k)$, i.e. $y(k) = r(k) - \Sigma_1 \mathcal{H}(z)u(k)$. In this way, it is possible to open the feedback loop and replace the original control problem by the problem of finding the optimal feedforward controller $u(k) = \mathcal{W}(z)y(k)$. Figure 3.7 provides a block-diagram of the equivalent feedforward AO control system. As can be easily verified, the feedforward controller $\mathcal{W}(z)$ and feedback controller $\mathcal{C}(z)$ are related

in the following way

$$\mathcal{C}(z) = \mathcal{W}(z)(I - \Sigma_1 \mathcal{H}(z)\mathcal{W}(z))^{-1}. \tag{3.28}$$

For stable $\mathcal{H}(z)$, the class of stable $\mathcal{W}(z)$ parametrizes the class of stabilizing closed-loop controllers $\mathcal{C}(z)$. This parametrization of all stabilizing controllers is the so-called Youla parametrization (Youla et al. 1976). Considering the performance output $e(k)$ of the generalized plant $\mathcal{P}$ as defined in the previous section, it follows again from the Parseval theorem that the relation between the cost function (3.5) and the $\mathcal{H}_2$-norm of the open-loop transfer function $\mathcal{P}_{ew} - \mathcal{P}_{eu}\mathcal{W}\mathcal{P}_{yw}$ is given by

$$J = \|\mathcal{P}_{ew} - \mathcal{P}_{eu}\mathcal{W}\mathcal{P}_{yw}\|_2^2, \tag{3.29}$$

where $\mathcal{P}_{yw}$ denotes the open-loop transfer function from the exogenous white noise input $w(k)$ to the reconstructed open-loop WFS signal $y(k)$. With a slight abuse of notation, $\mathcal{P}_{ew}$ is used to denote the transfer function describing the influence of $w(k)$ on the performance output $e(k)$. Given the transfer functions $\mathcal{P}_{eu}(z)$, $\mathcal{P}_{ew}(z)$ and $\mathcal{P}_{yw}(z)$, the solution to the problem of finding the optimal feedforward controller $\mathcal{W}(z)$ that minimizes cost function (3.29) is given by Lemma 3.1. If the transfer function $I + \mathcal{P}_{eu}(z)\mathcal{W}(z)$ is invertible, the feedback optimal stabilizing controller $\mathcal{C}(z)$ can be computed from the Youla parametrization (3.28).

Lemma 3.1 (Causal Wiener filter (Vidyasagar 1988)) *Let $\mathcal{P}_{ew}$, $\mathcal{P}_{eu}$ and $\mathcal{P}_{yw}$ belong to the set of asymptotically stable rational transfer functions and assume that $\mathcal{P}_{eu}(z)$ and $\mathcal{P}_{yw}(z)$ do not loose rank $\forall |z| = 1$. Then the optimal feedforward controller $\mathcal{W}(z)$ which is minimizing (3.29) is given by*

$$\mathcal{W} = -\mathcal{P}_{eu,o}^{\dagger} \left[\mathcal{P}_{eu,i}^* \mathcal{P}_{ew} \mathcal{P}_{yw,i}^*\right]_+ \mathcal{P}_{yw,o}^{\dagger} \tag{3.30}$$

with $[\cdot]_+$ the causality operator, $\mathcal{P}_{eu} = \mathcal{P}_{eu,i}\mathcal{P}_{eu,o}$ the inner-outer factorization of $\mathcal{P}_{eu}$, $\mathcal{P}_{yw} = \mathcal{P}_{yw,o}\mathcal{P}_{yw,i}$ the outer-inner factorization of $\mathcal{P}_{yw}$ and $(\cdot)^{\dagger}$ the left or right pseudo inverse. Here, $\mathcal{P}_{eu} = \mathcal{P}_{eu,i}\mathcal{P}_{eu,o}$ is an inner-outer factorization of $\mathcal{P}_{eu}$ if $\mathcal{P}_{eu,i}$ and $\mathcal{P}_{eu,o}$ are asymptotically stable, $\mathcal{P}_{eu,i}^* \mathcal{P}_{eu,i} = I$ and $\mathcal{P}_{eu,o}$ has a stable right inverse. Likewise, $\mathcal{P}_{yw} = \mathcal{P}_{yw,o}\mathcal{P}_{yw,i}$ is an outer-inner factorization of $\mathcal{P}_{yw}$ if $\mathcal{P}_{yw,o}$ and $\mathcal{P}_{yw,i}$ are asymptotically stable, $\mathcal{P}_{yw,i}\mathcal{P}_{yw,i}^* = I$ and $\mathcal{P}_{eu,o}$ has a stable left inverse.

As demonstrated in Vidyasagar (1988), the inner-outer and outer-inner factorizations of $\mathcal{P}_{eu}(z)$ and $\mathcal{P}_{yw}(z)$ can be expressed in terms of the solution of a Riccati equation. This implies that the approach of computing the optimal feedback controller via Youla parametrization and Wiener filtering, like the $\mathcal{H}_2$-optimal control approach in Theorem 3.2, generally involves the need to solve two Riccati equations. In fact it can be shown, that both approaches to compute the optimal feedback controller are completely equivalent (Fraanje 2004). Due to the special structure of the atmospheric disturbance model (3.4) and the input weighting (??) however, it is possible to directly give a valid outer-inner factorization of $\mathcal{P}_{yw}(z)$. The state-space realization of the open-loop transfer function $\mathcal{P}_{yw}(z)$ is given by

$$\mathcal{P}_{yw}(z) = \left[\left(\Sigma_1 C_d(zI - A_d)^{-1}K_d + I\right)R_v^{1/2} \quad 0\right] \tag{3.31}$$

Since the atmospheric disturbance model is in innovation form and $R_v^{1/2}$ is only a non-singular weighting matrix, the open-loop transfer function $\mathcal{P}_{yw}$ is minimum-phase. This implies that the outer factor can be chosen equal to the transfer function itself, which gives rise to the outer-inner factorization $\mathcal{P}_{yw} = \mathcal{P}_{yw,o}\mathcal{P}_{yw,i}$ with $\mathcal{P}_{yw,o} = \mathcal{P}_{yw}$ and $\mathcal{P}_{yw,i} = I$. As a result only the inner-outer factorization of $\mathcal{P}_{eu}$ has to be computed and at most one Riccati equation has to be solved in determining the optimal AO controller.

Like the open-loop transfer function $\mathcal{P}_{yw}(z)$, the following state-space realization of the open-loop transfer function $\mathcal{P}_{ew}(z)$ can be immediately read off from the atmospheric disturbance model (3.4) and the derived input weighting (**??**)

$$\mathcal{P}_{ew}(z) = \begin{bmatrix} C_d(zI - A_d)^{-1}K_d R_v^{1/2} + \Xi_1 & \Xi_2 \\ 0 & 0 \end{bmatrix}. \tag{3.32}$$

Furthermore, it follows from the particular choice of the performance output $e(k)$ that the transfer function $\mathcal{P}_{ew}(z)$ is equal to $\mathcal{P}_{eu}(z) = [\mathcal{H}^T(z)\ Q^{T/2}]^T$, where the DM model $\mathcal{H}(z)$ has a state-space realization (3.3). The above state-space realizations of $\mathcal{P}_{ew}$, $\mathcal{P}_{eu}$ and the outer-inner factorization of $\mathcal{P}_{yw}$ can be used to compute the optimal controller $\mathcal{C}(z)$.

Instead of considering the general state-space description (3.3) of the DM mirror, it is interesting to focus on the specific case where the DM can be considered as approximately static. Since the bandwidth of the DM is usually much larger than the control bandwidth (Le Roux et al. 2004), this is a reasonable assumption for a large class of AO systems. When the DM settling time can be neglected with respect to the WFS exposure-time, the dynamics of transfer function from $u(k)$ to $y(k)$ are completely determined by the zero-order hold (ZOH) nature of the digital-to-analog conversion, the integrating action related to the finite exposure-time of the WFS and the time delay caused by data acquisition and processing. With the WFS exposure-time being smaller than the sampling time T, this implies that the DM, including the WFS dynamics $g(z)$, can be described by means of the following finite impulse response (FIR) model (see for example Looze et al. 2003)

$$\mathcal{H}(z) = z^{-d}\left(H + \alpha z^{-1}H\right), \tag{3.33}$$

where $\alpha \in \mathbb{R}$, $d \in \mathbb{N}$ is an integer number of samples delay and $H \in \mathbb{R}^{m_p \times m_u}$ denotes the DM input-to-phase influence matrix. The number of samples delay should satisfy the condition $d \geq 1$ in order to account for the unavoidable one sample delay between measurement and correction. This delay is also necessary to avoid an algebraic loop.

Considering the above model structure for the DM model, it is possible to derive an analytical expression for the optimal feedback controller $\mathcal{C}(z)$. To this end it is useful to note that since the generalized plant $\mathcal{P}(z)$ is time-invariant, each of the channels of the performance output $e(k)$ can be delayed by an integer number of samples without modifying the cost function. By introducing a d-samples delay in the channels that correspond to the control effort weighting, i.e. $\mathcal{P}_{eu}(z) = [\mathcal{H}^T(z)\ \ z^{-d}Q^{T/2}]^T$, all channels of the transfer function $\mathcal{P}_{eu}(z)$ are delayed by the same amount and it is possible to factor out the delay z^{-d}. As can

be easily verified, this gives rise to a valid inner-outer factorization of $\mathcal{P}_{eu}(z)$ for $|\alpha| \leq 1$. A similar strategy can be followed for $|\alpha| > 1$. Since the value of the cost function is not influenced by an all-pass filter, the all-pass filter that contains the non minimum-phase zero of $\mathcal{H}(z)$ can be factored out. As a result the modified transfer function $\mathcal{P}_{eu}(z)$ has the following factorization

$$
\mathcal{P}_{eu}(z) = \begin{cases} \dfrac{1}{z^d}I & \times \begin{bmatrix} -(1+\alpha z^{-1})H \\ Q^{1/2} \end{bmatrix} \text{ for } |\alpha| \leq 1 \\[4mm] \dfrac{(z+\alpha)}{z^d(\alpha z+1)}I & \times \begin{bmatrix} -(\alpha + z^{-1})H \\ Q^{1/2} \end{bmatrix} \text{ for } |\alpha| > 1 \end{cases}, \tag{3.34}
$$

The derived expressions for the inner-outer and outer-inner factorizations of $\mathcal{P}_{eu}(z)$ and $\mathcal{P}_{yw}(z)$, can be used to compute the optimal feedforward and feedback controller via Lemma 3.1 and the Youla parametrization (3.28). This leads to the following theorem, which provides an analytical expression for the controllers $\mathcal{W}(z)$ and $\mathcal{C}(z)$ respectively.

Theorem 3.3 (Optimal control with quasi-static DM) *Let the wavefront distortions $\varphi(k)$ and open-loop WFS signal $y(k)$ be characterized by the regular stochastic process* (3.4) *with input covariance matrix $R_v > 0$. Furthermore, assume that the DM (including the WFS dynamics) can be modeled as* (3.33) *and that either H or Q has full column rank (i.e. $[H^T\ Q^{T/2}]^T$ is left invertible). Then the optimal feedforward controller $\mathcal{W}(z)$, which minimizes* (3.5)*, is given by*

$$
\begin{bmatrix} \hat{\chi}_1(k+1) \\ \hat{\chi}_2(k+1) \\ u(k) \end{bmatrix} = \left[\begin{array}{cc|c} \widetilde{A} & 0 & K_d \\ F\widetilde{A} & -G & FK_d \\ \hline F\widetilde{A} & -G & FK_d \end{array} \right] \begin{bmatrix} \hat{\chi}_1(k) \\ \hat{\chi}_2(k) \\ y(k) \end{bmatrix}, \tag{3.35}
$$

where the matrices $\widetilde{A}$, F, G are defined as $\widetilde{A} \doteq A_d - K_d\Sigma_1 C_d$, $F \doteq H_Q^\dagger C_d A_d^{d-1}$ and $G \doteq \gamma_\alpha H_Q^\dagger H$, and where

$$
H_Q^\dagger \doteq \begin{cases} (H^T H + Q)^{-1} H^T & \text{for } \alpha = 0 \\ \alpha \gamma_\alpha \left(\alpha^2 H^T H + \gamma_\alpha^2 Q\right)^{-1} H^T & \text{otherwise} \end{cases}, \tag{3.36}
$$

with

$$
\gamma_\alpha \doteq \begin{cases} \alpha & \text{for } |\alpha| \leq 1 \\ sgn(\alpha) & \text{for } |\alpha| > 1 \end{cases},
$$

can be interpreted as a regularized left pseudo-inverse of the DM influence matrix H. Furthermore, let the matrices L and M be defined as $L \doteq z^{-d+1} K_d \Sigma_1 H (\alpha I - G)$ and $M \doteq \widetilde{A} + z^{-d+1} K_d \Sigma_1 H F$. Then the corresponding optimal feedback controller $\mathcal{C}(z)$ has a state-space representation

$$
\begin{bmatrix} \hat{\xi}_1(k+1) \\ \hat{\xi}_2(k+1) \\ u(k) \end{bmatrix} = \left[\begin{array}{cc|c} -G & F & 0 \\ L & M & K_d \\ \hline FL+G^2 & FM-GF & FK_d \end{array} \right] \begin{bmatrix} \hat{\xi}_1(k) \\ \hat{\xi}_2(k) \\ r(k) \end{bmatrix}. \tag{3.37}
$$

Theorem 3.3 provides an analytical solution to the AO optimal control problem in the case that the DM model and the WFS dynamics can be described by equation (3.33). As already pointed out, this assumption on the DM model structure is quite general in the sense that it holds for a large class of AO systems. Given the matrix Σ_1 composed of the nonzero singular values of geometry matrix G and the DM influence matrix H, Theorem 3.3 can be used to compute the optimal controller from the system matrices $\widetilde{A}$, $\widetilde{C}$ and $\widetilde{K}$. By combining the subspace identification algorithm presented in Section 3.3 and Theorem 3.3, this gives rise to a direct and non-iterative way to go from open-loop measurement data to closed-loop controller design. The resulting closed-loop controller design procedure is entirely based on standard matrix operations. Furthermore, it is interesting to note that due to the presence of the $d-1$ samples delay in the state-update equations for ξ_2, the optimal feedforward and feedback controllers in Theorem 3.3 are effectively of order $n_d + m_u(d-1)$. A well-known property of $\mathcal{H}_2$-optimal control design is it that leads to a controller of the same dimension as the generalized plant (see also Theorem 3.2). From the DM model (3.33) and the state-space representation of the generalized plant, one would therefore expect the controller to be of order $n_d + m_u d$. This discrepancy in model order is caused by the special structure of the generalized plant. Due to this structure, the optimal controller computed by straightforward application of $\mathcal{H}_2$-optimal control theory is non-minimal. By removing the unobservable modes it is possible to arrive at a controller of reduced order. An additional advantage of the analytical solution is that the model reduction step has already been performed.

To obtain some more insight in the structure of the analytical expressions for the optimal controller, it is useful to consider the case where the combined DM and WFS dynamics consists of a pure delay. This is to say that the DM mirror model reduces to $\mathcal{H}(z) = z^{-d}H$, which is equivalent to choosing $\alpha = 0$ in equation (3.33). Physically this situation can be achieved by accurate synchronization of the DM digital-to-analog converters and the WFS exposure time. When the DM settling time is negligible and the digital-to-analog converters are synchronized in such a way that the ZOH output does not chance during the CCD exposure time, the only dynamics that are left is a pure delay. The expressions for the optimal feedforward and feedback controller obtained by considering this multi-sample delay case are summarized in Corollary 3.1. Note that the more restrictive case of a unit-sample delay has been elaborated in our previous work (Hinnen et al. 2005). In contrast to the derivation presented here however, the optimal controller was found by direct application of Theorem 3.2.

Corollary 3.1 (Multiple-sample delay) *Consider an AO system where the only dynamics exhibited by the WFS and DM is an integer number of samples delay, i.e. $\mathcal{H}(z) = z^{-d}H$, with $d \in \mathbb{N}$. Furthermore, assume that the conditions as stated in Theorem 3.3 are satisfied, then the feedforward controller $\mathcal{W}(z)$ which is (3.5) is given by*

$$\begin{bmatrix} \hat{\chi}_1(k+1) \\ \hline u(k) \end{bmatrix} = \begin{bmatrix} \widetilde{A} & K_d \\ \hline F\widetilde{A} & FK_d \end{bmatrix} \begin{bmatrix} \hat{\chi}_1(k) \\ \hline y(k) \end{bmatrix}. \tag{3.38}$$

Furthermore, the optimal feedback $\mathcal{C}(z)$ which internally stabilizes $\mathcal{P}(z)$ has a state-space

representation

$$
\begin{bmatrix} \hat{\xi}_1(k+1) \\ \hline u(k) \end{bmatrix} = \left[\begin{array}{c|c} \widetilde{A} + z^{-d+1}K_d\Sigma_1 HF & K_d \\ \hline F(\widetilde{A} + z^{-d+1}K_d\Sigma_1 HF) & FK_d \end{array} \right] \begin{bmatrix} \hat{\xi}_1(k) \\ \hline r(k) \end{bmatrix}, \tag{3.39}
$$

where the matrices $\widetilde{A}$ and F are defined as in Theorem 3.3, and $H_Q^\dagger$ reduces to the expression $H_Q^\dagger = (H^T H + Q)^{-1} H^T$.

Proof: The above result follows immediately from Theorem 3.3 by choosing $\alpha = 0$. Note that for $\alpha = 0$, $L = 0$ and $G = 0$. The latter equality implies that the state vectors $\hat{\chi}_2$ and $\hat{\chi}_1$ of the feedforward optimal controller (3.35) differ only by a multiplicative factor F, i.e. $\hat{\chi}_2 = F\hat{\chi}_1$. Using this relation to eliminate $\hat{\chi}_2$ from the state-space equations gives the desired expression for $\mathcal{W}(z)$. Furthermore, by substituting $G = 0$ and $L = 0$ in the state-space representation for the optimal feedback controller (3.37), it is clear that the state $\hat{\xi}_1$ is unobservable. Eliminating this state gives the desired expression for $\mathcal{C}(z)$. $\qquad\square$

The state-space equations (3.38) of the optimal feedforward controller $\mathcal{W}(z)$ allows for a nice physical interpretation of the optimal controller. To make this more clear, consider the atmospheric disturbance model (3.4) and let $\hat{x}(k|k-1)$ denote the conditional mean of $x(k)$ given the past open-loop WFS data $\{y(j),\ j \le k-1\}$. By using the output equation for $y(k)$ to eliminate the white noise input $v(k)$ from the state-update equation, the conditional mean of the state at the next sample instant can be expressed as $\hat{x}(k+1|k) = \widetilde{A}\hat{x}(k|k-1) + K_d y(k)$. From this it is clear that the state in equation (3.38) can be interpreted as the conditional mean $\hat{x}(k|k-1)$ of the state $x(k)$. In fact the state corresponds to the state in the Kalman predictor model. The unpredictability of the white noise input $v(k)$ causes the optimal prediction of future states to be obtained by iterating the state-update equation (3.4) with the future white noise set to its expected value zero (Goodwin et al. 2001). This gives rise to the state estimate $\hat{x}(k+d|k) = A^{d-1}(\widetilde{A}\hat{x}(k|k-1) + K_d y(k))$. By comparing this with the output-equation of the feedforward optimal controller $\mathcal{W}(z)$, it is clear that the control signal can be expressed as $u(k) = H_Q^\dagger C\hat{x}(k+d|k)$. With $\hat{x}(k+d|k)$ being the conditional mean of the state of the atmospheric disturbance model (3.4), $\hat{\varphi}(k+d|k) \doteq C\hat{x}(k+d|k)$ can be interpreted as the conditional mean of the open-loop wavefront distortion $\varphi(k+d)$ given the past open-loop WFS data $\{y(j),\ j \le k\}$. On the other hand, the matrix $H_Q^\dagger$ can be seen as a regularized version of the pseudo-inverse of the DM influence matrix H and provides a projection of the estimated phase $\hat{\varphi}(k+d|k)$ on the DM actuator space. From this it is clear that the optimal feedforward controller $\mathcal{W}(z)$ decomposes in a multiple step ahead predictor, which is concerned with estimating the uncorrected wavefront $\varphi(k+d)$, and a static matrix projection. This can be well understood in the context of the feedforward control problem in Figure 3.7. Furthermore, since the feedforward and feedback control problem are equivalent through the Youla parametrization, the interpretation of the structure of the feedforward controller

$\mathcal{W}(z)$ can be extended to the feedback case. For this reason, the state in the state-space representation of the optimal feedback controller $\mathcal{C}(z)$ can be seen as the conditional mean $\hat{x}(k|k-1)$ given the closed-loop WFS measurements $r(j), j \leq k-1$.

The above interpretation of the optimal feedforward and feedback controller in Corollary 3.1 can be used to obtain some more insight in the structure of the optimal controller in the case that the DM model and WFS dynamics are described by (3.33). To this end note that the first state-update equation in (3.35) and the state-update equation in (3.38) are equal. This implies that the state $\hat{\chi}_1(k)$ can still be interpreted as the conditional mean $\hat{x}(k|k-1)$ of the state $x(k)$. Furthermore, since the second state-update equation is equal to the output equation, it is clear that the state $\hat{\chi}_2(k+1)$ is equal to the control signal $u(k)$. Using the above interpretations, the output $u(k)$ of the controller in Theorem 3.3 can be expressed as

$$u(k) = H_Q^\dagger \hat{\varphi}(k+d|k) - \gamma_a H_Q^\dagger H u(k-1) \tag{3.40}$$

$$= (I - \gamma_a H_Q^\dagger H z^{-1})^{-1} H_Q^\dagger \hat{\varphi}(k+d|k). \tag{3.41}$$

This shows that the optimal controller still consists of a part that is concerned with estimating the uncorrected wavefront $\hat{\varphi}(k+d|k)$ but that the static projection $H_Q^\dagger$ has been replaced by a dynamic filter. The dynamic filter makes a trade off between the problem of inverting the undelayed part of the DM model $(1 + \alpha z^{-1})H$ and the problem of minimizing the contribution of the control effort to the cost function. This can be easily seen for the case that $|\alpha| < 1$ and the control effort weighting Q is set to zero. For this specific case $\gamma_\alpha H_Q^\dagger H = \alpha I$, so that dynamic filter reduces to $(1 + \alpha z^{-1})^{-1} H^\dagger$, which is precisely the inverse of $(1 + \alpha z^{-1})H$. Due to the equivalence between the feedforward and feedback control problem, also the output of the feedback controller $\mathcal{C}(z)$ can be expressed as in equation (3.40), where the conditional mean $\hat{\varphi}(k+d|k)$ is now determined on the basis of the closed-loop WFS measurements $r(j), j \leq k-1$ instead of the open-loop data.

3.7 Numerical validation optimal control strategy

In this section, the closed-loop control design approach obtained by combining the subspace identification algorithm presented in Section 3.3 and the analytical solution to the $\mathcal{H}_2$-optimal control problem presented in Section 3.6, is demonstrated by means of a simulation example. In this example, the performance of the proposed control strategy is compared with a control law commonly in AO systems. Before discussing the simulation experiments, the common AO control approach will be briefly summarized. For conformity with the rest of the paper, the common AO control approach will be reviewed in terms of the reduced signals introduced in Section 3.2.

3.7.1 AO control law used for performance comparison

The common AO control approach consists of a cascade of a static matrix multiplication and a series of feedback loops (Roddier 1999). Given a new WFS mea-

surement $y(k)$, the static part is concerned with the problem of finding the DM actuator inputs $\delta u(k)$ that would provide the best fit to the wavefront. In this step the temporal dynamics of both the wavefront disturbance and the AO system are neglected. The static relation between the actuator input and the WFS sensor measurement is given by $\delta u(k) = Ry(k)$ and the DM is modeled as $\varphi_m(k) = H\delta u(k)$, where H is the input-to-phase influence matrix. Furthermore, the WFS model is described by the static WFS model obtained by setting $\mathcal{G}(z) = G$ in (3.1). Considering the above models, the problem of finding the matrix R that provides the best fit is formulated as the following optimization problem

$$R = \arg\min_R \mathcal{E}\{\|\varphi(k) - HRy(k)\|^2\}, \tag{3.42}$$

where $\|\cdot\|$ denotes the $\mathcal{L}_2$-norm and the $\mathcal{E}$ is the conditional expectation given the WFS measurement $y(k)$. Let the covariance matrices of the measured wavefront $\varphi(k)$ and the measurement noise $\nu(k)$ be defined as $C_\varphi \doteq \mathcal{E}\{\varphi(k)\varphi^T(k)\}$ and $C_\nu \doteq \mathcal{E}\{\nu(k)\nu^T(k)\}$ respectively. Then, under the assumption that $\nu(k)$ and $\varphi(k)$ are uncorrelated and $C_\nu = \sigma_\nu^2 I$, the maximum a posteriori estimate of the matrix R is given by

$$R = (H^T H)^{-1} H^T (\Sigma_1^2 + \sigma_\nu^2 C_\varphi^{-1})^{-1} \Sigma_1,$$

where σ_ν^2 denotes the variance of the measurement noise. The signal $\delta u(k)$ computed from the reconstruction process cannot be directly used as the control signal. Since the AO system is operated in closed-loop, the WFS measures the residual wavefront $\varepsilon(k) = \varphi(k) - \varphi_m(k)$ instead of the uncorrected wavefront $\varphi(k)$. This implies that the signal obtained from the static reconstruction $\delta u(k) = Rr(k)$ provides only an estimate of the correction that has to be applied to current actuator commands. The parallel feedback loops are responsible for stability and closed-loop performance and have to posses integrating action to overcome this shortcoming. In this paper, the following control law has been used for performance comparison

$$u(k) = \frac{c_1}{1 - c_2 z^{-1}} \delta u(k) = \frac{c_1}{1 - c_2 z^{-1}} Rr(k), \tag{3.43}$$

where the integrator gain $c_1 \in \mathbb{R}$ and the loss factor $c_2 \in \mathbb{R}$ and are user defined control parameters. The change in the wavefront covariance matrix C_φ due to closed-loop operation is neglected in the common AO control approach.

3.7.2 Simulation procedure and performance measures

The data-driven optimal control approach obtained by combining the subspace identification algorithm and the analytical expressions for the optimal controller has been validated using the same open-loop WFS data set from the AO laboratory setup, as described in Section 3.4. Just as before, the total data set consisting of $N_t = 1.0 \times 10^4$ is divided in a $N_i = 8000$ samples identification set and data set consisting of $N_v = 2000$ samples reserved for performance evaluation. Furthermore, the number of past p and f future block-rows used in the identification

of the atmospheric disturbance model is again 15 and the model order is chosen equal to $n_d = 300$. This implies that the closed-loop simulation experiments can be directly compared to the results presented in Section 3.4. To facilitate this comparison, the performance of the closed-loop controllers will be evaluated in terms of the performance criteria as introduced in Section 3.4.1. More specifically, the performance of the controller will be evaluated by simulating the closed-loop residual wavefront error using the open-loop wavefront distortions reconstructed from the validation set as disturbance input. The residual wavefront error is then used to the normalized mean-square residual phase error J_1 and normalized averaged power spectrum $P(\omega)$ in accordance with equations (3.25) and (3.26), where $\varepsilon(k)$ now denotes the residual wavefront error.

The combined subspace identification and $\mathcal{H}_2$-optimal control approach has been validated for the quasi-static case where the DM mirror and WFS can be considered to be static apart from an unit-sample delay, i.e. $\mathcal{H}(z) = z^{-1}H$. Two different simulation scenarios have been elaborated. The first scenario consists of closed-loop simulations with an ideal DM. Here, the term ideal is used to refer to a hypothetical DM that is able to take the shape of the estimated wavefront without introducing a fitting error. In terms of the considered DM model structure, this means that the DM influence matrix H is assumed to have full row rank. The ideal DM should therefore have at least as many actuators as the number of independent WFS channels in the system. Even though this condition is hardly ever satisfied for realistic systems, it is still interesting to consider the ideal DM as it provides a better insight in error sources other than the DM fitting error which itself is not influenced by the controller. Note that from the physical interpretation of the controller structure in the previous section it is clear that in this case the residual phase error is equal to the wavefront prediction error, i.e. $\varepsilon(k) = \varphi(k) - \hat{\varphi}(k|k-1)$. The second simulation scenario considers a more realistic DM model. In this case the DM model is obtained by identifying the influence matrix H of the mirror used in the AO test bench. The AO test bench is equipped with a 37-actuator electrostatic membrane mirror provided by OKO technologies in the Netherlands. The mirror is operated around an offset and is almost completely linear with the applied voltage squared. The influence matrix H has been estimated from a least squares fit on the reconstructed wavefront data obtained by measuring the steady state WFS response y_m to a set of predefined inputs u.

In the simulation experiments, the performance of the proposed control design strategy has been compared with the common AO control law in (3.43). The control parameters c_1 and c_2 have been tuned so as to minimize the cost function over the identification data set, which resulted in the values $c_1 = 1.31$ and $c_2 = 0.997$. The covariance matrix C_φ has been computed by assuming a perfect Kolmogorov spatial distribution with a Fried parameter r_0 satisfying the specification of the turbulence simulator. The variance of the measurement noise σ_ν^2 has been estimated by computing the variance of the WFS measurements $y(k)$ for a static wavefront distortion $\varphi(k)$ generated by the turbulence simulator. The control-effort weighting in cost function (3.5) is neglected by choosing $Q = 0$.

3.7.3 Closed-loop simulation results

Figures 3.8 and 3.9 show the normalized averaged power spectra $P(\omega)$ of the open-loop wavefront distortions (dashed-dotted line), of the residual wavefront error using the common AO control approach (solid line) and of the residual wavefront error using the data-driven optimal control approach (dashed line), for respectively the ideal and the estimated DM influence matrix. For the ideal DM, the normalized averaged residual power spectrum corresponding to proposed control approach is approximately white. This means that, at least on average, there is no temporal correlation in the residue that can be used to further improve the performance of the controller. The residue obtained with the common AO control law on the other hand has a strong coloring and shows that there is still plenty room for improvement. From the power spectra it is already clear that the residual phase error obtained with the optimal control approach is much smaller than with the common AO control approach. This is confirmed by the normalized performance index J_1. The normalized reductions obtained in the simulations with the ideal DM are $J_1 = 3.27 \times 10^{-3}$ for the common control approach and $J_1 = 1.52 \times 10^{-3}$ for the optimal control approach. This is a reduction of 53.5%. The corresponding values for the simulations with estimated influence matrix of the AO test bench mirror are $J_1 = 3.13 \times 10^{-2}$ and $J_1 = 1.92 \times 10^{-2}$, which corresponds to a reduction of 38.6%.
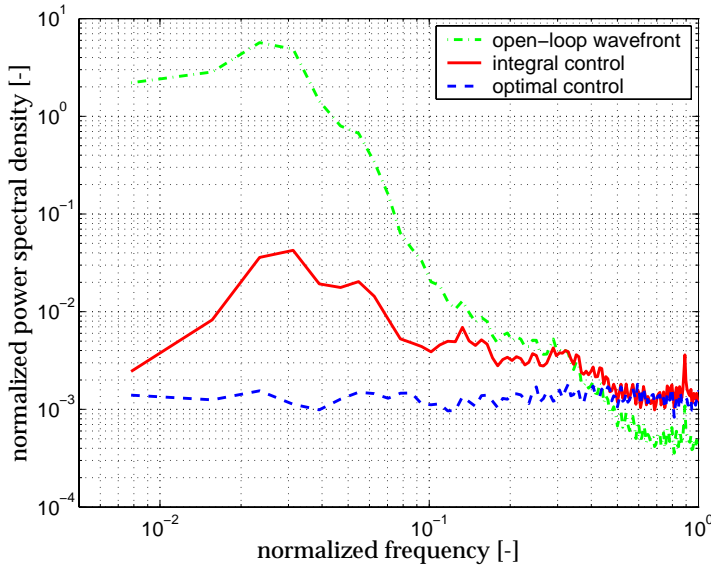

Figure 3.8: Normalized averaged power spectrum $P(\omega)$ of residual wavefront error $\varepsilon(k)$ for closed-loop simulations with an ideal DM.

The simulations show that the performance improvement for the simulation with the estimated DM influence matrix is considerably smaller. Since both simulations differ only in the DM influence matrix, it is clear that in the second scenario
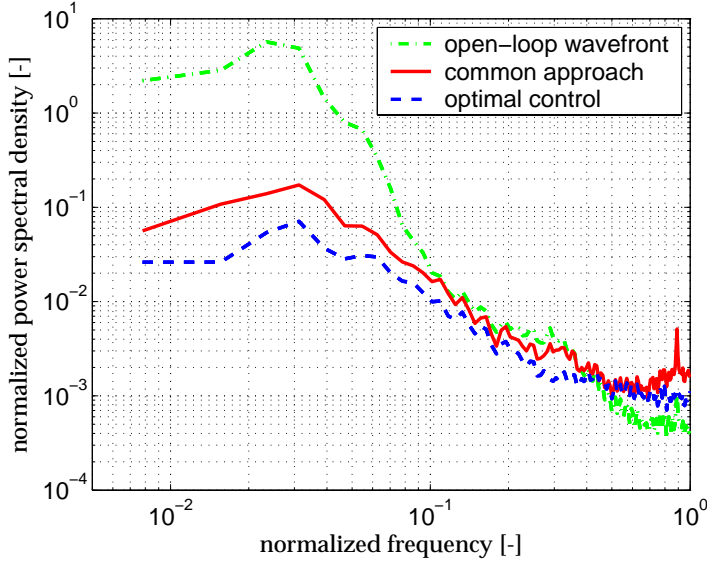
Figure 3.9: Normalized averaged power spectrum $P(\omega)$ of residual wavefront error $\varepsilon(k)$ for closed-loop simulations with realistic DM model.

the DM fitting error is the limiting factor. As the DM fitting error is not influenced by the controller, it is to be expected that an AO system can only benefit from an advanced control strategy if the DM fitting error is small compared to other error sources. The simulations with the ideal DM show that the proposed control strategy is able to significantly reduce the contribution due to the temporal error. By predicting the future wavefront distortions as $\hat{\varphi}(k + 1|k)$, the temporal error caused by the finite time delay between measurement and corrections is reduced. In the AO test bench, the ratio between the fitting and temporal error is very unfavorable because of the relatively small number of DM actuators compared to the number of WFS channels. It is therefore to be expected that the performance gain is larger in AO systems with a smaller fitting error and in situations where the controller related error is more dominant for instance due to an increased turbulence wind speed or higher levels of measurement noise.

Finally note that the normalized mean-square residual phase error J_1 obtained for optimal controller and the ideal DM is precisely the same as the normalized prediction error on the validation set in Section 3.4. This is no coincidence. From the discussion in Section 3.6 it should be clear that the closed-loop optimal control problem for a quasi-static DM mirror model, of which the only dynamics is an unit-sample delay, is precisely equivalent to the one-step ahead prediction problem in open loop. Also the performance of the random walk predictor and the common AO control approach are very close to one another. The slightly better performance of the common AO control law can be explained by the use of the minimum-variance wavefront reconstructor and the first-order lag filter in the common AO control approach. The closed-loop equivalent of the random walk

predictor would a least-squares reconstructor followed by a pure integrator.

3.8 Conclusions

In this paper we have presented a data-driven approach to design a controller for rejecting the wavefront distortions in an AO system. The proposed control design strategy is able to take full advantage of the spatio-temporal correlation in the wavefront distortion and consists of two major steps. In the first step open-loop wavefront sensor (WFS) measurements are used to identify a multi-variable atmospheric disturbance model. In the second step the identified atmospheric disturbance model is used to compute the optimal controller.

To identify the multi-variable atmospheric disturbance model from the open-loop WFS data, a dedicated subspace identification algorithm has been developed. An important advantage of the proposed subspace algorithm is that it avoids the need for spectral factorization by directly estimating the system matrices of the Kalman predictor model corresponding to minimum-phase spectral factor. Since AO systems typically have a large number of WFS channels and the atmospheric disturbance model should describe the full spatio-temporal correlation without assuming any form of decoupling, computational efficiency is an important issue. For this reason special attention has been paid to reduce the computational demands of the algorithm. The different steps of the algorithm are expressed in terms of the R factor of a single RQ factorization of the stacked block Hankel matrices of past and future data, which is used for data compression. This leads to an efficient implementation both in terms of the number of flops and required memory storage. The proposed subspace identification algorithm can be used to identify an atmospheric disturbance model for small to medium sized AO systems and has been demonstrated by means simulation examples on open-loop WFS data. In these simulations data sets from an AO laboratory setup as well as measurements from a real telescope have been considered. For each of the considered data sets, the proposed subspace identification algorithm is able to identify an accurate atmospheric disturbance model that can be used to reduce the one-step ahead prediction error over the random walk approach. Furthermore, the simulations on real telescope data confirm that the atmosphere can be reasonable well modeled as a regular stochastic process on a time scale of a few minutes.

An important aspect of the AO control problem is that there is a difference between the objective of minimizing the mean-square residual WFS signal and the actual cost function. Given the identified atmospheric disturbance model, the $\mathcal{H}_2$-optimal control framework provides an attractive way to deal with this discrepancy. Formulating the AO problem as a $\mathcal{H}_2$-optimal control problem provides a general strategy for computing the optimal controller. Computing the $\mathcal{H}_2$-optimal controller typically involves the solution to two Riccati equations. By using the Youla parametrization to render the AO control problem into an equivalent feedforward problem, it has been shown that due to the minimum-phase property of the atmospheric disturbance model one of the Riccati equations can be avoided. Furthermore, it has been shown that in the special case that the DM settling time

can be neglected with respect to the WFS exposure time an analytical expression for the optimal controller can be derived. The analytical expressions have been used to interpret the optimal controller as a wavefront estimation problem followed by a regularized inversion of the DM model.

The closed-loop control design procedure obtained by combining the proposed subspace identification algorithm and the analytical solutions to the $\mathcal{H}_2$-optimal control problem, is entirely based on standard matrix operations and provides a non-iterative way to go from open-loop measurement data to closed-loop controller design. The proposed control strategy has been demonstrated by means of numerical validation experiments on open-loop WFS data obtained from an experimental setup. The validation experiments show a performance improvement with respect to the common AO control approach. Under the assumption that the DM is able to take the shape of the estimated wavefront, the use of the proposed control strategy leads to a reduction of the mean-square residual phase error by more than 70%. Using a realistic DM model, the gain in performance for considered experimental setup reduces to about 14%. The rather drastic reduction in performance improvement can be explained by the relatively small number of DM actuators, which give rise to a large fitting error. In situations where the DM fitting error is not limiting, a large gain in performance is to be expected.

References

Arnold, W. F., Laub, A. J., Dec. 1984. Generalized eigenproblem algorithms and software for algebraic Riccati equations. Proceedings of the IEEE 72 (12), 1746–1754.

Bauer, D., Aug. 2004. Choosing integer parameters in subspace identification: A survey on asymptotic results. In: Proceedings of the IFAC Symposium on System Identification. Rotterdam.

Born, M., Wolf, E., 1999. Principles of Optics: Electromagnetic Theory of Propagation, seventh (expanded) edition Edition. Cambridge University Press.

Chen, T., Francis, B., 1995. Optimal Sampled-Data Control Systems. Springer-Verslag, London, Berlin.

Dessenne, C., Madec, P.-Y., Rousset, G., Oct. 1997. Modal prediction for closed-loop adaptive optics. Optics Letters 22 (20), 1535–1537.

Ellerbroek, B. L., Van Loan, C., Pitsianis, N. P., Plemmons, R. J., Nov. 1994. Optimizing closed-loop adaptive-optics performance with use of multiple control bandwidths. Journal of the Optical Society of America A 11 (11), 2871–2886.

Fraanje, P. R., May 2004. Robust and fast schemes in broadband active noise and vibration control. Ph.D. thesis, University of Twente, the Netherlands.

Gavel, D. T., Wiberg, D., Feb. 2003. Towards strehl-optimizing adaptive optics controllers. In: Wizinowich, P. L., Bonaccini, D. (Eds.), Adaptive Optical System Technologies II. Vol. 4839 of Proceedings of SPIE. Hawaii, pp. 890–901.

Gendron, E., Léna, P., 1994. Astronomical adaptive optics. Astronomy and Astrophysics: I. Modal control optimization 291, 337–347.

Goodwin, G. C., Graebe, S. F., Salgado, M. E., 2001. Control System Design. Prentice Hall, Upper Saddle River, New Jersey.

Hardy, J. W., 1998. Adaptive Optics for Astronomical Telescopes. Oxford series in optical and imaging sciences. Oxford University Press, New York.

Hinnen, K. J. G., Doelman, N. J., Verhaegen, M., Aug. 2005. $\mathcal{H}_2$-optimal control of an adaptive optics system: Part II, closed-loop controller design. In: Tyson, R. K., Lloyd-Hart, M. (Eds.), Astronomical Adaptive Optics Systems and Applications II. Vol. 5903 of Proceedings of SPIE. San Diego, California, United States, pp. 86–99.

Jansson, M., Aug. 2003. Subspace identification and ARX modeling. In: 13th IFAC Symposium on System Identification. Rotterdam, the Netherlands, pp. 1625–1630.

Kailath, T., Sayed, A. H., Hassibi, B., 2000. Linear Estimation. Prentice Hall Information ans Systems Sciences Series, Upper Saddle River, New Jersey.

Law, N. F., Lane, R. G., May 1996. Wavefront estimation at low light levels. Optics Communications 126, 19–24.

Le Roux, B., Conan, J.-M., Kulcsár, C., Raynaud, H.-F., Mugnier, L. M., Fusco, T., Jul. 2004. Optimal control law for classical and multi-conjugate adaptive optics. Journal of the Optical Society of America. A 21 (7), 1261–1276.

Lindquist, A., Picci, G., May 1996. Canonical correlation analysis, approximate covariance extension and identification of stationary time series. Automatica 32 (5), 709–733.

Ljung, L., 1999. System Identification: Theory for the User, 2nd Edition. Prentice Hall PRT, Upper Saddle River, NJ.

Ljung, L., Glad, T., 1994. Modeling of Dynamic Systems. Prentice Hall PRT, Englewood Cliffs.

Looze, D. P., Becker, O., Kasper, M., Hippler, S., Dec. 1999. Optimal compensation and implementation for adaptive optics systems. In: Proceedings of the 38th Conference of Decision & Control. Phoenix, Arizona, United States, pp. 1715–1720.

Looze, D. P., m. Kasper, Hippler, S., ans R. Weiss, O. B., 2003. Optimal compensation and implementation for adaptive optics systems. Experimental Astronomy 15 (2), 67–88.

Mari, J., Stoica, P., McKelvey, T., Jul. 2000. Vector ARMA estimation: A reliable subspace approach. IEEE Transactions on Signal Processing 48 (7), 2092–2104.

Mastronardi, N., Kressner, D., Sima, V., Dooren, P. V., Huffel, S. V., Jul. 2001. A fast algorithm for subspace state-space system identification via exploitation of the displacement structure. Journal of Computational and Applied Mathematics 132 (1), 71–81.

Morari, M., Zafiriou, 1989. Robust Process Control. Prentice-Hall, Englewood Cliffs, New Jersy, United States.

Paschall, R. N., Anderson, D. J., Nov. 1993. Linear quadratic gaussian control of a deformable mirror adaptive optics system with time-delayed measurements. Applied Optics 32 (31), 6347–6358.

Peternell, K., Scherrer, W., Deistler, M., 1996. Statistical analysis of novel subspace identification methods. Signal Processing 52, 161–177.

Roddier, F., Jul. 1998. Maximum gain and efficiency of adaptive optics systems. Publications of the Astronomical Society of the Pacific 110, 837–840.

Roddier, F., 1999. Adaptive Optics in Astronomy. Cambridge University Press.

Saint-Jacques, D., Dec. 1998. Astronomical seeing in space and time. Ph.D. thesis, University of Cambridge.

Shu, H., Chen, T., Dec. 1995. State-space approach to discrete-time $\mathcal{H}_2$ optimal control with a causality constraint. In: Proceedings of the 34th Conference on Decision and Control. New Orleans, LA, pp. 1927–1932.

Tyson, R. K., 1998. Principles of Adaptive Optics, 2nd Edition. Academic Press.

van Dam, M. A., Le Mignant, D., Macintosh, B. A., Oct. 2004. Performance of Keck observatory adaptive optics system. Applied Optics 43 (29), 5458–5467.

Van Overschee, P., De Moor, B., May 1993. Subspace algorithms for the stochastic identification problem. Automatica 29 (3), 648–660.

Verhaegen, M., Jan. 1994. Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. Automatica 30 (1), 61–74.

Vidyasagar, M., Sep. 1988. Control systems synthesis: a factorization approach. MIT Press, Cambridge, Massachusetts, United States.

Wild, W. J., Sep. 1996. Predictive optimal estimators for adaptive-optics systems. Optics Letters 21 (18), 1433–1435.

Wild, W. J., Apr. 1998. Lyapunov stability criteria for zonal adaptive-optics systems. Optics Letters 23 (8), 570–572.

Youla, D. C., Bongiorno, J. J., Jabr, H. A., Jun. 1976. Modern wiener-hopf design of optimal controllers–part II: The multivariable case. IEEE Transactions on Automatica Control 21 (3), 319–338.

Zhou, K., Doyle, J. C., Glover, K., 1996. Robust and Optimal Control. Prentice Hall, Upper Saddle River, New Jersy.

Experimental results: Exploiting the Spatio-Temporal Correlation

In this paper, a recently proposed data-driven $\mathcal{H}_2$-optimal control approach is demonstrated on an laboratory setup. Most adaptive optics (AO) systems are based on a control law that neglects the temporal evolution of the wavefront. The proposed control approach, on the other hand, is able to exploit the spatio-temporal correlation in the wavefront without assuming any form of decoupling. It is shown that for a static wavefront correction device, the necessary conditions for having an analytical expression for the optimal controller are satisfied. In this way, the standard and cumbersome problem of solving two Riccati equation is bypassed. The performance of the optimal control approach is compared to the standard common method. A detailed analysis of the dominant error sources shows that optimal control may lead to a significant reduction in the temporal error. Since the temporal error increases with the Greenwood to sample frequency ratio, the performance gain is especially large at large ratios.

Keywords: Adaptive optics; Data-driven disturbance modeling; Stochastic identification; Optimal control; Experimental validation

4.1 Introduction

Adaptive optics (Hardy 1998; Roddier 1999) is a well established technique for real-time compensation of the optical wavefront distortions introduced by a turbulent medium. It has found widespread application in ground-based astronomical imaging, where it is used to counteract the devastating effect of atmospheric turbulence on the angular resolution. In this paper we concentrate on the control aspects of adaptive optics (AO). Most AO systems are based on a control strategy

that is not able to exploit the spatio-temporal correlation in the wavefront. Usually, the control law (see e.g. van Dam et al. 2004) consists of a cascade of a static part, concerned with the problem of finding the actuator inputs that provide the best fit to the wavefront, and a series of parallel feedback loops responsible for stability and closed-loop performance. In the simplest case, the static wavefront reconstruction and fitting step is formulated as a matrix inversion problem. To improve the accuracy, both maximum likelihood and maximum *a posteriori* techniques have been used to include prior knowledge on the spatial correlation of the wavefront (van Dam et al. 2004; Law and Lane 1996). Prior knowledge on the temporal evolution of the wavefront is usually not included in the control design. Each of the parallel feedback loops typically consist of a first-order lag filter or proportional-integral (PI) controller, of which the parameters are tuned to make a trade off between disturbance rejection, noise propagation and closed-loop stability. To relax the trade off between these conflicting requirements, modal control optimization has been proposed (Ellerbroek et al. 1994; Gendron and Léna 1994). In this approach the wavefront is decomposed in a set of modes of which the corresponding servo gains are optimized.

The separation of the control law into static wavefront reconstruction and temporal compensation is based on the assumption that the spatial and temporal dynamics can be decoupled. The Taylor hypothesis (Taylor 1938; Gendron and Léna 1996), which states that the atmospheric turbulence evolves at a time scale that is long compared to the time it takes for the wind-blown inhomogeneities to cross the line of sight, shows that this is typically not the case. When the turbulence can be considered as a frozen layer, there exists a strong correlation between the spatial and temporal dynamics which may be used to the benefit of the controller. By including a priori knowledge on the spatio-temporal correlation, wavefront sensor (WFS) measurements from the past and neighboring channels may be used to predict future wavefront distortions. In this way, it should be possible to reduce the temporal error due to delayed system response. Also, the sensitivity to measurement noise may be reduced. As such, advanced control may help to improve the performance, either in terms of the ability to suppress wavefront distortions, or in terms of the limiting magnitude of the required guide star.

Since the above control laws are composed of a static wavefront reconstruction step and a series of independent servo loops, they are based on the implicit assumption that spatial and temporal dynamics can be fully decoupled. The Taylor hypothesis (Taylor 1938), which states that the atmospheric turbulence evolves at a time scale that is long compared to the time it takes for the wind-blown inhomogeneities to cross the line of sight, clearly shows that this is typically not the case. When the atmospheric turbulence can be considered as a frozen layer moving across the telescope aperture, there exists a strong correlation between the spatial and temporal dynamics of the wavefront, which may be used at the benefit of the controller (Gendron and Léna 1996). By including a priori knowledge on the spatio-temporal correlation of the wavefront, WFS measurements from the past and neighboring channels may be used to anticipate future wavefront distortions. In this way, it should be possible to reduce the effect of the delayed response associated with the temporal error. Also the sensitivity to measurement noise may be

reduced as photons collected at different time instants and WFS channels may all used to improve the wavefront estimate at a certain position in the aperture plane. As a result, the performance of an AO system may benefit from a control strategy that is able to account for the spatio-temporal correlation in the wavefront. Such a control strategy may lead to an improved performance, either in terms of the ability to suppress the incoming wavefront distortions, or in terms of the limiting magnitude of the guide star needed for the observations.

To exploit the spatio-temporal correlation in the wavefront, we have recently proposed a data-driven $\mathcal{H}_2$-optimal control strategy consisting of two steps (Hinnen et al. 2006). First, a dedicated subspace-identification algorithm is used to identify a full multi-variable atmospheric disturbance model on the basis of open-loop WFS data. The identified model is then used to compute the optimal controller by formulating the control problem in an $\mathcal{H}_2$-optimal control framework. This formulation is closely related to the use of Linear Quadratic Gaussian (LQG) control for AO (Paschall and Anderson 1993; Looze et al. 1999, 2003; Le Roux et al. 2004). In fact, it can be shown that the LQG and the $\mathcal{H}_2$-optimal control framework are equivalent in the sense that an LQG problem can be recast in an $\mathcal{H}_2$-optimal and *visa versa*. The $\mathcal{H}_2$-optimal control framework however, provides a more elegant way of dealing with the fact that there is a discrepancy between the WFS measurements and the phase signal that we try to minimize. The LQG framework does not make a distinction between reference and performance outputs.

The main difference between the proposed $\mathcal{H}_2$-optimal control design strategy and the LQG based approaches is not in the framework of analysis, but in modeling the atmospheric turbulence and the way of computing the optimal controller. The subspace-identification algorithm provides an efficient way of identifying an atmospheric disturbance model, without assuming any form of decoupling. The existing LQG approaches, on the other hand, are based on an atmospheric disturbance model that either assumes modal decoupling, or consists only of a first-order auto-regressive (AR) model. As a consequence, the LQG approaches are not able to fully exploit the spatio-temporal correlation. Since the $\mathcal{H}_2$-optimal control design strategy does not assume any form of decoupling, it is sufficiently general to exploit the spatio-temporal correlation imposed by the Taylor hypothesis, without being dependent on it. To the best of our knowledge, only the control approach proposed by Gavel and Wiberg (2003) is explicitly based on a frozen-flow model of the atmosphere. This control approach, which is especially developed to account for the spatio-temporal correlation imposed by the Taylor hypothesis, requires accurate knowledge of the wind speed and direction of the frozen-flow. A disadvantage however is that the control approach is not applicable if the Taylor hypothesis is not or only partially satisfied, if the turbulence is distributed over different layers or if the relevant physical parameters, like the wind speed and direction, are not accurately known.

Apart from the extended model structure there is another important difference between the proposed $\mathcal{H}_2$-optimal control approach and existing LQG based approaches. Computing the $\mathcal{H}_2$-optimal controller (and also the LQG controller), generally involves the numerical solution of two Riccati equation. This may become computationally very intensive, especially when considering a full multi-

variable atmospheric disturbance model. In previous work (Hinnen et al. 2006) we have shown that due to the special structure of the identified disturbance model, at most one Riccati equation needs to be solved. Furthermore, it has been shown that if each channel of the transfer function from control input to WFS output exhibits the same scalar dynamics consisting of a two taps impulse response and an integer number of samples delay, then the $\mathcal{H}_2$-optimal control can be computed analytically. Together with the subspace identification algorithm, this results in a non-iterative way to go from open-loop WFS data to closed-loop controller design. The goal of this paper is twofold. First, by analyzing the dynamic behavior of an AO system, it will be shown that the above requirements on the transfer function from control input to WFS output are quite general and hold for any AO system in which the wavefront correction device can be considered to be static. The second goal of this paper is to demonstrate the data-driven $\mathcal{H}_2$-optimal control approach on an experimental setup. The proposed control approach provides a proof of concept that clearly shows the relevance of accounting for the spatio-temporal correlation in the AO controller design.

The remainder of this paper is organized as follows. In Section 4.2 we will first provide a brief description of the AO laboratory setup that has been used to validate the proposed optimal control approach. Subsequently, in Section 4.3 we will have a closer look at the problem of modeling the different components in the experimental setup. After introducing the required notation, the dynamics of the AO system will be considered in more detail. It will be shown that in the case that the wavefront correction device can be considered static, the transfer function from actuator inputs to WFS outputs can be modeled as an integer number of samples delay and in impulse response of two taps. Using this knowledge, a data-driven identification is developed for modeling the AO system. The identified model is precisely in the right form for the proposed $\mathcal{H}_2$-optimal control strategy that is validated in this paper. Section 4.4 provides a brief outline of the main steps in the proposed control approach. The performance of the optimal controller will be compared with a conventional AO control law, which will be briefly reviewed in Section 4.5. After a brief overview of the dominant error sources, this section will also provide an overview of the criteria used for performance evaluation. The outcome of the different experiments is described in Section 4.6. By classifying the different error contributions in the AO system, it will be shown that optimal control is indeed able to reduce the effect of the temporal error. Furthermore, the error classification provides more insight in the conditions under which the optimal control is to be preferred over the common AO control approach. The paper concludes with a short discussion in Section 4.7.

4.2   The experimental setup

This section considers the AO laboratory setup used to test the proposed control approach in an experimental setting. The AO setup, at TNO Science and Industry, the Netherlands, is depicted in Figure 4.1. A schematic representation of the layout of the optical test bench, is provided in Figure 4.2.   In the setup, light
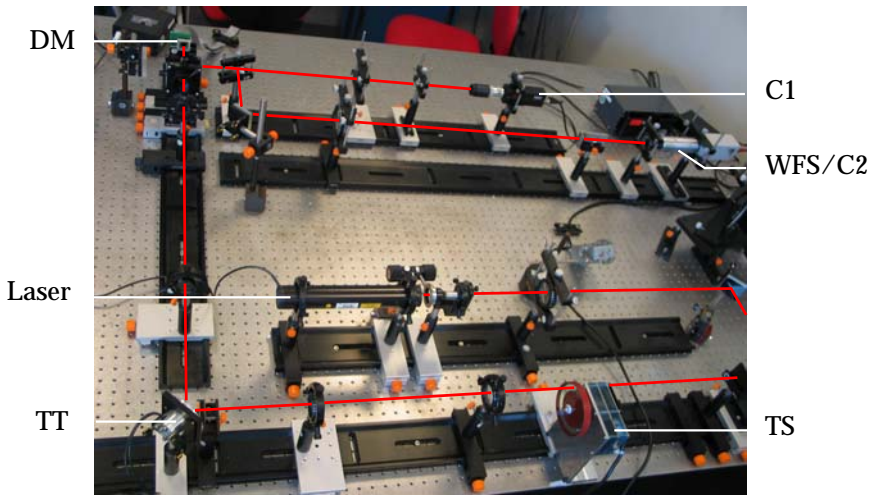
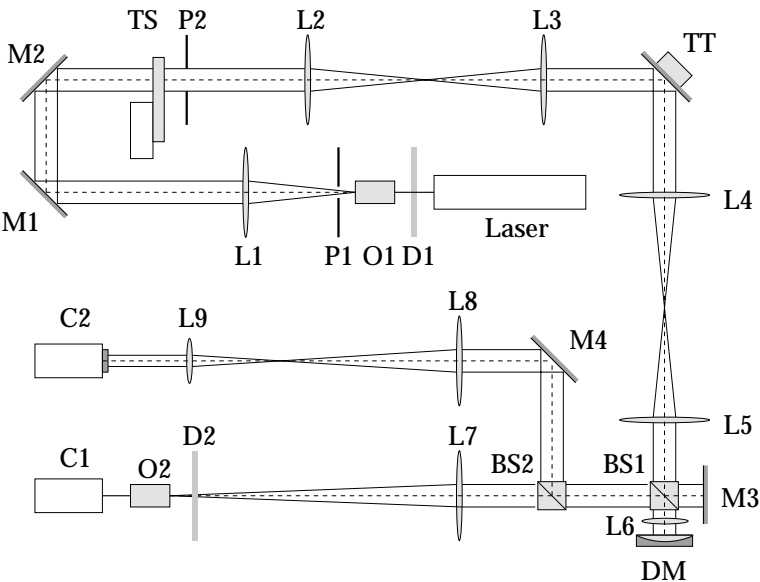Figure 4.1: The adaptive optics laboratory set-up at TNO Science and Industry, Delft, the Netherlands.

Figure 4.2: Schematic representation of the optical layout of the AO test bench at TNO Science and Industry, Delft, the Netherlands.

form a HeNe laser ($\lambda = 633\,\text{nm}$) is focused on a $20\,\mu\text{m}$ pinhole P1 and is then collimated by the lens L1 to mimic a distant point source. As the laser produces a polarized beam, the intensity of the light source can be adjusted by means of the polarizer D1. The atmospheric turbulence is simulated by a turbulence simulator TS consisting of a circular plan parallel glass plate that is rotated through the collimated beam. One side of the glass plate has been machined in such a way that the resulting wavefront distortions have a spatial Kolmogorov distribution. The distortions are characterized by a turbulence coherence length, or Fried parameter (Hardy 1998), of $r_0 = 2\,\text{mm}$. With an entrance pupil of $D = 10\,\text{mm}$, this gives rise to a D/r_0 of 5. By adjusting the rotational speed of the glass plate it is possible to simulate different wind speeds of a single layer of frozen turbulence.

The distorted light is directed to a tip-tilt mirror (TT) that is conjugated to the entrance pupil P2. The TT-mirror consists of a flat mirror mounted on a Physik Instrumente S-330.10 piezo tip-tilt stage. Separate tip-tilt compensation is important since compensation by the deformable mirror (DM) would demand too much of its dynamic range. Via the beam splitter BS1, the entrance pupil is re-imaged on both the DM and the calibration mirror M3. During normal operation the mirror M3 is shielded; it is only used to calibrate the wavefront sensor (WFS). The DM is a 37-channel electrostatic membrane mirror provided by OKO technologies (Vdovin and Sarro 1995). The mirror has a clear aperture of $15\,\text{mm}$ in diameter, and the electrostatic actuators are arranged in a hexagonal grid with an inter-actuator spacing of 1.8mm. A disadvantage of electrostatic actuation is that the actuators are only able to apply a pulling force on the membrane. To allow bi-directional actuation, a bias voltage is applied to each of the actuators. The bias introduces additional focus, and is tuned in such a way that if fully compensates the negative lens L6 in front of the DM.

The second beam splitter cube BS2 divides the light reflected from the DM in the WFS path and the science path. The science path is used to visualize the image after wavefront correction. The lens L7 focuses the light on the objective O2 to produce an enlarged image of the point source on the science camera C1. A neutral density filter D2 adapts the light intensity to the dynamic range of the camera. In the WFS path a Shack-Hartmann sensor is used to probe the residual phase errors. The WFS signal forms the input to the controller, which is responsible for determining the actuator commands to the DM and TT-mirror. The Shack-Hartmann sensor consists of a hexagonal array of 127 micro-lenses ML with a focal distance of $15\,\text{mm}$ and a pitch of $300\,\mu\text{m}$. Like the DM the micro-lens array is provided by OKO technologies. The lenses L8 and L9 reduce the beam size to $3.3\,\text{mm}$ and ensure that also the micro-lens array is conjugated to the entrance pupil P2. The spot pattern formed by the mirco-lens array is imaged on the WFS camera C2. Both camera's are digital progressive scan cameras obtained from SVS-Vistek GmbH. The science camera C1 is a SVS085 Color camera with 1280×1024 square pixels of $6.7\,\mu\text{m}$ and a maximum frame-rate of $13\,\text{Hz}$. The camera C2 used for wavefront sensing is a SVS204MFCP, monochrome camera with 1024×768 square pixels of $4.65\,\mu\text{m}$ and a maximum frame-rate of $50\,\text{Hz}$. Figure 4.3 provides an impression of the sensor-actuator layout of the system. The figure shows a false color scale image of the illuminated WFS spots in the case that there are no atmospheric wave-

front distortions. The yellow crosses on top of the image denote the approximate positions of the DM actuators in the WFS image plane.
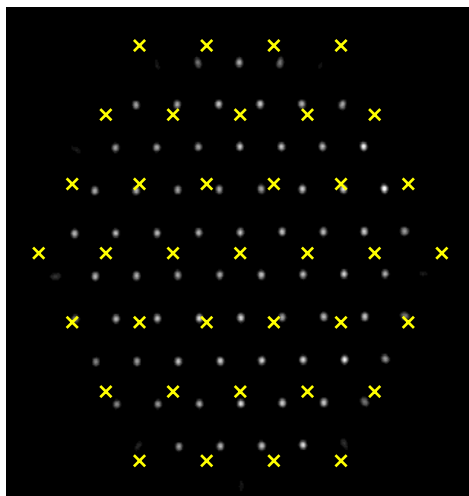

Figure 4.3: Impression of the sensor-actuator layout.

The control computer is a general purpose PC with a $3\,\mathrm{GHz}$ Intel Pentium IV processor and 512 Mb of internal memory, running under real-time Linux. This provides a flexible environment for developing, implementing and testing different kinds of control algorithms. Apart from implementing the control algorithm, the control computer is also responsible for processing the frames from the WFS camera. Using a standard center-of-mass type of algorithm with background compensation and an adjustable threshold level (Hardy 1998; Thomas 2004), the control computer has to estimate the deviation of the spots from their nominal position. Both the centroid algorithm and the control algorithm are implemented in the form of a C-routine, which is executed for each new WFS frame. Using Remote Data Access (RDA) library developed by TNO Science and Industry gives on-line access to various control parameters and signals from Matlab. A second general purpose PC is used for simultaneous recording the images from the camera's C1 and C2. The recorded images are only used for performance evaluation.

Remark 4.1 *To have a well established time reference, the framegrabbers of both cameras as well as the real-time control computer with digital to analogue converters are triggered by an external pulse generator. A precise timing of the sampling and the control action is of utmost importance in demonstrating the optimal control approach. In particular, when using data-driven identification to estimate the transfer function from control input to WFS output, it is important to ensure that synchronization during the identification experiments is precisely the same as during control. Jitter on the sample frequency or an inaccurate synchronization of the WFS process and the digital to analogue converters causes a mismatch between the expected and actual time delay between measurement and correction. Since the optimal control strategy is trying to compensate this delay by means of prediction, jitter and inaccurate synchronization will inevitably lead to a loss*

of performance. Previous experiments with an analogue WFS camera that was not syn-chronized with the control computer, showed a performance comparable to common AO control approach as a result of this effect.

4.3   Modeling the AO system

Computing the $\mathcal{H}_2$-optimal controller requires a control relevant model of both the AO system and the atmospheric wavefront distortions. In this section we will consider the problem of modeling the discrete-time transfer function from DM and TT-mirror actuator inputs to WFS outputs. After deriving the appropriate model structure, a data-driven identification approach is developed to estimate the relevant parameters. In this paper it will be assumed that the wavefront dis-tortion profile can be represented by a finite-dimensional vector signal $\phi(\cdot) \in \mathbb{R}^{m_\phi}$. Whether the signal $\phi(\cdot)$ provides a zonal or modal representation of the wavefront is irrelevant. The only requirement, is that the mean-square error of the vector representation provides a good approximation of the mean-square wavefront er-ror. A similar representation will be used for the phase correction applied by the DM and TT-mirror $\phi_m(\cdot)$ and the residual wavefront error $\epsilon(\cdot) = \phi(\cdot) - \phi_m(\cdot)$. Furthermore, for notional convenience we will use the argument of a signal (e.g. $\phi(\cdot)$) to distinguish between its continuous-time ($\phi(t), t \in \mathbb{R}$) and discrete-time counterpart ($\phi(k), k \in \mathbb{N}$). If no argument is specified, the difference between the continuous-time and discrete-time version of a signal should be clear from the context. A similar convention will be used to distinguish between the Laplace transform ($\phi(s)$) and the z-transform ($\phi(z)$).

4.3.1   The wavefront sensor model

The proposed $\mathcal{H}_2$-optimal control approach requires a discrete-time model of the AO system seen by the controller. In order to derive the model structure for the discrete-time transfer function from control inputs to WFS outputs, we will first consider the relation between the continuous-time versions of these signals. To this end, consider the block-scheme of the Shack-Hartmann WFS in Figure 4.4. The WFS is not able to directly measure the wavefront $\phi(\cdot)$, but provides a signal $s(k)$ that is a filtered version of its slope. The optical transformation from phase
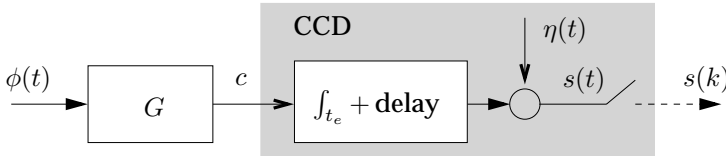

Figure 4.4: Schematic representation of Shack-Hartmann WFS.

$\phi(\cdot)$ to slopes or spot positions $c(\cdot)$ is modeled by the static mapping $c = G\phi$, with G the so called phase-to-slope geometry matrix. The shaded block in Figure 4.4

models the dynamics introduced by the CCD camera. The spot positions on the CCD camera, cannot be observed instantaneously but the camera integrates the image over an exposure time $t_e \in \mathbb{R}$. At the end of the integration, the image is read from the CCD camera, which is then again reset to zero. Furthermore, the time required to read and process the frames, introduces a time delay $t_d \in \mathbb{R}$. With the measurement noise represented by the additive zero-mean white noise term $\eta(t)$, the relation between $s(t)$ and $\phi(t)$ is hence given by

$$s(t) = \frac{1}{t_e} \int_{t-t_e-t_d}^{t-t_d} G\phi(\tau)\mathrm{d}\tau + \eta(t). \tag{4.1}$$

Strictly speaking, the above equation is only valid at the discrete sample instants $t = kT, \; k \in \mathbb{N}$ at which the image is read from the camera. For the moment, however, it is useful to consider the integrating action of the CCD and sampling process separately. Without loss of generality, the discrete WFS output is described as a sampled version of the fictitious continuous-time signal $s(t)$ as defined in equation (4.1). Since at the sampling instants $t = kT, \; k \in \mathbb{N}$ the signal is well defined, considering the integrating action and sampling process separately will not effect the discrete-time output. In the Laplace domain, the above WFS model can be expressed as

$$s = \mathcal{G}(s)\phi + \eta, \quad \text{where} \quad \mathcal{G}(s) = \frac{1 - e^{-st_e}}{st_e} \, e^{-st_d} G = g(s)G. \tag{4.2}$$

Since the delay and the integrating action affect all channels in a similar way, the WFS dynamics are fully decoupled. Furthermore, it is clear that the dynamics can be modeled as an LTI system.

An important complication in the AO control problem is that it is generally not possible to reconstruct the entire wavefront $\phi(t)$ from $s(t)$. This can be easily seen by noting that only the part of $\phi(t)$ that is in the row space of G is able to contribute to $s(t)$. With the phase-to-slope geometry matrix G being tall, the number of modes that cannot be reconstructed from $s(t)$ is equal to the rank deficiency of G. Another consequence of the rank deficiency of G is that, apart from the measurement noise, the channels of $s(t)$ are linear dependent. These observations can be used to introduce a reduced basis that parametrizes only the informative part of $s(t)$. The main advantage of such a basis is that it reduces the effective number of WFS channels that have to be modeled. Moreover, it improves the numerical conditioning of the identification and control problem by removing the linear dependence between the components of the signal $G\phi(t)$. The reduced basis is obtained by considering the singular value decomposition (SVD)

$$G = U\Sigma V^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}, \tag{4.3}$$

where U and V are partitioned such that the matrix Σ_1 contains all nonzero singular values. By substituting the SVD in equation (4.1) and exploiting the orthogonality of U, i.e. $U_1 U_1^T + U_2 U_2^T = I$, the signal $s(t)$ can be decomposed as

$$\begin{aligned} s(t) &= U_1 \Sigma_1 \varphi(t) + (U_1 U_1^T + U_2 U_2^T)\eta(t) \\ &= U_1 y(t) + U_2 U_2^T \eta(t), \end{aligned}$$

where $\varphi(t) \doteq V_1^T \frac{1}{t_e} \int_{t-t_d-t_e}^{t-t_d} \phi(\tau)d\tau$, $y(t) \doteq \Sigma_1\varphi(t) + \nu(t)$ and $\nu(t) \doteq U_1^T\eta(t)$. Since the spaces spanned by U_1 and U_2 are orthogonal and $\eta(t)$ is a zero-mean white noise process uncorrelated to $\phi(t)$, the second term is not related to the turbulence process. For this reason, it is possible to replace the WFS signal $s(t)$ with the lower dimensional signal $y(t)$ without loosing any relevant information. Furthermore, by multiplying equation (4.4) from the left by U_1^T it is clear that the signal $y(t)$ can be simply obtained as $y(t) \doteq U_1^T s(t)$. Since the wavefront cannot be measured directly, the only way to relate the $\phi(t)$ to $y(t)$ via the reduced WFS model

$$y = \Sigma_1\varphi + \nu. \tag{4.4}$$

Because of the orthogonality of V, $\phi(t)$ can be decomposed $\phi(t) = V_1 V_1^T \phi(t) + V_2 V_2^T \phi(t)$. By substituting this in the definition of $\varphi(t)$ it is clear that only the first term, i.e. $V_1 V_1^T \phi(t)$, can be reconstructed from the measurements. The signal $\varphi(t)$ can hence be interpreted as a filtered reduced representation of the observable part of $\phi(t)$ (see Hinnen et al. 2006). The signal $y(t)$, on the other hand, can be regarded as a reduced representation of $s(t)$. Furthermore, since the signals $V_1 V_1^T \phi(t)$ and $\varphi(t)$ have the same 2-norm, the control problem can be reformulated as finding the controller that minimizes the variance of $\varphi(t)$.

In accordance with the above definitions, the reduced representation of the applied phase correction is defined as $\varphi_m(t) \doteq g(t)V_1^T\phi_m(t)$, while the corresponding residual phase error is defined as $\varepsilon(t) \doteq \varphi(t) - \varphi_m(t)$. Since the WFS is linear in its input, the output corresponding to $\varepsilon(t)$ can be expressed as $r(t) \doteq y(t) - y_m(t)$, where $y(t)$ and $y_m(t)$ denote the contributions due to $\varphi(t)$ and $\varphi_m(t)$, respectively. Note that by definition of $\varphi(t)$ and $\varphi_m(t)$, the WFS model (4.4) becomes static as the WFS dynamics are included in the mirror model.

4.3.2 Linearization of the deformable mirror

In the proposed $\mathcal{H}_2$-optimal control approach it is assumed that the transfer function from control input $u(k)$ to WFS output $y(k)$ can be described by an LTI system. This implies that the DM and TT-mirror are assumed to be linear in their actuator inputs. Apart from a possible offset this holds for the TT-mirror, but the DM still does not satisfy this assumption. For this reason an additional linearization step has to be performed. As described in Section 4.2, the DM is an electrostatic actuated membrane mirror and the actuators are almost linear with the applied voltage squared. Furthermore, to allow bi-directional actuation, a bias voltage has to be applied to each of the actuators. Taking into account this bias, the DM can be linearized by defining a new artificial control input $u_{dm}(k) \in \mathbb{R}^{m_d}$, which is related to the actuator input $u_l(k)$ that is actually applied to DM as

$$u_l(k) = \sqrt{u_{dm}(k) + u_b^{dm}}, \tag{4.5}$$

where $u_b^{dm} \in \mathbb{R}^{m_d}$ denotes the bias and $m_d \in \mathbb{N}$ is the number of DM actuators. The above operation on the control input makes the DM linear in $u_{dm}(k)$. The bias u_b^{dm} on the actuator inputs has to be chosen in such a way that the initial

shape of the DM is compensated by the additional negative bias lens in front of the mirror. The required actuator bias u_b^{dm} is found by a calibration procedure. The calibration procedure consists of an iterative algorithm that searches for the actuator bias u_b^{dm} that minimizes the measured wavefront. In a similar way, the offset on the TT-mirror is removed by adding a bias u_b^{tt} to the TT-mirror control input $u_{tt}(k)$. In the proposed $\mathcal{H}_2$-optimal control approach there is no separate loop for controlling the TT-mirror. Controlling the TT-mirror is seen as an integral part of the control design problem. The DM and TT-mirror are described by a integrated model with a control input $u(k)$ obtained by stacking $u_{dm}(k)$ and $u_{tt}(k)$. As we will see in Section 4.6, the above procedure linearizes the transfer function for control input $u(k)$ to WFS output $y(k)$ to such an extend that the system can be considered LTI for the purposes of this paper.

4.3.3 Discrete-time active mirror and WFS model structure

The WFS model in equation (4.4) provides a continuous-time model of the relation between the open-loop wavefront distortion $\varphi(t)$ and the corresponding WFS measurement signal $y(t)$. The continuous-time signal $y(t)$, however, is physically non-existent as the WFS provides only a sampled data output. Furthermore, the control design strategy requires a discrete-time description of the AO system. To derive such a model, consider Figure 4.5. It provides a schematic representation of the relation between the discrete-time control input $u(k) \in \mathbb{R}^{m_u}$ and the corresponding WFS output $y_m(k)$.
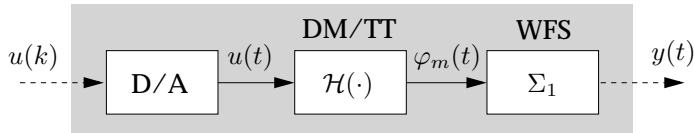
Figure 4.5: Schematic representation of AO system as seen from the controller.

The wavefront correction $\varphi_m(t)$, generated by the DM and TT-mirror, depends on the continuous-time actuator input $u(t) \in \mathbb{R}^{m_s}$. This signal is in turn generated by a digital-to-analog (D/A) converter with discrete-time control input $u(k) \in \mathbb{R}^{m_s}$. As a result, the contribution of $y_m(k)$ to the closed-loop WFS signal $r(k)$ can be seen as the output of the discrete-time system formed by the cascade of D/A converter, active mirrors and WFS. In the proposed $\mathcal{H}_2$-optimal control approach there is no separate loop for controlling the TT-mirror. Controlling the TT-mirror is an integral part of the control design problem. The DM and TT-mirror, also denoted by active mirrors, are described by a single integrated model $\varphi_m(s) = \mathcal{H}(s)u(s)$.

Given the model $y_m(s) = \Sigma_1 \mathcal{H}(s)u(s)$, the equivalent discrete-time transfer function $\Sigma_1 \mathcal{H}(z)$ from $u(k)$ to $y_m(k)$ can be computed using the step-invariant transformation (Åström and Wittenmark 1997; Chen and Francis 1995). Looze (2005), for instance, has used the step-invariant transformation to derive the equivalent discrete-time transfer function for a system with ZOH input and discrete-

time measurements based on a CCD camera with an exposure time equal to the sampling interval. Even though $\Sigma_1 \mathcal{H}(z)$ could be derived in this way, the analysis will be performed in the time-domain as this provides more insight.

The active mirrors used in the experimental setup have a time constant that is short compared to CCD exposure-time and the sampling period of the system. Considering a step response, the DM reaches 80% of its maximum value within $500\,\mu s$ (Vdovin and Sarro 1995). The first resonance frequency of the DM is located around the $2\,kHz$ (Weyrauch et al. 2001), while the resonance frequency of the TT mirror is specified at $2.4\,kHz$. These resonance frequencies are very high compared to the maximum sample frequency of $20\,Hz$ considered in the validation experiments. Also the used CCD exposure time ($t_e = 5\,ms$) is short compared to the rise time of the DM. Hence, from a practical point of view, the mirrors can be considered to be static and the only dynamics in transfer function from $u(k)$ to $y_m(k)$ derives from the D/A converter and the WFS sampling process. The projected wavefront $V_1^T \phi_m(t)$ can be hence expressed as $V_1^T \phi_m(t) = \bar{H} u(t)$, with $\bar{H} \in \mathbb{R}^{m_y \times m_y}$ a static influence matrix. By pre-multiplying both sides of equation (4.1) with U_1^T, and substituting this relation, the WFS output at time instant $t = kT$, $k \in \mathbb{N}$ can be expressed as

$$y_m(k) = \frac{1}{t_e} \int_{kT-t_e-t_d}^{kT-t_d} \Sigma_1 \bar{H} u(\tau) \mathrm{d}\tau, \tag{4.6}$$

where the measurement noise is temporarily left out of consideration as this will be accounted for in the WFS contribution due to atmospheric turbulence. The continuous-time actuator input $u(t)$, is obtained from a ZOH type of D/A-converter operated at the same sample frequency as the camera, i.e.

$$u(t) \doteq u(k) \quad \text{for} \quad kT \leq t < (k+1)T. \tag{4.7}$$

Due to the physical limitations of the CCD camera, the exposure time should always be in the range $0 < t_e \leq T$. According to equation (4.6), this implies that the output $y_m(k)$ depends at most on two past samples of $u(k)$. To elaborate the integral, divide the time-delay $t_d \in \mathbb{R}$ in an integer number samples delay $d \in \mathbb{N}$ and a remainder $\tau_d \in \mathbb{R}$ as $t_d = dT - \tau_d$, where $d \geq 1$ and $0 < \tau_d \leq T$. Furthermore, let us assume for the moment that $t_e > \tau_d$. Then using the ZOH-relation (4.7), the WFS output $y_m(k)$ can be expressed as

$$y_m(k) = \Sigma_1 \bar{H} \left(\frac{1}{t_e} \int_0^{\tau_d} u(k-d)\mathrm{d}\tau + \frac{1}{t_e} \int_{T+\tau_d-t_e}^{T} u(k-d-1)\mathrm{d}\tau \right) \tag{4.9a}$$

$$= \Sigma_1 \bar{H}(\alpha_1 u(k-d) + \alpha_2 u(k-d-1)), \tag{4.9b}$$

with $\alpha_1 = \tau_d/t_e$ and $\alpha_2 = (t_e - \tau_d)/t_e$ are real-valued coefficients. A similar analysis can be performed for $t_e \leq \tau_d$. In this case the limits of the first integral extents for 0 to t_e, while the contribution due to the second term is zero. As a result, the WFS output $y_m(k)$ can still be expressed as in (4.9b), but with $\alpha_1 = 1$ and $\alpha_2 = 0$. By comparing equation (4.4) and equation (4.9b) it is clear that for a static mirror, the discrete time transfer function from $u(k)$ to $\varphi_m(k)$ can be expressed as $\mathcal{H}(z) = g(z)\bar{H}$, with $g(z) = z^{-d}(\alpha_1 + \alpha_2 z^{-1})$. Hence, for a static mirror the scalar

dynamics $g(z)$ can be expressed as two taps finite impulse response (FIR) filter plus an integer number of samples delay. That the transfer function still includes dynamics, is caused by the particular choice of the reduced wavefront signal $\varphi(t)$. Due to the definition of the combined DM and TT-mirror, $\mathcal{H}(z)$ incorporates the scalar dynamics introduced by the D/A converter and WFS. By introducing the definitions $H \doteq \alpha_1 \bar{H}$ and $\alpha \doteq \alpha_2/\alpha_1$, this gives rise to the following model structure for the mirror

$$\mathcal{H}(z) = z^{-d}(H + \alpha z^{-1} H). \tag{4.10}$$

Considering α as a free parameter, the above model structure will also hold when the DM is not perfectly static but has a time-constant that is short compared to the WFS exposure time t_e. This is nicely illustrated by the simulation example in Looze (2005). If the time-constant is too large, the mirror model substituted in equation (4.6) should be replaced by a dynamic one. Because of this, the WFS output $y_m(k)$ will generally depend on more than two samples of $u(k)$ and so the required number of FIR taps for modeling $g(z)$ will increase.

4.3.4 Data-driven modeling of DM and TT-mirror

From the previous discussion it is clear that if the DM and TT-mirror can be considered to be static, the relation between the actuator inputs $u(k)$ and WFS outputs $y_m(k)$ reduces to a static influence matrix H followed by a scalar dynamics $g(z)$ transfer function consisting of an integer number of samples delay and an impulse response of two taps. Considering the derived model structure (4.10), a data-driven identification procedure has been developed to estimate the unknown coefficient α and the influence matrix H. Since the proposed model structure includes a priori knowledge on the system, imposing the structure is expected to improve the accuracy of the identified model. Another advantage of considering the specific model structure is that it allows for analytical solution of the $\mathcal{H}_2$ optimal control problem. This will be discussed in more detail in Section 4.4.

The unknown coefficient α and the influence matrix H are estimated on the basis of the WFS response $y_m(k)$ measured by exciting the DM and TT-mirror with a zero-mean white input sequence $u(k)$. To avoid actuator saturation, the input signal is generated according a Gaussian probability density function of which the tails are chopped off. Given a data batch of $N \in \mathbb{N}$ samples of the applied actuator input $u(k)$ and the measured WFS response $y_m(k)$, the problem of estimating the unknowns α and H is solved by minimizing a prediction error criterion (Ljung 1999). The main difficulty in estimating the unknowns is that the second term in the proposed model structure (4.10) contains the product of α and H, which gives rise to a non-convex optimization problem. Before focusing on the problem of identifying the unknowns α and H, it is therefore useful to consider the following more general FIR model structure for describing the DM and TT-mirror

$$\mathcal{H}(z) = \sum_{i=p_1}^{p_2} z^{-i} H_{i-p_1+1}, \tag{4.11}$$

where $p \doteq p_2 - p_1 + 1$ denotes the number of nonzero taps. The above model is linear in the matrices H_i, $i \in \{1, \ldots, p\}$ and will be used to obtain an initial estimate of the value optimal α in equation (4.10). Considering this model structure, the WFS output predicted on the basis of the applied input signal can be expressed as $\hat{y}_m(k) = \Sigma_1 \mathcal{H}(z) u(k)$. The problem of identifying the unknown coefficient matrices H_i is now formulated as the following optimization problem

$$\min_{\mathcal{H}(z)} \sum_{k=p_2}^{N-1} \|y_m(k) - \hat{y}_m(k)\|_2^2. \tag{4.12}$$

Since the predicted output $\hat{y}_m(k)$ is linear in the coefficient matrices H_i, the above optimization problem reduces to a standard linear regression problem. The solution to this problem can be computed analytically and is given by

$$\begin{bmatrix} \hat{H}_p & \ldots & \hat{H}_2 & \hat{H}_1 \end{bmatrix} = \Sigma_1^{-1} Y_{p_2, N-1} \left(U_{0,p,N-p_2} \right)^{\dagger}, \tag{4.13}$$

where

$$U_{0,p,N-p_2} \doteq \begin{bmatrix} u_m(0) & u_m(1) & \ldots & u_m(N-p_2-1) \\ u_m(1) & u_m(2) & \ldots & u_m(N-p_2) \\ \vdots & \vdots & \ddots & \vdots \\ u_m(p-1) & u_m(p) & \ldots & u_m(N-p_1), \end{bmatrix}, \tag{4.14}$$

and $Y_{p_2,N-1} \doteq \begin{bmatrix} y_m(p_2) & y_m(p_2+1) & \ldots y_m(N-1) \end{bmatrix}$ and $(\cdot)^{\dagger}$ is used to denote the (right) pseudo-inverse. By making a proper choice for p_1 and p_2, exploratory identification experiments can be used both to validate the proposed model structure and to obtain an initial estimate of the value of α. Also the number of samples delay d may be determined by exploratory identification experiments using the FIR model structure. Suppose that p_1 and p_2 are chosen such that $p_1 \leq d \leq p_2 - 1$. Then if the proposed model structure (4.10) is right, there are at most two FIR coefficient matrices H_i $i \in \{1, \ldots, p\}$, succeeding one another, with a norm significantly different from zero. Moreover, the index $j \in \mathbb{N}$ of the first of these coefficient is related to the number of samples delay as $d = j + p_1 - 1$.

From the derivation of model structure (4.10) it is clear that there are two cases to be considered. First, if $t_e \leq \tau_d$, the coefficient α is zero, which implies that the exploratory experiments should give rise to only one FIR coefficient significantly different from zero. If this is the case, the problem of identifying H reduces to a standard linear least squares problem and H can be estimated by setting $p_1 = p_2 = d$ in the general FIR model (4.13). On the other hand, if $t_e > \tau_d$, there should be precisely two FIR coefficient matrices H_{j+1} and H_{j+2} with a norm significantly different from zero. The ratio of these norms, i.e. $\|H_{j+2}\|_2/\|H_{j+1}\|_2$, provides an estimate of α. The so-computed α is used to initialize a numerical procedure for solving the optimization problem obtained by substituting the model structure (4.10) in the prediction error criterion (4.12). Using the block-Hankel notation, this gives rise to the following optimization problem in the unknowns

$$\min_{\alpha, H} \|Y_{d+1,N-1} - \Sigma_1 H[I \ \alpha I] U_{0,2,N-d-1}\|_F^2, \tag{4.15}$$

where the block Hankel matrix $U_{1,2,N-p_2}$ is defined in accordance with equation (4.14). Due to the imposed structure on the FIR coefficient matrix, the above optimization problem is no longer a convex function in the unknowns α and H. The resulting optimization problem however, can be efficiently solved by using the obtained initial estimate of α and the concept of separable least squares proposed by Golub and Pereyra (1973).

Separable least squares is a technique for solving nonlinear least squares optimization problems in which the parameters can be divided in a set that enters the function that is minimized in a linear way and a disjoint set that contains the remaining parameters. The parameters that enter in a linear way can be eliminated from the least-squares problem, resulting in a modified nonlinear optimization problem in a reduced parameter space. After solving the modified optimization problem, the linear parameters can be computed from a linear least squares problem. In the nonlinear least squares problem (4.15), only the product of α and H is responsible for a nonlinear dependence of prediction-error on the unknowns. For a fixed value of α, the nonlinear least squares problem (4.15) reduces to a linear least squares problem. As a result, the optimal matrix H for a given value of α can be computed analytically and is given by

$$H^*(\alpha) = \Sigma_1^{-1} Y_{d+1,N-1} \left([I \ \ \alpha I] U_{0,2,N-d-1} \right)^\dagger ,\qquad (4.16)$$

By substituting the above expression in (4.15) it is possible to eliminate the unknown H, which gives rise to a modified nonlinear optimization problem over $\alpha \in \mathbb{R}$. Since this is an optimization problem over one single parameter, it can be efficiently solved using standard numerical optimization tools. Furthermore, the exploratory experiments provide already a reasonably initial estimate of α, which reduces the risk of ending up in a local minimum. Given the value for α found in the numerical optimization process, α^*, the estimate for the influence matrix H can be computed from equation (4.16).

4.4 Data-driven optimal control for AO

This section provides a brief outline of the recently proposed data-driven $\mathcal{H}_2$-optimal control design strategy considered in (Hinnen et al. 2006) and clarifies the need for an accurate model of the AO system. More specifically, it will be shown that the model structure derived in the previous section is in harmony with the control approach. Indeed, the developed identification procedure for estimating the parameters in equation (4.10) can be seen as the completion of the proposed control approach, in line with the philosophy of using data-driven identification.

The proposed $\mathcal{H}_2$-optimal control design approach is based on the generalized plant model depicted in Figure 4.6. Its main ingredients, indicated by the shaded boxes, are a model of the atmospheric turbulence $\mathcal{S}(z)$ and the model of the AO system $\mathcal{H}(z)$ considered in the previous section. Starting from $\mathcal{H}(z)$, the first step of the proposed control design strategy is to identify $\mathcal{S}(z)$. Here, it is assumed that the second order statistics of the signal $y(k)$ can be described as the output

of a LTI system with a zero-mean white noise input $v(k) \in \mathbb{R}^{m_y}$ and covariance $R_v \doteq \mathcal{E}\{v^T v\}$. This in combination with equation (4.4), gives rise to the following model structure

$$\mathcal{S} : \begin{cases} x(k+1) = & A_d x(k) + K_d v(k) \\ y(k) = \Sigma_1 C_d x(k) + & v(k) \\ \varphi(k) = & C_d x(k) + \zeta(k) \end{cases}, \tag{4.17}$$

where $(A_d - K_d \Sigma_1 C_d) \in \mathbb{R}^{n_d \times n_d}$ and $A_d \in \mathbb{R}^{n_d \times n_d}$ are stable, and $\zeta(k) = \Sigma_1^{-1}(v(k) - \nu(k))$ is a zero-mean white noise sequence with covariance $R_\zeta \doteq \mathcal{E}\{\zeta^T \zeta\}$. Considering this model structure is reasonable if the statistical properties of the wavefront change on a time scale that is long to the time scale of the fluctuations themselves. Indeed, validation experiments on open-loop WFS data from the William Herschel Telescope have shown that a model of this form can be used to predict future wavefront distortions (Hinnen et al. 2005b). Like $\mathcal{H}(z)$, the atmospheric disturbance model $\mathcal{S}(z)$ is identified on the basis of open-loop WFS data $y(k)$. Data-driven modeling has the advantage that it provides a good match with the prevalent turbulence conditions and does not depend on restrictive assumptions like the frozen layer hypothesis. Moreover, since the model structure does not assume any form of decoupling, it is sufficiently general to capture the spatio-temporal correlation imposed by a frozen flow without being dependent on it. A consequence of this extensive description is that already relatively small AO systems give rise to a considerable identification problem. For this reason, a dedicated subspace-identification algorithm has been developed. One of the main advantages of this algorithm is that it provides an estimate of the system matrices without the need for spectral factorization.
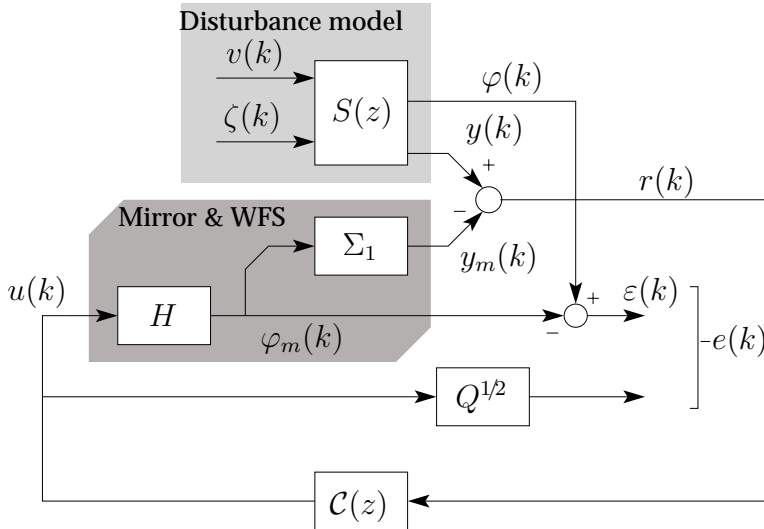

Figure 4.6: Schematic representation of the closed-loop AO system together with the defined performance outputs.

Given the identified atmospheric disturbance model $\mathcal{S}(z)$ and the transfer function $\mathcal{H}(z)$, the final step of the proposed control design strategy is to compute the

optimal controller. The control objective is to find the controller that minimizes the 2-norm of the performance output $e(k)$ composed of the sub-signals $\varepsilon(k)$ and $Q^{1/2}u(k)$, where $Q = Q^T \geq 0$ is typically chosen diagonal. As can be easily verified, this is equivalent to finding the controller $C(z)$ that minimizes the cost function

$$J = \mathcal{E}\left\{\varepsilon^T(k)\varepsilon(k)\right\} + \mathcal{E}\left\{u^T(k)Qu(k)\right\}, \tag{4.18}$$

where $\mathcal{E}$ denotes the conditional expectation. The regularization Q makes a trade-off between the objective of minimizing the expected residual wavefront error $\mathcal{E}(\varepsilon(k)^T\varepsilon(k))$ and the objective of minimizing the expected amount of control effort $\mathcal{E}(u(k)^Tu(k))$. By increasing the regularization matrix Q, it is possible to reduce the amount of energy dissipated by the DM and make the controller more robust to model uncertainties. In a normal AO system, the input regularization should be chosen as small as possible without sacrificing too much robustness in order to give priority to general objective of minimizing the aperture averaged mean-square residual wavefront error. The problem of finding the closed-loop optimal controller $C(z)$ can be conveniently expressed in a $\mathcal{H}_2$-optimal control framework. Even though straightforward application of standard $\mathcal{H}_2$-optimal control theory (see e.g. Chen and Francis 1995) provides a way to compute the optimal controller, the approach is unable to exploit the special structure in the AO control problem. Computing the $\mathcal{H}_2$-optimal controller generally involves the numerical solution of two algebraic Riccati equations. However, since the atmospheric disturbance model (4.17) is minimum-phase with respect to $y(k)$, the controller can be computed more efficiently. Indeed, if $\mathcal{H}(z)$ has the form (4.10), it is possible to derive an analytical expression for the optimal controller Hinnen et al. (2006). This result is summarized in the following theorem.

Theorem 4.1 (Optimal control with quasi-static DM) *Let the wavefront distortions $\varphi(k)$ and open-loop WFS signal $y(k)$ be characterized by the atmospheric disturbance model (4.17), with input covariance matrix $R_v > 0$. Assume that the transfer function from $u(k)$ to $\varphi_m(k)$ is of the form (4.10) and that either H or Q has full column rank. Furthermore, let the regularized left pseudo-inverse of the DM influence matrix H be defined as*

$$H_Q^\dagger \doteq \begin{cases} (H^TH + Q)^{-1}H^T & \text{for } \alpha = 0 \\ \alpha\gamma_\alpha \left(\alpha^2 H^TH + \gamma_\alpha^2 Q\right)^{-1}H^T & \text{otherwise} \end{cases} \tag{4.19}$$

with

$$\gamma_\alpha \doteq \begin{cases} \alpha & \text{for } |\alpha| \leq 1 \\ sgn(\alpha) & \text{for } |\alpha| > 1 \end{cases}.$$

Then the optimal controller $C(z)$, that is minimizing the cost function (4.18), is given by:

$$\begin{bmatrix} \hat{\xi}_1(k+1) \\ \hat{\xi}_2(k+1) \\ \hline u(k) \end{bmatrix} = \left[\begin{array}{cc|c} -G & F & 0 \\ L & M & K_d \\ \hline FL+G^2 & FM-GF & FK_d \end{array} \right] \begin{bmatrix} \hat{\xi}_1(k) \\ \hat{\xi}_2(k) \\ \hline r(k) \end{bmatrix}. \tag{4.20}$$

where the matrices F, G, L and M are defined as $F \doteq H_Q^\dagger C_d A_d^{d-1}$, $G \doteq \gamma_\alpha H_Q^\dagger H$, $L \doteq z^{-d+1}K_d\Sigma_1 H(\alpha I - G)$ and $M \doteq \widetilde{A} + z^{-d+1}K_d\Sigma_1 HF$, with $\widetilde{A} \doteq A_d - K_d\Sigma_1 C_d$.

The above theorem has first been proved for the special case of $d = 1$ and $Q = \rho I$, with $\rho \in \mathbb{R}$ (Hinnen et al. 2005a). A sketch of the proof for the more general case considered here, can be found in Hinnen et al. (2006). Having an analytical expression for the $\mathcal{H}_2$-optimal controller is useful as it leads to an efficient implementation of the proposed control design strategy. Together with the subspace identification algorithm, it gives rise to a non-iterative way to go from open-loop measurement data $y(k)$ to closed-loop controller design that is entirely based on standard matrix operations. Besides advantages in computational efficiency, the analytical solution is also attractive for numerical robustness. Since the poles of the atmospheric disturbance model typically cluster in neighborhood of the point $z = -1$, standard Riccati solvers may suffer from convergence problems and increased numerical sensitivity.

Finally, the analytical solution is useful as it allows for a nice interpretation of the optimal control design. By comparing the state-space equations of the optimal and the atmospheric disturbance model, it can be shown (Hinnen et al. 2006) that the control output can be expressed as

$$u(k) = (I - \gamma_\alpha H_Q^\dagger H z^{-1})^{-1} H_Q^\dagger \hat{\varphi}(k + d|k), \qquad (4.21)$$

where $\hat{\varphi}(k+d|k)$ denotes the conditional estimate of $\varphi(k+d)$ given the past closed-loop WFS data $r(j) = y(j) - y_m(j), j \leq k$. From this expression it is clear that the optimal controller consists of a part that is concerned with predicting future wavefront distortions, over a time horizon of d samples, and a dynamic filter that projects the estimated wavefront on the actuator space. In line with the cost function, the dynamic filter makes a trade off between the objective of finding the control input $u(k)$ that minimizes the difference with the estimated wavefront and the one that minimizes the control effort. Here it is interesting to note that the wavefront prediction part of the controller is unaffected by the input regularization applied in the cost function. In the special case that $\alpha = 0$, the dynamic filter reduces to the static matrix projection $H_Q^\dagger$. Furthermore, the above division of the optimal controller also shows the resemblance with the common AO control approach. As will be shown in Section 4.5, also the common AO control problem decomposes in a wavefront reconstruction step followed by a projection on the actuator space. The main difference between both approaches is that in the optimal control approach the static wavefront reconstruction is replaced by dynamic prediction, which compensates for the delay in the system.

4.5   Validation methods and performance measures

The proposed data-driven $\mathcal{H}_2$-optimal control approach, discussed in the previous section, will be compared with a regularized version of the control approach commonly used in AO. For conformity with the rest of the paper, the common (or classical) AO control approach will be briefly reviewed in terms of the reduced signals introduced in Section 4.3. The common AO control approach consists of a cascade of a static matrix multiplication and a series of parallel feedback loops (Roddier

1999). Given a new WFS measurement $y(k)$, the static part deals with the problem of finding the DM actuator inputs $u(k)$ that would provide the best fit to the wavefront. Let the static relation between $u(k)$ and $y(k)$ be given by $u(k) = Ry(k)$ and let the mirrors be modeled as $\varphi_m(k) = Hu(k)$. Then, with the static WFS model (4.4), the problem of finding the reconstruction matrix R is formulated as

$$R = \arg\min_{\widetilde{R}} \left(\mathcal{E}\{[\varphi(k) - H\widetilde{R}y(k)]^T[\varphi(k) - H\widetilde{R}y(k)]\} + \mathcal{E}\{u^T(k)Qu(k)\} \right), \quad \textbf{(4.22)}$$

where $\mathcal{E}$ denotes the conditional expectation with respect to the WFS measurement $y(k)$. In comparison with the usual minimum-variance formulation of the reconstruction problem, the above optimization problem includes an additional penalty on the control effort. This regularization is needed to enable a fair comparison between both control approaches. The dynamic range of the DM is rather small compared to the wavefront distortions introduced by the atmospheric turbulence simulator. Without input regularization, the actuators of the DM easily saturate, which may even provoke closed-loop instabilities. For both control approaches, significant input regularization is needed to ensure that the DM stays within its linear range. By noting that the first term in equation (4.22) can be interpreted as the residual fitting error, it is clear that there is a one to one correspondence with the cost function (4.18) as considered in the optimal control approach. However, a disadvantage of penalizing the control effort is that it also leads to an increase in the fitting error. Under the assumption that the wavefront $\varphi(k)$ and the measurement noise $\nu(k)$ are uncorrelated zero-mean stochastic processes with a Gaussian distribution, the maximum *a posteriori* estimate of R is given by

$$R = \underbrace{(H^T H + Q)^{-1} H^T}_{F} \underbrace{C_\varphi \Sigma_1 (\Sigma_1 C_\varphi \Sigma_1 + C_\nu)^{-1}}_{E}, \quad \textbf{(4.23)}$$

where $C_\varphi \doteq \mathcal{E}\{\varphi(k)\varphi^T(k)\}$ and $C_\nu \doteq \mathcal{E}\{\nu(k)\nu^T(k)\}$ denote the covariance matrices of $\varphi(k)$ and $\nu(k)$, respectively. Form the above equation it is clear that the reconstruction matrix R can be separated into a fitting operator F and a estimation operator E. The operator E provides a minimum-variance estimate of the wavefront, while the operator F is responsible for projecting this estimate onto the actuator space. Note that the fitting operator F is equal to, and has the same function as, the projection $H_Q^\dagger$ in the optimal control approach. Since the AO system is operated in closed-loop, the signal obtained from the static reconstruction $u(k) = Ry(k)$ provides only an estimate of the increment needed to the current actuator commands. In order to ensure stability and closed-loop performance, the parallel feedback loops have to possess integrating action. The following control law has been used for performance comparison

$$u(k) = \frac{c_1}{1 - c_2 z^{-1}} Rr(k), \quad \textbf{(4.24)}$$

where the gain factor c_1 and the loss factor c_2 are user defined control parameters. The modification of the covariance matrix C_φ, as a result of closed-loop operation, is neglected in computing the construction matrix R.

To obtain a better insight of the conditions under which the proposed data-driven $\mathcal{H}_2$-optimal control strategy should be able to outperform the classical

control approach, it is important to have a closer look at the error sources in an AO system. In order to achieve a performance improvement, at least one of the dominant error contributions has to be influenced by the control design. Examples of well-know error sources in AO include anisoplanatic errors, wavefront measurement errors, wavefront fitting errors, temporal errors and tilt related errors (Hardy 1998). From these error sources, in particular, the temporal error is strongly dependent on the control design. The temporal error is one of the major error sources (van Dam et al. 2004) for the AO system on the Keck telescope.

In the experimental setup, the total wavefront error is determined almost exclusively by the wavefront fitting error and the temporal error. The wavefront fitting error is the error caused by the fact that the active mirrors cannot take an arbitrary shape. Since the mirrors have only a limited degrees of freedom only a part of the wavefront can be compensated. The size of the fitting error depends on both the turbulence statistics and the design of the active mirrors. In the case of Kolmogorov turbulence, the mean-square fitting error is given by

$$\sigma_f^2 = a_f (d/r_0)^{5/3},$$

where $a_f \in \mathbb{R}$ is a fitting coefficient depending on the influence function and d denotes the inter-actuator spacing. With the Fried parameter r_0 of the atmospheric disturbance simulator being fixed, the fitting error can be considered constant.

The temporal error on the other hand is caused by the inability of the AO system to immediately respond to changes in the wavefront at the very moment of occurrence. Both bandwidth limitations and pure time delays contribute to a delayed response. All errors related to imperfections in the temporal compensation are attributed to the temporal error. Advanced control strategies, like the optimal control approach considered in this paper, should be able to reduce the effect of the time delays by predicting the wavefront distortion at the time of correction. The temporal error contribution depends on turbulence dynamics, the pure time delay in the system and the bandwidth limitations. Assume for the moment that the pure time delay is inversely proportional to the sample frequency and that the bandwidth limitations are proportional to the sample frequency. Furthermore, assume that the controller consists of decoupled control loops with a first-order order low pass characteristic, then for Kolmogorov turbulence the mean-square temporal can be approximated as

$$\sigma_t^2 = a_t \left(f_G/f \right)^{5/3},$$

where $a_t \in \mathbb{R}$ is a scaling constant depending on the precise form of the controller and bandwidth specifications and f_G is the so-called Greenwood frequency (Hardy 1998; Fried 1990). The Greenwood frequency can be seen as a characteristic frequency of the atmospheric turbulence. For a single frozen layer with wind velocity v it can be computed as $f_G = 0.427(v/r_0)$. From the above discussion it is clear that the temporal error scales as a power law of the Greenwood to sample frequency ratio f_G/f. Under the assumption that the error sources are uncorrelated, the total wavefront error is given by summing the variances, which forms the motivation for the following generalized error model

$$\sigma_\varepsilon^2 \approx a_0 + a_1 \left(f_G/f \right)^{a_2}, \tag{4.25}$$

where a_0 is a constant representing the fitting error and a_1 and $a_2 \geq 0$ are the scaling constant and exponent of the temporal error. For a controller consisting of first-order feedback loops, the constant a_2 should be close to $5/3$. Furthermore, it will be assumed that the residual wavefront error obtained with other control laws satisfy the same expression, possibly with a different constant a_2. Since the fitting error a_0 is independent from the control design, it is to be expected, that large performance improvements can only be achieved at large Greenwood to sample frequency ratios.

The performance of both controllers has been compared at different Greenwood to sample frequency ratios. At each Greenwood to sample frequency ratio, the mean-square residual wavefront error is estimated on the basis of $N_s = 5500$ samples of WFS signal $r(k)$, measured during closed-loop operation. Given this data, the sample estimate of the mean-square residual wavefront error is computed as

$$\hat{\sigma}_\epsilon^2 = \frac{1}{(N_s - 1)m_\phi} \sum_{k=1}^{N_s} \hat{\varepsilon}(k)^T \hat{\varepsilon}(k), \tag{4.26}$$

where $\hat{\varepsilon}(k) = \Sigma_1^{-1} r(k)$ denotes the reconstructed residual wavefront. By normalizing the sample estimate in this way, it is consistent with the observable part of the physical mean-square phase error.

To verify if the error model (4.25) indeed provides a good description of the residual wavefront error, it has been fitted to the observed values of $\hat{\sigma}_\epsilon^2$. For a fixed a_0, the problem of estimating the coefficients a_1 and a_2 boils down to fitting an exponential relation $a_1 (f/f_G)^{a_2}$ to the measurements $(\hat{\sigma}_\epsilon^2 - a_0)$. Such a fitting problem is conveniently solved on a logarithmic scale as this renders the fitting error linear in the unknowns $\log_{10}(a_1)$ and a_2. This forms the motivation for defining the following least squares problem to estimate a_0, a_1 and a_2:

$$\min_{a_0, a_1, a_2} \left\| \log_{10}(\bar{\sigma}_\varepsilon^2 - a_0) - \mathbb{1}\log_{10}(a_1) - a_2 \log_{10}\left(\bar{f}\right) \right\|_2^2, \tag{4.27}$$

where $\bar{\sigma}_\varepsilon^2$ and $\bar{f}$ are the vectors obtained by stacking the different observations of σ_ϵ^2 and the corresponding f_G/f ratios, respectively. Furthermore, $\mathbb{1}$ is a vector of the same dimension as $\bar{\sigma}_\varepsilon^2$ and $\bar{f}$, with all elements equal to 1. For a fixed value of a_0, the above optimization problem reduces to a standard least squares problem, which has a global optimum given by

$$\begin{bmatrix} \log_{10}(a_1^*) \\ a_2^* \end{bmatrix} = \begin{bmatrix} \mathbb{1} & \log_{10}(\bar{f}) \end{bmatrix}^\dagger \log_{10}(\bar{\sigma}_\varepsilon^2 - a_0), \tag{4.28}$$

where the asterisk is used to denote the optimal value. From the above equation it is clear that the optimal values for a_1 and a_2 are a function of a_0. As a result, it is again possible to use a separable least squares approach to simplify the optimization process. By substituting the optimal values for a_1 and a_2 in equation (4.27), the original optimization problem reducers to an optimization problem in a single parameter, which can easily be solved using a line search.

Apart from the estimated mean-square residual error $\hat{\sigma}_\varepsilon^2$, a number of other criteria have been used to characterize the performance of the AO system. Since

the optimal controller is designed to minimize the cost function (4.18), an obvious choice is to look at the relative improvement in the sample estimates this function. Given the reconstructed residual wavefront error $\varepsilon = \Sigma_1^{-1} r(k)$ and the applied control input $u(k)$ for $k \in \{1 \ldots N_s\}$, the sample estimate of the cost function is defined as

$$\hat{J} = \frac{1}{N_s - 1} \sum_{k=1}^{N_s} \hat{\varepsilon}^T(k)\hat{\varepsilon}(k) + u^T(k)Qu(k). \qquad (4.29)$$

Let $\hat{J}_c$ and $\hat{J}_o$ denote the sample estimates of the cost function J obtained with the common and optimal control approach respectively, then the relative performance improvement by the optimal control approach is defined as $\hat{J}_c/\hat{J}_o$. Both performance criteria $\hat{\sigma}_\varepsilon^2$ and $\hat{J}_c/\hat{J}_o$ are computed on the basis of the closed-loop WFS measurement signal $r(k)$. Since the ultimate objective of an AO system is to obtain a high resolution image of the science object, it is also interesting to have performance criteria that depend directly on the optical quality of the corrected image. In the following, the full-width of half maximum (FWHM), the normalized maximum encircled energy and the Strehl ratio, will be computed as additional performance measures. Of these measures, the Strehl ratio is the most commonly used metric to quantify the imaging quality of an AO system. The Strehl is defined as the peak intensity of the image of a point source, normalized to the diffraction limited peak intensity (Hardy 1998; Roddier 1999; Roberts Jr. et al. 2004).

To obtain an estimate of these quantities $N_f = 250$ frames of the science camera have been collected. Subsequently, the obtained images are averaged as

$$\bar{I}(p) = \frac{1}{N_f} \sum_{k=1}^{N_f} I(k, p),$$

where $I(k, \cdot)$ denotes the measured intensity at time k and the argument $p \in \mathbb{N}^2$ is used to discriminate between the different pixels of a CCD frame. In this way, a recording with a long exposure time is mimicked. Since the science camera has a fixed exposure time of $5\,\mathrm{ms}$, the total recording time of the averaged image is constant. If the wavefront would have been perfectly compensated, the image of the point source would have been diffraction limited. For this reason the FWHM will be estimated by fitting the averaged image $\bar{I}(p)$ to the theoretical Airy pattern for a diffraction limited spot. This is equivalent to the problem of minimizing the cost function

$$\Psi(p_0, g_1, g_2, g_3) = \frac{1}{\sum_p \left\| \bar{I}(p) \right\|_2^2} \sum_p \left\| \bar{I}(p) - g_2 \mathcal{A}(g_1, p_0) - g_3 \right\|_2^2, \qquad (4.30)$$

where $\mathcal{A}(g_1, p_0)$ denotes a normalized Airy spot with a FWHM g_1, centered around the point p_0. Furthermore, g_2 and g_3 are used to parametrize the peak and background intensity of the averaged frame. Since the fitted intensity pattern is linear in g_2 and g_3, separable least squares can be used again to reduce the number of parameters in the nonlinear least squares problem.

Also the normalized maximum encircled energy is computed on the basis of the time averaged image $\bar{I}(p)$. Let $\mathcal{B}_r(p_0) = \{p \in \mathbb{N}^2 | \, \|p - p_o\|_2 < r\}$ denote the set

of pixels $p \in N^2$ located in a circle with radius r around the point $p_o \in \mathbb{N}^2$, then we will adopt the following definition for normalized maximum

$$E(r) = \max_{p_o} \frac{\sum_{p \in \mathcal{B}_r(p_o)} \bar{I}(p) - \bar{I}_b(p)}{\sum_p \bar{I}(p) - \bar{I}_b(p)}, \tag{4.31}$$

where $\bar{I}_b(p)$ denotes the CCD dark pattern computed by averaging $N_f = 250$ background images. The maximum normalized encircled energy defined in this way provides a measure of the fraction of the total incident energy that is contained in the central core of the imaged spot. The better the wavefront correction, the more the energy is concentrated in the central spot and the faster the encircled energy will increase with the radius r. Due to the applied normalization, the encircled energy $E(r)$ will approach one for large values of r. Furthermore, as a result of the maximization over p_o, $E(r)$ is not only insensitive to tip-tilt errors but also to misalignment errors. The normalized maximum encircled energy $E(r)$ is computed for a number predefined values of the radius r. Note that by defining a function $M(p - p_o, r)$ that is one for $p \in \mathcal{B}_r(p_o)$ and zero otherwise, computing the normalized maximum encircled energy reduces to evaluating a convolution. To speed up the computations, the convolution is evaluated in the frequency domain, using precomputed templates of the Fourier transform of the masks $M(p, r)$.

Using the same notation, the Strehl ratio is computed by using the following procedure. Let q_0 denote the radius of the theoretical diffraction limited spot computed on the basis of the aperture size. Then the first step in estimating the Strehl is to extract a neighborhood with a radius $0.2q_0$ around the pixel $p_m \in \mathbb{N}$ with maximum intensity. This neighborhood is used to obtain refined estimates of the peak intensity I_q and position p_q by fitting a quadratic form to the background compensated measured intensity $\bar{I}(p) - \bar{I}_b(p)$, $p \in \mathcal{B}_{0.2q_0}(p_m)$. With the refined estimate of the peak position a larger neighborhood $\mathcal{B}_{2p_0}(p_q)$ is extracted over which the total flux is computed as $I_t = \sum_{p \in \mathcal{B}_{2q_0}(p_q)} \bar{I}(p)$. Also the diffraction limited total flux I_d on this neighborhood is computed from the theoretical diffraction pattern with unit peak intensity by oversampling the pixels by a factor 8. In the final step, the Strehl is computed by weighting the estimated peak intensity by the ratio of the measured and the theoretically computed total flux, i.e. $S = (I_q/I_t)I_d$.

4.6 Experimental results

Both the $\mathcal{H}_2$-optimal control approach and the common AO control law have been implemented on the experimental setup. After aligning the WFS, only 69 of the 127 micro-lenses are illuminated sufficiently to be used for wavefront sensing (see also Figure 4.3). This implies that the unreduced WFS signal consists of 138 channels. The geometry matrix G, specifying the relation between slope measurements and phase, is defined by adapting the well-know Fried configuration for a hexagonal grid. Just as for a rectangular grid, the position of the phase points is determined by shifting the hexagonal grid over half the pitch size. Projecting out the unobservable modes, leads to a reduced WFS signal of $n_s = 88$ independent

channels. During the experiments, the WFS exposure time has been adjusted to $t_e = 5\,\mathrm{ms}$. The WFS gain has been calibrated using the TT-mirror.

The real-time software is implemented in such a way that, for the given exposure time and sample frequencies, the condition $t_e \leq \tau_d$ is always satisfied with $d = 2$. From subsection 4.3.3 it is clear that, under these conditions, $\alpha = 0$ and the transfer function $\mathcal{H}(z)$ reduces to a pure delay, i.e. $\mathcal{H}(z) = z^{-2}H$. The choice for the reduced model structure for $\mathcal{H}(z)$ is supported by the exploratory FIR identification experiments, which show only one FIR coefficient that differs significantly from zero. An even more important argument for using the reduced model structure is the close agreement between measured and predicted WFS output. Validation experiments on a data set of $N = 1000$ samples show a mean *variance accounted for* (VAF) (Verdult and Verhaegen 2001) of more than 99.2% over the WFS channels. This implies that more than 99.2% of the variance of the measured WFS response $y(k)$ can be explained by the identified model. The identified transfer function provides therefore an accurate description of the true system.

The wavefront fitting error, in the experimental setup, is rather dominant. The DM has only 37 actuators and can only compensate the lowest spatial frequencies of the wavefront. Besides the small number of actuators, also the limited dynamic range and the applied input regularization contribute to a large fitting error. As a result, optimal control can only be expected to achieve a performance improvement at large Greenwood to sample frequency rations. Both control algorithms have been tested for Greenwood to sample frequency ratios in the range $f_G/f \in [0.015\ 0.43]$. The upper-bound on this range is rather large as a $f_G/f = 0.5$ implies that the characteristic frequency of the turbulence is equal to the Nyquist frequency. To verify if the residual wavefront error indeed depends only on the ratio of f_G and f, and not on their respective values, the f_G/f ratios have been obtained by considering different Greenwood and sample frequency combinations. Some of these combinations give rise to the same or approximately the same f/f_G ratio. The sample rates used to obtain the different f/f_G ratios include $f \in \{4.44, 6.44, 8.33, 10.41, 12.5, 14.29, 16.67, 20.0\}$ Hz.

At each Greenwood to sample frequency ratio, the performance of the common control approach is determined using the same value for c_1 and c_2. The control parameters have been tuned to minimize $\hat{\sigma}_\epsilon^2$ at a f_G/f ratio of 0.043. This resulted in the values $c_1 = 0.48$ and $c_2 = 0.98$. In evaluating the reconstruction matrix R, the covariance matrix C_φ is computed theoretically assuming a perfect Kolmogorov spatial distribution (see Wallner 1983). Furthermore, it is assumed that the measurement noise has a covariance matrix of the form $C_\nu = \sigma_\nu^2 I$, where the variance of the noise σ_ν^2 is estimated from open-loop WFS data $y(k)$ from a static distortion. For a fair comparison, the same input regularization matrix Q has been used in both control approaches. The regularization matrix is chosen diagonal, i.e. $Q = \mathrm{diag}\{q_1, q_2, \ldots, q_{m_u}\}$, and the control effort weighting $q_i, i \in \{1, 2, \ldots, m_u\}$ on each of the actuators is tuned to avoid actuator saturation and to ensure that the mirrors stay within a linear range. This results in an additional weight on the actuators with a small dynamic range, like the actuators near the edge of the DM. The TT-mirror, on the other hand, does not require regularization. In the optimal control approach, the atmospheric disturbance model

$\mathcal{S}$ is identified on the basis of $N_s = 5500$ samples of open-loop WFS data $y(k)$. The number of block-rows of past and future measurement data, as used in the subspace algorithm, is 20 and the model order is chosen equal to $n_d = 256$.

The estimated mean-square residual phase error $\hat{\sigma}_\epsilon^2$ obtained in the different experiments is depicted in Figure 4.7. Each circle and each cross is the result of an experiment at a specified Greenwood to sample frequency ratio using the optimal and common control approach, respectively. As expected from the discussion in
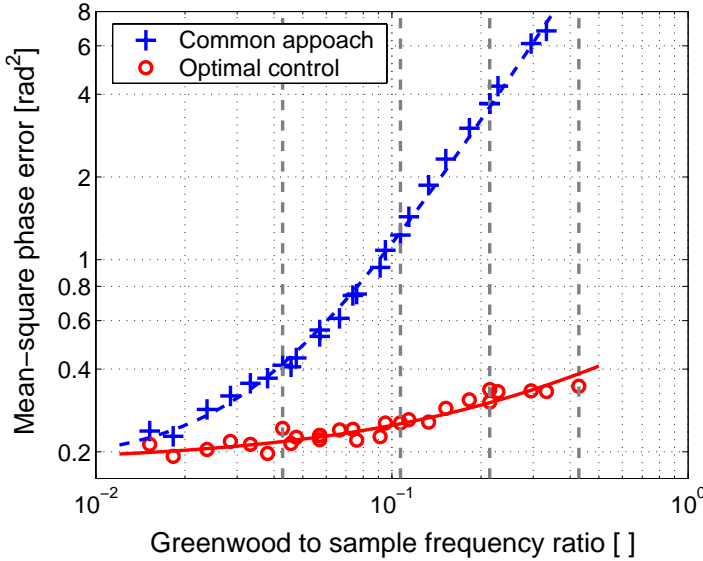

Figure 4.7: Mean-square phase error as a function of the Greenwood to sample frequency ratio f_G/f.

Section 4.5, the performance improvement increases with the Greenwood to sample frequency ratio. Whereas the improvement is rather low at low Greenwood to sample frequency ratios, a considerable improvement is observed at high ratios. The dashed line in Figure 4.7 represents the fit of the error model (4.25) to the residual phase error $\hat{\sigma}_\varepsilon^2$ obtained with the common control approach. The corresponding fit for the optimal control approach is depicted by the solid line. Since the mean-square residual phase error for the optimal control approach shows a rather weak dependence on the Greenwood to sample frequency ratio, the estimate of the parameters a_1 and a_2 is quite sensitive to a slight variation in a_0. For this reason, optimization over a_0 is omitted and its value is fixed to the value of a_0 found for the common control approach. This is a reasonable assumption, since the fitting error in both approaches should be the same. The fitted error model

obtained for both control approaches is given by

$$(\text{common}) \quad \hat{\sigma}_\varepsilon^2 \approx 0.1833 + 44.21 \left(\frac{f_G}{f} \right)^{1.6616}$$

$$(\text{optimal}) \quad \hat{\sigma}_\varepsilon^2 \approx 0.1833 + 0.3853 \left(\frac{f_G}{f} \right)^{0.7642}$$

Figure 4.7 shows that the above relations provide an accurate fit to the mean-square residual phase error observed in the experiments. Furthermore, the fitted exponent a_2 for the common control approach is in close agreement with the theoretical value $5/3$ (≈ 1.667). These observations support the error model (4.25), and show that optimal control is indeed effective in reducing the temporal error. Since at low Greenwood to sample frequency ratios the fitting error becomes the limiting factor, only little can be gained by optimal control in this regime.

To prove that the performance at low Greenwood to sample frequency ratios is indeed limited by the fitting error, it is useful to estimate the error on the basis of the available open-loop data $y(k)$. This is achieved by first reconstructing the uncorrected wavefront as $\hat{\varphi}(k) = \Sigma_1^{-1} y(k)$. Since the operator $H_Q^\dagger$ can be interpreted as the projection of the open-loop wavefront estimate on the actuator space, the actuator commands corresponding to this wavefront estimate can be computed as $\hat{u}(k) = H_Q^\dagger u(k)$. The computed actuator commands $\hat{u}(k)$ are then used to determine the applied wavefront correction $\hat{\varphi}_m(k) = H\hat{u}(k)$. Neglecting all dynamics, the fitting error can now estimated as the mean-square error of $\hat{\varepsilon}(k) = \hat{\varphi}(k) - \hat{\varphi}_m(k)$. To demonstrate the effect of the input regularization, $\hat{u}(k)$ has been computed both with and without input regularization. Furthermore, the effect actuator saturation has been investigated by chopping off the signals that are out off range. The averaged estimated mean-square fitting error as well as the standard deviation over the different data sets are shown in Table 4.1.

regularization	saturation	$\hat{\sigma}_f^2$	std
no	no	0.087	0.003
no	yes	0.170	0.009
yes	yes	0.197	0.007

Table 4.1: Estimate of DM fitting error

The table shows that when accounting for both the input regularization and the actuator saturation the estimated fitting error is in close agreement with the constant offset a_0 obtained from the error model. The fitting error estimated on the basis of the open-loop WFS data is within 2 standard deviations of the estimated value of a_0. Furthermore, the Table 4.1 shows that actuator saturation almost doubles the observed fitting error. The additional increase of the fitting error by accounting for both regularization and saturation, is rather small since both error contributions are strongly correlated in the sense that the regularization parameter Q has tuned to avoid actuator saturation. Finally, it is clear that in

AO systems where actuator saturation is no issue, the fitting error is smaller, and optimal control may already become favorable at lower Greenwood to sample frequency ratios.

As pointed out in the previous section, the relative improvement in cost function $\hat{J}_c/\hat{J}_o$, the FWHM of the corrected image and Strehl have been used as additional measures to compare the performance of both control algorithms. The performance measures have been computed for the Greenwood to sample frequency ratios indicated by the vertical dashed lines in Figure 4.7 and are summarized in Table 4.2. Furthermore, also the mismatch error $\Psi(p_0, g_1, g_2, g_3) \times 100\%$, obtained by fitting the observed intensity to the Airy pattern is included in the table. The averaged or long exposure images $\bar{I}(p)$ used in computing the FWHM and the Strehl ratio are depicted in Figure 4.8. Since the science object in the setup resembles a point source, the averaged image can be interpreted as the point spread function (PSF) obtained with AO wavefront correction.

f_G/f	Optimal control			Common approach			Relative
	FWHM	Ψ	Strehl	FWHM	Ψ	Strehl	$\hat{J}_c/\hat{J}_o$
$(\times 10^{-1})$	[px]	[%]	[ ]	[px]	[%]	[ ]	[ ]
0.43	14.5	0.32	0.85	14.5	0.33	0.79	1.70
1.1	14.5	0.26	0.85	14.5	6.14	0.57	4.66
2.1	14.5	0.28	0.81	27.6	12.5	0.20	11.7
4.3	14.6	0.29	0.78	45.1	6.18	0.12	19.3

Table 4.2: Performance comparison control algorithms

The results in Table 4.2 and Figure 4.8 are consistent with the previous observations. They show that optimal control is able to achieve a considerable performance improvement at large Greenwood to sample frequency ratios, while the performance improvement at low Greenwood to sample frequency ratios is rather modest. When considering the different performance measures, especially the relative improvement in cost function $\hat{J}_c/\hat{J}_o$ is large. This is to be expected as the $\mathcal{H}_2$-optimal controller is especially designed to minimize this criterion. Moreover, this suggests that the performance gain with respect to the other criteria may be even larger if no input regularization is needed. The table shows that the optimal control is also able to achieve a performance improvement with respect to the optical performance criteria. Especially at large Greenwood to sample frequency ratios, optimal control gives rise to a considerable improvement of the observed long exposure image. The observed image has a much smaller FWHM, a higher Strehl and smaller mismatch error with the theoretical Airy pattern. This is confirmed by the images depicted in Figure 4.8.

Figure 4.9 shows the normalized encircled energy achieved with both the optimal and the common AO control approach for different Greenwood to sample frequency ratios. Also this figure shows that optimal control is able to improve the science image and that the gain in performance increases with the Greenwood
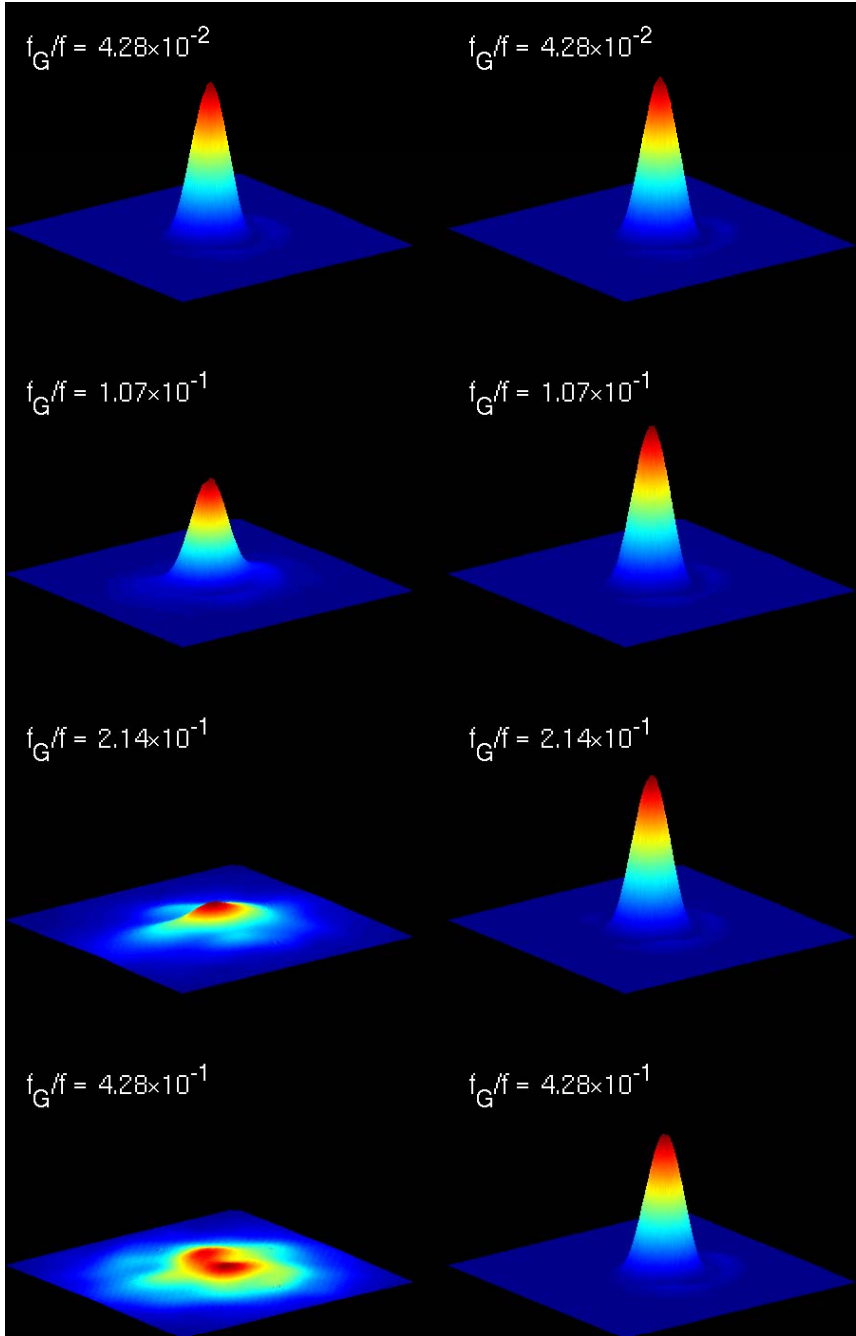
Figure 4.8: The averaged image $\bar{I}(p)$ obtained from the science camera when using AO, recorded at different Greenwood to sample frequency rations f_G/f. The left column shows the images obtained with the classical AO control approach, while the images in the right column correspond to the optimal control approach. Since the science object in the setup resembles a point source, the averaged image can be interpreted as the point spread function obtained with AO wavefront correction.

Figure 4.9: Normalized encircled energy for different Greenwood to sample frequency ratios.

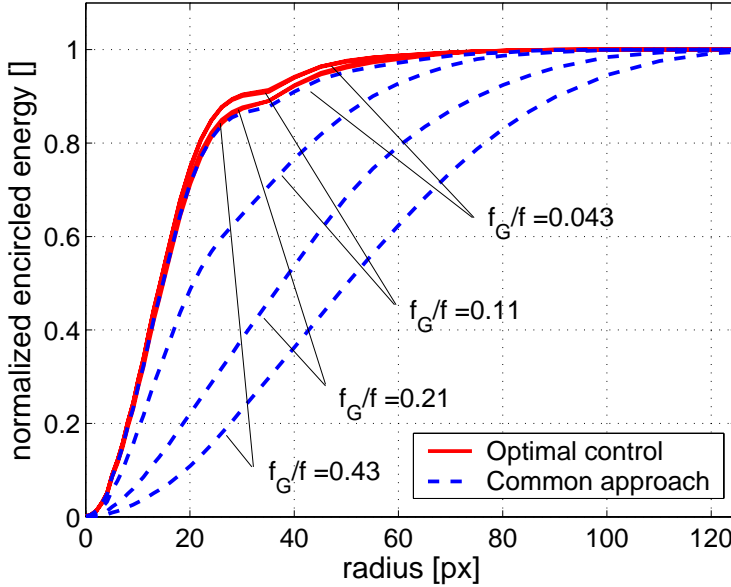to sample frequency ratio. Whereas the encircled energy curves obtained with the common control approach indicate at a rather spread out intensity pattern, optimal control gives rise to a science image in which a larger fraction of the incident energy is concentrated in the central core. Finally note that both the FWHM and the normalized encircled energy, obtained with the optimal control approach, are hardly influenced by Greenwood to sample frequency ratio. This is in close agreement with the observation that the mean-square residual phase error shows a weaker dependence on the Greenwood to sample frequency ratio for the optimal than the for the common control approach. Optimal control is therefore also attractive from the viewpoint of performance robustness to variations in atmospheric turbulence conditions.

4.7   Conclusions

In this paper, we have demonstrated a recently proposed data-driven $\mathcal{H}_2$-optimal control approach (Hinnen et al. 2006) on an experimental setup. In contrast to existing AO control approaches, this approach does not assume any form of decoupling and has the potential to exploit the spatio-temporal correlation imposed by the Taylor hypothesis. Computing the optimal controller requires an accurate model of the transfer function from actuator inputs to WFS outputs. By analyzing the dynamic behavior of the WFS, it has been shown that if the wavefront correction device can be considered to be static, the scalar-dynamics of the discrete-time transfer function from actuator inputs to WFS output can be modeled as an integer

number of samples delay followed by a two taps impulse response. This model structure fits perfectly in the proposed $\mathcal{H}_2$-optimal control strategy as it allows the optimal controller to be computed analytically. A data-driven identification approach, based on separable least squares optimization, has been developed to identify a transfer function of the desired structure from input-output data.

The performance of the data-driven $\mathcal{H}_2$-optimal control approach has been compared with a commonly applied AO control law. In this comparison different performance criteria based on measurements from both the science camera and WFS have been used. The considered performance criteria include an estimate of the mean-square residual wavefront error, the FWHM, the Strehl and normalized maximum encircled energy. The experiments show that optimal control is able to achieve a performance improvement, with respect to each of these criteria, and that the gain in performance increases with the Greenwood to sample frequency ratio. A careful analysis of the dominant error source in the setup reveals that the gain in performance can be attributed to a reduction of the temporal error. The total mean-square wavefront error can be decomposed in a wavefront fitting error and a temporal error. The wavefront fitting error depends on the turbulence conditions and the applied wavefront correction device and cannot be influenced by the controller. The temporal error, on the other hand, is strongly influenced by the controller. By exploiting the spatio-temporal correlation, the optimal controller is able to reduce the effect of temporal error. Since the temporal error is an exponentially increasing function of the Greenwood to sample frequency ratio, there is a lot to be gained at high ratios. At low ratios the performance is almost completely determined by the fitting error. Even though the measurement noise in the experimental setup is almost negligible, it may be an error source of considerable importance in real-life telescope systems. Since exploiting the spatio-temporal correlation is expected to reduce the effect of measurement noise, optimal control may also be beneficial in low signal to noise situations. The gain in performance under these conditions remains a point of further investigation.

References

Åström, K. J., Wittenmark, B., 1997. Computer-Controlled Systems, Theory and Design. Prentice-Hall International Inc.

Chen, T., Francis, B., 1995. Optimal Sampled-Data Control Systems. Springer-Verslag, London, Berlin.

Ellerbroek, B. L., Van Loan, C., Pitsianis, N. P., Plemmons, R. J., Nov. 1994. Optimizing closed-loop adaptive-optics performance with use of multiple control bandwidths. Journal of the Optical Society of America A 11 (11), 2871–2886.

Fried, D. L., Jul. 1990. Time-delay-induced mean-square error in adaptive optics. Journal of the Optical Society of America A 7 (7), 1224–1225.

Gavel, D. T., Wiberg, D., Feb. 2003. Towards strehl-optimizing adaptive optics controllers. In: Wizinowich, P. L., Bonaccini, D. (Eds.), Adaptive Optical System Technologies II. Vol. 4839 of Proceedings of SPIE. Hawaii, pp. 890–901.

Gendron, E., Léna, P., 1994. Astronomical adaptive optics. Astronomy and Astrophysics: I. Modal control optimization 291, 337–347.

Gendron, E., Léna, P., Sep. 1996. Single layer atmospheric turbulence demonstrated by adaptive optics observations. Astrophysics and Space Science 239 (2), 221–228.

Golub, G. H., Pereyra, V., Apr. 1973. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables are separate. SIAM Journal on Numerical Analysis 10 (2), 413–432.

Hardy, J. W., 1998. Adaptive Optics for Astronomical Telescopes. Oxford series in optical and imaging sciences. Oxford University Press, New York.

Hinnen, K. J. G., Doelman, N. J., Verhaegen, M., Aug. 2005a. $\mathcal{H}_2$-optimal control of an adaptive optics system: Part II, closed-loop controller design. In: Tyson, R. K., Lloyd-Hart, M. (Eds.), Astronomical Adaptive Optics Systems and Applications II. Vol. 5903 of Proceedings of SPIE. San Diego, California, United States, pp. 86–99.

Hinnen, K. J. G., Verhaegen, M., Doelman, N. J., Aug. 2005b. $\mathcal{H}_2$-optimal control of an adaptive optics system: Part I, data-driven modeling of the wavefront disturbance. In: Tyson, R. K., Lloyd-Hart, M. (Eds.), Astronomical Adaptive Optics Systems and Applications II. Vol. 5903 of Proceedings of SPIE. San Diego, California, United States, pp. 75–85.

Hinnen, K. J. G., Verhaegen, M., Doelman, N. J., Mar. 2006. A data-driven $\mathcal{H}_2$-optimal control approach for adaptive optics, submitted to IEEE Transactions on Control Systems Technology.

Law, N. F., Lane, R. G., May 1996. Wavefront estimation at low light levels. Optics Communications 126, 19–24.

Le Roux, B., Conan, J.-M., Kulcsár, C., Raynaud, H.-F., Mugnier, L. M., Fusco, T., Jul. 2004. Optimal control law for classical and multi-conjugate adaptive optics. Journal of the Optical Society of America. A 21 (7), 1261–1276.

Ljung, L., 1999. System Identification: Theory for the User, 2nd Edition. Prentice Hall PRT, Upper Saddle River, NJ.

Looze, D. P., Nov. 2005. Realization of systems with CCD-based measurements. Automatica 41 (11), 2005–2009.

Looze, D. P., Becker, O., Kasper, M., Hippler, S., Dec. 1999. Optimal compensation and implementation for adaptive optics systems. In: Proceedings of the 38th Conference of Decision & Control. Phoenix, Arizona, United States, pp. 1715–1720.

Looze, D. P., m. Kasper, Hippler, S., ans R. Weiss, O. B., 2003. Optimal compensation and implementation for adaptive optics systems. Experimental Astronomy 15 (2), 67–88.

Paschall, R. N., Anderson, D. J., Nov. 1993. Linear quadratic gaussian control of a deformable mirror adaptive optics system with time-delayed measurements. Applied Optics 32 (31), 6347–6358.

Roberts Jr., L. C., Perrin, M. D., Marchis, F., Sivaramakrishnan, A., Makidon, R., Christou, J. C., Macintosh, B. A., Poyneer, L. A., van Dam, M. A., Troy, M., 2004. Is that really your strehl ratio? In: Calia, D. B., Ellerboek, B. L. (Eds.), Advancements in Adaptive Optics. Vol. 5490 of Proceedings of SPIE. pp. 504–512.

Roddier, F., 1999. Adaptive Optics in Astronomy. Cambridge University Press.

Taylor, G. I., Feb. 1938. The spectrum of turbulence. Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences 164 (919), 476–490.

Thomas, S., Oct. 2004. Optimized centroid computing in a shack-hartmann sensor. In: D. B. Calia, B. L. Ellerbroek, R. R. (Ed.), Advances in Adaptive Optics. Vol. 5490 of Proceedings of SPIE. Glasgow, Scotland, United Kingdom, pp. 1238–1246.

van Dam, M. A., Le Mignant, D., Macintosh, B. A., Oct. 2004. Performance of Keck observatory adaptive optics system. Applied Optics 43 (29), 5458–5467.

Vdovin, G., Sarro, P. M., Jun. 1995. Flexible mirror micro-machined in silicon. Applied Optics 34 (16), 2969–2972.

Verdult, V., Verhaegen, M., Dec. 2001. Identification of multivariable bilinear state space systems based on subspace techniques and separable least squares optimization. International Journal of Control 74 (18), 1824–1836.

Wallner, E. P., Dec. 1983. Optimal wave-front correction using slope measurements. Journal of the Optical Society of America 73 (12), 1771–1776.

Weyrauch, T., Vorontsov, M., Bifano, T., Hammer, J., Cohen, M., Cauwenberghs, G., Aug. 2001. Microscale adaptive optics: wave-front control with a μ-mirror array and a vlsi stochastic gradient descent controller. Applied Optics 40 (24), 4243–4251.

5 | CHAPTER

Conclusions and Recommendations

This final chapter consists of two sections. In the first section we will briefly summarize the main conclusions that can be drawn from the research presented in this thesis. This discussion gives rise to a number of suggestions and recommendations for further research, which will be elaborated in the subsequent section.

5.1 Conclusions

In this thesis we have developed an optimal control strategy for adaptive optics (AO). The AO control design problem has been approached as multi-variable disturbance rejection problem, using general concepts from the field of systems and control. More specifically, the followed control strategy is one of identifying dynamic models for the atmospheric wavefront distortions and the AO system, followed by an minimum-variance optimal control design. It has been shown, that such an optimal control strategy can be used to improve the wavefront correction performance of current AO systems. Since the applied strategy is very general, many of the developed concepts can also be applied outside the field of AO.

The problem of finding a disturbance model that accurately describes the turbulence dynamics is the most fundamental step in determining the optimal controller. From a theoretical point of view, the atmospheric distortions can be described in terms of frozen layers with a Kolmogorov spatial distribution. Considering a single point in the aperture, this implies that the power spectrum of the temporal evolution of the wavefront distortions follows a $-11/3$ power law. Since the theoretical model provides a description in terms of power spectra, a subspace based algorithm for estimating the minimum-phase spectral factor from spectral data has been derived. The proposed spectral factor approximation algorithm has been successfully applied to approximate the non-rational Kolmogorov spectrum over a certain frequency region of interest, specified by a set of equidistantly spectral samples. Such a finite-dimensional approximation of the Kolmogorov spec-

trum could be useful in replacing the conventional design procedure of the parallel feedback loops in classical AO control with a model based approach that explicitly accounts for the temporal dynamics in the wavefront. A full multi-variable optimal control design procedure, however, requires an atmospheric disturbance model that also describes the mutual relation between the WFS channels. Even though the proposed spectral factor approximation algorithm is suitable for matrix valued power spectra, its computational complexity prohibits one to consider more than five to ten independent channels using a general purpose PC. The main complication in this approach is to guarantee the preservation of the positive realness property which is necessary to perform the spectral factorization step. Despite the improved efficiency achieved by using conic linear programming, the spectral factorization remains too time consuming to scale the algorithm to the number of channels used in a typical AO system.

To avoid the problem of spectral factorization, a dedicated subspace identification algorithm that estimates the atmospheric disturbance model directly from open-loop WFS data, has been developed. The use of data-driven identification is attractive as it yields a good match with the prevalent turbulence conditions without the need for accurate estimates of physical parameters, such as the wind speed and wind direction. One of the key features of the proposed algorithm is that it provides a direct estimate of the system matrices of the one-step ahead predictor model that corresponds to the minimum-phase spectral factor. In this approach the minimum-phase requirement on the atmospheric disturbance model is translated into a stability requirement on the system matrices of the identified Kalman predictor model. This requirement can be easily checked and if necessary enforced by using Schur re-stabilization.

Before modeling the WFS signal it is represented in an reduced basis, which parametrizes only the observable part of the wavefront. This re-parametrization of the sensor space has the advantage that it reduces the dimension of the identification problem and improves the numerical conditioning. To arrive at an efficient implementation, both in terms of the number of flops and the memory requirements, the different steps of the algorithm have been expressed in terms of the R-factor of a single RQ factorization used for data-compression. A further efficiency improvement is achieved by exploiting the displacement structure of the block-Hankel matrix in computing the R-factor. The resulting subspace identification algorithm is sufficiently efficient to identify a full multi-variable ARMA disturbance model for AO systems with up to a few hundred degrees of freedom on a general purpose PC with a $3\,\mathrm{GHz}$ Intel Pentium IV processor and 512 Mb of internal memory. The spatio-temporal correlation imposed by atmospheric wavefront distortions that satisfy the Taylor hypothesis can be described as a special case of a such general multi-variable disturbance model. The proposed subspace identification algorithm has been successfully demonstrated on both open-loop WFS data obtained from an experimental setup as well as on real turbulence measurements obtained from the JOSE seeing-monitor at the William Herschel Telescope.

Besides an atmospheric disturbance model, the optimal control design strategy also requires an accurate model of the transfer function from actuator inputs to WFS outputs. By analyzing the dynamic behavior caused by the integrating

action of the WFS camera, it has been shown that if the deformable mirror (DM) has a time constant that is short compared to the WFS exposure time, the scalar dynamics of this transfer function reduces to an integer number of samples delay followed by a two taps impulse response. A data-driven identification algorithm, based on separable least squares optimization, has been developed for identifying the model parameters of the derived model structure. In the case that the DM dynamics cannot be neglected, more general identification tools have been used to identify the AO system dynamics.

Given the identified model for the wavefront distortions and the AO system dynamics, the AO control problem has been expressed in an $\mathcal{H}_2$-optimal control framework. By making a clear distinction between performance and measurement outputs, $\mathcal{H}_2$-optimal control provides an attractive framework to account for the fact that the WFS is not able to directly measure the phase, while it is the mean-square residual phase error which needs to be minimized. Computing the $\mathcal{H}_2$-optimal controller generally involves the numerical solution of two Riccati equations. By using the Youla parametrization to open the control loop, it has been shown that as a result of the minimum-phase property of the identified disturbance model at least one of the Riccati equations can be eliminated. Also the second Riccati equation can be eliminated if the model of the AO system dynamics has a known inner-outer factorization. It has been shown that this is in particular the case for the previously derived model structure of an AO system in which the only dynamics can be traced back to the WFS camera. The so obtained analytical expressions, provide a straightforward way to compute the optimal controller from the atmospheric disturbance model.

Furthermore, by comparing the optimal controller with the equivalent optimal feedforward controller, it has been shown that the controller can be interpreted as wavefront predictor followed by a dynamic filter. The task of the wavefront predictor is to compensate for pure time delays, while the dynamic filter projects the estimate on the actuator space and makes a trade off between the objective of minimizing the fitting error and the control effort. This interpretation of the optimal controller is in particular useful as it shows that optimal control approach basically reduces to a modeling and identification problem for wavefront prediction. Hence, from a theoretical point of view, the wavefront prediction step on the basis of the identified atmospheric disturbance model could be in principle replaced by any other suitable wavefront prediction scheme. The proposed subspace identification algorithm, together with the analytical expression for the optimal controller, provides a non-iterative way to go from open-loop WFS data to closed-loop controller design.

The proposed data-driven $\mathcal{H}_2$-optimal control design approach has been demonstrated on an AO laboratory setup with a single layer Kolmogorov type of turbulence simulator. In this validation study the performance of the optimal controller has been compared with a classical control law consisting of a minimum-variance wavefront reconstructor followed by a first-order lag filter as a temporal compensator. The validation experiments demonstrate that optimal control is effective in reducing the temporal error, which is the error caused by both pure time delays and bandwidth limitations in the control system. Both control strategies

have been compared over a wide range for Greenwood to sample frequency ratios, by estimating the mean-square residual phase error. The experiments show a considerable performance improvement at large Greenwood to sample frequency ratios, while the gain at low ratios is modest.

This behavior has been explained by decomposing the residual wavefront error in the setup in two main contributions, i.e. the wavefront fitting error and the temporal error. Since the spatial distribution of the phase distortions of the turbulence simulator are fixed, the wavefront fitting error can be considered constant. The temporal error on the other hand, is an exponentially increasing function of the Greenwood to sample frequency. At low Greenwood to sample frequency ratios, the fitting error is the dominant error source so that there is only little room for improvement by control. When the sample frequency is sufficiently high and measurement noise is no issue, the change in wavefront from sample to sample is small, which renders the classical control law close to optimal. However, since for large Greenwood to sample frequency ratios the temporal error becomes the dominant error source there is a lot to be gained by optimal control. By exploiting the spatio-temporal correlation in the wavefront, optimal control is able to anticipate future wavefront distortions and to reduce the temporal error. Optimal control does not only achieve an performance improvement with respect to the mean-square residual phase error, but also with respect to optical performance criteria such as the Strehl ratio, the FWHM and the encircled energy of the corrected spot. Also with respect to these criteria, the benefit of optimal control becomes more and more apparent at higher Greenwood to sample frequency ratios.

In conclusion, the validation experiments have demonstrated that optimal control is especially effective in reducing the temporal error, which becomes the dominant error at high Greenwood to sample frequency ratios. From a practical point of view, this implies that there is potentially a lot to gain in heavy turbulence conditions including high wind speeds and small Fried parameters, and under low light level conditions where measurement noise prohibits the use of a high sampling frequency. Even though the effect of measurement noise has not explicitly been considered in the validation experiments, it is clear that optimal control may also be beneficial in reducing this error source. Since optimal control includes the spatio-temporal correlation, past and neighboring WFS measurements may all contribute to improve the prediction of the wavefront distortion. The Greenwood to sample frequency ratios considered in the validation experiments are relatively high compared to those encountered in real telescopes. Considering typical values for the wind speed, Fried parameter and sample frequency of respectively $\bar{v} = 10\,\mathrm{m/s}$, $r_0 = 0.2\,\mathrm{m}$, and $f = 500\,\mathrm{Hz}$, gives rise to a Greenwood to sample frequency ratio of $f_G/f = 0.04$. This is indeed at the lower end of the considered Greenwood to sample frequency range.

The reason for considering these high ratios is twofold. First and foremost, since the experimental setup uses a DM with only a small number of actuators the fitting error is relatively large. Also the input regularization used to avoid actuator saturation contributes to a higher fitting error. The relatively large fitting error implies that it takes higher Greenwood to sample frequency ratios before the effect of the temporal error becomes noticeable. To demonstrate the benefit of optimal

control it is therefore necessary to consider higher ratios. Since most real telescopes have more actuators and actuator saturation is usually not a problem, the contribution of the fitting error will be smaller, and optimal control is expected to become already favorable at lower and hence more realistic Greenwood to sample frequency ratios. The temporal error is indeed one of the dominant error sources in for instance the Keck telescope (van Dam et al. 2004). Furthermore, note that even on the experimental setup the reduction in mean-square residual wavefront error achieved by optimal control at $f_G/f = 0.04$ is still in the order 45%.

The second reason to compare both algorithms over a wide range of ratios is that it is interesting to consider the possible increase in Greenwood to sample frequency ratio for achieving a certain level of performance. On the experimental setup, optimal control achieves the same level of performance at a Greenwood to sample frequency ratio of $f_G/f = 0.45$ as the classical approach at a ratio of $f_G/f = 0.04$. This implies that, at least in this particular case, optimal control is able to achieve the same performance at less than one tenth of the sample frequency. Lowering the sample frequency is especially attractive as this enables longer WFS exposure times, which will reduce the measuring noise. By reducing the effect of measurement noise, optimal control is very promising in raising the limiting magnitude of the guide star needed for closing the control loop. In this way, control may help to improve the sky coverage of current AO systems.

Finally, it is interesting to note that the performance achieved with the optimal control approach shows a much weaker dependence on the Greenwood to sample frequency ratio. Optimal control may therefore also be used to improve the performance robustness of the AO system with respect to fluctuations in turbulence conditions. It is very difficult to predict how these results precisely translate to real telescopes, however the experiments have clearly demonstrated the benefit of using optimal control.

5.2   Recommendations

Even though the results presented in this thesis have clearly demonstrated the opportunities and advantages of improved control for AO, the research within this area is far from finished. In view of experience and insights gained over the course of time, the following suggestions are made for future research:

- The data-driven $\mathcal{H}_2$-optimal control approach presented in this thesis uses a reduced representation of the WFS signal space that parametrizes only the observable part of the wavefront. To achieve a further reduction in the computational complexity, it may be useful to follow a similar procedure to remove the uncontrollable part of the wavefront. For a static mirror, this can be achieved in a completely analogous way by deriving a projection matrix from the SVD of the influence matrix. Introducing a minimal representation for the WFS signal in the intersection of the observable and controllable part of the wavefront may lead to a considerable reduction in the number of

WFS signals that has to be modeled. Note that after removing the uncontrollable modes, the dimension of the WFS is always smaller than the number of control inputs and the number of reconstructed wavefront points. Since the AO system is not able to influence the uncontrollable part of the wavefront, this part of the wavefront does not need to be predicted and can hence be disregarded without any loss of performance.

- Another interesting extension of the proposed $\mathcal{H}_2$-optimal control strategy is related to the analytical expression for the closed-loop optimal controller. This expression has been derived for the case that the AO system can be characterized by a scalar dynamic transfer function consisting of an integer number of samples delay and a two taps impulse response. This model structure has shown to be valid for AO systems where the DM dynamics can be neglected and the only dynamics derives from the integrating action of the WFS camera. However, in deriving the analytical expression for the optimal controller the only property required to avoid the second Riccati equation is that the transfer function from control input to WFS output has a known inner-outer factorization. This implies that the class of AO systems for which the optimal controller can be computed analytically is larger than that of AO systems with a static DM. In particular, all AO systems with a transfer function that is minimum-phase can be considered in this way. Also for all AO systems with a scalar dynamic transfer function with known zeros it is possible to derive an analytical expression for the optimal controller, and if the zeros are unknown, the large scale Riccati equation can be replaced by a Riccati equation for a SISO system to yield an semi-analytical solution. Note that this is particular useful for AO systems of which the DM dynamics can be described by a common transfer function for each of the actuators.

- For the practical application of the proposed control approach it would be interesting to make the subspace algorithm suitable for closed-loop identification. In its current configuration the algorithm uses open-loop WFS data to identify the atmospheric disturbance model. A first important reason for considering closed-loop identification might be the limited dynamic range of the WFS. In many AO systems, the WFS has been optimized for sensitivity rather than for a large linear dynamic range. This implies that for these systems it is impossible to collect open-loop WFS without being distorted. Since the WFS excursions during closed-loop operation are much smaller, closed-loop identification may provide a way out to resolve this issue. In this approach one could think of using the classical AO control as initialization for collecting closed-loop data. The closed-loop data are then used to identify an atmospheric disturbance model and to update the controller.

  A second reason for considering closed-loop identification is that it provides a way to deal with slow variations in the turbulence conditions. As pointed out before, the turbulence statistics may change on a time-scale of a few minutes. To guarantee close to optimal performance over longer time scales, it becomes necessary to update the atmospheric disturbance model on a regular basis. A closed-loop identification scheme may enable a regular model update without interrupting the observations. In this perspective, it would

also be interesting to compare the algorithm with the adaptive control strategies discussed in Section 1.4.5. A quasi-adaptive scheme based on the offline methods developed in this thesis might be attractive as it uses a more general model structure and is not sensitive to convergence problems. Adaptive control schemes on the other hand might be more suitable for tracking fast changes in the turbulence conditions.

- From an implementational point of view, the proposed control strategy consists of an off-line part concerned with the control design and an on-line part that implements the designed controller. Especially the computational complexity of the off-line part increases rapidly with the number of degrees of freedom of the AO system, i.e. the number of WFS and DM channels. An important point for further research is therefore to increase the number of channels that can be handled by the optimal control approach. The introduction of the reduced basis as discussed in the first item of this list, may well lead to an efficiency improvement but is certainly not be able to resolve the issue. As explained before, the problem of identifying an suitable atmospheric disturbance model is currently the most computational intensive step. For this reason most of the research effort in scaling up the control strategy should be devoted to the modeling and identification part.

  Even though the proposed subspace algorithm provides very satisfying results for the considered size of AO systems, it is not very suitable for dealing with very large systems. Like any subspace identification algorithm, the algorithm is based on standard matrix operations performed on a block-Hankel matrix filled with measurement data. Under the assumption that the choice for the number of block-rows of past and future data, p and f, scales linearly with the number of channels m_s, both dimensions of the block-Hankel matrix grow quadratically in m_s. With the efficient RQ-factorization, this gives rise to a total computational complexity that scales as the fifth power of the number of channels, i.e. $\mathcal{O}(m_s^5)$. Furthermore, note that under these assumptions also the number of time samples for filling the block-Hankel matrix increases quadratically with m_s. This implies that more and more data are needed for identifying the atmospheric disturbance model which may eventually lead to a contradiction with the stationarity assumption. The above arguments demonstrate that, even though there might be still some room for improvement, it is not very likely that subspace algorithms can be scaled to AO systems with hundreds to thousands of degrees of freedom. For such a system, the identification of a central atmospheric disturbance model is no longer feasible and a fundamentally different strategy has to be used. For large AO systems and high sample frequencies also the on-line computational complexity will become an issue. The on-line computational complexity is that of implementing a state-space system and scales therefore in the same way as the classical AO control approach when implemented as a matrix-vector multiply.

  One interesting approach to deal with both the on-line as well as the off-line increase in computational complexity might be the development of algorithms to identifying structured models that can be used in a decentralized

or distributed control design approach. Important research questions are then to quantify the loss in performance compared to the central solution and to find out if there is still any room for improvement over the classical control approach using an advanced minimum-variance wavefront reconstructor based on sparse matrix techniques.

- Apart from the simulations on the William Herschel telescope WFS data, all simulations and experiments have been performed on the basis of an atmospheric turbulence simulator which behaves as perfectly frozen. Since the proposed control strategy does not depend on the Taylor hypothesis, it would be interesting to gain more insight in its performance when the turbulence deviates from being frozen. Here one could think of simulations or experiments with multiple layers of frozen turbulence, moving with different wind speeds in different directions. More generally speaking, the AO control community may benefit from a generally available and well accepted simulation toolbox for testing control algorithms. Such a toolbox should contain well validated dynamic models of both the AO system components and the atmospheric wavefront distortions as well as tools for performance analysis. By defining a number of realistic and relevant simulation scenarios it becomes possible to compare the performance of different new developed control algorithms.

- The proof of the pudding in the eating. The research results presented in this thesis have demonstrated the opportunities and possible advantages of a control strategy that is able to account for the spatio-temporal correlation in the wavefront. Now the proposed data-driven $\mathcal{H}_2$-optimal control strategy has been successfully demonstrated on an AO laboratory setup, the ultimate test is to apply it on a real telescope. Even though the above research suggestions might be useful to improve different aspects of the developed techniques, in the end the only way to find out if advanced control can live up to the expectations is to demonstrate it in reality.

A | APPENDIX

$\mathcal{H}_2$-Optimal Control with a Quasi-Static Mirror

The goal of this appendix is to formalize the proof of Theorem 3.3. In Chapter 3 we have provided a brief outline of the line of reasoning used to arrive at Theorem 3.3. This outline was mainly intended to provide insight and can be seen as a sketch of the proof. In this appendix we will actually elaborate the intermediate steps used in deriving the analytical expressions for the optimal feedforward and feedback controller. To simplify the derivation, we will rely on a mixture of state-space and transfer function representations.

The remainder of this appendix is organized as follows. In Section A.1, we will first provide an overview of the different ways used to represent a linear time invariant (LTI) system. Furthermore, this section provides a review of a number of useful operations on state-space realizations. These operations will be used in Section A.2 to derive a state-space realization for the optimal feedforward controller $\mathcal{W}(z)$. Finally, in Section A.3, the derived expression for $\mathcal{W}(z)$ is used to compute the optimal feedback controller $\mathcal{C}(z)$.

A.1 Useful operations on state-space realizations

Before defining some operations useful in deriving the analytical expressions for the optimal feedforward and feedback controller, we will first introduce some additional notation. Let P denote a LTI system with state-space representation

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k), \end{aligned} \tag{A.1}$$

then the following notation will be used to characterize its input-output behavior. First of all, we will use the quadruple (A, B, C, D) to compactly refer to the state-space realization in (A.1). Furthermore, the transfer function corresponding to P

is given by $P(z) = D + C(zI - A)^{-1}B$, where z denotes the unit shift-forward operator in the time-domain and the complex indeterminate in the z-domain. For notational convenience, this transfer function will also be denoted as

$$P = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]. \tag{A.2}$$

Note that since (A.2) is defined as an alternative representation of the transfer function $P(z)$, the different entries are not restricted to pure rational matrices but are allowed to be rational functions of z themselves. For the interpretation of such a transfer function, one should resort to the definition of $P(z)$. Using the above definitions, the basic operation useful in deriving the state-space expressions for the optimal controller have been summarized in the list below. An more extensive overview of basic operations on state-space realization linear systems is given in Fraanje (2004). The first three operations can be proved quite straightforwardly by performing basic algebraic manipulations on the state-space equations (A.1) of the respective systems. The last operation, on the other hand, is a special case of the causal anti-causal split in Verhaegen (1997).

Similarity transformation: Consider a LTI system P with state-space realization (A, B, C, D) and let the system P' be defined as

$$P' = \left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array} \right], \tag{A.3}$$

with T a non-singular matrix. Then P and P' have the same input-output behavior and are said to be equal up to the similarity transformation T.

Pseudo-inverse of system: Consider a LTI system P with state-space realization (A, B, C, D) and let the system $P^\dagger$ be defined as

$$P^\dagger = \left[\begin{array}{c|c} A - BD^\dagger C & BD^\dagger \\ \hline -D^\dagger C & D^\dagger \end{array} \right], \tag{A.4}$$

then the following cases can be distinguished:

1. *The direct feed-trough matrix D is non-singular.* In this case, $D^\dagger = D^{-1}$ and $P^\dagger = P^{-1}$ is the inverse of P, i.e. $PP^{-1} = P^{-1}P = I$.
2. *The matrix $D^T D$ is non-singular.* In this case $D^\dagger = (D^T D)^{-1} D^T$ and $P^\dagger$ is the left pseudo-inverse of P, i.e. $P^\dagger P = I$.
3. *The matrix DD^T is non-singular.* In this case $D^\dagger = D^T (DD^T)^{-1}$ and $P^\dagger$ is the right pseudo-inverse of P, i.e. $PP^\dagger = I$.

Cascade of two systems: Consider two LTI systems, P_1 and P_2, with state-space realizations (A_1, B_1, C_1, D_1) and (A_2, B_2, C_2, D_2), respectively. Furthermore, let the number of inputs to P_1 be equal to the number of outputs of P_2, then the cascade connection $P_1 P_2$ has a state-space realization

$$P_1 P_2 = \left[\begin{array}{cc|c} A_2 & 0 & B_2 \\ B_1 C_2 & A_1 & B_1 D_2 \\ \hline D_1 C_2 & C_1 & D_1 D_2. \end{array} \right] \tag{A.5}$$

Causality operator for system products: Let P_1 and P_2 be two strictly stable LTI systems with state-space realization (A_1, B_1, C_1, D_1) and (A_2, B_2, C_2, D_2), i.e. all eigenvalues of A_1 and A_2 are located inside the unit-disk. Then causal part of the cascade of P_1 and the adjoint P_2^* is given by

$$[P_2^* P_1]_+ = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline D_2^T C_1 + B_2^T X A_1 & D_2^T D_1 + B_2^T X B_1 \end{array} \right] \qquad \text{(A.6)}$$

where, $[\cdot]_+$ denotes the causality operator and where X is the solution to the Sylvester equation $X - A_1 X A_2^T = B_1 B_2^T$.

The above operations in principle suffice to derive the analytical expressions for the optimal feedforward and feedback controller. This, however, would require the state-space realization of all relevant transfer function, i.e. of both $\mathcal{H}(z)$, $\mathcal{P}_{ew}(z)$ and the inner-outer and outer-inner factorizations of $\mathcal{P}_{eu}(z)$ and $\mathcal{P}_{yw}(z)$, respectively. Since the transfer matrix $\mathcal{P}_{ew}(z)$ and the outer-inner factorization $\mathcal{P}_{yw}(z) = \mathcal{P}_{yw,o}(z)\mathcal{P}_{yw,i}(z)$ as derived in Chapter 3 are in the form $P(z) = D + C(zI - A)^{-1}B$, their respective state-space realization can be easily read off. Also in deriving a state-space realization for the transfer functions $\mathcal{H}(z)$ and $\mathcal{P}_{eu}(z)$, there is no fundamental problem. However, since these transfer functions both depend on an undetermined delay z^{-d} (i.e. d is fixed, but still unknown), this would lead to a very cumbersome notation as well as a substantial increase in the order of systems used in the intermediate steps of the computation. For this reason, it would be attractive if the need for an explicit state-space realization of z^{-d} could be avoided. Indeed, this can be achieved by using a mixture of transfer function and state-space notation. The following Lemma provides a convenient expression for the causal part of a strictly stable LTI system P of which the outputs are advanced d samples in time. This result can be seen as a special case of the causality rule for system products (A.6) with $P_2 = z^{-d}$, but avoids the need for a state-space realization of z^{-d}.

Lemma A.1 (Causal part of time-shifted system) *Let P be a strictly stable LTI system with state-space realization (A, B, C, D). Then the causal part of the system obtained by shifting the system output $d \in \mathbb{N}, d > 0$ steps forward in time is given by*

$$\left[z^d P \right]_+ = \left[\begin{array}{c|c} A & B \\ \hline CA^d & CA^{d-1}B \end{array} \right] . \qquad \text{(A.7)}$$

Proof: Even though the lemma could be proved via the causality rule for system products (A.6), we will not pursuit this approach for the reasons explained above. Instead, the lemma is proved in a transfer function setting. Since the system P is strictly stable, $(zI - A)$ is boundedly invertible for $|z| \geq 1$. Hence, the transfer function $z^d P$ can be expanded into an infinite Neumann series as

$$\begin{aligned} z^d P &= z^d \left(D + C(zI - A)^{-1}B \right) \\ &= z^d D + z^{d-1}C(I - z^{-1}A)^{-1}B \\ &= z^d D + z^{d-1}C \sum_{i=0}^{\infty} z^{-i} A^i B \end{aligned}$$

Considering this expansion, applying the causality operator is equivalent to collecting the causal terms with z^i, $i \leq 0$. By selecting the proper terms and factoring out the common factor, we can write

$$
\begin{aligned}
\left[z^d P \right]_+ &= C \sum_{i=d-1}^{\infty} A^i B \\
&= C A^{d-1} B + C \sum_{i=0}^{\infty} z^{-i} A^{d+i} B \\
&= C A^{d-1} B + C A^d (zI - A)^{-1} B
\end{aligned}
$$

Expressing this transfer function in the form (A.2) finally gives desired result. □

A.2 Optimal feedforward controller or Wiener filter

In outlining the procedure used to arrive at Theorem 3.3, we have explicitly derived expressions for the transfer function $\mathcal{P}_{ew}(z)$, the inner-outer factorization of $\mathcal{P}_{eu}(z)$ and the outer-inner factorization of $\mathcal{P}_{yw}(z)$. In this section, these expressions will be used to compute the optimal feedforward controller $\mathcal{W}(z)$. By applying the operations introduced in the previous section, the different terms in the expression for the Causal Wiener filter in Lemma 3.1, i.e. equation (3.30), will be elaborated one by one to arrive at the state-space expression of $\mathcal{W}(z)$.

First consider the pseudo-inverse of the outer factor $\mathcal{P}_{yw,o}$, i.e. $\mathcal{P}_{yw,o}^{\dagger}$. In Chapter 3 it has been argued that since $\mathcal{P}_{yw}(z)$ is minimum-phase, it has the following valid inner-outer factorization $\mathcal{P}_{yw} = \mathcal{P}_{yw,o} \mathcal{P}_{yw,i}$, where $\mathcal{P}_{yw,o} = \mathcal{P}_{yw}$ and $\mathcal{P}_{yw,i} = I$. Hence, computing $\mathcal{P}_{yw,o}^{\dagger}$ is equivalent to computing $\mathcal{P}_{yw}^{\dagger}$. By transforming the transfer function (3.31) into the form (A.2) and applying the rule for inverting a system (A.4), $\mathcal{P}_{yw,o}^{\dagger}$ can be expressed as

$$
\mathcal{P}_{yw,o}^{\dagger} = \left[\begin{array}{c|c} A_d - K_d \Sigma_1 C_d & K_d \\ \hline R_v^{-1/2} \Sigma_1 C_d & R_v^{-1/2} \\ 0 & 0 \end{array} \right]. \tag{A.8}
$$

Let us now consider the causal factor $\left[\mathcal{P}_{eu,i} \mathcal{P}_{ew} \mathcal{P}_{yw,i} \right]_+$. From the previous discussion it is clear that $\mathcal{P}_{yw,i} = I$. Furthermore, the outer-inner factorization of $\mathcal{P}_{eu}$ as provided in equation (3.34) shows that two separate cases have to be considered, i.e. $|\alpha| \leq 1$ and $|\alpha| > 1$. Assume for the moment that $|\alpha| \leq 1$. In this case the adjoint inner-factor $\mathcal{P}_{eu,i}^* = z^d$ and therefore $\left[\mathcal{P}_{eu,i}^* \mathcal{P}_{ew} \mathcal{P}_{yw,i}^* \right]_+ = \left[z^d \mathcal{P}_{ew} \right]_+$. By transforming the transfer function (3.32) into the form (A.2) and applying Lemma A.1, we can hence write

$$
\left[\mathcal{P}_{eu,i}^* \mathcal{P}_{ew} \mathcal{P}_{yw,i}^* \right]_+ = \left[\begin{array}{c|ccc} A_d & K_d R_v^{1/2} & 0 \\ \hline C_d A_d^d & C_d A_d^{d-1} K_d R_v^{1/2} & 0 \\ 0 & 0 & 0 \end{array} \right]. \tag{A.9}
$$

In the case that $|\alpha| > 1$, computing the causal factor $[\mathcal{P}_{eu,i}\mathcal{P}_{ew}\mathcal{P}_{yw,i}]_+$ is a bit more involved. To avoid the notational problems in representing $\mathcal{P}_{eu,i}(z)$ in state-space form, the pure delay z^{-d} is split off and considered separately. In this way, we obtain

$$\mathcal{P}_{eu,i}(z) \quad = \quad \frac{z+\alpha}{z^d(\alpha z+1)}I = \frac{1}{z^d}\left(1/\alpha + \frac{1-\alpha^2}{z+1/\alpha}\right)I \qquad (A.10)$$

$$= \quad z^{-d}\left[\begin{array}{c|c} -(1/\alpha)I & I \\ \hline (1-1/\alpha^2)I & (1/\alpha)I \end{array}\right] \qquad (A.11)$$

Since $\left[\mathcal{P}_{eu,i}^*\mathcal{P}_{ew}\mathcal{P}_{yw,i}^*\right]_+ = [z^d[z^{-d}\mathcal{P}_{eu,i}^*\mathcal{P}_{ew}\mathcal{P}_{yw,i}^*]_+]_+$, the causal factor can be computed by first neglecting the delay and then applying Lemma A.1 to account for it afterwards. Applying the causality operator for system products (A.9) to the delay free part of (A.11) it gives rise to the following expression

$$\left[\left(\frac{z+\alpha}{\alpha z+1}\right)^*\mathcal{P}_{ew}\right]_+ = \left[\begin{array}{c|cc} A_d & K_d R_v^{1/2} & 0 \\ \hline 1/\alpha\, C_d + XA_d & * & 0 \\ 0 & 0 & 0 \end{array}\right], \qquad (A.12)$$

where X denotes the solution to the Sylvester equation $(1 + 1/\alpha)XA_d = (1 - 1/\alpha^2)C_d$ and $*$ denotes a matrix irrelevant for the remainder of the derivation as it drops out in accounting for the delay. By solving for XA_d and accounting for the delay, the state-space realization of $\left[\mathcal{P}_{eu,i}^*\mathcal{P}_{ew}\mathcal{P}_{yw,i}^*\right]_+$ appears to be precisely equal to the result obtained before, i.e. for all $\alpha \in \mathbb{R}$ the causal factor is given by (A.9).

Finally consider the pseudo-inverse of the outer factor $\mathcal{P}_{eu,o}$, i.e. $\mathcal{P}_{eu,o}^\dagger$. Also in this case we have to distinguish between the conditions $|\alpha| \leq 1$ and $|\alpha| > 1$. Like $\mathcal{P}_{yw,o}^\dagger$, the transfer function $\mathcal{P}_{eu,o}^\dagger$ is computed by first finding a state-space realization for the outer factor $\mathcal{P}_{eu,o}$ given in equation (3.34) and then applying the rule for system inversion (A.4). For $|\alpha| \leq 1$ this gives rise to

$$\mathcal{P}_{eu,o}^\dagger \quad = \quad \left[\begin{array}{c|c} 0 & I \\ \hline -\alpha H & -H \\ 0 & Q^{1/2} \end{array}\right]^\dagger \qquad (A.13)$$

$$= \quad \left[\begin{array}{c|cc} -G & -H_Q^\dagger & * \\ \hline -G & -H_Q^\dagger & * \end{array}\right], \qquad (A.14)$$

where $H_Q^\dagger \doteq (H^T H + Q)^{-1}H^T$ and $G \doteq -\alpha H_Q^\dagger H$. Like before, the asterisks $*$ are used to denote matrices that do not influence the final solution. Here, knowledge about these matrices is not required since the second part of the output equation of (A.9) is zero and we are only interested in the product of $\mathcal{P}_{eu,o}^\dagger$ and the causal factor. In a similar way, $\mathcal{P}_{eu,o}^\dagger$ for $|\alpha| > 1$ can be expressed as

$$\mathcal{P}_{eu,o}^\dagger(z) = \left[\begin{array}{c|c} 0 & I \\ \hline -H & -\alpha H \\ 0 & Q^{1/2} \end{array}\right]^\dagger . \qquad (A.15)$$

By introducing the definitions $\beta \doteq 1/\alpha$ and $\bar{H} \doteq \alpha H$ it is clear that equations (A.14) and (A.15) have precisely the same structure. As a result, it is possible to express the pseudo-inverse of $\mathcal{P}_{eu,o}$ for both $|\alpha| \leq 1$ and $|\alpha| > 1$ in the same way. More specifically, by extending the above definition of $H_Q^{\dagger}$ as in (3.36), the pseudo-inverse of $\mathcal{P}_{eu,o}$ for all $\alpha \in \mathbb{R}$ is given by (A.14).

After having derived the state-space realizations (A.8), (A.9) and (A.14) for $\mathcal{P}_{yw,i}^{\dagger}$, $\left[\mathcal{P}_{eu,i}^{*}\mathcal{P}_{ew}\mathcal{P}_{yw,i}\right]_{+}$ and $\mathcal{P}_{eu,o}^{\dagger}$, respectively, Lemma 3.1 can now be used to derive the state-space realization for the optimal feedforward controller $\mathcal{W}(z)$. By applying the cascade rule for two systems (A.5), the optimal feedforward controller $\mathcal{W}(z)$ can be expressed as

$$\mathcal{W}(z) = -\mathcal{P}_{eo,o}^{\dagger} \left[\begin{array}{c|c} \widetilde{A} & K_d \\ \hline C_d A_d^{d-1} \widetilde{A} & C_d A_d^{d-1} K_d \\ 0 & 0 \end{array} \right] = - \left[\begin{array}{cc|c} \widetilde{A} & 0 & K_d \\ F\widetilde{A} & -G & FK_d \\ \hline F\widetilde{A} & -G & FK_d \end{array} \right] , \quad \text{(A.16)}$$

where the matrix F is defined as $F \doteq H_Q^{\dagger} C_d A_d^{d-1}$. The derived expression for the optimal feedforward controller $\mathcal{W}(z)$ is in agreement with Theorem 3.3.

A.3   State-space realization feedback controller

In the previous section we have derived a state-space realization for the optimal feedforward controller $\mathcal{W}(z)$. As explained in Chapter 3, this can be used together with the Youla parametrization

$$\mathcal{C}(z) = \mathcal{W}(z)(I - \Sigma_1 \mathcal{H}(z)\mathcal{W}(z))^{-1}, \quad \text{(A.17)}$$

to derive a state-space realization for the optimal feedback controller $\mathcal{C}(z)$. To elaborate equation (A.17) it is useful to consider the system obtained by delaying the output of $\mathcal{W}(z)$ by one sample. Since the second part of the state-update equation and the output equation in (A.16) are equal, the one sample delay can be introduced by choosing the output equal to second part of the state, i.e.

$$z^{-1}\mathcal{W}(z) = \left[\begin{array}{cc|c} \widetilde{A} & 0 & K_d \\ F\widetilde{A} & -G & FK_d \\ \hline 0 & I & 0 \end{array} \right]. \quad \text{(A.18)}$$

Note that this way of incorporating the one sample delay has the advantage that it does not increase the system order. Using the above expression for $z^{-1}\mathcal{W}(z)$ and applying the cascade rule for two systems (A.5), we can write

$$I - \Sigma_1 \mathcal{H}(z)\mathcal{W}(z) = I - z^{-d+1} \left[\begin{array}{c|c} 0 & I \\ \hline \alpha\Sigma_1 H & \Sigma_1 H \end{array} \right] \mathcal{W}(z) \quad \text{(A.19)}$$

$$= \left[\begin{array}{ccc|c} \widetilde{A} & 0 & 0 & K_d \\ F\widetilde{A} & -G & 0 & FK_d \\ 0 & I & 0 & 0 \\ \hline 0 & -z^{-d+1}\Sigma_1 H & -z^{-d+1}\alpha\Sigma_1 H & I \end{array} \right] \quad \text{(A.20)}$$

where the $d - 1$ samples delay have been included in operator form in the output equation of the system. The above state-space realization for $I - \Sigma_1 \mathcal{H}(z)\mathcal{W}(z)$ is however not minimal as can be seen by explicitly writing out the state update and output equations. To this end let $x_1(k)$, $x_2(k)$ and $x_3(k)$ denote the three components of the state-vector which is partitioned in accordance with the partitioning of the state-space matrices in equation (A.20). Using this notation, we infer from the third state update equation that $x_3(k + 1) = x_2(k)$. By introducing a time-shifted version of the state x_2, i.e. $\bar{x}_2(k + 1) \doteq x_2(k)$, it is now possible to express the output $y(k)$ of the state-space equations as

$$
\begin{aligned}
y(k) &= -z^{-d+1}\Sigma_1 H x_2(k) - z^{-d+1}\alpha\Sigma_1 H x_2(k - 1) + u(k) \quad &\text{(A.21)} \\
&= -z^{-d+1}\Sigma_1 H \left(\bar{x}_2(k + 1) + \alpha\bar{x}_2(k)\right) + u(k), \quad &\text{(A.22)}
\end{aligned}
$$

where $u(k)$ is used to denote the input to the state-space system. Furthermore, by comparing the first and second state update equation it is clear that $x_1(k)$ and $x_2(k)$ are related as $x_2(k) = Fx_1(k) - Gx_2(k - 1)$. This in combination with the definition $\bar{x}_2(k + 1) \doteq x_2(k)$ gives rise to the following state update equation for the time shifted state

$$
\bar{x}_2(k + 1) = Fx_1(k) - G\bar{x}_2(k).
$$

Both the above state-update equation and output equation (A.22) do no longer depend on the state $x_3(k)$, which can hence be disregarded. This in combination with the state update equation for $x_1(k)$ gives rise to the following result

$$
I - \Sigma_1\mathcal{H}(z)\mathcal{W}(z) = \left[\begin{array}{cc|c} \widetilde{A} & 0 & K_d \\ F & -G & 0 \\ \hline -z^{-d+1}\Sigma_1 HF & -z^{-d+1}\Sigma_1 H(\alpha I - G) & I \end{array} \right]
$$

By applying the rule for system inversion (A.4), it is possible to derive a state-space realization for $(I - \Sigma_1\mathcal{H}(z)\mathcal{W}(z))^{-1}$. The optimal feedback controller $\mathcal{C}(z)$ can now be computed by substituting the obtained state-space realizations for $\mathcal{W}(z)$ and $(I - \Sigma_1\mathcal{H}(z)\mathcal{W}(z)^{-1}$ in (A.17)

$$
\begin{aligned}
\mathcal{C}(z) &= \left[\begin{array}{cc|c} \widetilde{A} & 0 & K_d \\ F\widetilde{A} & -G & FK_d \\ \hline F\widetilde{A} & -G & FK_d \end{array} \right] \left[\begin{array}{cc|c} M & L & K_d \\ F & -G & 0 \\ \hline z^{-d+1}\Sigma_1 HF & z^{-d+1}\Sigma_1 H(\alpha I - G) & I \end{array} \right] \\
&= \left[\begin{array}{cccc|c} M & L & 0 & 0 & K_d \\ F & -G & 0 & 0 & 0 \\ z^{-d+1}K_d\Sigma_1 HF & L & \widetilde{A} & 0 & K_d \\ z^{-d+1}FK_d\Sigma_1 HF & FL & F\widetilde{A} & -G & FK_d \\ \hline z^{-d+1}FK_d\Sigma_1 HF & FL & F\widetilde{A} & -G & FK_d \end{array} \right],
\end{aligned}
$$

where the matrices L and M are defined as $L \doteq z^{-d+1}K_d\Sigma_1 H(\alpha I - G)$ and $M \doteq \widetilde{A} + z^{-d+1}K_d\Sigma_1 HF$. Even though this is already a state-space realization for the optimal feedback controller $\mathcal{C}(z)$, the above expression can still be simplified since

the above state-space realization is not minimal. To make this clear consider the similarity transformation

$$
T = \begin{bmatrix} I & 0 & -I & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}. \tag{A.23}
$$

By applying this transformation on the state-space realization (A.23), it is possible to single out one of the states that cannot be controlled from the input. Removing this state gives rise to the following reduced state-space realization for $\mathcal{C}(z)$

$$
\mathcal{C}(z) \;=\; \left[\begin{array}{ccc|c} -G & F & 0 & 0 \\ L & M & 0 & K_d \\ FL & FM & -G & FK_d \\ \hline FL & FM & -G & FK_d \end{array} \right].
$$

A further reduction of the state dimension can be obtained by following a similar approach as used to reduce the order of equation (A.20). Let x_1, x_2 and x_3 again denote the three components of the state vector which is partitioned in accordance with the partitioning of the state-space matrices, and let $u(k)$ and $y(k)$ denote the input and output signals to the system. Then by comparing the second and third state component the state update equations can be expressed as

$$
\begin{aligned}
x_1(k+1) &= -Gx_1(k) + \;\;\; Fx_2(k) \\
x_2(k+1) &= \;\;\; Lx_1(k) + \;\;\; Mx_2(k) \;\; + K_du(k) \;. \\
x_3(k+1) &= -Gx_3(k) + Fx_2(k+1)
\end{aligned} \tag{A.24}
$$

From the above expression it is clear that the states x_1 and x_3 are equal except for a one sample delay, i.e. $x_3(k) = x_1(k+1)$. It is hence possible to eliminate the state $x_3(k)$ from the state-space equations. By noting that $y(k) = x_3(k+1)$, the output of the system can be written as

$$
\begin{aligned}
y(k) &= \;\; Fx_2(k+1) - Gx_1(k+1) \\
&= \;\; (FL+G^2)x_1(k) + (FM - GF)x_2(k) + FK_du(k),
\end{aligned}
$$

where in the last step we have substituted the state update equations (A.24) for x_1 and x_2. The output equation and state update for x_1 and x_2 do no longer depend on x_3 and hence the optimal feedback controller has a state-space representation

$$
\mathcal{C}(z) = \left[\begin{array}{cc|c} -G & F & 0 \\ L & M & K_d \\ \hline (FL+G^2) & (FM - GF) & FK_d \end{array} \right], \tag{A.25}
$$

which completes the formal proof of Theorem 3.3.

Bibliography

Aitken, G. J. M. and D. McGaughey (1996). Predictability of atmospherically distorted stellar wavefronts. In M. Cullum (Ed.), *Proceedings European Southern Observatory (ESO) Conference on Adaptive Optics*, Volume 54, Garching, Germany (May), pp. 89–94.

Aoki, M. (1990). *State Space Modeling of Time Series.* Springer.

Arnold, W. F. and A. J. Laub (1984). Generalized eigenproblem algorithms and software for algebraic Riccati equations. *Proceedings of the IEEE* **72**(12), 1746–1754.

Åström, K. J. and B. Wittenmark (1997). *Computer-Controlled Systems, Theory and Design.* Prentice-Hall International Inc.

Babcock, H. W. (1953). The possibility of compensating astronomical seeing. *Publications of the Astronomical Society of the Pacific* **65**(386), 229–236.

Bakut, P. A., V. E. Kirakosyants, V. A. Loginov, C. J. Solomon and J. C. Dianty (1994). Optimal wavefront reconstruction from a shack-hartmann sensor by use of a bayesian algorithm. *Optics Communications* **109**, 10–15.

Bauer, D. (2004). Choosing integer parameters in subspace identification: A survey on asymptotic results. In *Proceedings of the IFAC Symposium on System Identification*, Rotterdam (August).

Beckers, J. M. (1989). Detailed compensation of atmospheric seeing using multi-conjugate adaptive optics. In F. J. Roddier (Ed.), *Active Telescope Systems*, Volume 1114 of *Proceedings of SPIE*, Orlando, Florida, United States (September), pp. 215–217.

Beckers, J. M. (1993). Adaptive optics for astronomy: Principles, performance, and application. *Annual review of astronomy and astrophysics* **31**, 13–62.

Ben-Tal, A. and A. Nemirovski (2001). *Lectures on modern convex optimization: analysis, algorithms, and engineering applications.* SIAM.

Benner, P., A. J. Laub and V. Mehrmann (1997). Benchmarks for the numerical solution of algebraic Riccati equations. *IEEE Control Systems Magazine* **17**(5), 18–28.

Born, M. and E. Wolf (1999). *Principles of Optics: Electromagnetic Theory of Propagation* (seventh (Expanded) Edition ed.). Cambridge University Press. ISBN 0-521-64222-1.

Brase, J. M., J. An, K. Avicola, B. V. Beeman, D. T. Gavel, R. Hurd, B. Johnston, H. Jones, T. Kuklo, C. E. Max, S. S. Olivier, K. E. Waltjen and J. Watson (1998). The wavefront control system for the keck telescope. In *Adaptive Optical System Technologies*, Volume 3353, Kona, Hawaii (March), pp. 517–521. Symposium on Astronomical Telescopes and Instrumentation.

Brigantic, R. T., M. C. Roggemann, B. M. Welsh and K. W. Bauer (1998). Optimization of adaptive-optics systems closed-loop bandwidth settings to maximize imaging-system performance. *Applied Optics* **37**(5), 848–855.

Bruls, J., C. T. Chou, B. R. J. Haverkamp and M. Verhaegen (1999). Linear and non-linear system identification using separable least-squares. *European Journal of Control* **5**, 116–128.

Burl, J. B. (1999). *Linear Optimal Control: $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Methods.* California, United States: Addison-Wesley.

Chang, C.-C and J. S. Gibson (2000). Parallel control loops based on spatial subband processing for adaptive optics. In *Proceedings of the American Control Conference*, Chicago, Illinois (June), pp. 2113–2117.

Chen, T. and B. Francis (1995). *Optimal Sampled-Data Control Systems.* London, Berlin: Springer-Verslag. ISBN 3-540-19949-7.

Conan, J.-M., G. Rousset and P.-Y. Madec (1995). Wave-front temporal spectra in high-resolution imaging through turbulence. *Journal of the Optical Society of America A* **12**(7), 1559–1570.

Dahlén, A. (2001). *Identification of stochastic systems: Subspace methods and covariance extension.* Ph. D. thesis, Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden.

Dessenne, C., P.-Y. Madec and G. Rousset (1997). Modal prediction for closed-loop adaptive optics. *Optics Letters* **22**(20), 1535–1537.

Dessenne, C., P.-Y Madec and G. Rousset (1998). Optimization of a predictive controller for closed-loop adaptive optics. *Applied Optics* **37**(21), 4823–4633.

Dessenne, C., P.-Y. Madec and G. Rousset (1999). Sky implementation of model predictive control in adaptive optics. *Optics Letters* **24**(5), 339–341.

Dicke, R. H. (1975). Phase-contrast detection of telescope seeing errors and their correction. *Astrophysical Journal* **198**, 605–615.

Doyle, J. C., K. Glover, P. P. Khargonekar and B. A. Fracis (1989). State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems. *IEEE Transactions on Automatic Control* **34**(8), 831–847.

Ellenbroek, R. M. L., M. Verhaegen, N. Doelman, R. F. M. M. Hamelinck, P. C. J. N. Rosielle and M. Steinbuch (2006). Distributed control in adaptive optics - deformable mirror and turbulence modeling. In B. L. Ellerbroek and D. B. Calia (Eds.), *Advances in Adaptive Optics II*, Volume 6272 of *Proceedings of SPIE*, Orlando, Florida, United States (May), pp. 62723K–1–12.

Ellerbroek, B. and F. Rigaut (2000). Optics adapt to the whole sky. *Nature* **403**, 25–26.

Ellerbroek, B. L. (1994). First-order performance evaluation of adaptive-optics systems for atmospheric-turbulence compensation in extended-field-of-view astronomical telescopes. *Journal of the Optical Society of America A* **11**(2), 783–805.

Ellerbroek, B. L. (2002). Efficient computation of minimum-variance wave-front reconstructors with sparse matrix techniques. *Journal of the Optical Society of America A* **19**(9), 1803–1816.

Ellerbroek, B. L. and T. A. Rhoadarmer (1998). Real-time adaptive optimization of wave-front reconstruction algorithms for closed-loop adaptive-optical systems. In D. Bonaccini and R. K. Tyson (Eds.), *Adaptive Optical System Technologies*, Volume 3353 of *Proceeding of SPIE*, Kona, Hawaii (September), pp. 1174–1185.

Ellerbroek, B. L. and T. A. Rhoadarmer (2001). Adaptive wavefront control algorithms for closed loop adaptive optics. *Mathematical and Computer Modelling* **33**(1-3), 145–158.

Ellerbroek, B. L., C. Van Loan, N. P. Pitsianis and R. J. Plemmons (1994). Optimizing closed-loop adaptive-optics performance with use of multiple control bandwidths. *Journal of the Optical Society of America A* **11**(11), 2871–2886.

Fraanje, P. R. (2004). *Robust and fast schemes in broadband active noise and vibration control*. Ph. D. thesis, University of Twente, the Netherlands.

Fraanje, R., M. Verhaegen, N. Doelman and A. Berkhoff (2004). Optimal and robust feedback controller estimation for a vibrating plate. *Control Engineering Practice* **12**(8), 1017–1027.

Francis, B. A. and W. M. Wonham (1976). The internal model principle of control theory. *Automatica* **12**(5), 457–465.

Fried, D. L. (1977). Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements. *Journal of the Optical Society of America* **67**(3), 370–375.

Fried, D. L. (1990). Time-delay-induced mean-square error in adaptive optics. *Journal of the Optical Society of America A* **7**(7), 1224–1225.

Fugate, R. Q., D. L. Fried, G. A. Ameer, B. R. Boeke, S. L. Browne, P. H. Roberts, R. E. Ruane, G. A. Tylor and L. M. Wopat (1991). Measurement of atmospheric wavefront distortion using scattered light from a laser guide-star. *Nature* **353**, 144–146.

Fusco, T., J.-M. Conan, G. Rousset, L. M. Mugnier and V. Michau (2001). Optimal wave-front reconstruction strategies for multi-conjugate adaptive optics. *Journal of the Optical Society of America A* **18**(10), 2527–2538.

Gavel, D. T. and D. Wiberg (2003). Towards strehl-optimizing adaptive optics controllers. In P. L. Wizinowich and D. Bonaccini (Eds.), *Adaptive Optical System Technologies II*, Volume 4839 of *Proceedings of SPIE*, Hawaii (February), pp. 890–901.

Gendron, E. and P. Léna (1994). Astronomical adaptive optics. *Astronomy and Astrophysics: I. Modal control optimization* **291**, 337–347.

Gendron, E. and P. Léna (1995). Astronomical adaptive optics. II experimental results of an optimized modal control. *Astronomy & Astrophysics Supplement Series* **111**, 153–167.

Gendron, E. and P. Léna (1996). Single layer atmospheric turbulence demonstrated by adaptive optics observations. *Astrophysics and Space Science* **239**(2), 221–228.

Gibson, J. S., C. Chang and B. L. Ellerbroek (2000). Adaptive optics: wave-front correction by use of adaptive filtering and control. *Applied Optics* **39**(16), 2525–2537.

Gibson, J. S. and C.-C. Chang (1999). Adaptive optics: Wavefront reconstruction by adaptive filtering and control. In *Proceedings of the 38th Conference on Decision & Control*, Phoenix, Arizona, United States (December), pp. 761–766. IEEE.

Gibson, J. S., C.-C. Chang and N. Chen (2001). Adaptive optics with a new modal decomposition of actuator and sensor spaces. In *Proceedings of the 2001 American Control Conference*, Volume 6, Arlington, VA (June), pp. 4619–4625.

Gibson, J. S., C.-C. Chang and B. L. Ellerbroek (1999). Adaptive optics: Wavefront reconstruction by adaptive filtering and control. In *Proceedings of the 38th Conference on Decision & Control*, Phoenix, Arizona, United States (December), pp. 761–765.

Gilles, L. (2003). Order-N sparse minimum-variance open-loop reconstructor for extreme adaptive optics. *Optics Letters* **28**(20), 1927–1929.

Gilles, L., B. L. Ellerbroek and C. R. Vogel (2003). Preconditioned conjugate gradient wave-front reconstructors for multi-conjugate adaptive optics. *Applied Optics* **42**(26), 5233–5250.

Gilles, L., C. R. Vogel and B. L. Ellerbroek (2002). Multi-grid preconditioned conjugate-gradient method for large-scale wave-front reconstruction. *Journal of the Optical Society of America A* **19**(9), 1817–1822.

Glindemann, A., S. Hippler, T. Berkefeld and W. Hackenberg (2000). Adaptive optics on large telescopes. *Experimental Astronomy* **10**(1), 5–49.

Golub, G. H. and V. Pereyra (1973). The differentiation of pseudo-inverses and nonlinear least squares problems whose variables are separate. *SIAM Journal on Numerical Analysis* **10**(2), 413–432.

Goodman, J. W. (1985). *Statistical Optics*. Wiley-Interscience.

Goodwin, G. C., S. F. Graebe and M. E. Salgado (2001). *Control System Design*. Upper Saddle River, New Jersey: Prentice Hall.

Greenway, A. and J. Burnett (2004). *Industrial and Medical Applications of Adaptive Optics*. Technology tracking. IOP Publishing Ltd. ISBN 0-7503-0850-8.

Greenwood, D. P. (1977). Bandwidth specification for adaptive optics systems. *Journal of the Optical Society of America* **67**, 390–393.

Greenwood, D. P. and D. L. Fried (1976). Power spectra requirements for wave-front-compensative systems. *Journal of the Optical Society of America* **66**, 193–206.

Hamelinck, R. F. M. M., P. C. J. N. Rosielle, J. P. Kappelhof, B. Snijders and M. Steinbuch (2004). Large adaptive deformable membrane mirror with high actuator density. In D. B. Calia, B. L. Ellerbroek, and R. Ragazzoni (Eds.), *Advancements in Adaptive Optics*, Volume 5490 of *Proceedings of SPIE*, Glasgow, Scotland, United Kingdom (June), pp. 1482–1492.

Hamelinck, R. F. M. M., P. C. J. N. Rosielle, M. Steinbuch and N. Doelman (2005a). Large adaptive deformable membrane mirror with high actuator density: design and first actuator tests. In W. Jiang (Ed.), *5th Workshop on Adaptive Optics for Industry and Medicine*, Volume 6018 of *Proceedings of SPIE*, Beijing, China (August), pp. 60180Y–1–13.

Hamelinck, R. F. M .M., P. C. J. N. Rosielle, M. Steinbuch and N. Doelman (2005b). Large adaptive deformable mirror: design and first prototypes. In M. T. Gruneisen, J. D. Gonglewski, and M. K. Giles (Eds.), *Optics and Photonics 2005 - Advanced Wavefront Control: Methods, Devices and Applications III*, Volume 5894 of *Proceedings of SPIE*, San Diego, California, United States (August), pp. 589418–1–11.

Hamelinck, R. F. M. M., P. C. J. N. Rosielle, M. Steinbuch, R. M. L. Ellenbroek, M. Verhaegen and N. Doelman (2006). Actuator tests for a large deformable membrane mirror. In B. L. Ellerbroek and D. B. Calia (Eds.), *Astronomical telescopes and instrumentation - advances in adaptive optics*, Volume 6272 of *Proceedings of SPIE*, Orlando, Florida, United States (June), pp. 627225–1–9.

Hanzon, B. and R. L. M. Peeters (2000). Balanced parameterization of stable SISO all-pass systems in discrete time. *Math. Control Signals Systems* **13**, 240–276.

Hardy, J. W. (1998). *Adaptive Optics for Astronomical Telescopes*. Oxford series in optical and imaging sciences. New York: Oxford University Press.

Hassibi, B., A. H. Sayed and T. Kailath (1999). *Indefinite-Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories*. Philadelphia: SIAM Studies in Applied and Numerical Mathematics.

Hayes, M. H. (1996). *Statistical digital signal processing and modeling*. New York: John Wiley & Sons, Inc. ISBN 0-471-59431-8.

Hecht, E. (1987). *Optics* (second edition ed.). Addison-Wesley Publishing Company.

Herrmann, J. (1980). Least-squares wave front errors of minimum norm. *Journal of the Optical Society of America* **70**(1), 28–35.

Herrmann, J. (1992). Phase variance and strehl ratio in adaptive optics. *Journal of the Optical Society of America A* **9**(12), 2257–2258.

Hinnen, K. J. G., N. J. Doelman and M. Verhaegen (2005). $\mathcal{H}_2$-optimal control of an adaptive optics system: Part II, closed-loop controller design. In R. K. Tyson and M. Lloyd-Hart (Eds.), *Astronomical Adaptive Optics Systems and Applications II*, Volume 5903 of *Proceedings of SPIE*, San Diego, California, United States (August), pp. 86–99.

Hinnen, K. J. G., M. Verhaegen and N. J. Doelman (2005). $\mathcal{H}_2$-optimal control of an adaptive optics system: Part I, data-driven modeling of the wavefront disturbance. In R. K. Tyson and M. Lloyd-Hart (Eds.), *Astronomical Adaptive Optics Systems and Applications II*, Volume 5903 of *Proceedings of SPIE*, San Diego, California, United States (August), pp. 75–85.

Hinnen, K. J. G., M. Verhaegen and N. J. Doelman (2006). A data-driven $\mathcal{H}_2$-optimal control approach for adaptive optics. Submitted to IEEE Transactions on Control Systems Technology.

Ho, B. L. and R. E. Kalman (1966). Effective construction of linear time state-variable models from input/output functions. *Regelungstechnik* **14**, 545–548.

Hudgin, R. H. (1977). Wave-front reconstruction for compensated imaging. *Journal of the Optical Society of America* **67**(3), 375–378.

Jansson, M. (2003). Subspace identification and ARX modeling. In *13th IFAC Symposium on System Identification*, Rotterdam, the Netherlands (August), pp. 1625–1630.

Johansson, R., M. Verhaegen, C. T. Chou and A. Robertsson (2001). Residual models and stochastic realization in state-space identification. *International Journal of Control* **74**(10), 988–995.

Kailath, T., A. H. Sayed and B. Hassibi (2000). *Linear Estimation*. Upper Saddle River, New Jersey: Prentice Hall Information ans Systems Sciences Series. ISBN 0-13-022464-2.

Kirkegaard, P. H. and P. Andersen (1997). State space identification of civil engineering structures from output measurements. In *Fracture and Dynamics*, Volume 3089 of *15th International Modal Analysis Conference*, Orlando, Florida, United States (February), pp. 889–895.

Kolmogorov, A. N. (1960). *Dissipation of Energy in Locally Isotropic Turbulence; translation in: Turbulence, Classic papers on statistical theory*, pp. 156–158. New York, London: Interscience Publishers, Inc. from Comptes rendus de l'academie des science de l'U.R.S.S., Vol. 31, 1941.

Larimore, W. E. (1990). Canonical variate analysis in identification, filtering, and adaptive control. In *29th IEEE Conference on Decision & Control*, Honolulu, Hawaii, United States (December), pp. 569–604.

Law, N. F. and R. G. Lane (1996). Wavefront estimation at low light levels. *Optics Communications* **126**, 19–24.

Le Roux, B., J.-M. Conan, C. Kulcsár, H.-F Raynaud, L. M. Mugnier and T. Fusco (2004). Optimal control law for classical and multi-conjugate adaptive optics. *Journal of the Optical Society of America. A* **21**(7), 1261–1276.

Léna, P. (1997). An introduction to atmospheric turbulence. In A. M. Lagrange, D. Mourand, and P. Léna (Eds.), *High Angular Resolution in Astrophysics: Proceedings of the NATO Advanced Study Institute*, Volume 501 of *NATO Asi Series. Series C, Mathematical and Physical Sciences*, Les Houches, France (April). Kluwer, Academic.

Lindquist, A. and G. Picci (1996). Canonical correlation analysis, approximate covariance extension and identification of stationary time series. *Automatica* **32**(5), 709–733.

Liu, Y-T. and S. Gibson (2004). Adaptive optics with adaptive filtering and control. In *Proceedings of the American Control Conference*, Volume 4, Boston, Massachusetts (July), pp. 3176–3179.

Ljung, L. (1999). *System Identification: Theory for the User* (2nd edition ed.). Upper Saddle River, NJ: Prentice Hall PRT.

Ljung, L. and T. Glad (1994). *Modeling of Dynamic Systems*. Englewood Cliffs: Prentice Hall PRT. ISBN 0-13-597097-0.

Lloyd-Hart, M. and P. McGuire (1995). Spatio-temporal prediction for adaptive optics wavefront reconstructors. In *Proc. Adaptive Optics Topical Meeting*, Garching, Germany (October), pp. 95.

Looze, D. P. (2005a). Minimum variance control structure for adaptive optics systems. In *American Control Conference*, Portland, Oregon, United States (June), pp. 1466–1471.

Looze, D. P. (2005b). Realization of systems with CCD-based measurements. *Automatica* **41**(11), 2005–2009.

Looze, D. P. (2006). Minimum variance control structure for adaptive optics systems. *Journal of the Optical Society of America A* **23**(3), 603–612.

Looze, D. P., O. Becker, M. Kasper and S. Hippler (1999). Optimal compensation and implementation for adaptive optics systems. In *Proceedings of the 38th Conference of Decision & Control*, Phoenix, Arizona, United States (December), pp. 1715–1720.

Looze, D. P., m. Kasper, S. Hippler and O. Beker ans R. Weiss (2003). Optimal compensation and implementation for adaptive optics systems. *Experimental Astronomy* **15**(2), 67–88.

MacMartin, D. G. (2003). Local, hierarchic, and iterative reconstructors for adaptive optics. *Journal of the Optical Society of America A* **20**(6), 1084–1093.

Madec, P. Y. (1999). *Adaptive Optics in Astronomy*, Chapter Control techniques, pp. 131–153. Cambridge University Press.

Mari, J., P. Stoica and T. McKelvey (2000a). Vector ARMA estimation: A reliable subspace approach. *IEEE Transactions on Signal Processing* **48**(7), 2092–2104.

Mari, J., P. Stoica and T. McKelvey (2000b). Vector ARMA estimation: A reliable subspace approach. *IEEE Transactions on Signal Processing* **48**(7), 2092–2104.

Mastronardi, N., D. Kressner, V. Sima, P. Van Dooren and S. Van Huffel (2001). A fast algorithm for subspace state-space system identification via exploitation of the displacement structure. *Journal of Computational and Applied Mathematics* **132**(1), 71–81.

McGuire, P. C., T. A. Rhoadarmer, H. A. Coy, J. R. P. Angel and M. Lloyd-Hart (2000). Linear zonal atmospheric prediction for adaptive optics. In P. L. Wizinowich (Ed.), *Adaptive Optical Systems Technology*, Volume 4007 of *Proceedings of SPIE*, Munich, Germany (July), pp. 682–691.

Morari, M. and Zafiriou (1989). *Robust Process Control*. Englewood Cliffs, New Jersey, United States: Prentice-Hall.

Paschall, R. N. (1991). Design of a linear quadratic gaussian controller for an adaptive optics system. In *Proceedings of the 30th Conference on Decision and Control*, Volume 2, Brighton, United Kingdom (December), pp. 1761–1769. IEEE.

Paschall, R. N. and D. J. Anderson (1993). Linear quadratic gaussian control of a deformable mirror adaptive optics system with time-delayed measurements. *Applied Optics* **32**(31), 6347–6358.

Peternell, K., W. Scherrer and M. Deistler (1996). Statistical analysis of novel subspace identification methods. *Signal Processing* **52**, 161–177.

Petit, C., J. M. Conan, C. Kulcár, H. F. Raynaud, T. Fusco, J. Montri and D. Rabaud (2006). First laboratory demonstration of closed-loop kalman based optimal control for vibration filtering and simplified MCAO. In B. L. Ellerbroek and D. B. Calia (Eds.), *Advances in Adaptive Optics II*, Volume 6272 of *Proceedings of SPIE*, Orlando, Florida, United States (June).

Petit, C., F. Quiros-Pacheco, J. M. Conan, C. Kulcsár, H. F. Raynaud, T. Fusco and G. Rousset (2004). Kalman-filter-based control for adaptive optics. In D. B. Calia, B. L. Ellerbroek, and R. Ragazzoni (Eds.), *Advancements in Adaptive Optics*, Volume 5490 of *Proceedings of SPIE*, Glasgow, Scotland, United Kingdom (June), pp. 1414–1425.

Poyneer, L. A., D. T. Gavel and J. M. Brase (2002). Fast wave-front reconstruction in large adaptive optics systems with use of the fourier transform. *Journal of the Optical Society of America A* **19**(10), 2100–2111.

Ragazzoni, R., E. Marchetti and G. Valente (2000). Adaptive-optics corrections available for the whole sky. *Nature* **403**, 54–56.

Ribarits, T., M. Deistler and B. Hanzon (2003). Separable least squares data driven local coordinates. In *13th IFAC Symposium on Systems Identification*, Rotterdam (August), pp. 1922–1927.

Rigaut, F. and E. Genrdon (1992). Laser guide star in adaptive optics - the tilt determination problem. *Astronomy and Astrophysics* **261**(2), 677–684.

Roberts Jr., L. C., M. D. Perrin, F. Marchis, A. Sivaramakrishnan, R.B. Makidon, J. C. Christou, B. A. Macintosh, L. A. Poyneer, M. A. van Dam and M. Troy (2004). Is that really your strehl ratio? In D. B. Calia and B. L. Ellerboek (Eds.), *Advancements in Adaptive Optics*, Volume 5490 of *Proceedings of SPIE*, pp. 504–512.

Roddier, F. (1981). *The effects of atmospheric turbulence in optical astronomy*, Volume 19 of *Progress in optics*, pp. 281–376. North-Holland, New York: Elsevier.

Roddier, F. (1998). Maximum gain and efficiency of adaptive optics systems. *Publications of the Astronomical Society of the Pacific* **110**, 837–840.

Roddier, F. (1999). *Adaptive Optics in Astronomy*. Cambridge University Press. ISBN 0-521-55375-X.

Rousset, G., J. C. Fontanella, P. Kern and F. Rigaut (1990). First diffraction-limited astronomical images with adaptive optics. *Astronomy and Astrophysics* **230**(2), L29–L32.

Ruhe, A. and P. A. Wedin (1980). Algorithms for separable nonlinear least squares problems. *SIAM Review* **22**(3), 318–337.

Saint-Jacques, D. (1998). *Astronomical Seeing in Space and Time*. Ph. D. thesis, University of Cambridge.

Sayed, A. H. and T. Kailath (2001). A survey of spectral factorization methods. *Numerical linear algebra with applications* **8**(6-7), 467–496.

Schöck, M. (1998). *An analysis of turbulent layers with a wavefront sensor*. Ph. D. thesis, University of Wyoming.

Schöck, M. and E. J. Spillar (2000). Method for a quantitative investigation of the frozen flow hypothesis. *Journal of the Optical Society of America A* **17**(9), 1650–1658.

Shanmugan, K. S. and A. M. Breipohl (1988). *Random Signals: Detection Estimation and Data Analysis.* John Wiley & Sons. ISBN 0-471-81555-1.

Shu, H. and T. Chen (1995). State-space approach to discrete-time $\mathcal{H}_2$ optimal control with a causality constraint. In *Proceedings of the 34th Conference on Decision and Control*, New Orleans, LA (December), pp. 1927–1932.

Sjöberg, J. and M. Viberg (1997). Separable non-linear least squares minimization - possible improvements for neural net fitting. In *Proceeding of IEEE Workshop in Neural Networks for Signal Processing*, Amelia Island Plantation, Florida (September), pp. 345–354.

Sturm, J. F. (2002). Implementation of interior point methods for mixed semidefinite and second order cone optimization problems. *Optimization Methods and Software* **17**(6), 1105–1154.

Tatarskii, V. I. (1971). *The Effects of the Turbulent Atmosphere on Wave Propagation*. Israel Program for Scientific Translations Ltd. Jarusalem: Keter Press.

Taylor, G. I. (1938). The spectrum of turbulence. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* **164**(919), 476–490.

Thomas, S. (2004). Optimized centroid computing in a shack-hartmann sensor. In R. Ragazzoni D. B. Calia, B. L. Ellerbroek (Ed.), *Advances in Adaptive Optics*, Volume 5490 of *Proceedings of SPIE*, Glasgow, Scotland, United Kingdom (October), pp. 1238–1246.

Tyler, G. A. (1994). Bandwidth considerations for tracking through turbulence. *Journal of the Optical Society of America A* **11**(1), 358–367.

Tyson, R. K. (1998). *Principles of Adaptive Optics* (2nd edition ed.). Academic Press. ISBN 0-12-705902-4.

Tyson, R. K. (2000). *Introduction to Adaptive Optics*. SPIE - The International Society of Optical Engineering.

van Dam, M. A., D. Le Mignant and B. A. Macintosh (2004). Performance of Keck observatory adaptive optics system. *Applied Optics* **43**(29), 5458–5467.

Van Overschee, P. and B. De Moor (1993a). Subspace algorithms for the stochastic identification problem. *Automatica* **29**(3), 649–660.

Van Overschee, P. and B. De Moor (1993b). Subspace algorithms for the stochastic identification problem. *Automatica* **29**(3), 648–660.

Van Overschee, P. and B. De Moor (1996). *Subspace Identification For Linear Systems, Theory - Implementation - Applications*. Dordrecht: Kluwer Academic Publishers.

Van Overschee, P., B. De Moor, W. Dehandschutter and J. Swevers (1997). A subspace algorithm for the identification of discrete time frequency domain power spectra. *Automatica* **33**(12), 2147–2157.

Vdovin, G. and P. M. Sarro (1995). Flexible mirror micro-machined in silicon. *Applied Optics* **34**(16), 2969–2972.

Verdult, V. and M. Verhaegen (2001). Identification of multivariable bilinear state space systems based on subspace techniques and separable least squares optimization. *International Journal of Control* **74**(18), 1824–1836.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica* **30**(1), 61–74.

Verhaegen, M. (1997). Weighted sensitivity minimization for causal linear, discrete time-varying system. *SIAM Journal on Control and Optimization* **35**(3), 791–809.

Viberg, M. (1995). Subspace-based methods for the identification of linear time-invariant systems. *Automatica* **31**(12), 1835–1851.

Vidyasagar, M. (1988). *Control systems synthesis: a factorization approach*. Cambridge, Massachusetts, United States: MIT Press. ISBN 0-262-72012-4.

Wallner, E. P. (1983). Optimal wave-front correction using slope measurements. *Journal of the Optical Society of America* **73**(12), 1771–1776.

Wang, J. Y. and J. K. Markey (1978). Modal compensation of atmospheric turbulence phase distortions. *Journal of the Optical Society of America* **68**, 78–87.

Weyrauch, T., M.A. Vorontsov, T.G. Bifano, J.A. Hammer, M. Cohen and G. Cauwenberghs (2001). Microscale adaptive optics: wave-front control with a μ-mirror array and a vlsi stochastic gradient descent controller. *Applied Optics* **40**(24), 4243–4251.

Wiberg, D. M., C. E. Max and D. T. Gavel (2004). A spatial non-dynamic LQG controller: Part II, theory. In *43rd IEEE Conference on Decision and Control*, Atlantic, Paradise Island, Bahamas (December), pp. 3333–3338.

Wiberg, D. M., C. E. Max and D. T. Gavel (2005). Geometric view of adaptive optics control. *Journal of the Optical Society of America A* **22**(5), 870–880.

Wild, W. J. (1996). Predictive optimal estimators for adaptive-optics systems. *Optics Letters* **21**(18), 1433–1435.

Wild, W. J. (1998). Lyapunov stability criteria for zonal adaptive-optics systems. *Optics Letters* **23**(8), 570–572.

Wonham, W. M. (1968). On the separation theorem of stochastic control. *SIAM Journal on Control and Optimization* **6**(2), 312–326.

Youla, D. C., J. J. Bongiorno and H. A. Jabr (1976). Modern wiener-hopf design of optimal controllers–part II: The multivariable case. *IEEE Transactions on Automatica Control* **21**(3), 319–338.

Zhou, K., J. C. Doyle and K. Glover (1996). *Robust and Optimal Control*. Upper Saddle River, New Jersy: Prentice Hall. ISBN 0-13-456567-3.

List of Abbreviations

AO	Adaptive Optics
AR	Auto-Regressive
ARMA	Auto-Regressive Moving Average
ARX	Auto-Regressive with eXogenous input
CCA	Canonical Correlation Analysis
CCD	Charge coupled device
CLP	Conic Linear Programming
DM	Deformable Mirror
DTFT	Discrete-Time Fourier Transform
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FoV	Field of View
FWHM	Full-Width of Half Maximum
IDFT	Inverse Discrete Fourier Transform
IMC	Internal Model Control
LFT	Linear Fractional Transformation
LGS	Laser Guide Star
LMI	Linear Matrix Inequality
LQG	Linear Quadratic Gaussian
LQR	Linear Quadratic Regulator
LTI	Linear Time-Invariant
MAP	Maximum a posteriori estimator
MCAO	Multi Conjugate Adaptive Optics
MIMO	Multiple Input Multiple Output
MOESP	Multi-variable Output Error State sPace

MSE	Mean Square Error
OTF	Optical Transfer Function
RDA	Remote Data Access
RLS	Recursive Least Squares
SISO	Single Input Single Output
SNR	Signal to Noise Ratio
SSAR	Subspace algorithm using AR model initial step
SSARX	Subspace algorithm using ARX model initial step
SVD	Singular Value Decomposition
TT	Tip-tilt mirror
VAF	Variance Accounted For
VAR	Variance
WFS	Wavefront Sensor
WHT	William Herschel Telescope
WSS	Wide Sense Stationary
ZOH	Zero-order hold

List of Publications

Accepted and submitted journal papers

Hinnen, K. J. G., R. Fraanje and M. Verhaegen (2004). The application of initial state correction in iterative learning control and the experimental validation on a piezoelectric tube scanner. *Proceedings of the I MECH E Part I Journal of Systems & Control Engineering* **16**, 503–511.

Hinnen, K., M. Verhaegen and N. Doelman (2005). Robust spectral factor approximation of discrete-time frequency domain power spectra. *Automatica* **41**(10), 1791–1798.

Hinnen, K. J. G., M. Verhaegen and N. J. Doelman (2006b). A data-driven $\mathcal{H}_2$-optimal control approach for adaptive optics. Submitted to IEEE Transactions on Control Systems Technology.

Hinnen, K., M. Verhaegen and N. J. Doelman (2006a). Exploiting the spatio-temporal correlation in adaptive optics using data-driven $\mathcal{H}_2$-optimal control. Submitted to Journal of the Optical Society of America A.

Conference proceedings

Hinnen, K. J. G., R. Fraanje and M. Verhaegen (2003). The application of initial state-correction in iterative learning control and the experimental validation on a piezoelectric tube scanner. In R. M. Parkin, A. El-Habaibeh, and M. R. Jackson (Eds.), *ICOM 2003 - International Conference on Mechatronics*, Loughborough, UK (June), pp. 223–228. Wiley.

de Boer, M., K. Hinnen, M. Verhaegen, R. Fraanje, G. Vdovin and N. Doelman (2003). Control of a thermal deformable mirror: Correction of a static disturbance with limited sensor information. In U. Wittrock (Ed.), *Adaptive Optics for Industry and Medicine Proceedings of the 4th International Workshop*, Volume 102 of *Springer Proceedings in Physics*, Münster, Germany (October), pp. 61–72.

Doelman, N. J., K. J. G. Hinnen, F. J. G. Stoffelen and M. Verhaegen (2004). Optimal control strategy to reduce the temporal wavefront error in AO systems. In D. B. Calia, B. L. Ellerbroek, and R. Ragazzoni (Eds.), *Advancements in Adaptive Optics*, Volume 5490 of *Proceedings of SPIE*, Glasgow, UK (October), pp. 1426–1437.

Hinnen, K. J. G., M. Verhaegen and N. J. Doelman (2005). $\mathcal{H}_2$-optimal control of an adaptive optics system: Part I, data-driven modeling of the wavefront disturbance. In R. K. Tyson and M. Lloyd-Hart (Eds.), *Astronomical Adaptive Optics Systems and Applications II*, Volume 5903 of *Proceedings of SPIE*, San Diego, USA (August), pp. 75–85.

Hinnen, K. J. G., N. J. Doelman and M. Verhaegen (2005). $\mathcal{H}_2$-optimal control of an adaptive optics system: Part II, closed-loop controller design. In R. K. Tyson and M. Lloyd-Hart (Eds.), *Astronomical Adaptive Optics Systems and Applications II*, Volume 5903 of *Proceedings of SPIE*, San Diego, USA (August), pp. 86–99.

Hinnen, K. J. G., M. Verhaegen and N. J. Doelman (2006). Adaptive optics $\mathcal{H}_2$-optimal control design applied on an experimental setup. In D. B. Calia and B. L. Ellerbroek (Eds.), *Advances in Adaptive Optics II*, Volume 6272 of *Proceedings of SPIE*, Orlando, USA (May), pp. 62722S–1–12.

Curriculum Vitae

Karel Hinnen was born on July 23, 1977 in Enschede, the Netherlands. He received his secondary education respectively at the "Scholengemeenschap Het Assink" (HAVO) in Haaksbergen and at "het Graafschap College" (VWO) in Doetinchem. From 1996 to januari 2002 he studied Applied Physics at University of Twente in Enschede. Because of his study results, the faculty of Applied Physics offered him in 2000 to visit the "37th Culham Plasma Physics Summer School" in Oxford, United Kingdom. He graduated cum laude, with a major in control system engineering and a minor in biophysics. His Master's thesis dealt with the development of an advanced iterative learning control scheme for improving the tracking performance of piezoelectric tube scanners in atomic force microscopy. After his graduation he started as a research assistant at Control Systems Engineering laboratory of the Faculty of Information Technology and Systems, Delft University of Technology. For a few months, he participated in a project concerned with making an inventory of control strategies for smart dynamic rotor control of large offshore wind turbines. In 2002 he started as Ph D student in the same group, which later merged into the Delft Center for Systems and Control. Under the supervision of Prof. Michel Verhaegen, he worked on control algorithms for adaptive optics systems. His research was sponsored by the TNO Institute of Applied Physics. From 2002 to 2003 he took graduate courses at the Dutch Institute of Systems and Control (DISC), for which he received a DISC certificate. During his Ph D project, he participated in several international conferences, workshops and symposia. As of September 2006 he plans to work as a researcher at the Delft Center for Systems and Control. In his spare time he is a passionate sailor and the analogy between sailing and research inspires him to find the best solution for making progress.