

UNIVERSITY OF VERONA
DEPARTMENT OF COMPUTER SCIENCE

GRADUATE SCHOOL OF NATURAL AND ENGINEERING SCIENCES

DOCTORAL PROGRAM IN COMPUTER SCIENCE

CYCLE XXXII

Development of Advanced Control Strategies
for Adaptive Optics Systems

S.S.D. INF/01

Coordinator: _____

Prof. Massimo Merro




Tutor: _____

Dr. Riccardo Muradore

Doctoral Student: _____

Jacopo Mocci

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License, Italy. To read a copy of the licence, visit the web page:
<http://creativecommons.org/licenses/by-nc-nd/3.0/>

-  **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
-  **NonCommercial** — You may not use the material for commercial purposes.
-  **NoDerivatives** — If you remix, transform, or build upon the material, you may not distribute the modified material.

Development of Advanced Control Strategies for Adaptive Optics Systems – Jacopo Mocci

PhD thesis
Verona, April 19, 2020

*to Lucilla,
which means "light" and is needed both in astronomy and in my life.*

Abstract

Atmospheric turbulence is a fast disturbance that requires high control frequency. At the same time, celestial objects are faint sources of light and thus WFSs often work in a low photon count regime. These two conditions require a trade-off between high closed-loop control frequency to improve the disturbance rejection performance, and large WFS exposure time to gather enough photons for the integrated signal to increase the Signal-to-Noise ratio (SNR), making the control a delicate yet fundamental aspect for AO systems.

The AO plant and atmospheric turbulence were formalized as state-space linear time-invariant systems. The full AO system model is the ground upon which a model-based control can be designed.

A Shack-Hartmann wavefront sensor was used to measure the horizontal atmospheric turbulence. The experimental measurements yielded to the Cn2 atmospheric structure parameter, which is key to describe the turbulence statistics, and the Zernike terms time-series.

Experimental validation shows that the centroid extraction algorithm implemented on the Jetson GPU outperforms (i.e. is faster) than the CPU implementation on the same hardware. In fact, due to the construction of the Shack-Hartmann wavefront sensor, the intensity image captured from its camera is partitioned into several sub-images, each related to a point of the incoming wavefront. Such sub-images are independent each-other and can be computed concurrently.

The AO model is exploited to automatically design an advanced linear-quadratic Gaussian controller with integral action. Experimental evidence shows that the system augmentation approach outperforms the simple integrator and the integrator filtered with the Kalman predictor, and that it requires less parameters to tune.

Contents

| | |
|--|-----------|
| 1 Introduction | 5 |
| 1.1 Thesis Outline | 7 |
| 2 Background | 11 |
| 2.1 Introduction | 12 |
| 2.2 Zernike Polynomials | 14 |
| 2.3 Strehl Ratio | 17 |
| 2.4 State-space Formalization | 19 |
| 2.4.1 Atmospheric Turbulence | 20 |
| 2.4.2 Wavefront Sensor | 22 |
| 2.4.3 Deformable Mirrors | 24 |
| 2.4.4 AO Controller | 26 |
| 2.4.5 Discrete SCAO Block Diagram | 28 |
| 2.5 Conclusion | 28 |
| 3 PhotonLoop AO Controller | 31 |
| 3.1 Introduction | 32 |
| 3.2 Control Layer | 32 |
| 3.3 Monitor Layer | 34 |
| 3.4 AO System Static Calibration | 37 |
| 3.5 Experimental verification | 40 |
| 3.5.1 Setup Calibration | 42 |
| 3.5.2 Aberrations Correction | 42 |
| 3.6 Conclusion | 45 |
| 4 Atmospheric Turbulence Characterization over an Horizontal Path | 47 |
| 4.1 Introduction | 48 |
| 4.2 Atmospheric turbulence | 48 |
| 4.3 Cn2 Index for Horizontal Turbulence | 49 |
| 4.4 Proposed Method and Setup | 53 |
| 4.5 Experimental Results | 54 |
| 4.5.1 Atmospheric Turbulence Measurement | 54 |

| | | |
|------------|--|------------|
| 4.5.2 | Statistics of the Zernike modes | 56 |
| 4.5.3 | Atmospheric Turbulence Correction | 58 |
| 4.6 | Conclusion | 62 |
| 5 | SH-WFS on Embedded GPU | 67 |
| 5.1 | Introduction | 68 |
| 5.2 | Moment-based Centroid Extraction | 69 |
| 5.3 | Parallel Implementation | 73 |
| 5.3.1 | Data Levels | 74 |
| 5.3.2 | Optimized GPU Data Transfer through Coalesced Memory Accesses | 75 |
| 5.3.3 | Data Reduction | 77 |
| 5.3.4 | Kernel Algorithm | 79 |
| 5.4 | Experiments | 79 |
| 5.5 | Conclusion | 80 |
| 6 | State of the Art on AO Model-Based Control Strategies | 85 |
| 6.1 | Introduction | 86 |
| 6.2 | System Models in Modal Space | 86 |
| 6.2.1 | Deterministic subsystem | 87 |
| 6.2.2 | Stochastic subsystem | 88 |
| 6.2.3 | Full System Model | 88 |
| 6.2.4 | Modal Phase Representation Error | 89 |
| 6.3 | Kalman Filter | 90 |
| 6.3.1 | Kalman Filter in AO | 91 |
| 6.4 | Linear Quadratic Control | 93 |
| 6.4.1 | Linear Quadratic Control in AO | 94 |
| 6.5 | \mathcal{H}_2 Control | 96 |
| 6.5.1 | \mathcal{H}_2 Control in AO | 97 |
| 6.6 | Model Predictive Control | 99 |
| 6.6.1 | Model Predictive Control in AO | 100 |
| 6.7 | Conclusion | 101 |
| 7 | PI-Shaped LQG Control | 103 |
| 7.1 | Introduction | 104 |
| 7.2 | PI-Shaped LQG controller | 104 |
| 7.2.1 | State Estimator | 105 |
| 7.2.2 | PI-shaped LQ state-feedback matrix | 107 |
| 7.2.3 | Full controller | 108 |
| 7.3 | Simulation Results | 109 |
| 7.4 | Experimental Results | 115 |
| 7.5 | Conclusion | 116 |
| 8 | Conclusions | 119 |
| A | Publications | 121 |
| References | | 126 |

Introduction

Adaptive Optics (AO) is a technique to manipulate the light beams by exploiting its optical properties. AO is integrated into optical setups to achieve different kinds of optics-related goals:

- To enhance the *spatial resolution* of an observed object image (e.g. astronomy);
- To optimize the *contact area* of a laser beam over a surface (e.g. laser cutting);
- To maximize the *bandwidth* of an optical communication.

It is not exact to associate light beam shaping to AO only. In fact, the boundaries of the light shaping field of study are rather large. For example, humans can qualitatively measure whether an object is out of focus or not, when observed through naked eyes. This is considered as the *sensing* part of light shaping, with the goal being to maximize the object image resolution. The human eye corrects for the focus through the *accomodation*, which is the process of optimizing the optical power of the eye to achieve higher spatial resolution. When the eye reaches its correction limit, then humans have to *adjust the distance* from the observed object, either by moving the object or himself. All of this qualifies as the optical *compensation* of the aberrations.

The community agrees into splitting the light optical manipulation into AO and *active optics*. This taxonomic distinction is loosely based upon the light shaper *response time*. The optical correction loop assumes that the light beam shape is first *sensed* and then *compensated* for, therefore implying a time lag between the two phases. Hence, the two classes are given as a function of the correction loop frequency [117]:

- If the correction loop frequency is lower than $10Hz$ (e.g., eye accomodation, bulky auxiliary astronomical deformable mirrors etc.), we refer to *active optics*;
- If the correction loop frequency is higher than $10Hz$ (e.g., fast deformable mirrors, optical communication devices etc.) we are within the field of *adaptive optics*.

Since its inception in the 1953 by Babcock [7], where it was only theorized, AO evolved and made its way into a number of scientific fields. An AO system is usually composed of a *sensor*, a *corrector* and a *controller*. The technological advances over the years gave rise to breakthroughs in the design and capability of

all such AO components. This was possible with the joint advances of the physical optical phenomena understanding, optical components manufacturing and computing techniques. Initially developed for military applications, AO is successfully used nowadays in any scientific telescope to provide diffraction-limited images with resolution exceeding those of the Hubble Space Telescope [40, 101] (i.e. with a telescope outside the atmosphere).

In the absence of aberrations, telescope resolution is limited by diffraction of light. With a circular aperture telescope without obstructions, the image of an ideal point-like source at infinity is called Point Spread Function (PSF) whose size (expressed in radians) is approximately λ/D , where λ is the wavelength at which the object is observed and D the diameter of the telescope. When turbulence affects light propagation distorting the incoming wavefront, the PSF becomes larger and less intense, leading to a loss of resolution. It is not sufficient to increase the telescope diameter D to achieve higher image resolution. In fact, the spatio-temporal variance of the atmospheric turbulence effect over the transversal section of the optical path increases as well (see Chapter 4).

AO is the key to provide almost diffraction-limited images for ground-based astronomical telescopes. In the 1990, Rousset et al. [103] experimentally demonstrated that retrofitting a 1.52m aperture diameter telescope with the COMEON AO system yielded to diffraction-limited celestial images. Nowadays, all large ground-based astronomical telescopes take AO into account as part of their design [106]:

- The 8.2m Very Large Telescope (VLT) in Chile [1]
- The 8.4m Large Binocular Telescope (LBT, dual aperture) in Arizona, USA [2]
- The 10m Keck telescopes (twin) in Hawaii [3]
- The 10.4m Gran Telescopio Canarias (GTC) in the Canary islands [4]
- The 8.4m Giant Magellan Telescope (GMT) in Chile [5]
- The 30m Thirty Meter Telescope (TMT) in Hawaii [6]
- The 39m European Extremely Large Telescope (E-ELT) in Chile [7]

Comparing ideal and aberrated PSFs, it is possible to determine a figure of merit often used by astronomers to quantify the performance of a telescope, the so-called Strehl Ratio (SR). It is defined as the ratio of the maximum intensity between the aberrated and the ideal PSF [36]. To this regard, a better control performance in the AO system means a higher SR which ultimately implies a better quality of object images and scientific throughput.

1 www.eso.org
 2 www.lbto.org
 3 www.keckobservatory.org
 4 www.gtc.iac.es
 5 www.gmto.org, in construction
 6 www.tmt.org, in construction
 7 www.eso.org, in construction

1.1 Thesis Outline

Atmospheric turbulence is a fast disturbance that requires high control frequency. At the same time, celestial objects are faint sources of light and thus WFSs often work in a low photon count regime. These two conditions require a trade-off between high closed-loop control frequency to improve the disturbance rejection performance, and large WFS exposure time to gather enough photons for increasing the Signal-to-Noise ratio (SNR) [100], making the control a delicate yet fundamental aspect for AO systems.

This PhD thesis tackles the AO atmospheric turbulence compensation problem from the controller design point of view:

1. The AO system and atmospheric disturbance are modeled using the *state-space control theory* framework to develop modern control strategies;
2. The controller logic is implemented into *advanced computing architectures* to boost the control loop frequency of small/medium AO setups;
3. A *model-based control technique* is designed to improve upon state-of-the-art AO controllers.

The chapters outline is shown in Figure 1.1

Chapter 2 briefly recalls the background on optics and control theory needed in the following Chapters. The light is modeled as an electro-magnetic wave, of which only the phase component can be effectively measured and manipulated by an AO system. The Zernike polynomials are chosen to describe the wavefront shape and, more importantly, how they are used to reconstruct the wavefront from the phase measurements. The Strehl Ratio, a measure of how well resolved in space is an observed point light source, is calculated from the Zernike expansion. Then, the key elements of any AO system are presented: the wavefront sensor, deformable mirror and controller are modeled as discrete linear time-invariant systems, yielding to the complete system diagram where the AO plant and the atmospheric turbulence models are interconnected.

Chapter 3 describes the PhotonLoop software where the AO controller will be implemented [70]. This tool has been updated over the years to meet the requirements of optical setups belonging to different research areas. PhotonLoop is written in the C++ programming language with the aid of the *Qt* libraries. Beyond the fundamental wavefront sensing and compensation functions, it hosts a comprehensive static calibration routine. State-space based control strategies can be loaded as well as standard PID controllers. The TCP-IP communication can interface with other scientific tools (e.g. *Matlab*) to perform lengthy experimental runs, and the logging feature stores the internal state of each discrete sample to disk. Almost all of the results shown in this PhD thesis are gathered from experiments using the PhotonLoop software.

The requirements for the design of an AO system depends on the characterization of the optical aberration. Using Kolmogorov model, the atmospheric turbulence over a vertical path (e.g. ground-based telescopes observing celestial objects) can be described by the atmospheric structure parameter from look-up tables. However, when the optical path is horizontal (e.g. ground to ground communication), the atmospheric turbulence cannot be characterized by look-up

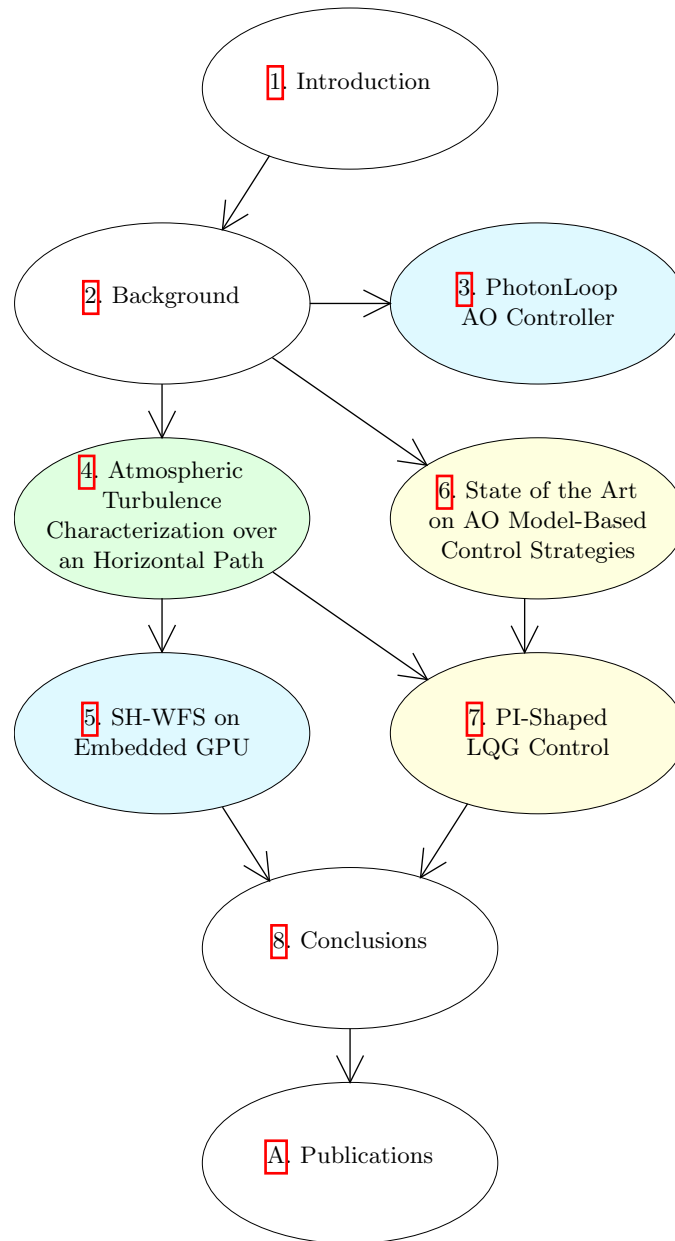


Fig. 1.1: Outline of the chapters in this PhD thesis.

tables and hence must be profiled in a different way. Chapter 4 describes the vertical atmospheric turbulence and proposes a method to characterize the horizontal atmospheric turbulence. Experimental results of the correction of the measured aberration time series reproduced into an optical test-bench are reported as well.

The spatial and temporal resolution requirements for the turbulence compensation pose a challenge for the AO setup design. The wavefront sensor is often the AO system bottleneck, as it must sample the incoming wavefront with high spatial resolution within real-time constraints, which respectively means high data bandwidth and low latency. Standard architectures rely on low-latency FPGA architectures for the implementation of such algorithms. However, FPGA development effort is considerably high. Chapter 5 illustrates how the CPU/GPU edge-computing architecture from Nvidia (the Jetson architecture line-up) can be exploited for AO purposes. While the achieved latency is not as low as the FPGA solution, GPU still compares favorably to on-board CPU and meets closed-loop frequency requirements for small-medium Adaptive Optics systems. This is possible thanks to the unified memory scheme of the Jetson architecture which avoids memory copy between CPU and GPU memory addressing spaces. Detailed experimental results show the effectiveness of the proposed architecture.

AO systems are usually designed with hundreds of measurement inputs and actuator outputs, depending on the correction requirements. Hence, traditional output-feedback control strategies such as the PID regulators must be calibrated almost manually for obtaining good correction performance. On the other hand, model-based control strategies automatically synthesize optimal controllers. Chapter 6 reviews the state of the art of modern model-based control strategies for the single conjugate AO systems. Each control strategy is described and then applied to the AO system within the state-space framework presented in Chapter 2.

Chapter 7 describes the improvement over the modal-based control strategy given in Chapter 6. The AO system described in Chapter 2 is partitioned into mutually independent parallel sub-systems by projecting the wavefront into Zernike modes. Each mode of the turbulence can be approximated as the output of a low-pass filter with a Gaussian white-noise as input (as modeled in Chapter 2 and measured in Chapter 4). For each mode a PI-shaped LQG controller is designed. Finally, the synthesized modal controllers are connected in parallel resulting into the full control strategy.

In Chapter 8 conclusion are drawn and the plan on future work is sketched. Thanks to the joint collaboration with the CNR-IFN of Padova, the PhotonLoop software has been integrated in several small to medium-sized optical setups, resulting in publications on a wide range of research fields (AO instrumentation, laser communication, high energy laser, ophthalmology, microscopy and astronomy) as listed in Appendix A.

Background

Summary. The design of Adaptive Optics requires the joint knowledge of physical optics, optical manufacturing and control theory. In this chapter, the conjugation principle is presented as the back-bone of any Adaptive Optics system. The light beam is modeled as an electromagnetic wave whose wavefront is represented in modal terms with the Zernike polynomials. Then, the most common hardware design for wavefront sensors, deformable mirrors and controller architectures are reviewed. Those components are modeled as discrete-time Linear Time-Invariant (LTI) Multiple-Input Multiple-Output (MIMO) state-space systems. This modeling will be exploited in Chapter [6](#) for designing the optimal controller.

2.1 Introduction

Adaptive Optics (AO) is a technique used in optical systems to correct for optical aberrations that otherwise will limit their performance. This technique, which was originally developed to reduce the effect of atmospheric perturbations in ground-based telescopes [100] [11] [28], is nowadays exploited in many applications where light propagation is affected by ambient turbulence or inhomogeneities in the medium through which light propagates such as: free space laser communication [61] [62], microscopy [46] [15] [14], ophthalmology [48] [88] and beam shaping in high power laser systems [87] [17] [53]. In Single Conjugate AO (SCAO) systems, aberrations are corrected by measuring the wavefront phase with a WaveFront Sensor (WFS) and optically conjugating it with a Deformable Mirror (DM), as seen in Figure 2.1. These two elements are connected to a control system which, in order to perform an optimal correction, must work in real-time and have low latency to achieve a cutoff frequency higher than the aberration bandwidth. Furthermore, since AO systems are usually employed to improve the performance of scientific instruments, they should provide robust behavior and easy-to-use interface to the operators.

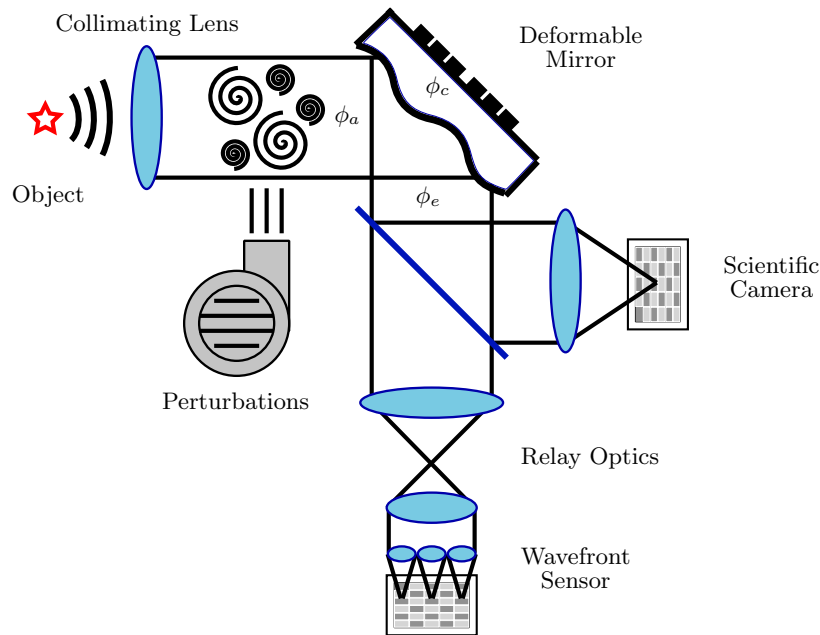


Fig. 2.1: Simplified schema of an Adaptive Optics system: a light source (laser or scientific object) is observed on a Scientific Camera after the correction with a Deformable Mirror (DM). The wavefront distortion is measured by a Wavefront Sensor (WFS) whose lenslets are optically conjugated to the deformable mirror by means of a relay optics.

An optical imaging system is composed by a series of optical elements (such as lenses and mirrors) that capture the light coming from an object and recreate its image in the image plane. Light is defined as a bundle of electro-magnetic waves, each having amplitude A and phase ϕ [100]

$$\Psi = Ae^{i\phi}, \quad (2.1)$$

traveling through the optical path. Due to the wave-like behavior of light, optical systems are characterized by a *Point Spread Function* (PSF) that is the image of an observed point-like source. The image of any observed object is the convolution of the geometrical image of the object with the PSF [116]. In case of circular apertures, such as those typically employed in optics, the PSF image is defined as the Airy Disk (Figure 2.2), having 84% of its energy in the central spot and the remaining energy in regularly spaced rings around the center. In ideal, aberration-free optical systems, the diffraction limit of optical elements determines the minimum size of the PSF and hence the upper limit of the optical system resolution.

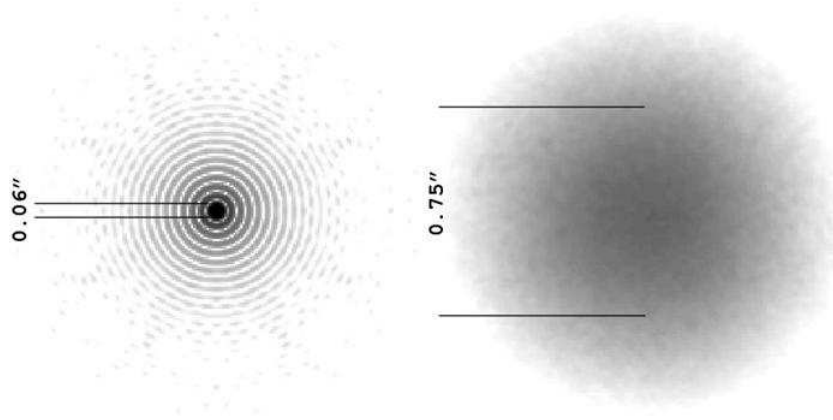


Fig. 2.2: Left: diffraction-limited Airy disk, i.e. no aberrations. Most of the energy is contained into the inner ring. Right: Airy disk aberrated by atmospheric turbulence. The energy is scattered, leading to loss of spatial resolution.

AO allows to correct the aberrations of optical systems and to increase their resolution by applying the *phase conjugation principle*, which consists in reversing the phase of optical aberrations to compensate for the wavefront distortions [116]. This task is performed by measuring the wavefront phase with a wavefront sensor and correcting it with a phase modulation element such as a deformable mirror. Usually, these two elements are placed in planes that are optically conjugated to the ones where aberrations are generated.

In a typical closed-loop AO architecture, the WFS observes the residual wavefront after the DM compensation of the incoming aberrations. The involved signals are

- ϕ_a : the incoming aberrated wavefront;
- ϕ_c : the DM-generated correction wavefront;
- ϕ_e : the residual wavefront after the DM correction,

as shown in Fig. 2.1.

The wavefront phase is a continuous function of time and space. Let $t \in \mathbb{R}$ be the time and $\mathbf{r} \in \mathbb{R}^2$ a point on the support \mathbb{A} . The signal $\phi(t, \mathbf{r})$ completely describes the spatial/temporal evolution of the wavefront phase. In the common case of a circular aperture of radius r , \mathbb{A} is expressed in Cartesian coordinates as

$$\mathbb{A} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r^2\}, \quad (2.2)$$

or, alternatively, in polar coordinates as

$$\mathbb{A} = \{(\rho, \theta) \in \mathbb{R}^2 \mid 0 \leq \rho \leq r; 0 \leq \theta < 2\pi\}. \quad (2.3)$$

The wavefront ϕ is continuous in every point, hence it is not trivial to compute its discrete analytical approximation [54] [43]. We will start introducing the modal expansion using Zernike polynomials and then we will explain how a WFS computes its discretization in time and space.

2.2 Zernike Polynomials

The wavefront phase can be represented in Cartesian space by means of a suitable basis $\{Z_j\}$ such that:

$$\phi(t; \mathbf{r}) = \sum_{j=1}^{+\infty} w^{(j)}(t) Z_j(\mathbf{r}), \quad (2.4)$$

with Z_j being the j -th spatial mode and $w^{(j)}$ its coefficient. Many bases can be used to represent $\phi(t; \mathbf{r})$. The most popular one uses the Zernike polynomials because of their straightforward optical interpretation and orthogonality property over a continuous-space circular support [117] [124].

The Zernike polynomials are formulated by Noll [82] and use a double indexing scheme for radial and azimuthal orders:

$$Z_n^m = \begin{cases} N_n^m R_n^m(\rho) \cos(m\theta), & m \geq 0 \\ -N_n^m R_n^m(\rho) \sin(m\theta), & m < 0 \end{cases}, \quad (2.5)$$

where

$$N_n^m = \begin{cases} \sqrt{n+1}, & m = 0 \\ \sqrt{2n+2}, & m \neq 0 \end{cases} \quad (2.6)$$

and

$$R_n^m(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s}, \quad (2.7)$$

with the support defined in the polar coordinates ρ and θ as in (2.3).

Throughout this thesis, the Zernike polynomials numbering is defined by the single indexing scheme (ANSI standard). Nonetheless, the single and double indexing are related by the bijective function

$$j = \frac{n(n+1)}{2} + \frac{n+m}{2}. \quad (2.8)$$

Zernike modes are grouped by radial orders, with higher orders corresponding to higher spatial frequency. Since the DM is unable to reproduce high order Zernike modes due to its finite number of actuators, the expansion introduced in Eq. (2.4) is truncated to the first N_w terms and hence the phase ϕ is approximated by

$$\phi(t; \mathbf{r}) \sim \bar{\phi}(t; \mathbf{r}) = \sum_{j=0}^{N_w} w^{(j)}(t) Z_j(\mathbf{r}). \quad (2.9)$$

A list of the first Zernike terms is reported in Table 2.1 and the corresponding wavefronts are shown in Figure 3.8.

Table 2.1: First Zernike polynomial expressions according to ANSI notation, expressed in Cartesian coordinates: i is the ANSI index, n and m are the radial and azimuthal degrees.

| Z_j | n | m | Name | Formula |
|-------|-----|-----|------------------------|-------------------------------|
| Z_0 | 0 | 0 | Piston | 1 |
| Z_1 | 1 | -1 | Tilt | $2y$ |
| Z_2 | 1 | 1 | Tip | $2x$ |
| Z_3 | 2 | -2 | Astigmatism 45° | $2\sqrt{6}xy$ |
| Z_4 | 2 | 0 | Defocus | $\sqrt{3}(2x^2 + 2y^2 - 1)$ |
| Z_5 | 2 | 2 | Astigmatism 0° | $\sqrt{6}(x^2 - y^2)$ |
| Z_6 | 3 | -3 | Trefoil 0° | $\sqrt{8}(3x^2y - y^3)$ |
| Z_7 | 3 | -1 | Coma 0° | $\sqrt{8}(3x^2y + 3y^3 - 2y)$ |
| Z_8 | 3 | 1 | Coma 90° | $\sqrt{8}(3x^3 + 3xy^2 - 2x)$ |
| Z_9 | 3 | 3 | Trefoil 30° | $\sqrt{8}(x^3 - 3xy^2)$ |

The Zernike polynomials are analytic functions in their support. Since several wavefront sensors are only able to measure wavefront phase gradients (slopes), this means that the wavefront phase can be reconstructed in the least squares sense using the modal estimation [116]. The relationship between phase gradient and Zernike terms differentiated over space \mathbf{r} is:

$$\frac{d}{d\mathbf{r}} \bar{\phi}(t, \mathbf{r}) = \frac{d}{d\mathbf{r}} \left(\sum_{j=1}^{N_w} w^{(j)}(t) Z_j(\mathbf{r}) \right) \quad (2.10)$$

$$= \sum_{j=1}^{N_w} w^{(j)}(t) \frac{d}{d\mathbf{r}} Z_j(\mathbf{r}). \quad (2.11)$$

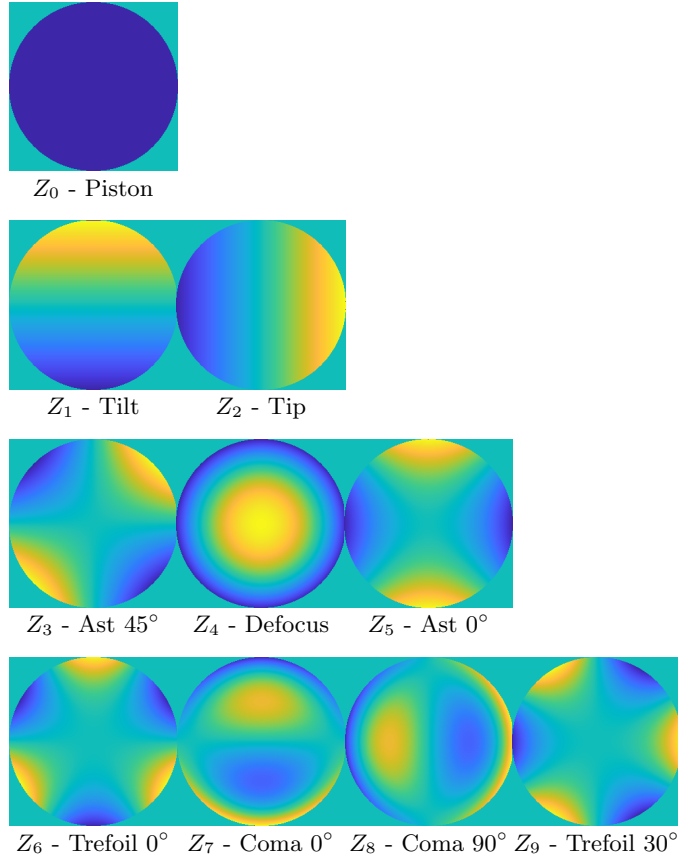


Fig. 2.3: Wavefront representation of the first 10 Zernike terms, ANSI ordered from top-left to bottom-right, grouped by radial order.

where the Piston term Z_0 is dropped¹.

Let $\mathbf{y} \in \mathbb{R}^{2L}$ be the discrete slope measurements vector defined in Cartesian coordinates as:

$$\mathbf{y}(k) = [y_1^x(k) \ y_2^x(k) \ \dots \ y_L^x(k) \ | \ y_1^y(k) \ y_2^y(k) \ \dots \ y_L^y(k)]^\top, \quad (2.12)$$

where each pair $(y_l^x(k); y_l^y(k))$ is the l -th local derivative of the wavefront at time $t = kT_s$ at the Cartesian coordinates (x_l, y_l) . Then, the slopes are related to Zernike derivatives shown in (2.11) as:

¹ The Piston term Z_0 is the average height of the wavefront surface. it is undetectable by the majority of wavefront sensors and is therefore removed. This has no impact on the performance of the AO correction.

$$y_l^x(k) \simeq \sum_{j=1}^{N_w} w^{(j)}(k) \left. \frac{dZ_j(x, y)}{dx} \right|_{\substack{x=x_l \\ y=y_l}}; \quad (2.13)$$

$$y_l^y(k) \simeq \sum_{j=1}^{N_w} w^{(j)}(k) \left. \frac{dZ_j(x, y)}{dy} \right|_{\substack{x=x_l \\ y=y_l}}, \quad (2.14)$$

which written in matrix form yields to

$$\mathbf{y}(k) = \Pi_{w2y} \mathbf{w}(k) \quad (2.15)$$

where the coefficient vector is

$$\mathbf{w}(k) \in \mathbb{R}^{N_w} = [w_1(k) \dots w_{N_w}(k)]^T \quad (2.16)$$

and the Zernike derivative projector mapping $\mathbf{w}(k)$ into slopes $\mathbf{y}(k)$ is

$$\Pi_{w2y} \in \mathbb{R}^{N_w \times 2L} = \begin{bmatrix} \frac{dZ_1(x_1, y_1)}{dx} & \dots & \frac{dZ_1(x_L, y_L)}{dx} & \left| \right. & \frac{dZ_1(x_1, y_1)}{dy} & \dots & \frac{dZ_1(x_L, y_L)}{dy} \\ \vdots & \ddots & \vdots & \left| \right. & \vdots & \ddots & \vdots \\ \frac{dZ_{N_w}(x_1, y_1)}{dx} & \dots & \frac{dZ_{N_w}(x_L, y_L)}{dx} & \left| \right. & \frac{dZ_{N_w}(x_1, y_1)}{dy} & \dots & \frac{dZ_{N_w}(x_L, y_L)}{dy} \end{bmatrix}^T. \quad (2.17)$$

The least-squares inversion of Π_{w2y} yields to the Zernike reconstructor $\Pi_{y2w} \in \mathbb{R}^{2L \times N_w}$ that projects slopes into Zernike coefficients:

$$\mathbf{w}(k) = \Pi_{y2w} \mathbf{y}(k). \quad (2.18)$$

Since higher order Zernike terms are characterized by higher spatial resolution, there is a threshold for the number of Zernike terms N_w below which spatial aliasing does not occur. On the other hand, the wavefront reconstructed from too few modal terms can have large reconstruction error. Therefore, the number of Zernike terms N_w is optimized to minimize the wavefront fitting error $\sigma_{>N_w}^2$ while avoiding aliasing artifacts. This is of the utmost importance when designing control strategies that rely on the modal decomposition to describe the wavefront to compensate for.

It is worth noting that the Zernike basis is orthogonal in space but not in time, i.e. $\mathbf{E}[w^{(j)}w^{(i)}] \neq \sigma_i^2 \delta_{i,j}$. A basis which is statistically orthogonal over both time and space is the Karhunen-Loève expansion. However, such basis cannot be written analytically but only computed numerically starting from a spatio-temporal model of the atmospheric turbulence. On the other hand, the analytical description of the Zernike polynomials and their derivatives are readily available.

Other bases can be chosen depending on the spatial shape of the wavefront support, like the annular Zernike [65] or the rectangular Legendre terms [125].

2.3 Strehl Ratio

Optical systems are affected by aberrations that, by modifying the wavefront phase, lead to a broadening of the PSF and thus to a decreasing of the system resolution. A figure of merit indicating the quality of an optical system is

the *Strehl Ratio* (SR), i.e. the ratio between peak intensity of the aberrated PSF over a perfect PSF. Since the PSF is the image of the observed point-like reference light source, near the desired object of interest (i.e. suffering from the same optical aberrations), it can be measured by a scientific far-field camera as shown in Figure 2.1. An alternative method to calculate the SR from wavefront phase aberrations, valid when the real SR is larger than 30%, is given by Maréchal approximation:

$$\text{SR} = 1 - \frac{2\pi}{\lambda}(\Delta\phi)^2 \quad (2.19)$$

$$\cong \exp\left[-\frac{2\pi}{\lambda}(\Delta\phi)^2\right], \quad (2.20)$$

where ϕ is the wavefront phase, λ is the beam wavelength and

$$(\Delta\phi)^2 = \frac{\iint_{\mathbb{A}} (\phi - \bar{\phi})^2 dx dy}{\iint_{\mathbb{A}} dx dy} \quad (2.21)$$

is the phase variance over the optical system aperture \mathbb{A} .

A property of the Zernike coefficients is that the wavefront variance at the sample time k can be approximated as the squared sum of all the reconstructed Zernike terms [116]

$$(\Delta\phi(k))^2 \sim \sum_{j=1}^{N_w} w^{(j)}(k)^2. \quad (2.22)$$

Then, the Strehl ratio can be estimated by substituting (2.22) into (2.19). The approximation also holds for time series of N_k samples:

$$(\Delta\phi)^2 \sim \frac{1}{N_k + 1} \sum_{k=0}^{N_k} \sum_{j=1}^{N_w} w^{(j)}(k)^2. \quad (2.23)$$

In astronomy, there is a difference between the SR calculated from short and long exposures. Usually, the reference guide star (or laser guide star) observed by the wavefront sensor is bright enough to guarantee high signal-to-noise ratio even at short exposure. However, the light incoming from faint celestial objects observed by scientific cameras must be exposed within larger times to collect enough photons. Hence, the wavefront sensor for compensating the atmospheric turbulence and the scientific camera operate at different frequencies. For this reason we are interested in both Short Exposure (SE) and Long Exposure (LE) Strehl ratio, for measuring AO performance and the scientific camera image resolution, respectively. In particular:

- the *Short Exposure (SE) Strehl ratio* is calculated from the RMS computed as in (2.22). In this case a new value for the SE Strehl is available any time a new set of slopes is measured;
- the *Long Exposure (LE) Strehl ratio* is calculated from the sum of the variance of the Zernike coefficients time series $w^{(j)}(\cdot)$ as in (2.23).

2.4 State-space Formalization

The whole AO system, together with the atmospheric turbulence, can be formalized within the modern control theory framework. This approach brings several advantages:

- Within a good approximation, the AO system can be regarded as a *discrete-time MIMO LTI system*, which greatly simplifies the analysis;
- The modern state-space control theory provides reliable routines to design *optimal estimators and controllers*.

The components that form the AO system are rather heterogeneous. The incoming light seen as an electromagnetic wave is continuous both in space and time, whereas measurements and command signals are discrete. In fact, the WFS samples the wavefront in space and time. Moreover, the disturbance processes are inherently stochastic, in contrast to the AO setup which is strictly deterministic.

From our purpose, it is sufficient to model the AO system using a space- and time-discrete system. The state-space realization of a discrete-time Linear Time-Invariant (LTI) system \mathbf{G} of size N , with P inputs and M outputs is

$$\mathbf{G} : \begin{cases} \mathbf{x}(k+1) &= A\mathbf{x}(k) + B\mathbf{u}(k) \\ \mathbf{y}(k) &= C\mathbf{x}(k) + D\mathbf{u}(k) \end{cases}, \quad (2.24)$$

where:

- $\mathbf{x}(k) \in \mathbb{R}^N$, $\mathbf{u}(k) \in \mathbb{R}^P$ and $\mathbf{y}(k) \in \mathbb{R}^M$ are the state, input and output vectors, respectively;
- $A \in \mathbb{R}^{N \times N}$ is the system matrix which determines the evolution of the state over time;
- $B \in \mathbb{R}^{N \times P}$ is the input matrix;
- $C \in \mathbb{R}^{M \times N}$ is the output matrix;
- $D \in \mathbb{R}^{M \times P}$ is the feedthrough matrix mapping the input directly into the output.

The system (2.24) can also be written as the tuple $\mathbf{G} = \{A, B, C, D\}$ or as

$$\mathbf{G} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] : \mathbf{u}(k) \mapsto \mathbf{y}(k), \quad (2.25)$$

where the state vector is implicit.

The transfer function of a state-space realization is

$$\mathbf{G}(z) = C(zI - A)^{-1}B + D, \quad (2.26)$$

where I is the identity and z is the variable of the \ddagger -transform. Signals and systems will also be occasionally written as a mix-up of functions in k (time) and z (frequency) to take advantage of both writing conventions:

$$\mathbf{y}(k) = \mathbf{G}(z)\mathbf{u}(k) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] \mathbf{u}(k). \quad (2.27)$$

The block diagram in Figure 2.4 shows the SCAO configuration as a connection of continuous- and discrete-time systems:

- ϕ_a, ϕ_c and $\phi_e \in \mathbb{R} \times \mathbb{A}$ are the continuous time signals representing the atmospheric, controlled and residual wavefronts, respectively;
- ϕ_e is sampled into the discrete variable $\mathbf{y}(k)$ by the WFS, which is also affected by Gaussian zero-mean white measurement noise $\boldsymbol{\eta}(k)$ related to the exposure (integration) time T_e ;
- $\mathbf{u} \in \mathbb{R}^P$ is the discrete Zero-Order Hold (ZOH) command signal fed to the P -actuators of the DM;
- $\mathbf{v}(k)$ is the Gaussian white noise of covariance Σ_v driving the atmospheric system.

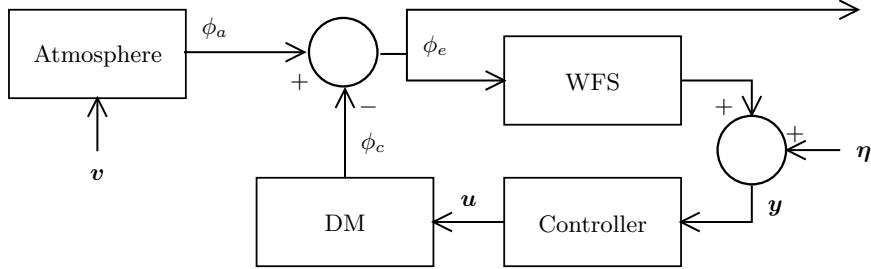


Fig. 2.4: Control Loop block diagram. The atmosphere is assumed to be a stochastic process driven by Gaussian noise \mathbf{v} . ϕ_a , ϕ_c and ϕ_e are the aberrated, controlled and residual wavefronts respectively. \mathbf{y} is the error vector measured by the WFS, combined with additive Gaussian noise $\boldsymbol{\eta}$. \mathbf{u} is the command vector sent to the DM. The dashed lines represent the discrete nature of the signals.

2.4.1 Atmospheric Turbulence

The atmospheric turbulence is a highly non-linear physical phenomenon which alters the refractive index of the optical path. Only one transversal plane of the optical path can be conjugated with the DM of a SCAO system. Since such plane collects the whole aberration effects from the source object to the plane itself, it is collocated to the primary mirror of the telescope.

The wavefront ϕ_a of the conjugated plane represents the atmospheric disturbance. Despite the fact that the atmospheric turbulence model lies in the continuous-time part of the system, we will model it using the discrete-time and discrete-space approximation $\mathbf{y}_a(k) \in \mathbb{R}^M$ as in Figure 2.5.

The evolution of $\mathbf{y}_a(k)$ is modeled as the output of a stochastic discrete-time transfer function

$$\mathbf{G}_a(z) = \begin{bmatrix} A_a & B_a \\ C_a & 0 \end{bmatrix} : \mathbf{v}(k) \mapsto \mathbf{y}_a(k) \quad (2.28)$$

$$= \begin{cases} \mathbf{x}_a(k+1) & = A_a \mathbf{x}_a(k) + B_a \mathbf{u}(k) \\ \mathbf{y}_a(k) & = C_a \mathbf{x}_a(k) + D_a \mathbf{v}(k) \end{cases}, \quad (2.29)$$

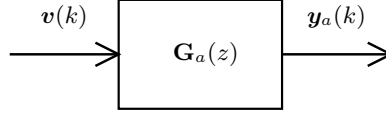


Fig. 2.5: The atmospheric turbulence discrete model \mathbf{G}_a . The zero-mean, white Gaussian noise $\mathbf{v}(k)$ is the process noise and is assumed with unit variance. The output $\mathbf{y}_a(k)$ is the wavefront represented as WFS measurements.

where A_a is Hurwitz² and $\mathbf{v}(k)$ is Gaussian, zero-mean white process noise.

The atmospheric turbulent covariance Σ_ϕ is approximated by the covariance matrix of $\mathbf{y}_a(k)$:

$$\Sigma_{\mathbf{y}_a} = \mathbb{E} [\mathbf{y}_a(k) \mathbf{y}_a(k)^\top], \quad (2.30)$$

where $\mathbb{E}[\cdot]$ is the expectation operator. Let

$$\Sigma_{\mathbf{x}_a} = \mathbb{E} [\mathbf{x}_a(k) \mathbf{x}_a(k)^\top]; \quad (2.31)$$

$$\Sigma_{\mathbf{v}} = \mathbb{E} [\mathbf{v}(k) \mathbf{v}(k)^\top] = I, \quad (2.32)$$

be the covariance matrices of the state $\mathbf{x}_a(k)$ and process noise $\mathbf{v}(k)$, respectively. If the ergodic theorem holds, then $\Sigma_{\mathbf{y}_a}$ is given by

$$\Sigma_{\mathbf{y}_a} = C_a \Sigma_{\mathbf{x}_a} C_a^\top, \quad (2.33)$$

where

$$\Sigma_{\mathbf{x}_a} = A_a \Sigma_{\mathbf{x}_a} A_a^\top + B_a \Sigma_{\mathbf{v}} B_a^\top. \quad (2.34)$$

It is useful to model the atmosphere as an equivalent system $\bar{\mathbf{G}}_a$ on the Zernike basis. The dynamic relationship among the N_w turbulence modes projected into the Zernike space are:

$$\bar{\mathbf{x}}_a(k) = \begin{bmatrix} \mathbf{x}_a^{(1)}(k) \\ \mathbf{x}_a^{(2)}(k) \\ \vdots \\ \mathbf{x}_a^{(N_w)}(k) \end{bmatrix}, \quad \bar{A}_a = \begin{bmatrix} A_a^{(11)} & A_a^{(12)} & \dots & A_a^{(1N_w)} \\ A_a^{(21)} & A_a^{(22)} & & \\ \vdots & & \ddots & \vdots \\ A_a^{(N_w 1)} & & \dots & A_a^{(N_w N_w)} \end{bmatrix}, \quad (2.35)$$

where $\bar{A}_a \in \mathbb{R}^{N_w T \times N_w T}$ is composed of sub-matrices $A_a^{(jj')} \in \mathbb{R}^{T \times T}$ defining the Zernike evolution, each modeled as a T -order Auto-Regressive (AR) process. The state $\bar{\mathbf{x}}_a(k) \in \mathbb{R}^{N_w T}$ is the composition of the N_w sub-states $\mathbf{x}_a^{(j)}(k) \in \mathbb{R}^T$. Entries of \bar{A}_a outside the diagonal are related to the cross-correlation between j -th and j' -th Zernike modes. Fedrigo *et al.* [28] consider the turbulence as a connection of independent Karhunen-Love modes with no cross-correlation. In that case, since the modal base would be statistically orthogonal, we will end up with $A_a^{(jj')} = \mathbf{0}$, $\forall j \neq j'$.

² A Hurwitz matrix has its eigenvalues within the unit circle, i.e. the corresponding modes are asymptotically stable.

Taking into account the Zernike projector into phase from Eq. (2.36), the process noise $\mathbf{v}(k)$ is related to $\mathbf{y}_a(k)$ by the following equation:

$$\mathbf{y}_a(k) = \Pi_{\mathbf{w}2\mathbf{y}} \bar{\mathbf{G}}_a(z) \mathbf{v}(k). \quad (2.36)$$

The simplest dynamic description for the modes of the atmospheric turbulence is by using low-order AR models (i.e. AR1) with no cross-correlation among them. For example, the AR1 SISO input-output equation for the j -th Zernike mode is given by

$$w_a^{(j)}(k+1) = A_a^{(jj)} w_a^{(j)}(k) + B_a^{(j)} v^{(j)}(k), \quad (2.37)$$

where $A_a^{(jj)}$ is the AR pole and $B_a^{(j)}$ is the input gain related to the variance of $v^{(j)}(k)$ in (2.34). The state-space system of the AR1 atmospheric turbulence model equals to

$$\begin{cases} \mathbf{w}_a(k+1) &= \bar{A}_a \mathbf{w}_a(k) + \bar{B}_a \mathbf{v}(k) \\ \mathbf{y}_a(k) &= \bar{C}_a \mathbf{w}_a(k+1) \end{cases}, \quad (2.38)$$

where

$$\mathbf{w}_a(k) = \begin{bmatrix} w_a^{(1)}(k) \\ w_a^{(2)}(k) \\ \vdots \\ w_a^{(N_w)}(k) \end{bmatrix}, \quad \bar{A}_a = \begin{bmatrix} A_a^{(1)} & & & \mathbf{0} \\ & A_a^{(j)} & & \\ & & \ddots & \\ \mathbf{0} & & & A_a^{(N_w)} \end{bmatrix}, \quad (2.39)$$

$$\bar{B}_a = \begin{bmatrix} B_a^{(1)} \\ B_a^{(2)} \\ \vdots \\ B_a^{(N_w)} \end{bmatrix}, \quad \bar{C}_a = \Pi_{\mathbf{w}2\mathbf{y}}. \quad (2.40)$$

The underlying assumption in (2.39) is that we neglect the temporal correlation among the N_w Zernike modes. As we will show later in Chapter 4, this assumption allows to easily identify each AR model by time series of the Zernike coefficient $w^{(j)}(k)$, $j = 1, \dots, N_w$.

2.4.2 Wavefront Sensor

A Wavefront Sensor (WFS) is a device that measures the incoming light wavefront, that is, maps the continuous phase $\phi(t, \mathbf{r})$ into the discretized measurement $\mathbf{y}(k)$, as shown in Figure 2.6. Since current technology is unable to directly measure the *phase* component of the wavefront, the wavefront distortion must be estimated from the measured *intensity* maps acquired by a CCD or CMOS camera. Hence, all WFSs share the same principle [18]:

1. The incoming light is captured by the camera pixels during the *exposure* time;
2. A suitable algorithm reconstructs the wavefront from the measured intensity;

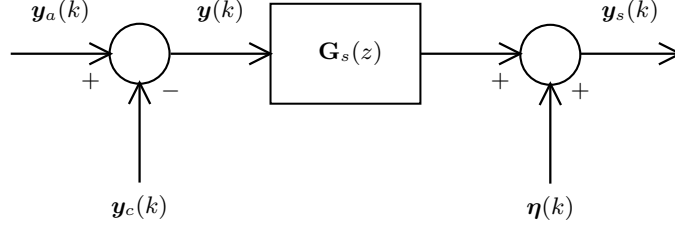


Fig. 2.6: The WFS model. The input signal $\mathbf{y}(k)$ is the residual wavefront resulting from the difference of the aberrated wavefront $\mathbf{y}_a(k)$ and the DM correction $\mathbf{y}_c(k)$. The measurement noise $\boldsymbol{\eta}(k)$ is a zero-mean, white Gaussian noise and is added to the WFS output yielding to the measurement output $\mathbf{y}_s(k)$.

3. Measurements of the distortion are provided.

Let T_s and T_e be the sampling period and exposure time, respectively, with $T_e < T_s$. The WFS can be modeled as a Sample-And-Hold (S&H) device of period T_s mapping $\phi_e(t, \mathbf{r})$ into the vector $\mathbf{y}(k) \in \mathbb{R}^M$:

$$\mathbf{y}(k) = \text{WFS}(\phi(t, \mathbf{r}), kT_s) \quad (2.41)$$

$$= D \left\{ \frac{1}{T_e} \int_{kT_s - T_e}^{kT_s} \phi(\tau, \mathbf{r}) d\tau \right\} + \boldsymbol{\eta}(k) \quad (2.42)$$

$$= [y_1(k) \ y_2(k) \ \dots \ y_M(k)]^\top, \quad (2.43)$$

with discrete time sample $k \in \mathbb{Z}$ and discrete space sample $m = 1, \dots, M$. The spatial discretization is described by the spatial sampling operator D . The Gaussian zero-mean measurement noise $\boldsymbol{\eta} \in \mathbb{R}^M$ with covariance $\Sigma_{\boldsymbol{\eta}}$ depends upon the operator D and the exposure time T_e . Since atmosphere, WFS and DM are related by

$$\phi_e(t, \mathbf{r}) = \phi_a(t, \mathbf{r}) - \phi_c(t, \mathbf{r}), \quad (2.44)$$

the discrete time measurement equation is given by

$$\mathbf{y}(k) = \mathbf{y}_a(k) - \mathbf{y}_c(k) + \boldsymbol{\eta}(k), \quad (2.45)$$

where $\mathbf{y}_a(k)$ is the sampled atmosphere (modeled as in (2.39) and (2.38)) and $\mathbf{y}_c(k)$ the sampled correction vectors both belonging to \mathbb{R}^M . The overall impact of the measurement error is encapsulated into the WFS error σ_{WFS}^2 , whereas the finite WFS sampling period T_s adds to the temporal correction error $\sigma_{T_s}^2$.

The most used WFS is the Shack-Hartmann WFS (SHWFS). By focusing the incoming light with a dense array of micro-lenslets, the beam is effectively spatially partitioned into several spots over the detector pixel plane. Each spot position over the pixels plane is related to the local wavefront gradient (*slopes*). The SHWFS is an improvement over the Hartmann plate test, in which light passes through a perforated sheet mask.

For challenging astronomy applications, the pyramid WFS is also used [45]. The beam is focused into the vertex of a square pyramid prism, which scatters the incoming light in four directions. This has the effect of imaging the beam into

four distinct detector pixel areas. The slope at a given position over the wavefront is calculated from the measurements of the intensity at the relative positions over the four imaged beams.

Other WFS technologies are:

- The holographic WFS, which spatially samples the wavefront into modes;
- The curvature sensor, which measures the wavefront intensities at two out-of-focus planes;
- The pseudo-direct WFS approaches, which estimate the wavefront by extracting the image properties of the PSF (using deep learning or other techniques).

2.4.3 Deformable Mirrors

A Deformable Mirror (DM) compensates for the incoming wavefront aberrations. By shaping its surface, the DM produces a wavefront phase $\phi_c(t, \mathbf{r})$. The inputs are the driving commands $\mathbf{u}(k) \in \mathbb{R}^P$ sent to the DM actuators for shaping the mirror to cancel out the incoming aberrated wavefront. The DM is shown in Figure 2.7 where $\mathbf{y}_c(k)$ is the discrete-time and -space output. Different technologies bring

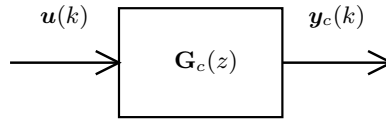


Fig. 2.7: The DM model. The DM changes its shape resulting into the correction ϕ_c as in its discrete-time and -space approximation $\mathbf{y}_c(k)$ upon receiving the command set $\mathbf{u}(k)$.

to different DM. The DM technology depends on the AO error budget of a given scenario. With the notable exception of very specific DM, such as the auto-focus element of a DLSR camera, a DM has at least two or more *actuators* that operate jointly to shape the wavefront surface. Therefore, a DM is characterized by the following spatio-temporal properties:

- *Number of actuators*;
- *Actuator pitch*: distance between neighbouring actuators;
- *Actuator stroke* (mechanical): vertical displacement along the axis of the actuator;
- *Actuator dynamics*: rise time, DM resonances and non-linearities.
- *Optical surface*: continuous, segmented or refractive (for lenses instead of mirrors).

These properties give information about the achievable *spatial resolution* (e.g. Zernike modes) and *temporal resolution* (e.g. control loop bandwidth) that determine the DM wavefront fitting error σ_{fit}^2 , i.e. how well the DM reconstructs the correction wavefront. In addition, there are several DM aspects that should not be overlooked. For example, AO systems for high-power laser operate under harsh environment. If the DM materials are sensitive to temperature variations, then an

efficient heat dissipation technique is required. Also, while a small percentage of failing actuators in a DM with high actuators count does not severely hinder the AO correction performance, a DM with replaceable actuators is preferable in some situations.

The dominant DM technologies are:

- *Stacked array DM*: Such DM are made by assembling a reflective optical surface on top of the actuators array (Figure 2.8). The actuators can be piezo-electric or electrostrictive and are fixed to a rigid base plate. Each actuator longitudinally elongates when stimulated by the electric field induced by applied voltage.

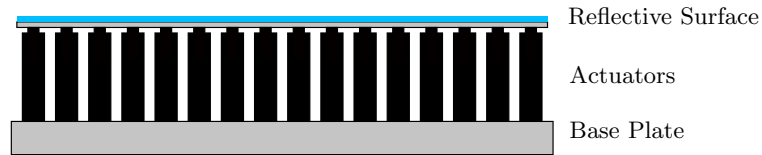


Fig. 2.8: Stacked array DM. The actuators, attached to the base plate, move longitudinally and shape the reflective surface.

The stacked array DM is desirable for the flexible spatial resolution that can be adapted to many AO setups, aside from non-linearities and manufacturing costs. Also, the actuators offer large strokes and higher resonant frequencies compared to other technologies.

- *Bimorph DM*: while in stacked array DM the actuators are placed longitudinally, in bimorph DM the actuators are sandwiched into two transversal layers of piezo-electric material. By the transverse piezoelectric effect principle, the voltage applied to an electrode between the two layers alters the local curvature of the optical plate. Bimorph DM are less expensive than stacked array DM, but require higher driving voltages. In addition, the mechanical mount exhibits resonant behaviours at lower frequencies.
- *Voice-coil DM*: instead of physically coupling the optical element to the actuators, voice-coil DM lets the reflective surface to float on the magnetic field generated by voice-coil actuators as shown in Figure 2.9. By adjusting the current flowing in an actuator, the generated magnetic field locally shape the optical surface. A very fast servo loop measures the effective displacement of the optical surface and dampens unwanted resonancies.

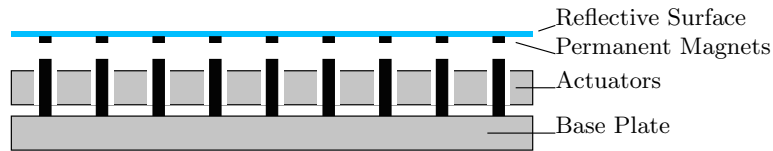


Fig. 2.9: Voice-coil DM. The permanent magnets attached to the reflective surface are influenced by the magnetic field generated by the actuators.

The voice-coil DM offers higher actuated strokes than the other technologies and is more maintainable since the actuators are contactless. However, there are constraints on the actuators pitch.

- *MEMS DM*: Micro Electro-Mechanical System (MEMS) DM exploits semi-conductors technology to miniaturize the actuators. The electrostatic (or electromagnetic) fields actuate a flexible layer on which the optical element lies. The miniaturization guarantees a large amount of actuators, tight pitch and no hysteresis. In the case of electromagnetic actuation, the stroke is comparable to those of stacked arrays.
- *Photo-controlled DM*: the photo-controlled DM uses a photoconductive substrate actuated by a display underneath. The photons emitted by the display are converted by the substrate into an electric field that shapes the mirror layer.

Since each pixel of the display is an actuator, the photo-controlled DM avoid the large number of wires of other DMs and offers very dense actuators pitch. On the other side, the stroke is limited and the desired wavefront must be carefully encoded into the pixel image.

- *Refractive Lens*: the refractive (or multi-actuator) lens shapes the wavefront by changing the refractive index of the optical path. A liquid optical medium with refractive index higher than air is placed between two lenses. A circular array of piezo-electric actuators is attached to the outer rim of the lenses. When an actuator elongates, the lens changes its curvature toward the lens center and thus alters the distance from the other lens. Hence, light passes through the liquid medium with uneven volume, altering its wavefront. By using refractive instead of reflective optical elements, the refractive DM can be placed almost seamlessly into any optical setup.

2.4.4 AO Controller

In the SCAO configuration, the controller $\mathbf{C}(z)$ minimizes the measured residuals $\mathbf{y}(k)$ by actuating the DM with the commands $\mathbf{u}(k)$:

$$\mathbf{C}(z) = \begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix} : \mathbf{y}(k) \mapsto \mathbf{u}(k), \quad (2.46)$$

where

$$\mathbf{y}(k) = \mathbf{y}_a(k) - \mathbf{y}_c(k) + \boldsymbol{\eta}(k), \quad (2.47)$$

with $\boldsymbol{\eta}(k)$ the measurements error.

The most commonly used AO controller is the Proportional-Integrative (PI) compensator [70], whose difference equation is

$$\mathbf{u}(k) = \mathbf{u}(k-1) + K_P \Pi_{\mathbf{y}2\mathbf{u}}(\mathbf{y}(k) - \mathbf{y}(k-1)) + K_I T_s \Pi_{\mathbf{y}2\mathbf{u}} \mathbf{y}(k), \quad (2.48)$$

where

$$\Pi_{\mathbf{y}2\mathbf{u}} = \Pi_{\mathbf{u}2\mathbf{y}}^\dagger \quad (2.49)$$

is the so-called Control Matrix (CM) obtained by pseudo-inverting of the IM, and $K_P, K_I \in \mathbb{R}$ are the proportional and integrative matrix gains, respectively.

The optimal control framework is a powerful tool derived from modern control system theory [32]. It improves output-feedback control strategies by taking into account the knowledge of the plant model. Given:

- a discrete-time LTI MIMO system $\mathbf{G}(z)$;
- a performance index J , function of $\mathbf{G}(z)$ inputs, outputs and states variables;
- constraints Λ on the $\mathbf{G}(z)$ signals,

the minimization of J yields to the optimal model-based controller. We will show in Chapter 6 the state-of-the-art AO model-based controllers.

The connection of WFS, DM and the controller form the AO system. The input-output delay of the AO system leads to the AO delay error σ_{delay}^2 , which takes into account the WFS exposure, the computational and the DM actuation times.

In general, control system architectures can be implemented on both dedicated platforms such as FPGA (Field Programmable Gate Array) or general purpose platforms such as CPU (Central Processing Unit) and GPU (Graphics Processing Unit):

- *FPGA-based architectures* [13 67 66 19 113] are dedicated solutions that can work stand-alone or coupled with Digital Signal Processing (DSP) components or CPUs. They guarantee massive parallel throughput, low latency and deterministic behavior. Such systems yield excellent performance but are expensive, not flexible and require skilled programmers. The scalable generic platform for AO described in [110], for example, uses an FPGA to control an AO system. This solution well suits complex telescope systems with tight performance requirements. However, simple parameter modification requires careful reprogramming to maximize FPGA processing power.
- *GPU* [107] are commonly found in medium/high-end computers and are less expensive, easier to program and also deploy than FPGA, while still being throughput-oriented thanks to their highly parallelized internal architecture. The main drawback is the added latency overhead of transferring data from CPU RAM to GPU.
- *CPU-based architectures* [8 20] are developed for general-purpose computers. In this case, the benefits are the use of high-level programming languages, modest hardware requirements, small physical footprint compared to FPGA, direct access to memory and peripheral registers as opposed to GPU, and easy maintenance. However, the computing units in CPUs, although more complex and powerful, are not designed to process a large number of simple operations in parallel, leading to additional latency. The Durham Adaptive optics Real-time Controller (darc) [9] is an example of a modular, distributed CPU-based AO architecture that is optimized for the Extremely Large Telescope (ELT) systems size. Despite its fine-tuning, latency time is in the order of hundreds of μs , higher than FPGA architectures designed for the same applications.

2.4.5 Discrete SCAO Block Diagram

From the control point of view, it makes sense to model both WFS and DM as a single AO plant with input $\mathbf{u}(k)$ and output $\mathbf{y}_c(k)$:

$$\mathbf{G}_{ao}(z) = \left[\begin{array}{c|c} A_{ao} & B_{ao} \\ \hline C_{ao} & 0 \end{array} \right] : \mathbf{u}(k) \mapsto \mathbf{y}_c(k). \quad (2.50)$$

The dynamic model of $\mathbf{G}_{ao}(z)$ can be identified using MIMO system identification methods [21, 78]. However, if the dynamic transient of the DM actuators is smaller than the sampling time T_s , it is sufficient to approximate \mathbf{G}_{ao} with two delays

$$\mathbf{y}_c(k) = \Pi_{\mathbf{u}2\mathbf{y}} \mathbf{u}(k-2), \quad (2.51)$$

accounting for the sum of the latencies between the communication of the command $\mathbf{u}(k)$ to the actuator electronics, the physically actuated desired shape and the WFS sample time.

The matrix

$$\Pi_{\mathbf{u}2\mathbf{y}} = C_{ao}(I - A_{ao})^{-1}B_{ao} \quad (2.52)$$

is the static gain mapping the DM inputs into the WFS measurements, commonly called Interaction Matrix (IM).

Data-driven approaches to modeling the atmospheric turbulence $\mathbf{G}_a(z)$ can also be found in e.g. [42, 41, 27]. Identification of the atmospheric disturbance in modal space may also facilitate online (recursive) update of these models in order to account for time-varying statistics of the disturbance, as suggested in [27].

Connecting the models from (2.28), (2.50) and (2.46) yields to:

$$\begin{aligned} \mathbf{G}_a(z) : \quad \mathbf{y}_a(k) &= \left[\begin{array}{c|c} A_a & B_a \\ \hline C_a & \mathbf{0} \end{array} \right] \mathbf{v}(k); \\ \mathbf{G}_{ao}(z) : \quad \mathbf{y}_c(k) &= \left[\begin{array}{c|c} A_{ao} & B_{ao} \\ \hline C_{ao} & \mathbf{0} \end{array} \right] \mathbf{u}(k); \\ \mathbf{C}(z) : \quad \mathbf{u}(k) &= \left[\begin{array}{c|c} A_c & B_c \\ \hline C_c & D_c \end{array} \right] \mathbf{y}(k); \\ \mathbf{y}(k) &= \mathbf{y}_a(k) - \mathbf{y}_c(k) + \boldsymbol{\eta}(k), \end{aligned}$$

the whole discrete-time block-diagram of a SCAO system (shown in Figure 2.10).

2.5 Conclusion

In ground-based astronomy, the spatial resolution of the optical image of a distant object is degraded by the presence of atmospheric turbulence. Therefore, the wavefront phase when the object light enters the telescope is no longer flat. The wavefront phase aberrations $\phi_a(t; \mathbf{r})$ can be compensated for by a SCAO system made of a WFS, DM and a controller. The WFS samples the incoming residual

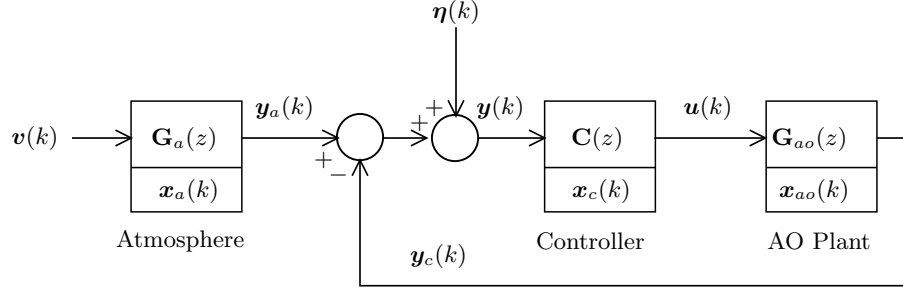


Fig. 2.10: Discrete SCAO block diagram. $\mathbf{G}_a(z)$, $\mathbf{G}_{ao}(z)$ and $\mathbf{C}(z)$ are the atmospheric, AO and controller state-space models, respectively, with their state notation; $\mathbf{v}(k)$ is the noise that drives the atmospheric turbulence model generating $\mathbf{y}_a(k)$, while $\boldsymbol{\eta}(k)$ is the measurement noise. The DM output is $\mathbf{y}_c(k)$, where $\mathbf{u}(k)$ is the command computed by the controller.

wavefront phase $\phi(t; \mathbf{r})$ into the measurement vector $\mathbf{y}(k)$. Then, the controller computes the control law, producing the DM commands $\mathbf{u}(k)$ from the WFS measurements. Finally, the DM, driven by the fed commands, shapes its surface to produce the conjugated wavefront phase $\phi_c(t; \mathbf{r})$ to cancel out the atmospheric turbulence aberrations.

The wavefront phase is described by the Zernike modal terms, which are suitable to most AO systems. The use of a modal basis brings advantages, such as efficient atmospheric turbulence modeling and control strategy design. Moreover, the Strehl ratio can be approximated from Zernike coefficient since the Zernike basis is formulated according to optical theory. The AO system (i.e. the connection of the AO components) is modeled as a state-space LTI MIMO system. Therefore, the controller can be designed using modern control theory results such as the optimal control framework.

For an AO system to perform well under given atmospheric turbulence characteristics, the AO system error budget σ_{AO}^2 must be investigated:


- The modal basis must be designed to minimize the uncorrected modes error $\sigma_{>N_w}^2$;
- By increasing the WFS exposure time, the measurement noise error σ_{WFS}^2 is lowered at the expense of higher delay error σ_{delay}^2 and loop frequency bandwidth error $\sigma_{T_s}^2$;
- The spatio-temporal discretization of the phase through the WFS may not respect the Nyquist-Shannon sampling theorem, leading to the aliasing error σ_{alias}^2 ;
- The spatial DM fitting error σ_{fit}^2 is determined not only by the DM capabilities but also by the WFS spatial resolution.

The overall AO budget error is then defined as the sum of all the aforementioned sources of error [54, 28]:

$$\sigma_{AO}^2 = \sigma_{>N_w}^2 + \sigma_{alias}^2 + \sigma_{WFS}^2 + \sigma_{T_s}^2 + \sigma_{fit}^2 + \sigma_{delay}^2. \quad (2.53)$$

PhotonLoop AO Controller

Summary. The PhotonLoop software is a multi-platform, CPU-based flexible framework that implements the regulator for controlling AO systems. It also provides the tools to tune the AO system, from wavefront measurement settings to control parameters. A logging feature allows in-depth offline data analysis, while scripting enables execution of batch experiments.

This chapter reviews PhotonLoop main features, with an explanation of the static calibration routine for the AO plant. The experimental results show that the proposed solution is able to correct aberrations of low to medium size SCAO systems using a consumer-grade notebook. 

¹ This chapter is based on the paper: J. Mocci, M. Quintavalla, C. Trestino, S. Bonora, and R. Muradore. “A Multiplatform CPU-Based Architecture for Cost-Effective Adaptive Optics Systems”. In: *IEEE Transactions on Industrial Informatics* 14.10 (Oct. 2018), pp. 4431–4439.

3.1 Introduction

In a SCAO system, aberrations are corrected by measuring the wavefront phase with a WFS and shaping accordingly the DM, as seen in Chapter 2. These two elements are connected through a real-time *control system* in order to close the feedback loop. The control loop cutoff frequency must match the cutoff bandwidth of the temporal dynamics of the atmospheric turbulence. Hence, the controller implementation is required to operate in real-time and have low latency. The control system can be implemented into different architectures, the choice of which depends on the AO system requirements. For small to medium AO setups, e.g., microscopy, ophthalmology and portable telescopes, the spatio-temporal constraints are not so strict to let the controller implementation be feasible on CPU. FPGA and GPU architectures are more commonly found in more challenging AO systems. Since AO systems are usually employed to improve the performance of scientific instruments, the controller should provide robust behavior and easy-to-use interface for the operators.

For the sake of fast development time we initially developed a CPU-based software called *PhotonLoop*. This tool is designed to achieve several goals:

- to deploy a stand-alone multi-platform software for controlling AO systems running on a consumer-grade notebook with USB interfaces;
- to guarantee the AO performance requirements of small to medium size systems (for example lab research and portable telescopes);
- to facilitate key operations such as deployment and tuning using a clear and friendly interface.

The proposed architecture is a software compiled for Windows and Unix operating systems. It consists of two layers of components, the *control layer* and the *monitor layer* as shown in Figure 3.1.

Components from both layers are implemented as threaded C++ objects and are connected to each other exclusively with the *Qt Company* framework's communication primitives [112], which allows for a decentralized, peer to peer design. The *Eigen* library [39] is the matrix calculation backbone of the software, offering a collection of fast C++ linear algebra routines derived from BLAS and LAPACK Fortran packages.

Small to medium AO systems are characterized by limited resolution camera and DM spatial resolution, hence resulting in vectors and matrices of negligible size which, in contrast, pose a challenge in larger systems. For this reason the data is internally stored as 64-bit doubles, favoring precision over space, except the camera spot image which is encoded as a matrix of integers.

3.2 Control Layer

The control layer is modeled according to the SCAO block diagram presented in Chapter 2. The Camera and DM Drivers, WFS and Controller components run in

² A more efficient GPU-based implementation is presented in Chapter 5.

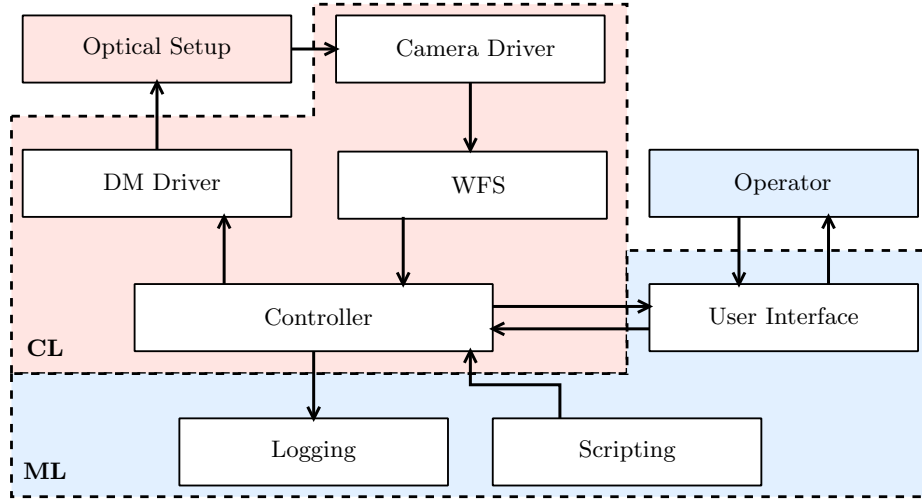


Fig. 3.1: PhotonLoop software architecture overview, with the Control Layer (CL)(red, upper), the Monitor Layer (ML)(blue, lower) and the physical environment (Optical Setup and Operator).

separate threads to optimize data processing. The main features of each subsystem are the following:

Camera Driver: Spots are captured from the WFS Camera, which transfers the spot image to the computer. The camera uses its internal clock source to determine the image capture framerate but can also be triggered by external inputs, for example when dealing with pulsed laser beams or other intermittent sources;

WFS: When the spot image transfer from the Camera Driver is completed, PhotonLoop immediately extracts the slopes. The camera image is partitioned in L sub-images related to the Shack-Hartmann lenslet array, each one corresponding to the l -th lenslet of the partitioning grid (see Chapter 5 for a more in-depth explanation on the working principle of this kind of wavefront sensor). Centroids positions c_l are calculated from the intensity map $I(x, y)$ using the Thresholded Weighted Center of Gravity method (TWCoG) [121], which improves the sensor's signal-to-noise ratio by thresholding and weighting the intensity to its cubic power:

$$I(x, y) = \begin{cases} I(x, y)^3, & \text{if } I(x, y) > \mathcal{T}_{\mathcal{P}}; \\ 0, & \text{else} \end{cases} \quad (3.1)$$

$$c_l = \begin{cases} c_l, & \text{if } \sum_{x, y \in l} I(x, y) > \mathcal{T}_{\mathcal{L}}, \\ \text{invalid}, & \text{else} \end{cases} \quad (3.2)$$

with $\mathcal{T}_{\mathcal{P}}$ the pixel threshold and $\mathcal{T}_{\mathcal{L}}$ the lenslet threshold. The centroids are stored by stacking the x and y components in a single vector. If the intensity in the cell is lower than $\mathcal{T}_{\mathcal{L}}$, or if the cell is manually excluded by the operator,

the corresponding centroid is removed from the vector to avoid unnecessary zeros in the subsequent matrix computation. The absolute reference centroids \bar{c}_l of any lenslet l are taken as the center point that satisfies the $\mathcal{T}_{\mathcal{L}}$ test. Alternatively, the reference centroids can be defined as the current centroids $\bar{c}_l = c_l$. The distance from c_l to \bar{c}_l yields the slopes s_l , ready to be sent to the Controller component.

After the definition of the reference centroids, the number of valid lenslets $\bar{L} \leq L$ is set, the Zernike projection matrix Π_{y2w} is built, and all the data structures are resized accordingly for efficient matrix multiplication.

The L sub-images are evenly mapped on the available CPU cores to speed-up the computation by parallelizing the TWCoG algorithm.

Remark. Centroids c_l can be extracted as soon as the pixels within the pixel area \mathbb{A}_l get transferred from the camera, without waiting for the whole image readout. This allows to reduce latency considerably [9]. At the best of our knowledge, no USB camera offers such interleaved image transfer yet, so this method is not implemented in PhotonLoop;

Controller: The controller component is in charge of calibrating the DM, decomposing the slopes in Zernike terms and computing the correction commands \mathbf{u} to be sent to the DM driver. The DM calibration parameters needed to compute the interaction matrix Π_{u2y} are:

- the pattern type (individual or Hadamard pokes);
- the actuator’s settling time;
- the poke amplitude;
- and the number of averaging frames.

It is also possible to skip the calibration phase by loading the interaction matrix computed in a previous session. In this case, the interaction matrix is expressed in Zernike terms instead of slopes.

Slopes can be edited to generate specific Zernike modes or to correct lens specific Zernike imperfection on the WFS optical path. It is possible to add offset $\bar{\mathbf{w}}$ to the measured Zernike coefficient vector \mathbf{w} :

$$\bar{\mathbf{y}} = \Pi_{w2y}(\mathbf{w} + \bar{\mathbf{w}}) \quad (3.3)$$

where $\bar{\mathbf{y}}$ are the edited slopes.

The closed-loop feedback controller produces the DM commands which minimize the slopes measurements by implementing an array of PI controllers or, in alternative, modal state-space control strategies.

3.3 Monitor Layer

The monitor layer is the interface towards the operator and data files. It is composed of the Graphical User Interface (GUI) and the scripting and logging components, and runs at lower priority in a CPU core not used by control layer components.

Operator Workflow: PhotonLoop can be adapted to specific AO systems by configuring drivers, graphical and off-line data parameters. The WFS slopes

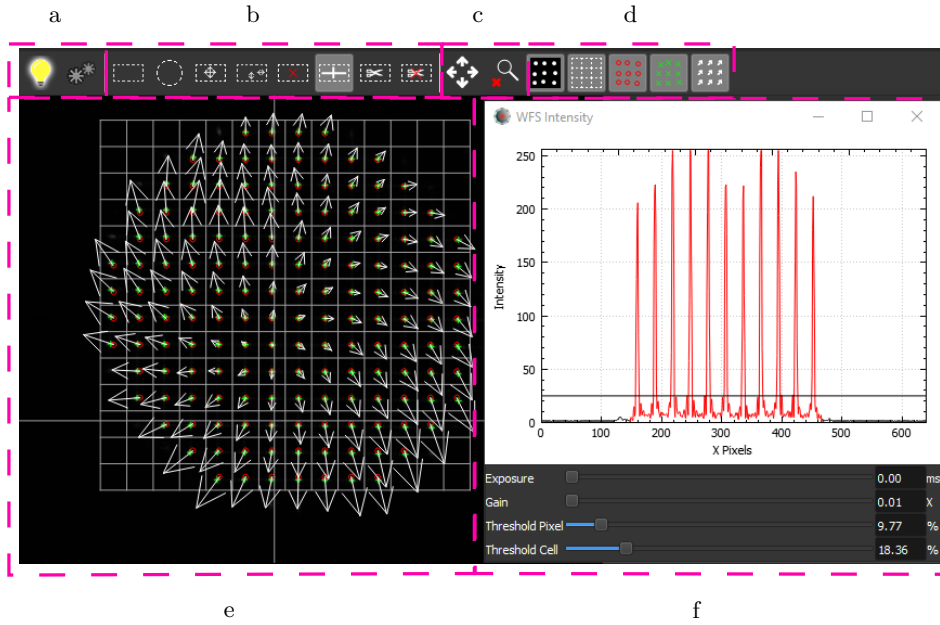


Fig. 3.2: PhotonLoop centroid extraction GUI. (a) WFS intensity and calibration toolbox; (b) centroiding grid manipulation; (c) fullscreen tools; (d) viewfinder items selection; (e) viewfinder; (f) WFS intensity toolbox, which shows the x pixels intensity profile centered on the detected centroid (to check sensor saturation) and the parameters for the Camera and the TWCoG algorithm.

extraction is configured in the Sensor page, as shown in Figure 3.2. The operator can crop the WFS camera image around the desired centroid aperture, draw and manipulate the extraction grid, and fine tune WFS and TWCoG parameters in the Intensity window. This window highlights the detected centroid intensity profile and the thresholds.

Figure 3.3 shows the Controller window with the global PhotonLoop GUI. The operator can choose the centroid reference, which can be absolute (measurements taken with respect to a flat wavefront) or relative to the current slope measurements, and automatically align the extraction grid to match the lenslet array.

The singular values coming from the SVD of the interaction matrix Π_{u2y} (i.e. system modes) are shown as a bar plot. Actuator responses and system modes can be promptly inspected as wavefront images. The number of singular values can be truncated by selecting the bar from the plot (more details are given in Section 3.4).

Once WFS and DM are calibrated, the controller performance is monitored by the Statistics display and a time-plot of the desired performance index, e.g. the Strehl ratio (estimated as in Eq. (2.19)), RMS or Peak-to-Valley value of the wavefront.

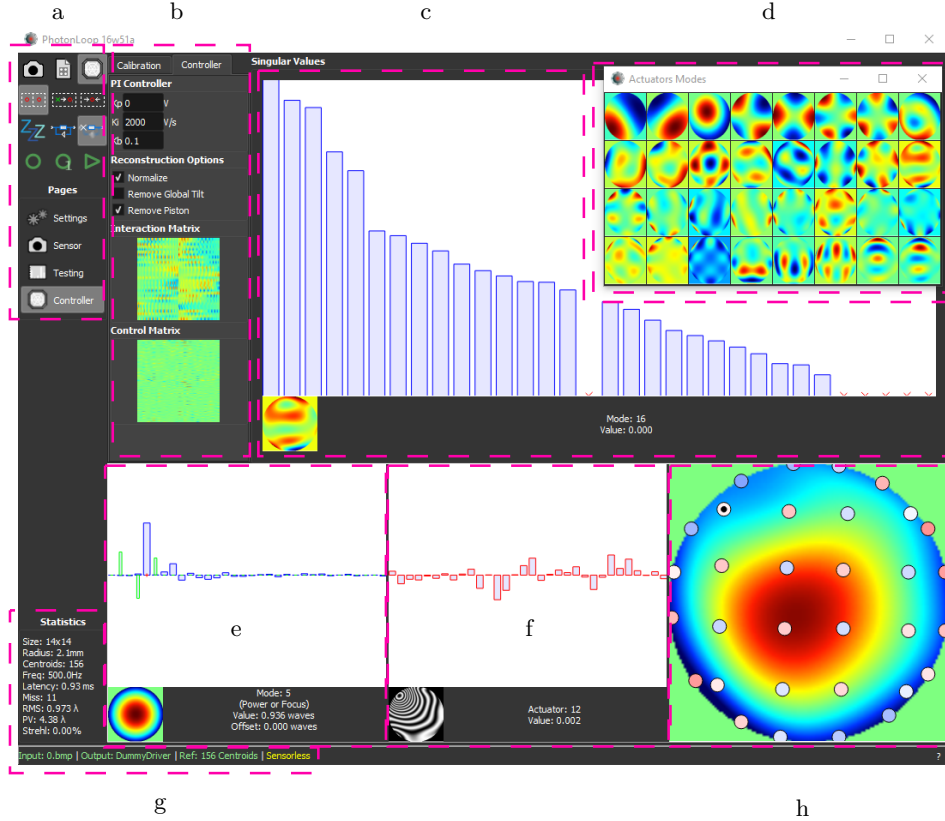


Fig. 3.3: PhotonLoop controller GUI. (a) PhotonLoop command center (open drivers, reference, controller modes, logging and scripting) and page selection; (b) controller window page, with PI parameters, interaction and control matrix visual representation; (c) singular values of the SVD; (d) SVD system modes as wavefronts; (e) Zernike coefficients (Blue: real, green: offsets); (f) DM commands; (g) Statistics (Strehl, centroids, performance, etc) and status bar; (h) reconstructed wavefront with superimposed actuator topology.

Input and Output Tools: PhotonLoop provides logging capabilities to save real-time data as time series. Off-line computations allow to reconstruct the controller state at every sample time from the recorded data. Within the scripting tool, it is possible to command the PhotonLoop components via JavaScript files, loaded at run-time, or via TCP-IP network protocol.

The integrated signal generator can load and playback time series of Zernike coefficients as a virtual aberration source. The time series interacts with the controller as shown in Figure 3.4, with the controller correcting for the simulated turbulence.

Scripting and signal generation are aspects of paramount importance when testing experimental AO hardware components or control algorithms, as they automate long repetitive experimental sessions and optimize logging.

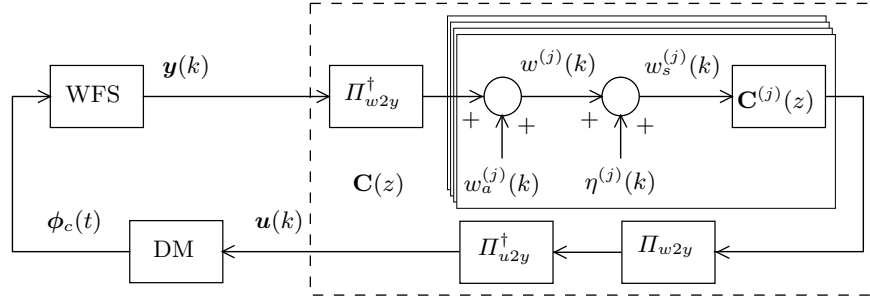


Fig. 3.4: The simulation test-bench. For each j -th mode, aberrations and noise time series are injected just before the j -th controller $\mathbf{C}^{(j)}(z)$. The real test-bench lies outside the dashed area. Since the incoming light has a perfectly planar wavefront, the controller commands the DM to be shaped to compensate for the simulated perturbations.

3.4 AO System Static Calibration

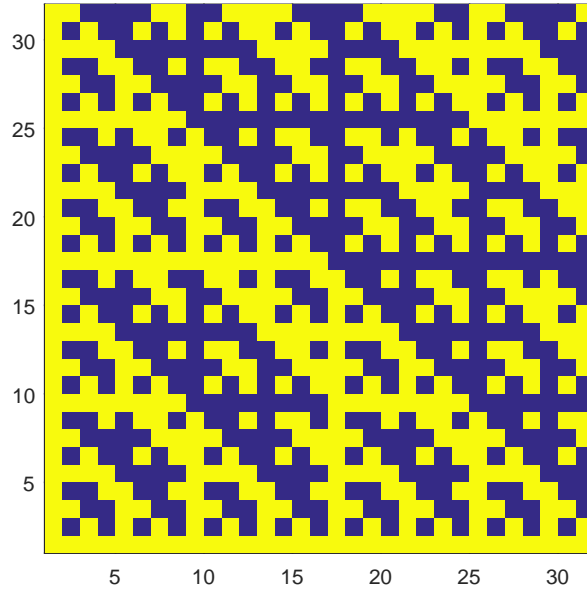
The AO calibration is a key phase since it provides the relation between the command $\mathbf{u}(k)$ and the DM output $\mathbf{y}_c(k)$. If the DM is assumed to have infinite dynamics with some delay units (usually a good approximation), then the calibration returns the DM static gain Π_{u2y} . More precisely, the calibration identifies the relations between the WFS measurements and the DM command $\mathbf{u}(z)$ of the AO setup [49].

The calibration routine is an automatized procedure to calculate the IM Π_{u2y} and, by pseudo-inverting it, the CM Π_{y2u} . The routine needs to be executed at the beginning of every experimental session or, at least, at every change in the input or output configuration. It is structured as follows:

1. Calculate the *command patterns* matrix U_{ao} . Each column of U_{ao} is a known command vector to be sent to the DM;
2. Measure the *actuator response* matrix Y_{ao} . Each column of Y_{ao} is the WFS measurement \mathbf{y}_s related to the DM command pattern matrix U_{ao} ;
3. Infer the static gain Π_{u2y} from the command matrix U_{ao} and the actuator matrix Y_{ao} .

The command pattern Y_{ao} must be sufficiently rich to stimulate all the actuators of the DM. Since the calibration is restricted to the static gain identification, it is sufficient for it to be orthogonal. There are two standard techniques used to build the command pattern matrix:

- The *Spatial pattern* matrix is the identity matrix with size equal to the number of DM actuators. Each pattern has the effect of poking just one actuator at a time;
- The *Hadamard patterns* matrix is made from a class of orthogonal matrices composed of only -1 and 1 values, (see Figure 3.5). Its column count is equal or greater than the DM actuator count. Each pattern pokes all actuators at once: half of them in one direction, the other half in the opposite direction.

Fig. 3.5: 32×32 Hadamard matrix

Hadamard patterns are preferred to Spatial patterns as they improve the signal-to-noise ratio and better capture cross-correlations between actuators. However, Hadamard matrices are not defined for all square sizes. When the Hadamard matrix does not match the number of DM actuators, then the nearest bigger Hadamard matrix is used instead. The excess rows are removed so that the columns have the same dimension as the DM command vector. The remaining columns within the Hadamard matrix are orthogonal.

The calibration routine uses the voltages in the columns of the pattern matrix U_{ao} to iteratively collect the actuator response matrix Y_{ao} . Since the AO model \mathbf{G}_{ao} is formed by the concatenation of the DM and WFS models, \mathbf{G}_c and \mathbf{G}_s respectively, there are several aspects to take into account. The actuator response to a poke pattern must be measured after the AO system settles due to the AO system delay. To improve the signal-to-noise ratio of the WFS measurements, the actuator responses are averaged over time.

Depending on the DM technology and the optical quality of the AO setup, the actuator responses to the DM zero command might not yield to a perfectly flat wavefront. Hence, the reference wavefront is the one obtained when the DM actuators are at their natural zero instead of the wavefront flat. Unmodeled nonlinearities such as piezo-electric creep and hysteresis are mitigated by *relaxing* the DM actuators between each poke iteration. The relax is a technique that lets the actuators to reset to their natural zero position, and is performed by sending a command time series made of a sinusoid whose amplitude decreases exponentially over time. A less effective but faster technique is to simply reset the reference

wavefront to the actuators zero at the end of each iteration. This prevents the actuator footprints to creep into the old reference wavefront.

The actuator responses might be non-linear, i.e. the absolute wavefront displacement achieved by poking in a direction is not equal when poking along the opposite direction. In that case, two measurements are taken for each actuator pattern. The first measurement is the response to the actual pattern, whereas the second one is the response to the opposite pattern. The recorded response is the average between the two measurements. The pattern matrix is scaled so that the DM and WFS do not saturate. The scaling factor is chosen depending on how the pattern matrix is defined. For example, the response to Hadamard patterns has larger values than Spatial patterns, given the same scaling factor.

Algorithm 1 describes the calibration steps, which results into the actuators response matrix Y_{ao} . The interaction matrix is then constructed from actuator patterns and response matrix as

$$\Pi_{u2y} = Y_{ao}U_{ao}^\dagger, \quad (3.4)$$

and its pseudoinverse

$$\Pi_{y2u} = \Pi_{u2y}^\dagger \quad (3.5)$$

is the control matrix.

Algorithm 1 Calibration routine

- 1: Build actuator pattern matrix U_{ao}
 - 2: Relax actuators
 - 3: Take wavefront reference
 - 4: **for all** patterns $\mathbf{u}_{ao} \in U_{ao}$ **do**
 - 5: Poke actuators with the \mathbf{u}_{ao} command
 - 6: Wait for settling time
 - 7: Collect the positive actuator response \mathbf{y}_{ao}^+ averaged over time
 - 8: Relax / Set to zero
 - 9: Take wavefront reference
 - 10: Poke actuators with the $-\mathbf{u}_{ao}$ command
 - 11: Wait for settling time
 - 12: Collect the negative actuator response \mathbf{y}_{ao}^- averaged over time
 - 13: Calculate the actuator response \mathbf{y}_{ao} as the average between \mathbf{y}_{ao}^+ and $-\mathbf{y}_{ao}^-$
 - 14: Concatenate the actuator response \mathbf{y}_{ao} to the actuator response matrix Y_{ao}
 - 15: Relax / Set to zero
 - 16: Take wavefront reference
 - 17: **end for**
 - 18: Relax actuators
-

The computation of the control matrix from the interaction matrix as in 3.5 is not straightforward because:

- Actuator cross-correlations and possible lack of information from some measurement elements make the Π_{u2y} not full-rank;

- The number of actuators and measurement elements is very likely not equal and thus Π_{u2y} is not square.

Hence, the Singular Values Decomposition (SVD) is used to calculate the pseudo-inverse in the least-squares sense (Moore-Penrose pseudo-inverse) [34]. This technique has the advantage to compute the rank of Π_{u2y} and hence select the number of *system modes* (SVD singular modes relative to the singular values) by removing modes that are either too noisy or require too much dynamics to be obtained, at the expense of the spatial resolution of the correction. Figure 3.6 shows the system mode wavefronts of a DM with 32 actuators, reconstructed from the left singular columns of Π_{u2y} .

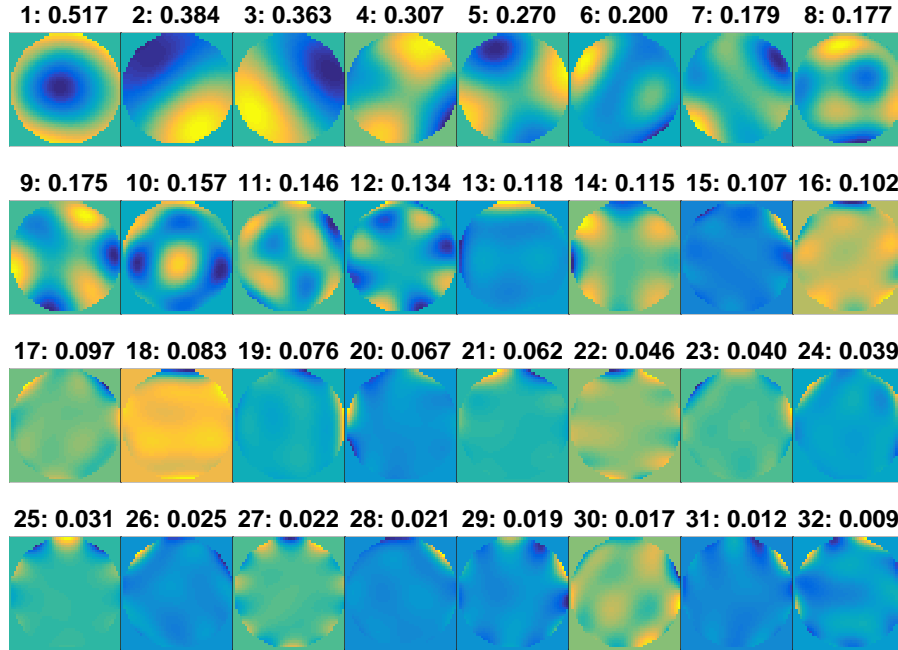


Fig. 3.6: System modes of a 32 elements DM, with each mode number (on the left) and its respective singular value (on the right).

3.5 Experimental verification

The control software is tested on the AO testbench shown in Figure 3.7. The light coming from a single emitter laser diode (Thorlabs CPS670, $\lambda = 670nm$, power $4.5mW$) is collimated by the lens l_1 to create a point-like source at infinity. The light beam is then directed on a $22mm$ aperture bimorph piezoelectric DM with 32 actuators which acts as the optical system pupil. The DM is developed by the CNR-IFN of Padova [12].

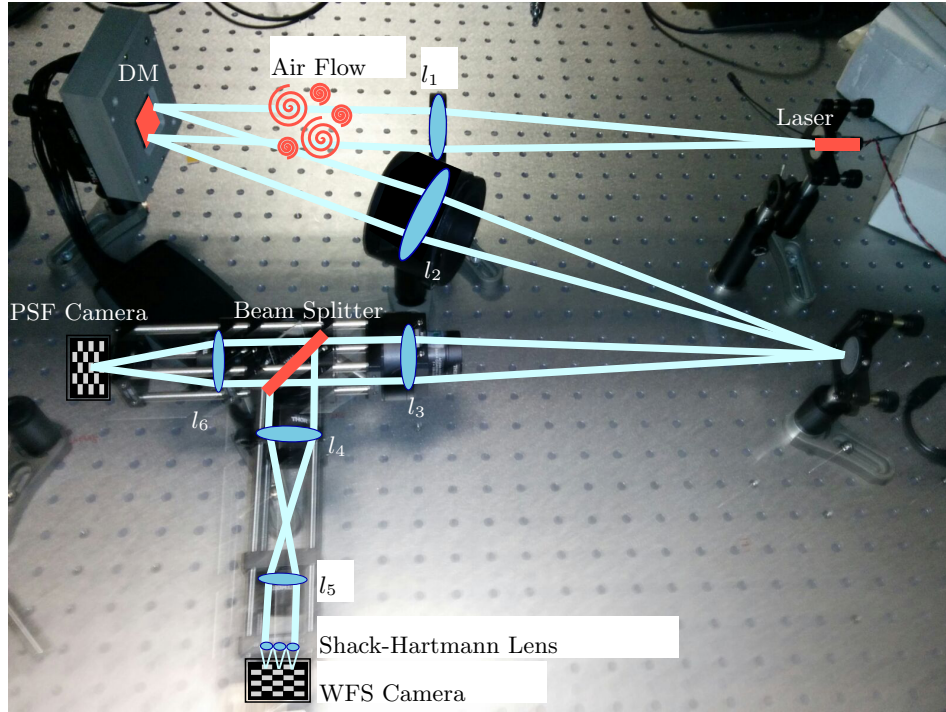


Fig. 3.7: The AO test-bench, consisting of laser emitter, DM, WFS and PSF (Scientific) Camera. l_1, \dots, l_6 indicate the lenses.

The DM is driven by an electronic controller interfaced to the computer by USB 2.0. The pupil is conjugated to the WFS by means of two telescopes (lenses l_2/l_3 and l_4/l_5 , respectively) that also reduce the beam size to make it smaller than the WFS aperture. Part of the beam is focused by the lens l_6 on a scientific camera where the light source is reimaged. The WFS camera is a USB 3.0 global-shutter CMOS type monochromatic sensor with square pixel size $\mathcal{P}_w = \mathcal{P}_h = 5.86\mu\text{m}$ manufactured by the *iDS* company. The lenslet array has lenslet pitch $\mathcal{L} = 150\mu\text{m}$ and focal $f = 6.4\text{mm}$. Each lenslet is mapped into a pixel area \mathbb{A}_l of 25×25 pixels. Spot images are transferred as 8-bit integer matrices representing the quantized pixel intensity I . Due to the camera transmission protocol, higher precisions (e.g. 10, 12, 16 bits) would be encoded in 16-bit integers halving the camera framerate. Since the software operative frequency depends on the WFS camera frame rate, which in turn depends on the considered Region of Interest (RoI), the optical beam directed to the WFS was adjusted to reach an acceptable compromise between speed and accuracy (number of lit lenses on the lenslet array). In this case, the beam diameter is set to 1.9mm . The wavefront is sampled by 121 subapertures, excluding the faint ones and those outside the unit circle. With such configuration the control loop frequency is 500Hz . WFS and DM are driven by the software controller, which runs on a 64bit Windows consumer-grade laptop with a dual core (four virtual cores) Intel i5-2410 CPU clocked at 2.30GHz . The computation time from image retrieval to actuator commands, excluding transfer times, is $40\mu\text{s}$.

3.5.1 Setup Calibration

Camera exposition time is tuned so that the intensity I doesn't saturate the sensor. Pixel threshold \mathcal{T}_P is set to 10% of the maximum achievable intensity to suppress background noise. No other image calibration is done. The mirror is calibrated with the Hadamard pattern. Voltages are scaled to 30% of the maximum value to avoid sensor saturation. Once the electrical voltage is applied, the mirror was let to stabilize for 0.2s after which 20 consecutive measurements were averaged to filter out sensor noise. From a total of 32 system modes, the last 9 singular values (i.e., the smaller ones) are set to zero to avoid actuator saturation.

3.5.2 Aberrations Correction

Closed-loop experiments are performed by blowing a hot air stream through the optical path before the DM (see Figure 3.7) to create a turbulence effect representative of the air turbulence that may affect AO systems. The optimal integral and proportional gains \mathcal{K}_P and \mathcal{K}_I are obtained by minimizing the RMS residual error.

The Zernike coefficients are calculated from the WFS slopes measurements as shown in Section 2.2. The histogram of the Zernike coefficients in Figure 3.8 shows the effect of the AO correction over the aberrated wavefront. The approximated wavefront RMS is calculated from the discrete-time series of Zernike coefficients $\mathbf{w}(k)$:

$$\text{RMS} = \sqrt{\frac{1}{N_k} \sum_{k=0}^{N_k} \mathbf{w}^2(k)}, \quad (3.6)$$

where N_k is the length of the window. The Strehl Ratio is approximated from the Zernike coefficients using the equations in Section 2.3.

The amplitude of every Zernike coefficient is strongly reduced, as the overall sum of the coefficients. The difference in terms of Strehl ratio between uncorrected and corrected wavefront is reported in Figure 3.9 an improvement from 5% to 60% can be seen. The same figure shows how the correction affects the PSFs when AO is switched on and off. This data clearly demonstrates that the AO system can effectively correct for dynamical aberrations, yielding a very low residual wavefront error.

It is worth computing the *Rejection Transfer Function* (RTF) (also called sensitivity function $\mathbf{S}(k)$) that describes how the incoming aberrated wavefront is attenuated by the AO system [31]. Using the discrete time formulation we have

$$\mathbf{S}(z) = \frac{\text{Closed-Loop Frequency Response}}{\text{Open-Loop Frequency Response}} \quad (3.7)$$

$$= \frac{1}{1 + \mathbf{G}_{ao}(z)\mathbf{C}(z)} \quad (3.8)$$

where z is the Zeta transform variable and $\mathbf{G}_{ao}(z)$, $\mathbf{C}(z)$ are the transfer functions for the AO plant and the controller, respectively. The magnitude of $\mathbf{S}(z)$ is estimated as the ratio of the Welch PSD of the residual and of the input for each Zernike mode j

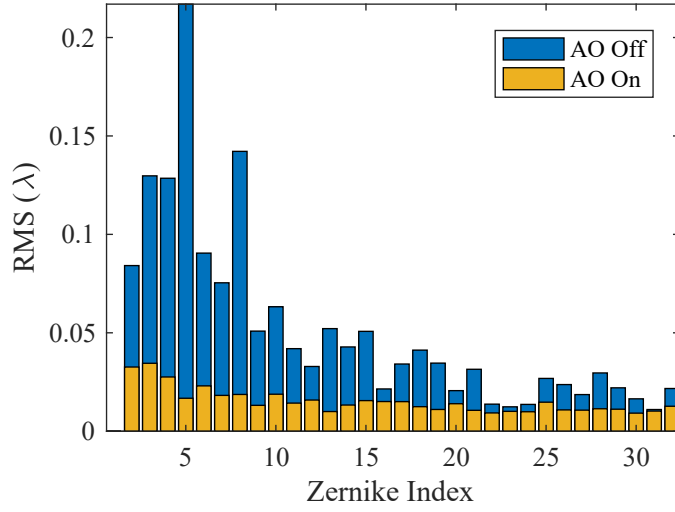


Fig. 3.8: Histogram of Zernike coefficients temporal RMS, measured over a time window of 3 seconds with AO correction on and off. RMS is given in waves of wavelength $\lambda = 670nm$.

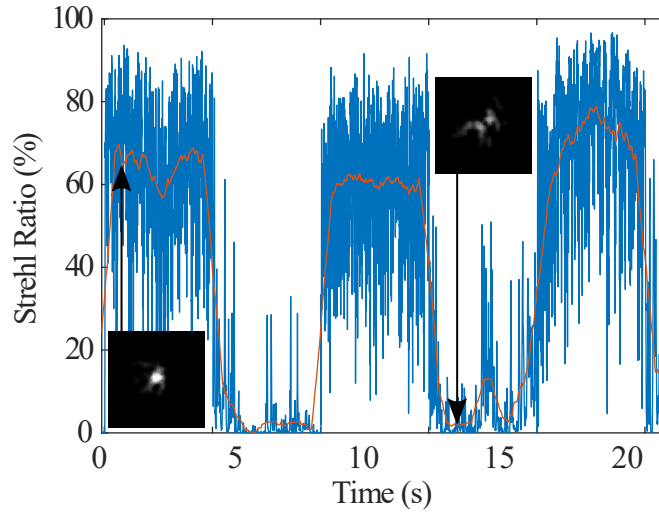


Fig. 3.9: Strehl ratio measurement (calculated according to Eq. 2.19) of the aberrated optical system, toggling AO correction between on and off every 5 seconds. Examples of measured PSF of both on and off cases are shown. In red: averaged Strehl ratio.

$$\mathbf{S}^{(j)}(z) \approx \frac{\text{PSD}^{(j)}(\text{residual})}{\text{PSD}^{(j)}(\text{input})}. \tag{3.9}$$

In Figure 3.10 the experimental rejection transfer functions are shown for the Tip, Tilt and Defocus Zernike modes. The AO system has a control frequency

bandwidth around 25Hz . Figure 3.11 shows the Power Spectral Density (PSD) of

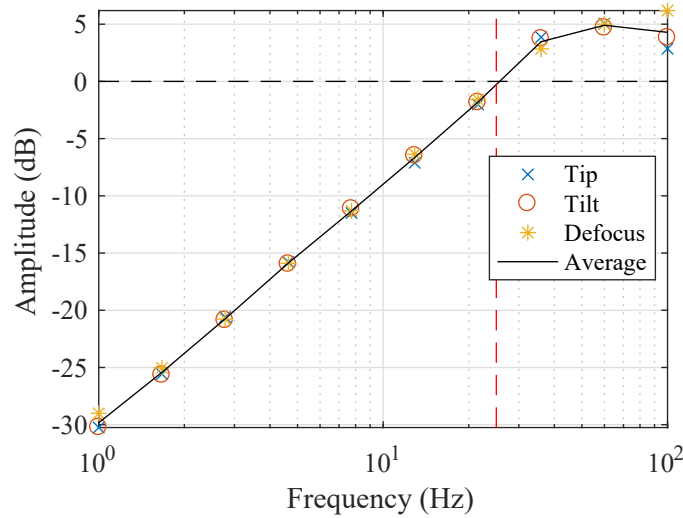


Fig. 3.10: Rejection transfer function of Tilt, Tip and Defocus Zernike modes. Dashed vertical line: cutoff frequency at 25Hz . Camera frequency is fixed at 500Hz .

Tip, Tilt and Defocus Zernike modes with the AO system switched on and off. The PSDs are strongly attenuated in the low frequency range as expected from the shape of the rejection transfer functions.

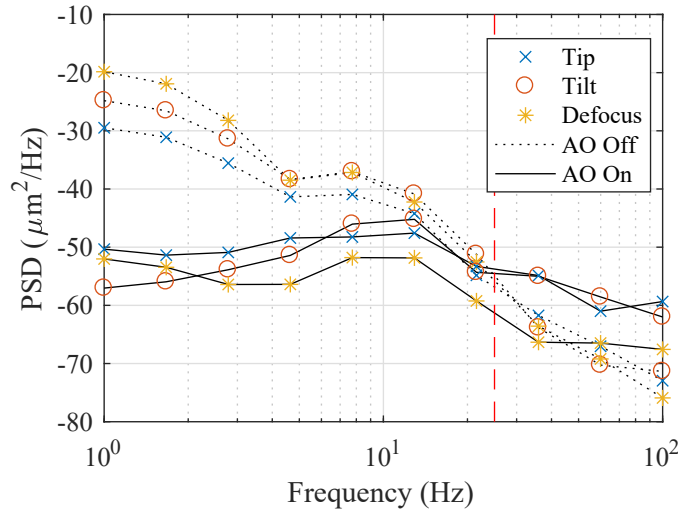


Fig. 3.11: Frequency content of Tilt, Tip and Defocus Zernike modes, measured in a time window of 3 seconds with AO correction on (solid lines) and off (dashed lines). Dashed vertical line: cutoff frequency at 25Hz .

System performance agrees with the heuristics given in [58]: the closed-loop correction bandwidth is $\approx 1/20$ of the control frequency $F_s = 1/T_s = 500Hz$, i.e. $25Hz$. The correction could be further improved by minimizing the communication latency, for example by using the faster USB 3.0 interface for both DM and WFS.

3.6 Conclusion

In this chapter we presented the Adaptive Optics software PhotonLoop that can be easily installed on a consumer grade computer without any dedicated hardware like image framegrabbers or GPUs. The strength of this approach is the favorable trade off between performance, easy implementation and tunability. This software has been easily integrated within research laboratory, medical or industrial instrumentation such as laser systems, ophthalmoscopes and small/medium size telescopes, as reported in Appendix A.

Atmospheric Turbulence Characterization over an Horizontal Path

Summary. When observing celestial objects from ground-based telescopes, the light beams go through the atmospheric turbulence layer. In this case, it is possible to describe the turbulence statistics using universally accepted look-up tables which values are related to the telescope location, inclination and time of the year. However, the image of an object through an horizontal turbulent path, e.g. a laser source placed far away from the telescope, has not received so much attention. This chapter presents a method to infer the atmospheric turbulence parameters from a Shack-Hartmann wavefront sensor measurements, hence without adding complexity to an Adaptive Optics setup. The horizontal turbulence parameters are estimated in a real-world scenario by observing a laser source with a telescope, positioned $1km$ apart over the sea. The measured time series are then injected back into an optical test-bench (as explained in Chapter [3](#)) and compensated for by the AO system. [1](#)

¹ This chapter is based on the conference paper: J. Mocci, M. Quintavalla, S. Firpi, L. Bancallari, S. Bonora, and R. Muradore. “Analysis of Horizontal Atmospheric Turbulence by using a Shack-Hartmann Wavefront Sensor”. In: AOIM XI, Murcia. 2018

4.1 Introduction

Atmospheric turbulence is a widely studied topic ([99], [123], [24]) in the context of astronomical applications. This is because AO systems must be designed accordingly to the observation site and the aberrations to be compensated. For AO over vertical optical paths, i.e. ground-based telescopes, *look-up tables* of a given site returns the atmospheric turbulence parameters as functions of the height h and time of the year [100]. Such look-up tables are either calculated by profiling the atmosphere over large periods of time, or obtained from mathematical models.

Unfortunately, there are no look-up tables available for horizontal optical paths (i.e. ground optical communication). Water bodies, industrial activities and other temporary heat exchange sources contribute to the horizontal atmospheric turbulence in unique ways. Hence, AO hardware must be chosen for the turbulence worst case scenario. As such, there is an interest into modeling the horizontal atmosphere, preferably using the same AO system equipment.

We will briefly report the basic facts about the atmospheric turbulence phenomenon and then we will focus on the horizontal turbulence (i.e. the characterization of the ground layer $h \simeq 0$) that is the target of this Chapter.

4.2 Atmospheric turbulence

The atmosphere is governed by fluid dynamics processes, where temperature variations result in wind changes (so-called eddies). The constant evolution of such processes is hence identified as turbulence. When a volume of air becomes hotter or colder, its density changes along with its refractive index [100].

Light emitted by a distant object (e.g., star or planet) goes through the atmospheric layer before reaching the telescope. Initially, the light wavefront is flat (i.e. plane). However, the non-homogeneous refractive index of the atmosphere alters the wavefront phase (aberration). Diverging beams leads to interferences in the light intensity (scintillation) that degrades the image resolution. Figure 4.1 shows the turbulence effect on the incoming wavefront.

The aberrated wavefront is described as an electromagnetic wave as

$$I(\mathbf{r}) = |A(\mathbf{r})| \exp[-i\phi(\mathbf{r})], \quad (4.1)$$

where the wavefront phase is defined as

$$\phi(\mathbf{r}) = \frac{2\pi}{\lambda} l(\mathbf{r}), \quad (4.2)$$

with λ being the wavelength and $l(\mathbf{r})$ the optical path. The atmospheric refractive index variations are considered achromatic and hence $l(\mathbf{r})$ does not depend on λ .

A reasonable assumption is to think of the atmospheric turbulence as free space lenses, translating and rotating over time. Each lens is generated by a turbulent eddies and aberrates the traveling wavefront by changing the local refractive index. Therefore, scintillation effects are regarded as second order phenomena, byproducts of the wavefront aberrations. The Kolmogorov model is based on the lens-like behavior of the atmospheric turbulence, and introduces further assumptions:

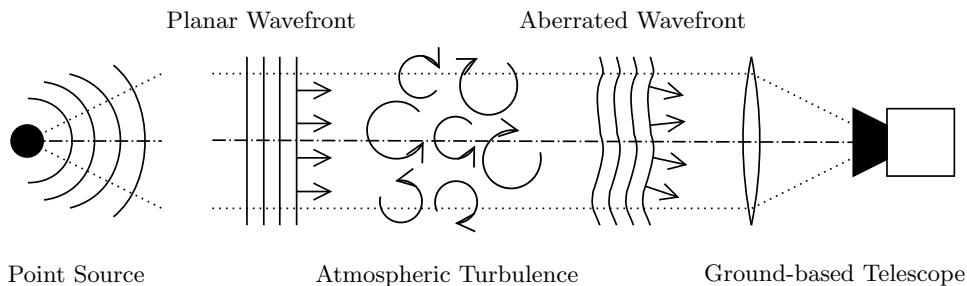


Fig. 4.1: Atmospheric turbulence diagram. Before entering the atmosphere, the point object wavefront is planar. The turbulent eddies act as moving lenses, which change the local refractive index and therefore aberrating the wavefront. Since the light rays are diverging, the image observed by the telescope suffers from intensity scintillations, degrading its resolution.

- *Local homogeneity*: the velocity difference between two points in space only depends on their distance;
- *Isotropic atmosphere*: the direction does not play a role in the velocity difference;
- *Incompressible atmosphere*: the volume is assumed to be constant.

Since the phase aberrations dominate over the intensity scintillation effects, the Kolmogorov model represents a good approximation of atmospheric turbulence.

4.3 Cn2 Index for Horizontal Turbulence

Since all the aberrations along the optical path sums up, in the SCAO scenario the analysis can be restricted to the wavefront phase instead of the entire optical path volume. It is therefore sufficient to measure and correct into the plane conjugated at the telescope primary mirror to compensate for atmospheric turbulence.

The atmospherically induced covariance between two phase points over the wavefront, \mathbf{r} and \mathbf{r}' , is described by the *structure function*

$$C_\phi(\mathbf{r}, \mathbf{r}') = \mathbf{E}[\phi(\mathbf{r}), \phi(\mathbf{r}')]. \quad (4.3)$$

Since the atmosphere is modeled with Kolmogorov, the structure function is only dependent on the distance between phase points:

$$C_\phi(\mathbf{r}, \mathbf{r}') = C_\phi(|\mathbf{r} - \mathbf{r}'|^2). \quad (4.4)$$

The refractive index structure constant $C_n^2(h)$ is the structure function depending only on the h vertical height [5]. $C_n^2(h)$ -based profiles for the turbulence strength are readily available for vertical or slanted optical paths. Many $C_n^2(h)$ profile models have been developed from measurements over the years, with vertical height being the only parameter. Vertical and slanted optical paths were the only modeled $C_n^2(h)$ profiles as the horizontal path case is independent of h .

Several practical atmospheric turbulence parameters can be inferred from the $C_n^2(h)$ profiles, along with the telescope to conjugated plane distance L [100]:

- The *Fried parameter*

$$r_0 = \left[0.422 \left(\frac{2\pi^2}{\lambda} \right) \sec(z) \int_0^L C_n^2(h) dh \right]^{-3/5} \quad (4.5)$$

is the atmospheric turbulence coherence length over which the wavefront aberrations are averaged to 1 rad^2 spatial variance, and is lower when the wavefront aberration spatial variance between two points is higher. Therefore, it is extremely useful when designing the AO setup since r_0 defines the maximum actuator pitch of the DM, as shown in Figure 4.2. Since ground-based telescopes are characterized by large apertures, their DM must be designed with more actuators than the one needed in smaller telescopes to achieve diffraction limited image resolution.

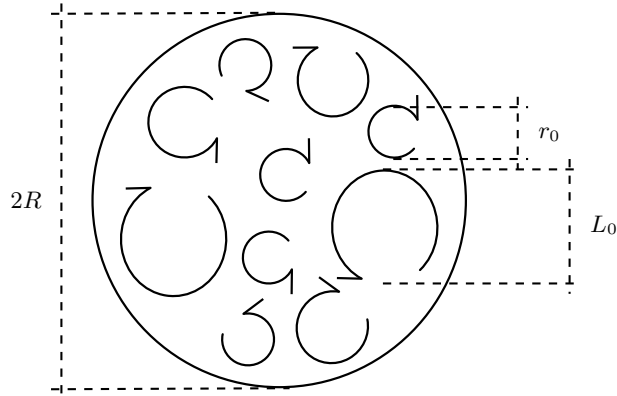


Fig. 4.2: Wavefront aberrations induced by the atmospheric turbulence. The Fried parameter r_0 is the inner turbulence scale, whereas L_0 is the outer turbulence scale. Such parameters, along with the telescope aperture radius R , determine how many DM actuators are needed to correct for the wavefront aberrations.

The size of turbulent eddies is not constant. They can break down into smaller eddies, eventually reaching the minimum size given by the Fried parameter. The maximum size of a turbulent eddy is modeled by the *outer turbulence scale* L_0 .

- The *anisoplanatic angle*

$$\theta_0 = \left[2.91 \left(\frac{2\pi^2}{\lambda} \right) \sec^{8/3}(z) \int_0^L C_n^2(h) h^{5/3} dh \right]^{-3/5} \quad (4.6)$$

is the maximum separation between two light sources after which the wavefront is considered different. For example, as the SCAO system is able to correct for the aberrations of a single reference object, the *field of view* in which the wavefront is corrected is given by the anisoplanatic angle, as shown in Figure 4.3.

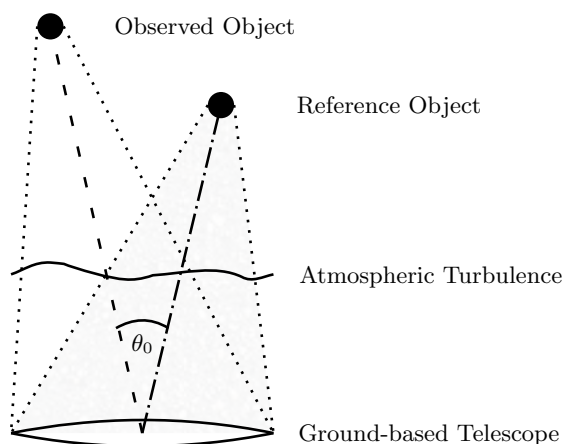


Fig. 4.3: The isoplanatic angle θ_0 determines how far the observed object can be with respect to the reference object for them to share the same wavefront .

If the Fried parameter is known and C_n^2 is assumed to be constant over the optical path (e.g. averaged), then the anisoplanatic angle can be calculated as

$$\theta_0 \simeq 0.6 \left(\frac{r_0}{L} \right). \quad (4.7)$$

- The *Greenwood frequency* (or temporal coherence)

$$\tau_0 = 2.31\lambda^{-6/5} \left[\sec(z) \int_0^L C_n^2(h) v_w^{5/3}(h) dh \right]^{3/5} \quad (4.8)$$

is the time scale in which the wavefront changes its shape due to the turbulent eddies. It is dependent on the wind velocity $v_w(h)$ which is often found in look-up tables next to the $C_n^2(h)$ parameter.

As with the isoplanatic angle, by knowing the Fried parameter and keeping both $C_n^2(h)$ and $v_w(h)$ constant over the optical path, the Greenwood frequency can be approximated as

$$\tau_0 \simeq 0.43 \left(\frac{v_w}{r_0} \right). \quad (4.9)$$

By exploiting the Rytov method, it is possible to derive C_n^2 from experimental scintillation data [6]. The scintillation index is defined by

$$\sigma_I^2(\mathbf{r}, L) = B_I(\mathbf{r}, L) = \frac{\langle I^2(\mathbf{r}, L) \rangle}{\langle I(\mathbf{r}, L) \rangle^2} - 1 \quad (4.10)$$

$$= \frac{\langle I^2(\mathbf{r}, L) \rangle - \langle I(\mathbf{r}, L) \rangle^2}{\langle I(\mathbf{r}, L) \rangle^2} \quad (4.11)$$

where $\langle \cdot \rangle$ is the ensemble average and $I(\mathbf{r}, L)$ is the sampled intensity as a function of the observation point \mathbf{r} and the propagation distance L . Since only the variance is needed, it is not necessary to calibrate the intensity measurements.

The link between scintillation index and C_n^2 is given by the Rytov variance

$$\sigma_1^2 = K C_n^2 \left(\frac{2\pi}{\lambda} \right)^{7/6} L^{11/6} \quad (4.12)$$

where K is related to the type of light source ($K = 1.23$ for plane waves, $K = 0.5$ for spherical waves) and L is the propagation length. In the case of weak fluctuations (when the scintillation index σ_I^2 is less than one), Rytov approximation allows to state the following equality:

$$\sigma_I^2 = \sigma_1^2, \quad \sigma_1^2 < 1. \quad (4.13)$$

Scintillation measurements to compute the parameter C_n^2 are obtained with dedicated instruments such as scintillometers (SCIDAR) or Shack-Hartmann wavefront sensors (SLODAR). Many variants have been proposed exploiting, for example, the correlation between different sources or autocorrelation between scintillation and slopes measurements [23].

In this work, we derived the scintillation measurements by images collected with a single Shack-Hartmann WaveFront Sensor (SHWFS). SHWFS spatially samples the wavefront by focusing the light beam with a lenslet array to the camera pixel plane. The image captured by the camera is a grid arrangement of light spots. their positions inside the grid cells give information on the local tilts of the wavefront.

Since (4.10) doesn't require the exact scale of intensity measurements, the scintillation index is calculated for each spot independently and outliers are removed by thresholding. The terms in (4.10) are then approximated by

$$\langle I(\mathbf{r}, L) \rangle \simeq \langle I \rangle \quad (4.14)$$

$$\langle I^2(\mathbf{r}, L) \rangle \simeq \langle I^2 \rangle \quad (4.15)$$

where $\langle I \rangle$ and $\langle I^2 \rangle$ are the ensemble averages on the number of lens within the SHWFS and on the samples for each time series [23].

Substituting the measurements of the scintillation index

$$\sigma_1^2 = \frac{\langle (I(x) - \langle I \rangle)^2 \rangle}{\langle I \rangle^2} \quad (4.16)$$

in (4.12), it is possible to compute the value of the C_n^2 that characterizes the horizontal turbulence

$$C_n^2 = \frac{\langle I^2 \rangle / \langle I \rangle^2 - 1}{0.5k^{7/6} L^{11/6}} \quad (4.17)$$

as shown by Andrews in [6].

In the literature there is another formulation for the C_n^2 proposed by Parry [83] assuming a log-normal distribution of the scintillation index over the atmosphere [111]

$$C_n^2 = \frac{\ln(\langle I^2 \rangle / \langle I \rangle^2)}{0.5k^{7/6} L^{11/6}}. \quad (4.18)$$

For the sake of completeness we have computed both values in the next Section.

4.4 Proposed Method and Setup

In the previous sections we showed that the required measurements to calculate C_n^2 for horizontal turbulence are the intensity time series, while for Zernike frequency content are the wavefront gradient's time series. The core idea of the proposed method is to use the Shack-Hartmann Wavefront Sensor (SHWFS) camera to capture both (relative) intensity and slope measurements from a known light source placed far away.

The experimental setup and a few frames captured from SHWFS and PSF cameras are shown in Figure 4.4. The parameters are listed in Table 4.1. The optical setup consists of:

- A Celestron C11 telescope with a lens assembly;
- A SHWFS;
- A point-like LED light source.

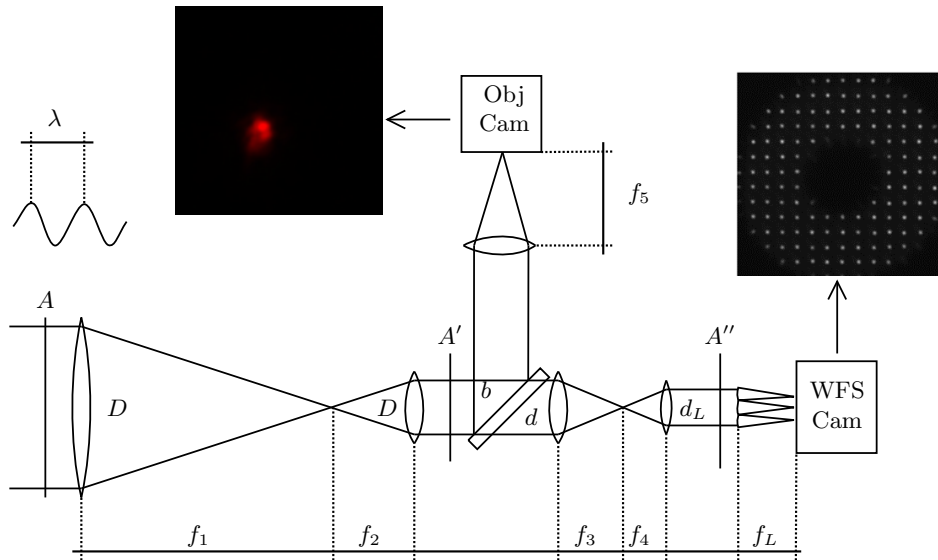


Fig. 4.4: Experimental setup. A , A' and A'' are the conjugated image planes; D , d and d_L are the lens diameters; b is the beam splitter; f_1, \dots, f_5 are the lens focal lengths; λ is the LED beam wavelength. The Obj Cam acquires the LED PSF, whereas the WFS Cam measures the wavefront.

The telescope (Figure 4.5) and light source were positioned 1km apart on the ground, with the optical path in-between being over the sea. The LED image plane (far field) was conjugated with both the object camera and SHWFS camera. The images captured by the camera are corrupted by the horizontal turbulence and so the Point Spread Function (PSF) cannot be seen any longer.

Datasets were taken on the 19th of July 2016 at Arsenale Militare of La Spezia, Italy. The weather was very sunny and windy, so the telescope was covered to minimize telescope alignment drifts due to temperature variations.

| Parameter | Value |
|-----------|------------------------------|
| D | 279.4mm |
| $F_{\#}$ | 10 |
| f_1 | 2794mm ($D \cdot F_{\#}$) |
| f_2 | 75mm |
| f_3 | 100mm |
| f_4 | 30mm |
| f_L | 5.2mm |
| d | 7.5mm ($D \cdot f_2/f_1$) |
| d_L | 2.25mm ($d \cdot f_4/f_3$) |
| λ | 650nm |

Table 4.1: Parameters of the optical setup.



Fig. 4.5: The Celestron C11 telescope used in the horizontal atmospheric turbulence measurements. The optics are attached on a tray over the telescope.

4.5 Experimental Results

4.5.1 Atmospheric Turbulence Measurement

Figure 4.6 shows the values of the C_n^2 (Andrews) computed on 7 experiments collected during the measurement campaign. Both the Andrews and Parry formulation for the C_n^2 are listed in Table 4.2. The differences in average and standard deviation are small.

The values in Table 4.2 are the C_n^2 means and standard deviations over the length of the experiments. To have an idea about the temporal variation of the C_n^2 , the intensity values I have been averaged on a moving window of 300 samples. Figure 4.7 shows the computed time series for one of the dataset: the C_n^2 is quite stable over a short time period (~ 2 minutes). The C_n^2 values computed with the Andrews formulation are slightly higher than the ones computed with the Parry formulation, as shown in Figure 4.8.

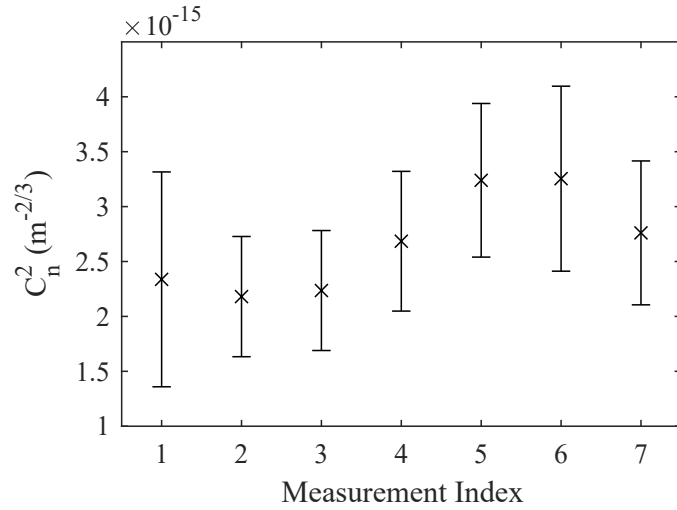


Fig. 4.6: Summary of averaged C_n^2 measurements. The error bars represent the standard deviation of the C_n^2 measurements.

| Recording date | C_n^2 Parry | | C_n^2 Andrews | |
|----------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | μ | σ | μ | σ |
| 14h46m | $2.337e^{-15} m^{-2/3}$ | $9.782e^{-16} m^{-2/3}$ | $2.417e^{-15} m^{-2/3}$ | $1.051e^{-15} m^{-2/3}$ |
| 14h49m | $2.180e^{-15} m^{-2/3}$ | $5.474e^{-16} m^{-2/3}$ | $2.242e^{-15} m^{-2/3}$ | $5.797e^{-16} m^{-2/3}$ |
| 15h23m | $2.235e^{-15} m^{-2/3}$ | $5.463e^{-16} m^{-2/3}$ | $2.299e^{-15} m^{-2/3}$ | $5.809e^{-16} m^{-2/3}$ |
| 15h25m | $2.684e^{-15} m^{-2/3}$ | $6.362e^{-16} m^{-2/3}$ | $2.776e^{-15} m^{-2/3}$ | $6.849e^{-16} m^{-2/3}$ |
| 15h27m | $3.239e^{-15} m^{-2/3}$ | $6.993e^{-16} m^{-2/3}$ | $3.372e^{-15} m^{-2/3}$ | $7.558e^{-16} m^{-2/3}$ |
| 15h28m | $3.254e^{-15} m^{-2/3}$ | $8.422e^{-16} m^{-2/3}$ | $3.391e^{-15} m^{-2/3}$ | $9.211e^{-16} m^{-2/3}$ |
| 15h30m | $2.760e^{-15} m^{-2/3}$ | $6.551e^{-16} m^{-2/3}$ | $2.858e^{-15} m^{-2/3}$ | $7.020e^{-16} m^{-2/3}$ |

Table 4.2: C_n^2 mean μ and standard deviation σ for every dataset.

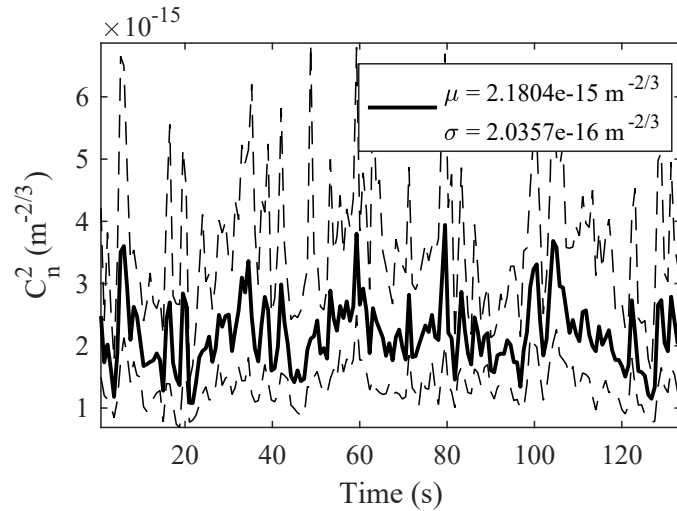


Fig. 4.7: C_n^2 measurement. The solid line is the mean between C_n^2 measured from different lenslets. Dashed lines are the maximum and minimum measured values.

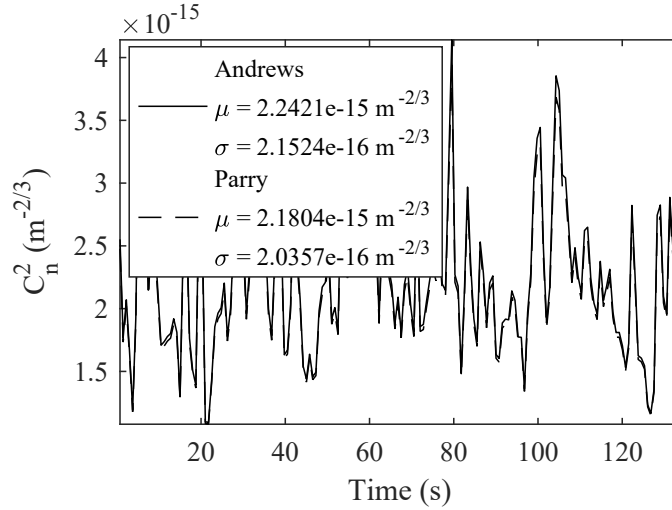


Fig. 4.8: Comparison between Andrews and Parry C_n^2 formulation, complete with mean μ and standard deviation σ .

4.5.2 Statistics of the Zernike modes

While the C_n^2 parameter yields to useful information about the atmospheric turbulence model, for the design of the control system it is sufficient to have an estimate on how the input variables evolve. Zernike modes provide another way to evaluate and measure the horizontal atmospheric turbulence (see Chapter 2). It is worth highlighting that we now use the Zernike decomposition to describe W and not the phase ϕ as before.

The number of modes taken into account depends on the spatial resolution of the SHWFS. Figure 4.9 shows the time series for the first 5 modes $w^{(j)}(t)$ of the Zernike decomposition as a function of time. It is easy to see that the first two modes, tip and tilt, explain most of the variance of the horizontal atmospheric turbulence (this also happens for the vertical turbulence).

To better understand the relationship among the different modes, the histogram of each mode and the Welch's Power Spectral Density (PSD) are reported in Figures 4.10 and 4.11, respectively. Since the datasets were recorded either at 400 or 500Hz, the Welch PSD is evaluated up to 200Hz to ease the comparison.

Figure 4.12 shows the cumulative PSDs of the PSDs in Figure 4.11; these plots allow us to easily see the frequency content mode-by-mode. They are normalized for the sake of readiness: the lower the mode order, the lower the frequency range where most of the power is. This data is important for the design of the control architecture and for the manufacturing of the deformable mirror that should compensate for the atmospheric turbulence.

Remark 4.1. In our experimental setup, the pupil is obscured in the center due to the secondary mirror of the telescope in Figure 4.5. In the computation of the Zernike modes and the slopes we slightly modified the expression for $Z_n(x, y)$ and $Z_m(x, y)$ to work with orthonormal annular Zernike modes [65].

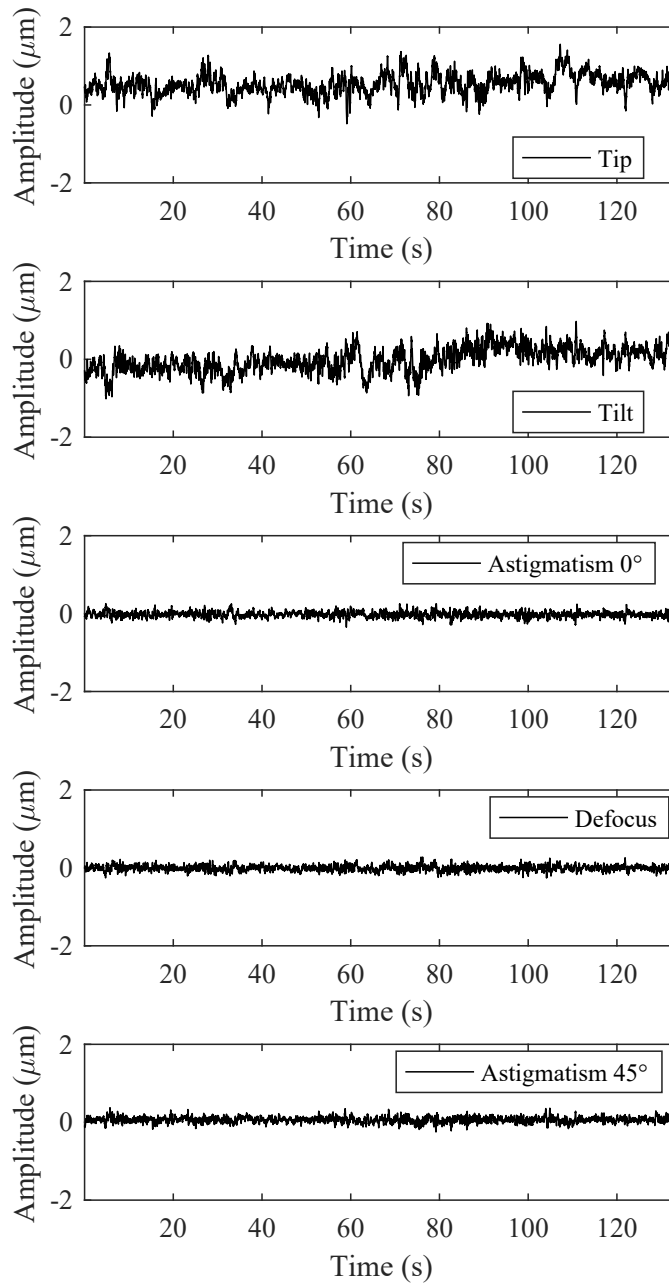


Fig. 4.9: Zernike time series of the first and second order.

SE and LE Strehl ratio can be calculated from the Zernike time series measured from the WFS measurements. For the LE Strehl ratio, the time series is averaged by a moving window of 300 samples. Figures [4.13](#) and [4.14](#) show the time series for the SE and LE Strehl ratio for a dataset. In particular, the LE Strehl ratio

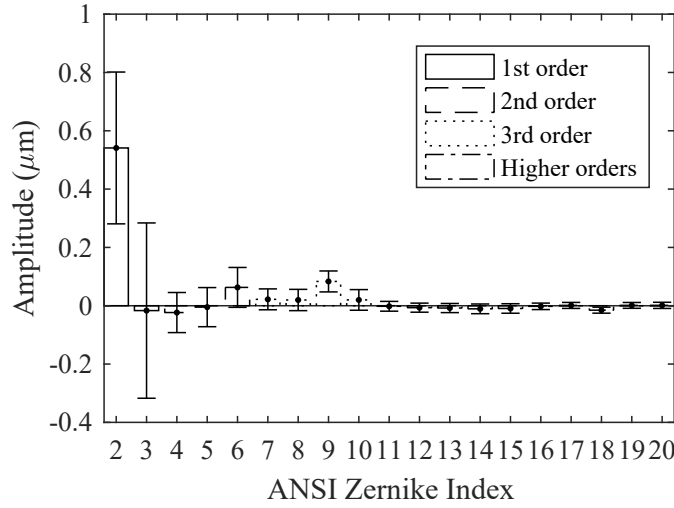


Fig. 4.10: Histogram representation of the mean and variance for each mode of the measured Zernike time series.

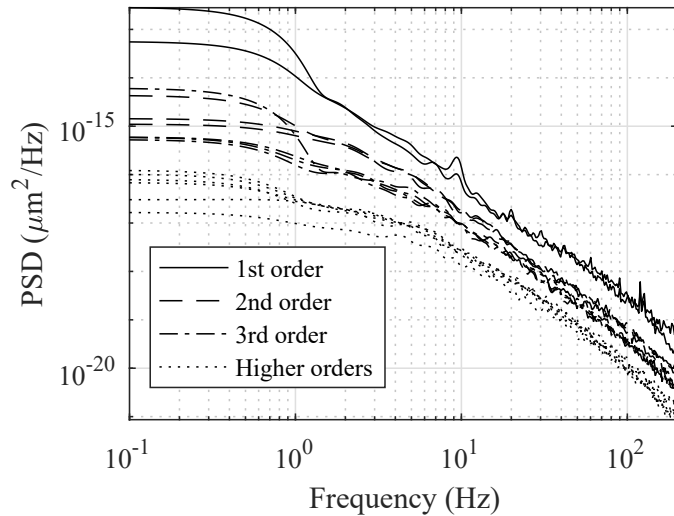


Fig. 4.11: Power Spectral Density of the measured Zernike time series.

shows that the horizontal turbulence is extremely strong and thus requires a very fast DM. The SE and LE Strehl ratio shown are only qualitative, as both are too small for the Marechal approximation to apply.

4.5.3 Atmospheric Turbulence Correction

In this section, the horizontal turbulence in the dataset 19d07m2016y, 14h49m (400Hz as the F_s) is corrected using the experimental setup available at the CNR-IFN facility. Because of the thermal drift that affects the atmospheric turbulence,

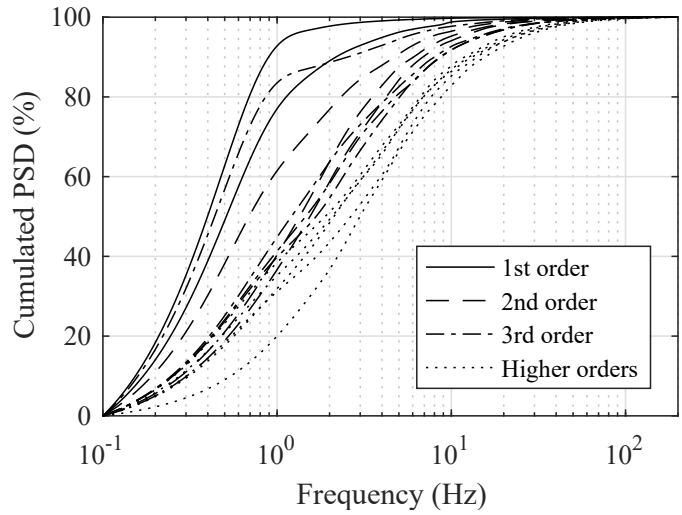


Fig. 4.12: Normalized Cumulative Power Spectral Density of the measured Zernike time series.

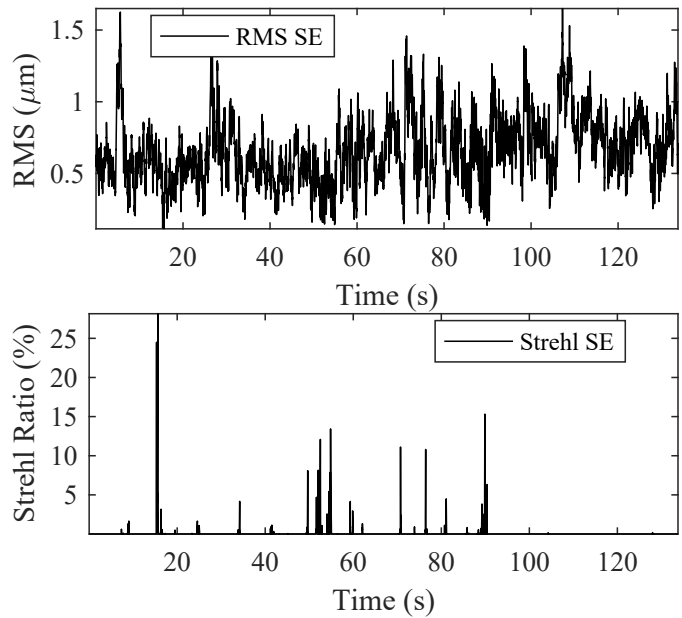


Fig. 4.13: RMS (upper) and Short Exposure Strehl Ratio (lower).

some modal coefficient keep growing. Hence, we will show in the following figures the correction for two time intervals, taking the first sample as reference: $0s - 1.5s$ and $30s - 31.5s$.

The control block diagram is shown in Figure 2.4. The deformable mirror is custom-made and has 32 piezo-electric actuators. In this particular experiment,

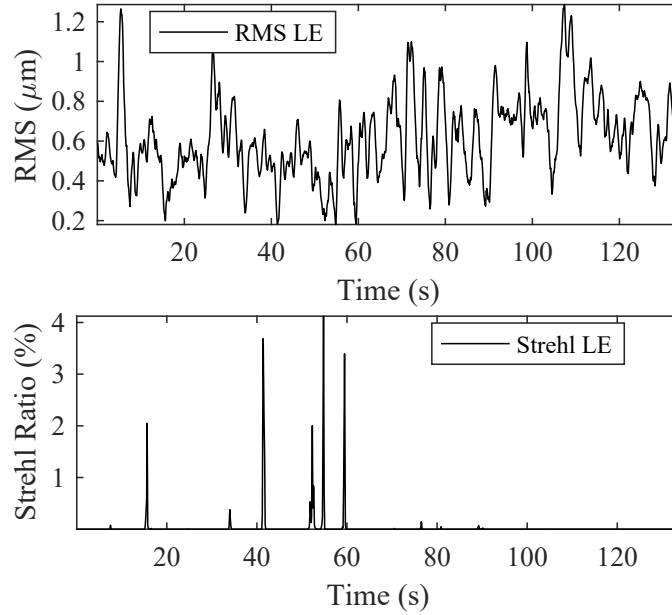


Fig. 4.14: RMS (upper) and Long Exposure Strehl Ratio (lower).

the regulator implemented in the PhotonLoop software is a pure integrative compensator running at $F_s = 400\text{Hz}$ with a computational latency of 0.7ms .

The turbulence is simulated by sending the time series of the Zernike coefficients $w^{(j)}(t)$ as offset to the controller, as explained in Chapter 3. This way, the AO controller behaves the same way as when compensating for an incoming wavefront w . The Zernike decomposition is truncated at $N_w = 20$ terms because the variance of higher order Zernike modes shown in Figure 4.10 is small. In this sense, the wavefront reconstruction error of uncontrolled modes $\sigma_{>N_w}^2$ is negligible. An accurate analysis of the fitting error σ_{fit}^2 involves dedicated phase measurement tools (i.e. independent from the WFS) that were not available.

For each interval we reported

- The Zernike coefficients of the input wavefront and the DM correction (Figures 4.16 and 4.19);
- The RMS of the residual e (Figures 4.17 and 4.20);
- The SE, LE and smoothed LE Strehl ratios (Figures 4.18 and 4.21).

Figure 4.15 shows the Rejection Transfer Functions for the different modes (calculated as in Chapter 3). As expected they are very similar. The differences at the low frequency range are probably due to numerical errors related to the computation of the PSD with the Welch algorithm.

The RMS and Strehl ratio are calculated over the time series (Figures 4.16 and 4.19) using the equations (2.22) and (2.19) in Chapter 2, respectively. Corrected RMS is compared to the input RMS (Figures 4.17 and 4.20), while Strehl ratio is shown in Short Exposure, Long Exposure ($N_k = 200$ samples integration) and smoothed ($N_k = 200$ samples mean).

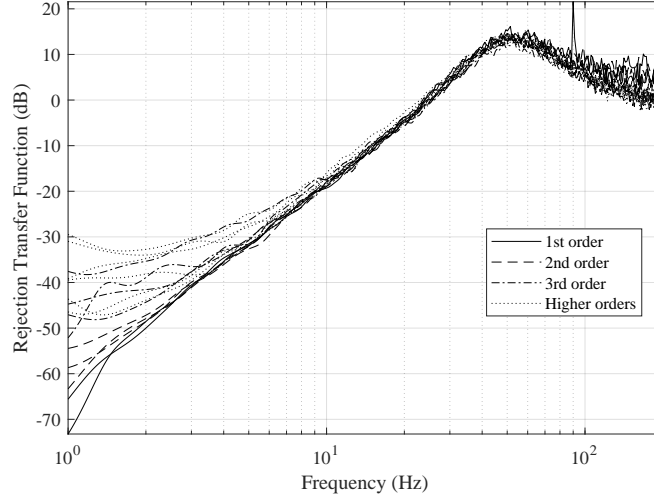


Fig. 4.15: Rejection Transfer Function of the system.

Since the turbulent Zernike time series is injected into the WFS, the regulator actually introduces the aberrations into the AO system, hence resulting in an aberrated PSF. Therefore, the only way to estimate the Strehl ratio is by using the Zernike terms.

Corrections for the first 10 Zernike modes are listed in Table 4.3. The RMS correction ratio is given by:

$$\text{RMS}_{\%} = \left(1 - \frac{\text{RMS}_{\text{error}}}{\text{RMS}_{\text{input}}} \right) \times 100 \tag{4.19}$$

whereas the standard deviation correction ratio by:

$$\sigma_{\%} = \left(1 - \frac{\sigma_{\text{error}}}{\sigma_{\text{input}}} \right) \times 100. \tag{4.20}$$

| Zernike | RMS _{input} (m) | RMS _{error} (m) | RMS _% | σ _{input} (m) | σ _{error} (m) | σ _% |
|-------------|--------------------------|--------------------------|------------------|------------------------|------------------------|----------------|
| Tilt | 7.638e-05 | 5.141e-06 | 93.27 | 4.837e-14 | 1.099e-15 | 97.73 |
| Tip | 4.431e-05 | 5.455e-06 | 87.69 | 4.579e-14 | 1.240e-15 | 97.29 |
| Ast 45° | 1.123e-05 | 2.441e-06 | 78.26 | 4.888e-15 | 2.466e-16 | 94.95 |
| Defocus | 1.003e-05 | 2.446e-06 | 75.60 | 4.167e-15 | 2.477e-16 | 94.06 |
| Ast 0° | 1.463e-05 | 2.562e-06 | 82.48 | 4.323e-15 | 2.735e-16 | 93.67 |
| Trefoil 0° | 6.309e-06 | 1.768e-06 | 71.98 | 1.215e-15 | 1.278e-16 | 89.48 |
| Coma 0° | 6.499e-06 | 1.918e-06 | 70.48 | 1.334e-15 | 1.350e-16 | 89.88 |
| Coma 45° | 1.384e-05 | 1.966e-06 | 85.80 | 1.180e-15 | 1.558e-16 | 86.80 |
| Trefoil 45° | 6.327e-06 | 1.735e-06 | 72.58 | 1.184e-15 | 1.216e-16 | 89.73 |

Table 4.3: Correction in closed loop.

The results in Table 4.3 show that the SCAO system is performing very well, reflecting into a Strehl ratio of 70%.

The LE Strehl ratio reaches the Smoothed SE Strehl after a time transient (Figure 4.18) and follows the latter thereafter (Figure 4.21). Most of the turbulence variance is in the Tilt-Tip modes and is well compensated (>87%) by the feedback loop, whereas the compensation factor of higher order modes is more modest.

4.6 Conclusion

In this Chapter, a technique for estimating the atmospheric structure constant C_n^2 and wavefront spatio-temporal dynamics, represented in the Zernike polynomials basis, has been described. The proposed method exploits the Shack-Hartmann WFS intensity and slope measurements, mimicking the SCIDAR principle while using less optical components. Information about the atmospheric turbulence can be inferred from the C_n^2 parameter (i.e., Fried spatial coherence parameter r_0 , Greenwood temporal coherency frequency τ_0 , isoplanatic angle θ_0). This information can be used as a baseline to design the AO system taking into consideration the AO error budget σ_{AO}^2 introduced in Chapter 2. Furthermore, the Zernike terms calculated from the measured slopes can be fitted into an AR process to model the atmospheric turbulence mode by mode. Converting each obtained AR model into a stochastic SISO LTI state-space system is useful in the synthesis of optimal model-based controllers as explained in Chapter 6.

The compensation performance could be improved by:

- Using more powerful hardware architectures to increase the sampling frequency and reduce the latency (e.g. using a GPU-based architecture as shown in Chapter 5), reducing the bandwidth error $\sigma_{T_s}^2$ and delay error σ_{delay}^2 , respectively;
- Exploiting the atmospheric turbulence statistics to design a model-based control strategy that predicts the step-ahead aberration time sample to further reduce the delay error σ_{delay}^2 , as shown in Chapter 7.

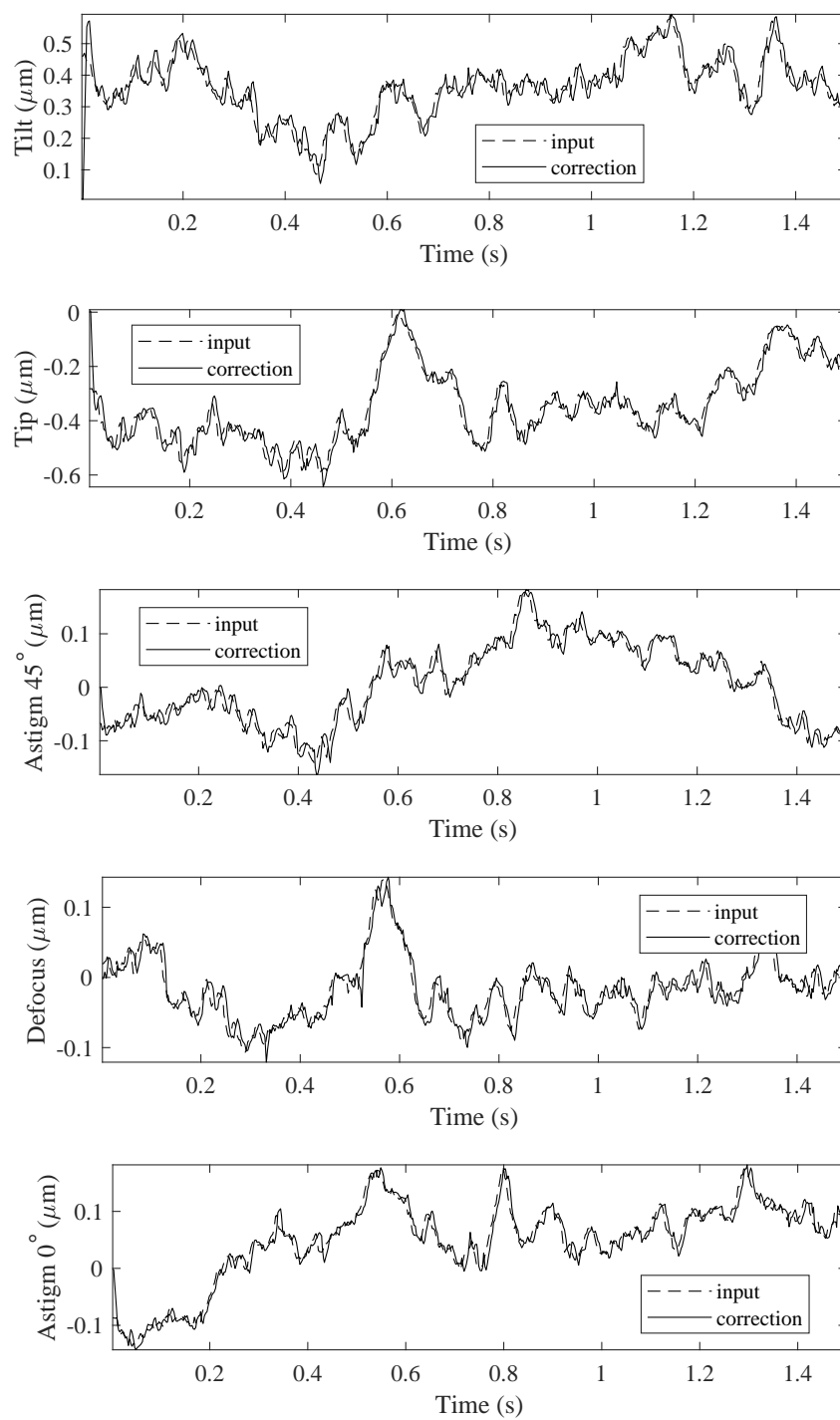


Fig. 4.16: 0s – 1.5s: Correction of the Zernike time series.

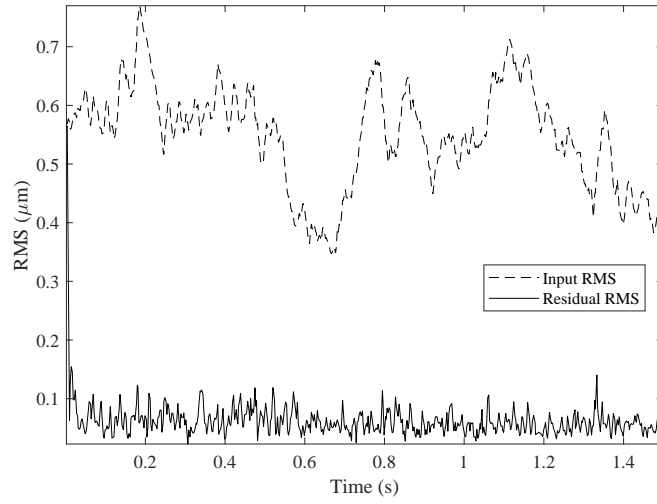


Fig. 4.17: 0s – 1.5s: RMS of the corrected time series.

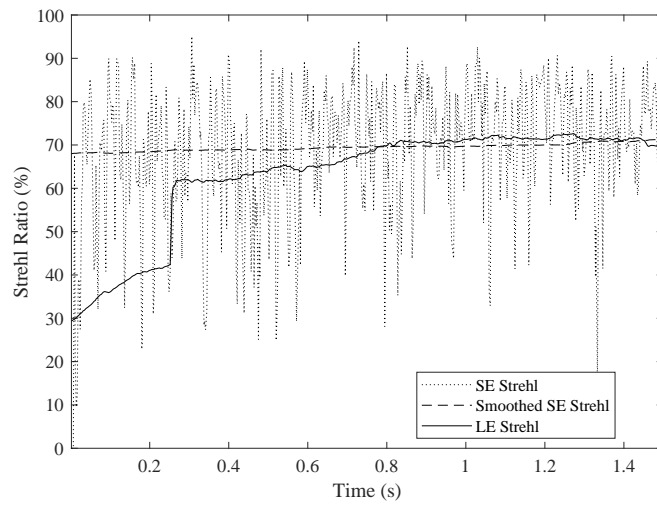


Fig. 4.18: 0s – 1.5s: Strehl of the corrected time series.

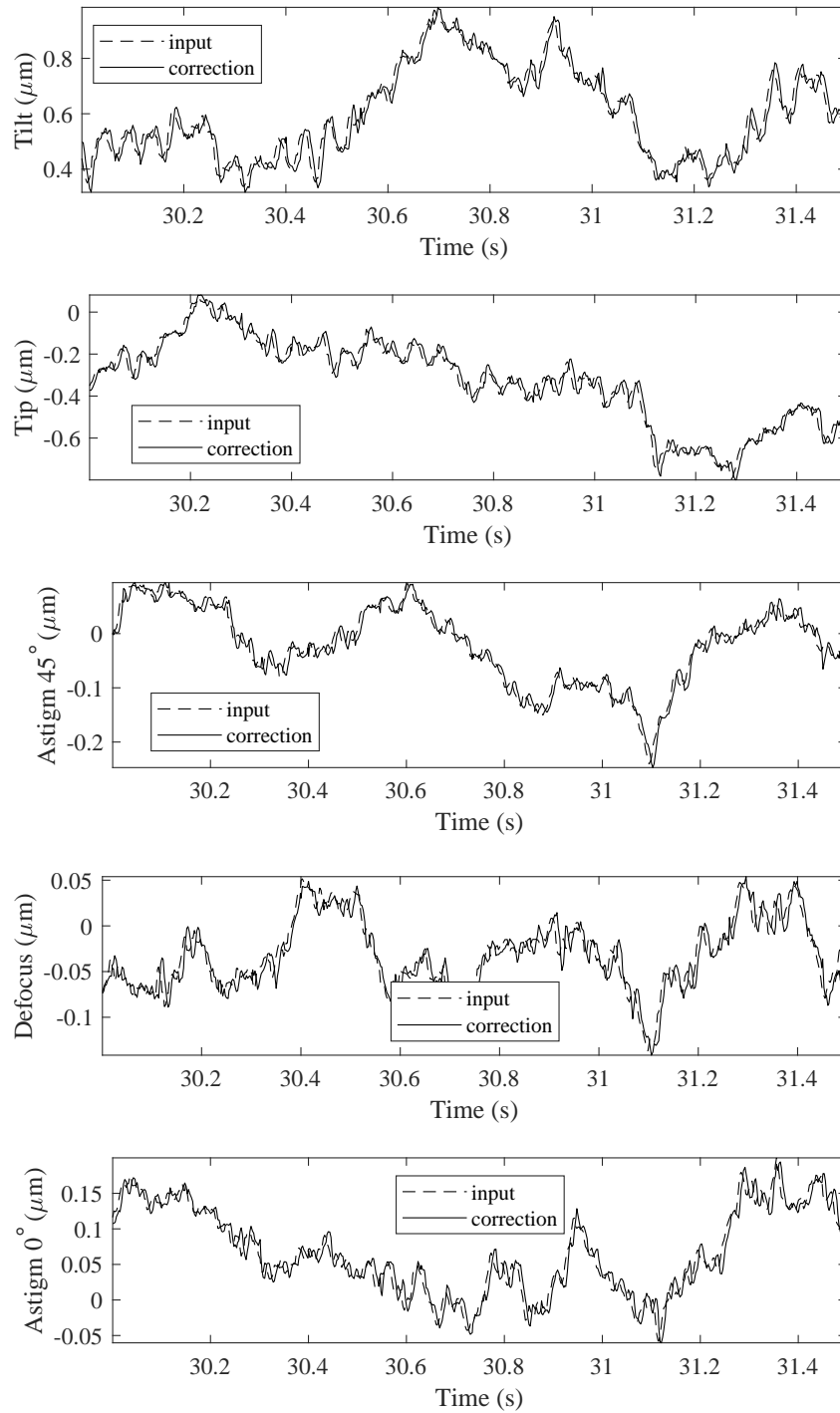


Fig. 4.19: 30s – 31.5s: Correction of the Zernike time series.

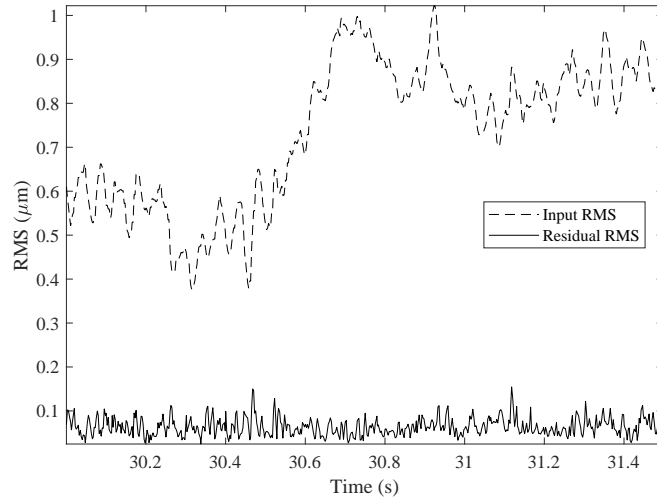


Fig. 4.20: 30s – 31.5s: RMS of the corrected time series.

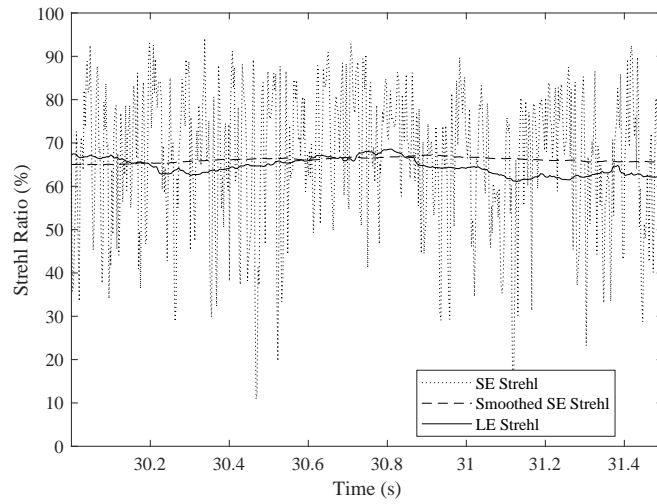



Fig. 4.21: 30s – 31.5s: Strehl of the corrected time series.

SH-WFS on Embedded GPU

Summary. One of Adaptive Optics key components is the Wavefront Sensor, which is typically implemented by a Shack-Hartmann sensor capturing images related to the aberrated wavefront. To compensate for the wavefront aberration temporal dynamics, the control loop latency must be minimized. GPU-based architectures are comparable to FPGA for data throughput and to CPU for programming ease. We propose a SH-WFS centroid extraction algorithm that is tailored for the NVIDIA Jetson TX2. Since this device is characterized by a unified CPU/GPU memory, it allows to sensibly reduce useless CPU/GPU data transfers. Being a compact, power-efficient embedded board, it can be fitted as part of a smart camera. The experiments over images up to 1000×1000 pixels wide show that the computational latency is compatible with the AO closed-loop latency constraint of 2ms for low-cost small/medium sized AO systems. 

¹ This chapter is based on the paper: J. Mocci, F. Busato, N. Bombieri, S. Bonora, and R. Muradore. “An efficient implementation of the Shack-Hartmann wavefront reconstruction for edge computing” (submitted).

5.1 Introduction

The Shack-Hartmann WFS (SH-WFS) is one of the most used wavefront sensor in AO thanks to its effectiveness and relative simplicity. A SH-WFS measures the local gradients of the incoming wavefront by spatially sampling it with a lenslet array. Each lenslet focuses the local subaperture into a CCD or CMOS camera pixel array. The displacement of the focused spot centroids from their reference positions are related to the wavefront gradient.

The centroids extraction algorithm can be divided into three phases:

1. Identify the spots in the image;
2. Calculate the spots positions;
3. Associate each spot position to its reference position.

Several methods can be used to identify and associate the image spots, like the spiral algorithm developed by Mauch et al. [68]. However, when the centroids spatial dynamic is small (e.g., almost flat wavefront) then each spot position is considered to be contained in a known closed, non-overlapping portion of the camera pixels, avoiding the spot identification step. There are three main classes of centroids extraction algorithms: moment, quad-cell and correlation-based [115] [50] [121]. The moment-based algorithms exploit the Center of Gravity (CoG) of the spot, and is well suited to Gaussian-shaped spots. Improvements to it include weighting and thresholding the intensities. A special case of the CoG algorithm is the Quad Cell (QC), where each spot is represented by just 4 pixels. The correlation-based algorithms first correlate the spot with a reference spot, and then calculate the centroid of the resulting correlation map. Due to the added computational complexity, correlation is prevalently used when dealing with extended sources.

Since the AO feedback control loop should be fast and reactive enough to compensate for wavefront aberrations dynamics, the WFS measurement latency must be minimized by either sacrificing wavefront spatial resolution (e.g., smaller beam aperture leads to higher photon flux on fewer pixels to be transmitted and computed) or by choosing a suitable architecture to implement a fine-tuned extraction algorithm. It is natural to associate the SH-WFS centroid extraction problem to a throughput-oriented Single-Instruction Multiple-Data (SIMD) paradigm, so that the centroids can be computed in parallel using the same set of instructions.

Graphic Processing Units (GPU) and Field-Programmable-Gate-Array (FPGA) architectures are examples of SIMD devices. While the FPGA excels in raw performance compared to the GPU, the advantage of GPU over FPGA is the *programming flexibility*, which is almost as high as the latency-oriented Central Processing Units (CPU). Some AO-related algorithms implemented on GPU are demonstrated to perform even better than their FPGA counterparts. For example, the GPU implementation of a correlation-based algorithm compares favorably with respect to the FPGA implementation [120]. The big amount of data coming from the SH-WFS in the Very and Extremely Large Telescopes can be tackled by GPU clusters to accelerate the gradients extraction. Perret et al. take advantage of DMA mechanisms between FPGA and GPU to minimize latency [85], whereas Lainée et al. implement the whole control software on GPU [57]. Beside from the real-time computation, introductory work on the tomographic reconstruction of the atmospheric

turbulence through neural networks on GPU has also been done [35] and several GPU-accelerated optics simulation frameworks have been developed [30] [10] [102].

However, little attention has been given to the GPU implementation of the SH-WFS algorithms for *small to medium AO systems*. The bottleneck is mainly in the memory transfer overhead between GPU and the rest of the architecture data-flow which adds to the overall latency. For that reason, the low-latency characteristic of CPU-based strategies [70] [9] and FPGA designed as so-called Wavefront Processing Units (WPU) [47] [89] [110] [68] is still desirable. The recent hybrid CPU/GPU Jetson architectures from NVIDIA overcome the memory transfer latency overhead by unifying CPU and GPU memories. Also, such devices can be integrated in *edge-computing* systems thanks to their compact foot-print, power-efficiency and competitive High-Performance Computing (HPC) capabilities as pointed out by Ukidave et al. [119]. Edge-computing is an emerging paradigm where data are elaborated as close as possible to their source. In the AO field, this approach could be used to further miniaturize the AO controller of small-size telescopes [93]. Pairing an FPGA to a camera sensor yields to the so-called *smart-camera* [26], which is able to process the wavefront with minimal latency [97] [114]. Nonetheless, the Jetson architecture can substitute FPGA smart cameras when dealing with hard to implement visual computing techniques [60] [16] [129].

5.2 Moment-based Centroid Extraction

The SH-WFS is a sensor that measures the wavefront distortions by computing its local gradients. Figure 5.1 shows the optical principle that allows the SH-WFS to sample the incoming wavefront in time and space. The l -th lenslet of the lenslet array focuses the local wavefront into a light spot on the *pixel array* of the capture device. Assuming a point-like light source (e.g., a distant star), the *centroid position* $c_l \in \mathbb{R}^2$ of the spot is related to the spatial displacement of the incoming aberrated wavefront with respect to the flat wavefront (no aberrations).

Let $D \in \mathbb{R}$ be the diameter of the telescope circular aperture and L be the grid-shaped lenslet array in which the incoming wavefront is inscribed. The number of lenslets in each row of the lenslet grid is given by

$$w_L = \left\lfloor \frac{D}{d_L} \right\rfloor, \quad (5.1)$$

where $d_L \in \mathbb{R}$ is the size of each lenslet as shown in Figure 5.1.

The lenslets focal $f \in \mathbb{R}$ determines the size of the pixel region in which the l -th spot is focused into and is tuned to trade-off dynamic range and sensitivity. Spots might be imaged anywhere in the pixel array, resulting in arbitrarily large and overlapping pixel regions. However, if f is sufficiently small in relation to the expected incoming wavefront spatial variance, then the pixel regions are disjoint. This is a valid assumption when measuring the wavefront compensated by an AO control system. Figure 5.2 shows how the intensity values of a lenslet spot are distributed in a pixel region.

Each pixel region is a square since the lenslet is arranged in a square grid. The number of pixels contained in a row of a pixel region is

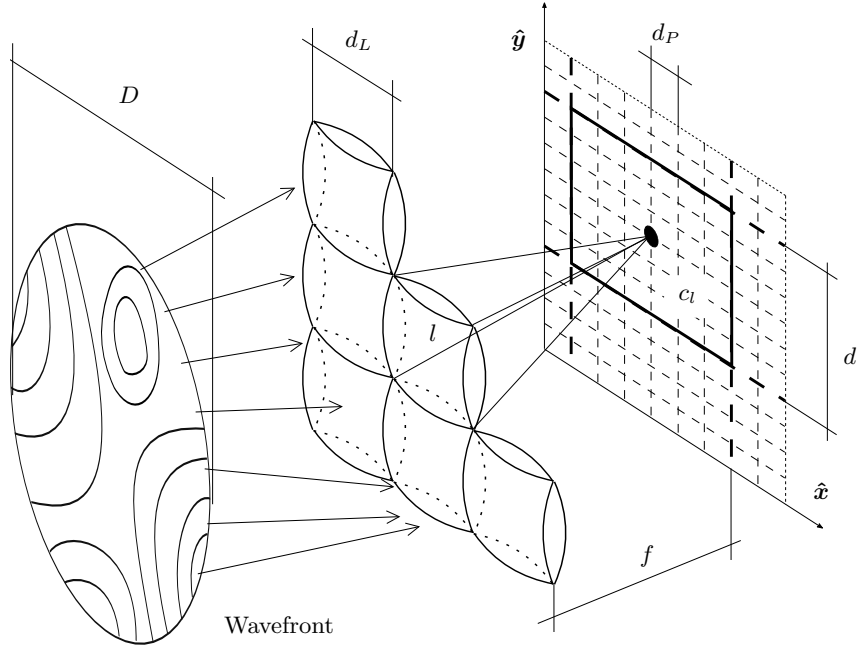


Fig. 5.1: Operation principle of the Shack-Hartmann WFS.

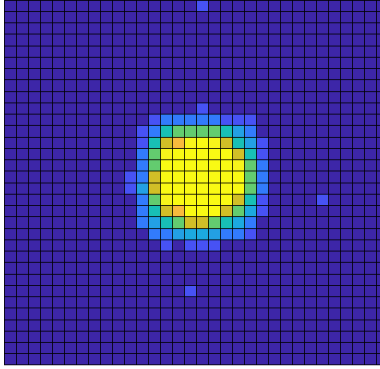


Fig. 5.2: Pixel region close-up. . Yellow dots represents higher intensity values.

$$d = \frac{d_L}{d_P}, \quad (5.2)$$

where $d_P \in \mathbb{R}$ is the pixel width (assuming no *dead zone* among regions).

Let $p = (x_p, y_p)$ be a pixel position, with $x_p, y_p \in \mathbb{N}$ being the Cartesian coordinates of the pixel in the pixel array. Each lenslet l focuses the spot into the pixels in the l -th pixel region

$$P_l = \{p \mid \lfloor x_l \rfloor \leq x_p < \lfloor x_l + d \rfloor, \lfloor y_l \rfloor \leq y_p < \lfloor y_l + d \rfloor\}, \quad (5.3)$$

where $(x_l, y_l) \in \mathbb{R}^2$ are the pixel coordinates of the bottom-left corner of the pixel region. The mapping from the lenslet array into the pixel array is calibrated by identifying the pixel position (x_0, y_0) located at the bottom-left pixel of the bottom-left lenslet in the lenslet grid. Then, the bottom-left pixel coordinates of the l -th pixel region are calculated as

$$(x_l, y_l) = \left(\left\lfloor x_0 + d \left\lfloor \frac{l}{w_L} \right\rfloor \right\rfloor, \left\lfloor y_0 + d(l \bmod w_L) \right\rfloor \right). \quad (5.4)$$

Joining all pixel regions yields to the Region of Interest (RoI) P , which is a square region having width

$$w_P = dw_L, \quad (5.5)$$

as illustrated in Figure 5.3

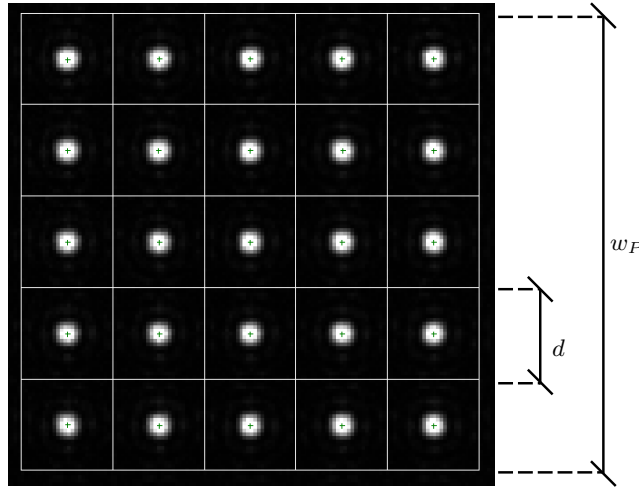


Fig. 5.3: SH-WFS spot image. The pixel regions P_L of size $d \times d$ are adjacent each-other, and together form the pixel RoI P of size $w_P \times w_P$.

Comparing the measured centroid position c_l with its *reference position*, i.e. the centroid measured when the wavefront is flat, yields to the wavefront phase gradients (slopes) from which the wavefront phase can be reconstructed by zonal or modal techniques (see Chapter 2).

Let $I(p) \in \mathbb{R}$ be the measured intensity value at a RoI pixel $p \in P$. The image moments for the l -th spot image are defined as

$$m_l^{ij} = \sum_{p \in P_l} x_p^i y_p^j I(p) \quad (5.6)$$

where $i, j \in \mathbb{N}$ are the Cartesian directions order, respectively. Using the image moments, the l -th centroid position c_l is the CoG of the l -th spot calculated as

$$c_l = \left(\frac{m_l^{10}}{m_l^{00}}, \frac{m_l^{01}}{m_l^{00}} \right). \quad (5.7)$$

The centroids can be extracted from the WFS spot image by moment-based methods [115], which are used to calculate the Center of Gravity (CoG) of each spot on the pixel array. To enhance the detection accuracy when the camera sensor Signal-to-Noise Ratio (SNR) is poor, the CoG is improved using a Thresholded Weighted CoG (TWCoG) to reject noisy pixels (thresholding) and favor brighter intensities (weighting).

Using the moments-based centroids calculation described in (5.6) and (5.7) with the pixel region defined as in (5.3), all the centroids of the spot image can be extracted pixel-wise as explained in Algorithm 2 or lenslet-wise as in Algorithm 3 [70].

Algorithm 2 Pixel-wise Centroids Extraction

```

for all  $l \in \mathbb{L}$  do
   $m_l^{00}, m_l^{10}, m_l^{01} \leftarrow 0$ 
end for
for all  $(x_p, y_p) \in \mathbb{P}$  do
   $l = \left\lfloor \frac{x_p - x_0}{d} \right\rfloor + w_L \left\lfloor \frac{y_p - y_0}{d} \right\rfloor$ 
   $m_l^{00} \leftarrow m_l^{00} + I(x_p, y_p)$ 
   $m_l^{10} \leftarrow m_l^{10} + (x_p - x_l)I(x_p, y_p)$ 
   $m_l^{01} \leftarrow m_l^{01} + (y_p - y_l)I(x_p, y_p)$ 
end for
for all  $l \in \mathbb{L}$  do
   $c_l \leftarrow \left( \frac{m_l^{10}}{m_l^{00}}, \frac{m_l^{01}}{m_l^{00}} \right)$ 
end for

```

Algorithm 3 Lenslet-wise Centroids Extraction

```

for all  $l \in L$  do
   $m_l^{00}, m_l^{10}, m_l^{01} \leftarrow 0$ 
  for all  $p \in P_l$  do
     $m_l^{00} \leftarrow m_l^{00} + I(p)$ 
     $m_l^{10} \leftarrow m_l^{10} + x_p I(p)$ 
     $m_l^{01} \leftarrow m_l^{01} + y_p I(p)$ 
  end for
   $c_l \leftarrow \left( \frac{m_l^{10}}{m_l^{00}}, \frac{m_l^{01}}{m_l^{00}} \right)$ 
end for

```

While both algorithms produce the correct result, the pixel topology is queried only once per lenslet in the lenslet-wise algorithm instead of once per pixel as in the pixel-wise one. Since there are less lenslets than pixels, the lenslet-based algorithm requires less operations to be done. Therefore, the lenslet-based algorithm has lower computational complexity than the pixel-based algorithm. Both algorithms

benefit from calculating *a priori* the lenslet or pixel topology (e.g. look-up tables) to shift from temporal to spatial complexity.

The lower bound of the computational cost for the centroids extraction algorithm [3](#) is given by the analysis of the two nested for-loops. Since the pixel regions P_l covering the lenslet array L are adjacent, disjoint and completely contained in the pixel RoI P , the inner loop iterates over each pixel $p \in P$. The pixel RoI P is square and hence its size is $\lfloor dw_L \rfloor^2$. A total of 5 operations are needed for each pixel p to update the moments. The outer loop iterates over each lenslet $l \in L$, where the lenslet grid L is square having size w_L^2 . Each centroid c_l requires 2 instructions to be calculated from the moments.

Therefore, the lower bound complexity depends on the lenslet grid width w_L and the pixel region width d :

$$\Omega(w_L, d) = 2w_L^2 + 5 \lfloor dw_L \rfloor^2. \quad (5.8)$$

Since neighboring lenslets are disjoint and the addends of [\(5.6\)](#) are mutually independent, all the addends can be computed concurrently at once and then summed up into the respective lenslet moments.

5.3 Parallel Implementation

CUDA and OpenCL are the two dominant frameworks for parallel programming. They offer the same capabilities, e.g., exploiting unified memory, albeit with different hardware terminology and code syntax. OpenCL is open-source and compatible with a wide range of GPU and multi-core CPU architectures, whereas CUDA is proprietary and only compatible on NVIDIA architectures. On the other hand, since the CUDA framework is specific to NVIDIA architectures, it guarantees tighter integration to the underlying hardware and more reliable development tools.

We rely on the CUDA parallel framework to fully exploit the Jetson architecture. In the CUDA programming language, a *device* (GPU) utilizes its *threads* and *memory* to execute *kernels* (i.e., functions) called by the *host* (CPU):

- Threads are organized in *blocks*;
- Blocks are contained in a *grid*;
- The scheduler maps blocks into multiple *streaming multiprocessors*;
- Each thread in a block is mapped into a *core*;
- Up to 32 threads (i.e., one *warp* of threads) can be scheduled to concurrently execute the same instructions (i.e. SIMD).

Each thread can access *local* or *shared* memory. While the former is limited to the thread scope, the latter is available to all threads in a block and, hence, can be used for efficient communication among threads. Local and shared memories are device-side resources meant to store temporary results. The host can access the device *global* memory to store the dataset to be processed by the kernel and read the results.

5.3.1 Data Levels

Since the spot image is stored into memory as a *row-major* linear array, the pixel expressed in Cartesian coordinate $p = (x_p, y_p)$ is mapped into the array index as

$$\phi = x_p + w_P y_p. \quad (5.9)$$

By accessing contiguous elements along the rows, neighboring data chunks are cached in fast memory leading to smaller transfer latency time. To leverage the memory cache, the spot image is partitioned into data levels:

- The topmost level is the entire *spot image* in which the pixels RoI P , i.e., the lenslet grid, is immersed;
- The lenslet grid is divided into rows that are spanned by *lenslet groups*. Each lenslet group \bar{l} has the same power-of-two size, up to 16 lenslets wide, to satisfy the condition for the reduction operation on its elements. Multiple lenslet groups are concatenated to cover all lenslets in a row. Eventual remainder lenslets in the last lenslet group are padded with zeros;
- A lenslet l is partitioned in a power-of-two number of stacked *row groups*. The size of each row group $\phi_{\bar{r}}^l$ is given by the rounded up ratio between the number of pixel rows in a lenslet and the number of row groups to be assigned for a lenslet. Partial moments associated to the remainder rows of the last row group are outside the lenslet and hence set to zero to remove them from computation;
- Each *row* $\phi_{\bar{r}}^{l,\bar{r}}$ in a row group is composed of *pixels*;
- Pixels in a row are covered by *pixel groups*, with each pixel group $\phi_{\bar{p}}^{l,\bar{r},r}$ containing 4 pixels $\phi_p^{l,\bar{r},r,\bar{p}}$ to optimize memory transfer. Since pixel groups are aligned to the beginning of the spot image array, leading and trailing pixel groups may contain pixels outside the scope of the actual lenslet. The intensities of such pixels are set to zero so that they don't appear as terms of the partial moments.

Each level calculates the *partial moment* from the contained elements:

- The *pixel partial* of the pixel $\phi_p^{l,\bar{r},r,\bar{p}}$ is

$$\mu_p^{l,\bar{r},r,\bar{p}} = \{m^{00}(\phi_p^{l,\bar{r},r,\bar{p}}); m^{10}(\phi_p^{l,\bar{r},r,\bar{p}}); m^{01}(\phi_p^{l,\bar{r},r,\bar{p}})\}, \quad (5.10)$$

with the moments m^{00} , m^{10} and m^{01} calculated at the array index $\phi_p^{l,\bar{r},r,\bar{p}}$;

- Summing together all the pixel partials associated to a pixel group yields to the *pixel group partials* $\mu_{\bar{p}}^{l,\bar{r},r}$;
- The *row partial* $\mu_{\bar{r}}^{l,\bar{r}}$ of a row group is obtained by adding the pixel group partials calculated on its underlying pixel groups;
- The *row group partial* $\mu_{\bar{r}}^l$ of a lenslet is given by summing up all the row partials in a row group;
- Finally, the *lenslet moment* μ_l of the lenslet grid is the sum of the row group partials contained in the l -th lenslet.

Figure 5.4 describes how a single pixel p of the spot image is indexed through the data levels. The lenslet grid is 7 lenslets wide and the lenslet size is 10×10 pixels. With the lenslet groups size fixed to 4 lenslets, it takes 2 lenslet groups to cover a row of lenslets, with one lenslet as remainder. In the figure, the 5-th

lenslet group accesses lenslets 18, 19 and 20, zero padding the remainder. Each lenslet is partitioned into 2 row groups and hence each row group spans 5 rows, with no remainder. Depending on the memory alignment, up to 4 pixel groups can be concatenated to span each lenslet row. In the figure, the 2-nd pixel group reads the remainder pixel intensities as zeros (the 3-rd pixel group is only zeros and hence not drawn). The selected pixel p is indexed as $\phi_1^{20,1,4,2}$ and produces the partial pixel moment $\mu_1^{20,1,4,2}$. Adding together the partial moments from the bottom to the top of the data hierarchy levels yields to the lenslet moment μ_{20} .

5.3.2 Optimized GPU Data Transfer through Coalesced Memory Accesses

In most architectures, the global memory and the CPU memory are physically decoupled. This means that data has to be transferred between memories, with each transfer increasing the temporal overhead over the execution time. To overcome such overheads, latency-hiding techniques can be exploited, e.g., overlapping transfer and computation phases by extracting centroids as soon as the pixels regions are transferred. However, the kernel execution time must be comparable to the transfer time to take advantage of those techniques. Since host and device on a Jetson architecture physically share the same global memory, allocated memory can be addressed both by host and device by *pinning* it (page-locked mapped memory, also called Zero-Copy), hence avoiding any transfer overhead.

Remark 5.1. Page-locked memory on the Jetson TX2 GPU is not cached when accessed by the CPU. As a consequence, host-side memory reading is not optimized. Nevertheless, since the CPU only accesses the memory to write the intensity values acquired from the sensor, there are no performance penalty using page-locked memory instead of other addressing options.

Up to 128 Bytes of data in global memory can be accessed in one transaction. To fully exploit the cache, threads in a warp should ideally access consecutive single precision words (4 Bytes) starting from a 128 Bytes-aligned address to realize a *coalesced* memory transfer. Since the spot image is encoded into a row-major 8-bit array with stride divisible by 4, each thread of a warp reads adjacent words of 4 Bytes, hence accessing a full pixel group at once.

The highest bandwidth throughput is achieved when a warp reads 32 pixel groups, i.e. 128 Bytes. However, since the implementation is limited to lenslets having pitch smaller than 30 pixels, no more than 8 pixel groups are needed to cover an entire lenslet row. To maximize caching performance, the warp operates on the contiguous rows of a lenslet group instead of a single lenslet row. By defining the number of packets needed to cover a lenslet row as

$$n_{\text{packets}} = \left\lceil \frac{d+6}{4} \right\rceil, \quad (5.11)$$

the optimal size of the lenslet group is

$$d_{\bar{L}} = \exp_2 \left\lceil \log_2 \left\lceil \frac{32}{n_{\text{packets}}} \right\rceil \right\rceil. \quad (5.12)$$

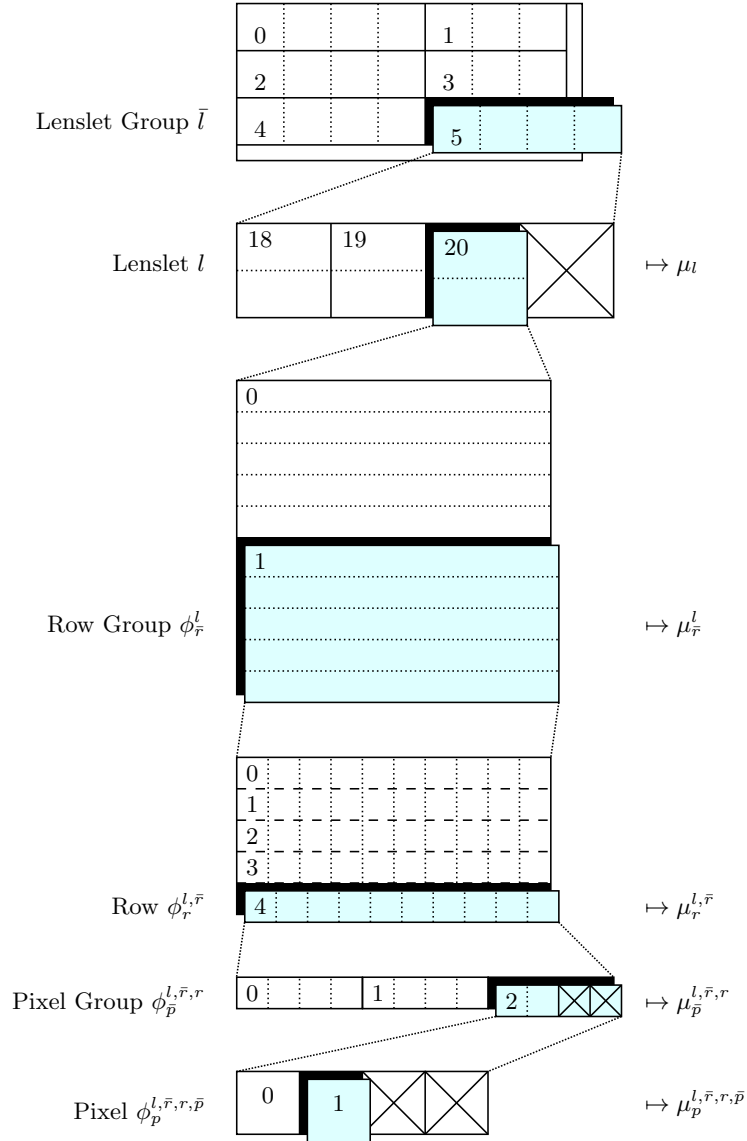


Fig. 5.4: Diagram of the data hierarchy levels as seen from the CUDA extraction algorithm. On the right: partials obtained by reducing the current level data. With the exception of the lenslet moment, numerical indexes are relative to the level.

Since the lenslet groups are consecutively stacked to cover a row of lenslets of the lenslet grid starting from the leftmost lenslet, the lenslets of the last rightmost lenslet group outside the lenslet grid are ignored.

5.3.3 Data Reduction

Parallel data reduction summarizes (by using a commutative binary operator) all the homogeneous data of a dataset by exploiting communication and synchronization primitives among concurrent execution units. Given 2^n elements in a dataset mapped to a pool of 2^n execution units, $n \in \mathbb{N}$, the elements of each disjoint pair of execution units in the pool are reduced concurrently into intermediate results, which are half the size of the original dataset. This process is iterated over such intermediate results, with each iteration halving their size. The reduction ends when there are no pairs left, and the value returned is the one calculated in the last iteration. Since operations are done concurrently in a *logarithmic-tree* fashion, the computational cost to reduce 2^n elements is $O(n)$. If the number of elements is not power-of-two, then the dataset is padded with elements whose value is neutral with respect to the binary operator considered.

In the CUDA programming model, the parallel reduction can be efficiently implemented at warp level through *shuffle* primitives. The shuffle instructions let threads access the registers of other threads scheduled in the same warp, despite registers being local to the threads. By exploiting such primitives, the parallel reduction operates over registers, which are the fastest type of memory in the CUDA architecture.

Since a warp consists of 32 concurrent threads, the dataset to be reduced must be 32 elements large to achieve the peak efficiency. However, several smaller datasets can be operated at once by combining them into a 32-elements dataset. To do so, the datasets must have the same power-of-two size, eventually padding remainder elements with neutral elements. Then, they are concatenated into the full dataset, which must be also padded with neutral elements if its size is not power-of-two. To avoid cross-talking between intermediate reduction results from different datasets, elements are accessed via the *butterfly* (i.e. XOR) scheme. Assuming 2^m datasets, the reduction algorithm returns the results after $\log_2(32) - m + 1$ iterations.

Figure 5.5 shows how the pixels of the spot image are reduced to lenslet moments μ_l . A warp covers the adjacent pixel rows of the lenslets contained in the lenslet group \bar{l} , one row per lenslet. Each thread in the warp sequentially reduces the pixel partials $\mu_p^{l,\bar{r},r,\bar{p}}$ into the pixel group partial $\mu_{\bar{p}}^{l,\bar{r},r}$. Then, the warp performs a parallel shuffle reduction over the dataset built from the pixel group partials, yielding to the row partials $\mu_{\bar{r}}^{l,\bar{r}}$.

Row partials within a row group $\phi_{\bar{r}}^l$ are operated one after the other by a single warp. To increase *occupancy*, multiple warps can be associated to a lenslet group, one for each row group. To do so, threads in a block are partitioned into warps by using *cooperative groups*, a CUDA implementation feature that lets warps to be synchronized independently. All row group partials $\mu_{\bar{r}}^l$ are then stored in shared memory and reduced in parallel into lenslet partials μ_l .

In the example in Figure 5.5, warps are limited to 8 threads for the sake of space. The pixel groups of the 2 lenslet rows are demuxed and summed into 8 pixel group partials, the last of which is zero. The shuffle reduction yields to the row partials of the lenslets 0 and 1, which are accumulated to the previously calculated partials resulting into the 0-th row group partials of both lenslets. Such row group partials are stored in shared memory. Since the lenslet is partitioned

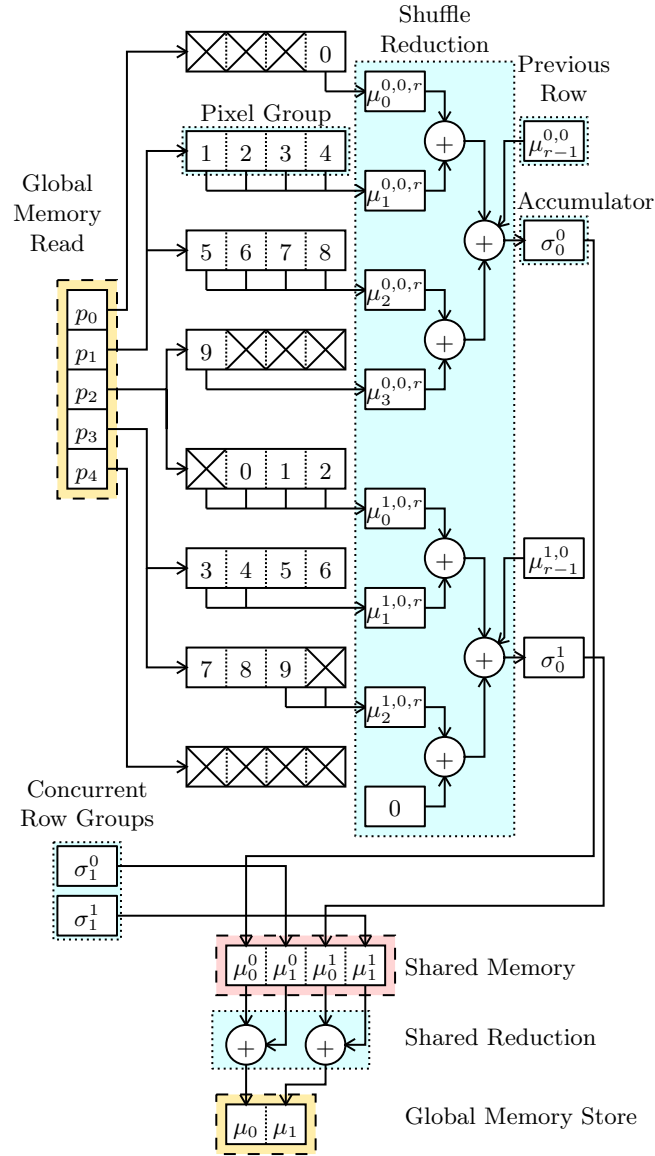


Fig. 5.5: Diagram of the data reduction. In this example, the pixel region width is $d = 10$, requiring 4 pixel groups. There are 2 row groups with 5 rows each.

into 2 row groups, the number of threads scheduled for a block is set to 16 so that two cooperative groups of 8 threads can be formed. Hence, the 1-st row group partials are calculated concurrently. Reducing the row group partials stored in shared memory yields to the two lenslet moments.

5.3.4 Kernel Algorithm

Algorithm 4 lists the steps of each concurrent thread to extract the centroids of a spot image. Lines 6-13 calculate the pixel group partials. To obtain the row partial, the butterfly shuffle reduction is carried out in Line 14. It is worth remarking that also rows from the other lenslets in the lenslet group are reduced at the same time. The rows in the row group are scanned in Lines 3-16 accumulating the row partials into the row group partial at each iteration. The shared reduction over the row group partials performed in Line 21 yields to the lenslet moments.

Algorithm 4 CUDA Centroids Extraction

```

1: Pixel group  $\phi_p^{l,\bar{r},r} \leftarrow$  index inferred from thread context
2: Row group partial  $\mu_{\bar{r}}^l \leftarrow 0$ .
3: for all Rows  $\phi_r^{l,\bar{r}} \in$  row group  $\phi_{\bar{r}}^l$  do
4:   Pixel group  $\phi_p^{l,\bar{r},r} \leftarrow$  global memory
5:   Pixel group partial  $\mu_p^{l,\bar{r},r} \leftarrow 0$ 
6:   for all Pixels  $\phi_p^{l,\bar{r},r,\bar{p}} \in$  pixel group  $\phi_p^{l,\bar{r},r}$  do
7:     if Pixel  $\phi_p^{l,\bar{r},r,\bar{p}}$  outside  $l$  then
8:       Pixel partial  $\mu_p^{l,\bar{r},r,\bar{p}} \leftarrow 0$ 
9:     else
10:      Pixel partial  $\mu_p^{l,\bar{r},r,\bar{p}} \leftarrow$  moments calculated from intensity value  $I(\phi_p^{l,\bar{r},r,\bar{p}})$ 
11:    end if
12:    Pixel group partial  $\mu_p^{l,\bar{r},r} \leftarrow \mu_p^{l,\bar{r},r} + \mu_p^{l,\bar{r},r,\bar{p}}$ 
13:  end for
14:  Row partial  $\mu_{\bar{r}}^{l,\bar{r}} \leftarrow$  butterfly shuffle sum reduction over pixel group partials  $\mu_p^{l,\bar{r},r}$ 
15:  Row group moment accumulator  $\sigma_{\bar{r}}^l \leftarrow \sigma_{\bar{r}}^l + \mu_{\bar{r}}^{l,\bar{r}}$ 
16: end for
17: if Pixel group index  $\bar{p} \neq 0$  then
18:   return
19: end if
20: Row group moment accumulator  $\sigma_{\bar{r}}^l \rightarrow$  shared memory  $\mu_{\bar{r}}^l$ 
21: Lenslet moments  $\mu_l \leftarrow$  shared memory sum reduction over row group partials  $\mu_{\bar{r}}^l$ 
22: if Row group index  $\bar{r} \neq 0$  then
23:   return
24: end if
25: Centroid  $c_l \leftarrow \left( \frac{m_l^{10}}{m_l^{00}}, \frac{m_l^{01}}{m_l^{00}} \right)$ 
26: Centroid  $c_l$ , lenslet moments  $\mu_l \rightarrow$  global memory

```

5.4 Experiments

In these experiments the centroid extraction is performed on images stored in memory. This choice is motivated by the fact that the exposure time and transfer delay of the image are performed before processing, and hence they have no impact on the execution time of the centroid extraction implementation.

The moment-based centroid extraction routine assumes that the captured image contains one intensity spot per lenslet, as shown in Figure 5.3. However, since all pixels must be accessed and processed (with no additional data-dependent conditions), the information contained in the spot image does not impact the execution time. A pool of images is fed to both the proposed implementation of Algorithm 4 and the sequential implementation of Algorithm 3 to test the correctness of our approach. The spot image used in the experiments is synthesized as a 8-bit region picture with each pixel having maximum intensity ($I = 255$). This means that the centroid position of each lenslet of the over-imposed lenslet grid is the center of the lenslet pixel region. The main benefit of using this kind of spot image is that the extraction grid can be arbitrarily cropped to benchmark a wide variety of use-cases.

The platform under use is the NVIDIA Jetson TX2, which integrates a 256-cores Pascal GPU, a dual-core NVIDIA Denver 2 CPU and a quad-core ARM Cortex-A57 CPU. The test-bench runs on the stock Linux distribution that comes with the NVIDIA Jetpack 4.2.1 firmware. The GPU frequency is locked at 1.3 GHz and each CPU core frequency is locked at 2 GHz.

A test run measures the time elapsed from the issue of the extraction command to the transfer of all extracted centroids, averaged over 50 executions on the same spot image. Different WFS optical aperture diameters ρ , lenslet sizes d_L and pixel sizes d_P are taken into account with the pixel region width d and pixels RoI width w_P parameters (as shown in Figure 5.3). The RoI ranges from 100×100 to 1000×1000 pixels, while the pixel region resolution ranges from 3×3 to 28×28 pixels. The result of each test run is presented in Table 5.1 where the CPU and GPU execution time t_{CPU}, t_{GPU} for a given combination of d and w_P are compared to calculate the speed-up:

$$\text{Speed-up} = \frac{t_{CPU}}{t_{GPU}}. \quad (5.13)$$

For every configuration except for the 100×100 RoI size, the GPU implementation results in a speed-up over the CPU implementation from 2 up to 13 times. Figures 5.6, 5.7, 5.8, 5.9 and 5.10 show the execution time for all parameters combinations. The algorithm execution time takes less than 1 ms for pixel region widths d larger than 5 pixels. Small pixel regions means that more lenslets fit the same RoI and hence lead to more centroids, increasing the execution time. In the case of small RoI size (Figure 5.6) the overhead latency when issuing a CUDA kernel launch (experimentally measured to be $50\mu s$ on average with an empty kernel) completely dominates the GPU execution time. The relatively high execution time variance for the CPU implementation is due to the underlying OS task scheduler behavior.

5.5 Conclusion

The NVIDIA Jetson platform is a CPU/GPU hybrid platform which, while compact and power-efficient, is powerful enough to justify its use in edge-computing and HPC. Due to its unified memory architecture, the latency introduced by copying data is avoided. Hence, images are processed as soon as they are transferred

| $w_P[\text{px}]$ | $d[\text{px}]$ | w_L^2 | $t_{CPU}[\mu\text{s}]$ | $t_{GPU}[\mu\text{s}]$ | Speed-up |
|------------------|----------------|---------|------------------------|------------------------|----------|
| 100 | 3.8 | 676 | 54 | 104 | 0.5192 |
| 100 | 11 | 81 | 44 | 59 | 0.7458 |
| 100 | 20 | 25 | 10 | 57 | 0.1754 |
| 100 | 29 | 9 | 1 | 57 | 0.0175 |
| 200 | 3.8 | 2704 | 308 | 151 | 2.0397 |
| 200 | 11 | 324 | 186 | 57 | 3.2632 |
| 200 | 20 | 100 | 158 | 57 | 2.7719 |
| 200 | 29 | 36 | 102 | 57 | 1.7895 |
| 500 | 3.8 | 17161 | 2095 | 615 | 3.4065 |
| 500 | 11 | 2025 | 1268 | 208 | 6.0962 |
| 500 | 20 | 625 | 1220 | 165 | 7.3939 |
| 500 | 29 | 289 | 1106 | 107 | 10.3364 |
| 700 | 3.8 | 33856 | 4194 | 1090 | 3.8477 |
| 700 | 11 | 3969 | 2475 | 331 | 7.4773 |
| 700 | 20 | 1225 | 2405 | 265 | 9.0755 |
| 700 | 29 | 576 | 2228 | 162 | 13.7531 |
| 1000 | 3.8 | 69169 | 8922 | 2307 | 3.8674 |
| 1000 | 11 | 8100 | 5158 | 631 | 8.1743 |
| 1000 | 20 | 2500 | 4266 | 481 | 8.8690 |
| 1000 | 29 | 1156 | 3210 | 317 | 10.1262 |

Table 5.1: Execution times of CPU and GPU (t_{CPU}, t_{GPU}) and relative speed-up for the configurations of RoI width w_P and pixel region width d , along with the number of extracted centroids w_L^2 .

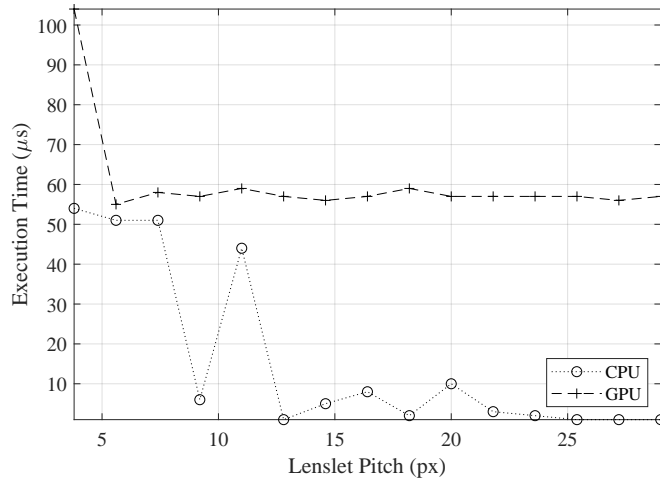


Fig. 5.6: Centroid extraction execution times for a 100×100 RoI of the spot image.

from the camera. The Jetson platform is an alternative to FPGA-based smart cameras, with the advantage of being easier to program and more flexible.

In small-scale AO systems, the SH-WFS is implemented on CPU or FPGA architectures. However, despite being low-latency, CPU solutions are not portable

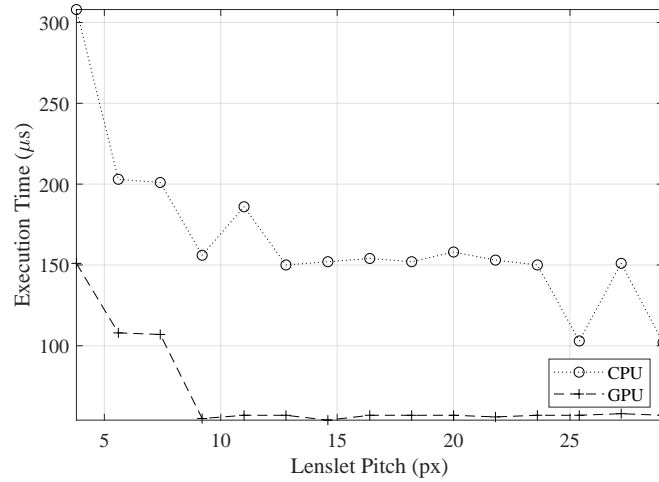


Fig. 5.7: Centroid extraction execution times for a 200×200 ROI of the spot image.

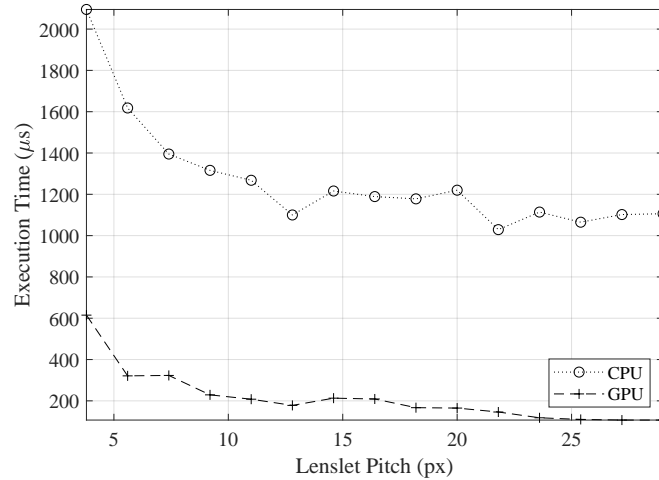


Fig. 5.8: Centroid extraction execution times for a 500×500 ROI of the spot image.

and FPGA design leads to long development time. The experimental results carried on the Jetson CPU/GPU platform show that the time required for the centroid extraction is less than 1 ms given a pixel region width larger than 5 pixels, and hence compatible with the AO closed-loop latency constraint.

The proposed approach suggests that an embedded GPU architecture is a valid alternative to FPGA-based SH-WFS solutions. The parallel capabilities of the device can be leveraged to develop advanced wavefront reconstruction schemes, e.g., extended source and high dynamics sensing. Furthermore, the spatio-temporal dynamics of atmospheric aberrations can be learned and predicted by on-board machine learning algorithms, routinely updating the controller parameters. Then, the

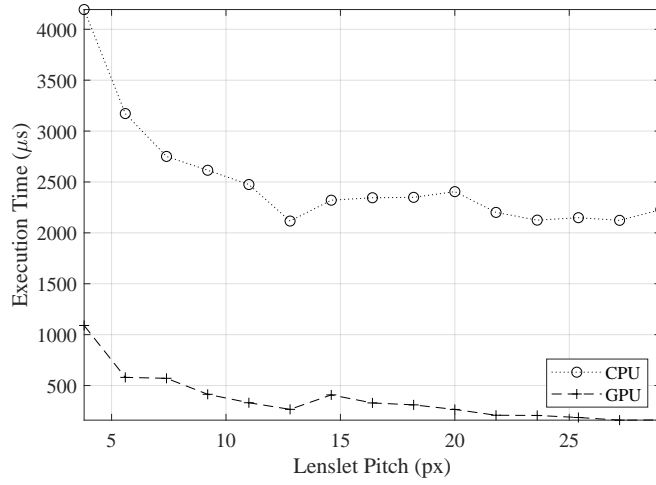


Fig. 5.9: Centroid extraction execution times for a 700×700 RoI of the spot image.

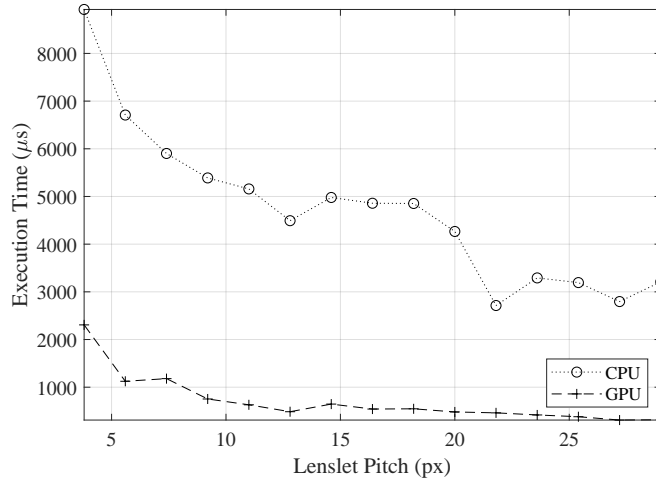


Fig. 5.10: Centroid extraction execution times for a 1000×1000 RoI of the spot image.

full AO control loop can be implemented with state-of-the-art control techniques like modal control and Predictive/Optimal Control.

By substituting the wavefront sensor implementation of the PhotonLoop software controller with the proposed hybrid GPU/CPU approach, the atmospheric turbulence correction performance results shown in Chapter 4 can be improved.

State of the Art on AO Model-Based Control Strategies

Summary. Optimal model-based control strategies are one of the most powerful results in control theory. Given a comprehensive model of the system, the optimal regulator is synthesized automatically. The optimization problem consists of an objective (cost) function, a target goal and a set of constraints. In this chapter we will review the Kalman filter, Linear Quadratic (LQ) control, \mathcal{H}_2 control and Model Predictive Control (MPC). Additionally, the Zernike modal-space formulation of the SCAO system is given to provide a unified state-space framework between the deterministic and stochastic signals into play.

6.1 Introduction

State-of-the-art control architectures currently implemented in AO systems are mostly based on classical Proportional-Integral (PI) controllers [70, 9], whose control law is composed of two terms: one is proportional to the measured residual phase, the other to its integral. While straightforward, the PI controller is sub-optimal and hard to tune for large MIMO systems.

More advanced controllers can be designed by exploiting the AO system model:

- The *Kalman filter* controller estimates the state variable of a system given its past state. Therefore, the noise component of the input can be filtered out before the integrator control law. By using a state-space estimator such as the Kalman filter, more complex AO configurations such as the multiple WFS present in MCAO systems can be handled [104]. Moreover, it can be conveniently designed into modal space where atmospheric turbulence are described mode-by-mode. By predicting one-step ahead the turbulent process, the AO delay error σ_{delay}^2 is mitigated [90];
- The *Linear-Quadratic Gaussian (LQG)* controller [2] was first proposed in AO in [84] and then followed by many others [27, 76, 63, 64, 54]. LQG control uses the system model to estimate the control objective that is then optimized to achieve the desired closed loop performance. The mirror dynamics are natively accounted for by describing them into the AO system model [25]. Parasitic vibrations of the AO system that degrade larger telescope aperture can be elegantly formalized as well [77, 86]. ;
- The \mathcal{H}_2 controller [126, 109, 44] is an alternative formulation to the LQG controller. It adds *virtual performance indexes* that make for a better interpretation of the weights to be applied into the cost function. This way, tuning complex MCAO systems is easier. A variant of the \mathcal{H}_2 controller is the \mathcal{H}_∞ [108, 1, 118] that prioritize robustness over performance ;
- The *Model Predictive Control* [91, 52, 51, 37] is a more general model-based framework in which the time horizon is limited and the signals are constrained. In the AO field, it can mitigate the non-linearities deriving from saturating actuators.

To overcome the heterogeneous nature of the signals (continuous wavefront phase and discrete measurements), the AO system is often formulated through a frequentist approach such as using the Zernike modes.

6.2 System Models in Modal Space

The components of the SCAO system can be modeled as equivalent discrete-time LTI systems with sampling time T_s . Each system is then partitioned into N_w independent SISO systems, one for each mode of the Zernike basis. WFS and DM belong to the *deterministic* subsystem, whereas the atmospheric turbulence to the *stochastic* subsystem.

6.2.1 Deterministic subsystem

A common discrete-time approximation of the DM dynamics is a d -samples delay system

$$\text{DM}(z) = \Pi_{u2y} z^{-d}, \quad (6.1)$$

where the matrix $\Pi_{u2y} \in \mathbb{R}^{N_u \times 2N_y}$ is the so-called *interaction matrix* mapping commands into slopes. It can be estimated via a so-called calibration phase measuring the DM response to each actuator motion (see Chapter 3). This simple model is accurate enough to describe the behavior of small DM. Instead, should the time constants of the DM become comparable to the (inverse of) desired control bandwidth, the DM behavior could be estimated e.g. by subspace identification methods [22].

Using the modal projector and the pseudo-inverse of the interaction matrix $\Pi_{c2s}^\dagger \in \mathbb{R}^{2N_y \times N_u}$ defined in Chapters 2 and 3, respectively, the DM model in (6.1) becomes

$$\mathbf{G}_c(z) = \Pi_{w2y}^\dagger \text{DM}(z) \Pi_{u2y}^\dagger \Pi_{w2y} \quad (6.2)$$

$$= I_{N_w} z^{-d}, \quad (6.3)$$

and therefore $\mathbf{G}_c(z)$ is partitioned into a set of N_w independent systems corresponding to the N_w modal terms:

$$\mathbf{G}_c(z) = \left\{ \mathbf{G}_c^{(1)}(z) \dots \mathbf{G}_c^{(N_w)}(z) \right\}. \quad (6.4)$$

The j -th state-space realization¹ is

$$\mathbf{G}_c^{(j)} : z^{-d} w_u^{(j)}(k) \mapsto w_c^{(j)}(k) \quad (6.7)$$

$$= \left[\begin{array}{c|c} A_c^{(j)} & B_c^{(j)} \\ \hline C_c^{(j)} & D_c^{(j)} \end{array} \right] : w_u^{(j)} \mapsto w_c^{(j)}, \quad (6.8)$$

with the j -th modal command $w_u^{(j)}(k) \in \mathbb{R}$ as input and the modal command $w_c^{(j)}(k) \in \mathbb{R}$ as output.

The Shack-Hartmann WFS is modeled as a unitary delay accounting for the exposure time (i.e. T_s). The measurements are affected by an additive noise which variance is related to the source magnitude [117]. The Gaussian distribution is a good approximation of such a noise.

¹ The discrete state-space model

$$\mathbf{G} : \begin{cases} \mathbf{x}(k+1) & = A\mathbf{x}(k) + B\mathbf{u}(k) \\ \mathbf{y}(k) & = C\mathbf{x}(k) + D\mathbf{u}(k) \end{cases} \quad (6.5)$$

is written in compact form as

$$\mathbf{G} : \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] : \mathbf{u} \mapsto \mathbf{y}. \quad (6.6)$$

The WFS transfer function in modal basis $\mathbf{G}_s(z)$ is partitioned into a set of N_w SISO systems

$$\mathbf{G}_s(z) = \left\{ \mathbf{G}_s^{(1)}(z) \dots \mathbf{G}_s^{(N_w)}(z) \right\}, \quad (6.9)$$

whose j -th state-space realization is

$$\mathbf{G}_s^{(j)} : z^{-1}w^{(j)}(k) + \sqrt{\Sigma_\eta^{(j)}}\eta^{(j)}(k) \mapsto w_s^{(j)}(k) \quad (6.10)$$

$$= \left[\begin{array}{c|c} A_s^{(j)} & B_{sy}^{(j)} \\ \hline C_s^{(j)} & \mathbf{0} \end{array} \right] : \begin{bmatrix} w^{(j)} \\ \eta^{(j)} \end{bmatrix} \mapsto w_s^{(j)} \quad (6.11)$$

with $w^{(j)}(k), w_s^{(j)}(k) \in \mathbb{R}$ the residual and measured mode coefficients, respectively, and $\eta^{(j)}(k) \in \mathbb{R}$ a zero mean Gaussian noise with covariance $\Sigma_\eta = I$. The matrix $D_{s\eta}^{(j)}$ is the square root of the j -th diagonal element $\Sigma_\eta^{(j)}$ of the variance Σ_η defined as

$$\Sigma_\eta = \Pi_{w2y}^\dagger \sigma_\eta^2 \Pi_{w2y}^{\dagger\top} \in \mathbb{R}^{N_w \times N_w}. \quad (6.12)$$

The covariance matrix Σ_η can also be calculated from the measurement errors $\eta^{(j)}$ estimated by the PSD analysis of the Zernike time series [29].

6.2.2 Stochastic subsystem

Following from Section 2.4.1, the atmospheric turbulence for the j -th modal term is modeled as a second-order low-pass filter in continuous time

$$\mathbf{H}_a^{(j)}(s) = \frac{\alpha_a^{(j)}}{\left(1 + \frac{1}{2\pi f_a^{(j)}}s\right)^2}, \quad (6.13)$$

with cutoff frequency $f_a^{(j)} \in \mathbb{R}$ and gain $\alpha_a^{(j)} \in \mathbb{R}$. Its discretization is given by

$$\mathbf{G}_a^{(j)} = \left[\begin{array}{c|c} A_a^{(j)} & B_a^{(j)} \\ \hline C_a^{(j)} & D_a^{(j)} \end{array} \right] : v^{(j)} \mapsto w_a^{(j)}, \quad (6.14)$$

where $v^{(j)}(s)$ is a zero-mean Gaussian white noise with covariance $\Sigma_v = I$. Such transfer function fit the Power Spectral Density (PSD) of the j -th measured time series. It is equivalent to the discrete-time AR2 process commonly used to approximate the turbulence dynamics while keeping complexity as low as possible [116].

6.2.3 Full System Model

The interconnection of the atmospheric turbulence model $\mathbf{G}_a^{(j)}(z)$ and the AO plant models $\mathbf{G}_c^{(j)}(z)$ and $\mathbf{G}_s^{(j)}(z)$, yields the full open-loop state-space system relative to the j -th modal term

$$\mathbf{G}_f^{(j)} = \left[\begin{array}{c|c|c|c} A_f^{(j)} & B_{fv}^{(j)} & B_{fu}^{(j)} & \mathbf{0} \\ \hline C_f^{(j)} & \mathbf{0} & \mathbf{0} & D_{f\eta}^{(j)} \end{array} \right] : \begin{bmatrix} v^{(j)} \\ w_u^{(j)} \\ \eta^{(j)} \end{bmatrix} \mapsto w_s^{(j)} \quad (6.15)$$

$$= \left[\begin{array}{ccc|ccc} A_a^{(j)} & \mathbf{0} & \mathbf{0} & B_a^{(j)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_c^{(j)} & \mathbf{0} & \mathbf{0} & B_c^{(j)} & \mathbf{0} \\ \hline B_{sy}^{(j)} C_a^{(j)} & -B_{sy}^{(j)} C_c^{(j)} & A_s^{(j)} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & C_s^{(j)} & \mathbf{0} & \mathbf{0} & D_{s\eta}^{(j)} \end{array} \right] : \begin{bmatrix} v^{(j)} \\ w_u^{(j)} \\ \eta^{(j)} \end{bmatrix} \mapsto w_s^{(j)}, \quad (6.16)$$

as shown in the block diagram of Figure 6.1

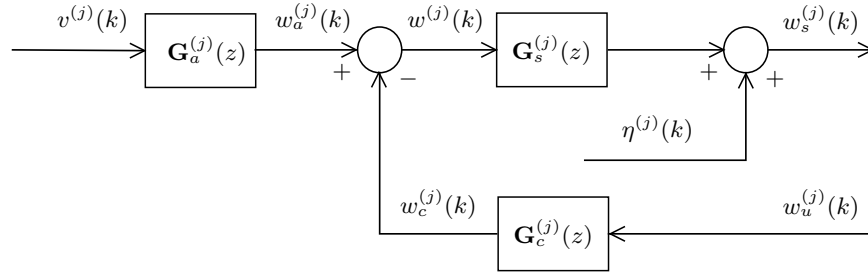


Fig. 6.1: Block diagram of the open-loop AO system for the j -th modal term.

Let

$$\mathbf{w}_s(k) = \left[w_s^{(1)}(k) \dots w_s^{(N_w)}(k) \right]^\top; \quad (6.17)$$

$$\mathbf{w}_u(k) = \left[w_u^{(1)}(k) \dots w_u^{(N_w)}(k) \right]^\top, \quad (6.18)$$

be the vectorization of all N_w measured and commanded modal terms $w_s^{(j)}(k)$ and $w_u^{(j)}(k)$ of the full modal system $\mathbf{G}_f^{(j)}(z)$. Using the projection from modal terms into slopes Π_{w2y} and the DM interaction matrix Π_{u2y} , the slope measurements $\mathbf{y}(k)$ and DM commands $\mathbf{u}(k)$ of the SCAO system shown in Chapter 2 are completely described by:

$$\mathbf{y}(k) = \Pi_{w2y} \mathbf{w}_s(k); \quad (6.19)$$

$$\mathbf{u}(k) = \Pi_{u2y}^\dagger \Pi_{w2y} \mathbf{w}_u(k). \quad (6.20)$$

These equations describe how the input-output signals of the N_w systems in the modal basis are mapped into the input-output signals of the original AO plant.

6.2.4 Modal Phase Representation Error

The main advantage of the Zernike base to represent the atmospheric turbulence is the straightforward interpretation of its modes in terms of optical aberrations, [82, 124]. Now that all the elements of the AO system are formalized into the modal

space framework by using the relationships from (6.19) and (6.20), the impact that the wavefront representation error has onto the controller performance can be analyzed by following Kúlcser et al. [54] and the references therein.

The WFS samples the incoming residual wavefront phase $\phi_e(t, \mathbf{r})$ into the discrete measurement vector $\mathbf{y}(k)$. Therefore, the turbulence can be partitioned into

$$\phi_e = \phi_e^{\mathcal{O}} + \phi_e^{\mathcal{O}\perp}, \quad (6.21)$$

where $\phi_e^{\mathcal{O}}$ is observable turbulence from the WFS (i.e. that can be reconstructed from $\mathbf{y}(k)$) and $\phi_e^{\mathcal{O}\perp}$ is the unobservable turbulence (i.e. the waffle mode of the DM).

The calibrated interaction matrix Π_{u2y} maps the WFS space into the DM space and, as such, is subject to the spatial sampling of the WFS and the fitting error of the DM. Such errors trickle down into the achievable correction phase produced by the DM:

$$\phi_c = \phi_c^{\mathcal{C}} + \phi_c^{\mathcal{C}\perp}, \quad (6.22)$$

where $\phi_c^{\mathcal{C}}$ is the wavefront corrected by the DM commands \mathbf{u} and $\phi_c^{\mathcal{C}\perp}$ is the component that can not be corrected. In system theory terminology, (6.21) refers to the observable/unobservable subsystems whereas (6.22) refers to the reachable/unreachable subsystems.

In the modal framework truncated to N_w terms, the projector Π_{y2w} that maps phase measurement \mathbf{y}_a into modes \mathbf{w}_a is a further approximation of the wavefront phase ϕ_a . Modal controllers are thus only able to compensate for the part of atmosphere approximated by Π_{y2w} . Other basis can be considered in relation to the particular AO problem (Section 2.2).

The basis using the system modes of Π_{u2y} discussed in Section 3.4 is implicitly built taking into account the observable and reachable subsystem.

6.3 Kalman Filter

State-feedback controllers need the state variable $\mathbf{x}(k)$ to calculate the command $\mathbf{u}(k)$. Let $\hat{\mathbf{x}}(k|h)$ be the state estimate at time sample k based on past measurements till time h $\mathbf{y}(0), \mathbf{y}(1), \dots, \mathbf{y}(h)$. The framework used to compute the optimal estimator/predictor is the Kalman filtering. Let

$$\mathbf{G} = \left[\begin{array}{c|c} A & B \quad I \quad \mathbf{0} \\ \hline C & D \quad \mathbf{0} \quad I \end{array} \right] : \begin{bmatrix} \mathbf{u}(k) \\ \mathbf{v}(k) \\ \boldsymbol{\eta}(k) \end{bmatrix} \mapsto \mathbf{y}(k), \quad (6.23)$$

be a state-space realization of a generic MIMO LTI system where $\mathbf{v}(k) \in \mathbb{R}^N$ is the model error and $\boldsymbol{\eta}(k) \in \mathbb{R}^M$ is the measurement additive noise. $\mathbf{v}(k)$ and $\boldsymbol{\eta}(k)$ are modeled as uncorrelated Gaussian zero-mean white noise signals with covariance $\Sigma_{\mathbf{v}}$ and $\Sigma_{\boldsymbol{\eta}}$, respectively. A more general formulation can be found in Anderson-Moore [4].

Under those hypotheses, the optimal estimator of the state $\hat{\mathbf{x}}(k+1|k+1)$ is the conditional expectation of the state \mathbf{x} at time $k+1$ on the past and present measurements

$$\hat{\mathbf{x}}(k+1|k+1) = \mathbb{E}[\mathbf{x}(k+1)|\mathbf{y}(k+1), \mathbf{y}(k), \dots, \mathbf{y}(0)]. \quad (6.24)$$

The state estimate is calculated in two steps:

1. (Prediction Step) (6.25)

$$\begin{cases} \hat{\mathbf{x}}(k+1|k) = A\hat{\mathbf{x}}(k|k) \\ Y(k+1|k) = AY(k|k)A^\top + \Sigma_v; \end{cases}$$

2. (Filtering Step) (6.26)

$$\begin{cases} \hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \\ \quad K(k+1)(\mathbf{y}(k+1) - C\hat{\mathbf{x}}(k+1|k)) \\ Y(k+1|k+1) = Y(k+1|k) - \\ \quad Y(k+1|k)C^\top (CY(k+1|k)C^\top + \Sigma_\eta)^{-1} CY(k+1|k), \end{cases}$$

where $Y(k+1|k)$ is the solution of the Discrete Riccati Equation

$$Y(k+1|k) = AY(k|k-1)A^\top + \quad (6.27)$$

$$\Sigma_v - AY(k|k-1)C^\top (CY(k|k-1)C^\top + \Sigma_\eta)^{-1} CY(k|k-1)A^\top \quad (6.28)$$

and

$$K(k+1) = Y(k+1|k)C^\top (CY(k+1|k)C^\top + \Sigma_\eta)^{-1} \quad (6.29)$$

is the Kalman gain for the predictor. The steady-state gain K_∞ can be calculated as the solution of the Discrete Algebraic Riccati Equation (DARE):

$$Y_\infty = AY_\infty A^\top - AY_\infty C^\top (CY_\infty C^\top + \Sigma_\eta)^{-1} CY_\infty A^\top + \Sigma_v, \quad (6.30)$$

which leads to the state estimator

$$\hat{\mathbf{x}}(k+1|k+1) = A\hat{\mathbf{x}}(k|k) + K_\infty(\mathbf{y}(k+1) - CA\hat{\mathbf{x}}(k|k)). \quad (6.31)$$

with

$$K_\infty = Y_\infty C^\top (CY_\infty C^\top + \Sigma_\eta)^{-1}. \quad (6.32)$$

The solution of Y_∞ is unique, positive definite, and stabilizing if and only if (A, C) is detectable and $(A, \Sigma_v^{1/2})$ is stabilizing [4].

6.3.1 Kalman Filter in AO

The typical use of the Kalman estimator in the SCAO configuration is to filter out the WFS measurement noise $\boldsymbol{\eta}(k)$ while predicting the one-step ahead atmospheric turbulence $\hat{\mathbf{y}}_a(k+1|k)$, as shown in Fig. 6.2. The most convenient atmospheric turbulence model is given in Zernike terms as described in Section 6.2.

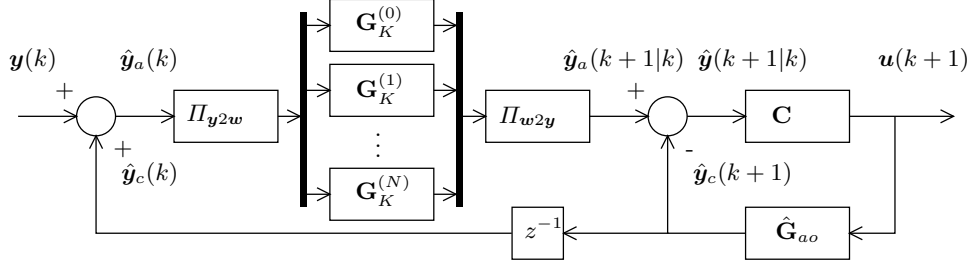


Fig. 6.2: SCAO Kalman prediction diagram. The Kalman filter bank \mathbf{G}_K predicts the one-step ahead atmospheric turbulence $\mathbf{y}_a(k+1)$ to compensate for the measurement unitary delay. The estimated AO plant $\hat{\mathbf{G}}_{ao}$ response $\hat{\mathbf{y}}_c(k+1)$ infers the residual $\hat{\mathbf{y}}(k+1|k)$, which is used by the controller \mathbf{C} , and the measured atmosphere $\mathbf{y}_a(k)$.

The vector of the coefficients of Zernike modes

$$\mathbf{w}_a(k) = \left[w_a^{(0)}(k) \ w_a^{(1)}(k) \ \dots \ w_a^{(N)}(k) \right]^\top \quad (6.33)$$

is obtained by projecting the atmospheric turbulence and the AO correction due to the DM into the modal space

$$\mathbf{w}_a(k) = \Pi_{\mathbf{y}2\mathbf{w}}\mathbf{y}(k) + \Pi_{\mathbf{y}2\mathbf{w}}\mathbf{y}_c(k). \quad (6.34)$$

If the DM is modeled as a simple delay, we end up with

$$\mathbf{w}_a(k) = \Pi_{\mathbf{y}2\mathbf{w}}\mathbf{y}(k) + \Pi_{\mathbf{y}2\mathbf{w}}\Pi_{\mathbf{u}2\mathbf{y}}\mathbf{u}(k-1). \quad (6.35)$$

The j -th Kalman predictor $\mathbf{G}_K^{(j)}(z)$ is given by

$$\mathbf{G}_K^{(j)}(z) = \left[\begin{array}{c|c} A_a^{(j)} - K_\infty^{(j)} C_a^{(j)} & K_\infty^{(j)} \\ \hline C_a^{(j)} & \mathbf{0} \end{array} \right] w_a^{(j)}(k) \mapsto w_a^{(j)}(k+1|k) \quad (6.36)$$

where $K_\infty^{(j)}$ is the j -th steady-state Kalman gain calculated as in Eq. (6.30).

In the SCAO configuration, the atmospheric turbulence $\mathbf{y}_a(k)$ cannot be directly measured. Instead, it has to be inferred from the current residual measurement $\mathbf{y}(k)$ and the previous estimated response $\hat{\mathbf{y}}_c(k)$. The Pseudo-Open Loop (POL) is:

$$\hat{\mathbf{y}}_a(k) = \mathbf{y}(k) + \hat{\mathbf{y}}_c(k) \quad (6.37)$$

$$= \mathbf{y}(k) + \left[z^{-1} \hat{\mathbf{G}}_{ao} \mathbf{C} \right] \mathbf{u}(k). \quad (6.38)$$

The same consideration holds when retrieving the predicted measurement $\hat{\mathbf{y}}(k+1|k)$:

$$\hat{\mathbf{y}}(k+1|k) = \hat{\mathbf{y}}_a(k+1|k) + \left[\hat{\mathbf{G}}_{ao}(z) \mathbf{C}(z) \right] \mathbf{u}(k+1), \quad (6.39)$$

where the unitary delay is suppressed and the DM is delayed by one sample.

The choice of the controller \mathbf{C} is independent of the design of the Kalman filter due to the separation principle. A common choice is to implement the controller as the traditional output-feedback integral compensator \mathbf{G}_I introduced in (2.48), or as a state-feedback, model-based control strategy.

6.4 Linear Quadratic Control

The Linear Quadratic (LQ) control is a strategy that is optimal with regard to a performance criterion to be minimized or maximized [31][54][55][98][86][28]. The generic LQ performance index on an interval $[0, K]$ is formulated as

$$J_K(\mathbf{u}, \mathbf{x}_0) = \sum_{k=0}^K [\mathbf{x}^\top(k)Q\mathbf{x}(k) + \mathbf{u}^\top(k)R\mathbf{u}(k)] + \mathbf{x}^\top(K)S\mathbf{x}(K) \quad (6.40)$$

where Q and S are symmetric positive semi-definite matrices, R is a symmetric positive definite matrix, $\mathbf{u}(k)$ is the vector command at time k and \mathbf{x}_0 is the initial state of the LTI system.

J_K has a global minimum and can be interpreted from an energetic point of view:

- $Q \in \mathbb{R}^{n \times n}$ weights the distance of the states from the origin;
- $R \in \mathbb{R}^{m \times m}$ weights the control effort;
- $S \in \mathbb{R}^{n \times n}$ weights the distance of the final state to the origin.

The optimal control at time k which minimizes J is given by

$$\mathbf{u}(k) = - (R + B^\top X(k+1)B)^{-1} B^\top X(k+1)A\mathbf{x}(k) \quad (6.41)$$

$$= F(k)\mathbf{x}(k), \quad (6.42)$$

where $X(k)$ solves the Discrete Riccati Equation

$$\begin{cases} X(k) &= Q + A^\top X(k+1)A - A^\top X(k+1)A (R + B^\top X(k+1)B)^{-1} A^\top X(k+1)A \\ X(K) &= S \end{cases} \quad (6.43)$$

Such equation is derived by solving the associated linear quadratic problem [3].

The corresponding infinite horizon case with performance index

$$J_\infty(\mathbf{u}, x_0) = \lim_{K \rightarrow +\infty} \sum_{k=0}^K [\mathbf{x}^\top(k)Q\mathbf{x}(k) + \mathbf{u}^\top(k)R\mathbf{u}(k)], \quad (6.44)$$

is the constant state feedback matrix F_∞

$$F_\infty = - (R + B^\top X_\infty B)^{-1} B^\top \quad (6.45)$$

where X_∞ is the unique positive definite solution of the Discrete Algebraic Riccati Equation (DARE)

$$X_\infty = Q + A^\top X_\infty A - A^\top X_\infty B (R + B^\top X_\infty B)^{-1} B^\top X_\infty A. \quad (6.46)$$

The optimal control with the steady-state LQ feedback matrix is

$$\mathbf{u}(k) = F_\infty \mathbf{x}(k), \quad (6.47)$$

and the controlled system is asymptotically stable if the pair (A, B) is stabilizing and the pair $(A, Q^{1/2})$ is detectable.

Computing the optimal command requires the knowledge of the state $\mathbf{x}(k)$, which is usually estimated with a Kalman filter. The combination of a Linear Quadratic controller and a Kalman estimator, both optimal with respect to their objectives, is called Linear Quadratic Gaussian (LQG) regulator [3]. Thanks to the separation principle, they can be designed independently.

6.4.1 Linear Quadratic Control in AO

Let \mathbf{G}_f be the LTI state-space system

$$\mathbf{G}_f = \left[\begin{array}{c|c} A_f & B_f \\ \hline C_f & \mathbf{0} \end{array} \right] : \left[\begin{array}{c} \mathbf{u}(k) \\ \mathbf{v}(k) \end{array} \right] \mapsto \mathbf{y}(k) \quad (6.48)$$

of the atmosphere and AO plant open loop connection, where

$$A_f = \begin{bmatrix} A_a & \mathbf{0} \\ \mathbf{0} & A_{ao} \end{bmatrix}, \quad (6.49)$$

$$B_f = \begin{bmatrix} \mathbf{0} & B_a \\ B_{ao} & \mathbf{0} \end{bmatrix}, \quad (6.50)$$

$$C_f = [C_a \ -C_{DM}]. \quad (6.51)$$

Assuming a Kolmogorov model for the turbulence, the ergodic theorem holds (expectation is equal to the sample mean) and then the residual phase $\mathbf{y}(k)$ has zero mean

$$\begin{aligned} \mathbb{E}[\mathbf{y}(k)] &= \lim_{K \rightarrow +\infty} \frac{1}{K} \sum_{k=0}^K \mathbf{y}(k) \\ &= \mathbf{0}, \end{aligned} \quad (6.52)$$

and variance

$$\begin{aligned} \mathbb{E}[(\mathbf{y}(k) - \mathbb{E}(\mathbf{y}(k)))^2] &= \mathbb{E}[\mathbf{y}^2(k)] \\ &= \lim_{K \rightarrow +\infty} \frac{1}{K} \sum_{k=0}^K \mathbf{y}^2(k) \\ &= \Sigma_{\mathbf{y}}, \end{aligned} \quad (6.53)$$

using the sample mean and the sample variance.

The residual wavefront variance in (6.53) is a quadratic function and is directly proportional to the continuous residual phase $\phi_e(t, \mathbf{r})$. Therefore, it can be used

as the performance cost as proposed by many authors, e.g. [54, 63]. The state variable $\mathbf{x}_f(k)$ of the full AO system \mathbf{G}_f is defined as:

$$\mathbf{x}_f(k) = \begin{bmatrix} \mathbf{x}_a(k) \\ \mathbf{x}_{ao}(k) \end{bmatrix}, \quad (6.54)$$

and the performance cost is formulated as

$$\begin{aligned} J(\mathbf{u}(k)) &= \mathbb{E} [\mathbf{x}_f^\top(k) Q \mathbf{x}_f(k) + \mathbf{u}^\top(k) R \mathbf{u}(k)] & (6.55) \\ &= \lim_{K \rightarrow +\infty} \frac{1}{K} \sum_{k=0}^K (\mathbf{x}_f^\top(k) Q \mathbf{x}_f(k) + \mathbf{u}^\top(k) R \mathbf{u}(k)) \\ &= \lim_{K \rightarrow +\infty} \frac{1}{K} \sum_{k=0}^K (\mathbf{y}_a^\top(k) Q \mathbf{y}_a(k) + \mathbf{y}_c^\top(k) Q \mathbf{y}_c(k) \\ &\quad - 2\mathbf{y}_c^\top(k) Q \mathbf{y}_a(k) + \mathbf{u}^\top(k) R \mathbf{u}(k)) \\ &= \lim_{K \rightarrow +\infty} \frac{1}{K} \sum_{k=0}^K ((C_a \mathbf{x}_a(k))^\top Q C_a \mathbf{x}_a(k) + (C_{ao} \mathbf{x}_{ao}(k))^\top Q C_{ao} \mathbf{x}_{ao}(k) \\ &\quad - 2(C_a \mathbf{x}_{ao}(k))^\top Q C_{ao} \mathbf{y}_a(k) + \mathbf{u}^\top(k) R \mathbf{u}(k)), \end{aligned}$$

showing the relation between atmosphere, correction and command vectors. Q and R weight the residual correction and control effort, respectively, and are chosen such that the matrix

$$\begin{bmatrix} Q & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix} \quad (6.56)$$

is positive definite.

Assuming null initial condition for the DM, the performance criterion is function of the input command $\mathbf{u}(k)$ only. The discrete-time state-space system of the AO plant $\mathbf{G}_{ao}(z)$ is formulated in Eq. (2.50) and reduces to

$$\mathbf{y}_c(k) = \Pi_{u2y} \mathbf{u}(k-1) \quad (6.57)$$

in case of fast DM. The input $\mathbf{u}(k)$ minimizes the cost (6.55) and is thus calculated as

$$\mathbf{u}(k) = \arg \min_{\mathbf{u}} J_\infty(\mathbf{u}) \quad (6.58)$$

$$= F_\infty \mathbf{x}_f(k), \quad (6.59)$$

with F_∞ being the constant steady-state LQ feedback matrix given in Eq. (6.45).

Combining the LQ control strategy with the Kalman filter yields to the LQG controller. The control system $\mathbf{C}(z)$ is synthesized as

$$\mathbf{C}(z) = \left[\begin{array}{c|c} A_f - B_f F_\infty - K_\infty C_f & K_\infty \\ \hline F_\infty & \mathbf{0} \end{array} \right] : \mathbf{y}(k) \mapsto \mathbf{u}(k), \quad (6.60)$$

where K_∞ is the steady-state Kalman matrix gain calculated in Eq. (6.32) and F_∞ is the constant state-feedback matrix gain.

6.5 \mathcal{H}_2 Control

A drawback of the LQ control strategy is that the matrix weights Q and R often lack physical meaning, which plays a crucial role in tuning the system response. To overcome this limitation, the performance criterion can be formulated in terms of the \mathcal{H}_2 system norm, or the \mathcal{H}_∞ norm tailored for robust control and/or performance.

Consider a state-space system of the form

$$\mathbf{G}_{\mathbf{u}\mathbf{v}\mathbf{2}\mathbf{y}} = \left[\begin{array}{c|cc} A & B_{\mathbf{u}} & B_{\mathbf{v}} \\ \hline C & D_{\mathbf{u}} & D_{\mathbf{v}} \end{array} \right] : \begin{bmatrix} \mathbf{v}(k) \\ \mathbf{u}(k) \end{bmatrix} \mapsto \mathbf{y}(k) \quad (6.61)$$

which maps input $\mathbf{u}(k)$ and process noise $\mathbf{v}(k)$ into the output $\mathbf{y}(k)$. $\mathbf{G}_{\mathbf{u}\mathbf{v}\mathbf{2}\mathbf{y}}$ can be augmented to explicit a performance variable $\mathbf{z}(k)$:

$$\mathbf{G}_{\mathbf{u}\mathbf{v}\mathbf{2}\mathbf{z}\mathbf{y}} = \begin{bmatrix} \mathbf{G}_{\mathbf{v}\mathbf{2}\mathbf{z}} & \mathbf{G}_{\mathbf{u}\mathbf{2}\mathbf{z}} \\ \mathbf{G}_{\mathbf{v}\mathbf{2}\mathbf{y}} & \mathbf{G}_{\mathbf{u}\mathbf{2}\mathbf{y}} \end{bmatrix}, \quad (6.62)$$

of which a generic realization is:

$$\mathbf{G}_{\mathbf{u}\mathbf{v}\mathbf{2}\mathbf{z}\mathbf{y}} = \left[\begin{array}{c|cc} A & B_{\mathbf{u}} & B_{\mathbf{v}} \\ \hline \begin{array}{c} C_{\mathbf{z}} \\ C \end{array} & \begin{array}{c} D_{\mathbf{u}\mathbf{2}\mathbf{z}} \\ D_{\mathbf{u}\mathbf{2}\mathbf{y}} \end{array} & \begin{array}{c} D_{\mathbf{v}\mathbf{2}\mathbf{z}} \\ D_{\mathbf{v}\mathbf{2}\mathbf{y}} \end{array} \end{array} \right] : \begin{bmatrix} \mathbf{v}(k) \\ \mathbf{u}(k) \end{bmatrix} \mapsto \begin{bmatrix} \mathbf{z}(k) \\ \mathbf{y}(k) \end{bmatrix}. \quad (6.63)$$

Consider now the feedback system $\bar{\mathbf{G}}_{\mathbf{v}\mathbf{2}\mathbf{z}}$ from the process noise $\mathbf{v}(k)$ to the performance variable $\mathbf{z}(k)$:

$$\bar{\mathbf{G}}_{\mathbf{v}\mathbf{2}\mathbf{z}} = \mathbf{G}_{\mathbf{v}\mathbf{2}\mathbf{z}} + \mathbf{G}_{\mathbf{u}\mathbf{2}\mathbf{z}} \mathbf{C} (\mathbf{I} - \mathbf{G}_{\mathbf{u}\mathbf{2}\mathbf{y}} \mathbf{C})^{-1} \mathbf{G}_{\mathbf{v}\mathbf{2}\mathbf{y}} \quad (6.64)$$

where \mathbf{C} is the controller which maps the input $\mathbf{y}(k)$ into the output $\mathbf{u}(k)$. Then, the performance cost can be stated as the \mathcal{H}_2 norm of the feedback system, function of the controller \mathbf{C} :

$$J(\mathbf{C}) = \|\bar{\mathbf{G}}_{\mathbf{v}\mathbf{2}\mathbf{z}}\|_{\mathcal{H}_2}^2. \quad (6.65)$$

Assuming $\mathbf{v}(k)$ to be a Gaussian zero-mean white noise, $J_{\mathbf{C}}$ is equivalent to

$$J(\mathbf{C}) = \lim_{K \rightarrow +\infty} \frac{1}{K} \sum_{k=0}^K [\mathbf{z}^\top(k) \mathbf{z}(k)]. \quad (6.66)$$

The following lemma provides the optimal output feedback controller $\mathbf{C}(z)$ which minimizes the performance cost $J(\mathbf{C})$.

Lemma 6.1 (State-space solution to discrete-time \mathcal{H}_2 -optimization problem [43]). *Given the state-space discrete system $\mathbf{G}_{\mathbf{u}\mathbf{v}\mathbf{2}\mathbf{z}\mathbf{y}}(z)$, and assuming that*

1. $(A, B_{\mathbf{u}})$ is stabilizable and (A, C) is detectable,
2. $D_{\mathbf{u}\mathbf{2}\mathbf{z}}^\top D_{\mathbf{u}\mathbf{2}\mathbf{z}}$ and $D_{\mathbf{v}\mathbf{2}\mathbf{y}}^\top D_{\mathbf{v}\mathbf{2}\mathbf{y}}$ are positive definite,

3. the matrices

$$\begin{bmatrix} A - \lambda I & B_u \\ C_z & D_{u2z} \end{bmatrix}, \begin{bmatrix} A - \lambda I & B_v \\ C_y & D_{v2y} \end{bmatrix} \quad (6.67)$$

have full rank for $\{\lambda \mid \lambda \in \mathbb{C}, |\lambda| = 1\}$;

then the optimal output feedback controller \mathbf{C} that minimizes the criterion $J(\mathbf{C})$ is

$$\mathbf{C} = \left[\frac{A + B_v L_0 C - B_c F - LC}{F - L_0 C} \middle| \frac{B_u L_0 - L}{L_0} \right] : \mathbf{y}(k) \mapsto \mathbf{u}(k), \quad (6.68)$$

where the matrices

$$F = (B_u^\top Y B_u + D_{u2z}^\top D_{u2z})^{-1} (B_u^\top Y A + D_{v2y}^\top C_z) \quad (6.69)$$

$$F_0 = (B_u^\top Y B_u + D_{u2z}^\top D_{u2z})^{-1} (B_u^\top Y B_v + D_{zu}^\top D_{v2z}) \quad (6.70)$$

$$L = (A X C^\top + B_v D_{v2y}^\top) (C X C^\top + D_{v2y} D_{v2y}^\top)^{-1} \quad (6.71)$$

$$L_0 = (F X C^\top + F_0 D_{v2y}^\top) (C X C^\top + D_{v2y} D_{v2y}^\top)^{-1} \quad (6.72)$$

are calculated using the unique solutions of the Discrete Algebraic Riccati Equations (DARE)

$$X = A X A^\top - (A X C^\top + B_v D_{v2y}^\top) (C X C^\top + D_{v2y} D_{v2y}^\top)^{-1} (\cdot)^\top + B_v B_v^\top \quad (6.73)$$

$$Y = A^\top Y A - (A^\top Y B_u + C_z^\top D_{u2z}) (B_u^\top Y B_u + D_{u2z}^\top D_{u2z})^{-1} (\cdot)^\top + C_z^\top C_z. \quad (6.74)$$

The \mathcal{H}_2 controller is equivalent to the one synthesized by LQG method on an infinity horizon. The proof relies on the reformulation of the performance cost. The matrices C_z and D_{u2z} can be defined as

$$C_z = \begin{bmatrix} Q^{1/2} \\ 0 \end{bmatrix}, D_{u2z} = \begin{bmatrix} 0 \\ R^{1/2} \end{bmatrix} \quad (6.75)$$

where $Q^{1/2}$ and $R^{1/2}$ are the Cholesky factorization of the LQG weights Q and R . LQG and \mathcal{H}_2 are then related by:

$$\mathbf{z}^\top(k) \mathbf{z}(k) = \mathbf{x}^\top(k) Q \mathbf{x}(k) + \mathbf{u}^\top(k) R \mathbf{u}(k) \quad (6.76)$$

$$= (C_z \mathbf{x}(k) + D_{zu} \mathbf{u}(k))^\top (C_z \mathbf{x}(k) + D_{zu} \mathbf{u}(k)). \quad (6.77)$$

6.5.1 \mathcal{H}_2 Control in AO

In the \mathcal{H}_2 framework, the block diagram for the SCAO system is built as in Fig. 6.3 [43]. The performance variable is chosen as

$$\mathbf{z}(k) = \left[\mathbf{y}(k)^\top \quad (Q^{1/2} \mathbf{u}(k))^\top \right]^\top, \quad (6.78)$$

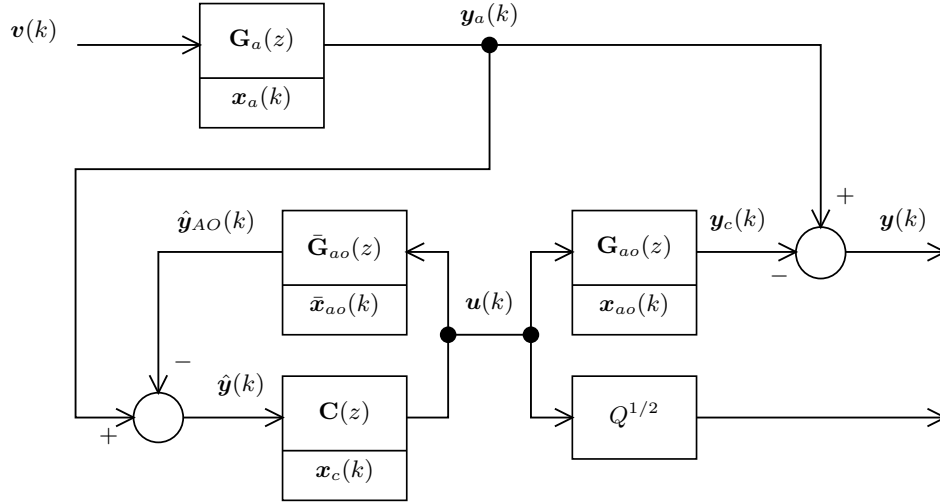


Fig. 6.3: Diagram of the open-loop control SCAO configuration in the \mathcal{H}_2 framework. The system $\bar{\mathbf{G}}_{ao}(z)$ is the replicated model of the AO plant. The signals $\mathbf{y}(k)$ and $Q^{1/2}\mathbf{u}(k)$ are components of the controlled variable $\mathbf{z}(k)$.

to include both the residual $\mathbf{y}(k)$ and command effort $\mathbf{u}(k)$, with $Q^{1/2}$ the weight matrix within the performance cost.

The overall state-space system is hence expressed as

$$\mathbf{G}_f = \left[\begin{array}{cc|cc} A_a & \mathbf{0} & B_a \Sigma_v^{1/2} & \mathbf{0} \\ \mathbf{0} & A_{ao} & \mathbf{0} & B_{ao} \\ \hline C_a & -C_{ao} & \Sigma_v^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & Q^{1/2} \\ \hline C_a & -C_{ao} & \Sigma_v^{1/2} & \mathbf{0} \end{array} \right] : \begin{bmatrix} \mathbf{v}(k) \\ \mathbf{u}(k) \end{bmatrix} \mapsto \begin{bmatrix} \mathbf{z}(k) \\ \mathbf{y}(k) \end{bmatrix}, \quad (6.79)$$

where $\Sigma_v^{1/2}$ represents the square root of the process noise variance. This means that $\mathbf{v}(k)$ is a Gaussian zero-mean unit-variance white noise.

With respect to the same assumptions made in the LQ control strategy case, the DM dynamics can be approximated by a simple delay, which leads to the following transfer functions:

$$W_{\mathbf{u}2\mathbf{z}}(z) = \begin{bmatrix} -z^{-1}C_a \\ Q^{1/2} \end{bmatrix}, \quad (6.80)$$

$$W_{\mathbf{v}2\mathbf{y}}(z) = \Sigma_v^{1/2} + C_a(zI - A_a)^{-1}B_a\Sigma_v^{1/2}. \quad (6.81)$$

Then, the analytical solution is given by the following state-space representation

43

$$\mathbf{C} = \left[\begin{array}{c|c} \tilde{A}_a + B_a(C_{ao})\tilde{C}_{ao}C_a & B_a \\ \hline \tilde{C}_{ao}C_a(\tilde{A}_a + B_aC_{ao}\tilde{C}_{ao}C_a) & \tilde{C}_{ao}C_aB_a \end{array} \right] : \mathbf{u}(k) \mapsto \mathbf{y}(k), \quad (6.82)$$

with

$$\begin{aligned}\tilde{A}_a &= A_a - B_a C_a; \\ \tilde{C}_{ao} &= (C_{ao}^\top C_{ao} + Q)^{-1} C_{ao}^\top.\end{aligned}\quad (6.83)$$

6.6 Model Predictive Control

Both LQ and \mathcal{H}_2 control strategies compute the command $\mathbf{u}(k)$ to be optimal with respect to a performance cost J . However, such strategies don't take into consideration the constraints on the state and on the command. The Model Predictive Control (MPC) aims to solve constrained problems introducing the concept of temporal-receding horizon.

The performance criterion of the MPC strategy for a state-space system $\mathbf{G}(z)$ is

$$J_{N_p, N_c}(k) = \sum_{i=0}^{N_p} \hat{\mathbf{x}}(k+i|k)^\top Q_i \hat{\mathbf{x}}(k+i|k) + \sum_{i=0}^{N_c} \hat{\mathbf{u}}(k+i|k)^\top R_i \hat{\mathbf{u}}(k+i|k), \quad (6.84)$$

where $\hat{\mathbf{x}}(k+i|k)$ and $\hat{\mathbf{u}}(k+i|k)$ are the i -steps ahead predictions of state and control, and N_p and N_c are the state and control horizons, respectively. Q_i and R_i are the weight matrices for the i -steps ahead predictions of state and control.

Let $\hat{X}_{N_p}(k)$ and $\hat{U}_{N_c}(k)$ be the matrix-vector forms of the state and command on the horizons N_p and N_c

$$\hat{X}_{N_p}(k) = \begin{bmatrix} \hat{\mathbf{x}}(k+1|k) \\ \hat{\mathbf{x}}(k+2|k) \\ \vdots \\ \hat{\mathbf{x}}(k+N_p|k) \end{bmatrix}, \quad \hat{U}_{N_c}(k) = \begin{bmatrix} \hat{\mathbf{u}}(k|k) \\ \hat{\mathbf{u}}(k+1|k) \\ \vdots \\ \hat{\mathbf{u}}(k+N_c-1|k) \end{bmatrix}, \quad (6.85)$$

and Q and R be the overall weight matrices

$$Q = \begin{bmatrix} Q_0 & & \\ & \ddots & \\ & & Q_{N_p} \end{bmatrix}, \quad R = \begin{bmatrix} R_0 & & \\ & \ddots & \\ & & R_{N_c} \end{bmatrix}. \quad (6.86)$$

Under the assumption that the state at time k is measurable ($i=0$), we have $\hat{\mathbf{x}}(k|k) = \mathbf{x}(k)$.

The performance cost (6.84) can be re-written equivalently as

$$J_{N_p, N_c}(k) = \mathbf{x}(k)^\top Q_0 \mathbf{x}(k) + \hat{X}_{N_p}(k)^\top Q \hat{X}_{N_p}(k) + \hat{U}_{N_c}(k)^\top R \hat{U}_{N_c}(k). \quad (6.87)$$

The evolution of the vector $\hat{X}_{N_p}(k)$ given the control vector \hat{U}_{N_c} is expressed as

$$\hat{X}_{N_p}(k) = \Gamma \hat{U}_{N_c}(k) + \Omega \mathbf{x}(k)^\top \quad (6.88)$$

where the matrices Γ and Ω define the forced and free evolution

$$\Gamma = \begin{bmatrix} B & \dots & \mathbf{0} \\ AB & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ A^{N_p-1}B & \dots & A^{N_p-N_c}B \end{bmatrix}, \quad \Omega = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^{N_p} \end{bmatrix}. \quad (6.89)$$

Substituting Eq. (6.88) in the performance criterion yields

$$\begin{aligned} J_{N_p, N_c}(k) &= \mathbf{x}(k)^\top Q_0 \mathbf{x}(k) + \hat{U}_{N_c}(k)^\top (\Gamma^\top Q \Gamma + R) \hat{U}_{N_c}(k) + 2\hat{U}_{N_c}(k)^\top \Gamma^\top Q \Omega \mathbf{x}(k) \\ &= \mathbf{x}(k)^\top Q_0 \mathbf{x}(k) + \hat{U}_{N_c}(k)^\top \mathbb{H} \hat{U}_{N_c}(k) + 2\hat{U}_{N_c}(k)^\top \mathbb{F} \mathbf{x}(k) \end{aligned} \quad (6.90)$$

where the first term is not affected by the commands $\hat{U}_{N_c}(k)$ and

$$\mathbb{H} \triangleq \Gamma^\top Q \Gamma + R, \quad \mathbb{F} \triangleq \Gamma^\top Q \Omega. \quad (6.91)$$

Typical MPC constraints are upper and lower bounds on the range of admissible values for the state and command variables for each time sample. They are expressed as matrix inequalities

$$\underline{U} \leq \hat{U}_{N_c}(k) \leq \bar{U} \quad (6.92)$$

$$\underline{X} - \Omega \hat{\mathbf{x}}(k|k)^\top \leq \Gamma \hat{U}_{N_p}(k) \leq \bar{X} - \Omega \hat{\mathbf{x}}(k|k)^\top, \quad (6.93)$$

or in compact form as

$$\Xi \hat{U}_{N_c}(k) \leq \Lambda \quad (6.94)$$

where Ξ incorporates all the inequality relationships, and Λ contains the minimum and maximum values.

The optimal input series is then obtained by minimizing the cost function with constraints:

$$\hat{U}_{N_c}(k) = \arg \min_{\hat{U}} \left(\hat{U}_{N_c}(k)^\top \mathbb{H} \hat{U}_{N_c}(k) + \hat{U}_{N_c}(k)^\top \mathbb{F} \mathbf{x}(k) \right) \quad (6.95)$$

$$\text{with constraints } \Xi \hat{U} \leq \Lambda. \quad (6.96)$$

To obtain the command $\mathbf{u}(k)$ at time k , MPC computes the following steps:

1. get the new state measurement $\mathbf{x}(k)$;
2. minimize the performance cost J_{N_p, N_c} by computing the optimal predicted states $\hat{X}_{N_p}(k)$ and predicted commands $\hat{U}_{N_c}(k)$ subject to constraints;
3. take the first predicted command $\hat{\mathbf{u}}(k|k)$ from $\hat{U}_{N_c}(k)$, as shown in Fig. 6.4.

6.6.1 Model Predictive Control in AO

The MPC to the SCAO problem is conceptually similar to a LQ problem solved every sample time where, for example, maximum/minimum admissible voltages for the DM actuators can be easily expressed as constraints on the output command vector.

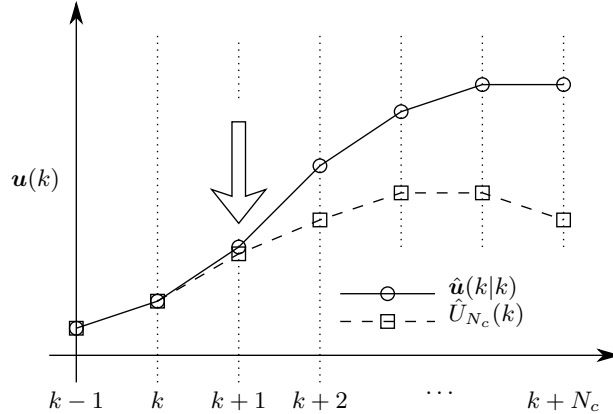


Fig. 6.4: Example of the optimal control $\mathbf{u}(k)$ in the MPC framework: only the one-step ahead command $\hat{\mathbf{u}}(k|k)$ is applied from the receding horizon command vector $\hat{U}_{N_c}(k)$.

The matrices \mathbb{H} and \mathbb{F} within the quadratic programming minimization may bring to ill conditioned numerical solutions. DM actuator coupling and dynamics affect the condition number which, in turn, decreases the convergence time for gradient-based numerical methods, as stated by Konnik *et al.* [51]. The condition number can be improved by simplifying the actuator dynamic model and exploiting the banded character of \mathbb{H} .

Choosing the most suitable horizons N_p and N_c is not trivial: it is a trade-off between better optimization with respect to the constraints when horizons are larger, or computationally tighter closed-loop control when horizons are smaller.

It is worth highlighting that the prediction of $\hat{\mathbf{x}}$ for N_p -ahead steps required by the MPC algorithm is usually computed by an optimal Kalman predictor.

6.7 Conclusion

The design of a AO control strategy must take into account the hybrid deterministic-stochastic and continuous-discrete characteristics of the AO configuration, along with the large number of WFS inputs and DM outputs. Since the atmospheric turbulence energy is concentrated on the low spatio-temporal frequency range, and WFS measurement noise mostly affects higher frequencies, the integral regulator is essential in the closed-loop regulator (Chapter 3). The most used AO controller is an array of PI regulators placed after the control matrix (Chapter 3) and proved to achieve diffraction-limited images on ground-based telescopes. However, tuning the PI parameters of a large array of PI controllers is problematic for modern AO systems and is sub-optimal with respect to the incoming turbulence. Since the AO system can be modeled as a MIMO state-space system, advanced control strategies can be exploited to design optimal regulators. Furthermore, decoupling the MIMO system into several SISO sub-systems has a decisive impact on the control strategy implementation, as each SISO sub-system can be developed independently.


This Chapter reviewed the most common model-based control strategies implemented in AO systems. The regulator consists of two components:

- The *observer* that estimates the AO system state from the WFS measurement, implemented by the Kalman filter;
- The *controller* that produces the DM command from the AO system state, designed using LQ, \mathcal{H}_2 or \mathcal{H}_∞ approach.

The Model Predictive Controller (MPC) is introduced as a formulation that allows to integrate in the optimization problem constraints on the DM stroke and actuator saturation. Moreover, such constraints can be efficiently expressed on the modal-based design. At the time of writing, the MPC is still too computationally demanding to be deployed for large systems.

The following Chapter [7](#) presents a modally decoupled LQG shaped like a PI controller (LQGI).

PI-Shaped LQG Control

Summary. The goal of an Adaptive Optics system is to control a deformable mirror to compensate for the distortions introduced by atmospheric turbulence in the incoming wavefront. Therefore, a modal PI-shaped LQG controller is designed, taking advantage of model knowledge. This approach compares favorably with standard PI and Kalman-based controllers used in state-of-the-art control systems. The architecture is validated with simulations and experiments on a laboratory adaptive optics test-bench. 

¹ This chapter is based on the paper: J. Mocci, M. Quintavalla, A. Chiuso, S. Bonora, and R. Muradore. “PI-Shaped LQG Control Design for Adaptive Optics Systems”. In: *Control Engineering Practice* (2020) (submitted).

7.1 Introduction

The LQG regulator is optimal with regard to a LTI state-space description of the system to control. This approach uses the modeled system to describe the performance index to be optimized (i.e. minimization of the residual phase variance) and then design a feedback loop with the desired properties.

The integral action in PI controllers has proved to be extremely useful in AO systems for improving atmospheric turbulence rejection and for correcting almost static aberrations. However, the classical LQG regulator can not force the integral action in the regulator. Hence, a PI-shaped LQG modal controller is proposed, based on a deterministic MIMO model of the AO system augmented with an integral term (as shown in Figure 7.1), which shares the properties of the PI controller [105].

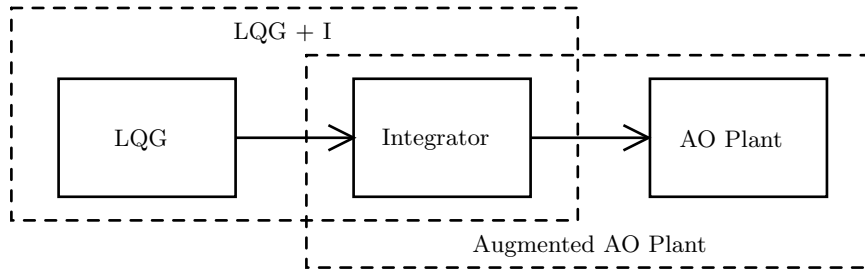


Fig. 7.1: Conceptual LQGI strategy scheme. The full AO plant is augmented with an integrator. The cost function of the LQG controller takes the integrator into account, yielding to the LQGI control strategy.

Modal control [59, 31] is also effective when dealing with AO systems, as it projects the original MIMO system into a set of decoupled SISO systems by exploiting a suitable modal basis. This way, a PI-shaped LQG regulator can be designed and tuned individually for each modal term.

The proposed control strategy is implemented on the PhotonLoop software [70] and evaluated both in simulation and using the optical testbench available at *CNR-IFN* of Padova. The AO system is composed of a DM with 32 actuators, an USB3 camera measuring 25 modal terms and a controller, with closed-loop frequency of 200Hz and latency of $\sim 10\text{ms}$. PhotonLoop allows to *play* the same atmosphere multiple times, making the comparison of the results much easier and fair. The proposed scheme compares favorably with standard controllers used in state-of-the-art telescopes.

7.2 PI-Shaped LQG controller

The proposed controller is an improvement of the standard LQG controllers. It is well known that, by using the separation principle [2], the controller consists of the cascade of a Kalman filter (see subsection 7.2.1) and an LQ state-feedback

controller (see subsection 7.2.2). However, standard LQ controllers would not, in general, include integral actions which are needed to ensure asymptotic rejection of constant disturbances. In order to do so, following [2], the LQ design is performed on the full system model augmented with an integral component. This lets the integral action to appear into the performance index as a weighted term.

The LQGI regulator is designed using the AO system formulated into the Zernike modal-space presented in Section 6.2. Being spatially orthogonal, control in the Zernike modal space can be performed separately for each of the N_w modal terms in a parallel fashion. The outcome of this design will be an Optimal Modal Gain (OMG) controlled system in which each SISO LQG control problem related to the system $\mathbf{G}_f^{(j)}$ computes the controller $\mathbf{C}_j(z)$. The N_w controllers are then combined into $\mathbf{C}(z)$ to form the complete control strategy (see Figure 7.2):

$$\mathbf{C}(z) = \Pi_{u2y}^\dagger \Pi_{u2y} [\mathbf{C}_1(z) \dots \mathbf{C}_{N_w}(z)]^\top \Pi_{w2y}^\dagger. \quad (7.1)$$

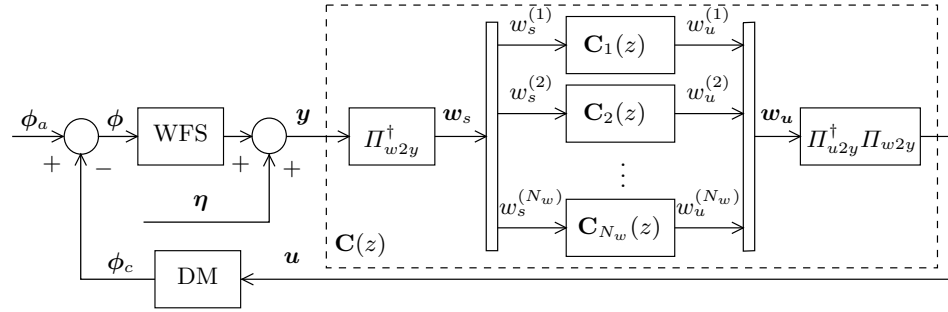


Fig. 7.2: Modal control architecture.

7.2.1 State Estimator

Let $\mathbf{x}_f^{(j)}(k)$ be the state vector of the j -th full system realization $\mathbf{G}_f^{(j)}$ defined in Eq. (6.15)

$$\mathbf{x}_f^{(j)}(k) = \begin{bmatrix} \mathbf{x}_a^{(j)}(k) \\ \mathbf{x}_{ao}^{(j)}(k) \end{bmatrix}, \quad (7.2)$$

where $\mathbf{x}_a^{(j)}(k)$ is the state vector of the atmospheric turbulence system $\mathbf{G}_a^{(j)}(z)$ and $\mathbf{x}_{ao}^{(j)}(k)$ is the state vector of the joint DM and WFS system

$$\mathbf{G}_{ao}^{(j)} = \mathbf{G}_c^{(j)} \mathbf{G}_s^{(j)} \quad (7.3)$$

$$= \begin{bmatrix} A_{ao}^{(j)} & B_{ao}^{(j)} \\ C_{ao}^{(j)} & \mathbf{0} \end{bmatrix} : w_u^{(j)} \mapsto w_c^{(j)}. \quad (7.4)$$

The state vector $\mathbf{x}_a^{(j)}(k)$ of the atmospheric turbulence is estimated using the *current* Kalman estimator [33]:

$$\mathbf{G}_K^{(j)} = \left[\begin{array}{c|c} A_a^{(j)} - K^{(j)}C_a^{(j)} & K^{(j)} \\ \hline I - K_x^{(j)}C_a^{(j)} & K_x^{(j)} \end{array} \right] : w_a^{(j)} \mapsto \hat{\mathbf{x}}_a^{(j)}, \quad (7.5)$$

where

$$K^{(j)} = A_a^{(j)}Y^{(j)}C_a^{(j)\top} \left(C_a^{(j)}Y^{(j)}C_a^{(j)\top} + r_K^{(j)} \right); \quad (7.6)$$

$$K_x^{(j)} = Y^{(j)}C_a^{(j)\top} \left(C_a^{(j)}Y^{(j)}C_a^{(j)\top} + r_K^{(j)} \right), \quad (7.7)$$

are the steady-state filter and predictor Kalman gains, respectively. Both are calculated from the unique solution $Y^{(j)}$ of the associated Discrete Algebraic Riccati Equation (DARE):

$$Y^{(j)} = A_a^{(j)}Y^{(j)}A_a^{(j)\top} - \frac{A_a^{(j)}Y^{(j)}C_a^{(j)\top}C_a^{(j)}Y^{(j)}A_a^{(j)\top}}{C_a^{(j)}Y^{(j)}C_a^{(j)\top} + r_K^{(j)}} + Q_K^{(j)}. \quad (7.8)$$

The weights $Q_K^{(j)}$ and $r_K^{(j)}$ are inferred from the knowledge of atmospheric turbulence $\mathbf{G}_a^{(j)}$ dynamics and WFS noise, and are thus defined as

$$Q_K^{(j)} = \rho_K B_a^{(j)} B_a^{(j)\top}; \quad (7.9)$$

$$r_K^{(j)} = \sigma_\eta^{2(j)}, \quad (7.10)$$

where $\rho_K \in \mathbb{R}$ is a parameter to be tuned.

Since $\mathbf{G}_{ao}^{(j)}$ is modeled as a cascade of unitary delays, it is sufficient to estimate the state $\bar{\mathbf{x}}_{ao}^{(j)}(k)$ by an open-loop state observer built with the AO system matrices and fed with the command $u^{(j)}(k)$:

$$\mathbf{G}_{aoe}^{(j)} = \left[\begin{array}{c|c} A_{ao}^{(j)} & B_{ao}^{(j)} \\ \hline I & \mathbf{0} \end{array} \right] : w_u^{(j)}(k) \mapsto \bar{\mathbf{x}}_{ao}^{(j)}. \quad (7.11)$$

The Kalman estimator in (7.5) receives the measurement $w_a^{(j)}(k)$ as input. However, the WFS is only able to measure the closed-loop residual $w_s^{(j)}(k)$, i.e. the difference between incoming atmospheric turbulence and DM correction. Hence, the output of the AO estimator in (7.11)

$$\bar{w}_c^{(j)}(k) = C_{ao}^{(j)} \bar{\mathbf{x}}_{ao}^{(j)}(k) \quad (7.12)$$

is used to compute the Pseudo-Open Loop (POL) [28]

$$\bar{w}_a^{(j)}(k) = w_s^{(j)}(k) - \bar{w}_c^{(j)}(k), \quad (7.13)$$

by adding the estimated DM correction $\bar{w}_c^{(j)}(k)$ to the measured residual. The POL approximation of the true atmosphere is good at low temporal frequency, where the DM compensates for the atmospheric turbulence and the signal is not corrupted by high frequency measurement noise.

7.2.2 PI-shaped LQ state-feedback matrix

The design of the optimal state feedback matrix is obtained solving a slightly modified LQ problem. Since an integrator is required in the control loop to ensure good performance (e.g. rejection of constant disturbances), the LQ performance index must include an integral term of the compensation error. Let

$$x_I^{(j)}(k+1) = x_I^{(j)}(k) + w_s^{(j)}(k) \quad (7.14)$$

$$= x_I^{(j)}(k) + C_f^{(j)} \mathbf{x}_f^{(j)}(k), \quad (7.15)$$

be the state vector of the j -th system $\mathbf{G}_I^{(j)} : w_s^{(j)} \mapsto w_I^{(j)}$ which computes the backward Euler integral $w_I^{(j)}(k)$ of the measured residual $w_s^{(j)}(k)$ of the full system $\mathbf{G}_f^{(j)}$. By augmenting the update equation of the full system $\mathbf{G}_f^{(j)}$ with $x_I^{(j)}(k)$ we end up with

$$\mathbf{G}_+^{(j)} = \begin{cases} \mathbf{x}_f(k+1) = A_f^{(j)} \mathbf{x}_f(k) + B_f^{(j)} \left[v^{(j)}(k) w_u^{(j)}(k) \eta^{(j)}(k) \right]^\top \\ x_I(k+1) = C_f^{(j)} \mathbf{x}_f(k) + x_I(k) \end{cases}, \quad (7.16)$$

where

$$\mathbf{x}_+^{(j)}(k) = \begin{bmatrix} \mathbf{x}_f^{(j)}(k) \\ x_I^{(j)}(k) \end{bmatrix}; A_+^{(j)} = \begin{bmatrix} A_f^{(j)} & \mathbf{0} \\ C_f^{(j)} & 1 \end{bmatrix}; B_+^{(j)} = \begin{bmatrix} B_f^{(j)} \\ \mathbf{0} \end{bmatrix}, \quad (7.17)$$

is the state vector and the state-space matrices of the augmented system $\mathbf{G}_+^{(j)}$ (the output equation is omitted).

Our goal is to keep the variances of both $w_y^{(j)}(k)$ and its integral small. Mathematically this can be formalized by saying that, for all modal terms, $\mathbf{x}_f^{(j)\top} C_f^{(j)\top} C_f^{(j)} \mathbf{x}_f^{(j)}(k)$ and $x_I(k)^\top x_I^{(j)}(k)$ have to be small. Therefore it is reasonable to consider the following (PI-shaped) LQ cost

$$\min_{u^{(j)}} \lim_{N_k \rightarrow +\infty} \frac{1}{N_k} \sum_{k=0}^{N_k} \left(\mathbf{x}_+^{(j)}(k)^\top Q_F^{(j)} \mathbf{x}_+^{(j)}(k) + w_u(k)^\top r_F w_u^{(j)}(k) \right) \quad (7.18)$$

where

$$Q_F^{(j)} = \begin{bmatrix} \rho_P C_f^{(j)\top} C_f^{(j)} & 0 \\ 0 & \rho_I I \end{bmatrix} \quad (7.19)$$

and $r_F^{(j)} \in \mathbb{R}$ are the weights, and $\rho_P, \rho_I \in \mathbb{R}$ are the proportional and integrative tunable parameters, respectively.

From the solution $X^{(j)}$ of the DARE associated to the j -th LQ problem

$$X^{(j)} = A_+^{(j)\top} X^{(j)} A_+^{(j)} - \frac{A_+^{(j)\top} X^{(j)} B_+^{(j)} B_+^{(j)\top} X^{(j)} A_+^{(j)}}{B_+^{(j)\top} X^{(j)} B_+^{(j)} + r_F^{(j)}} + Q_F^{(j)} \quad (7.20)$$

the state feedback matrix

$$F^{(j)} = \left(r_F + B_+^{(j)\top} X^{(j)} B_+^{(j)} \right)^\dagger B^{(j)\top} X^{(j)} = \begin{bmatrix} F_a^{(j)} & F_{ao}^{(j)} & F_I^{(j)} \end{bmatrix}, \quad (7.21)$$

gives the PI-shaped control strategy

$$w_u^{(j)}(k) = -F_a^{(j)} \hat{\mathbf{x}}_a^{(j)}(k) - F_{ao}^{(j)} \hat{\mathbf{x}}_{ao}^{(j)}(k) - F_I^{(j)} x_I^{(j)}(k). \quad (7.22)$$

Tuning of the parameters ρ_I , ρ_P and r_F is not a trivial task, see [2] and the references therein; however even the suboptimal choice of setting the same weigh for all N_w controllers, outperforms the standard *PI* and the Kalman-based controller, as we will show later.

Remark 7.1. It is worth remarking that what astronomers are really interested in is to reduce the effects of the atmosphere turbulence *before* the WFS, i.e. the residual ϕ . The reason is due to the fact that the science camera integrates the residual at exactly this point.

7.2.3 Full controller

The Kalman observer $\mathbf{G}_K^{(j)}$, AO estimator $\mathbf{G}_{aoe}^{(j)}$, integrator $\mathbf{G}_I^{(j)}$ and PI-shaped LQI regulator gain F are interconnected together to yield the complete controller $\mathbf{C}_j(z)$ for the j -th modal term as shown in Figure 7.3

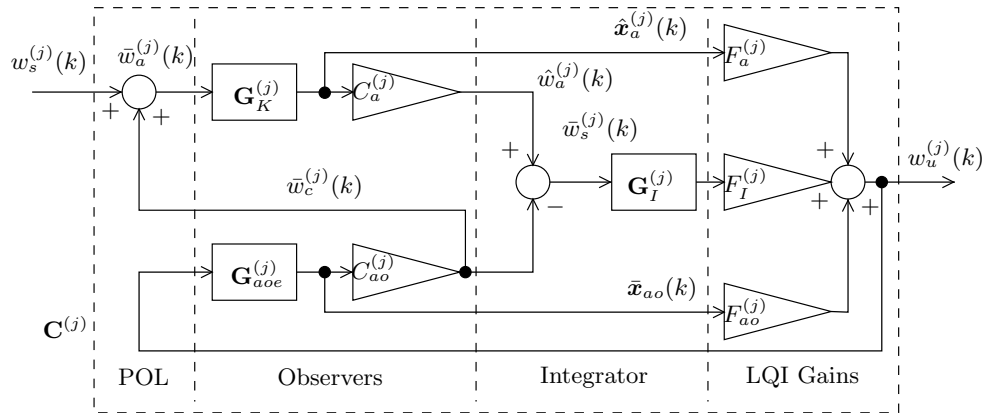


Fig. 7.3: Block diagram of the j -th mode PI-shaped LQG regulator.

Exploiting the separation principle, the state feedback in (7.22) is computed using the estimated state $\hat{\mathbf{x}}_+^{(j)}(k|k)$, i.e.

$$w_u^{(j)}(k) = -F_a^{(j)} \hat{\mathbf{x}}_a^{(j)}(k|k) - F_{ao}^{(j)} \hat{\mathbf{x}}_{ao}^{(j)}(k) - F_I^{(j)} \hat{x}_I^{(j)}(k|k). \quad (7.23)$$

Since the AO estimator $\mathbf{G}_{aoe}^{(j)}$ is modeled as a delay, the feed-forward term due to the *current* Kalman state estimator $\mathbf{G}_K^{(j)}$ in (7.5) does not lead to an algebraic loop.

7.3 Simulation Results

The simulations are carried out in Matlab, modeling the AO system for each modal term as shown in Figure 6.1. We use the orthogonal Zernike polynomials basis truncated to $N_w = 25$ terms. The simulation testbench does not include the modal projector Π_{w2y} and the interaction matrix Π_{u2y} in the loop, since all the signals are already expressed as modal coefficients.

Atmospheric turbulence affecting each j -th Zernike mode is simulated by the output of the discrete turbulence model $\mathbf{G}_a^{(j)}(z)$ described in (6.13) and (6.14), fed by Gaussian white noise process with zero mean $v^{(j)}(k)$. Since high-order Zernike modes have higher cut-off frequency $f_a^{(j)}$ and lower amplitude $\alpha_a^{(j)}$ than lower order modes [100], we select the parameters accordingly to the values listed in Table 7.1

| j | [1, 2] | [3, 5] | [6, 9] | [10, 14] | [15, 20] | [21, 25] |
|------------------|--------|--------|--------|----------|----------|----------|
| $f_a^{(j)}$ (Hz) | 1 | 2.5 | 4 | 5.5 | 7 | 8.5 |
| $\alpha_a^{(j)}$ | 5.0000 | 1.5811 | 0.5000 | 0.1581 | 0.0500 | 0.0158 |

Table 7.1: Cutoff $f_a^{(j)}$ and gain $\alpha_a^{(j)}$ parameters of the atmospheric turbulence model for the first $N_w = 25$ Zernike modal terms.

The measurement noise $\eta^{(j)}(k)$ added to each j -th mode is modeled as a zero-mean, white Gaussian process, whose variance $\sigma_\eta^{2(j)}$ is inferred from the WFS measurement noise variance Σ_η , as described in (6.12).

The DM model $\mathbf{G}_c(z)$ defined in (6.1) and (6.7) is designed with $d = 1$ sample delay, which approximates the actual DM used in the testbench.

We compare our PI-shaped LQGI control strategy $\mathbf{C}(z)$ with a classical PI controller $\mathbf{C}^{PI}(z)$ and a PI controller with an array of Kalman filters/estimators $\mathbf{C}^K(z)$ to filter out the noise [117].

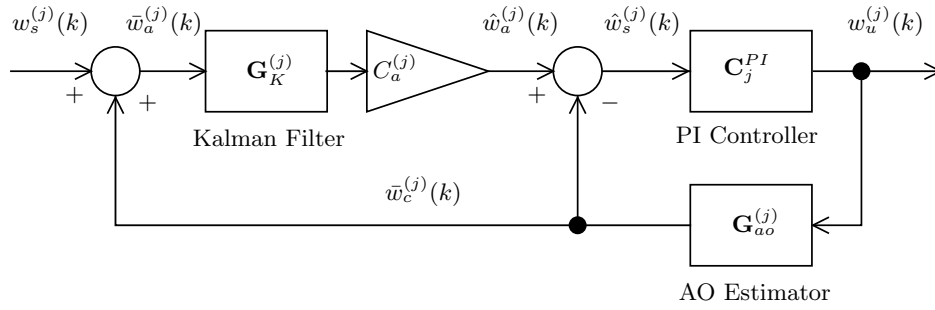
PI strategy: For the j -th mode, the PI control law is:

$$\mathbf{C}_j^{PI}(z) = \frac{k_I z^{-1}}{1 - z^{-1}} + k_P, \quad (7.24)$$

where $k_P, k_I \in \mathbb{R}$ are the proportional and integrative gain, respectively.

It can be shown that the integral part of this strategy implicitly contains an observer of the turbulence model which, in modal control, can be tuned accordingly to obtain good results [56]. However, the lack of more in-depth knowledge of the system like the exact turbulence model and the measurement noise variance makes the tuning hard and time consuming;

Kalman+PI strategy: An improvement over the PI strategy is to filter the measurement $w_s^{(j)}(k)$ before the control action by using the POL estimator $\mathbf{G}_{aoe}^{(j)}(z)$ with the Kalman estimator $\mathbf{G}_K^{(j)}(z)$ described in Section 7.2.1 [28]. The system connecting the Kalman filter, PI controller and POL estimator is shown in Figure 7.4.

Fig. 7.4: Kalman + PI strategy for the j -th modal term.

| PI | PI-Kalman | PI-LQGI |
|-------------------|---------------------|---------------------------------|
| $k_P \in [.6, 1]$ | $k_P \in [.8; 1.2]$ | $\rho_P \in [1e^0, 1e^3]$ |
| $k_I \in [.6, 1]$ | $k_I \in [.6, .8]$ | $\rho_I \in [1e^{-4}, 1e^{-2}]$ |

(a) $\sigma_\eta^2 = 0.1$

| PI | PI-Kalman | PI-LQGI |
|--------------------|-------------------|---------------------------------|
| $k_P \in [.3, .6]$ | $k_P \in [.6; 1]$ | $\rho_P \in [1e^0, 1e^3]$ |
| $k_I \in [.3, .6]$ | $k_I \in [.6, 1]$ | $\rho_I \in [1e^{-4}, 1e^{-2}]$ |

(b) $\sigma_\eta^2 = 1$

| PI | PI-Kalman | PI-LQGI |
|---------------------|-------------------|---------------------------------|
| $k_P \in [.01, .4]$ | $k_P \in [.6; 1]$ | $\rho_P \in [1e^0, 1e^3]$ |
| $k_I \in [.01, .4]$ | $k_I \in [.6, 1]$ | $\rho_I \in [1e^{-4}, 1e^{-2}]$ |

(c) $\sigma_\eta^2 = 10$ Table 7.2: Minimum and maximum parameter ranges for different controllers and different measurement noise variances σ_η^2 .

The parameter ranges for each control strategy and different noise variances are given in Table 7.2.

We now introduce an approximation of the Strehl Ratio that will be used to compare the different control strategies. Let

$$(\Delta\phi)^2 \approx \frac{1}{N_k + 1} \sum_{k=0}^{N_k} \sum_{j=1}^{N_w} w^{(j)}(k)^2 \quad (7.25)$$

be the temporal mean over N_k samples of the wavefront phase variance estimated from the residual Zernike coefficients $w^{(j)}(k)$ [117]. The outcome of each simulation run is the Maréchal's approximation of the Strehl ratio [117] computed as:

$$\text{SR} \approx \exp\left(-\left(\frac{2\pi}{\lambda}\right)^2 (\Delta\phi)^2\right), \quad (7.26)$$

where the wavelength $\lambda = 633\text{nm}$. This approximation is valid for Strehl ratio larger than 20%.

Surfaces in Figure 7.5 show the Strehl ratio for different tuning of controller parameters and different measurement noise variance σ_η^2 . While the classical PI controller must be tuned for each variance σ_η^2 from scratch, the LQGI achieves good results with the same parameter configuration in all simulated variances. It can be observed that adding the integrator to the LQG strategy means that for some weights the optimization problem has no solutions that guarantees the stability of the closed-loop system, as noted by Guaracy *et al* 38. Therefore, the parameter space is explored in the simulated environment to gather the subset of parameters that makes the closed-loop system asymptotically stable while focusing on the achieved Strehl. The corresponding phase and gain margins can be calculated from the regulator and plant transfer functions.

Figure 7.6 shows that the PI-LQGI approach outperforms the other strategies in all cases, achieving higher Strehl ratio.

The rejection transfer functions for several Zernike modes of different order are shown in Figure 7.7. Since the PI regulator parameter set for a given measurement noise variance is optimized for all modes as a whole, its response is the same for all modes as well, yielding to sub-optimal performance when considering each mode independently. On the other hand, the frequency responses of the Kalman+PI regulator adapt to the atmospheric turbulence characteristics of each mode. This is because the turbulence model is embedded into the Kalman predictor. However, the rejection transfer functions exhibit a resonant peak between 40 and 50 Hz. Furthermore, the integral action is severely reduced at the lowest measurement noise variance ($\sigma_\eta^2 = 0.1$). The LQGI regulator shows a better frequency responses over the Kalman+PI regulator. Since the WFS and DM delays are embedded into the AO model, the LQGI rejection transfer functions avoid the resonant spikes.

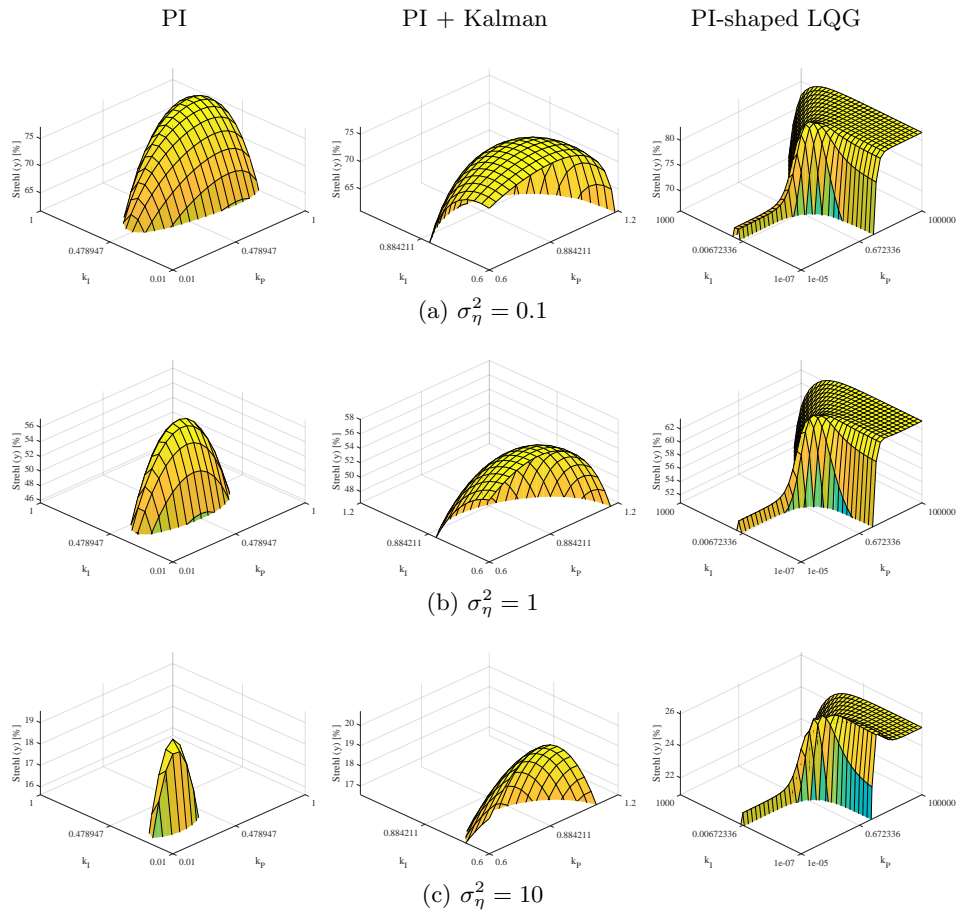
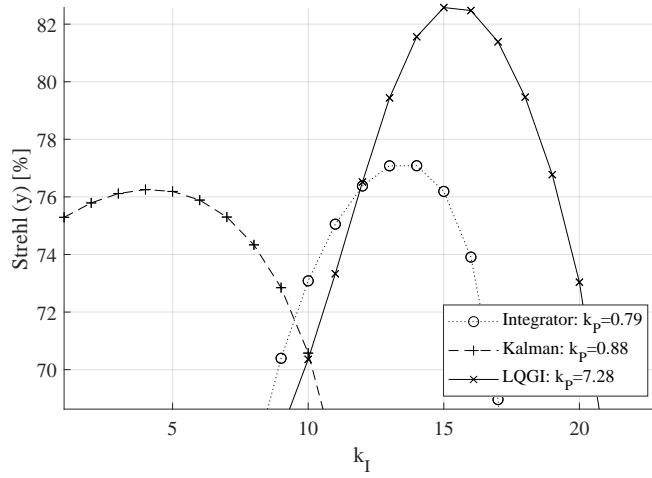
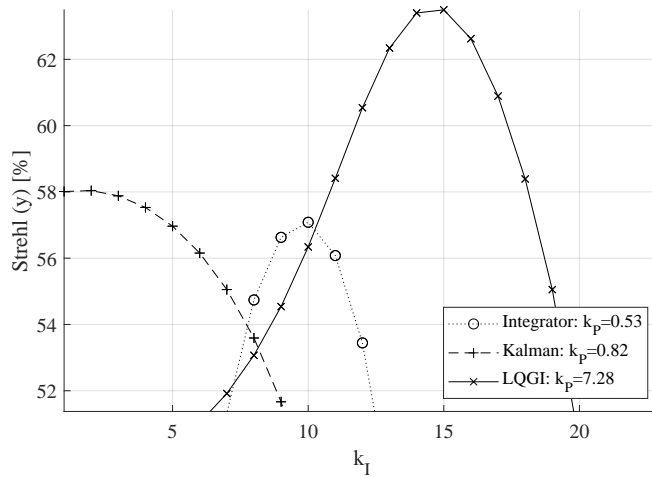


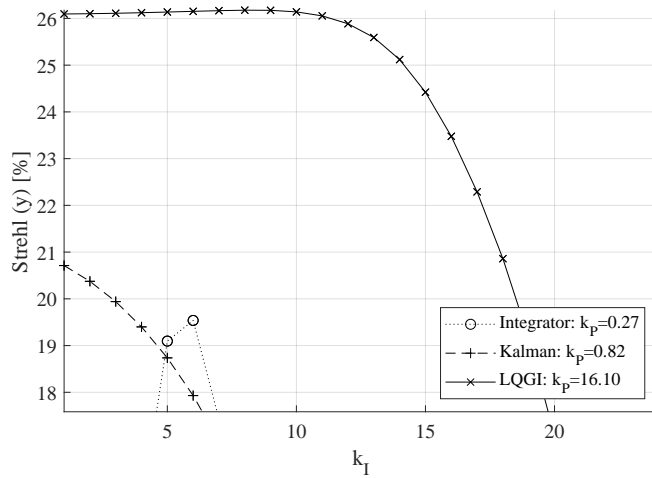
Fig. 7.5: Values of k_I (ρ_I) and k_P (ρ_P) for the best runs of each regulator for the given measurement noise σ_η^2 .



(a) $\sigma_\eta^2 = 0.1$



(b) $\sigma_\eta^2 = 1$



(c) $\sigma_\eta^2 = 10$

Fig. 7.6: Values of k_I (ρ_I) for the best k_P (ρ_P) of each regulator for the given measurement noise variance σ_η^2 .

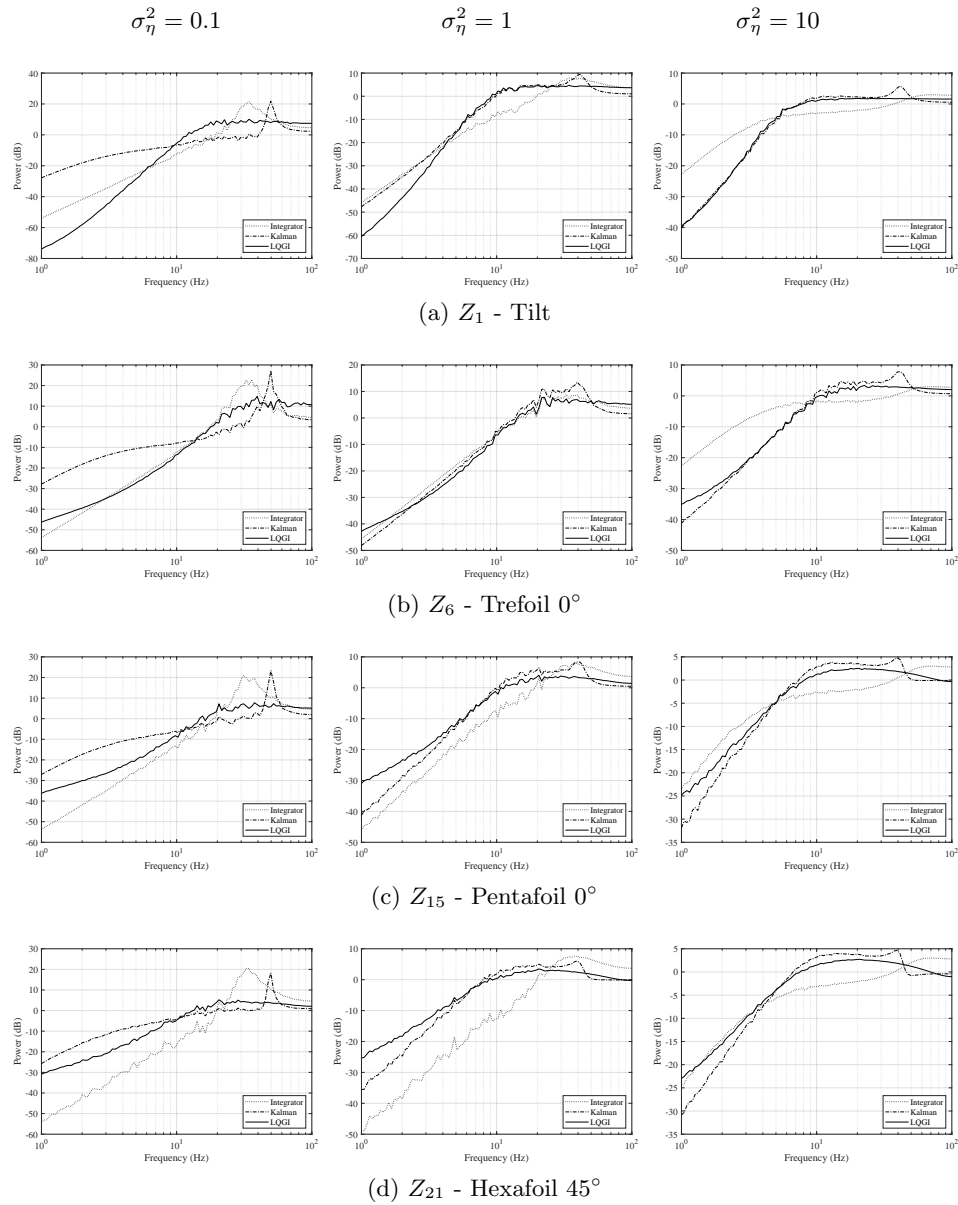


Fig. 7.7: Rejection functions for the optimal values of each regulator. From left to right: increasing measurement noise variance. From top to bottom: increasing Zernike orders.

7.4 Experimental Results

We carried out the experiments on the AO testbench available at the *CNR-IFN* laboratory of Padova (see Figure 7.8).

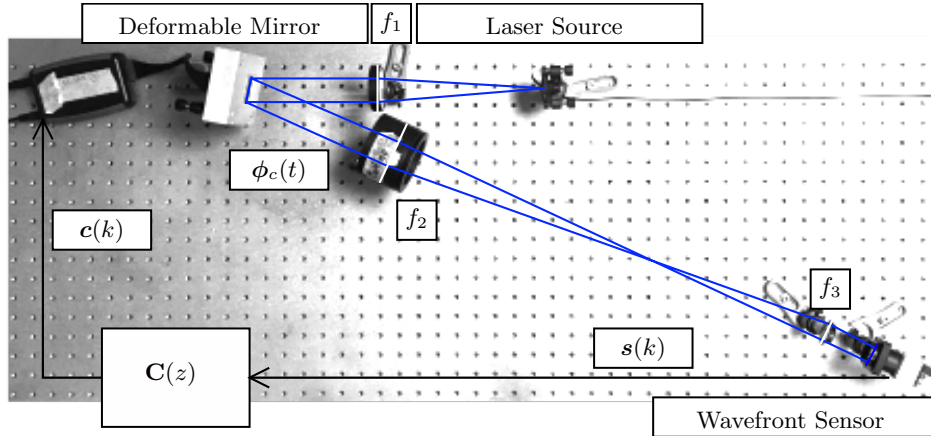


Fig. 7.8: AO real testbench: the controller $C(z)$ is implemented in a laptop.

A laser source ($\lambda = 633nm$) collimated by the lens $f_1 = 250mm$ is sent on a 30 mm diameter, 32 actuators piezoelectric bimorph DM, manufactured at *CNR-IFN* of Padova. The reflected light beam is collected by the WFS through lenses $f_2 = 400mm$ and $f_3 = 100mm$ that reduce the beam size to a diameter of 4mm and provide the optical conjugation of the DM surface to the WFS. The Shack-Hartmann WFS is composed by a lenslets array (focal length of 5.139mm and 150 μm pitch) and a CMOS camera (*iDS UI-3060CP*, pixel size 5.86 \times 5.86 μm). In this configuration, the wavefront is sampled over a grid of 27 \times 27 subapertures for a total of 576 subapertures inside the DM circular pupil, corresponding to a Region Of Interest (ROI) of 682 \times 687 pixels on the camera.

The controller $C(z)$ is implemented on the *PhotonLoop* software [70] installed on a *Lenovo Thinkpad* laptop with an *i7* CPU from *Intel*. DM and WFS are connected via USB2.0 and USB3.0, respectively. In order to maximize bandwidth throughput and minimize latency and jitter effects of the communication between the controller and the physical devices, the controller frequency is set to 200Hz and the camera image bit-depth to 8-bit. The total latency of the AO control loop is 2 frames, i.e. $\sim 10ms$, due to the WFS exposure and transfer times and DM response time.

After setting the WFS parameters and powering the DM, the camera's ROI is automatically centered on the optical beam so that the centroids extraction grid is correctly located over the Shack-Hartmann spots, removing the global tilt.

The exposure time of the camera is set to 0.32ms in order to avoid saturation of the WFS detector. The spots reference positions, from which slopes are calculated, are set from the measured centroids when the DM is in its relaxed position, i.e. $\mathbf{u}(k)$ is zero.

The interaction matrix Π_{u2y} is built by collecting the measured slopes produced by poking the DM actuators with Hadamard patterns [49]. Since the coefficients of Hadamard patterns are either 1's or -1's, all actuators are poked at once, hence improving the Signal-to-Noise Ratio (SNR) and capturing eventual cross-correlations between neighbouring actuators. By using the Singular Values Decomposition (SVD) to pseudo-inverting the interaction matrix into the control matrix Π_{u2y}^\dagger , noisy or hardly obtainable system modes are filtered out. In our case we keep 23 system modes out of 32.

To keep the experiments fair, the time series of the atmospheric turbulence $w_a^{(j)}(k)$ is simulated once using the model $\mathbf{G}_a^{(j)}(z)$ for each of the N_w modal terms. Since there is no real turbulence $\phi_a(t)$ in the optical testbench, the WFS measures only what is produced by the DM. Therefore, the synthesized turbulence $\mathbf{w}_a(k)$ is *played back* by the *PhotonLoop* software, i.e. it is added to the measured modal terms, to calculate the residual modes $\mathbf{w}(k)$ (as shown in Chapter 3). As the intensity of the laser beam is high enough to make the WFS measurement noise negligible, a zero-mean Gaussian process signal $\eta^{(j)}(k)$ of variance $\sigma_\eta^{2(j)}$ is injected after the residual $w^{(j)}(k)$ to simulate different amounts of measurement noise.

The experimental results shown in Figure 7.10 and summarized in Table 7.9 confirms that the PI-LQGI strategy outperforms the other control strategies in all working conditions.

The Strehl values are higher than the ones obtained in simulation: this is likely due to the closed-loop latency model mismatch, which in the real case is a value between 1 and 2 samples.

| σ_η^2 | PI | PI-Kalman | PI-LQGI |
|-----------------|--------|-----------|---------------|
| 0.1 | 92.99% | 92.71% | 93.47% |
| 1 | 82.39% | 83.08% | 84.32% |
| 10 | 48.24% | 51.42% | 56.21% |

Fig. 7.9: Summary of the best Strehl results from the testbench experiments.

7.5 Conclusion

With respect to standard controllers, the proposed model-based PI-shaped LQG control architecture provides the best correction in the considered cases and so the highest Strehl ratio. Such results have been confirmed by deploying and testing the control strategy on a real AO testbench.

The main advantages of the PI-shaped LQG strategy are:

1. automatic procedure to design the controller;
2. parallel computation of each $\mathbf{C}_j(z)$ due to their mutual independence;
3. possibility to exploit modal control in more complicated AO systems, i.e. woofer-tweeter AO systems;
4. simple integration of additional modules such as vibration compensation [77].

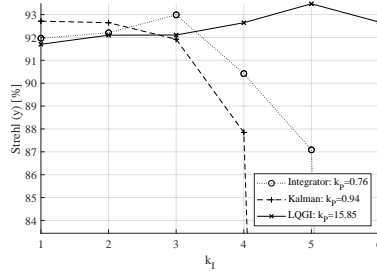
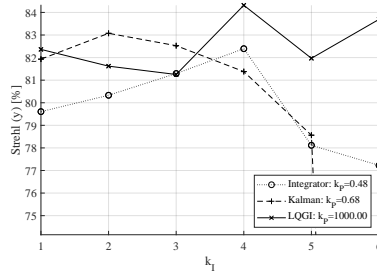
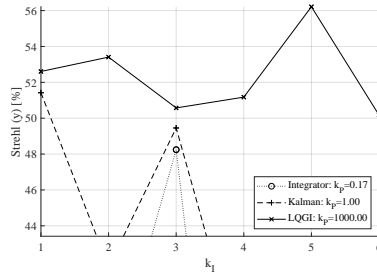
(a) $\sigma_\eta^2 = 0.1$ (b) $\sigma_\eta^2 = 1$ (c) $\sigma_\eta^2 = 10$

Fig. 7.10: Experimental values of k_I (ρ_I) for the best k_P (ρ_P) of each regulator for the given measurement noise variance σ_η^2 .

The PI-shaped LQG strategy can also be paired with an on-line identification procedure for the atmospheric turbulence and AO plant models, i.e. periodic identification using subspace methods, to guarantee optimality when the model of the controlled system evolves over time.

The modal-space design of the proposed LQGI controller, while easier to understand and tune, introduces reconstruction error $\sigma_{>N_w}^2$. If the Zernike basis requires N_w terms to reconstruct the wavefront with small error but the low spatial resolution of the WFS makes for a large aliasing error, then a more suitable basis can be chosen. The aliasing error σ_{alias}^2 that might arise due to the WFS not respecting the Nyquist-Shannon theorem (e.g., reduced number of subapertures because the incoming photon flux is too weak) leads inevitably to modal coupling. We observed

that the SISO regulators correct each-other, therefore mitigating the issue (albeit non optimally).

Conclusions

Adaptive Optics is a rather large, constantly growing topic that embraces several scientific fields. One of the challenges tackled by Adaptive Optics is the compensation for the atmospheric turbulence, which degrades the spatial resolution of the observed object. Aside from the appeal of Adaptive Optics to the astronomical community, for which it is considered a key component part of any large telescope, it is now so widely used that it gave rise to optical technologies for a wide range of applications. For example, a flying unmanned drone can be powered by feeding it a high power laser from a ground station. There, Adaptive Optics is able to accurately compensate for the atmospheric turbulence between the ground station and the drone, hence manipulating the laser beam such that the energy does not get excessively scattered. Moreover, there are reports of a solar photothermal plant which increased its power output by controlling each of its glass panels to compensate for the aberrations - much alike to the deformable mirror of an Adaptive Optics system.

This PhD thesis is motivated by the need of advanced control strategy for designing high-performance controllers. The AO plant and atmospheric turbulence were formalized as state-space linear time-invariant systems in Chapter 2. The full AO system model is needed to exploit model-based control.

A Shack-Hartmann wavefront sensor was used to measure the horizontal atmospheric turbulence (Chapter 4). The experimental measurements yielded to the C_n^2 atmospheric structure parameter, which is essential in the derivation of turbulence parameters, and the Zernike terms time-series.

Experimental validation in Chapter 5 shows that the centroid extraction algorithm implemented on the Jetson GPU outperforms (i.e. it is faster) than the CPU implementation. In fact, due to the working principle of the Shack-Hartmann wavefront sensor, the intensity image captured from its camera is partitioned into several sub-images, each related to a point of the incoming wavefront. Such sub-images are independent each-other and can be analyzed concurrently.

In Chapter 7, the AO model discussed in Chapter 2 is exploited to automatically design an advanced linear-quadratic Gaussian controller with integral action. Experimental evidence shows that the system augmentation approach outperforms the simple integrator and the integrator with the Kalman predictor.

The outcome of this PhD thesis are:

- The implementation of the WFS calculation in a compact, portable GPU architecture, which reduces the latency;
- The design of a modern control strategy that uses the augmented AO model to automatically synthesize the optimal regulator.

Future works will focus on the following topics:

- The atmospheric turbulence model can be inferred by means of *convolutional neural networks*, with the added benefit of *multi-sample* predictions;
- The full identification and regulator strategies can be integrated into a hybrid, compact CPU/GPU architecture yielding to a *standalone* AO controller which can fit into a *smart camera*. Coupled with the deformable lenses, instead of deformable mirrors, this solution allows high performance AO within minimal space and cost requirements;
- One of the model-based strategies discussed in Chapter 6 is the Model Predictive Control, which is yet to be implemented in an AO system because of the challenges it poses for its real-time implementation. However, the GPU architecture reviewed in Chapter 5 can *accelerate* the computation and hence make it feasible within the AO temporal constraints.

A

Publications

- J. Mocci, M. Quintavalla, C. Trestino, S. Bonora, and R. Muradore. “A Multiplatform CPU-Based Architecture for Cost-Effective Adaptive Optics Systems”. In: *IEEE Transactions on Industrial Informatics* 14.10 (2018), pp. 4431–4439

This paper reviews the AO software *PhotonLoop*, highlighting its architecture and measuring the compensation performance in a real optical test-bench. The software is written in the C++ language using the Qt and Eigen libraries. To comply with different research needs, PhotonLoop is designed to be low-latency, multi-platform and scriptable. It can be implemented on embedded architectures (e.g. ARM) and interfaced with scientific software (e.g. Matlab).

The experimental results shown in the presented thesis are carried exclusively on the PhotonLoop software or its routines.

- M. Quintavalla, J. Mocci, R. Muradore, and R. Bonora. “Comparison of adaptive free form lens with deformable mirrors”. In: *Imaging Systems and Applications*. Optical Society of America. 2019, JW2A–38

The deformable mirror compensates for aberrations by shaping its reflective surface. However, the incoming wavefront can also be corrected by altering the refractive index of an optical medium such as a deformable lens. In this paper, the static and dynamic performance of the deformable lens developed at the CNR-IFN of Padova is compared to commercially available deformable mirror solutions.

The PhotonLoop software was used to gather data, with the scripting component being exploited to automate very long identification procedures.

- N. Leonhard, R. Berlich, S. Minardi, A. Barth, S. Mauch, J. Mocci, M. Goy, M. Appelfelder, E. Beckert, and C. Reinlein. “Real-time adaptive optics testbed to investigate point-ahead angle in pre-compensation of Earth-to-GEO optical communication”. In: *Optics express* 24.12 (2016), pp. 13157–13172

The optical communication between Earth and Geostationary (GEO) satellites is negatively affected by atmospheric turbulence. This paper showcases a rapid control prototyping breadboard to pre-compensate for aberrations. Results demonstrate that this is a viable technique to enhance Earth-to-GEO optical communication.

The thesis author participated in the work to design the feedback loop of the tip-tilt woofer DM along with tuning of the tweeter DM regulator.

- S. A. Moosavi, M. Quintavalla, J. Mocci, R. Muradore, H. Saghaffar, and S. Bonora. “Improvement of coupling efficiency in free-space optical communication with a multi-actuator adaptive lens”. In: *Optics letters* 44.3 (2019), pp. 606–609
- S. A. Moosavi, M. Quintavalla, J. Mocci, R. Muradore, H. Saghaffar, and S. Bonora. “FSO communication system with multi actuator adaptive lens”. In: *2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*. IEEE. 2019, pp. 1–1

In free-space optical communication, the propagation of a laser beam through the atmosphere causes wavefront distortions that decrease the Coupling Efficiency (CE) from free space to single-mode fiber. This tremendously degrades the performance of the communication channel even in the case of weak turbulence regime. In this letter and conference paper, it is demonstrated that a multi-actuator adaptive lens working in closed loop with a wavefront sensor can strongly reduce the effect of turbulence while reducing the system complexity with respect to correction systems using deformable mirrors or liquid crystal spatial light modulators.

- M. Quintavalla, A. Ciriolo, J. Mocci, M. Negro, M. Devetta, R. Muradore, C. Vozzi, S. Bonora, and S. Stagira. “Boosting a high-energy IR OPA for Attosecond Science with high-speed adaptive deformable lenses”. In: *2018 International Conference Laser Optics (ICLO)*. IEEE. 2018, pp. 187–187
- M. Quintavalla, A. G. Ciriolo, J. Mocci, M. Negro, M. Devetta, R. Muradore, S. Stagira, S. Bonora, and C. Vozzi. “Stabilization of a high-energy optical parametric amplifier with high-speed adaptive deformable lenses”. In: *High Intensity Lasers and High Field Phenomena*. Optical Society of America. 2018, HT1A–3
- M. Negro, M. Quintavalla, J. Mocci, A. G. Ciriolo, M. Devetta, R. Muradore, S. Stagira, C. Vozzi, and S. Bonora. “Fast stabilization of a high-energy ultrafast OPA with adaptive lenses”. In: *Scientific reports* 8 (2018)
- M. Negro, M. Quintavalla, J. Mocci, R. Muradore, A. G. Ciriolo, M. Devetta, S. Stagira, C. Vozzi, and S. Bonora. “High-energy OPA stability control with adaptive lenses”. In: *Laser Resonators, Microresonators, and Beam Control XX*. Vol. 10518. International Society for Optics and Photonics. 2018, 105181B
- M. Negro, M. Quintavalla, J. Mocci, A. G. Ciriolo, M. Devetta, R. Muradore, S. Stagira, C. Vozzi, and S. Bonora. “High-speed adaptive deformable lens for boosting an high-energy optical parametric amplifier”. In: *2017 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*. IEEE. 2017, pp. 1–1

The use of fast closed-loop adaptive optics has improved the performance of optical systems since its first application. These publications demonstrate the amplitude and carrier-envelope phase stabilization of a high energy IR optical parametric amplifier devoted to Attosecond Science exploiting two high speed adaptive optical systems for the correction of static and dynamic instabilities. The exploitation of multi actuator adaptive lenses allowed for a minimal impact on the optical setup.

- P. Zhang, D. J. Wahl, J. Mocci, S. Manna, R. K. Meleppat, S. Bonora, M. V. Sarunic, E. N. Pugh Jr, and R. J. Zawadzki. “Adaptive optics with combined optical coherence tomography and scanning laser ophthalmoscopy for in vivo mouse retina imaging”. In: *Ophthalmic Technologies XXVIII*. Vol. 10474. International Society for Optics and Photonics. 2018, p. 1047427

Optical coherence tomography (OCT) and scanning laser ophthalmoscopy (SLO) are two state-of-the-art imaging technologies commonly used to study retina. Adaptive Optics (AO) methodologies enable high-fidelity correction of ocular aberrations, resulting in improved resolution and sensitivity for both SLO and OCT systems. Here we present work integrating OCT into a previously described mouse retinal AO-SLO system, allowing simultaneous reflectance and fluorescence imaging. The new system allows simultaneous data acquisition of AO-SLO and AO-OCT, facilitating registration and comparison of data from both modalities. The system has data acquisition speed of 200 kHz A-scans/pixel, and high volumetric resolution.

- D. J. Wahl, P. Zhang, J. Mocci, M. Quintavalla, R. Muradore, Y. Jian, S. Bonora, M. V. Sarunic, and R. J. Zawadzki. “Adaptive optics in the mouse eye: wavefront sensing based vs. image-guided aberration correction”. In: *Biomedical optics express* 10.9 (2019), pp. 4757–4774

Adaptive Optics (AO) is required to achieve diffraction limited resolution in many real-life imaging applications in biology and medicine. AO is essential to guarantee high fidelity visualization of cellular structures for retinal imaging by correcting ocular aberrations. Aberration correction for mouse retinal imaging by direct wavefront measurement has been demonstrated with great success. However, for mouse eyes, the performance of the wavefront sensor (WFS) based AO can be limited by several factors including non-common path errors, wavefront reconstruction errors, and an ill-defined reference plane. Image-based AO can avoid these issues at the cost of algorithmic execution time. Furthermore, image-based approaches can provide improvements to compactness, accessibility, and even the performance of AO systems. Here, we demonstrate the ability of image-based AO to provide comparable aberration correction and image resolution to the conventional Shack-Hartmann WFS-based AO approach. The residual wavefront error of the mouse eye was monitored during a wavefront sensorless optimization to allow comparison with classical AO. This also allowed us to improve the performance of our AO system for small animal retinal imaging.

- P. Zhang, J. Mocci, D. J. Wahl, R. K. Meleppat, S. K. Manna, M. Quintavalla, R. Muradore, M. V. Sarunic, S. Bonora, E. N. Pugh Jr, et al. “Effect of a contact lens on mouse retinal in vivo imaging: Effective focal length changes and monochromatic aberrations”. In: *Experimental eye research* 172 (2018), pp. 86–93

For in vivo mouse retinal imaging, especially with Adaptive Optics instruments, application of a contact lens is desirable, as it allows maintenance of cornea hydration and helps to prevent cataract formation during lengthy imaging sessions.

However, since the refractive elements of the eye (cornea and lens) serve as the objective for most in vivo retinal imaging systems, the use of a contact lens, even with 0 Dpt. refractive power, can alter the system's optical properties. In this investigation we examined the effective focal length change and the aberrations that arise from use of a contact lens. First, focal length changes were simulated with a Zemax mouse eye model. Then ocular aberrations with and without a 0 Dpt. contact lens were measured with a Shack-Hartmann wavefront sensor (SHWS) in a customized AO-SLO system. Total RMS wavefront errors were measured for two groups of mice (14-month, and 2.5-month-old), decomposed into 66 Zernike aberration terms, and compared. These data revealed that vertical coma and spherical aberrations were increased with use of a contact lens in our system. Based on the ocular wavefront data we evaluated the effect of the contact lens on the imaging system performance as a function of the pupil size. Both RMS error and Strehl ratios were quantified for the two groups of mice, with and without contact lenses, and for different input beam sizes. These results provide information for determining optimum pupil size for retinal imaging without adaptive optics, and raise critical issues for design of mouse optical imaging systems that incorporate contact lenses.

- J. Mocci, M. Cua, S. Lee, Y. Jian, P. Pozzi, M. Quintavalla, C. Trestino, H. Verstraete, D. Wahl, R. Muradore, et al. "Wavefront control with a multi-actuator adaptive Lens in imaging applications". In: *Adaptive Optics and Wavefront Control for Biological Systems III*. Vol. 10073. International Society for Optics and Photonics. 2017, p. 1007302

The use of adaptive lenses instead of deformable mirrors can simplify the implementation of an adaptive optics system. The recently introduced Multi-actuator Adaptive Lens (MAL) can be used in closed loop with a wavefront sensor to correct for time-variant wavefront aberrations. The MAL can guarantee a level of correction and a response time similar to the ones obtained with deformable mirrors. The adaptive lens is based on the use of piezoelectric actuators and, without any obstruction or electrodes in the clear aperture, can guarantee a fast response time, less than 10ms. Our tests show that the MAL can be used both in combination with a wavefront sensor in a classical adaptive optics closed loop, or in a wavefront sensorless configuration. The latter has allowed us to design more compact and simple imaging systems for different microscopy platforms. We will show that the Multi-actuator Adaptive Lens has been successfully used for in-vivo OCT ophthalmic imaging in both mice and humans, as well as confocal and two photon microscopy. We tested and compared different optimization strategies such as coordinate search and the DONE algorithm. The results suggest that the MAL optimization can correct for eye aberrations with a pupil of 5mm or sample induced aberrations in microscopy.

- M. Quintavalla, J. Mocci, R. Muradore, A. Corso, and S. Bonora. "Adaptive optics on small astronomical telescope with multi-actuator adaptive lens". In: *Free-Space Laser Communication and Atmospheric Propagation XXX*. Vol. 10524. International Society for Optics and Photonics. 2018, p. 1052414

Adaptive Optics (AO) is a key technology for ground-based astronomical telescopes, allowing to overcome the limits imposed by atmospheric turbulence and obtain high resolution images. This technique however, has not been developed for small size telescopes, because of its high cost and complexity. We realized an AO system based on a Multi-actuator Adaptive Lens and a Shack-Hartmann wavefront sensor (WFS), allowing for a great compactness and simplification of the optical design. The system was integrated on a 11 telescope and controlled by a consumer-grade laptop allowing to perform Closed-Loop AO correction up to 400 Hz.

References

- [1] G. Agapito et al. “Automatic Tuning of the Internal Position Control of an Adaptive Secondary Mirror”. In: *European Journal of Control* 17.3 (2011), pp. 273–289. ISSN: 09473580.
- [2] B. Anderson and J. Moore. *Optimal control: linear quadratic methods*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1990.
- [3] B. D. O. Anderson and J. B. Moore. *Optimal Control: Linear Quadratic Methods (Dover Books on Engineering)*. Dover Publications, 2007. ISBN: 1558505113.
- [4] B. D. Anderson and J. B. Moore. “Optimal filtering”. In: *Englewood Cliffs* 21 (1979), pp. 22–95.
- [5] L. Andrews. *Field Guide to Atmospheric Optics*. Field Guide Series. SPIE Press, 2004. ISBN: 9780819453181.
- [6] L. Andrews, R. Phillips, and C. Hopen. *Laser Beam Scintillation with Applications*. SPIE Press monograph. Society of Photo Optical, 2001. ISBN: 9780819441034.
- [7] H. W. Babcock. “The possibility of compensating astronomical seeing”. In: *Publications of the Astronomical Society of the Pacific* 65.386 (1953), pp. 229–236.
- [8] A. G. Basden. “Investigation of POWER8 processors for astronomical adaptive optics real-time control”. In: *Mon. Not. R. Astron. Soc.* 452.2 (July 2015), pp. 1694–1701.
- [9] A. G. Basden and R. M. Myers. “The Durham adaptive optics real-time controller: capability and Extremely Large Telescope suitability”. In: *Monthly Notices of the Royal Astronomical Society* 424.2 (June 2012), pp. 1483–1494.
- [10] J. Beck and J. P. Bos. “Open source acceleration of wave optics simulations on energy efficient high-performance computing platforms”. In: *Long-Range Imaging II* 10204 (2017), 102040F.
- [11] J. M. Beckers. “Adaptive Optics for Astronomy: Principles, Performance, and Applications”. In: *Annual Review of Astronomy and Astrophysics* 31.1 (Sept. 1993), pp. 13–62.
- [12] S. Bonora et al. “Design of deformable mirrors for high power lasers”. In: *High Power Laser Science and Engineering* 4 (2016), p. 16.
- [13] S. Bonora et al. “Fast wavefront active control by a simple DSP-Driven deformable mirror”. In: *Review of Scientific Instruments* (2006).
- [14] M. J. Booth. “Adaptive optical microscopy: the ongoing quest for a perfect image”. In: *Light: Science & Applications* 3.4 (Apr. 2014), e165.
- [15] D. Burke et al. “Adaptive optics for single molecule switching nanoscopy”. In: *2015 Conference on Lasers and Electro-Optics (CLEO)*. May 2015, pp. 1–2.
- [16] M. Carraro, M. Munaro, and E. Menegatti. “Cost-efficient RGB-D smart camera for people detection and tracking”. In: *Journal of Electronic Imaging* 25.4 (2016), p. 041007.
- [17] A. Cartella et al. “Pulse shaping in the mid-infrared by a deformable mirror”. In: *Optics Letters* 39.6 (Mar. 2014), p. 1485.

- [18] G. Chanan. “Principles of wavefront sensing and reconstruction”. In: *Proceedings: Summer School on Adaptive Optics; Center for Adaptive Optics (CfAO): Santa Cruz, CA, USA* (2000), pp. 5–40.
- [19] C. Chang et al. “Easily implementable field programmable gate array-based adaptive optics system with state-space multichannel control”. In: *Rev. Sci. Instrum.* 84.9 (2013), p. 095112.
- [20] S. Chen et al. “A compact multi-core CPU based adaptive optics real-time controller”. In: *7th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Large Mirrors and Telescopes*. Ed. by W. Jiang, M. K. Cho, and F. Wu. SPIE-Intl Soc Optical Eng, Sept. 2014.
- [21] A. Chiuso, R. Muradore, and E. Fedrigo. “Adaptive optics systems: a challenge for closed loop subspace identification”. In: *Proceedings of the 26th American Control Conference ACC*. 2007.
- [22] A. Chiuso, R. Muradore, and E. Marchetti. “Dynamic Calibration of Adaptive Optics Systems: A System Identification Approach”. In: *IEEE Transactions on Control Systems Technology* 18.3 (May 2010), pp. 705–713.
- [23] D. Coburn, D. Garnier, and J. Dainty. “A single star SCIDAR system for profiling atmospheric turbulence”. In: *Remote Sensing*. International Society for Optics and Photonics. 2005, pp. 59810D–59810D.
- [24] J. M. Conan, G. Rousset, and P. Y. Madec. “Wave-front temporal spectra in high-resolution imaging through turbulence”. In: *J. Opt. Soc. Am. A* 12 (1995), pp. 1559–1570.
- [25] C. Correia et al. “Accounting for Mirror Dynamics in Optimal Adaptive Optics Control”. In: *Proceedings of the European Control Conference*. 2009, pp. 3515–3520.
- [26] F. Dias et al. “Hardware, design and implementation issues on a FPGA-based smart camera”. In: *2007 First ACM/IEEE International Conference on Distributed Smart Cameras*. IEEE. 2007, pp. 20–26.
- [27] N. Doelman et al. “Adaptive and Real Time Optimal Control for Adaptive Optics Systems”. In: *Proceedings of the European Control Conference*. 2009, pp. 3527–3532.
- [28] E. Fedrigo, R. Muradore, and D. Zilio. “High performance adaptive optics system with fine tip/tilt control”. In: *Control Engineering Practice* 17.1 (Jan. 2009), pp. 122–135.
- [29] E. Fedrigo et al. “Real-time Control of ESO Adaptive Optics Systems (Echtzeitsteuerung der ESO Adaptive Optik Systeme)”. In: *at - Automatisierungstechnik* 53.10/2005 (Jan. 2005).
- [30] F. Ferreira et al. “COMPASS: An efficient GPU-based simulation software for adaptive optics systems”. In: *Proceedings - 2018 International Conference on High Performance Computing and Simulation, HPCS 2018* (2018), pp. 180–187.
- [31] J.-P. Folcher et al. “Adaptive Optics Feedback Control”. In: *EAS Publications Series* 59 (2013). Ed. by D. Mary, C. Theys, and C. Aime, pp. 93–130.
- [32] E. Fornasini. *Appunti di teoria dei sistemi*. Progetto Editore, 2011. ISBN: 978-8896477328.

- [33] G. F. Franklin, J. D. Powell, and M. L. Workman. *Digital Control of Dynamic Systems (3rd Edition)*. Addison-Wesley, 1997. ISBN: 0201820544.
- [34] G. H. Golub and C. F. V. Loan. *Matrix Computations*. Johns Hopkins University Press, 2012. ISBN: 1421408597.
- [35] C. Gonz et al. “Using GPUs to Speed up a Tomographic Reconstructor Based on Machine Learning”. In: 527 (2017).
- [36] J. Goodman. *Introduction to Fourier optics*. Roberts & Company Publishers, 2005.
- [37] G. C. Goodwin, J. A. D. Doná, and M. M. Seron. *Constrained Control and Estimation*. Springer London, 2005.
- [38] F. H. D. Guaracy, D. L. F. da Silva, and L. H. C. Ferreira. “On the plant augmentation by integrators in the discrete-time LQG/LTR control”. In: *2014 IEEE Conference on Control Applications (CCA)*. IEEE, Oct. 2014.
- [39] G. Guennebaud and B. J. et al. *Eigen v3*. 2010.
- [40] J. Hardy. *Adaptive optics for astronomical telescopes*. Oxford University Press, USA, 1998.
- [41] K. Hinnen, M. Verhaegen, and N. Doelman. “A Data-driven H_2 -optimal Control approach for Adaptive Optics”. In: *IEEE Transactions on Control Systems Technology* 16.3 (2008), pp. 381–395.
- [42] K. Hinnen, M. Verhaegen, and N. Doelman. “Exploiting the spatiotemporal correlation in adaptive optics using data-driven H_2 -optimal control”. In: *Journal of the Optical Society of America A* 24.6 (2007), pp. 1714–1725.
- [43] K. Hinnen, N. Doelman, and M. Verhaegen. “ \mathcal{H}_2 -optimal control of an adaptive optics system: part II, closed-loop controller design”. In: *Astronomical Adaptive Optics Systems and Applications II*. Ed. by R. K. Tyson and M. Lloyd-Hart. SPIE, Aug. 2005.
- [44] K. Hinnen, M. Verhaegen, and N. Doelman. “A Data-Driven \mathcal{H}_2 -Optimal Control Approach for Adaptive Optics”. In: *IEEE Transactions on Control Systems Technology* 16.3 (2008), pp. 381–395.
- [45] V. Hutterer, R. Ramlau, and I. Shatokhina. “Real-time adaptive optics with pyramid wavefront sensors: a theoretical analysis of the pyramid sensor model”. In: *arXiv preprint arXiv:1810.00682* (2018).
- [46] N. Ji. “Adaptive optical fluorescence microscopy”. In: *Nature Methods* 14.4 (Mar. 2017), pp. 374–380.
- [47] J. Jia et al. “Design of parallel wave-front restoration algorithm based on FPGA”. In: *Optik* 176.July 2018 (2019), pp. 168–174. ISSN: 00304026.
- [48] Y. Jian et al. “Lens-based wavefront sensorless adaptive optics swept source OCT”. In: *Scientific Reports* 6.1 (June 2016).
- [49] M. Kasper et al. “Fast calibration of high-order adaptive optics systems”. In: *Journal of the Optical Society of America A* 21.6 (June 2004), p. 1004.
- [50] F. Kong, M. C. Polo, and A. Lambert. “Centroid estimation for a Shack–Hartmann wavefront sensor based on stream processing”. In: *Applied Optics* 56.23 (2017), p. 6466. ISSN: 1559-128X.
- [51] M. V. Konnik and J. D. Dona. “Feasibility of Constrained Receding Horizon Control Implementation in Adaptive Optics”. In: *IEEE Transactions on Control Systems Technology* 23.1 (Jan. 2015), pp. 274–289.

- [52] M. V. Konnik, J. D. Doná, and J. S. Welsh. “On application of constrained receding horizon control in astronomical adaptive optics”. In: *Adaptive Optics Systems III*. Ed. by B. L. Ellerbroek, E. Marchetti, and J.-P. Véran. SPIE, Sept. 2012.
- [53] A. V. Kudryashov and V. V. Samarkin. “Control of high power CO₂ laser beam by adaptive optical elements”. In: *Optics Communications* 118.3-4 (July 1995), pp. 317–322.
- [54] C. Kulcsár et al. “Minimum variance prediction and control for adaptive optics”. In: *Automatica* 48.9 (Sept. 2012), pp. 1939–1954.
- [55] C. Kulcsár et al. “Optimal control, observers and integrators in adaptive optics”. In: *Optics Express* 14.17 (Aug. 2006), p. 7464.
- [56] C. Kulcsár et al. “Optimal control, observers and integrators in adaptive optics”. In: *Optics express* 14.17 (2006), pp. 7464–7476.
- [57] M. Lainée et al. “A GPU based RTC for E-ELT Adaptive optics : Real Time Controller prototype”. In: January (2018).
- [58] Y. Landau. *Digital control systems: design, identification and implementation*. Berlin New York: Springer, 2006. ISBN: 978-1-84628-056-6.
- [59] B. Le Roux et al. “Optimal control law for classical and multiconjugate adaptive optics”. In: *Journal of the Optical Society of America A* 21.7 (2004), pp. 1261–1276.
- [60] S.-H. Lee and C.-S. Yang. “A real time object recognition and counting system for smart industrial camera sensor”. In: *IEEE Sensors Journal* 17.8 (2017), pp. 2516–2523.
- [61] N. Leonhard et al. “Real-time adaptive optics testbed to investigate point-ahead angle in pre-compensation of Earth-to-GEO optical communication”. In: *Optics express* 24.12 (2016), pp. 13157–13172.
- [62] J. Li, J. Sun, and W. Chen. “Performance of atmospheric coherent laser receiver and bandwidth for adaptive optics system”. In: *2015 IEEE International Conference on Space Optical Systems and Applications (ICSOS)*. Oct. 2015, pp. 1–1.
- [63] D. Looze. “Structure and Approximation of LQG controllers Based on Hybrid AO System Model”. In: *Proceedings of the European Control Conference*. 2009, pp. 3521–3526.
- [64] D. Looze et al. “Optimal compensation and implementation for adaptive optics systems”. In: *Experimental Astronomy* 15.2 (2003), pp. 67–88.
- [65] V. N. Mahajan. “Zernike annular polynomials for imaging systems with annular pupils”. In: *JOSA* 71.1 (1981), pp. 75–85.
- [66] S. Mauch and J. Reger. “Real-Time Spot Detection and Ordering for a Shack–Hartmann Wavefront Sensor With a Low-Cost FPGA”. In: *IEEE Transactions on Instrumentation and Measurement* 63.10 (Oct. 2014), pp. 2379–2386.
- [67] S. Mauch et al. “FPGA-accelerated adaptive optics wavefront control”. In: *MEMS Adaptive Optics VIII*. Ed. by T. G. Bifano, J. Kubby, and S. Gigan. SPIE-Intl Soc Optical Eng, Mar. 2014.
- [68] S. Mauch and J. Reger. “Real-time implementation of the spiral algorithm for Shack-Hartmann wavefront sensor pattern sorting on an FPGA”. In:

- Measurement: Journal of the International Measurement Confederation* 92 (2016), pp. 63–69. ISSN: 02632241.
- [69] J. Mocci et al. “Wavefront control with a multi-actuator adaptive Lens in imaging applications”. In: *Adaptive Optics and Wavefront Control for Biological Systems III*. Vol. 10073. International Society for Optics and Photonics. 2017, p. 1007302.
- [70] J. Mocci et al. “A Multiplatform CPU-Based Architecture for Cost-Effective Adaptive Optics Systems”. In: *IEEE Transactions on Industrial Informatics* 14.10 (Oct. 2018), pp. 4431–4439.
- [71] J. Mocci et al. “A Multiplatform CPU-Based Architecture for Cost-Effective Adaptive Optics Systems”. In: *IEEE Transactions on Industrial Informatics* 14.10 (2018), pp. 4431–4439.
- [72] J. Mocci et al. “Analysis of Horizontal Atmospheric Turbulence by using a Shack-Hartmann Wavefront Sensor”. In: AOIM XI, Murcia. 2018.
- [73] J. Mocci et al. “PI-Shaped LQG Control Design for Adaptive Optics Systems”. In: *Control Engineering Practice* (2020).
- [74] S. A. Moosavi et al. “FSO communication system with multi actuator adaptive lens”. In: *2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*. IEEE. 2019, pp. 1–1.
- [75] S. A. Moosavi et al. “Improvement of coupling efficiency in free-space optical communication with a multi-actuator adaptive lens”. In: *Optics letters* 44.3 (2019), pp. 606–609.
- [76] R. Muradore, E. Fedrigo, and C. Correia. “LQ control design for adaptive optics systems based on MIMO identified model”. In: *Advances in Adaptive Optics II*. Ed. by B. L. Ellerbroek and D. B. Calia. SPIE-Intl Soc Optical Eng, June 2006.
- [77] R. Muradore et al. “An application of adaptive techniques to vibration rejection in adaptive optics systems”. In: *Control Engineering Practice* 32 (2014), pp. 87–95.
- [78] R. Muradore et al. “Multi-input multi-output identification for control of adaptive optics systems”. In: *Adaptive Optics Systems IV*. Ed. by E. Marchetti, L. M. Close, and J.-P. Véran. SPIE-Intl Soc Optical Eng, Aug. 2014.
- [79] M. Negro et al. “Fast stabilization of a high-energy ultrafast OPA with adaptive lenses”. In: *Scientific reports* 8 (2018).
- [80] M. Negro et al. “High-energy OPA stability control with adaptive lenses”. In: *Laser Resonators, Microresonators, and Beam Control XX*. Vol. 10518. International Society for Optics and Photonics. 2018, 105181B.
- [81] M. Negro et al. “High-speed adaptive deformable lens for boosting an high-energy optical parametric amplifier”. In: *2017 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*. IEEE. 2017, pp. 1–1.
- [82] R. Noll. “Zernike polynomials and atmospheric turbulence”. In: *JOSA* 66.3 (1976), pp. 207–211.

- [83] G. Parry. “Measurement of atmospheric turbulence induced intensity fluctuations in a laser beam”. In: *Journal of Modern Optics* 28.5 (1981), pp. 715–728.
- [84] R. Paschall and D. Anderson. “Linear quadratic Gaussian control of a deformable mirror adaptive optics system with time-delayed measurements”. In: *Applied optics* 32.31 (1993), pp. 6347–6358.
- [85] D. Perret et al. “Bridging FPGA and GPU technologies for AO real-time control”. In: *Adaptive Optics Systems V* 9909 (2016), p. 99094M.
- [86] C. Petit et al. “First laboratory validation of vibration filtering with LQG control law for Adaptive Optics”. In: *Optics Express* 16.1 (2008), p. 87.
- [87] J. Pilar et al. “Design and optimization of an adaptive optics system for a high-average-power multi-slab laser (HiLASE)”. In: *Applied Optics* 53.15 (May 2014), p. 3255.
- [88] M. Pircher and R. J. Zawadzki. “Review of adaptive optics OCT (AO-OCT): principles and applications for retinal imaging [Invited]”. In: *Biomedical Optics Express* 8.5 (Apr. 2017), p. 2536.
- [89] M. C. Polo, F. Kong, and A. Lambert. “FPGA Implementations of Low Latency Centroiding Algorithms for Adaptive Optics”. In: 2018 (2018), OTh3E.3.
- [90] L. Poyneer and J.-P. Véran. “Predictive wavefront control for adaptive optics with arbitrary control loop delays”. In: *Journal of the Optical Society of America A* 25.7 (June 2008), p. 1486.
- [91] D. Quaglia and R. Muradore. “Communication-Aware Bandwidth-Optimized Predictive Control of Motor Drives in Electric Vehicles”. In: *IEEE Transactions on Industrial Electronics* 63.9 (Sept. 2016), pp. 5602–5611.
- [92] M. Quintavalla et al. “Stabilization of a high-energy optical parametric amplifier with high-speed adaptive deformable lenses”. In: *High Intensity Lasers and High Field Phenomena*. Optical Society of America. 2018, HT1A–3.
- [93] M. Quintavalla et al. “Adaptive optics on small astronomical telescope with multi-actuator adaptive lens”. In: *Free-Space Laser Communication and Atmospheric Propagation XXX*. Vol. 10524. International Society for Optics and Photonics. 2018, p. 1052414.
- [94] M. Quintavalla et al. “Adaptive optics on small astronomical telescope with multi-actuator adaptive lens”. In: *Free-Space Laser Communication and Atmospheric Propagation XXX*. Vol. 10524. International Society for Optics and Photonics. 2018, p. 1052414.
- [95] M. Quintavalla et al. “Boosting a high-energy IR OPA for Attosecond Science with high-speed adaptive deformable lenses”. In: *2018 International Conference Laser Optics (ICLO)*. IEEE. 2018, pp. 187–187.
- [96] M. Quintavalla et al. “Comparison of adaptive free form lens with deformable mirrors”. In: *Imaging Systems and Applications*. Optical Society of America. 2019, JW2A–38.
- [97] R. Ragazzoni et al. “A smart fast camera”. In: *Ground-based Instrumentation for Astronomy*. Vol. 5492. International Society for Optics and Photonics. 2004, pp. 121–127.

- [98] H.-F. Raynaud et al. “The control switching adapter: a practical way to ensure bumpless switching between controllers while AO loop is engaged”. In: *Adaptive Optics Systems V*. Ed. by E. Marchetti, L. M. Close, and J.-P. Véran. SPIE, July 2016.
- [99] F. Roddier. *Adaptive Optics in Astronomy*. Cambridge University Press, 1999.
- [100] F. Roddier. *Adaptive Optics in Astronomy*. Cambridge University Press, 2004.
- [101] M. Roggemann and B. Welsh. *Imaging through Turbulence*. CRC Press, 1996.
- [102] F. L. Rosa, J. G. Marichal-Hernandez, and J. M. Rodriguez-Ramos. “Wavefront phase recovery using graphic processing units (GPUs)”. In: *Optics in Atmospheric Propagation and Adaptive Systems VII* 5572 (2004), p. 262.
- [103] G. Rousset et al. “First diffraction-limited astronomical images with adaptive optics”. In: *Astronomy and Astrophysics* 230 (1990), pp. L29–L32.
- [104] B. L. Roux et al. “Optimal control law for classical and multiconjugate adaptive optics”. In: *Journal of the Optical Society of America A* 21.7 (July 2004), p. 1261.
- [105] D. Ruscio. “Discrete LQ optimal control with integral action: A simple controller on incremental form for MIMO systems”. In: *Modeling, Identification and Control: A Norwegian Research Bulletin* 33 (Apr. 2012).
- [106] M. Seeds and D. Backman. *Horizons: Exploring the Universe*. Cengage Learning, 2016. ISBN: 9781305960961.
- [107] A. Sevin et al. “Enabling technologies for GPU driven adaptive optics real-time control”. In: *Adaptive Optics Systems IV*. Ed. by E. Marchetti, L. M. Close, and J.-P. Véran. SPIE-Intl Soc Optical Eng, July 2014.
- [108] D. Song, X. Li, and Z. Peng. “Design and simulation of adaptive optics controller based on mixed sensitivity \mathcal{H}_∞ control”. In: *8th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Large Mirrors and Telescopes*. Ed. by M. K. Cho and B. Fan. SPIE, Oct. 2016.
- [109] H. Song et al. “Controller Design for a High-Sampling-Rate Closed-Loop Adaptive Optics System with Piezo-Driven Deformable Mirror”. In: *European Journal of Control* 17.3 (2011), pp. 290–301. ISSN: 09473580.
- [110] A. Surendran et al. “Development of a scalable generic platform for adaptive optics real time control”. In: *International Conference on Optics and Photonics 2015*. Ed. by K. Bhattacharya. SPIE-Intl Soc Optical Eng, June 2015.
- [111] V. I. Tatarskii. *Wave Propagation in Turbulent Medium*. McGraw-Hill, 1961.
- [112] The Qt Company. *Qt*. Version 5.5. 2015.
- [113] M. Thier et al. “Low-Latency Shack-Hartmann Wavefront Sensor Based on an Industrial Smart Camera”. In: *IEEE Transactions on Instrumentation and Measurement* 62.5 (May 2013), pp. 1241–1249.
- [114] M. Thier et al. “Low-latency Shack-Hartmann wavefront sensor based on an industrial smart camera”. In: *IEEE transactions on instrumentation and measurement* 62.5 (2012), pp. 1241–1249.

- [115] S. Thomas et al. “Comparison of centroid computation algorithms in a Shack-Hartmann sensor”. In: *Monthly Notices of the Royal Astronomical Society* 371.1 (2006), pp. 323–336. ISSN: 00358711.
- [116] R. Tyson. *Adaptive Optics Engineering Handbook (Optical Science and Engineering)*. CRC Press, 1999. ISBN: 0824782755.
- [117] R. Tyson. *Principles of adaptive optics, Fourth Edition*. Boca Raton, FL: CRC Press, 2015. ISBN: 978-1439808580.
- [118] R. K. Tyson. “Theory and operation of a robust controller for a compact adaptive optics system”. In: *Optical Engineering* 43.12 (Dec. 2004), p. 2912.
- [119] Y. Ukidave et al. “Performance of the NVIDIA Jetson TK1 in HPC”. In: *2015 IEEE International Conference on Cluster Computing*. IEEE. 2015, pp. 533–534.
- [120] V. Venugopalan. “Evaluating latency and throughput bound acceleration of FPGAs and GPUs for adaptive optics algorithms”. In: *2014 IEEE High Performance Extreme Computing Conference, HPEC 2014* 1 (2014), pp. 1–6.
- [121] A. Vyas, M. B. Roopashree, and B. R. Prasad. “Optimization of existing centroiding algorithms for Shack Hartmann sensor”. In: *arXiv preprint arXiv:0908.4328* (2009).
- [122] D. J. Wahl et al. “Adaptive optics in the mouse eye: wavefront sensing based vs. image-guided aberration correction”. In: *Biomedical optics express* 10.9 (2019), pp. 4757–4774.
- [123] D. Winker. “Effect of a finite outer scale on the Zernike decomposition of atmospheric optical turbulence”. In: *Optical Society of America, Journal, A: Optics and Image Science* 8 (1991), pp. 1568–1573.
- [124] J. C. Wyant and K. Creath. “Basic Wavefront Aberration Theory for Optical Metrology”. In: *Applied Optics and Optical Engineering, Volume XI*. Ed. by R. R. Shannon and J. C. Wyant. Vol. 11. 1992, p. 2.
- [125] J. Ye et al. “Comparative assessment of orthogonal polynomials for wavefront reconstruction over the square aperture”. In: *JOSA A* 31.10 (2014), pp. 2304–2311.
- [126] C. Yu and M. Verhaegen. “Structured Modeling and Control of Adaptive Optics Systems”. In: *IEEE Transactions on Control Systems Technology* 26.2 (Mar. 2018), pp. 664–674.
- [127] P. Zhang et al. “Effect of a contact lens on mouse retinal in vivo imaging: Effective focal length changes and monochromatic aberrations”. In: *Experimental eye research* 172 (2018), pp. 86–93.
- [128] P. Zhang et al. “Adaptive optics with combined optical coherence tomography and scanning laser ophthalmoscopy for in vivo mouse retina imaging”. In: *Ophthalmic Technologies XXVIII*. Vol. 10474. International Society for Optics and Photonics. 2018, p. 1047427.
- [129] Z. Zhao et al. “Ecrt: An edge computing system for real-time image-based object tracking”. In: *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. ACM. 2018, pp. 394–395.