FOCAL PLANE WAVEFRONT SENSING AND CONTROL FOR VARIOUS OPTICAL SYSTEMS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Duan Li

August 2025

FOCAL PLANE WAVEFRONT SENSING AND CONTROL FOR VARIOUS OPTICAL SYSTEMS

Duan Li, Ph.D.

Cornell University 2025

High-precision optical systems are essential in fields ranging from autonomous manufacturing to astronomical imaging. However, their performance is often limited by optical aberrations arising from misalignments, environmental disturbances, and internal system imperfections. This thesis investigates the principles and practical implementations of wavefront sensing and control — primarily focal-plane techniques — to enhance optical system performance in two key domains: automated alignment and high-contrast exoplanet imaging with the Gemini Planet Imager.

In the context of automated optical alignment, we developed and tested a novel alignment framework for a double off-axis parabolic mirror system — an inherently challenging configuration due to its sensitivity to aberrations and nonlinear misalignment dynamics. Leveraging focal-plane wavefront sensing to eliminate non-common-path errors, and applying model-based nonlinear state estimation and control, we demonstrate micron-level alignment precision. A key discovery was the multi-state coupling effect, in which different misalignment states interact in ways that interfere with estimator performance. This phenomenon was further analyzed through observability analysis.

For high-contrast imaging, we contributed to the ongoing upgrade of the Gem-

ini Planet Imager to improve its capability to directly image and characterize exoplanets. In the adaptive optics control pipeline, we established design tools based on system stability metrics and error transfer function analysis. We further implemented a focal-plane wavefront control algorithm based on classical speckle nulling to effectively mitigate speckle noise caused by post-coronagraph aberrations and internal optical imperfections — a historical performance limitation of GPI. These results represent a significant improvement over GPI's baseline and will be integrated into its next-generation upgrade to enable detection of lower-mass exoplanets at smaller inner working angles.

By combining optical engineering, estimation theory, and control design, this thesis highlights the versatility and critical importance of focal-plane wavefront control in advancing both laboratory-based and astronomical optical systems.

BIOGRAPHICAL SKETCH

Duan Li was born in Nanjing, China and developed an early interest in optical instrumentation. She completed her undergraduate studies with double major in Mechanical Engineering and Physics at Johns Hopkins University, where she gained foundational knowledge in optics, astronomy and hardware.

In 2018, Duan began her PhD journey at Cornell University, joining the Space Imaging and Optical Systems Lab. Her doctoral research focuses on focal plane wavefront sensing and control for optical systems. Over the course of her PhD, Duan has contributed to innovating automated optical alignment methods and upgrading Gemini Planet Imager to higher contrast through publications and conference presentations.

Outside of research, Duan enjoys wildlife photography and dancing. She currently works on auto-focus and optical image stabilization as a camera systems engineer in industry and aspires to advance camera technologies.

This document is dedicated to all Cornell graduate students.

ACKNOWLEDGEMENTS

First and foremost, I express my deepest gratitude to my advisor, Dmitry Savransky, for their unwavering guidance, insightful advice, and relentless support throughout my doctoral journey. Their expertise and encouragement have been instrumental in shaping my research and professional growth.

I extend my heartfelt thanks to my dissertation committee members, Douglas MacMartin and James Lloyd, and the GPI collaboration, for the time they dedicated to reviewing my work, valuable feedbacks and thought-provoking discussions. Their perspectives have greatly enriched my research.

I am also deeply thankful to my colleagues and lab members, whose camaraderie, collaboration, and constructive feedback have made this journey intellectually stimulating and enjoyable. I am grateful for their support, expertise and friendship.

To my cat, family and friends, I owe immeasurable gratitude for their unconditional love, patience, and encouragement. Thank you for believing in me, even during moments of self-doubt, and for being my constant source of strength.

Finally, I want to acknowledge Cornell University and NRC Canada for providing the financial and institutional support that made this research possible.

Completing this PhD has been a challenging yet rewarding experience, and I am grateful to everyone who contributed to this journey, directly or indirectly. Thank you for walking with me on this path.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Wavefront sensing and control

Modern optical applications require high precision and resolution, but wavefront aberrations - arising from optical imperfections, misalignment, atmospheric turbulence, thermal gradients and vibrations - can severely degrade optical system performance. Moreover, these aberrations are often dynamic, thus requiring real-time corrections. Wavefront sensing and control addresses these challenges by manipulating the phase and amplitude of light, and provides improved image quality, enhanced contrast, and optimized system performance.

Wavefront sensing and control has transformative impact across disciplines: In astronomy, wavefront control is prominently featured in adaptive optics systems. Ground-based astronomical telescopes use adaptive optics with deformable mirrors to counteract atmospheric-turbulence-induced wavefront errors in real time, achieving near-diffraction-limited resolution for deep-sky observations. Combined with coronagraphs, specialized adaptive optics systems with high actuator density and fast control loops make it possible to detect extremely faint companions, e.g. exoplanets, close to bright stars[40, 97, 77]. In ophthalmology, wavefront technology helps map wavefront errors in the patient's eyes, guide treatments such as wavefront-guided LASIK and cataract surgery, improve post-operative visual acuity and reduce side effects such as halos or glare[120, 69, 103]. In manufacturing, adaptive optics enhance three-dimensional fabrication with ultrafast lasers by modulating the phase, ampli-

tude and polarization of the laser beam, tailoring the focal intensity distribution, and enabling parallelization to reduce processing times[99]. In lithography, because small phase errors can lead to significant pattern distortions, high-resolution wavefront manipulators are designed to minimize aberrations, and maintain precise overlay and focus performance when imaging nanoscale features on semiconductor wafers[110]. In free-space optical communications, real-time wavefront control helps manage beam dispersion, beam wander and scintillations, improve signal quality and expand data-carrying capacity[118]. In biomedical imaging, wavefront control is used to correct for refractive index changes in tissue, thereby improving the quality of deep-tissue imaging and minimizing the phototoxic effects often associated with high-intensity illumination[9].

Wavefront sensing and control is achieved through a combination of three key components: wavefront sensors, adaptive elements and control algorithms. The design of robust wavefront sensing and control architectures must balance hardware complexity, computational demands, and the specific temporal and spatial requirements of each application.

Wavefront sensors detect phase distortions by translating the incoming light into measurable intensity signals. Wavefront sensing techniques can be classified into two main categories: pupil plane sensing and focal plane sensing. Pupil plane wavefront sensors directly measure the local slope or curvature of the wavefront at or near the entrance pupil. Common examples include: the Shack–Hartmann wavefront sensor, widely used in astronomy and ophthalmology due to its simplicity and robustness[88, 95]; the pyra-

mid wavefront sensor, effective in extreme adaptive optics because of its high sensitivity and resolution[21]; curvature sensing, optimal for detecting low-order aberrations at defocused pupil planes[98]; and interferometric techniques such as Twyman–Green or Fizeau interferometers offer precise and highly accurate wavefront measurements essential for optical metrology and surface characterization[70]. In contrast, focal plane sensing techniques estimate the wavefront directly from intensity distributions in the image plane, rather than from phase derivatives at the pupil. This approach eliminates non-common-path errors, thereby improving overall system efficiency, reducing hardware complexity, and delivering correction better aligned with the science instrument's performance metric. Commonly implemented focal-plane sensing algorithms include the Gerchberg–Saxton algorithm, which reconstructs wavefront phase through iterative Fourier domain calculations[28]; phase diversity, which utilizes multiple defocused images to computationally retrieve wavefront phase[32]; electric field conjugation, which aims to eliminate unwanted starlight in the focal plane by directly estimating and manipulating the complex electric field; and stochastic parallel gradient descent[29, 37], a robust optimization-based method that iteratively improves wavefront quality by optimizing focal-plane image metrics such as sharpness, Strehl ratio, or contrast[116].

Adaptive elements. Both phase and amplitude distortions in the wavefront can degrade image quality, but in adaptive optics and optical alignment applications, which are the primary focus of this thesis, phase aberrations are typically the dominant contributor. Phase fluctuations are commonly corrected using adaptive elements such as deformable mirrors, tip-tilt mirrors, and motor-

ized stages, which introduce phase shifts by producing geometric optical path differences. Deformable mirrors are preferred in astronomical adaptive optics over birefringent electro-optical phase correctors which produce refractive index differences, because deformable mirrors offer fast response, large strokes, polarization insensitivity, wavelength-independent phase shifts, and high, uniform reflectivity[105]. In contrast, spatial light modulators or other liquid crystal devices can modulate not only phase but also amplitude and polarization, enabling dynamic wavefront shaping across a broad range of applications including microscopy[76], optical trapping[79], beam shaping[117], holography and laser processing[41].

Control algorithms are usually implemented on a real time computer and translate wavefront sensing data into adaptive elements' adjustments with high speed and accuracy, with the goal of minimizing residual wavefront aberration. Closed-loop feedback control is often deployed, with the input and output being respectively the incoming wavefront phase perturbations and the residual phase after correction. Typical control strategies include optimized modal control[27] and Kalman filtering[54, 91, 33] coupled with linear-quadratic-Gaussian control for better rejection of mechanical vibrations mainly caused by cryo-coolers and electronics[51, 87]. In chapter 3, we will present how we extend the optimized modal gain integrator proposed by Gendron[27] to our applications on GPI.

This dissertation explores the principles, methodologies, and applications of wavefront sensing and control, in two key areas:
(1) Automated optical alignment

(2) High-contrast imaging - the Gemini Planet Imager

This work integrates concepts from optics, control theory, computational modeling and experimental validation and places particular emphasis on focal plane wavefront control. The following sections introduce these application domains, while subsequent chapters delve into the theoretical foundations and practical implementations of wavefront control in both contexts. Through this study, we aim to demonstrate the versatility and critical importance of focal plane wavefront control in advancing optical system performance.

1.2 Automated alignment

Precise alignment of optical systems is critical to achieve optimal optical throughput, minimize aberrations, and ensure system performance and functionality. However, traditional alignment methods are often labor-intensive, time-consuming, and reliant on manual intervention, which can introduce errors and limit repeatability. Automated alignment leverages advanced sensing, control algorithms, and actuation technologies to meet the need of increasingly complex optical systems and tighter tolerances. In Chapter 2, we present our research on automated alignment using focal plane sensing and Kalman filtering.

1.3 High-Contrast Imaging: Gemini Planet Imager

1.3.1 Exoplanet Detection and Direct Imaging

The discovery and study of exoplanets has revolutionized our understanding of planetary systems and the potential for life beyond Earth. Since the first confirmed detection of an exoplanet in 1995[78], advancements in techniques and

instrumentation have dramatically expanded the field, leading to the identification of thousands of exoplanets with diverse characteristics[81]. These discoveries have fueled a deeper investigation into the formation, evolution, and habitability of planetary systems, posing fundamental questions about our place in the universe.

A key challenge in exoplanet detection is the inherent difficulty of observing these distant objects, which are extremely faint and often obscured by the overwhelming brightness of their host stars. To overcome these obstacles, researchers have developed a variety of detection methods, each tailored to exploit specific physical effects or phenomena. Radial velocity measurements detect exoplanets by observing the periodic Doppler shift in a star's spectral lines caused by the gravitational force of an orbiting planet. This technique has been instrumental in detecting planets around nearby stars, particularly those with large masses or close-in orbits.[13, 24, 3] Transit photometry, on the other hand, identifies exoplanets by measuring the slight dimming of a star's light as a planet passes in front of it. This method has enabled the discovery of numerous Earth-sized exoplanets and has facilitated atmospheric characterization through transit spectroscopy.[16, 11, 94] Microlensing, based on the gravitational lensing effect predicted by Einstein's general theory of relativity, has proven effective in detecting exoplanets at greater distances from their stars and in diverse environments, including low-mass stars and free-floating planets.[8, 5]

Unlike the previous indirect methods which infer the presence of exoplanets through their effects on their host stars, direct imaging aims to capture the light emitted or reflected by the planet itself. This approach provides unparalleled

opportunities for characterizing planetary atmospheres, surface properties, and orbital dynamics, and offers key insights into planetary formation, evolution, and potential habitability. However, direct imaging represents an extraordinary technological challenge because stars are typically millions to billions of times brighter than their planets. The stark contrast and the small angular separation between a star and its planet demand exceptional spatial resolution and contrast performance from imaging instruments. [75, 52, 65]

To overcome these challenges, astronomers have developed a suite of advanced techniques and technologies. Coronagraphy[64, 35, 113, 77] and starshade[115], which are internal and external occulters respectively to block out the starlight to reveal nearby objects, are key tools for enhancing contrast. High-contrast imaging systems are further augmented by adaptive optics, which correct for atmospheric turbulence in real time, restoring the sharpness and clarity of ground-based observations. Post-processing algorithms, such as angular differential imaging[72] and principal component analysis[108], play a crucial role in extracting faint planetary signals from noisy datasets.

Ground-based observatories equipped with state-of-the-art instruments, such as the Very Large Telescope's SPHERE[7], the Gemini Planet Imager[66, 65], and Subaru's SCExAO[36], have made significant contributions to direct imaging of exoplanets. Complementing these efforts, space-based missions like the Hubble Space Telescope[42, 47] and the James Webb Space Telescope[85, 14, 80] offer unparalleled sensitivity and stability, free from the distortions of Earth's atmosphere. These facilities have enabled the detection and characterization of several directly imaged exoplanets, providing valuable data on their masses,

compositions, and thermal emissions. These facilities' instrument design and operation have informed the development of next-generation facilities, such as the Extremely Large Telescope's METIS instrument[12] and the Nancy Grace Roman Space Telescope's coronagraph[4], to achieve even higher contrast and sensitivity.

1.3.2 Gemini Planet Imager and its Upgrade

The Gemini Planet Imager (GPI) has revolutionized the field of exoplanet detection and characterization by enabling high-contrast imaging and spectroscopy of exoplanets. One of GPI's primary objectives is to directly image young, Jupiter-sized exoplanets that emit detectable infrared radiation as a result of their formation processes. The instrument's high sensitivity and resolution enables GPI to detect planets with contrasts as high as 10^{-6} relative to their host stars and characterize their atmospheric compositions, temperatures, and potential weather patterns. Since its first light in 2013 on the Gemini South Telescope in Chile, GPI has achieved the direct imaging of several exoplanets and brown dwarfs, as well as the discovery of circumstellar disks that provide insights into the environments where planets form. The GPI Exoplanet Survey, a large-scale observational program, has conducted systematic studies of nearby stars, significantly expanding our knowledge of giant planet demographics and their host systems. [18]

GPI's state-of-the-art design consists of adaptive optics, an Apodized Pupil Lyot Coronagraph (APLC) coronagraph, an infrared calibration system (CAL) with a wavefront sensor, an integral field spectrograph (IFS), and advanced data processing algorithms. Figure 1.1 shows GPI's schematic with light paths. GPI's

AO system exists to control both dynamic wavefront errors (WFEs) such as atmospheric turbulence and slowly-varying static WFEs monitored by CAL. CAL is a high-order wavefront sensing system downstream of the AO system and the coronagraph fore-optics, and plays a critical role in compensating for residual and non-common path aberrations (NCPAs). The APLC coronagraph effectively suppresses coherent light from the central star to reveal faint planetary companions. IFS captures spatially resolved spectral data, providing detailed information about the chemical and physical properties of detected exoplanets. [66]

After years of continuous operations, GPI is being relocated to Gemini North Telescope and undergoing a significant upgrade to its second-generation iteration, GPI 2.0, to expand its scientific capabilities in surveying cold-start planets, planets of low-mass stars and very young stars, asteroids and solar system objects, probing planet formation in disks, and studying planet variability and abundance. The instrument upgrade falls under the following categories and is illustrated in Figure 1.2. [18, 19, 20]

(1) AO: Replacing the original Shack-Hartmann wavefront sensor with a pyramid wavefront sensor, replacing the conventional CCD with a near-zero-noise EMCCD, and transitioning the Real-Time Computer to use Herzberg Extensible Adaptive Real-time Toolkit architecture. The upgrades will push the limiting magnitude to I=14, increase Strehl ratio, reduce latency and significantly enhance contrast.

(2) Coronagraph: Replacing the original APLC masks with a new, numerically-optimized design to improve diffraction suppression, inner working angles, throughput and contrast for observing fainter stars.

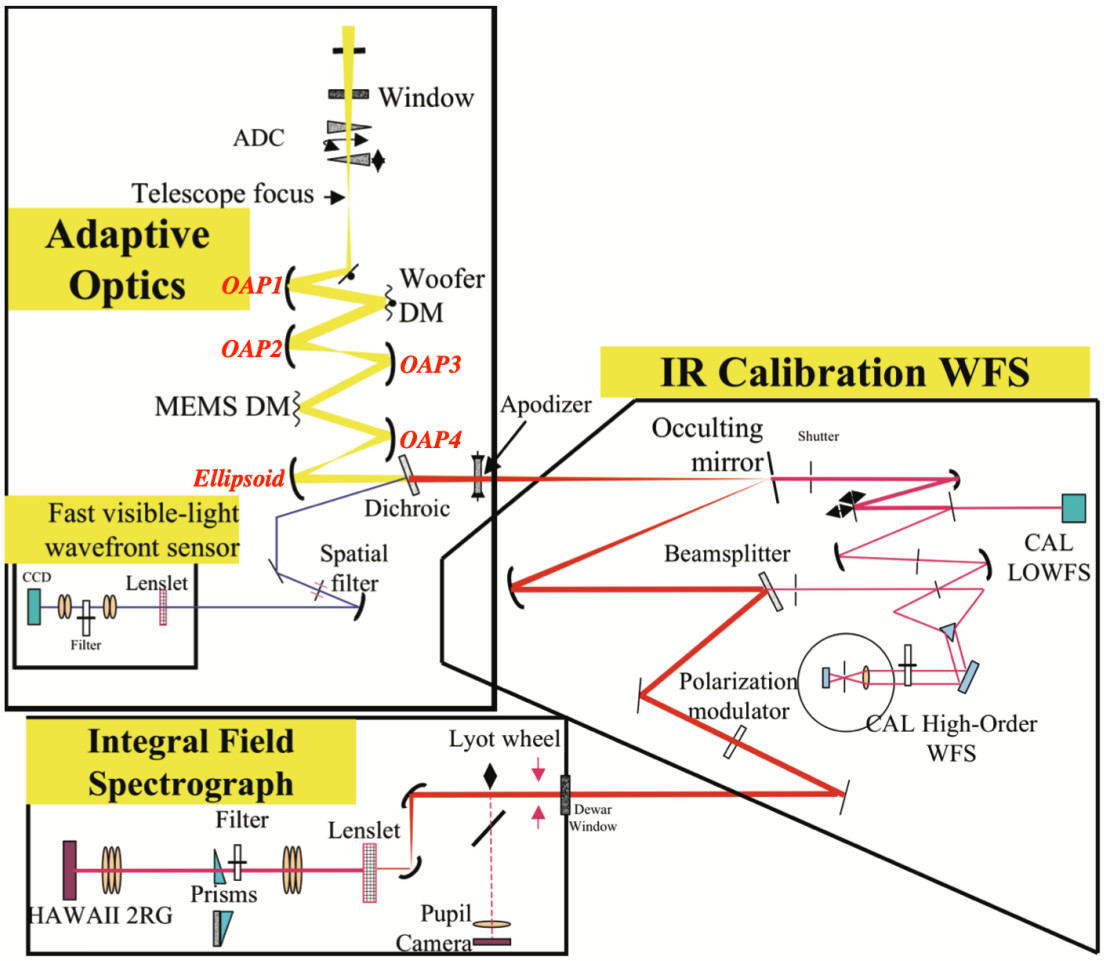

Figure 1.1: GPI schematic showing the light path and the three principal subsystems. [67]

(3) CAL 2.0: Replacing the high-order WFS with a Self-Coherent Camera, replacing the dual-arm interferometer design with a common-path design, and enabling improved aberration correction and stellar noise removal through focal plane wavefront sensing and control.

(4) IFS: Introducing new prisms for additional low-spectral-resolution mode to increase sensitivity for faint companions, broadening spectral range for improved calibration, and reducing telescope time cost for K-band observations.

(5) Software operation and usability: Enhancing software's robustness and user-

friendliness by automating user actions, reducing human intervention, improving alignment and calibration processes, transitioning the high-level control library from IDL to Python, introducing new data processing algorithms to improve precision and reliability, making the instrument more efficient for large-scale surveys.

In the following chapters, we present our research in AO control optimization for the AO upgrade in Chapter 3, followed by our work on focal plane wavefront sensing and control for the CAL 2.0 upgrade in Chapter 4.

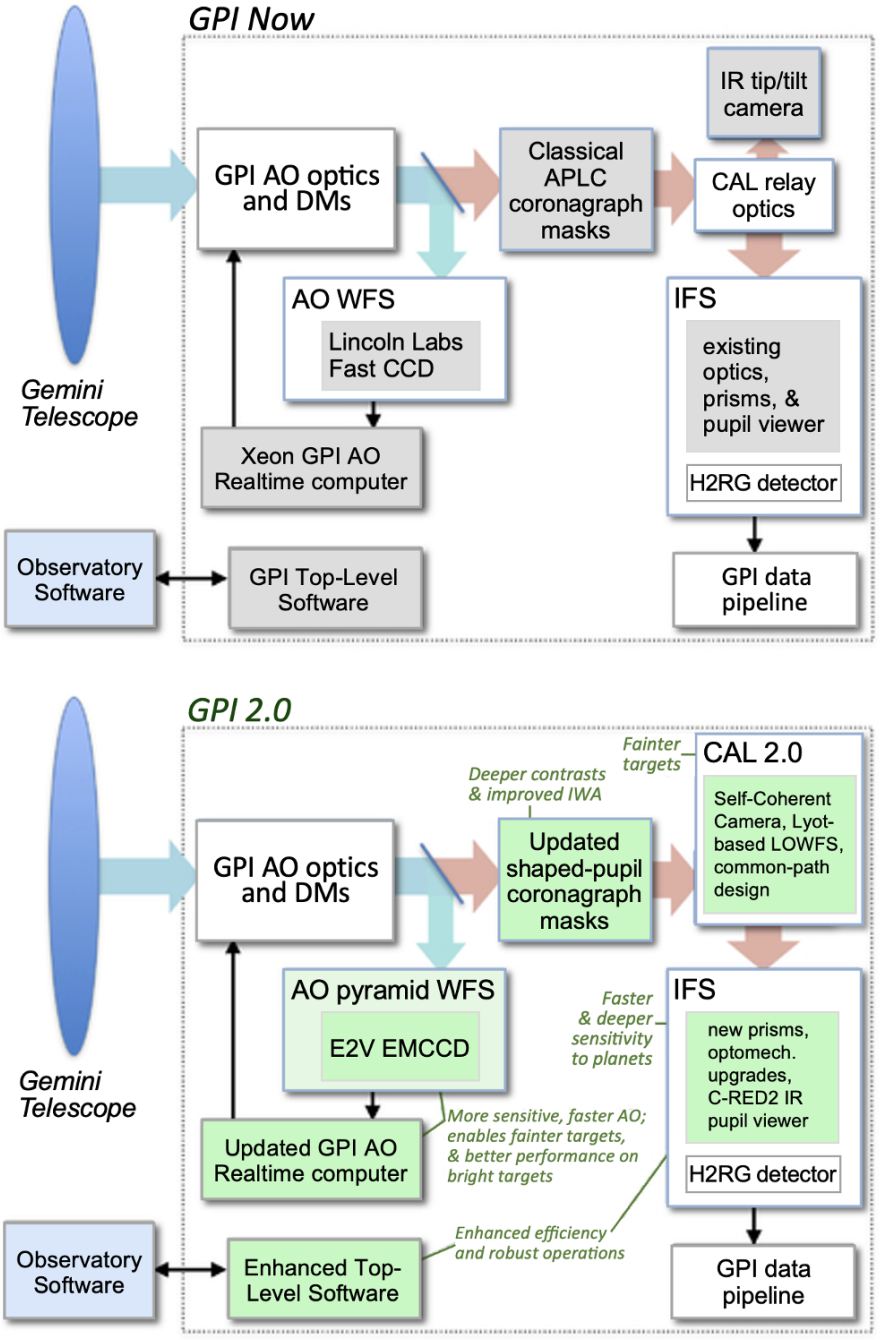

Figure 1.2: Diagram of GPI subsystems showing the original (left) and the upgraded (right) configurations. Adapted from [18].

CHAPTER 2

AUTOMATED ALIGNMENT OF OFF-AXIS ASPHERICS USING FOCAL PLANE WAVEFRONT SENSING AND KALMAN FILTERING

2.1 Introduction

As scientific and engineering optical systems become more sophisticated, there has been an increasing demand for automation in their mechanical operation. Various automated optical systems, which have some motorized degrees of freedom (DOFs), are under development or currently in use: adaptive optics systems employing mechanical DOFs to ensure coregistration between wavefront sensors and deformable mirrors in astronomy and ophthalmology [97], active optics that shape mirror geometry to prevent deformation due to environmental factors [39], corrective AR/VR systems that employ tunable-focus lenses for dynamic prescription vision correction [15] and industrial alignment of optical systems using robotic manipulators and wavefront filtering [104]. Of particular interest is the ability for optical systems with many DOFs to automatically align, so as to minimize manual effort in rapid assembly, iterative alignment, adjustment and maintenance. High performance optical systems often have strict demands on alignment precision, which may be violated due to thermal variation, structural deformation, air turbulence, vibration, and other disturbances. Automated alignment allows more efficient and flexible handling of optical systems, especially those that are not easily accessible, such as space instruments and autonomous, remote, ground-based systems.

Most of the current automated alignment methods map wavefront errors to

misalignments. Early computer-aided alignment methods used interferometers to measure the wavefront at multiple field points and adjust the mirrors based on Zernike sensitivity on misalignment to achieve the alignment goals [23, 122]. This method only works, however, in the linear range of the sensitivity matrix and will deteriorate in accuracy with large misalignments. To solve this problem, a reverse optimization method using merit function regression (MFR) was developed for faster and better convergence in iterative alignment processes. This approach used the embedded nonlinear optimization process in the optical design software ZEMAX to minimize the merit function representing the wavefront error measured by inteferometers [121, 50]. One can also use a Shack–Hartmann wavefront sensor for closed-loop alignment and control of an optical system in much the same way as an interferometer, with the advantage of vibration insensitivity, fast measurement, and ease of use [82], but with a trade-off in sampling size. These methods all use multi-field wavefront sampling and find reliability issues in multi-component alignment problems due to the intrinsic degeneracy and neglected non-linear coupling relationships between alignment state parameters. Differential wavefront sampling (DWS) was proposed to express the derivative of the wavefront with respect to the alignment states as a series expansion, where higher order terms contain nonlinear couplings between alignment parameters. By deliberately perturbing the system and solving a set of differential equations, a more accurate estimate of multi-component alignment states was achieved [55, 56]. This method was later integrated with multi-configuration MFR optimization to further improve its alignment estimation accuracy [84]. Hinrichs and Piotrowski [43] trained neural networks on the relationship between misalignment and wavefront errors. They added as-fabricated element wavefront errors into the optical model to help the

neural network distinguish between coupled DOFs such as tilt-induced decenter. Instead of taking a numerical approach, Gu et al. [34] developed an alignment model based on nodal aberration theory and derived analytical expressions for aberration field decenter vectors and boresight errors. Besides wavefront sensing, other automated alignment techniques exist for various optical systems. Computer vision has been used for repeated detection of semiregular features and continual realigning of coronagraph components [100]. Precision alignment of optical fibers utilizes optical power sensing or telecentric stereo microvision [17].

Although dedicated wavefront sensors or interferometric devices provide accurate multi-field wavefront measurements, they increase the cost and complexity of the optical system, decrease optical throughput, and most importantly, introduce non-common path error. Alternatively, wavefront sensing schemes employing internal sensors eliminate such concerns by trading off wavefront estimation accuracy. Common focal plane wavefront sensing methods involve pupil plane masking, multiple detectors, moving cameras or tunable lenses to introduce phase diversity and sense full-field wavefront aberrations [48]. Our technique is based on utilizing the existing focal plane sensor in an optical system for wavefront sensing so as to save resources and space. This also makes our alignment approach more easily integrated into other optical systems.

The aforementioned self-aligning methods, such as wavefront sensitivity on misalignment, DWS, and neural networks, all depend on some predetermined relationship between misalignment and wavefront errors and do not account well for model uncertainty, structural instability, real time wavefront variations,

positioning error and sensor noise. Therefore, an iterative misalignment identification and rectification process based on these approaches may become less stable when encountering such uncertainties and would be limited by additional alignment residuals. Riggs et al. [96] used an iterated extended Kalman filter to increase the system's robustness to large uncertainties and enable fast recursive wavefront correction in high contrast imaging. Lou et al. [62] used a Kalman filter to compensate for the system wavefront error and to improve the James Webb Space Telescope's (JWST) fine phasing process under large optical misalignments. This work likewise will use several variants of Kalman filtering to handle the nonlinearity in the misalignment to wavefront mapping and to achieve optimal estimation of misalignment states under various noise conditions and uncertainties.

Our lab has previously demonstrated an automated alignment technique on multi-element refractive systems using pure focal plane sensing and optimal state estimation [22]. However, because lenses and mirrors are very different in terms of geometry, actuation mechanisms, and sensitivity to aberrations, the previous technique does not directly apply to reflective systems. Off-axis and aspheric elements, in particular, impose extra difficulty in alignment, due to the lack of rotational symmetry, non-orthogonal adjustments, the cross-coupling between focus change and lateral image motion, and the optical axis being non-parallel to the gut ray. In this chapter, we adapt and extend these earlier refractive system techniques to align multi-element reflective systems with off-axis parabolic mirrors (OAPs).

2.2 Model and Methodology

Our automated alignment scheme is shown in Figure 2.1, expressed in the form of a control diagram. The plant is the optical system modeled in Figure 2.2, in which the OAPs are misaligned and need to be corrected. The sensor is the CCD camera in Figure 2.2 and captures focal plane images during the alignment process. Inputs $\boldsymbol{q}(t)$ and $\boldsymbol{v}(t)$ are the process noise and the sensor noise respectively. Images are processed to extract key features as the system performance measurements $\boldsymbol{y}$ (Equation 2.3). The Kalman filter takes in the measurements and provides optimal estimation of the misalignment states $\boldsymbol{x}$ (Equation 2.1). The optimal control inputs $\boldsymbol{u}$ (Equation 2.2) are then calculated by a linear quadratic regulator to correct for the misalignment. Overall, the system goes through a closed-loop, iterative misalignment identification and rectification process. In this section, we present the details of the optical model and alignment strategy based on Karhunen-Loève (KL) transform and optimal state estimation.

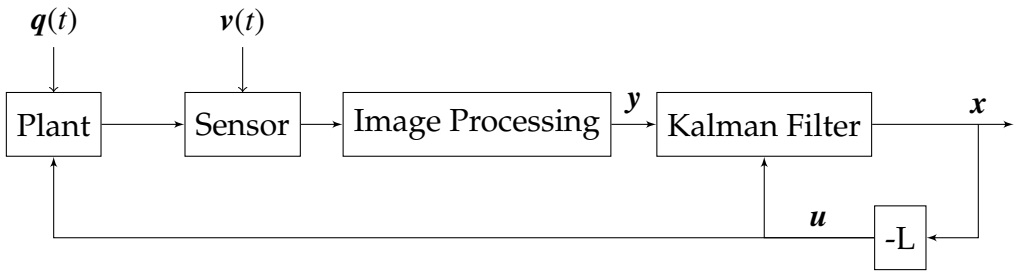

Figure 2.1: Schematic of our automated alignment technique using closed-loop control

2.2.1 Optical Model

Our optical system features two relaying OAPs and 10 DOFs. We modeled and optimized the virtual system in Zemax OpticStudio 2018 (Figure 2.2) and built the actual system in lab (Figure 2.13). The overall design represents a typical

reflective system and makes it applicable to generalize our alignment method to many other finite conjugate imaging applications.

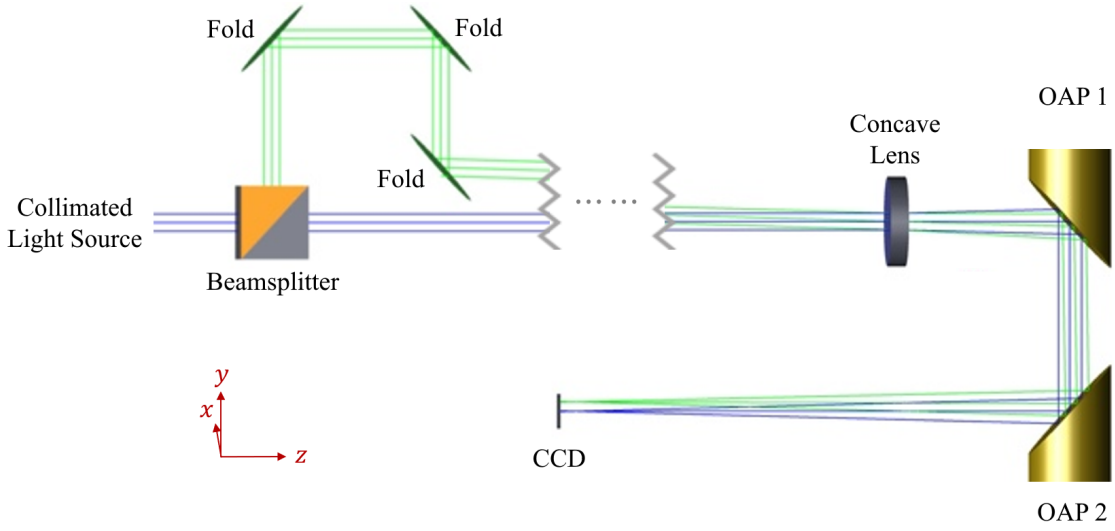

Figure 2.2: The optical system modeled in Zemax OpticStudio.

We mount the two OAPs on motorized translation and tip-tilt stages to control their position and orientation. Each OAP can move along all 3 axes and rotate about the x and z axes. We define the vector $\boldsymbol{x}$ to include all 10 misalignment states:

$$\boldsymbol{x} = [D_{x1}, D_{y1}, D_{z1}, T_{x1}, T_{z1}, D_{x2}, D_{y2}, D_{z2}, T_{x2}, T_{z2}]^T \tag{2.1}$$

where the decenter D_{ij} is OAP j's translation along the i-axis and the tip-tilt T_{ij} is OAP j's rotation abount the i-axis. Rolls T_{y1} and T_{y2}, which are the rotations about the y axis, are neglected because the OAPs are fixed in this direction. We assume the rest of the optical system is perfectly aligned and fixed while introducing an initial misalignment to $\boldsymbol{x}$. Our goal is to control the motorized stages so that $\boldsymbol{x}$ converges to $\mathbf{0}$ and the system realigns. We define the control command as

$$\boldsymbol{u} = [\Delta D_{x1}, \Delta D_{y1}, \Delta D_{z1}, \Delta T_{x1}, \Delta T_{z1}, \Delta D_{x2}, \Delta D_{y2}, \Delta D_{z2}, \Delta T_{x2}, \Delta T_{z2}]^T \tag{2.2}$$

The optical system is designed as follows: The collimated light source consists of a monochromatic laser source, a collimator and a neutral density filter. The beamsplitter splits the Gaussian beam into an on-axis beam (blue) and an off-axis beam (green). The off-axis beam forms an angle of 1.5108° with the on-axis beam and the center of the two beams rejoin at the center of the concave lens. The concave lens is positioned such that the virtual images of the two beams are in the focal plane of OAP1. The configuration ensures that both beams form diffraction-limited spots in OAP2's focal plane. The CCD is offset along the optical axis by 5 mm from OAP2's focal plane to introduce defocus for better imaging and feature detection. Currently, only the on-axis beam is used for image sensing and closed-loop alignment. The off-axis beam is not actively involved in the control loop, but rather is used for algorithm performance verification and repeatability testing. The optical components are modeled after off-the-shelf products summarized in Table 2.3. We choose two OAPs with different focal lengths to induce more distinctive behaviors from each and thus reduce the multi-state coupling effect (Section 2.5.1).

2.2.2 Image Processing

Misalignment of the OAPs introduces wavefront aberrations, which further cause intensity aberrations in images. To understand the image aberrations under stochastic optical misalignment, we simulate 10,000 focal plane images using Zemax Physical Optics Propagation (POP) by setting the misalignment vector $\boldsymbol{x}$ to random values using the uniform distribution $\{D_{ij}\} \sim U(-1, 1)$ mm for translations and $\{T_{ij}\} \sim U(-0.2, 0.2)^\circ$ for rotations. These ranges were chosen to allow maximum travel while confining the beam within the sensor field of view.

For each configuration $\boldsymbol{x}$, Zemax POP outputs the position (C_x, C_y) of the chief ray in the sensor frame to capture the lateral image motion, and a high resolution local frame image to capture the focal distortion. To quantify the beam variation, we interpolate the local frame images into 250×250 pixel images, max-min normalize them and apply Karhunen-Loève (KL) transform [93, 107] for dimension reduction (see Section 2B in Fang and Savransky 2016 [22] for mathematical details). We decompose each image into a weighted sum of KL modes and we choose the first 50 weights $\{w_i, i = 1, 2, ..., 50\}$ as the low dimensional representation of the local frame images to better match the dimension of $\boldsymbol{x}$.

The KL modes are a set of eigenimages of the optical system. Unlike Zernike or Fourier decomposition, the KL transform, being a variant of principal component analysis, enables image compression in the most energy compact way by minimizing the L2 loss function of the reconstructed image given the same number of modes used [93]. Due to the radial symmetry of the system, the major KL modes in this case closely resemble Zernike modes. Figure 2.3 shows the first 12 KL modes. We can see that modes 1 and 2 resemble defocus; modes 3 and 4 look like secondary astigmatism; modes 8 and 9 resemble higher order Zernike polynomials; modes 5 and 7 have oval envelopes that are a common feature of parabolic mirrors.

We define our system measurement vector $\boldsymbol{y}$ to include the spot center position and the KL weights:

$$\boldsymbol{y} = [C_x, C_y, w_1, w_2, ..., w_{50}]^T \tag{2.3}$$

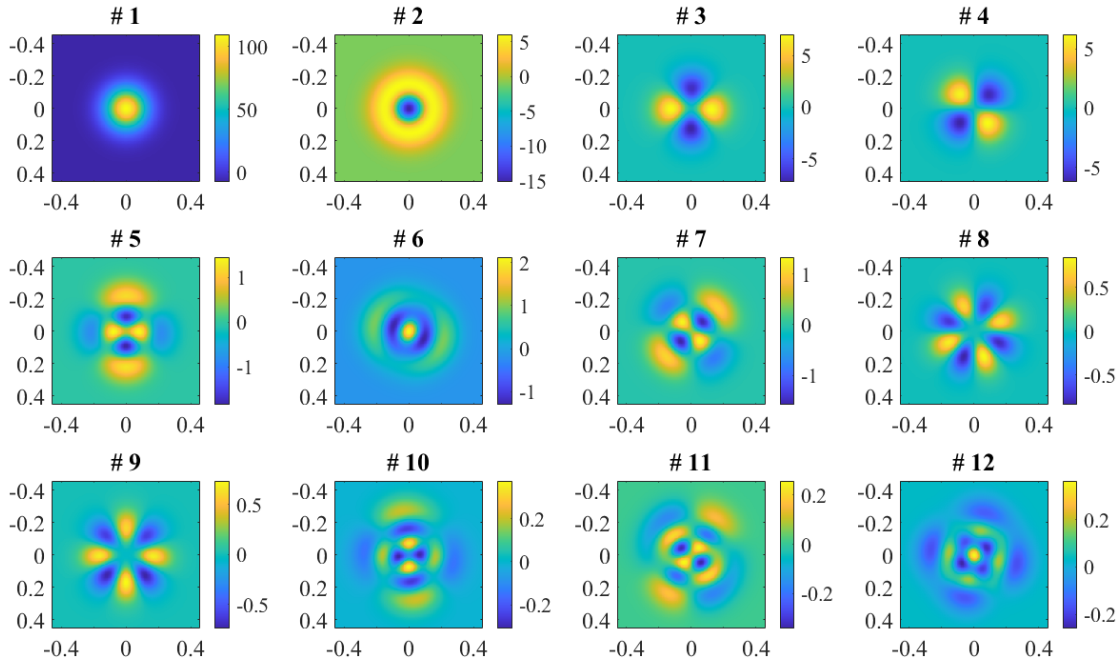

Figure 2.3: First 12 KL modes obtained from KL transform. The modes are in local frames with dimensions in mm. The colorbars indicate the irradiance in W/mm^2.

2.2.3 Measurement Function

Reverse optimization [63, 49] is the process of optimizing the system variables, in our case $\boldsymbol{x}$, by matching the system performance measurement, in our case $\boldsymbol{y}$, to a model. As part of our reverse optimization process using Kalman filtering (Section 2.3), we model the measurement function $\hat{\boldsymbol{y}} = \boldsymbol{h}(\boldsymbol{x})$, where the predicted value $\hat{\boldsymbol{y}}$ differs from the simulated value $\boldsymbol{y}$ due to a fitting error $\boldsymbol{e} = \hat{\boldsymbol{y}} - \boldsymbol{y}$.

To model the relationship between optical misalignment and focal plane images, we simulate another 20,000 images using the same method as in Section 2.2.2 and partition them into a 14,000-sample training set and a 6,000-sample testing set. Figure 2.4 shows a scatter plot between $\boldsymbol{x}$ and the first 8 elements in $\boldsymbol{y}$. To capture the nonlinearity in the scatter, we choose to use a second-order

polynomial

$$\hat{y}_j = h_j(\boldsymbol{x}) = \sum_{i=1}^{N}\sum_{l=1}^{N} a_{jil}x_i x_l + \sum_{i=1}^{N} b_{ji}x_i + c_j \tag{2.4}$$

where each nonlinear function h_j maps the misalignment states $\boldsymbol{x}$ to the j-th element in $\hat{\boldsymbol{y}}$, and $N = 10$ is the number of misalignment states. We perform a nonlinear least-squares fit on the training set using the Levenberg–Marquardt algorithm [92] to find $\boldsymbol{h}(\cdot)$.

To model the fitting error covariance, we assume $\boldsymbol{e}$ as unbiased and define an error matrix containing the error for each testing sample

$$\boldsymbol{E} = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \dots & \boldsymbol{e}_{n_{test}} \end{pmatrix} \tag{2.5}$$

where $n_{test} = 6000$ is the testing set size. The error covariance of the measurement function, $\boldsymbol{R}_h$, which will be part of the observation error covariance $\boldsymbol{R}$ (Equation 2.9) in the Kalman filter, is therefore

$$\boldsymbol{R}_h = \boldsymbol{E}\boldsymbol{E}^T/(n_{test} - 1) \tag{2.6}$$

2.3 Optimal Estimation and Control

We use Kalman filtering to achieve optimal estimation of the misalignment states and a linear quadratic regulator to achieve optimal control of the actuators.

Kalman filtering is widely used in optical state estimation and wavefront con-

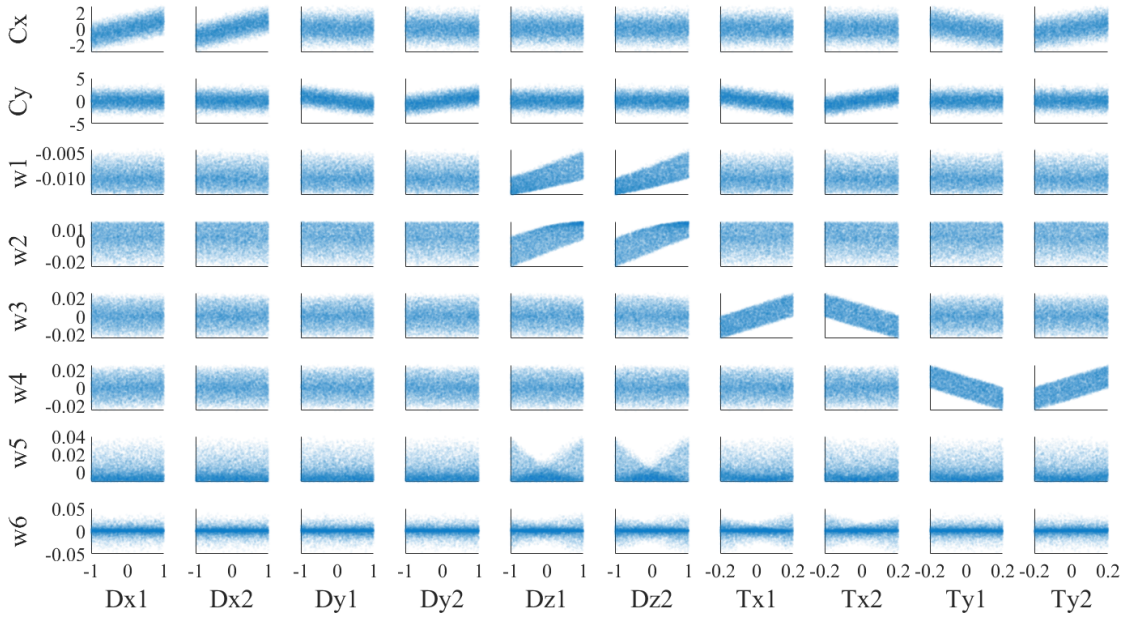

Figure 2.4: Scatter plot matrix between misalignment states $\boldsymbol{x}$ and measurements $\boldsymbol{y}$ in the training set. The units are mm for D_{ij} and degrees for T_{ij}.

trol to increase the system robustness to uncertainties, such as modeling error and structural instability. It uses a series of observed measurements, combined with statistical models of noise, to produce optimal estimates of unknown variables. The algorithm works recursively in a predicting and updating process. The filter first predicts the current state variables, along with their uncertainties, using the previous state estimate and known system dynamics. Once the measurement is observed, the filter updates its prediction using an uncertainty weighted average between dynamics and observations.

We assume the following discrete, time-invariant state-space representation of the system [111]

$$\boldsymbol{x}_k = \boldsymbol{F}\boldsymbol{x}_{k-1} + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{G}\boldsymbol{q}_k \tag{2.7}$$

$$\boldsymbol{y}_k = \boldsymbol{h}(\boldsymbol{x}_k) + \boldsymbol{D}\boldsymbol{r}_k \tag{2.8}$$

where k is the iteration step. The state transition matrix $\boldsymbol{F}$, the control input matrix $\boldsymbol{B}$, the process noise transition matrix $\boldsymbol{G}$ and the measurement noise transition matrix $\boldsymbol{D}$ in our case are all taken to be identity matrices. Kalman filtering assumes Gaussian noise, so we model the process noise as $\boldsymbol{q}_k \sim N(0, \boldsymbol{Q}_k)$ and the observation error as $\boldsymbol{r}_k \sim N(0, \boldsymbol{R})$. The process noise is associated with actuator precision and is modeled as a percentage error according to the stage specifications. The observation error consists of many error sources:

$$\boldsymbol{R} = \boldsymbol{R}_h + \boldsymbol{R}_{meas} + \boldsymbol{R}_{align} + \boldsymbol{R}_{optics} + \boldsymbol{R}_{sampling} + ... \tag{2.9}$$

where $\boldsymbol{R}_{meas}$ is caused by image noise, $\boldsymbol{R}_{align}$ is caused by the manual alignment error in the rest of the system, $\boldsymbol{R}_{optics}$ is caused by optics manufacturing error, and $\boldsymbol{R}_{sampling}$ is caused by finite image sampling in both simulation and experiments. $\boldsymbol{R}_{meas}$ is affected by the sensor noise $\boldsymbol{v}(t)$ (defined in Figure 2.1), laser fluctuation and vibration. We estimate $\boldsymbol{R}_{meas}$ by collecting images at different field points and analyzing the static variation. $\boldsymbol{R}_{align}$ is technically a bias but we model it as an error due to the lack of system identification tools. We estimate $\boldsymbol{R}_{align}$ by simulating the misalignment effect of the upstream optics. We neglect other less significant error sources.

Our dynamics (Equation 2.7) are linear but the measurement function (Equation 2.8) is not, so we use the nonlinear variants of the Kalman filter. The iterated extended Kalman filter [111] (IEKF) and the unscented Kalman filter [46, 45] (UKF) are extensions and generalizations of the original Kalman filter that work better with nonlinear state-space models. We apply and compare IEKF and UKF

on misalignment state estimation. We also apply the information filter [106] (IF) for observability analysis in Section 2.5.2. We introduce the three filters in the following sections.

2.3.1 Iterated Extended Kalman Filter

The IEKF uses local linearization of the state-space model (Equation 2.12) and follows the typical algorithm below. At each predicting step, we estimate the state $\boldsymbol{x}$ and its covariance $\boldsymbol{P}$ based on dynamics:

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{F}\hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{B}\boldsymbol{u}_k \tag{2.10}$$

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}^T + \boldsymbol{Q}_k \tag{2.11}$$

At each updating step, we linearize the measurement function $\boldsymbol{h}(\cdot)$ by taking its Jacobian $\boldsymbol{H}$, calculate the Kalman gain $\boldsymbol{K}$ and update the state and its covariance estimates based on observations:

$$\boldsymbol{H}_k = \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k-1}} \tag{2.12}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{R})^{-1} \tag{2.13}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\tilde{\boldsymbol{y}}_k \tag{2.14}$$

$$\boldsymbol{P}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{k|k-1} \tag{2.15}$$

where $\tilde{\boldsymbol{y}}_k$ is the actual measurement. The IEKF ensures convergence of $\hat{\boldsymbol{x}}_{k|k}$ during each update step by iterating through Equation 2.12 to Equation 2.15 until

the quadratic cost function, defined as

$$J = \frac{1}{2}(\hat{\boldsymbol{x}}_{k|k} - \hat{\boldsymbol{x}}_{k|k-1})^T \boldsymbol{P}_{k|k-1}^{-1}(\hat{\boldsymbol{x}}_{k|k} - \hat{\boldsymbol{x}}_{k|k-1}) + \frac{1}{2}(\tilde{\boldsymbol{y}}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_{k|k}))^T(\tilde{\boldsymbol{y}}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_{k|k})) \tag{2.16}$$

is minimized.

2.3.2 Square-Root Unscented Kalman Filter

The UKF uses the unscented transform to estimate the first two moments of the states without computing the Jacobians, so in theory the UKF has better performance than an IEKF when the system is highly nonlinear. We use the square-root UKF algorithm (SR-UKF) for state estimation [114] to enhance numerical stability and to ensure positive semi-definiteness of the covariance matrices. In Section 3, we present and compare filter performance on our system.

2.3.3 Information Filter

We choose to implement an information filter (IF), which is another variant of the KF and has more efficient observability computation. This is because the observability matrix for an IF can be computed offline and we have more measurements than states. The IF is carried out in the following recursive steps [106]:

$$\boldsymbol{x}_{k|k-1} = \boldsymbol{F}\boldsymbol{x}_{k-1|k-1} + \boldsymbol{B}\boldsymbol{u}_k \tag{2.17}$$

$$\boldsymbol{\mathcal{I}}_{k|k-1} = \boldsymbol{Q}_k^{-1} - \boldsymbol{Q}_k^{-1}\boldsymbol{F}(\boldsymbol{\mathcal{I}}_{k-1|k-1} + \boldsymbol{F}^T\boldsymbol{Q}_k^{-1}\boldsymbol{F})^{-1}\boldsymbol{F}^T\boldsymbol{Q}_k^{-1} \tag{2.18}$$

$$\boldsymbol{\mathcal{I}}_{k|k} = \boldsymbol{\mathcal{I}}_{k|k-1} + \boldsymbol{H}_k^T\boldsymbol{R}^{-1}\boldsymbol{H}_k \tag{2.19}$$

$$\boldsymbol{K}_k = \boldsymbol{\mathcal{I}}_{k|k}^{-1}\boldsymbol{H}_k^T\boldsymbol{R}^{-1} \tag{2.20}$$

$$\boldsymbol{x}_{k|k} = \boldsymbol{x}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{x}_{k|k-1}) \tag{2.21}$$

where k is the filter step. $\boldsymbol{u}$ is the control input. $\boldsymbol{y}$ are the measurements. $\boldsymbol{\mathcal{I}}$ is the information matrix and is defined to be the inverse of the covariance of $\boldsymbol{x}$. The state transition matrix $\boldsymbol{F}$ and the control input matrix $\boldsymbol{B}$ are taken to be identity matrices. $\boldsymbol{H}$ is Jacobian of the measurement function. $\boldsymbol{K}$ is the Kalman gain. $\boldsymbol{Q}$ is the process noise covariance and $\boldsymbol{R}$ is the measurement error covariance.

2.3.4 Linear Quadratic Regulator

Once we have the optimal state estimates at each iteration step, we apply the linear-quadratic regulator [111] as our optimal controller. We define the quadratic cost function as our control performance index

$$J = \sum_{k=1}^{\infty} \hat{\boldsymbol{x}}_{k|k}^T \boldsymbol{Q}' \hat{\boldsymbol{x}}_{k|k} + \boldsymbol{u}_k^T \boldsymbol{R}' \boldsymbol{u}_k \tag{2.22}$$

where $\boldsymbol{Q}'$ and $\boldsymbol{R}'$ are the weighting matrices on the costs of state deviations and control effort respectively, both defined to be identity matrices. The feedback control law that minimizes the cost is

$$\boldsymbol{u} = -\boldsymbol{L}\boldsymbol{x} \tag{2.23}$$

where $\boldsymbol{L}$ is the control gain matrix calculated by solving the algebraic Riccati equation related to Equation 2.22.

2.4 Simulation

2.4.1 Results

We simulate the alignment process in MATLAB R2021a and Zemax OpticStudio. The above mentioned image processing and optimal estimation and control algorithms are hard coded in MATLAB while the optical model is in Zemax. Matlab interacts with Zemax through Zemax API, including generating motions in the motorized degrees of freedom and retrieving focal plane images from POP.

Figure 2.5 shows the statistics of the simulated performance of our alignment algorithm using IEKF. SR-UKF and IF have similar performances. For each simulation, we introduce a random initial misalignment $\{D_{ij}\} \sim U(-1, 1)$ mm and $\{T_{ij}\} \sim U(-0.2, 0.2)^\circ$ to match the estimated manual alignment error. At early iteration steps ($k <= 25$), instead of feeding back the state estimate as the control input, we give random control commands to the motorized stages to increase phase diversity. During the random walk, the Kalman filter keeps predicting and updating and thus accumulates an internal knowledge of the phase. If we were to apply feedback control at the very beginning, the system would quickly converge to a local optimum and the magnitude of the state perturbations would be insufficient to produce enough phase diversity for correct wavefront reconstruction. We set the random walk steps to 25 and the magnitude to $\{D_{ij}\} \sim U(-1, 1)$ mm and $\{T_{ij}\} \sim U(-0.2, 0.2)^\circ$. For further discussion of this algorithm design see the observability analysis in Section 2.5.2.

Once the closed-loop linear-quadratic feedback control starts at step 26, the amount of misalignment decreases quickly by several orders of magnitude in

approximately 30 steps and then converges to a stable value in approximately 100 steps. The standard deviations follow the same trend. Table 2.1 shows the statistics of the final misalignment states. Both IEKF and SR-UKF are able to consistently regulate the linear misalignment down to < 5 μm and the angular misalignment down to < 6 arcsec in absolute values, and maintain the converged states in closed loop.

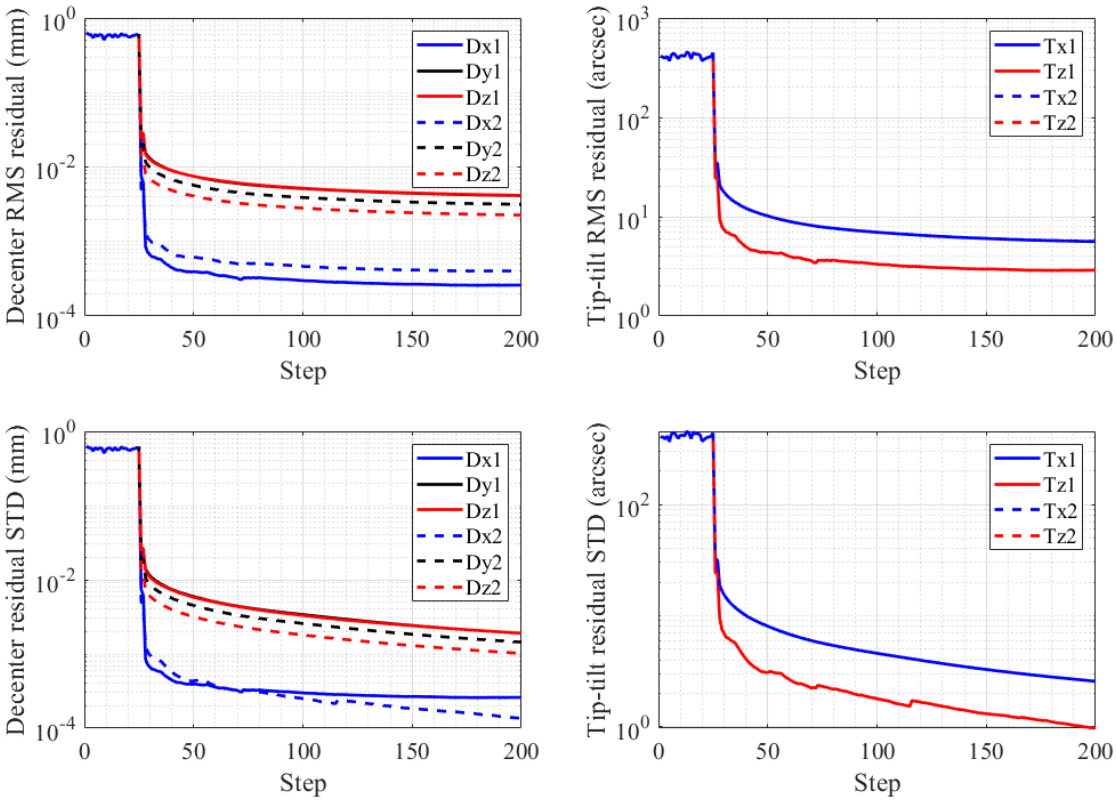

Figure 2.5: Simulated alignment performance, based on 67 tests using IEKF. (Top) The RMS values of true misalignment states ($\boldsymbol{x} = \mathbf{0}$ is the perfectly aligned case) at different iteration steps. (Bottom) The standard deviation of the residuals. Note that there is overlap between D_{y1} and D_{z1}, between T_{x2} and T_{x1} and between T_{z2} and T_{z1}. The random walk steps are only shown for D_{x1} and T_{x1} for neatness.

Table 2.1: Root-mean-square values and standard deviations of the misalignment states after 200 iteration steps for 67 IEKF tests and 5 UKF tests

State	IEKF RMS residual	SR-UKF RMS residual
D_{x1}	0.26 ± 0.09 μm	0.11 ± 0.075 μm
D_{y1}	4.1 ± 1.9 μm	4.2 ± 2.3 μm
D_{z1}	4.2 ± 1.9 μm	4.6 ± 2.8 μm
T_{x1}	5.6 ± 2.6 arcsec	5.7 ± 3.1 arcsec
T_{z1}	2.9 ± 1.0 arcsec	1.2 ± 0.80 arcsec
D_{x2}	0.40 ± 0.14 μm	0.17 ± 0.12 μm
D_{y2}	3.2 ± 1.4 μm	3.2 ± 1.7 μm
D_{z2}	2.3 ± 1.0 μm	2.4 ± 1.4 μm
T_{x2}	5.6 ± 2.6 arcsec	5.7 ± 3.1 arcsec
T_{z2}	2.9 ± 1.0 arcsec	1.2 ± 0.83 arcsec

2.4.2 Alignment Quality Assessment

We evaluate the simulated final alignment quality through analysis of image residuals and wavefront errors and summarize the statistics in Table 2.2, Figure 2.6 and Figure 2.7. Table 2.2 shows the lateral deviation of the spot in the global frame. Figure 2.6 shows the focal distortion of the on-axis and the off-axis spots in the local frame. The image residual is 4 orders of magnitude lower than the benchmark irradiance. Figure 2.7 shows the on-axis and the off-axis wavefront in the focal plane. The final wavefront aberration is dominated by piston, Y-tilt, defocus and vertical astigmatism, and is also 4 orders of magnitude lower than the benchmark wavefront.

Although we are only observing the on-axis spot during the closed-loop control,

Table 2.2: Statistics of the spot center deviations after 200 iteration steps for 67 IEKF tests

Measurement	RMS residual
$C_{x,\text{on-axis}}$	0.19 ± 0.19 nm
$C_{y,\text{on-axis}}$	20 ± 2.4 nm
$C_{x,\text{off-axis}}$	92 ± 31 nm
$C_{y,\text{off-axis}}$	66 ± 33 nm

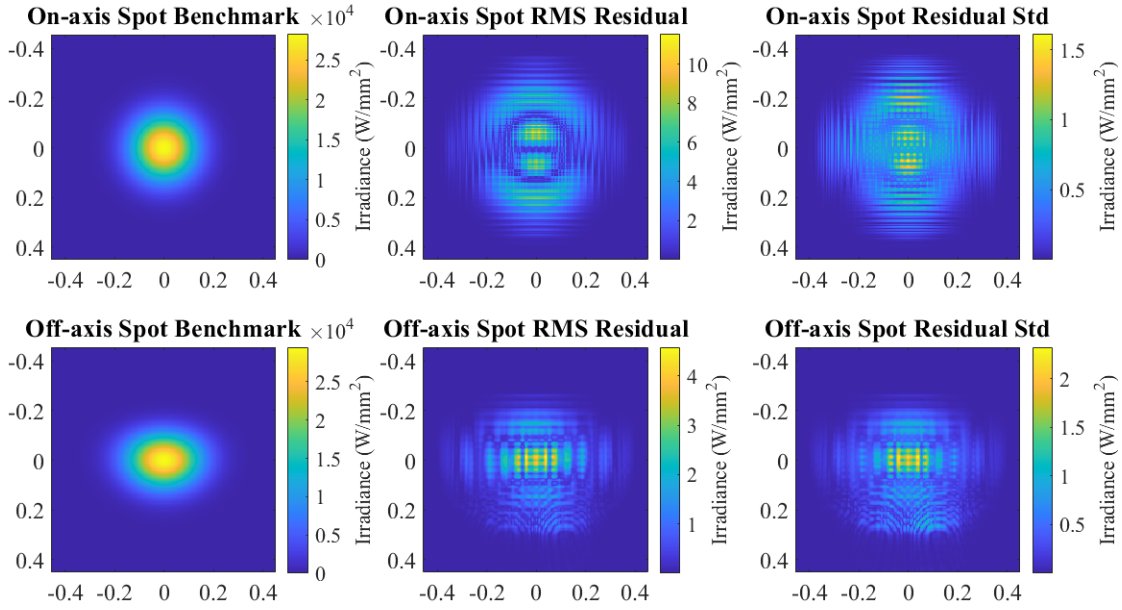

Figure 2.6: Final image quality. The benchmark (left column) is the perfectly aligned case. The RMS residual (middle column) and the standard deviation (right column) are evaluated after 200 iteration steps for 67 IEKF tests. The x and y axes are in units of mm.

the image sharpness of the off-axis spot indicates good image quality outside our narrow field of view. It also suggests that the on-axis spot alone is sufficient for focal plane sensing and alignment with linear accuracy of 5 μm and angular accuracy of 6 arcsec.

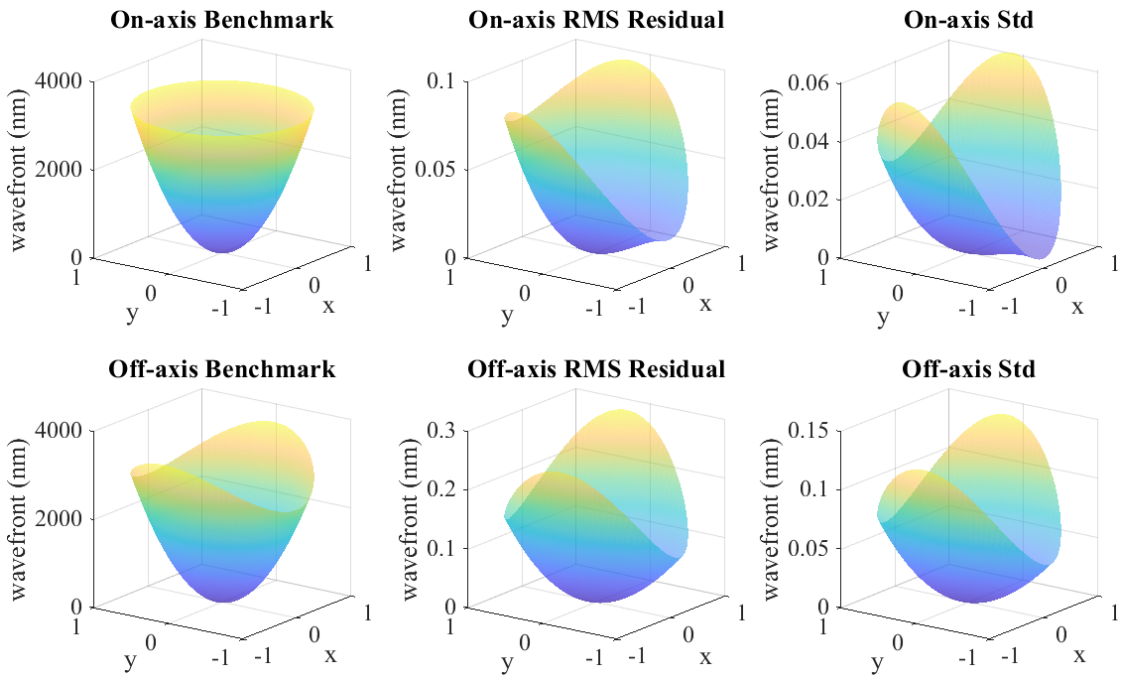

Figure 2.7: Final wavefront quality. The arrangement is the same as in Figure 2.6. x and y are in relative units.

2.5 Observability Analysis

2.5.1 Multi-State Coupling

Inspecting Table 2.1, Table 2.2 and Figure 2.7 brings up an important question: Why are the final image and wavefront accurate up to nanometers when the alignment is off by microns? The answer is a multi-state coupling effect [55] that causes different misalignment states to have compensating effects in measurements, thus making the system poorly observed.

If we look at a single test case and plot the estimated states and uncertainties compared to the true states, as shown in Figure 2.8, we can see the filter consistently underestimates the amount of misalignment and even the uncertainty

for T_{x1}. This means the filter has converged to a local optimum, which is nearly identical to the global optimum in the measurement space. Figure 2.9 shows the estimation error (the difference between the estimated state and the true state) of all misalignment states in the same test. We notice 2 groups of coupled states: D_{x1}, D_{x2}, T_{z1} and T_{z2} share a similar trend once the closed-loop feedback control starts. The same is true for the other 6 states: D_{y1}, D_{y2}, D_{z1}, D_{z2}, T_{x1} and T_{x2}. These couplings are also manifested in the overlapping between pairs $\{D_{y1}, D_{z1}\}$, $\{T_{x1}, T_{x2}\}$ and $\{T_{z1}, T_{z2}\}$ throughout the alignment iterations, as shown in Figure 2.5. Figure 2.4 shows the linear dependency between adjacent columns: pairs of $\{D_{x1}, D_{x2}\}$, $\{D_{y1}, D_{y2}\}$, $\{D_{z1}, D_{z2}\}$, $\{T_{x1}, T_{x2}\}$ and $\{T_{z1}, T_{z2}\}$ either have similar and opposite effects on $\boldsymbol{y}$. Though we deliberately designed the 2 OAPs to have different focal lengths, they still preserve a lot of symmetry in their DoFs along the same axis. The linear dependency implies that the measurement jacobian $\boldsymbol{H}$, which is derived from Figure 2.4 and used in the filtering algorithm, could be singular or close to singular for certain values of $\boldsymbol{x}$. As a result, the multi-state coupling effect introduces modeling error and is the dominating factor that prevents the Kalman filters from further convergence.

2.5.2 Observability

In order to further understand the multi-state coupling effect, we evaluate the observability [111] of our system. Observability is a measure of how well internal states of a system can be inferred from knowledge of its external outputs.

The observability matrix at time step k of an information filter for a nonlinear

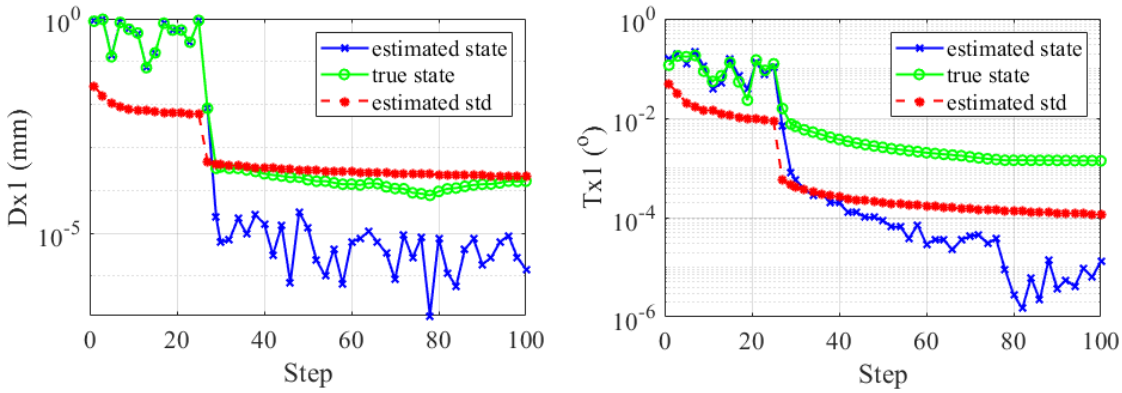

Figure 2.8: Filter performance in one typical IEKF test. The estimated misalignment values (blue), the estimated estimation error (red) and the true misalignment values (green) are shown for representative misalignment states. The other states have similar patterns. Absolute values are shown.

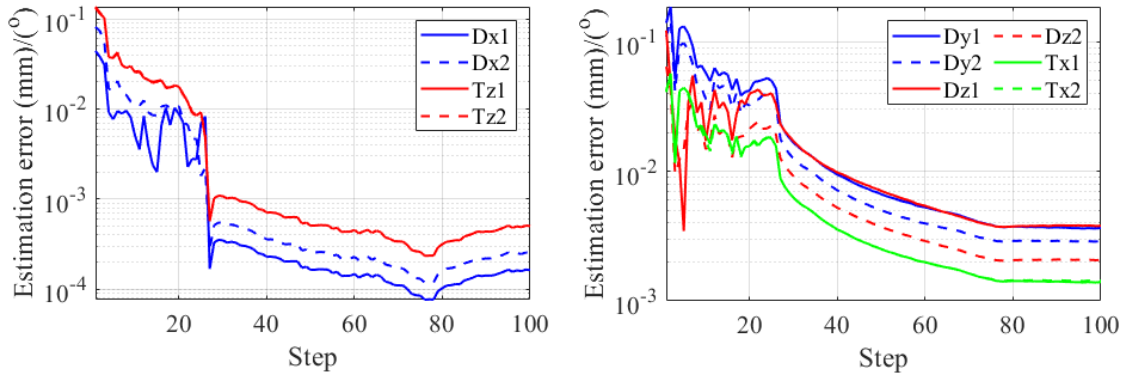

Figure 2.9: Absolute state estimation error in one typical IEKF test. There is overlap between T_{z1} and T_{z2} and between T_{x2} and T_{x1}.

system is given by [119]

$$\boldsymbol{O}_k = \begin{bmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2\boldsymbol{F}_1 \\ \dots \\ \boldsymbol{H}_k\boldsymbol{F}_{k-1}\dots\boldsymbol{F}_1 \end{bmatrix} \tag{2.24}$$

where $\boldsymbol{H}$ and $\boldsymbol{F}$ follow the definitions in Section 2.3. The scaled observability Gramian is defined as

$$\bar{\boldsymbol{O}}_k = \boldsymbol{O}_k^T \mathcal{R}_k^{-1} \boldsymbol{O}_k \tag{2.25}$$

where $\mathcal{R}_k$ is the augmented version of $\boldsymbol{R}_k$ to match the dimension of $\boldsymbol{O}_k$. [44] We take the eigenvalues and eigenvectors of $\bar{\boldsymbol{O}}^{-1}$ and rank the eigenvectors by their eigenvalues, among which large eigenvalues indicate poor observability for the corresponding eigenvector. [38] We plot the scatter of the eigenvectors with lowest observability in different IF tests (Figure 2.10) and notice some consistency regardless of the initial misalignment and the random walk. First, there are two groups of coupled states and their ratio follows the relationship:

$$\begin{aligned} &D_{x1} : D_{x2} : T_{z1} : T_{z2} = 1.0 : -1.6 : 3.0 : 3.0 \\ &D_{y1} : D_{y2} : D_{z1} : D_{z2} : T_{x1} : T_{x2} = 2.7 : 2.0 : 2.6 : 1.4 : -1.0 : -1.0 \end{aligned} \tag{2.26}$$

which matches the grouping and the relative magnitude of the estimation error in Figure 2.9. This means that this eigenvector indeed dominates the state estimation error. Second, the values of D_{y1}, D_{y2}, D_{z1}, D_{z2}, T_{x1} and T_{x2} are highly clustered, with dispersion 2 orders of magnitude lower than the stage movement. Thus we can express the eigenvector with lowest observability in the form of

$$\boldsymbol{\eta} \approx [\delta \text{ mm}, 0.57 \text{ mm}, 0.55 \text{ mm}, -0.21^\circ, 3\delta^\circ, -1.6\delta \text{ mm}, 0.43 \text{ mm}, -0.30 \text{ mm}, -0.21^\circ, 3\delta^\circ]^T \tag{2.27}$$

where $\delta < 0.012$ is a perturbation variable that matches the dispersion in D_{x1}, D_{x2}, T_{z1} and T_{z2}. When the stage movements are proportional to $\boldsymbol{\eta}$, the Kalman filter finds it hardest to capture the difference in measurements. This suggests

that to better observe the system, we can potentially break the coupling by maneuvering the states separately and by an amount disproportional to $\boldsymbol{\eta}$.

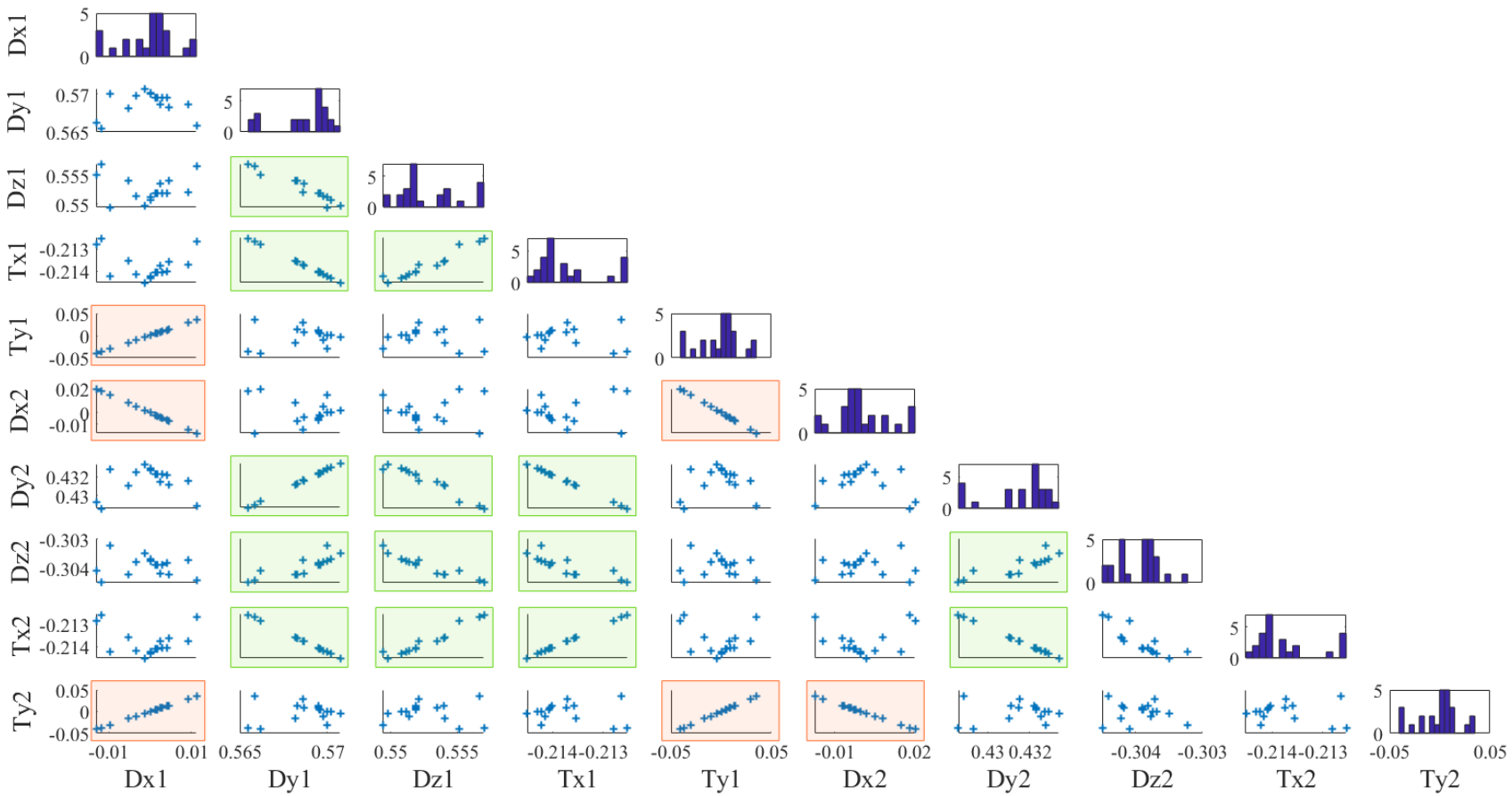

Figure 2.10: Scatter plot matrix of the state values in the eigenvectors with lowest observability within 50 iteration steps for 25 information filter tests. Linearity indicates coupling. The orange and the green boxes indicate 2 groups of states are coupled within each other. The units are mm for D_{ij} and degrees for T_{ij}.

We plot the histogram of the pairwise projections (dot products) between the worst observed eigenvectors in Figure 2.11. All the projection values are close to 1, meaning the most poorly observed eigenvector remains nearly constant regardless of the initial misalignment and the control history and is intrinsic to the system.

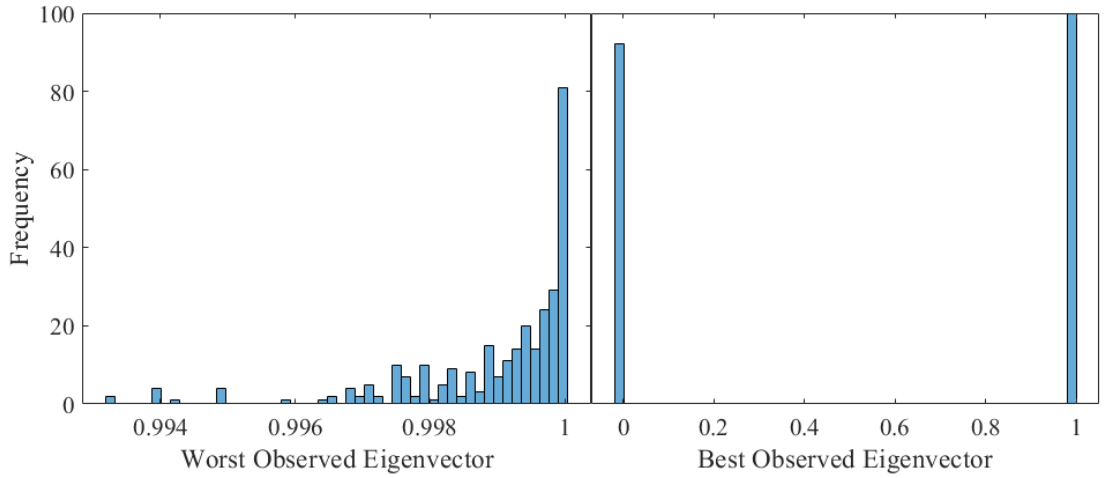

Figure 2.11: Histogram of pairwise dot products between the most poorly observed eigenvectors (left) and between the best observed eigenvectors (right) after 50 time steps for 25 filter test runs. A dot product of value 1 means the two eigenvectors are the same. A dot product of value 0 means the two eigenvectors are orthogonal.

We apply the same analysis to the eigenvectors with the highest degree of observability (Figure 2.11). Unlike the worst observed eigenvectors which are approximately equal to each other, the best observed eigenvectors have two clear clusterings because the projections have values of exactly 1 or 0. Moreover, the two eigenvectors have equal significance because of the similar frequencies at 1

and 0. The most commonly best observed eigenvectors are

$$\begin{aligned}\boldsymbol{\eta}_{maxOD,1} = [&-0.199 \text{ mm}, 0.016 \text{ mm}, -0.0012 \text{ mm}, 0.078^\circ, 0.593^\circ, \\ &- 0.332 \text{ mm}, -0.024 \text{ mm}, 0.0012 \text{ mm}, -0.088^\circ, -0.696^\circ]^T\end{aligned} \tag{2.28}$$

and

$$\begin{aligned}\boldsymbol{\eta}_{maxOD,2} = [&0.024 \text{ mm}, 0.131 \text{ mm}, -0.0095 \text{ mm}, 0.641^\circ, -0.073^\circ, \\ &0.040 \text{ mm}, -0.193 \text{ mm}, 0.0095 \text{ mm}, -0.721^\circ, 0.085^\circ]^T\end{aligned} \tag{2.29}$$

Stage movements proportional to $\boldsymbol{\eta}_{maxOD}$ would induce variation in measurements that are easiest to capture. To increase the overall observability of the system, we can potentially divide the states into different combinations without coupled pairs and observe the combinations separately. Alternatively, we can control the stage movements to be proportional to the well observed eigenvectors.

2.5.3 Algorithm Design

In our filter test runs, the observability matrix is full rank at every time step, which means all the states are observable throughout the filtering. However, considering the existence of the multi-state coupling effect, the rank does not tell how well the states are observed. Therefore, we assess the degree of observability OD using the scaled observability Gramian

$$OD_k = \frac{\sqrt{\lambda_{min}(\bar{\boldsymbol{O}}_k)}}{\sqrt{\lambda_{max}(\bar{\boldsymbol{O}}_k)}} \tag{2.30}$$

where $\lambda(\cdot)$ means the eigenvalue of $(\cdot)$ and $\boldsymbol{\mathcal{R}}_k$ is the augmented observation error covariance matrix matching the dimension of $\boldsymbol{O}_k$. $OD \in [0, 1]$ and a larger value of OD means better observability.

At early filtering steps, instead of feeding back the state estimate as the control input, we give random control commands to the motorized stages to increase phase diversity. If we were to feed back the state estimate as the control input at the very beginning, the system would quickly converge to a local optimum and the state perturbations would be insufficient to produce enough phase diversity for correct wavefront reconstruction. The amount of perturbations required depends on the initial guesses of the state and its covariance. In our algorithm, we set random control input $\{\Delta D_{ij}\} \sim U(-1, 1)$ mm and $\{\Delta T_{ij}\} \sim U(-0.2, 0.2)^\circ$ for the first 25 steps and start the closed-loop linear-quadratic feedback control at step 26. Such settings were based on our past simulation and experimental experience.

Figure 2.12 shows how the degree of observability change at different filter steps. Initially, the random walk quickly increases the OD by approximately 1.5 orders of magnitude, meaning the Kalman filter is accumulating an internal knowledge of the phase without correcting for the misalignment. Then the growth slows down and the OD universally passes the maximum value by step 25, meaning the system knowledge provided by the random walk has saturated. Therefore, the current settings for the random walk appear to be optimal. Once the feedback control starts, the misalignment states start to converge and the observability decreases monotonically, indicating that the random walk should not be ended too early.

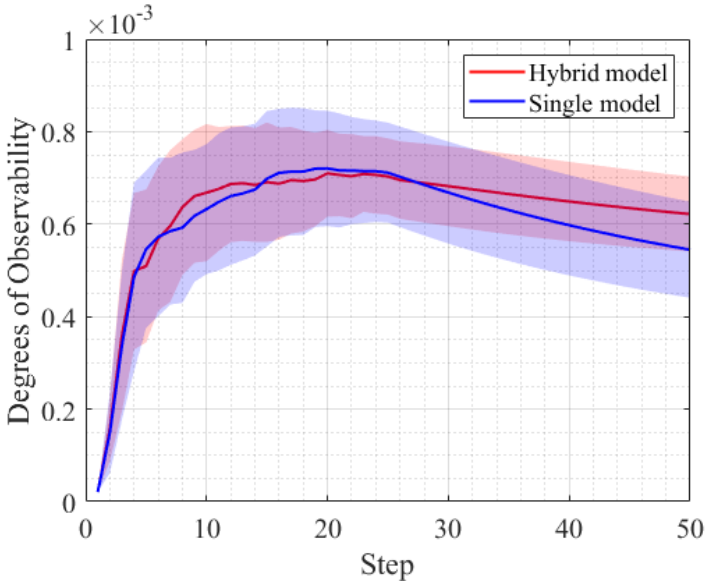

Figure 2.12: RMS degrees of observability with one-sigma confidence interval in 25 IF test runs using single measurement model and hybrid measurement model respectively.

Because the optical system is nonlinear, we use a hybrid model estimator instead of a single model estimator for better converge. The hybrid model estimator involves multiple measurement functions derived for different misalignment range and the filter linearizes different models depending on the current state estimate. Initially we use the measurement function derived for $\{D_{ij}\} \sim U(-1, 1)$ mm and $\{T_{ij}\} \sim U(-0.2, 0.2)^\circ$. Once the feedback control starts, we constantly check the state estimation. If the estimated states get within $\{D_{ij}\} \sim U(-0.01, 0.01)$ mm and $\{T_{ij}\} \sim U(-0.01, 0.01)^\circ$, we switch to the measurement function derived for this range. Figure 2.12 compares the OD between when using the hybrid model and when using the single model for $\{D_{ij}\} \sim U(-1, 1)$ mm and $\{T_{ij}\} \sim U(-0.2, 0.2)^\circ$ only. The hybrid model shows a slower decrease in OD, consistent with the fact that the hybrid model works better than the single model in our simulation. These two estimators have the same implementation of the random walk, so the OD difference in the first 25

steps are numerical variations. We can potentially add more models for smaller ranges of misalignment and push for further convergence.

2.6 Experiments

2.6.1 Experimental Setup

Figure 2.13 shows the experimental setup where all the labeled components are listed in Table 2.3. The setup meets the mechanical constraints such as the dynamic range of the actuators and avoids potential vignetting from optics mounting structures. We manually align all the components initially. In simulation, every other DoF besides $\boldsymbol{x}$ is perfectly aligned and fixed while an initial misalignment is introduced to $\boldsymbol{x}$ to match the estimated manual alignment error. In experiments, manual alignment error exists in the rest of the system and can theoretically be up to ±1 mm for positions and ±0.2° for orientations. Due to poor knowledge of the positioning ground truth, we need to evaluate the final automated alignment quality through repeatability tests by looking at the focal plane image.

MATLAB interacts with the camera through a client and server model, shown in Figure 2.14. The server initiates the camera once and remains on standby, constantly checking the registry until it receives a image request. The time cost is around $2.5s$ per image. Mex file runs C++ subroutine in Matlab, uses the API to control the camera and take images, but outputs image data in Matlab format. MATLAB interacts with the actuator controllers through ActiveX®control.

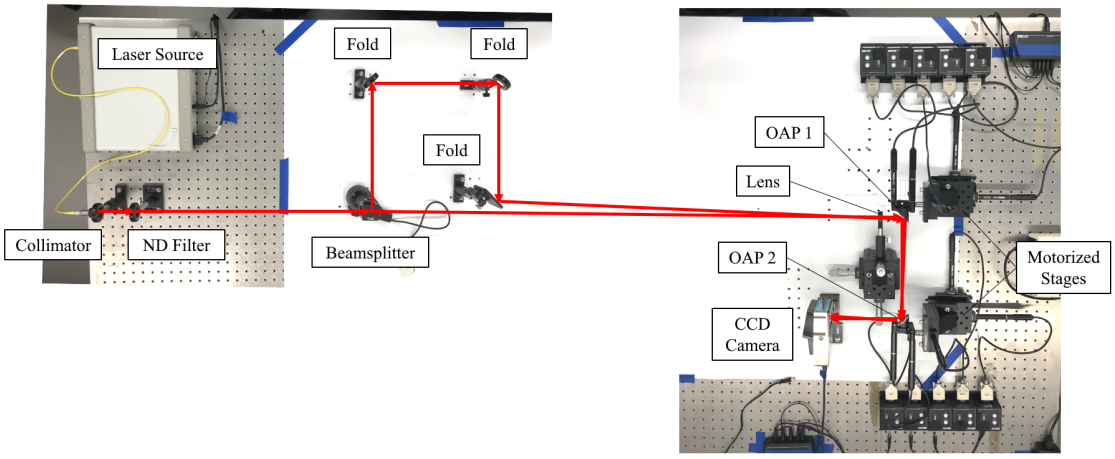

Figure 2.13: Experimental setup with red arrows tracing the beam path. The two parts of the optical system are photographed separately and stitched together. The spacing between the two photos is out of scale. [58]

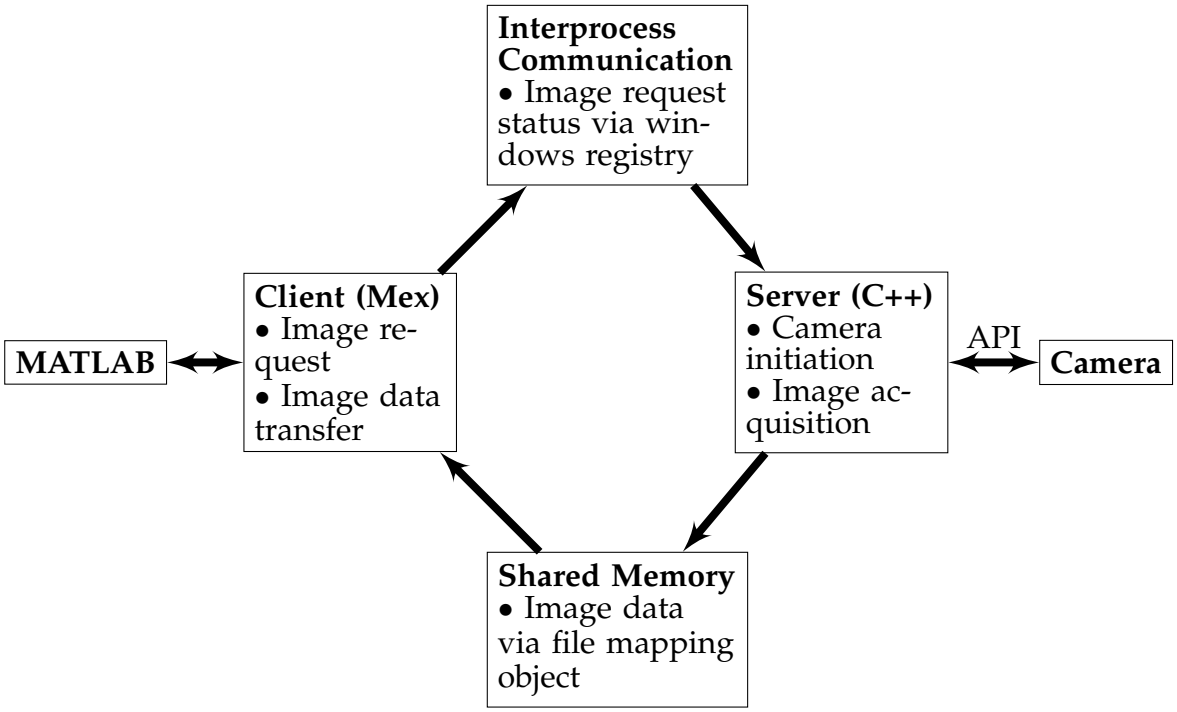

Figure 2.14: Diagram of camera hardware and software interface

Table 2.3: Specifications of the optical components and devices

Components	Model	Description
Laser	Thorlabs MCLS1-635	Wavelength: 635 nm
Collimator	Thorlabs TC25FC-633	Beam diameter ($1/e^2$): 4.67 mm
Beamsplitter	Thorlabs BS013	
Fold Mirrors	Thorlabs PF10-03-P01	Protected Silver Mirror
Concave Lens	Thorlabs LD1613-A	Focal length: -100.3 mm
OAP 1	Thorlabs MPD169-P01	Reflected focal length: 152.4 mm
OAP 2	Thorlabs MPD149-P01	Reflected focal length: 101.6 mm
Camera	Apogee Ascent A694	Array size: 2750 x 2200 pixels Pixel size: 4.54 × 4.54 μm
Motorized Translation Stages	Thorlabs MT3-Z9	
Tip-tilt Stages	Thorlabs KM100	
Tip-tilt Actuators	Thorlabs Z912B	

2.6.2 Experimental Limitations

There are many factors, such as manual misalignment error in stationary components and modeling errors, that contribute to focal plane image error. The filter attributes all these external error sources to misalignment in $\boldsymbol{x}$, resulting in large state estimation errors.

Among all error sources, experiments and simulation show the filter performance is extremely sensitive to the manual alignment error in T_{y1}. This is fundamentally different from our automated refractive systems alignment [22] where the corresponding optics are axisymmetric. Figure 2.15 shows how the automated alignment residual increases with manual alignment error in T_{y1} in simu-

lation when every other DoF is perfectly aligned. When T_{y1} is small, the filter is trying to compensate for T_{y1} by rearranging the OAPs into a slightly misaligned state to minimize the perturbation of the on-axis spot and the wavefront error in the system's narrow field of view, according to the measurement function [58]. However, when T_{y1} gets large, the amount of deviation required for the OAPs exceeds the geometric constraints, the on-axis spot goes out of the system's field of view and the filter has to terminate early. The terminating step gets progressively earlier as T_{y1} goes beyond 0.3°.

Currently, we are experiencing the same issue in experiments. The OAPs' DoFs T_{y1} and T_{y2} are neglected in our automated DoFs (Equation 2.1) because the current within-the-budget off-the-shelf hardware does not support them to be motorized. Instead, we put OAP 1 on a rotary plate to manually adjust T_{y1}. We evaluate the manual alignment of T_{y1} through its coupling with D_{x2}. By perturbing both states simultaneously and tracing the center of of on-axis spot in the focal plane, we try to constrain the manual alignment error in T_{y1}. However, because many factors are contributing to the state estimation error, we have not achieved the manual alignment accuracy required to replicate our simulation results in experiments. As a result, the filter currently has large estimation error and would command large movements to the OAPs, causing the laser beam to go off the optics. Once the filter completely loses observability, it has to abort. We implemented inequality constraints on the iterated extended Kalman filter using active set methods [106] to constrain the state estimation within reasonable range but there was little improvement. Therefore, we conclude that our experiment has reached its practical limits.

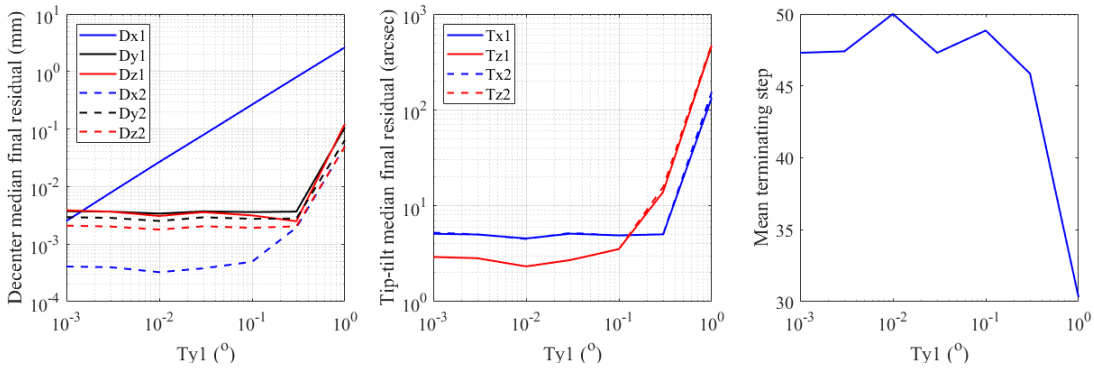

Figure 2.15: (Left and middle) Simulated final alignment residual (absolute value) when the filter terminates given different manual alignment error in T_{y1}. (Right) Simulated step at which the filter terminates or is forced to terminate given different manual alignment error in T_{y1}. The filter either terminate automatically at the maximum step number 50 or when the on-axis spot goes out of the field of view. The T_{y1} values are unsigned because it has nearly equal effect on the absolute residual disregarding the sign. The medians and the means are taken over 10 test runs for each T_{y1} value.

2.7 Conclusion

This study presents an innovative automated alignment method for reflective optical systems using focal plane sensing, aimed at enhancing efficiency and flexibility in optical assembly, alignment, adjustment, and maintenance. By leveraging the existing focal plane sensor in an optical system, our technique eliminates the need for dedicated wavefront sensors, reducing costs and complexity while improving optical throughput and avoiding non-common path errors. Notably, we applied our method to an optical system with multiple OAPs, which introduce significant challenges in alignment and modeling. Our wavefront control strategy employs several variants of Kalman filtering to effectively address modeling errors, nonlinearity, noise, and uncertainties. Simulation results demonstrate that our automated alignment algorithm achieved a linear ac-

curacy of $< 5\,\mu$m, angular accuracy of < 6 arcsec and wavefront error of < 0.3 nm ($< 10^{-4}$ waves). This approach holds potential for application in optical systems capable of tolerating such levels of alignment and wavefront errors. Furthermore, when combined with our lab's previous work on refractive systems, this method could be generalized to more complex optical systems comprising both reflective and refractive elements.

The current theoretical alignment accuracy is constrained by the multi-state coupling effect, wherein the measurement does not correspond to a unique set of misalignments. Future simulation work will address these challenges by:
(1) Reducing the multi-state coupling effect by decoupling the stage movements and assessing the performance through observability analysis.
(2) Improving the models by replacing the current measurement function with convolutional neural networks to capture additional image features.
(3) Incorporating the off-axis beam as supplementary measurement.

Experimental limitations stem from the manual alignment accuracy of DOFs not actively controlled by the algorithm. To mitigate this, we propose the following:
(1) Evaluating and characterizing manual misalignment errors in T_y and other system components using focal plane wavefront reconstruction methods such as the Gerchberg-Saxton algorithm or interferometry.
(2) Enhancing filter robustness to system errors and uncertainties by refining the process noise model and integrating sensor noise.
(3) Modifying the simulation to account for T_y as an additional DOF and employing a 6-DOF motorized stage, such as a hexapod, in future experiments.

In conclusion, our work demonstrates a promising step toward the automation of optical system alignment, paving the way for broader applications in advanced optical setups. By addressing current limitations and expanding the method's scope, we aim to further refine this technique for deployment in increasingly complex optical systems.

Publications. This work is published in Li and Savransky 2020 [58], Li and Savransky 2021 [59], and Li and Savransky 2023 [60]. Data underlying the results presented in this chapter are available in Li and Savransky 2023 [57].

CHAPTER 3

CONTROL SYSTEM DESIGN FOR THE GEMINI PLANET IMAGER 2.0'S ADAPTIVE OPTICS SYSTEM

3.1 Introduction

Adaptive optics (AO) enables real-time correction of wavefront distortions caused by atmospheric turbulence, thermal gradients, optical imperfections, and other environmental factors in astronomical telescopes, delivering diffraction-limited performance essential for high-contrast imaging. The compensation for wavefront aberrations is achieved through a combination of three key components: wavefront sensors, deformable mirrors, and control algorithms. Wavefront sensors detect distortions by analyzing the incoming light, while deformable mirrors adjust their surface shape to counteract these distortions. Control algorithms coordinate these components, translating sensor data into mirror adjustments with high speed and accuracy.

The effectiveness of an AO system heavily depends on the performance of its control algorithms, which must handle the inherent complexities of real-time operation in noisy, nonlinear, and uncertain environments. Over the years, numerous control techniques have been developed to optimize the performance of adaptive optics systems. Classical methods such as proportional-integral-derivative (PID) controllers and least-squares optimization have been widely used for their simplicity and effectiveness. However, as AO applications grow more demanding, advanced techniques are increasingly employed. Kalman filtering, model predictive control, and machine learning-based approaches

have demonstrated significant improvements in handling system uncertainties, noise, and computational constraints. These advanced methods enable AO systems to achieve higher correction bandwidths, better robustness, and greater adaptability to complex operating conditions.

3.2 System Transfer Function Characterization

For GPI 2.0 AO upgrade, we retain GPI 1.0's control philosophy and optimize the control gain for the new pyramid WFS and the faster real-time computer (RTC). Our hybrid continuous-discrete AO control system is represented by the block diagram in Figure 3.1. [97, 89] In this context, we use $\bullet(t)$ to denote variables in the continuous-time domain, $\bullet(kT)$ for the discrete-time domain, $\bullet(s)$ for the Laplace domain after the Laplace transform, and $\bullet(s)$ for the Z-domain after the Z-transform. In discrete-time domain notation, k refers to the step number and T refers to the frame time.

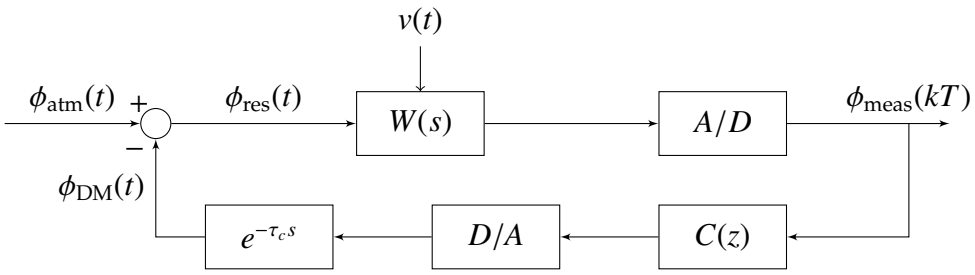

Figure 3.1: Adaptive optics system control block diagram

The input turbulent wavefront $\phi_{\text{atm}}(t)$ is continuously compensated by DMs' phase output $\phi_{\text{DM}}(t)$. The WFS measures the residual wavefront $\phi_{\text{res}}(t)$ with sensor noise $v(t)$. The WFS behavior can be characterized as an integration over one

sampling period T, with a transfer function expressed as

$$W(s) = \frac{1 - e^{-Ts}}{Ts} \tag{3.1}$$

The RTC records the WFS signals, introducing a pure time delay, $e^{-\tau_r s}$, due to the WFS read-out time τ_r. This process is represented by the A/D block, which converts the continuous input wavefront into a discrete wavefront measurement, $\phi_{\text{meas}}(kT)$. The discrete integral controller $C(z)$ processes these measurements and computes discrete control voltages for the tip-tilt stages and DMs. These devices hold the computed voltages constant during each sampling period, T. The behavior of the tip-tilt stages and DMs is modeled as a D/A converter with a zero-order hold transfer function

$$D(s) = \frac{1 - e^{-Ts}}{Ts} \tag{3.2}$$

Although $W(s)$ and $D(s)$ share the same mathematical form, they represent different physical processes. Additionally, the system's computational latency is modeled as $e^{-\tau_c s}$ where τ_c represents the total computational delay for one control loop iteration.

The AO operating conditions and performance specifications are summarized in Table 3.1. To meet these specifications, we utilize a discrete integral controller with transfer function

$$C(z) = \frac{g}{1 - cz^{-1}} \tag{3.3}$$

where c is the integrator constant and g is the control gain. Instead of a standard pure integrator ($c = 1$), we employ a leaky integrator with $c = 0.999$ to

Table 3.1: Compare of GPI's and GPI 2.0's AO control system specifications and estimated performance

	GPI as built	GPI 2.0 best case
Rate (kHz)	1	2
Read time (μs)	664	330
Comp time (AOC) (μs)	695	100
CT IC max gain	0.34723	0.48977
CT IC bandwidth (Hz)	50	139

gradually reduce the accumulation of the integral term over time. This method, commonly used in adaptive optics, provides several advantages, including mitigating integral windup, improving stability by reducing oscillations, enhancing system responsiveness, and increasing robustness against disturbances and noise. Meanwhile, g is mode-specific and optimized to meet stability margin requirements.

3.3 Control Gain Optimization

To analyze the system's stability margins, the system transfer function must be expressed in continuous form. This involves converting the controller's transfer function from the Z-domain to the Laplace domain using the matched pole-zero method, which approximates $z = e^{sT}$ [25]. We first rewrite the controller transfer function in a form suitable for transformation

$$C(z) = \frac{gz}{z - c} = \frac{gz}{z - e^{-aT}} \tag{3.4}$$

where $a = -\frac{log(c)}{T}$. Next, we apply the inverse Z-transform to obtain the discrete-time-domain representation of the controller

$$C(kT) = e^{-akT} \tag{3.5}$$

Finally, we apply the Laplace transform to derive the continuous-time-domain controller function

$$C(s) = \frac{1}{s + a} \tag{3.6}$$

It now becomes apparent that the effect of the controller is to add a pole at $s = -a$ in the Laplace domain, or equivalently, a pole at $z = c$ in the z-domain. By choosing $c < 1$ or equivalently $a > 0$, we place the pole within the region of stability in the z-plane (for which is a unit circle centered at the origin) and the s-plane (for which is the left-half plane).

We design and optimize the controller using classical frequency-domain control theory. The open-loop transfer function $L(s)$

$$L(s) = W(s)D(s)C(e^s)e^{-(\tau_r+\tau_c)s} \tag{3.7}$$

serves as a key tool for evaluating the stability and performance of the closed-loop system. A higher $L(s)$ value signifies improved tracking performance and disturbance rejection, while a lower $L(s)$ value indicates better sensor noise rejection and reduced actuator response. In our design, we prioritize tracking and disturbance rejection at low frequencies and noise rejection at high frequencies.

The error transfer function (ETF), in our case the ratio between the residual

wavefront and the input wavefront, characterizes the AO system performance at different temporal frequencies. The ETF is directly related to the open-loop transfer function

$$ETF(s) = \frac{\phi_{\text{res}}(s)}{\phi_{\text{atm}}(s)} = \frac{1}{1 + L(s)} \tag{3.8}$$

Since $L(s)$ and ETF are only mathematical models of the process, they exhibit discrepancy from the actual physical process, due to factors such as modeling error, noise, system aging and deterioration. However, the closed-loop feedback control inherently provides a level of robustness against such uncertainties and changes. Additional stability is ensured through gain margin and phase margin.

Our control system design objective is to determine the optimal control gain g for each mode while ensuring the system's robustness against modeling errors and noise. We choose our design specifications to maximize g, while maintaining a gain margin > 2.5 and a phase margin > 45°, for each possible value of the total time delay $\tau = \tau_r + \tau_c$. (Please refer to Section A.1 for the impact of stability margins on closed-loop system behaviors.) The analytical expressions for the open-loop gain and phase are

$$|L(i\omega)| = \frac{2g(1-\cos(\omega T))}{(\omega T)^2\sqrt{1-2c\cos(\omega T)+c^2}} \tag{3.9}$$

$$\angle L(i\omega) = \angle\left(\frac{-c\sin(\omega T)}{1-c\cos(\omega T)}\right) + 2\angle\left(\frac{\cos(\omega T)-1}{\sin(\omega T)}\right) - 2\pi\omega\tau \tag{3.10}$$

We formulate our constrained modal gain optimization problem as

$$\text{Objective:} \quad \min_g(-g) \tag{3.11}$$

$$\text{Constraint 1:} \quad |L(g,\omega,\tau)| > 2.5 \text{ or } |L(g,\omega,\tau)| < \tfrac{1}{2.5} \quad \forall \angle L(g,\omega,\tau) = 180^\circ \tag{3.12}$$

$$\text{Constraint 2:} \quad \angle L(g,\omega,\tau) > -135^\circ \quad \forall |L(g,\omega,\tau)| = 1 \tag{3.13}$$

We summarize the optimized gain values for different system time delays in Table 3.2. Figure 3.2 compares the Bode plot of the best-case ETF in GPI 2.0 to original GPI design. The results show an improvement in bandwidth from 50 Hz to 139 Hz, and a factor of 8 improvement in the rejection of slower signals in terms of power, which corresponds to reductions in mean-squared error and speckle intensity.

Table 3.2: Optimal loop gains at various delays

τ (ms)	0.800	0.758	0.717	0.676	0.635	0.594	0.553	0.512	0.471	0.430
g	0.306	0.319	0.333	0.349	0.367	0.385	0.407	0.431	0.457	0.488

3.4 Conclusion

In this work, we establish a control system design pipeline, where key features such as the error transfer function (ETF) and stability margins serve as essential tools for analyzing and fine-tuning the GPI 2.0 AO response. Furthermore, the results from this work are integrated into GPI 2.0 AO simulations [68], enabling more accurate predictions of real-time closed-loop behavior and system performance under the atmospheric conditions at the new site, Mauna Kea, Hawaii. By combining the ETF with an atmospheric model, we demonstrated its ability

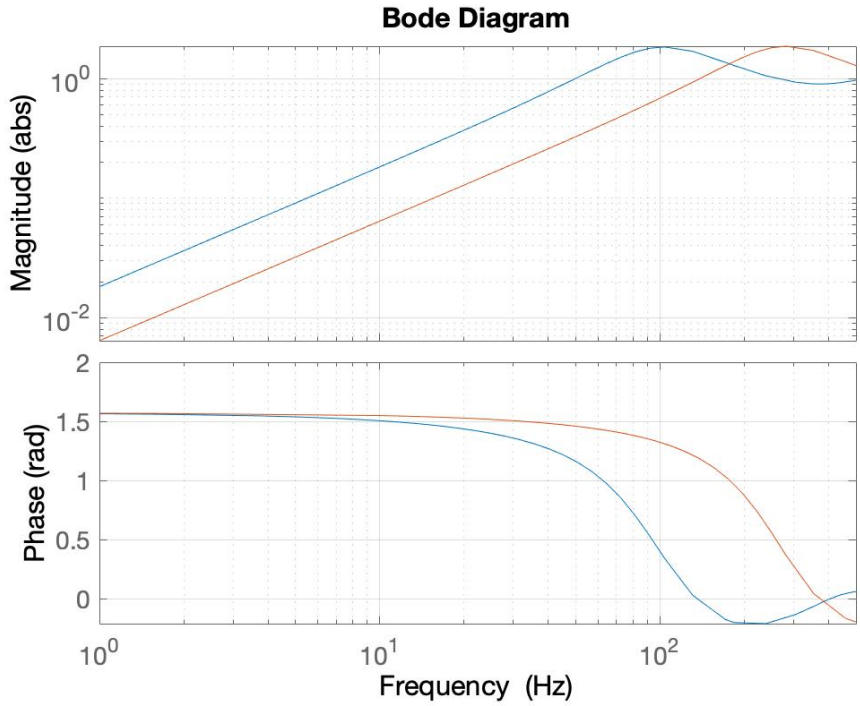

Figure 3.2: Bode plot for GPI and 2.0's error transfer functions. Blue: GPI as built. Red: GPI 2.0 best case.

to predict the residual wavefront. Specifically, the contrast enhancement can be quantified as follows:

$$\text{contrast enhancement} = \frac{\sum_i \sqrt{C^2_{N,\text{Hawaii}}(z_i)}\Delta z_i \times \left|\text{ETF}_{\text{GPI2}}\left(\frac{v_{\text{Hawaii}}(z_i)}{p}\right)\right|}{\sum_i \sqrt{C^2_{N,\text{Chile}}(z_i)}\Delta z_i \times \left|\text{ETF}_{\text{GPI}}\left(\frac{v_{\text{Chile}}(z_i)}{p}\right)\right|} \tag{3.14}$$

where $i = 0, 1, ..., 25$ represents the index of different atmosphere layers at altitudes z_i, C_N^2 denotes the structure function in the Kolmogorov power spectrum used to model atmospheric turbulence, v is the wind velocity, and p is speckle phase aberration mode length. This approach not only enhances our understanding of system performance but also provides a framework for optimizing

GPI 2.0 AO operation in diverse environmental conditions, with a particular focus on Mauna Kea's unique atmospheric dynamics. Future efforts may focus on refining these models and incorporating real-time data from Mauna Kea to further validate and optimize system performance.

Publications. This work is published in Madurowicz et al. 2020 [68].

CHAPTER 4

FOCAL PLANE WAVEFRONT CONTROL FOR THE GEMINI PLANET IMAGER 2.0'S CALIBRATION SYSTEM

4.1 Introduction

Adaptive optics (AO) systems are designed to correct wavefront distortions caused by atmospheric turbulence and optical imperfections, but residual wavefront error persists due to various practical limitations, such as system temporal lag, wavefront sensor noise, finite DM resolution, calibration errors, high-order aberrations and instrumental limitations. These residual aberrations, which can manifest as quasi-static, dynamic, or chromatic speckles in the image plane, represent a significant source of error and limit the achievable contrast of the instrument. ([10])

To address these limitations, it is essential to implement wavefront control directly at the final image plane. Various methods have been developed to achieve this goal, including: classical speckle nulling by [112] and [6], pair-wise probing with electric field conjugation (EFC) by [30] and [31], and implicit electric field conjugation (iEFC) by [37]. Speckle nulling requires minimal system modeling and calibration, but is time-consuming due to the need for numerous images, particularly when paired with conventional science instruments. EFC, while more efficient, relies on an end-to-end model of the optical system. Meanwhile, iEFC requires a linearized model between the DM modal coefficients and the modulated intensity measurements. Therefore, both EFC and iEFC are inherently susceptible to modeling error.

As we introduced in Section 1.3.2, GPI consists of three principal subsystems: the adaptive optics system (AO), the infrared calibration system (CAL) and the integral field spectrograph (IFS) science instrument. The CAL system is a high-order wavefront sensing system downstream of the AO system, and plays a critical role in compensating for AO residual wavefront error and non-common path aberrations (NCPAs).

As part of the GPI2 upgrade, CAL2 is getting upgraded from the dual-arm interferometer design with a high-order wavefront sensor (HOWFS) to a common-path interferometer design featuring a Fast Atmospheric Self-Coherent Camera (SCC). ([71], [73]) As shown in Fig. 4.1, in the original CAL1 design, the focal plane mask (FPM) was employed to pass the on-axis light (ref arm) while reflecting the off-axis (sci arm) light. The two beams would interfere at the HOWFS, which was implemented to operate at a few Hz. However, the HOWFS had very limited performance due to the non-common path vibrations introduced by the dual-arm configuration. The CAL2 upgrade will employ a new FPM design to reflect both on-axis and off-axis light along a single optical path, leading to substantial vibration resistance. Furthermore, the HOWFS will be replaced with a Fast Atmospheric Self-Coherent Camera (SCC). The SCC enables direct measurement of the focal plane electric field and integration of high-order wavefront control into AO loop. With its low noise and significantly faster integration time, the SCC could potentially enable speckle correction within tens of milliseconds and overcome the limitations of classical speckle nulling.

Therefore, we pursue classical speckle nulling as our method for focal-plane wavefront control in this paper. Our objective is to leverage only the intensity

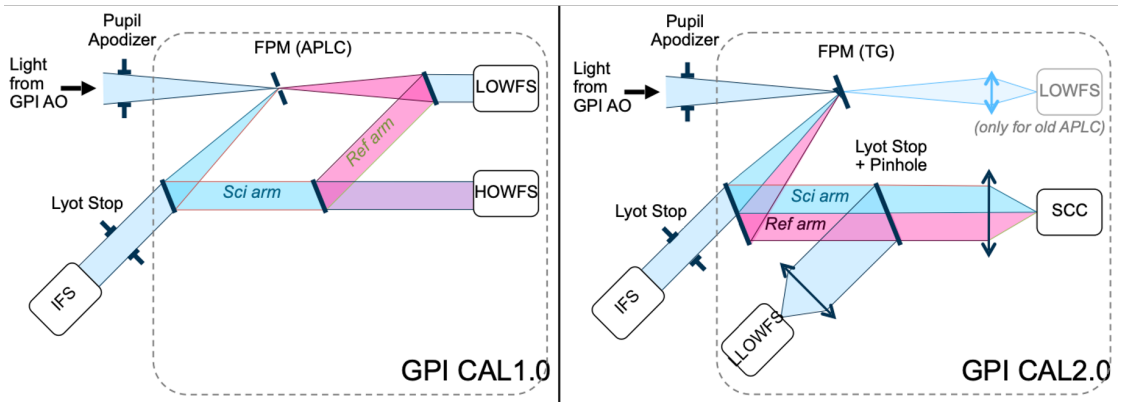

Figure 4.1: Simplified schematics of GPI CAL1 and CAL2 design for comparison. CAL1 employed a dual-arm interferometer design with a HOWFS, while CAL2 will pursue a common-path interferometer design featuring a SCC. ([71])

measurements provided by the SCC, and to establish a reliable backup strategy for the high-order wavefront sensing and control framework based on SCC electric field measurements. Section 4.2 outlines the theoretical foundation, incorporating principles from Fourier optics and feedback control theory. In Section 4.4, we describe the software tools used to simulate the residual wavefront errors (WFE), the GPI2 model and our wavefront sensing and control strategy. Section 4.5 evaluates the algorithm's performance in terms of contrast improvement and convergence speed. Finally, Section 4.6 discusses the algorithm's tolerance to misalignments and its optimization process.

4.2 Theory and Methodology

4.2.1 Focal-plane wavefront sensing

GPI follows the general formalism of an APLC which can be simplified into four successive optical planes, as shown in Figure 4.2: A. an apodized pupil

downstream of and conjugate to the telescope entrance pupil and DM planes, B. a focal plane with an occulting mask (FPM), C. a pupil with a Lyot stop, D. the final image plane.

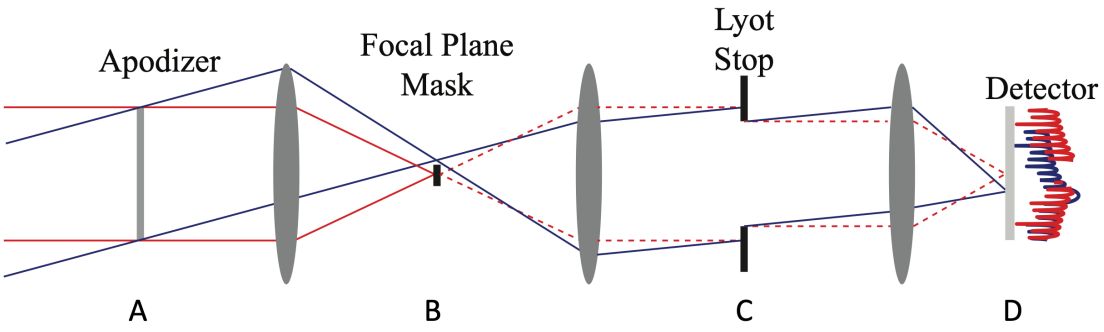

Figure 4.2: General formalism of APLC.

We define the telescope entrance pupil function Φ_{pupil}, the apodizer function $\Phi_{apodizer}$, the DM function Φ_{DM}, the FPM transmission function Φ_{FPM} and the Lyot stop transmission function Φ_{Lyot}. Assuming a wavefront Φ_{turb} arrives at the entrance pupil, the electric field in plane A is

$$\Psi_A \propto \Phi_{turb}\Phi_{pupil}\Phi_{DM}\Phi_{apodizer} \tag{4.1}$$

and the electric field in plane D is ([109])

$$\Psi_D \propto [\mathcal{F}(\Psi_A)\Phi_{FPM}] \circledast \mathcal{F}(\Phi_{Lyot}) \tag{4.2}$$

where $\mathcal{F}(\cdot)$ denotes Fourier transform and $\circledast$ denotes convolution. Therefore, the image in plane D is

$$I_D \propto |\Psi_D|^2 \propto |[\mathcal{F}(\Phi_{turb}\Phi_{pupil}\Phi_{DM}\Phi_{apodizer})\Phi_{FPM}] \circledast \mathcal{F}(\Phi_{Lyot})|^2 \tag{4.3}$$

The FPM function Φ_{FPM} and the Lyot stop function Φ_{Lyot} modify the Point Spread Function (PSF) only in terms of amplitude and phase symmetrically. If we decompose the turbulence Φ_{turb} into Fourier modes

$$\Phi_{turb} = \sum_i \boldsymbol{\phi}_{turb,i} \tag{4.4}$$

each sinusoidal wavefront $\boldsymbol{\phi}_{turb,i}$ would result in two symmetric peaks in I_D. In reverse, each speckle in the final image plane can be independently associated with a sinusoidal wavefront of spatial frequency (k_x, k_y) at the entrance pupil

$$\boldsymbol{\phi}_{turb,i}(\boldsymbol{r}) = B_i \, sin(k_{x,i}x + k_{y,i}y + \eta_i) \tag{4.5}$$

where $\boldsymbol{r} = (x, y)$ is the position vector in the optical plane. The turbulence Φ_{turb} can be nullified if the DM can shape the opposite wavefront

$$\Phi_{DM} = \sum_i \boldsymbol{\phi}_{DM,i} = \sum_i -\hat{\boldsymbol{\phi}}_{turb,i}(\boldsymbol{r}) = -\sum_i \hat{B}_i \, sin(\hat{k}_{x,i}x + \hat{k}_{y,i}y + \hat{\eta}_i) \tag{4.6}$$

where $\hat{(\cdot)}$ denotes the estimated values through wavefront sensing. Below we describe the process to estimate the four variables of each $\boldsymbol{\phi}_{DM,i}$: amplitude $\hat{B}_i$, spatial frequency $(\hat{k}_{x,i}, \hat{k}_{y,i})$ and phase $\hat{\eta}_i$ from one single focal-plane image.

According to Fourier optics, the mapping between spatial frequency $(\hat{k}_x, \hat{k}_y)$ and speckle position $(x_{speckle}, y_{speckle})$ can be approximated by a linear model

$$\begin{bmatrix} x_{speckle} \\ y_{speckle} \end{bmatrix} = \pm d \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} k_x \\ k_y \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \tag{4.7}$$

where d is a scaling factor internal to the optical system; the derotation angle θ and the optical axis offset (x_0, y_0) are due to misalignment between the DM and the image sensor. These parameters can be calibrated by inducing speckles from benchmark DM shapes, following the steps in ([101]). We can then reverse the equation to derive $(\hat{k}_x, \hat{k}_y)$ from $(x_{speckle}, y_{speckle})$.

Next, we relate amplitude $\hat{B}$ to the speckle peak intensity at its geometric center and by an exponential equation

$$I(x_{speckle} - x_0, y_{speckle} - y_0) = c_1' B^2 e^{c_2'(k_x^2 + k_y^2)} \tag{4.8}$$

to capture the fact that as the DM approaches higher-order modes, its phase shaping capability decreases due to its limited bandwidth in discrete sampling, so the energy of the reproduced opposite speckle decreases. Because of the linear relationship between (k_x, k_y) and $(x_{speckle}, y_{speckle})$ in Equation 4.7, we can rewrite Equation 4.8 in terms of radial distances in the image plane

$$I(\boldsymbol{r}_{speckle} - \boldsymbol{r_0}) = c_1 B^2 e^{c_2(x_{speckle}^2 + y_{speckle}^2)} = c_1 B^2 e^{c_2 r_{speckle}^2} \tag{4.9}$$

where c_1 and c_2 are scaling constants. We derive c_1 and c_2 by applying different DM sinusoidal patterns with the same amplitude to induce speckles that cover our region of interest (ROI) in the image plane. The radial distribution of speckle intensity is shown in Figure 4.3, along with the intensity mapping model derived from a least-squares fit. Now given the speckle intensity $I(x_{speckle}, y_{speckle})$ and position, we can reverse the equation to determine amplitude $\hat{B}$.

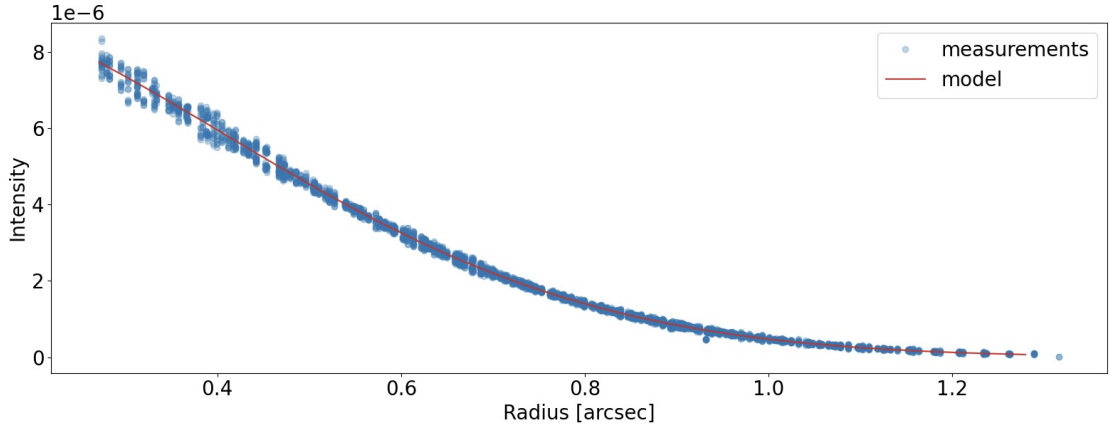

Figure 4.3: Radially-symmetric speckle intensity vs angular separation, induced by DM sinusoidal patterns with different spatial frequency and a constant amplitude of 10 nm. Blue dots are the measurements from simulated images. Red curve is the least-squares fitted model of Equation 4.9.

Last, the phase η for destructive interference can be estimated from phase probing. We place 8 trial DM shapes $\Phi_{dm,j}$ with 8 evenly-spaced trial phases η'_j between $[0, 2\pi)$

$$\Phi_{dm,j} = -\sum_i \hat{B}_i \, sin(\hat{k}_{x,i}x + \hat{k}_{y,i}y + \eta'_j) \quad , \quad j = 1, 2, ..., 8 \tag{4.10}$$
$$\eta'_j = 0, \pi/4, \pi/2, ..., 7\pi/4$$

We take a focal-plane image for each $\Phi_{dm,j}$. By relating the speckle intensity at each position $I_j(x_{speckle,i}, y_{speckle,i})$ to η'_j, we can interpolate for the phase $\hat{\eta}_i$ that causes the most destructive interference. ([101]).

However, because of DM's influence function and image sensor pixel crosstalk, neighboring speckles have influence over each other and should not be corrected simultaneously. Therefore, we iteratively prioritize the speckles based

on their intensity and correct a limited number of them each time. Our methodology resembles with Fourier modal control, and allows fast Fourier-transform reconstruction of the superposed sinusoidal wavefront patterns and direct control of specific regions in the image plane.

4.2.2 Closed-loop control

We apply a closed-loop wavefront sensing and control strategy, as illustrated in Figure. 4.2.2.

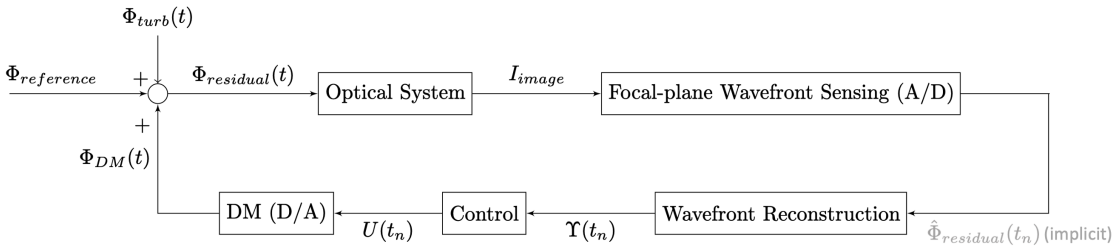

Figure 4.4: Closed-loop focal-plane wavefront sensing and control scheme

The control reference $\Phi_{reference}(t)$ is defined as a flat wavefront with no aberrations. The incoming dynamic phase aberrations $\Phi_{turb}(t)$, which arise from non-common-path errors and residual errors of the AO system, acts as a disturbance to the system. The deformable mirror introduces a corrective phase $\Phi_{DM}(t)$, and the superposition of all these three produces the residual wavefront $\Phi_{residual}(t)$. Although $\Phi_{DM}(t)$ has a plus sign in the summation block, the system operates as a negative feedback loop. This is because the DM applies a phase opposite to the disturbance, aiming to minimize the residual phase and ideally drive it toward zero.

This residual wavefront $\Phi_{residual}(t)$ propagates through the optical system and

produces an intensity pattern $I_{image}(t)$ in the image plane. This intensity measurement is then processed by the focal-plane wavefront sensing and reconstruction module, which converts the image-based measurements into theoretical DM commands $\Upsilon(t_n)$ intended to suppress the unwanted speckles created by the residual wavefront. Due to the finite integration time of the camera, this process effectively functions as an analog-to-digital conversion.

The controller then processes $\Upsilon(t_n)$ to compute the optimal DM command $U(t_n)$, which is sent to the DM hardware. The DM updates its surface accordingly and maintains its shape between updates, effectively acting as a digital-to-analog converter with a zero-order hold.

For the controller, we employ a leaky integrator architecture.

$$U(t_n) = f\ U(t_{n-1}) + g\ \Upsilon(t_n) \tag{4.11}$$

where f is the leaky factor and g is the integral gain.

The leaky integrator is commonly used in AO to reject large and slow disturbances like turbulence and compensate for modeling error, bias and uncertainties. The traditional pure integrator with integral gain $f = 1$ is suited to correcting for low order modes with high turbulence. Meanwhile, a leaky integrator with $f < 1$ can better correct for high order modes with low turbulence, more effectively reduce disturbance at low temporal frequencies and significantly reduce DM stroke without sacrificing the correction quality. ([1]) In our research, we optimize the integral gain to be $g = 0.6$ as described in Section 4.5.1 and

choose our legacy leaky factor $f = 0.999$ according to ([1]) to better suit our application.

4.3 Preliminary Experiment

To validate our theory, we traveled to NRC - Herzberg Astronomy and Astrophysics, Victoria, Canada, to conduct preliminary experiments on the under-development Subaru Pathfinder Instrument for Detecting Exoplanets and Retrieving Spectra (SPIDERS), because CAL2 will utilize a similar architecture to SPIDERS. [74] The main optical components on SPIDERS include a reproduced Subaru entrance pupil, an Apodized Pupil Lyot Coronagraph (APLC), a 468 actuator ALPAO DM (circular with 24-actuator diameter), a pupil-plane Lyot-based Low-Order Wavefront Sensor (LLOWFS) and an SCC. [53] Our speckle nulling algorithm is tested at 1.65 μm wavelength with a 2% bandwidth, and with an image integration time of 10 μs. The speckle nulling algorithm is run in parallel with the LLOWFS while the SCC is disabled. The system has a residual vibration at 200 Hz. The current initial contrast suffers from some defects in the off-axis ellipsoid mirrors which will be replaced for deployment.

Limited by circumstances in the preliminary experiment, we (1) Use (k, l) to annotate DM wavefront spatial frequency in the (x, y) direction, and measure the spatial frequencies in terms of cycles per pupil; (2) Restrict our control space to discrete, integer-valued spatial frequencies. Figure 4.5 shows the 1-to-2 discrete spatial mapping between the spatial frequency (k, l) in the DM plane and the speckle location (x, y) in the image plane, consistent with Equation 4.7). We are limited to spatial frequency $\in [-11, 11]$ to stay below SPIDERS DM's Nyquist frequency. Moreover, the color map of Figure 4.5 captures the relative intensity

scaling of each DM Fourier mode (Equation 4.8).

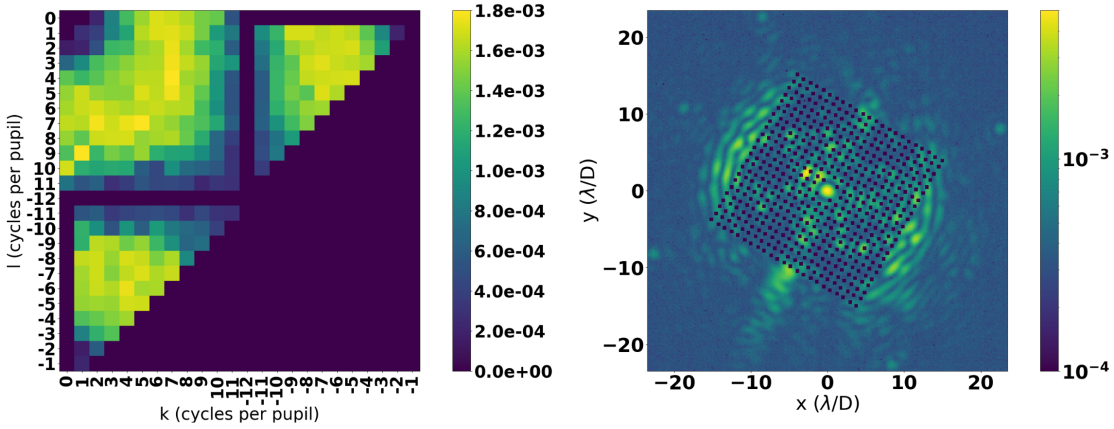

Figure 4.5: Fourier modal control space. (Left) DM frequency domain. The spatial frequencies are currently integer-valued in the range of [-11, 11] due to the number of DM actuators available. We do not attempt to control the highest or lowest spatial frequencies. The calibrated focal plane intensity associated with each sinusoid is calculated and shown in terms of 1σ contrast. (Right) Image plane. The dot grid shows the full spatial extent of the control space, where each dot is associated with one spatial frequency (k, l).

We further tailor the experiment for SPIDERS by (1) Constraining our control space to an IWA of 4 λ/D to avoid the central obscuration due to the APLC, and an OWA of 9 λ/D to ease the DM effort at high spatial frequencies; (2) Ignoring the rectangular regions that passes through the center which are contaminated by large amplitude errors uncorrectable by speckle nulling. Figure 4.6 shows our final control space in DM's frequency domain and the corresponding ROI in the image plane.

Figure 4.7 shows an example crosstalk mask we apply to avoid simultaneous control of the neighboring modes, and the mask translates to an annular region around the speckle in the image plane.

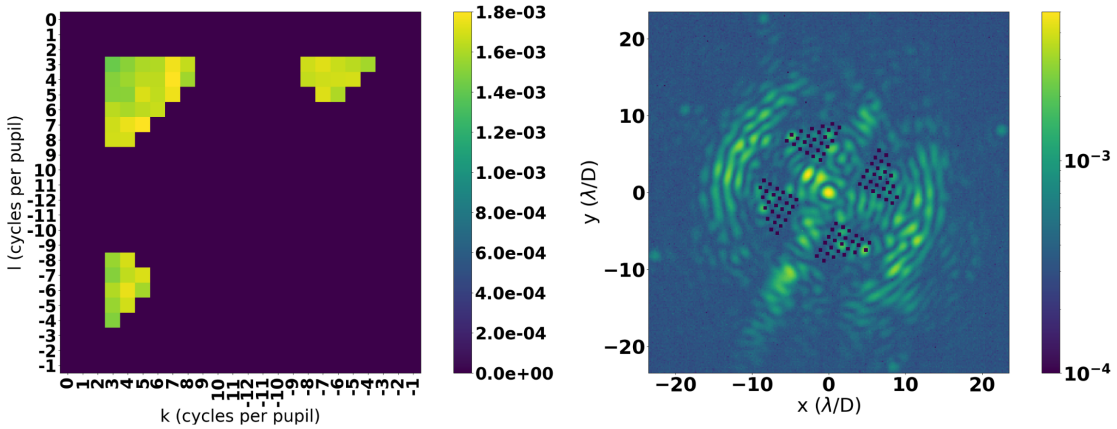

Figure 4.6: Further restricted control space from 4 to 9 λ/D. (Left) DM frequency domain. (Right) Image plane. The speckles with significant brightness in the tilted narrow rectangular regions that passes through the center are large amplitude errors not correctable by speckle nulling.

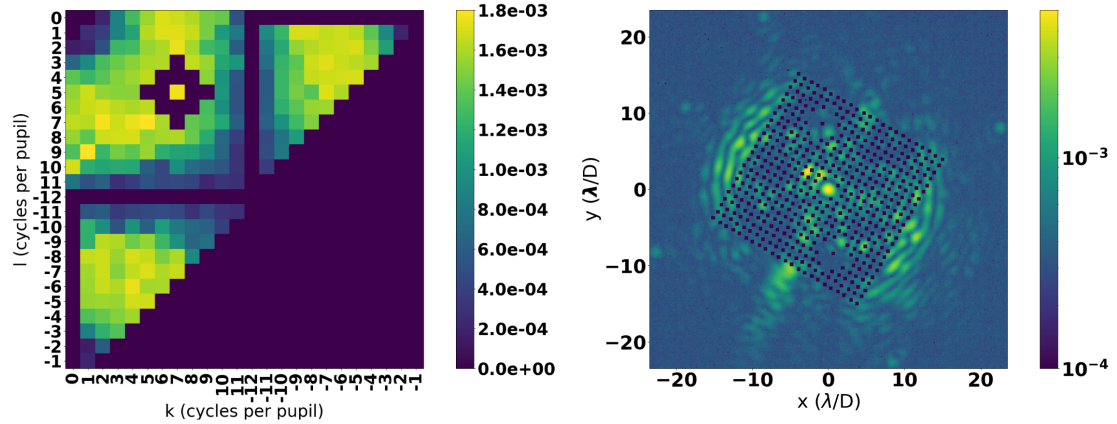

Figure 4.7: Example of blocking out neighboring modes around the mode being controlled to avoid crosstalk. (Left) DM frequency domain. (Right) Image plane.

We inject 5 pairs of speckles with purely real phases at the spatial frequencies shown in Figure 4.8. We apply a leaky integrator with a integral gain of 0.4 and an leaky factor of 0.999 to generate DM commands. Figure 4.9 shows the improvement in the point spread function (PSF) after 9 iterations, and Figure

4.10 shows the evolution of 1σ contrast throughout iterations. We achieved a maximal local contrast improvement factor of 3.2, validating that our algorithm is capable of mitigating speckle noise.

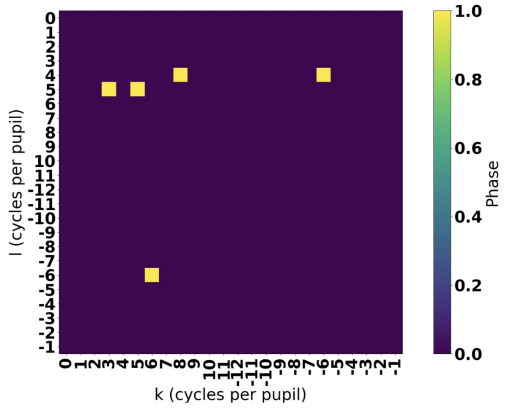

Figure 4.8: (Left) Injected speckles.

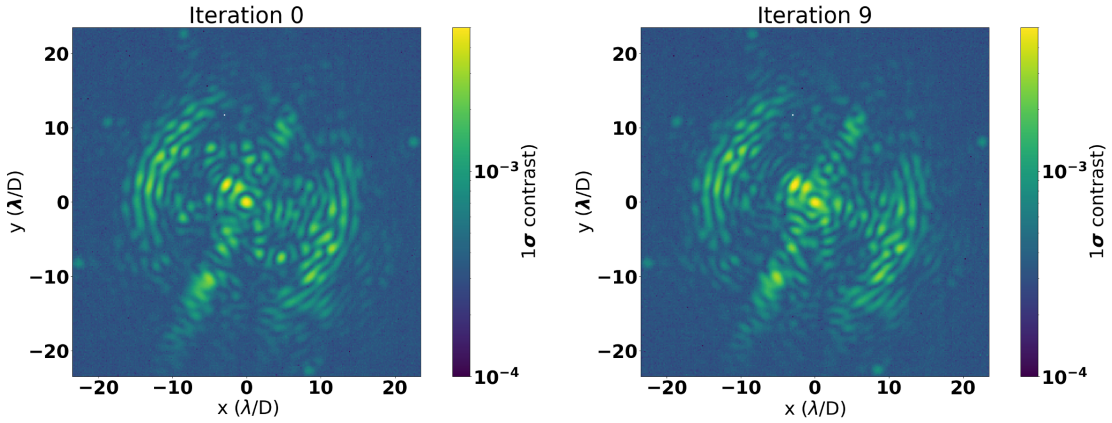

Figure 4.9: (Left) Initial PSF with injected speckles. (Right) PSF at speckle nulling iteration 9.

However, the amount of contrast improvement fell short of expectations, and the large oscillations in the contrast vs. iterations curve in Figure 4.10 indicates that our system is suboptimal and unstable. The major error source turns out to be the insufficient crosstalk mask size to fully occult a speckle. As a result, our algorithm would misidentify one speckle as multiple duplicates and reach wrong phase estimation. Moreover, discrete control space only allows coarse

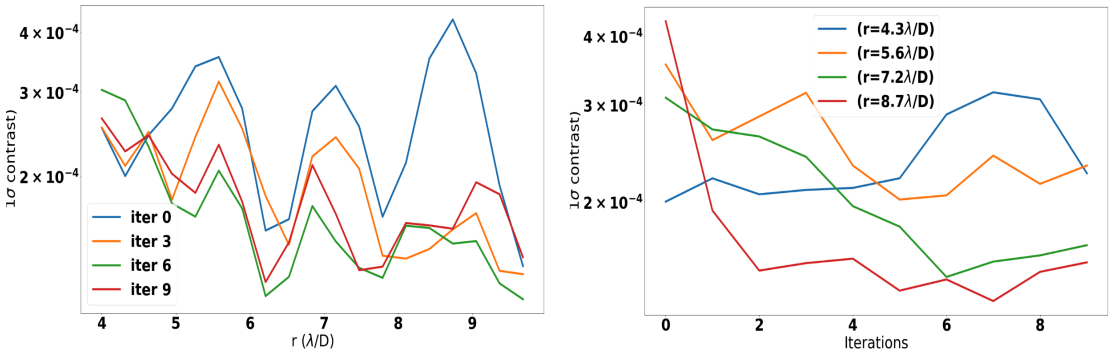

Figure 4.10: (Left) 1σ contrast as a function of angular separation for different iterations. (Right) 1σ contrast as a function of iterations for different angular separations.

control of the image plane and significantly limits the achievable contrast. Such findings led to our comprehensive system simulation, enhanced control strategy with continuous control space, algorithm optimization and improved results in the upcoming sections.

Publications. This work is published in Li et al. 2023 [61].

4.4 Simulation

GPI 2.0 coronagraph is modeled using GPIPSFS: GPI Point Spread Function Simulation Toolkit (developed by [86]) using Fraunhofer diffraction. The Gemini telescope primary mirror follows the standard design of a Cassegrain telescope. The DMs operate in a woofer-tweeter configuration. The woofer DM provides high-stroke, low-spatial-frequency correction, while the tweeter DM provides low-stroke, high-spatial-frequency correction. The woofer DM is a 9×9-actuator piezo-electric DM with a calibrated single-actuator stroke of ± 3.4 μm around bias voltage. The woofer is mounted on a tip–tilt stage with a cali-

brated single-axis stroke of ±1.8 arcsec around bias voltage. The tweeter DM is a 48×48 subregion of the 64×64 microelectromechanical system (MEMS) mirror, and has a calibrated single-actuator stroke of ±1.07 μm around bias voltage. ([90]) There are five dysfunctional actuators in the tweeter DM so we mask them out in the apodizer and the Lyot stop. The DM face sheet's profile is modeled as a 2D Gaussian sum over the actuator command. The GPI 2.0 apodizers and the Lyot stops are upgraded and numerically optimized for different bands by [83] to improve raw contrast at closer inner working angles and core throughput. Figure 4.11 shows GPI2's broadband optics as integrated in GPIPSFS. When there is no WFE, GPI2 in broadband could theoretically reach a raw contrast of 10^{-7} as close as 0.25 arcsec, as shown in Figure 4.17.

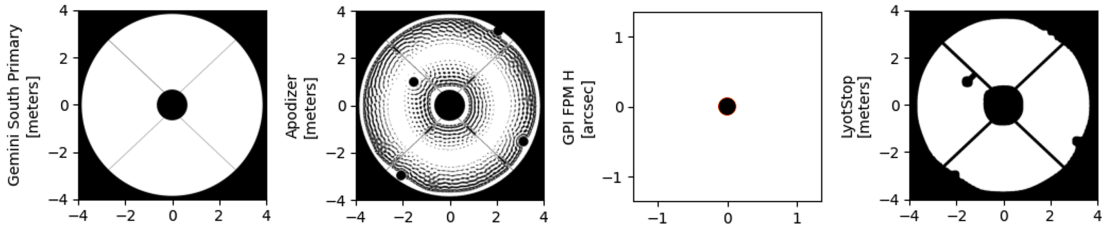

Figure 4.11: GPI2 broadband design featuring a new apodizer and a new Lyot mask. The apodizer and Lyot stop are illustrated with dimensions as if viewed from the telescope entrance pupil's optical space.

We simulate the input residual WFE $\Phi_{turb}(t)$ in Figure 4.2.2 using PyrAmid Simulator Software for Adaptive opTics Arcetri (PASSATA) developed by [2]. PASSATA is an IDL and CUDA based software capable of performing Monte-Carlo end-to-end simulations on multi-conjugate AO systems with pyramid wavefront sensors. Our system includes: a natural on-axis 8-magnitude source, a 7-layer atmosphere with 0.54" seeing, a 60x60 sub-aperture pyramid WFS, a classic CCD detector, a real-time computer - consisting of a slope computer, a

modal reconstructor and a modal integrator controller - a DM capable to control 1603 Zernike modes and a camera that computes H-band PSFs.

We simulate 10000 frames of GPI 2.0 AO residual WFE with 1-msec time steps for a 10-sec duration using PASSATA and select 25 evenly-spaced frames during 0:00:00.400 - 0:00:00.500 (H:MM:SS.MS)) as the input to our focal plane wavefront control simulation. We assume the residual WFE is static during each 1-msec time step. Figure 4.12 shows the residual WFE at time frames 0:00:00.400, 0:00:00.420, 0:00:00.440 and 0:00:00.480 as an example. Due to the residual WFE's high-spatial-frequency nature, we currently solely rely on the tweeter DM to compensate for the residual WFE. Figure 4.13 shows the comparison between GPI2's average PSFs with and without the injected residual WFE. When residual WFE exists in the system, the overall contrast degrades by more than 2 orders of magnitude.

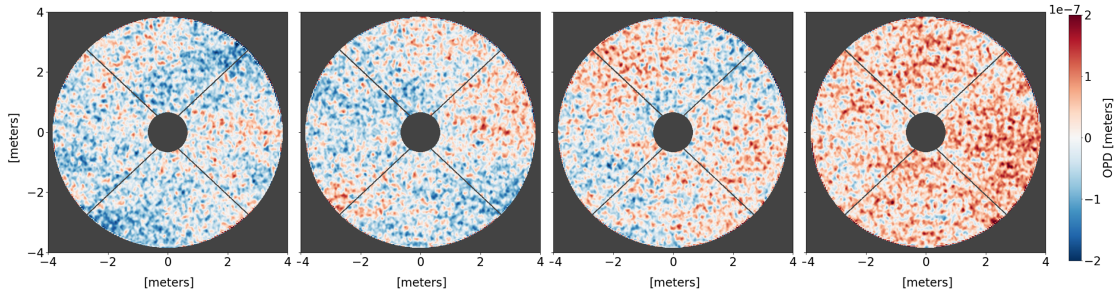

Figure 4.12: Quasi-static AO residual WFE input viewed in the pupil plane of the primary mirror. Here we show the simulated WFE at 4 different time frames H:MM:SS.MS = 0:00:00.400, 0:00:00.420, 0:00:00.440 and 0:00:00.480 as an example.

As described by Equation 4.7, there is a 1-to-2 continuous mapping between the spatial frequency (k_x, k_y) of a sinusoidal DM pattern and the speckle location $(x_{speckle}, y_{speckle})$ in the image plane. Figure 4.14 shows an example of the

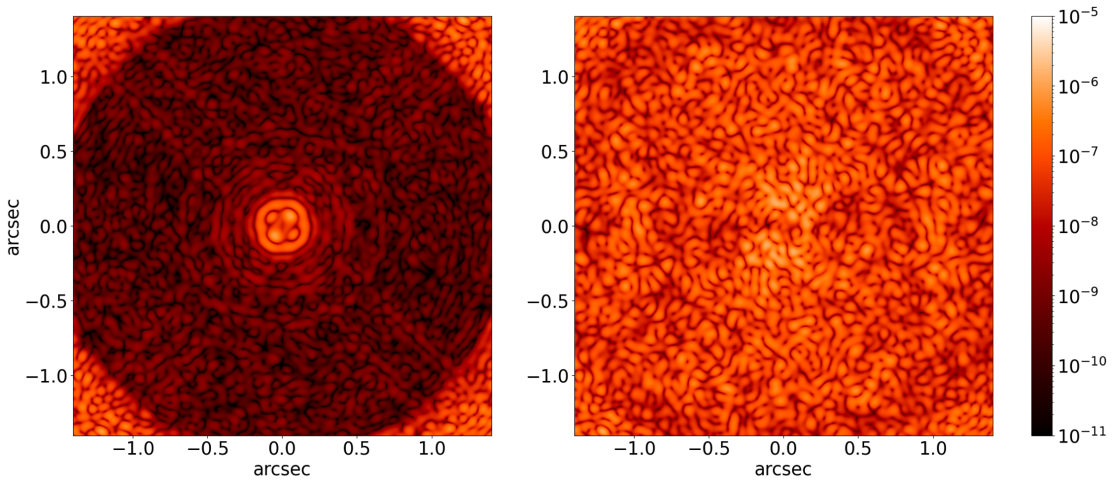

Figure 4.13: (left) The PSF with no WFE. (right) The PSF with WFE input at time frame 0:00:00.400.

two speckles induced when a single-spatial-frequency sinusoid is placed on the DM. The actual sinusoidal pattern loses some fidelity due to DM's limited bandwidth, quantization error and the face sheet's elasticity, causing the energy to leak into neighboring modes. This is shown by the fact that the speckles span multiple pixels in Figure 4.14. Therefore, for each speckle we are correcting, we mask out its surrounding region to avoid crosstalk with neighboring modes. The mask size is optimized in Section 4.6.2.

We limit our region of interest (ROI) to an annulus with angular separation $\in [0.27, 0.89]$ arcsec (DM spatial frequency $k_{x/y} \in [7, 23]$ cycles/pupil). The lower limit is to match the coronagraph's inner working angle, and the upper limit is to keep DM below its Nyquist frequency. Because of the slight asymmetry in the two speckles due to optical imperfections and numerical artifacts, we use measurements only in the upper half of the annular ROI to correct for the entire ROI.

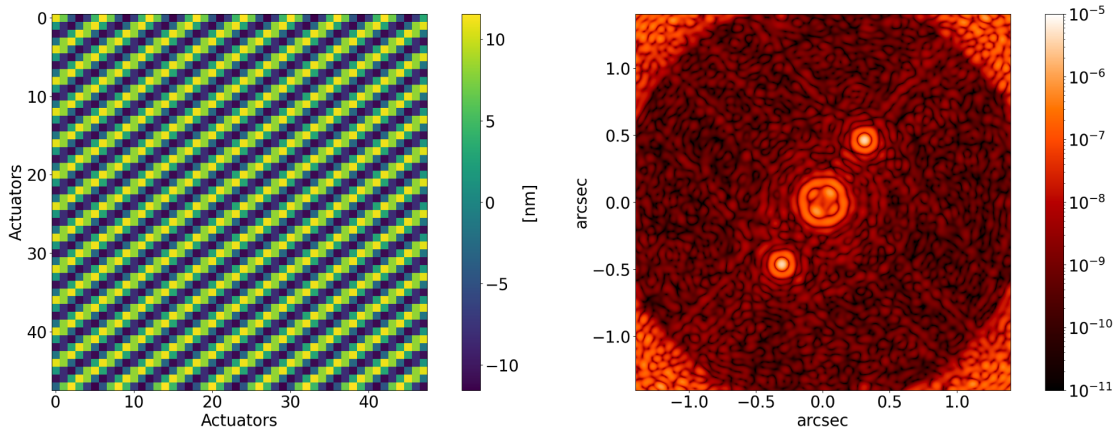

Figure 4.14: A single-spatial-frequency sinusoid carried out by the DM (left) and the pair of induced speckles (right). We use $(k_x, k_y) = (8, 12)$ cycles/pupil as an example.

4.5 Results

4.5.1 Monochromatic performance

We integrate our wavefront sensing and control scheme into GPIPSFS, and inject residual WFE from PASSATA.

We evaluate the monochromatic performance of our wavefront sensing and control algorithm at wavelength $\lambda = 1.65$ μm. Figure 4.15 shows the distribution of contrast before and after wavefront control. On average, our algorithm could improve the final raw contrast from $1.098 \times 10^{-5} \pm 5.999 \times 10^{-7}$ to $6.468 \times 10^{-7} \pm 1.743 \times 10^{-7}$, leading to an improvement ratio of 18.2 ± 5.3.

Our algorithm performance varies with WFE input. For each WFE input, our algorithm goes through several closed-loop iterations (Figure 4.2.2) before it achieves convergence in contrast improvement. Figure 4.16 shows the evolution

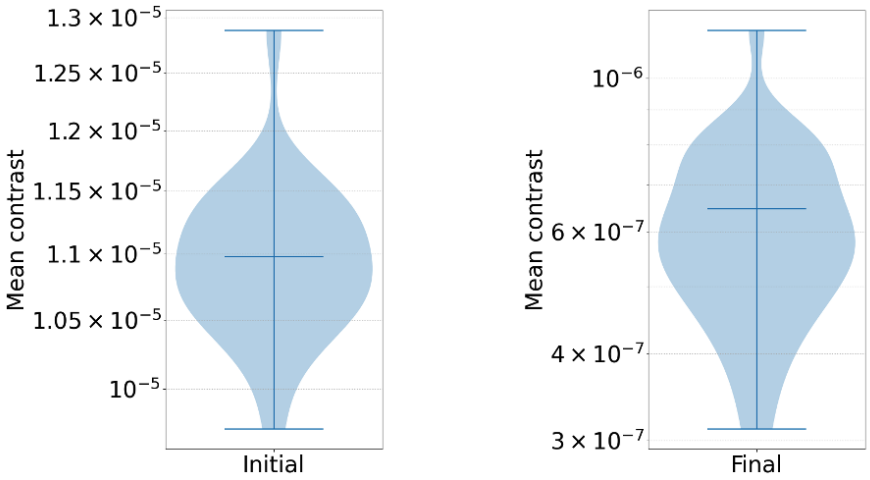

Figure 4.15: The distribution of contrast before and after wavefront control with 25 different WFE inputs. The reported contrast is the average raw contrast over the ROI. Across all time samples, our algorithm could improve the final raw contrast from $1.098 \times 10^{-5} \pm 5.999 \times 10^{-7}$ to $6.468 \times 10^{-7} \pm 1.743 \times 10^{-7}$, leading to an improvement ratio of 18.2 ± 5.3.

of contrast for the 25 different WFE inputs, with various convergence speed and final states. We end the algorithm halfway when there is no more continuous step improvement in contrast.

Figure 4.17 illustrates three representative cases of algorithm performance at three individual time samples. The top panel shows a case with a contrast improvement factor of 18.1, which is close to the average across all samples and thus serves as a representative average scenario. The middle panel highlights the best-case performance, achieving a 35-fold improvement in contrast. The bottom panel shows the worst-case scenario, with an 8-fold improvement, yet still surpassing the performance of GPI 1.0. These examples demonstrate that even in the least favorable conditions, the proposed method consistently yields substantial contrast improvement.

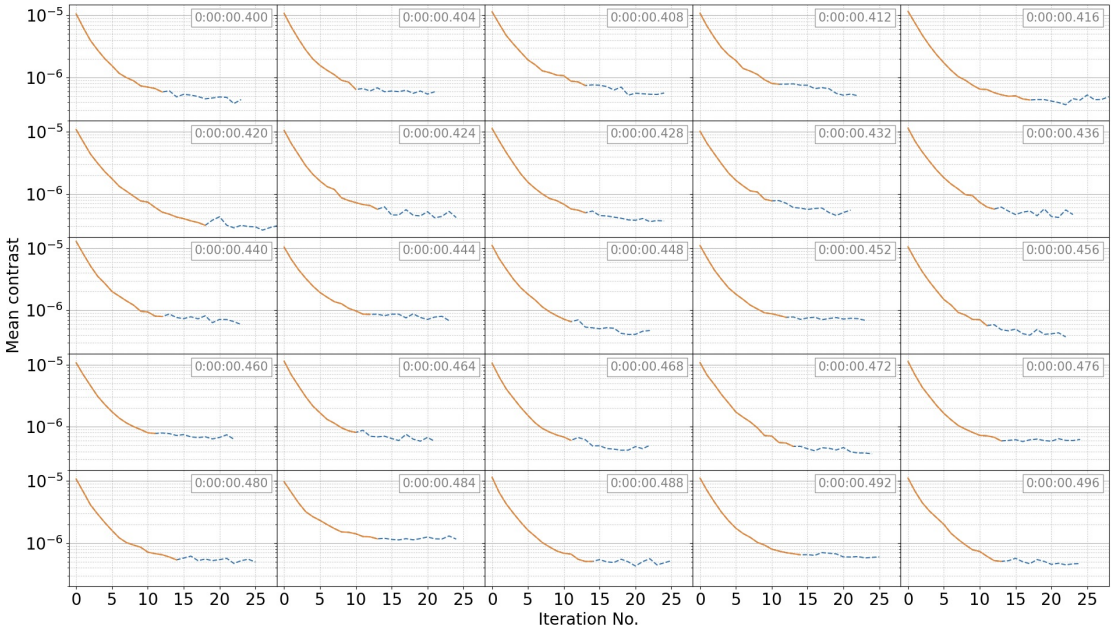

Figure 4.16: The evolution of contrast during our closed-loop, iterative wavefront sensing and control process. Plotted are 25 curves with different WFE inputs at different time frames. For the results reported in Section 4.5.1, we end the algorithm when there is no more continuous step improvement (contrast time evolution corresponding to the orange curve). If we run the algorithm for extended iterations, the contrast continues to decrease at a slower rate with damped oscillations (blue dashed curve).

4.5.2 Control loop timing and convergence

For the results reported in Section 4.5.1, the algorithm ends (Figure 4.16) when there is no more continuous step improvement in contrast. On average, the algorithm takes 13.5 closed-loop iterations to end. Each iteration requires 9 images: 1 image to determine $(\hat{k}_{x,i}, \hat{k}_{y,i}, \hat{B}_i)$ and 8 images to determine $\hat{\eta}_i$ (Section 4.2.1). Because the SCC has an integration time of 10 μsec, the total imaging latency is $13.5 \times 9 \times 10$ μsec = 1.215 msec. In comparison, the previous implementation on GPI 1.0 relied on the IFS, which required integration times of up

to 30 seconds. By adopting the SCC, we have effectively eliminated the latency overhead introduced by the imaging system. The total latency is now dominated by the AO control loop itself, significantly improving the system's speed and responsiveness for real-time wavefront correction.

Additionally, Figure 4.16 shows the contrast evolution over extended iterations after the ending criteria is met. The contrast evolutions resemble typical step responses of a closed-loop control system. The contrast continues to decrease at a slower rate after the ending criteria is met, with no overshoot but damped oscillations towards the steady state. Therefore, we can trade off latency for better contrast improvement. Figure 4.18 shows the average final contrast achieved over 25 iterations with its corresponding imaging latency. We currently choose to have an imaging latency of 1.215 msec, for an average final contrast of $6.468 \times 10^{-7} \pm 1.743 \times 10^{-7}$.

4.6 Discussion

4.6.1 Effect of optical misalignment

The simulation in Section 4.5.1 assumes the GPI 2.0 coronagraph is perfectly aligned. In reality, misalignment is mostly likely to occur on moveable components such as the apodizer and the Lyot stop, because they are mounted on rotating wheels to allow switching between different designs and bands. We estimate the misalignment to be $< 0.2\%$ of the pupil size [102].

To understand misalignment's effect on our wavefront sensing and control algorithm, we assume a conservative scenario and model the misalignment by shift-

ing the apodizer by < 0.4% off-axis in 45-deg direction and shifting the Lyot stop by the same amount in the opposite direction. The misalignment degrades the ideal contrast by a factor of 2 from 0.86×10^{-7} to 1.76×10^{-7}, as shown in Figure 4.19. We simulate with the 25 WFE inputs. Figure 4.20 shows the distribution before and after wavefront control for the two GPI 2.0 models: with and without misalignment. Figure 4.19 shows the contrast curves comparison for the average improvement case presented in Figure 4.17. When there is uncorrected WFE in the system, the effect of the misalignment is not as obvious: The two models have nearly identical distribution of initial contrast, with slight degradation in the misaligned model. The initial contrast is 1.107×10^{-5} with misalignment vs. 1.098×10^{-5} without misalignment. After wavefront control, the difference in the distribution amplifies: The misalignment limits the best achievable contrast to 4.2×10^{-7} and more tests ended with poorer contrast. Nonetheless, the final contrast are relatively close: 6.8×10^{-7} with misalignment vs. 6.5×10^{-7} without misalignment, and this corresponds to an improvement ratio of 17.6 with misalignment vs. 18.2 without misalignment.

Therefore, we conclude that our wavefront sensing and control algorithm is relatively robust against misalignment, and can partially correct for misalignment and close the gap between raw contrasts. We summarize the statistical results in Table 4.1.

4.6.2 Algorithm optimization

To achieve the results above, we optimize the algorithm by tuning a set of parameters, including speckle crosstalk mask size, number of speckles corrected each iteration, and closed-loop control integral gain.

Table 4.1: Performance statistics of misaligned vs. perfectly-aligned GPI 2.0 models

		No misalignment	Misalignment
Mean contrast ($\times 10^{-7}$)	Ideal	0.86	1.76
	Initial	109.8 ± 6.0	110.7 ± 5.7
	Final	6.5 ± 1.7	6.8 ± 2.0
Contrast improvement ratio	Avg. ± std.	18.2 ± 5.3	17.6 ± 5.0
	Max	35.0	26.9
	Min	8.3	8.6
Number of iterations		13.5 ± 1.9	13.0 ± 1.9

The speckle crosstalk mask size is crucial to the success of our algorithm: It must cover enough speckle energy while avoiding masking out too many neighboring speckles. Our wavefront control test runs show that the DM actuator stroke goes up to ±17.5 nm. According to the 1D speckle profile induced by a DM sinusoid with 17.5 nm amplitude (Figure 4.21), there are three mask size options, illustrated by the dotted gray lines: $\varnothing$ = 0.116 arcsec to mask out the central lobe, $\varnothing$ = 0.233 arcsec to mask out the first order side lobes, or $\varnothing$ = 0.310 arcsec to mask out the second order side lobes. We experiment with these three options and summarize the contrast improvement results in Figure 4.22. $\varnothing$ = 0.233 arcsec yields the best performance in both contrast improvement and convergence speed. On the contrary, $\varnothing$ = 0.116 arcsec could not suppress enough speckle energy leakage and crosstalk, causing the side lobes to be misidentified as other speckles and the algorithm to end immediately after 1 iteration. Meanwhile, $\varnothing$ = 0.310 arcsec is oversized and would mask out excessive neighboring speckles we could have corrected for in each turn, leading to algorithm inefficiency

and slow convergence.

The number of target speckles each iteration is not limited by DM's wavefront shaping ability, as our simulation results (Figure 4.23 shows. We achieve the best contrast improvement and convergence speed when we correct for as many speckles as possible in each iteration. Therefore, there's no penalty for pushing for a larger target set. The maximum number of speckles we could target after applying the crosstalk mask is approximately 70 in each iteration.

A large integral gain increases the convergence speed, but also amplifies modeling error and uncertainties. Figure 4.24 shows that as the integral gain g increases from 0.3 to 0.7, the system responds more quickly. However, beyond a certain threshold, further increase in g reduces stability and can ultimately lead to divergence. An integral gain in the range of 0.6 to 0.7 yields the best contrast improvement and a well-balanced convergence speed.

4.7 Conclusion

In this paper, we presented a classical speckle-nulling-based focal-plane wavefront control strategy designed to mitigate AO residual wavefront errors. This approach relies solely on intensity measurements provided by the SCC, serving as a robust fallback option for the primary high-order wavefront sensing and control system on GPI2. We developed an end-to-end simulation and evaluated the system performance at a wavelength of 1.65 microns. Our results demonstrated an average contrast improvement by a factor of 18.2 ± 5.3 at 823 Hz, with an imaging latency of 1.215 msec. We achieved an average final raw contrast of

6.468×10^{-7} within the angular separation range of $[0.27, 0.89]$ arcsec.

This work represents a significant advancement compared to prior studies on GPI1 ([101]) and SPIDERS ([61]). While GPI1's focal-plane wavefront control strategy theoretically achieved a contrast improvement ratio of 5.3 in simulations, it was not successful deployed on-sky due to IFS's long integration time (30 seconds to achieve adequate SNR) and CAL's dual-arm non-common path vibrations (Figure 4.1). Similarly, our earlier work on SPIDERS demonstrated only a 3.2-fold contrast improvement on injected bright speckles, limited by sub-optimal crosstalk mask dimensions and discrete modal control.

Despite the demonstrated performance, several sources of error continue to limit the system's overall capability. These include both modeling inaccuracies and uncorrected physical effects.

(1) The optical modeling used in GPIPSFS is based on Fraunhofer diffraction, which assumes far-field propagation. While this approximation is generally reasonable, it fails to capture near-field diffraction effects that can become significant in a complex optical system such as GPI.

(2) There are calibration errors in the DM model. Specifically, each actuator's influence function is approximated as Gaussian, which does not fully represent the physical response of the DM's facesheet.

(3) In the speckle phase probing process, limited spatial sampling and interpolation introduces additional uncertainty.

(4) The current wavefront control framework does not account for amplitude aberrations, which imposes a fundamental limit on the achievable contrast even with perfect phase correction.

To address these limitations and improve system performance, we have identified several key directions for future work:

(1) Hardware Validation: We plan to test and validate the optimized wavefront control algorithm directly on the GPI 2 hardware to assess its real-world performance and robustness.

(2) Improved Optical Modeling: We intend to incorporate Fresnel diffraction into the optical simulations to more accurately represent near-field propagation effects. In addition, we will simulate broadband performance to better predict contrast limitations and chromatic behavior.

(3) Latency Analysis: We aim to simulate the impact of system latency on closed-loop performance under dynamic aberrations. These studies will help define control requirements, such as bandwidth.

(4) Alternative Control Strategies: We are also exploring alternative wavefront control techniques that offer reduced latency, such as implicit Electric Field Conjugation (iEFC).

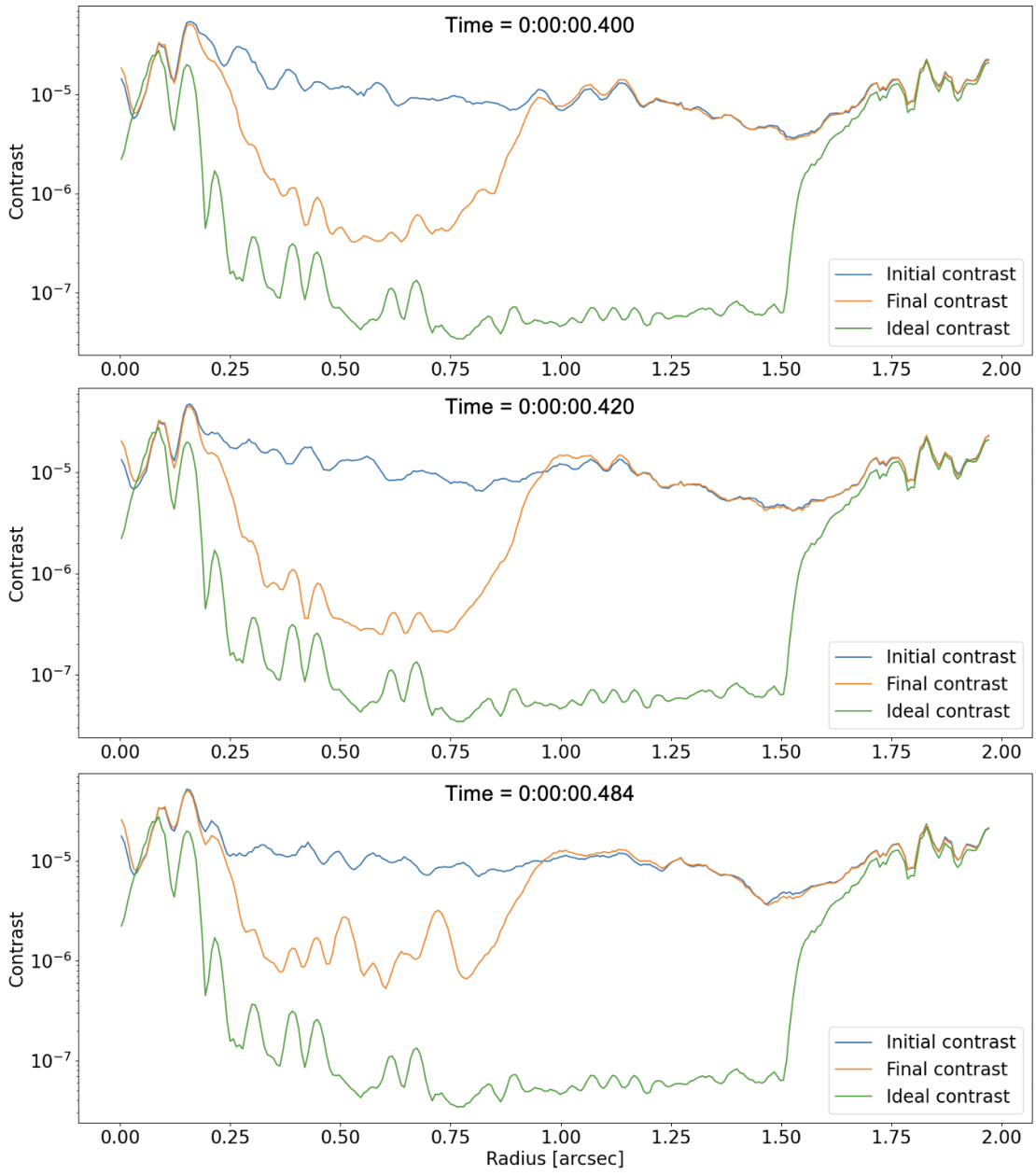

Figure 4.17: Average (top), maximal (middle) and minimal (bottom) contrast improvement achieved with the WFE inputs. The three cases correspond to contrast improvement ratios of 18.1, 35.0 and 8.3 respectively. Comparisons of initial contrast with injected WFE, final contrast after wavefront control, and raw contrast as a reference, are shown for each case.

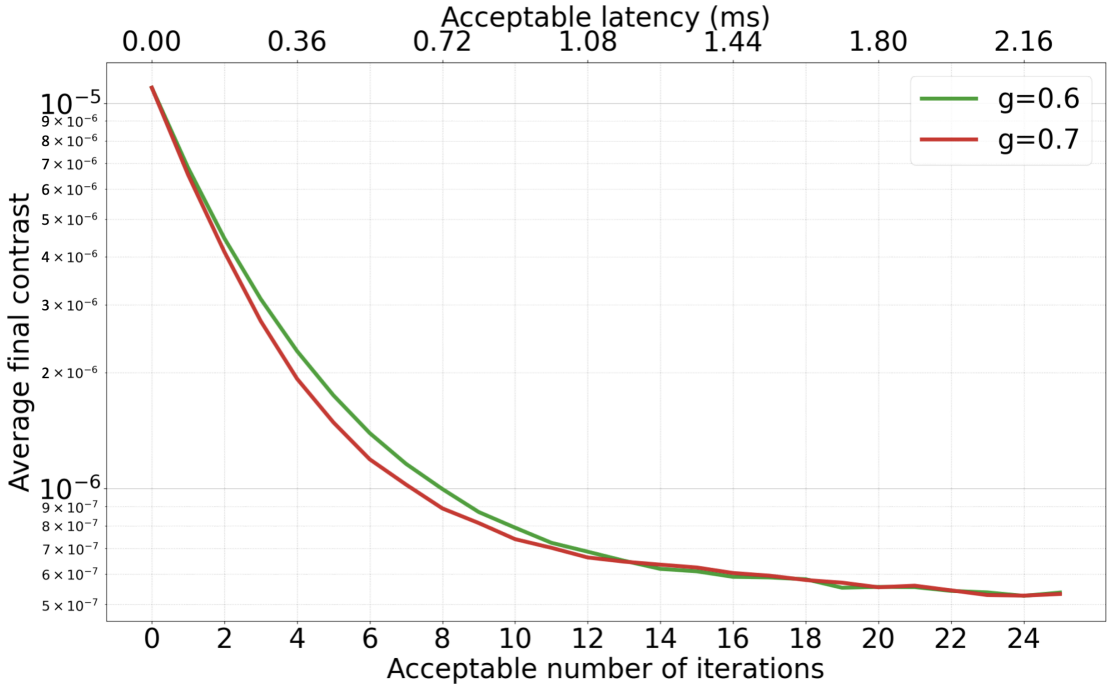

Figure 4.18: Average contrast evolution over 25 iterations, for integral gains of 0.6 and 0.7. The raw contrast is averaged across all time samples.

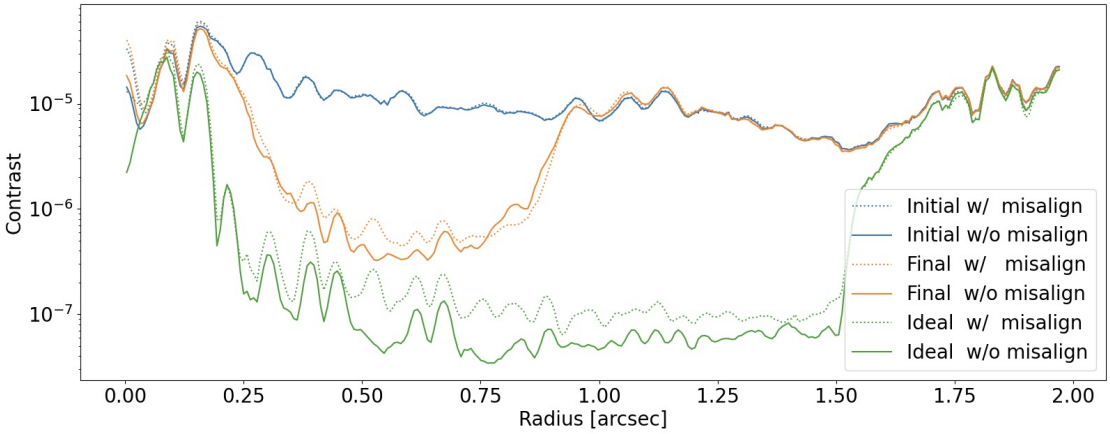

Figure 4.19: Contrast curves comparison between misaligned vs. perfectly-aligned GPI 2.0 models, using the WFE input from the average contrast improvement case presented in Figure 4.17. Comparisons of initial contrast with injected WFE, final contrast after wavefront control, and raw contrast are shown.

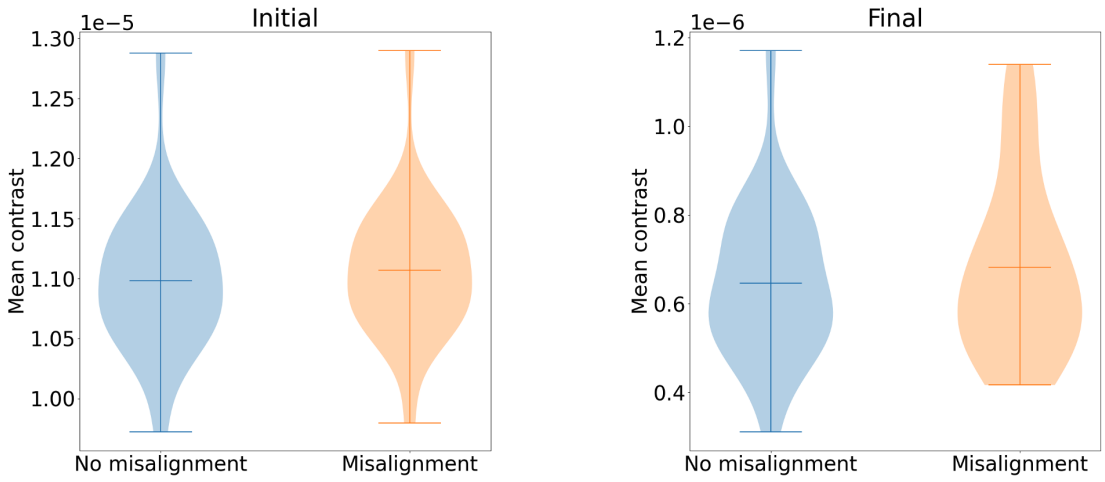

Figure 4.20: Comparison between misaligned vs. perfectly-aligned GPI 2.0 models: Distribution of contrast with 25 different WFE inputs before and after wavefront control.

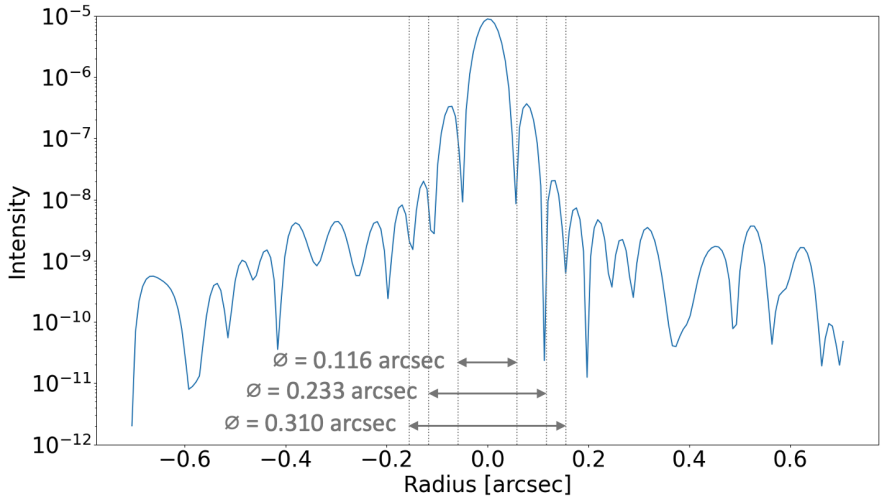

Figure 4.21: 1D profile of a speckle induced by shaping the DM into a sinusoid with 17.5 nm (max DM stroke observed) amplitude. The dotted gray lines illustrate 3 optional sizes of the speckle crosstalk mask, and the corresponding diameters are annotated.

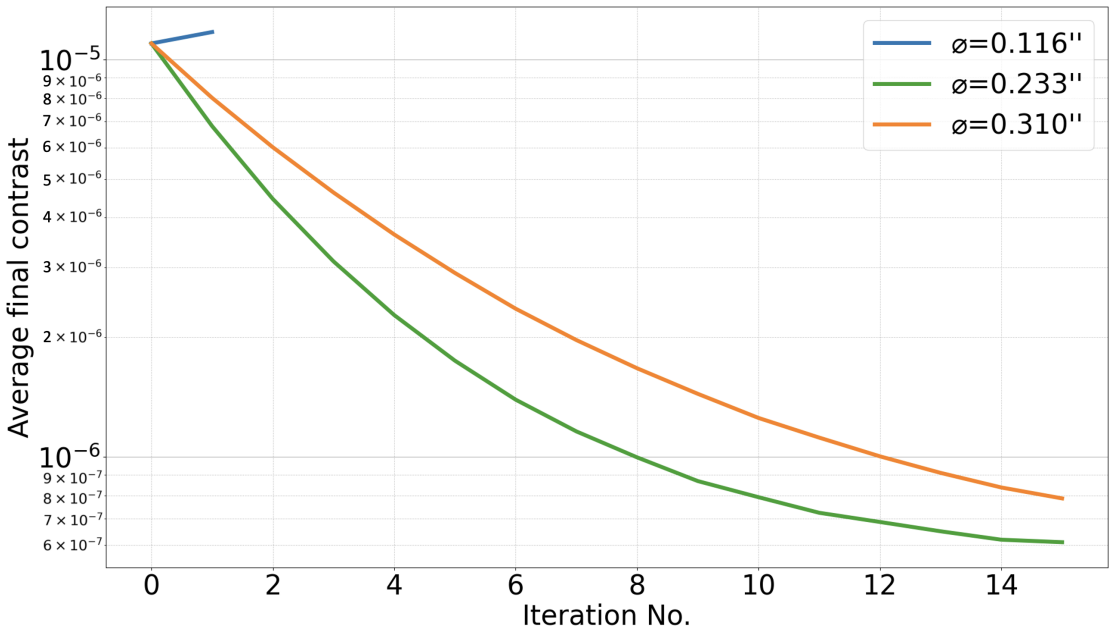

Figure 4.22: Average contrast evolution given different crosstalk mask diameters. The raw contrast is averaged across all time samples. Controlled variables include: number of target speckles per iteration = max, integral gain = 0.6.

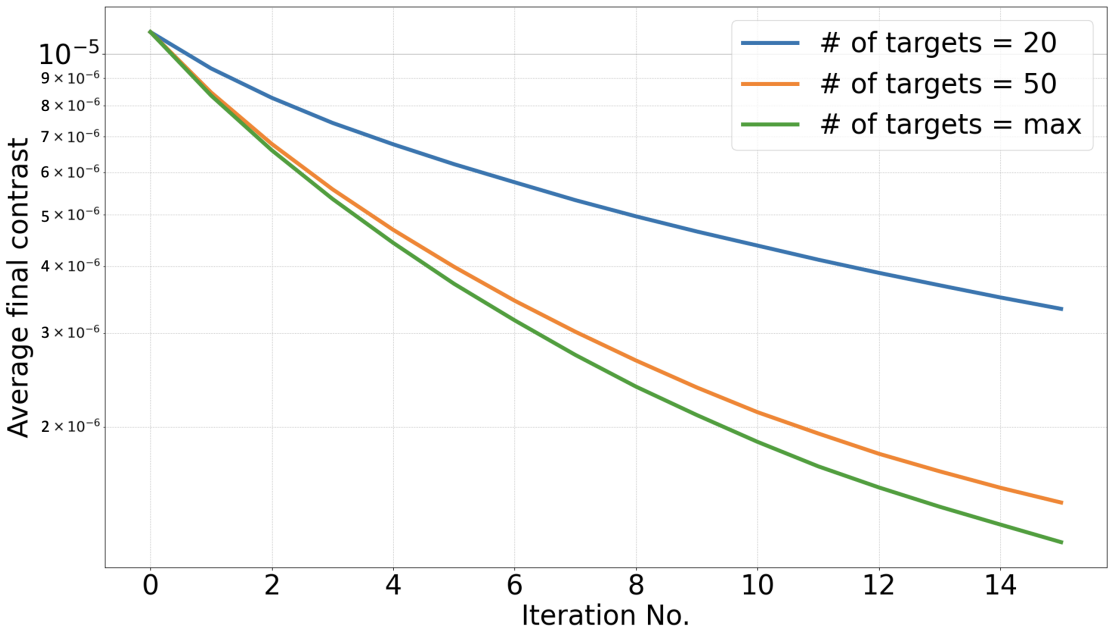

Figure 4.23: Average contrast evolution given different numbers of target speckles per iteration. Max means as many speckles as possible. Controlled variables include: crosstalk mask diameter = 0.233 arcsec, integral gain = 0.3.

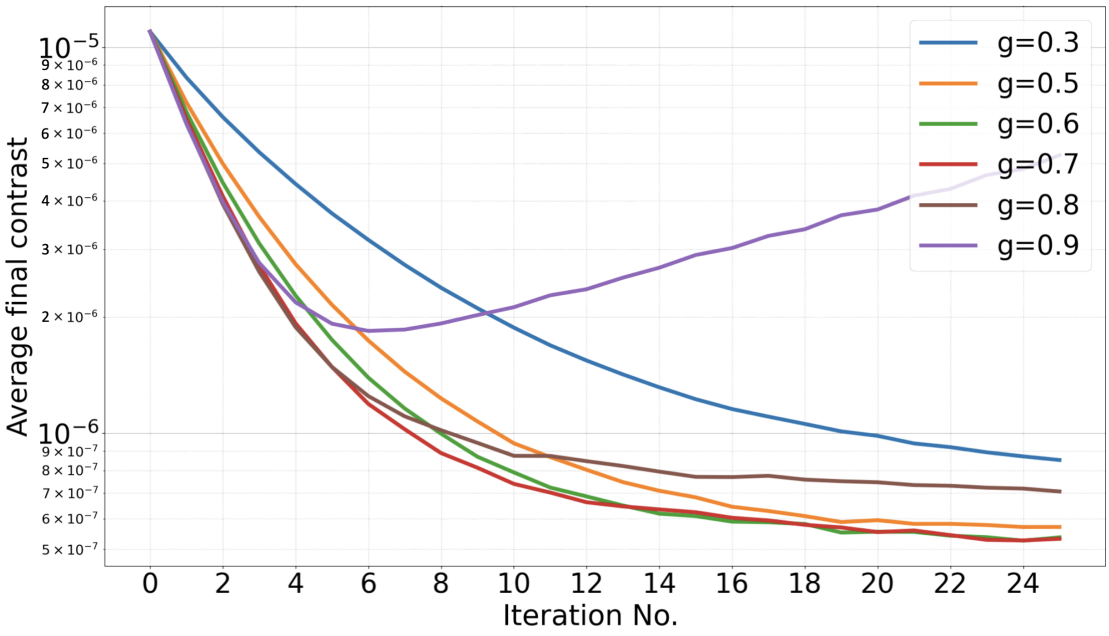

Figure 4.24: Average contrast evolution given different integral gains. Controlled variables include: crosstalk mask diameter = 0.233 arcsec, number of target speckles per iteration = max.

CHAPTER 5

CONCLUSION

In this thesis, we investigated the principles, methodologies, and applications of wavefront control in automated optical alignment and in high-contrast imaging with the Gemini Planet Imager, with a primary focus on focal-plane wavefront sensing and control.

In Chapter 2, we explored automated alignment techniques aimed at enhancing optical system precision and flexibility, while reducing the time and labor required for system setup and maintenance. We presented an automated alignment technique on a double off-axis parabolic mirror system, which poses challenging alignment problems due to the mirrors' high sensitivity to aberrations, rotational asymmetry and non-orthogonality in stage adjustments. Moreover, we used focal plane wavefront sensing to eliminate non-common path error, increase optical throughput and reduce cost and system complexity involving dedicated wavefront sensors. We incorporated model-based optimal estimation and control to better handle the nonlinearity, model uncertainty and noise. Using either an iterated extended Kalman filter or a square-root unscented Kalman filter as the optimal nonlinear misalignment state estimator, we were able to consistently reduce the linear misalignment from around 1 mm to < 5 μm and the angular misalignment from around 500 to < 6 arcsec in simulation, achieving a final wavefront error of $< 5 * 10^{-5}$ waves within field of view when tested at wavelength 635 nm. We discovered a multi-state coupling effect, which implies that different misalignment states have compensating effects on system measurements, thus interfering with the estimator's observation of misalign-

ment state changes. We further investigated the coupling's effects on alignment quality through observability analysis.

In Chapter 3 and 4, we presented our contributions to the Gemini Planet Imager's upgrade to advance its high-contrast imaging capabilities, enabling more effective detection and characterization of exoplanets. In Chapter 3, we facilitated the establishment of a control system design pipeline to address the challenges of real-time wavefront correction under dynamic aberrations. Our key metrics for analyzing and fine-tuning the GPI 2.0 AO response included the error transfer function and stability margins. In Chapter 4, we presented a focal-plane wavefront control strategy based on classical speckle nulling to address a long-standing limitation of the Gemini Planet Imager: performance degradation caused by speckle noise originating from post-coronagraph aberrations and internal optical imperfections. Through a combination of theoretical modeling, algorithm development, and preliminary experimental validation, we demonstrated contrast improvement by a factor of 18.2 ± 5.3 at 823 Hz on average, resulting in a final average contrast of 6.468×10^{-7} within an angular separation range of $[0.27, 0.89]$ arcsec, in simulation. These results mark a substantial improvement over GPI's previous performance and will be integrated into its second-generation upgrade. This advancement enables deeper contrast at smaller inner working angles, paving the way for the detection of lower-mass exoplanets.

Overall, this thesis integrates control theory and optical engineering and underscores the critical role of wavefront sensing and control in improving the fidelity and functionality of optical systems. While this study focused on specific con-

trol algorithms and system architectures, future research could explore hybrid control approaches, machine-learning-based optimization, and broader applications across optical instrumentation. Further hardware integration and on-sky testing would be valuable in validating the performance gains in operational environments.

APPENDIX A

A.1 Control Stability Margins

In the context of Section 3.3, the closed-loop system becomes unstable when the denominator of the ETF equates to zero, i.e.

$$1 + L(s) = 0 \tag{A.1}$$

This happens when the loop gain $|L(s)| = 1$ and the phase $\angle L(s) = 180°$. The gain margin is the amount $|L(s)|$ can be changed, when $\angle L(s) = 180°$, without reducing $1 + L(s)$ to zero. The phase margin is the amount of phase lag that can be added, when $|L(s)| = 1$, with the same purpose. Figure A.1 illustrates the margins on a typical Bode plot. The margins describe how much disturbances, changes or uncertainties the system can endure before getting unstable.

Figure A.2 shows a typical closed-loop-system step response with control metrics including rise time, overshoot, settling time and steady-state value. A large gain or phase margin often leads to small overshoot and little ringing (few resonant peaks, shown in Figure A.2). Figure A.3 shows how overshoot depends on phase margin and Figure A.4 shows how rise time and settling time depend on bandwidth for a phase margin of 60°.

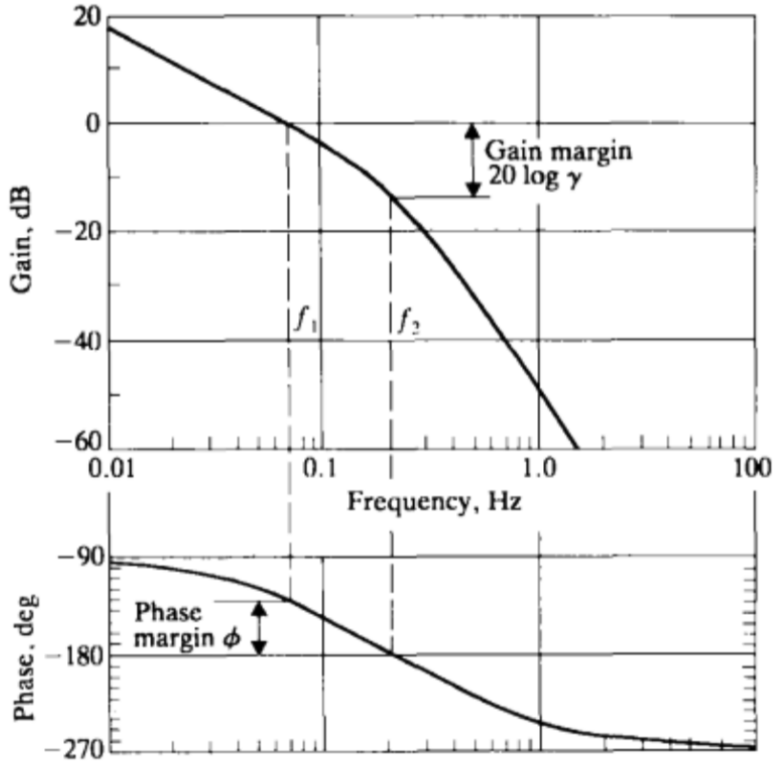

Figure A.1: Gain and phase margins defined on a Bode plot. [26]

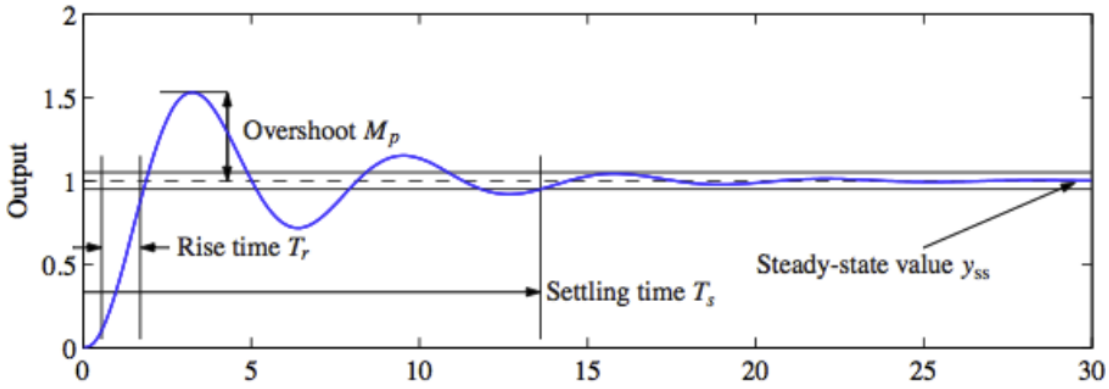

Figure A.2: Typical step response demonstrating control system performance metrics, including rise time, overshoot, settling time and steady-state value. Source: D. MacMartin, Cornell University, MAE5780 Lecture 8-1 course material.

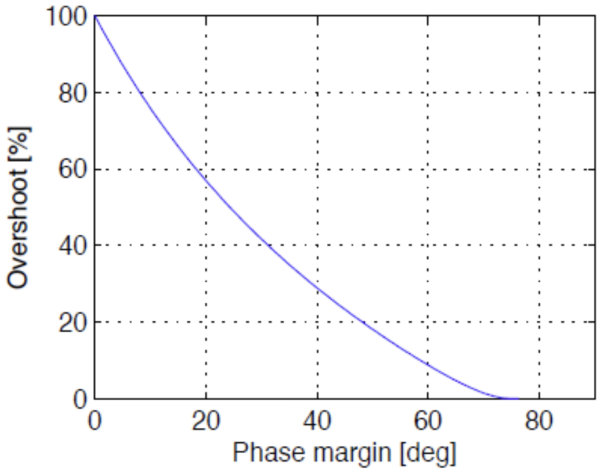

Figure A.3: Step response overshoot as a function of phase margin. Source: D. MacMartin, Cornell University, MAE5780 Lecture 8-1 course material.

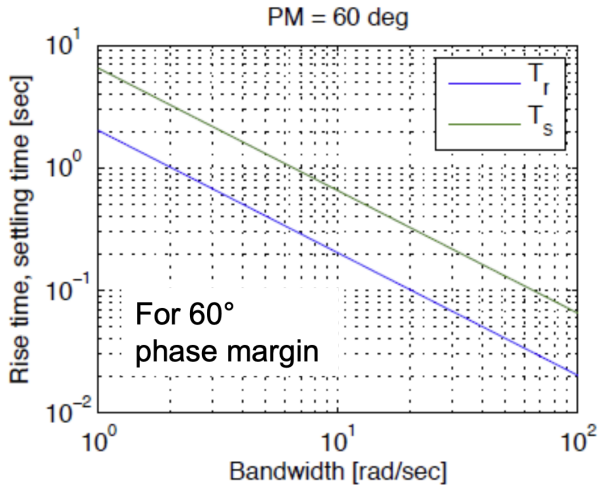

Figure A.4: Step response rise time and settling time as a function of bandwidth. Source: D. MacMartin, Cornell University, MAE5780 Lecture 8-1 course material.

APPENDIX B

B.1 PASSATA Simulation Parameters

Here is the parameters we used in PASSATA to simulate GPI's AO residual wavefront errors.

```
{
"MAIN": {
    "INSTRUMENT_NAME": "gpi",
    "ZENITHANNGLEINDEG": 30,     ; [deg] Zenith angle
    "TOTAL_TIME": 10,            ; [sec] Total simulation time
    "PIXEL_PUPIL": 176,          ; Linear dimiension of pupil phase
                                       array
    "PRECISION": 0,
    "PIXEL_PITCH": 0.045000002, ; [m] Pitch of the pupil phase array
    "TIME_STEP": 0.00100000      ; [sec] Simulation time step
    },
"WFS_SOURCE": {
    "HEIGHT": Inf,
    "MAGNITUDE": 8,              ; Source magnitude
    "WAVELENGTHINNM": 750,
    "POLAR_COORDINATE": [0.0000000, 30.000000]
                                 ; [arcec, deg] Source polar
                                   coodinates
    },
"PUPIL_STOP": {
    "MASK_DIAM": 1.0000000,
    "OBS_DIAM": 0.16200000       ; Pupil stop mask obstruction size
    },
"PYRAMID": {
```

```
    "PUP_DIAM": 60.000000,      ; Pupil diameter in subapertures
    "MOD_AMP": 3.0000000,       ; [lambda/D] Modulation radius
    "FOV": 2.0999999,           ; [arcsec] Requested field of view
    "OUTPUT_RESOLUTION": 140,   ; [px] Output sampling
    "FOV_ERRINF": 0.10000000,   ; Maximum error in reducing FoV
    "FFT_RES": 3.0000000,       ; [lambda/D] Pyramid focal-plane PSF
                                      sampling
    "WAVELENGTHINNM": 750,      ; [nm] Pyramid wavelength
    "PUP_DIST": 72.000000       ; Requested separation between pupil
                                      centers in subapertures
    },
"MODULATE_DIST": {
    "INFLUENCE_FUNCTION": "GPI_176pix_1603KLs3zern",
    "MODENUMBER": 4,
    "CARRIERAMPLITUDEINNM": 5.0000000,
    "CARRIERFREQUENCY": 80.000000,
    "HEIGHT": 0,
    "AMP": [0.0000000, 0.0000000, 0.0000000, 0.0000000, 5.0000000],
    "CONSTANT": [0.0000000, 0.0000000, 0.0000000, 0.0000000,
                 0.0000000],
    "FREQ": [0.0000000, 0.0000000, 0.0000000, 0.0000000, 80.000000],
    "FUNC_TYPE": "SIN",
    "NMODES": 5,
    "DT": 1.0000000
    },
"DETECTOR": {
    "SIZE": [140, 140],         ; [px] Detector size
    "PHOTON_NOISE": 1,          ; [binary] Enable photon noise
    "BANDW": 300,               ; Sensor bandwidth
    "SKY_BG_NORM": 0.0000000,   ; Sky background level
    "QUANTUM_EFF": 0.31999999,  ; Quantum efficiency * total
                                      transmission
```

```
    "BINNING": 1,                ; [binary]
    "DARKCURRENT_NOISE": 1,      ; [binary] Enable dark current
    "DT": 0.00100000,            ; Detector integration time
    "NAME": "OCAM2kSWspec",
    "BACKGROUND_NOISE": 0,       ; [binary] Disable background noise
    "READOUT_NOISE": 1,          ; [binary] Enable readout noise
    "READOUT_LEVEL": 0.40000000596046448,
                                 ; [e-/pix/frame] Readout noise
    "DARKCURRENT_LEVEL": 0.0013799999952316285,
    "BACKGROUND_LEVEL": 0.0000000,
    "EXCESS_NOISE": 1
    },
"CAMERA": {
    "WAVELENGTHINNM": 1650,      ; [nm] Imaging wavelength
    "ND": 4                      ; Padding coefficient for PSF
                                        computation
    },
"OPTICALGAIN_LOOP": {
    "OPTGAIN": 0.89999998,
    "GAIN": 0.30000001
    },
"SLOPEC": {
    "THR1": 0.30000001,
    "PUPDATA_TAG": "gpi_ps176p0.045_pyr60x60_wl750_fv2.1_ft3.
                                        0_bn1_th0.30a0.30b",
                                 ; Pyramid WFS pupils
    "THR2": 0.30000001,
    "SN_TAG": "gpi_ps176p0.045_pyr60x60_wl750_fv2.1_ft3.0_ma3_bn1_th0
                                        .30a0.30b"
                                 ; Slope reference vector
    },
"MODALANALYSIS": {
```

```
    "PHASE2MODES_TAG": "GPI_176pix_1603KLs3zern_inv"
                                ; Inverse of influence function
    },
"MODALREC": {
    "RECMAT_TAG": "gpi_ps176p0.045_pyr60x60_wl750_fv2.1_ft3.0_ma3_bn1
                  _th0.30a0.30b_mn1603",
                                ; Reconstruction matrix
    "NMODES": 1603,
    },
"NCPA": {
    "HEIGHT": 0,                ; Conjugation altitude of the
                                      aberration
    "INFLUENCE_FUNCTION": "GPI_176pix_1603KLs3zern",
    "FUNC_TYPE": "SIN",
    "NMODES": 100,
    "CONSTANT": [0.0000000 (repeat x99)]
    },
"DM": {
    "HEIGHT": 0,                ; [m] DM height
    "IFUNC_TAG": "GPI_176pix_1603KLs3zern"
                                ; Influence function
    },
"CONTROL": {
    "FF": [0.99991608(x2), 0.99995077(x3), 0.99997342(x4),
           0.99998719(x5), 0.99999475(x6), 0.99999833(x7),
           0.99999964(x8), 1.0000000(x30), 0.99999964(x12),
           ..., 0.89999998(x8)],
                                ; A forgetting factor vector of the
                                  same size as INT_GAIN to get a
                                  leaky integrator
    "DELAY": 2,                 ; [frames] Total temporal delay
    "INT_GAIN": [0.44999999(x500), 0.38249999(x500),
```

```
                0.31500000(x300), 0.17999999(x150),
                0.089999996(x153)],
                                ; Integrator gain vector (for 'INT'
                                  control)
    "TYPE": "INT",              ; Type of control (Integral control)
    "OPT_DT": 1.0000000,        ; [sec] Parameter for recursive
                                       optmization of modal gain
                                       vector
    "LIMIT_INC": 1
    },
"SEEING": {
    "CONSTANT": 0.54000002,     ; [arcsec] seeing
    "FUNC_TYPE": "SIN"
    },
"ATMO": {
    "L0": 30,                   ; [m] Outer scale
    "HEIGHTS": [0, 500, 1000, 2000, 4000, 8000, 16000],
                                ; Layer heights at 0 zenith angle
    "CN2": [0.45570001, 0.12950000, 0.044199999, 0.050600000,
            0.11670000, 0.092600003, 0.11070000],
                                ; Cn2 weights (sum up to 1)
    "MCAO_FOV": 120,            ; Multi-conjugate AO system FoV
    "WAVELENGTHINNM": 500,
    "SEED": 1
    }
"WIND_DIRECTION": {
    "CONSTANT": [0.0000000, 0.0000000, 0.0000000, 0.0000000,
                0.0000000, 0.0000000, 0.0000000],
                                ; Wind direction value
    "FUNC_TYPE": "SIN"
    },
"WIND_SPEED": {
```

```
    "CONSTANT": [5.5999999, 5.7700000, 6.2500000, 7.5700002,
                 13.310000, 19.059999, 12.140000],
                                    ; Wind speed value
    "FUNC_TYPE": "SIN"
    },
}
```

BIBLIOGRAPHY

[1] Guido Agapito, Enrico Pinna, Alfio Puglisi, and Fabio Rossi. Elephants, goldfishes and soul: a dissertation on forgetfulness and control systems. *arXiv preprint arXiv:1911.05989*, 2019.

[2] Guido Agapito, A Puglisi, and Simone Esposito. Passata: object oriented numerical simulation software for adaptive optics. In *Adaptive Optics Systems V*, volume 9909, pages 2164–2172. SPIE, 2016.

[3] Guillem Anglada-Escudé, Pedro J Amado, John Barnes, Zaira M Berdiñas, R Paul Butler, Gavin AL Coleman, Ignacio de La Cueva, Stefan Dreizler, Michael Endl, Benjamin Giesers, et al. A terrestrial planet candidate in a temperate orbit around proxima centauri. *nature*, 536(7617):437–440, 2016.

[4] Vanessa P Bailey, Eduardo Bendek, Brian Monacelli, Caleb Baker, Gasia Bedrosian, Eric Cady, Ewan S Douglas, Tyler Groff, Sergi R Hildebrandt, N Jeremy Kasdin, et al. Nancy grace roman space telescope coronagraph instrument overview and status. *Techniques and Instrumentation for Detection of Exoplanets XI*, 12680:283–293, 2023.

[5] J-P Beaulieu, David P Bennett, Pascal Fouqué, Andrew Williams, Martin Dominik, UG Jørgensen, Daniel Kubas, Arnaud Cassan, Christian Coutures, John Greenhill, et al. Discovery of a cool planet of 5.5 earth masses through gravitational microlensing. *Nature*, 439(7075):437–440, 2006.

[6] Ruslan Belikov, Amir Give'on, John T Trauger, Michael Carr, N Jeremy Kasdin, Robert J Vanderbei, Fang Shi, Kunjithapatham Balasubramanian, and Andreas Kuhnert. Toward 10 10 contrast for terrestrial exoplanet detection: demonstration of wavefront correction in a shaped-pupil coronagraph. In *Space Telescopes and Instrumentation I: Optical, Infrared, and Millimeter*, volume 6265, pages 383–392. SPIE, 2006.

[7] J-L Beuzit, Arthur Vigan, David Mouillet, Kjetil Dohlen, Raffaele Gratton, Anthony Boccaletti, J-F Sauvage, Hans Martin Schmid, Maud Langlois, Cyril Petit, et al. Sphere: the exoplanet imager for the very large telescope. *Astronomy & Astrophysics*, 631:A155, 2019.

[8] Ian A Bond, A Udalski, M Jaroszyński, NJ Rattenbury, B Paczyński, I Soszyński, L Wyrzykowski, MK Szymański, M Kubiak, O Szewczyk, et al. Ogle 2003-blg-235/moa 2003-blg-53: a planetary microlensing event. *The Astrophysical Journal*, 606(2):L155, 2004.

[9] Martin J Booth. Adaptive optics in microscopy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1861):2829–2843, 2007.

[10] Pascal J Bordé and Wesley A Traub. High-contrast imaging from space: speckle nulling in a low-aberration regime. *The Astrophysical Journal*, 638(1):488, 2006.

[11] William J Borucki, David Koch, Gibor Basri, Natalie Batalha, Timothy Brown, Douglas Caldwell, John Caldwell, Jørgen Christensen-Dalsgaard, William D Cochran, Edna DeVore, et al. Kepler planet-detection mission: introduction and first results. *Science*, 327(5968):977–980, 2010.

[12] Rory Bowens, Michael R Meyer, Christian Delacroix, Olivier Absil, Roy van Boekel, Sascha Patrick Quanz, M Shinde, Matthew Kenworthy, Brunella Carlomagno, G Orban De Xivry, et al. Exoplanets with elt-metis-i. estimating the direct imaging exoplanet yield around stars within 6.5 parsecs. *Astronomy & Astrophysics*, 653:A8, 2021.

[13] R Paul Butler, Geoffrey W Marcy, Eric Williams, Chris McCarthy, Preet Dosanjh, and Steven S Vogt. Attaining doppler precision of 3 m s-1. *Publications of the Astronomical Society of the Pacific*, 108(724):500, 1996.

[14] Aarynn L Carter, Sasha Hinkley, Jens Kammerer, Andrew Skemer, Beth A Biller, Jarron M Leisenring, Maxwell A Millar-Blanchaer, Simon Petrus, Jordan M Stone, Kimberly Ward-Duong, et al. The jwst early release science program for direct observations of exoplanetary systems i: High-contrast imaging of the exoplanet hip 65426 b from 2 to 16 μm. *The Astrophysical journal letters*, 951(1):L20, 2023.

[15] P. Chakravarthula, D. Dunn, K. Akşit, and H. Fuchs. Focusar: Auto-focus augmented reality eyeglasses for both real world and virtual imagery. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2906–2916, 2018.

[16] David Charbonneau, Timothy M Brown, David W Latham, and Michel Mayor. Detection of planetary transits across a sun-likestar. *The Astrophysical Journal*, 529(1):L45, 1999.

[17] Z. Chen, D. Zhou, H. Liao, and X. Zhang. Precision alignment of optical fibers based on telecentric stereo microvision. *IEEE/ASME Transactions on Mechatronics*, 21(4):1924–1934, 2016.

[18] Jeffrey Chilcote, Quinn Konopacky, Robert J De Rosa, Randall Hamper, Bruce Macintosh, Christian Marois, Marshall D Perrin, Dmitry Savransky, Rémi Soummer, Jean-Pierre Véran, et al. Gpi 2.0: upgrading the gemini planet imager. In *Ground-based and Airborne Instrumentation for Astronomy VIII*, volume 11447, pages 394–407. SPIE, 2020.

[19] Jeffrey Chilcote, Quinn Konopacky, Joeleff Fitzsimmons, Randall Hamper, Bruce Macintosh, Christian Marois, Dmitry Savransky, Remi Soummer, Jean-Pierre Véran, Guido Agapito, et al. Gpi 2.0: upgrade status of the gemini planet imager. In *Ground-based and Airborne Instrumentation for Astronomy IX*, volume 12184, pages 635–649. SPIE, 2022.

[20] Jeffrey Chilcote, Quinn Konopacky, Randall Hamper, Bruce Macintosh, Christian Marois, Dmitry Savransky, Rémi Soummer, Jean-Pierre Véran, Guido Agapito, Arlene Aleman, et al. Gpi 2.0: upgrade status of the gemini planet imager. In *Ground-based and Airborne Instrumentation for Astronomy X*, volume 13096, pages 2722–2735. SPIE, 2024.

[21] Simone Esposito and Armando Riccardi. Pyramid wavefront sensor behavior in partial correction adaptive optic systems. *Astronomy & Astrophysics*, 369(2):L9–L12, 2001.

[22] J. Fang and D. Savransky. Automated alignment of a reconfigurable optical system using focal-plane sensing and kalman filtering. *Applied Optics*, 55(22), 2016.

[23] J W Figoski, T E Shrode, and G F. Moore. Computer-Aided Alignment Of A Wide-Field, Three-Mirror, Unobscured, High-Resolution Sensor. In Robert E. Fischer and Richard C. Juergens, editors, *Recent Trends in Optical Systems Design and Computer Lens Design Workshop II*, volume 1049, pages 166 – 177. International Society for Optics and Photonics, SPIE, 1989.

[24] Debra A Fischer, Guillem Anglada-Escude, Pamela Arriagada, Roman V Baluev, Jacob L Bean, Francois Bouchy, Lars A Buchhave, Thorsten Carroll, Abhijit Chakraborty, Justin R Crepp, et al. State of the field: extreme precision radial velocities. *Publications of the Astronomical Society of the Pacific*, 128(964):066001, 2016.

[25] Gene F. Franklin, J. David Powell, and Emami-Naeini Abbas. *Feedback Control of Dynamic Systems*. Pearson, 2014.

[26] Bernard Friedland. *Control Systems Design: An Introduction to State-Space Methods*. McGraw-Hill Higher Education, 1985.

[27] Eric Gendron and Pierre Léna. Astronomical adaptive optics. 1: Modal control optimization. *Astronomy and Astrophysics (ISSN 0004-6361), vol. 291, no. 1, p. 337-347*, 291:337–347, 1994.

[28] Ralph W Gerchberg. A practical algorithm for the determination of plane from image and diffraction pictures. *Optik*, 35(2):237–246, 1972.

[29] Amir Give'on, B Kern, Stuart Shaklan, Dwight C Moody, and Laurent Pueyo. Electric field conjugation-a broadband wavefront correction algorithm for high-contrast imaging systems. In *American Astronomical Society Meeting Abstracts*, volume 211, pages 135–20, 2007.

[30] Amir Give'on, Brian D Kern, and Stuart Shaklan. Pair-wise, deformable mirror, image plane-based diversity electric field estimation for high contrast coronagraphy. In *Techniques and Instrumentation for Detection of Exoplanets V*, volume 8151, pages 376–385. SPIE, 2011.

[31] Amir Give'on, Ruslan Belikov, Stuart Shaklan, and Jeremy Kasdin. Closed loop, dm diversity-based, wavefront correction algorithm for high contrast imaging systems. *Optics Express*, 15(19):12338–12343, 2007.

[32] Robert A Gonsalves. Phase retrieval and diversity in adaptive optics. *Optical Engineering*, 21(5):829–832, 1982.

[33] Tyler D Groff and N Jeremy Kasdin. Kalman filtering techniques for focal plane¡? a3b2 show [pmg: line-break justify=" yes"/]?¿ electric field estimation. *Journal of the Optical Society of America A*, 30(1):128–139, 2012.

[34] Zhiyuan Gu, Changxiang Yan, and Yang Wang. Alignment of a three-mirror anastigmatic telescope using nodal aberration theory. *Optics express*, 23(19):25182–25201, 2015.

[35] Olivier Guyon. Phase-induced amplitude apodization of telescope pupils for extrasolar terrestrial planet imaging. *Astronomy & Astrophysics*, 404(1):379–387, 2003.

[36] Olivier Guyon, Frantz Martinache, Vincent Garrel, Frederic Vogt, Kaito Yokochi, and Takashi Yoshikawa. The subaru coronagraphic extreme ao (scexao) system: wavefront control and detection of exoplanets with coherent light modulation in the focal plane. In *Adaptive optics systems II*, volume 7736, pages 792–800. SPIE, 2010.

[37] SY Haffert, JR Males, K Ahn, K Van Gorkom, O Guyon, LM Close, JD Long, AD Hedglen, L Schatz, M Kautz, et al. Implicit electric field conjugation: Data-driven focal plane control. *Astronomy & Astrophysics*, 673:A28, 2023.

[38] Fredric M Ham and R Grover Brown. Observability, eigenvalues, and kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems*, (2):269–273, 1983.

[39] John W Hardy. Active optics: a new technology for the control of light. *Proceedings of the IEEE*, 66(6):651–697, 1978.

[40] John W Hardy. *Adaptive optics for astronomical telescopes*, volume 16. Oxford University Press, USA, 1998.

[41] Yoshio Hayasaki, Takashi Sugimoto, Akihiro Takita, and Nobuo Nishida. Variable holographic femtosecond laser processing by use of a spatial light modulator. *Applied Physics Letters*, 87(3), 2005.

[42] Sara R Heap, Don J Lindler, Thierry M Lanz, Robert H Cornett, Ivan Hubeny, Stephen P Maran, and Bruce Woodgate. Space telescope imaging spectrograph coronagraphic observations of β pictoris. *The Astrophysical Journal*, 539(1):435, 2000.

[43] Keith M. Hinrichs and John J. Piotrowski. Neural networks for faster optical alignment. *Optical Engineering*, 59(7):1 – 12, 2020.

[44] Haonan Jiang and Yuanli Cai. Adaptive fifth-degree cubature information filter for multi-sensor bearings-only tracking. *Sensors*, 18(10):3241, 2018.

[45] S. J. Julier and J. K. Uhlmann. New extension of the kalman filter to nonlinear systems. In *Signal processing, sensor fusion, and target recognition VI*, volume 3068, pages 182–193. International Society for Optics and Photonics, 1997.

[46] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.

[47] Paul Kalas, James R Graham, Eugene Chiang, Michael P Fitzgerald, Mark Clampin, Edwin S Kite, Karl Stapelfeldt, Christian Marois, and John Krist. Optical images of an exosolar planet 25 light-years from earth. *science*, 322(5906):1345–1348, 2008.

[48] Richard L Kendrick, Daniel S Acton, and AL Duncan. Phase-diversity wave-front sensor for imaging systems. *Applied Optics*, 33(27):6533–6546, 1994.

[49] Eugene D Kim, Young-Wan Choi, Myung-Seok Kang, and Se Chol Choi. Reverse-optimization alignment algorithm using zernike sensitivity. *Journal of the Optical Society of Korea*, 9(2):68–73, 2005.

[50] Seonghui Kim, Ho-Soon Yang, Yun-Woo Lee, and Sug-Whan Kim. Merit function regression method for efficient alignment control of two-mirror optical systems. *Opt. Express*, 15(8):5059–5068, Apr 2007.

[51] Caroline Kulcsár, Henri-François Raynaud, Cyril Petit, Jean-Marc Conan, and Patrick Viaris de Lesegno. Optimal control, observers and integrators in adaptive optics. *Optics express*, 14(17):7464–7476, 2006.

[52] A-M Lagrange, M Bonnefoy, G Chauvin, D Apai, D Ehrenreich, A Boccaletti, D Gratadour, D Rouan, D Mouillet, S Lacour, et al. A giant planet imaged in the disk of the young star β pictoris. *Science*, 329(5987):57–59, 2010.

[53] Olivier Lardière, Christian Marois, William Thompson, Garima Singh, Adam Johnson, Joeleff Fitzsimmons, Tim Hardy, Mamadou N'diaye, Wolfgang Heidrich, Qiang Fu, et al. Optical design of spiders, a subaru pathfinder instrument for detecting exoplanets and retrieving spectra. In *Adaptive Optics Systems VIII*, volume 12185, pages 1252–1262. SPIE, 2022.

[54] Brice Le Roux, Jean-Marc Conan, Caroline Kulcsár, Henri-François Raynaud, Laurent M Mugnier, and Thierry Fusco. Optimal control law for classical and multiconjugate adaptive optics. *Journal of the Optical Society of America A*, 21(7):1261–1276, 2004.

[55] Hanshin Lee, Gavin B Dalton, Ian AJ Tosh, and Sug-Whan Kim. Computer-guided alignment i: Phase and amplitude modulation of alignment-influenced optical wavefront. *Optics express*, 15(6):3127–3139, 2007.

[56] Hanshin Lee, Gavin B Dalton, Ian AJ Tosh, and Sug-Whan Kim. Computer-guided alignment ii: Optical system alignment using differential wavefront sampling. *Optics express*, 15(23):15424–15437, 2007.

[57] Duan Li and Dmitry Savransky. Dataset for image plane wavefront sens-

ing and kalman filtering for automated alignment of off-axis aspherics. https://doi.org/10.7298/49w5-d093.

[58] Duan Li and Dmitry Savransky. Automated reflective optical system alignment with focal plane sensing and Kalman filters. In *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation IV*, volume 11451, pages 932 – 941. International Society for Optics and Photonics, SPIE, 2020.

[59] Duan Li and Dmitry Savransky. Automated reflective optical system alignment: analysis and experiments. In Keith B. Doyle, Jonathan D. Ellis, José M. Sasián, and Richard N. Youngworth, editors, *Optomechanics and Optical Alignment*, volume 11816, page 118160B. International Society for Optics and Photonics, SPIE, 2021.

[60] Duan Li and Dmitry Savransky. Image plane wavefront sensing and kalman filtering for automated alignment of off-axis aspherics. *JOSA A*, 40(7):1327–1336, 2023.

[61] Duan Li, William Thompson, Dmitry Savransky, and Christian Marois. Focal plane wavefront control for the gemini planet imager 2.0 calibration system (cal2). In *Techniques and Instrumentation for Detection of Exoplanets XI*, volume 12680, pages 388–394. SPIE, 2023.

[62] John Z Lou, David Redding, Norbert Sigrist, Yan Zhang, and Scott Basinger. Jwst on-orbit multi-field wavefront control with a kalman filter. In *Optical Modeling and Performance Predictions II*, volume 5867, page 58670T. International Society for Optics and Photonics, 2005.

[63] Mark A Lundgren and William L Wolfe. Alignment of a three-mirror off-axis telescope by reverse optimization. *Optical engineering*, 30(3):307–311, 1991.

[64] Bernard Lyot. The study of the solar corona and prominences without eclipses (george darwin lecture, 1939). *Monthly Notices of the Royal Astronomical Society, Vol. 99, p. 580*, 99:580, 1939.

[65] B Macintosh, JR Graham, T Barman, RJ De Rosa, Q Konopacky, MS Marley, C Marois, EL Nielsen, L Pueyo, A Rajan, et al. Discovery and spectroscopy of the young jovian planet 51 eri b with the gemini planet imager. *Science*, 350(6256):64–67, 2015.

[66] Bruce Macintosh, James R Graham, Patrick Ingraham, Quinn Konopacky, Christian Marois, Marshall Perrin, Lisa Poyneer, Brian Bauman, Travis Barman, Adam S Burrows, et al. First light of the gemini planet imager. *proceedings of the National Academy of Sciences*, 111(35):12661–12666, 2014.

[67] Bruce A Macintosh, James R Graham, David W Palmer, René Doyon, Jennifer Dunn, Donald T Gavel, James Larkin, Ben Oppenheimer, Les Saddlemyer, Anand Sivaramakrishnan, et al. The gemini planet imager: from science to design to construction. In *Adaptive Optics Systems*, volume 7015, pages 315–327. SPIE, 2008.

[68] Alexander Madurowicz, Bruce Macintosh, Lisa Poyneer, Duan Li, Jean-Baptiste Ruffio, Jean-Pierre Veran, S Mark Ammons, Dmitry Savransky, Jeffrey Chilcote, Jerome Maire, et al. Gpi 2.0: Optimizing reconstructor performance in simulations and preliminary contrast estimates. In *Adaptive Optics Systems VII*, volume 11448, pages 526–545. SPIE, 2020.

[69] Naoyuki Maeda. Clinical applications of wavefront aberrometry–a review. *Clinical & experimental ophthalmology*, 37(1):118–129, 2009.

[70] Daniel Malacara. *Optical shop testing*. John Wiley & Sons, 2007.

[71] Christian Marois, B Gerard, O Lardière, A Anthony, C Bradley, J Dunn, Qiang Fu, T Hardy, Wolfgang Heidrich, G Herriot, et al. Upgrading the gemini planet imager calibration unit with a photon counting focal plane wavefront sensor. In *Adaptive Optics Systems VII*, volume 11448, pages 1366–1376. SPIE, 2020.

[72] Christian Marois, David Lafreniere, René Doyon, Bruce Macintosh, and Daniel Nadeau. Angular differential imaging: a powerful high-contrast imaging technique. *The Astrophysical Journal*, 641(1):556, 2006.

[73] Christian Marois, Olivier Lardière, Joeleff Fitzsimmons, Mathieu Demers, Adam B Johnson, Saugata Dutt, Patrick Nkwari, Benjamin Gerard, Jennifer Dunn, Duan Li, et al. Cal2: project update of the nrc canada facility-class focal plane wavefront sensor for the gemini planet imager 2 upgrade. In *Adaptive Optics Systems IX*, volume 13097, pages 33–46. SPIE, 2024.

[74] Christian Marois, Olivier Lardière, William Thompson, Garima Singh, Adam Johnson, Tim Hardy, Joeleff Fitzsimmons, Benjamin L Gerard, Suresh Sivanandam, Simon Thibault, et al. Deployment of focal plane wfs technologies on 8-m telescopes: from the subaru spiders pathfinder,

to the facility-class gpi 2.0 cal2 system. In *Adaptive Optics Systems VIII*, volume 12185, pages 594–603. SPIE, 2022.

[75] Christian Marois, Bruce Macintosh, Travis Barman, B Zuckerman, Inseok Song, Jennifer Patience, David Lafrenière, and René Doyon. Direct imaging of multiple planets orbiting the star hr 8799. *science*, 322(5906):1348–1352, 2008.

[76] Christian Maurer, Alexander Jesacher, Stefan Bernet, and Monika Ritsch-Marte. What spatial light modulators can do for optical microscopy. *Laser & Photonics Reviews*, 5(1):81–101, 2011.

[77] Dimitri Mawet, Laurent Pueyo, Peter Lawson, Laurent Mugnier, Wesley Traub, Anthony Boccaletti, John T Trauger, Szymon Gladysz, Eugene Serabyn, Julien Milli, et al. Review of small-angle coronagraphic techniques in the wake of ground-based second-generation adaptive optics systems. In *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, volume 8442, pages 62–82. SPIE, 2012.

[78] Michel Mayor and Didier Queloz. A jupiter-mass companion to a solar-type star. *nature*, 378(6555):355–359, 1995.

[79] David McGloin, Gabriel C Spalding, H Melville, Wilson Sibbett, and Kishan Dholakia. Applications of spatial light modulators in atom optics. *Optics express*, 11(2):158–166, 2003.

[80] Brittany E Miles, Beth A Biller, Polychronis Patapis, Kadin Worthen, Emily Rickman, Kielan KW Hoch, Andrew Skemer, Marshall D Perrin, Niall Whiteford, Christine H Chen, et al. The jwst early-release science program for direct observations of exoplanetary systems ii: A 1 to 20 μm spectrum of the planetary-mass companion vhs 1256–1257 b. *The Astrophysical journal letters*, 946(1):L6, 2023.

[81] NASA Exoplanet Archive. NASA Exoplanet Archive. `https://exoplanetarchive.ipac.caltech.edu/`, 2025. Accessed: 2025-04-28.

[82] Daniel R Neal and Justin Mansell. Application of shack-hartmann wavefront sensors to optical system calibration and alignment. In *Adaptive Optics For Industry And Medicine*, pages 234–240. World Scientific, 2000.

[83] Meiji N Nguyen, Bryony F Nickson, Emiel H Por, Rémi Soummer, John G Hagopian, Bruce Macintosh, Jeffrey Chilcote, Laurent Pueyo, Marshall

Perrin, and Quinn Konopacky. Gpi 2.0: optical designs for the upgrade of the gemini planet imager coronagraphic system. In *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation V*, volume 12188, pages 1447–1458. SPIE, 2022.

[84] Eun-Song Oh, Seonghui Kim, Yunjong Kim, Hanshin Lee, Sug-Whan Kim, and Ho-Soon Yang. Integration of differential wavefront sampling with merit function regression for efficient alignment of three-mirror anastigmat optical system. In *Optical System Alignment, Tolerancing, and Verification IV*, volume 7793, page 77930F. International Society for Optics and Photonics, 2010.

[85] Polychronis Patapis, Evert Nasedkin, Gabriele Cugno, Adrian Michael Glauser, Ioannis Argyriou, Niall P Whiteford, Paul Mollière, Alistair Glasse, and Sascha Patrick Quanz. Direct emission spectroscopy of exoplanets with the medium resolution imaging spectrometer on board jwst miri-i. molecular mapping and sensitivity to instrumental effects. *Astronomy & Astrophysics*, 658:A72, 2022.

[86] Marshall Perrin, Eric Gentry, and Robert De Rosa. gpipsfs: Gpi point spread function (psf) simulation toolkit. `https://github.com/geminiplanetimager/gpipsfs/tree/master`, 2020.

[87] Cyril Petit, Jean-Marc Conan, Caroline Kulcsár, Henri-Francois Raynaud, and Thierry Fusco. First laboratory validation of vibration filtering with lqg control law for adaptive optics. *Optics Express*, 16(1):87–97, 2008.

[88] Ben C Platt and Roland Shack. History and principles of shack-hartmann wavefront sensing, 2001.

[89] Lisa A. Poyneer, David W. Palmer, Bruce Macintosh, Dmitry Savransky, Naru Sadakuni, Sandrine Thomas, Jean-Pierre Véran, Katherine B. Follette, Alexandra Z. Greenbaum, S. Mark Ammons, Vanessa P. Bailey, Brian Bauman, Andrew Cardwell, Daren Dillon, Donald Gavel, Markus Hartung, Pascale Hibon, Marshall D. Perrin, Fredrik T. Rantakyrö, Anand Sivaramakrishnan, and Jason J. Wang. Performance of the gemini planet imager's adaptive optics system. *Appl. Opt.*, 55(2):323–340, Jan 2016.

[90] Lisa A Poyneer, David W Palmer, Bruce Macintosh, Dmitry Savransky, Naru Sadakuni, Sandrine Thomas, Jean-Pierre Véran, Katherine B Follette, Alexandra Z Greenbaum, S Mark Ammons, et al. Performance of the gemini planet imager's adaptive optics system. *Applied Optics*, 55(2):323–340, 2016.

[91] Lisa A Poyneer and Jean-Pierre Véran. Kalman filtering to suppress spurious signals in adaptive optics control. *Journal of the Optical Society of America A*, 27(11):A223–A234, 2010.

[92] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, third edition, 2007.

[93] K. R. Rao and P. C. Yip. *The Transform and Data Compression Handbook*. CRC Press, first edition, 2000.

[94] George R Ricker, Joshua N Winn, Roland Vanderspek, David W Latham, Gáspár Á Bakos, Jacob L Bean, Zachory K Berta-Thompson, Timothy M Brown, Lars Buchhave, Nathaniel R Butler, et al. Transiting exoplanet survey satellite. *Journal of Astronomical Telescopes, Instruments, and Systems*, 1(1):014003–014003, 2015.

[95] Françcois Rigaut, Brent L Ellerbroek, and Malcolm J Northcott. Comparison of curvature-based and shack–hartmann-based adaptive optics for the gemini telescope. *Applied optics*, 36(13):2856–2868, 1997.

[96] AJ Eldorado Riggs, N Jeremy Kasdin, and Tyler D Groff. Recursive starlight and bias estimation for high-contrast imaging with an extended kalman filter. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(1):011017, 2016.

[97] François Roddier. *Adaptive optics in astronomy*. Cambridge university press, 1999.

[98] Francois Roddier, Claude Roddier, and Nicolas Roddier. Curvature sensing: a new wavefront sensing method. In *Statistical Optics*, volume 976, pages 203–209. SPIE, 1988.

[99] Patrick S Salter and Martin J Booth. Adaptive optics in laser processing. *Light: Science & Applications*, 8(1):110, 2019.

[100] D. Savransky, S. J. Thomas, L. A. Poyneer, and B. A. Macintosh. Computer vision applications for coronagraphic optical alignment and image processing. *Applied Optics*, 52(14):3394–3403, 2013.

[101] Dmitry Savransky, Bruce A Macintosh, Sandrine J Thomas, Lisa A Poyneer, David W Palmer, Robert J De Rosa, and Markus Hartung. Fo-

cal plane wavefront sensing and control for ground-based imaging. In *Adaptive Optics Systems III*, volume 8447, pages 2218–2228. SPIE, 2012.

[102] Dmitry Savransky, Sandrine J Thomas, Lisa A Poyneer, and Bruce A Macintosh. Computer vision applications for coronagraphic optical alignment and image processing. *Applied Optics*, 52(14):3394–3403, 2013.

[103] Steven C Schallhorn, Ayad A Farjo, David Huang, Brian S Boxer Wachler, William B Trattler, David J Tanzer, Parag A Majmudar, and Alan Sugar. Wavefront-guided lasik for the correction of primary myopia and astigmatism: a report by the american academy of ophthalmology. *Ophthalmology*, 115(7):1249–1261, 2008.

[104] C. Schindlbeck, C. Pape, and E. Reithmeier. Process-integrated state estimation of optical systems with macro–micro manipulators based on wavefront filtering. *IEEE Robotics and Automation Letters*, 4(4):4078–4085, 2019.

[105] Marc Séchaud. Wave-front compensation devices. *Adaptive optics in astronomy*, page 57, 1999.

[106] D. Simon. *Optimal State Estimation*. John Wiley & Sons, Ltd, 2006.

[107] R. Soummer, L. Pueyo, and J. Larkin. Detection and characterization of exoplanets and disks using projections on karhunen-loève eigenimages. *Astrophysical Journal Letters*, 755(2), 2012.

[108] Rémi Soummer, Laurent Pueyo, and James Larkin. Detection and characterization of exoplanets and disks using projections on karhunen–loève eigenimages. *The Astrophysical Journal Letters*, 755(2):L28, 2012.

[109] Remi Soummer, Laurent Pueyo, Anand Sivaramakrishnan, and Robert J Vanderbei. Fast computation of lyot-style coronagraph propagation. *Optics Express*, 15(24):15935–15951, 2007.

[110] Frank Staals, Alena Andryzhyieuskaya, Hans Bakker, Marcel Beems, Jo Finders, Thijs Hollink, Jan Mulkens, Angelique Nachtwein, Rob Willekers, Peter Engblom, et al. Advanced wavefront engineering for improved imaging and overlay applications on a 1.35 na immersion scanner. In *Optical Microlithography XXIV*, volume 7973, pages 491–503. SPIE, 2011.

[111] R. F. Stengel. *Optimal Control and Estimation*. Dover, 2012.

[112] John T Trauger, Chris Burrows, Brian Gordon, Joseph Jacob Green, Andrew E Lowman, Dwight Moody, Albert F Niessner, Fang Shi, and Daniel Wilson. Coronagraph contrast demonstrations with the high-contrast imaging testbed. In *Optical, Infrared, and Millimeter Space Telescopes*, volume 5487, pages 1330–1336. SPIE, 2004.

[113] John T Trauger and Wesley A Traub. A laboratory demonstration of the capability to image an earth-like extrasolar planet. *Nature*, 446(7137):771–773, 2007.

[114] R. Van der Merwe and E. A. Wan. The square-root unscented kalman filter for state and parameter-estimation. In *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings*, volume 6, pages 3461–3464, 2001.

[115] Robert J Vanderbei, Eric Cady, and N Jeremy Kasdin. Optimal occulter design for finding extrasolar planets. *The Astrophysical Journal*, 665(1):794, 2007.

[116] Mikhail A Vorontsov and Viktor P Sivokon. Stochastic parallel-gradient-descent technique for high-resolution wave-front phase-distortion correction. *Journal of the Optical Society of America A*, 15(10):2745–2758, 1998.

[117] Andrew M Weiner. Femtosecond pulse shaping using spatial light modulators. *Review of scientific instruments*, 71(5):1929–1960, 2000.

[118] Thomas Weyrauch and Mikhail A. Vorontsov. *Free-space laser communications with adaptive optics: Atmospheric compensation experiments*, pages 247–271. Springer New York, New York, NY, 2008.

[119] William Whitacre and Mark E Campbell. Decentralized geolocation and bias estimation for uninhabited aerial vehicles with articulating cameras. *Journal of guidance, control, and dynamics*, 34(2):564–573, 2011.

[120] David R. Williams and Jason Porter. *Development of Adaptive Optics in Vision Science and Ophthalmology*, chapter 1, pages 1–29. John Wiley Sons, Ltd, 2006.

[121] Ho-Soon Yang, Sung-Hee Kim, Yun-Woo Lee, Jae-Bong Song, Hyug-Gyo Rhee, Hoi-Yoon Lee, Jae-Hyeob Lee, In-Won Lee, and Sug-Whan Kim. Computer aided alignment using zernike coefficients. In *Interferometry XIII: Applications*, volume 6293, page 62930I. International Society for Optics and Photonics, 2006.

[122] Bin Zhang, Xiaohui Zhang, Cheng Wang, and Changyuan Han. Computer-aided alignment of the complex optical system. In *Advanced Optical Manufacturing and Testing Technology 2000*, volume 4231, pages 67–72. International Society for Optics and Photonics, 2000.