Use of Light Coherence for Exoplanet Detection and Characterization

Thesis by
Yeyuan (Yinzi) Xin

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy in Physics

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2025
Defended May 28th, 2025

Yeyuan (Yinzi) Xin
ORCID: 0000-0002-6171-9081

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dimitri Mawet, who first took me on as a SURF student, then again as a PhD student. You have always been so supportive of me, giving me independence in my research and providing incredibly helpful advice and insight at many stages over these years. I always come out of our meetings with renewed direction as well as renewed motivation, and I am so grateful to have been part of your group.

I would also like to acknowledge the many people who have guided, supported, and enriched my journey in the field of high-contrast imaging. In chronological order, I would like to first thank Garreth Ruane and Jacques Robert-Delorme, who mentored me as an undergraduate student, an experience I greatly enjoyed and which led me to apply to graduate school in the field.

Next, thank you to Kerri Cahoy and Ewan Douglas, my mentors during my Master's program at MIT. It was really cool getting to work on a cubesat that a) made it into space and b) accomplished its goal! Ewan especially taught me so much of what I know about optics and astronomy, and always took care to include me in research activities I would find interesting. Thank you as well to fellow STARLab graduate students Rachel Morgan and Greg Allan, who brightened my time at MIT and remain great colleagues and friends.

I also want to thank Laurent Pueyo, who took me in as a visiting researcher at the STScI and gave me the support and mentorship I needed to finish my Master's project, as well as Jules Fowler, Iva Laginja, Scott Will, and Remi Soummer for creating such a lively and welcoming environment at STScI.

At Caltech, I would like to thank Kent Wallace for mentorship on an early project on the Zernike wavefront sensor, and Nem Jovanovic, who enabled my work in photonic lanterns, including me in an amazing collaboration and providing advice and support throughout the way. I want to thank Dan Echeverri, Jerry Xuan, Yoo Jung Kim, and Jon Lin, who have all helped me with my research in significant ways; Sunil Golwala, Ashley Baker, and Katie Bouman for their interest in my research and for serving on my committee; and Judy McClain for her patience and help with all of my conference travels. To former and current group mates: Jorge, Dan, Niyati, Jerry, Katelyn, Aniket, Elijah, Jason, JB, Ashley, Bin, Arielle, Susan, Yapeng, and Andrea — overlapping with you all has been such a great pleasure!

Thank you as well to the SCExAO team, in particular Julien Lozi, Sebastien Vievard, Olivier Guyon, and Vincent Deo, for supporting my lanterns work at such an amazing facility. I also wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Last but not least, I wish to thank those in my personal life. I am very grateful to my parents, whose hard work and sacrifices have enabled me to chase my dreams. To my little sister Faye, and my sworn sisters Carol, Stephanie, and Kathy — I am so glad to have all of you in my life. Finally, to my partner Willis, who has been by my side this whole time even when we were a hemisphere apart — I want to thank you for more than I can express in words, and I can't wait to experience our future adventures together.

ABSTRACT

Since the first detection of an exoplanet in 1992, over 5,000 exoplanets have now been found through a variety of methods, both indirect (such as the radial velocity or transit method) and direct (such as with imaging, coronagraphy, and interferometry). The direct imaging and spectroscopy of exoplanets in particular plays a key role in characterizing their atmospheres, which can help distinguish between different models of planet formation and detect molecular signatures associated with life. However, directly observing exoplanets is extremely difficult: the small angular separations between the star and the planet require large telescopes to resolve, and the flux ratios between a planet and its star range from 10^{-4} in the infrared for hot, young, massive planets to 10^{-10} in the optical for mature Earth-like planets. Photon noise from the star drowns out the planet signal in conventional imagers or spectrographs, so dedicated instruments are needed to remove the majority of the starlight before it reaches the detector. Additional wavefront sensing and control methods are also needed to compensate for aberrations in the system — from fast varying atmospheric fluctuations to slower quasi-static drifts in the instrument and telescope.

This thesis presents advances in instrumentation for directly characterizing exoplanets, focusing on exploiting the coherence properties of light to increase sensitivity. It presents the invention of the Photonic Lantern Nuller, which uses a multimode-to-single-mode demultiplexing waveguide to cancel out starlight while maintaining planet light, allowing for the direct characterization of planets at a telescope's diffraction limit. The PLN was experimentally characterized in the lab, enhanced using common-path wavefront sensing control techniques, and demonstrated on sky at the Subaru Telescope. This thesis also presents work on the Keck Planet Imager and Characterizer, a fiber-fed high-resolution spectrograph — namely, the on-sky demonstration of the speckle nulling technique to destructively interfere residual starlight. Future directions in optimal stellar suppression and instrument-informed data analysis techniques are also discussed.

The advances in instrumentation and methodology from this work have applications to giant planets on existing ten-meter class ground-based telescopes, Earth-like exoplanets on the planned Habitable Worlds Observatory space telescope, and many planets of interest on future thirty-meter class telescopes.

PUBLISHED CONTENT AND CONTRIBUTIONS

Xin, Yinzi, Daniel Echeverri, Nemanja J. Jovanovic, et al. (2025). "Implicit electric field conjugation with a photonic lantern nuller". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 11.2, p. 025004. DOI: 10.1117/1.JATIS.11.2.025004. URL: https://doi.org/10.1117/1.JATIS.11.2.025004.

Xin, Yinzi, Daniel Echeverri, Nemanja Jovanovic, et al. (2024a). "Laboratory demonstration of a Photonic Lantern Nuller in monochromatic and broadband light". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 10.2, p. 025001. DOI: 10.1117/1.JATIS.10.2.025001. URL: https://doi.org/10.1117/1.JATIS.10.2.025001.
Y.X. conducted the laboratory experiments to characterize the lantern and demonstrate its nulling properties, and wrote the paper.

– (2024b). "Wavefront sensing and control for a photonic lantern nuller for exoplanet characterization". In: *Optical and Infrared Interferometry and Imaging IX*. Ed. by Jens Kammerer, Stephanie Sallum, and Joel Sanchez-Bermudez. Vol. 13095. International Society for Optics and Photonics. SPIE, 130950S. DOI: 10.1117/12.3019168. URL: https://doi.org/10.1117/12.3019168.
Y.X. adapted the code used to run the experiment, performed the experiment, analyzed the data, and wrote the paper.

Xin, Yinzi, Laurent Pueyo, et al. (Mar. 2024). "Coronagraphic Data Post-processing Using Projections on Instrumental Modes". In: *The Astrophysical Journal* 963.2, 96, p. 96. DOI: 10.3847/1538-4357/ad1879. arXiv: 2401.04269 [astro-ph.IM].
Y.X. conceptualized the post-processing technique, formulated the mathematical model and projection matrix, demonstrated the technique using numerical simulations, characterized its performance using detection testing techniques, and wrote the paper.

Xin, Yinzi, Daniel Echeverri, Nemanja Jovanovic, et al. (2023). "Laboratory characterization of a mode-selective photonic lantern for exoplanet characterization". In: *Techniques and Instrumentation for Detection of Exoplanets XI*. Ed. by Garreth J. Ruane. Vol. 12680. International Society for Optics and Photonics. SPIE, p. 126800I. DOI: 10.1117/12.2676421. URL: https://doi.org/10.1117/12.2676421.
Y.X. conducted the laboratory experiments to characterize the lantern and demonstrate its nulling properties, and wrote the paper.

Xin, Yinzi, Jerry W. Xuan, Dimitri Mawet, Jason Wang, Garreth Ruane, Daniel Echeverri, Nemanja Jovanovic, Clarissa Do Ó, et al. (July 2023). "On-sky speckle nulling through a single-mode fiber with the Keck Planet Imager and Characterizer". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 9, 035001, p. 035001. DOI: 10.1117/1.JATIS.9.3.035001. arXiv: 2307.11893 [astro-ph.IM].

Y.X. wrote the code to perform speckle nulling with KPIC, tested the algorithm using the internal source, demonstrated it on-sky, performed the data analysis, and wrote the paper.

Xin, Yinzi, Jerry W. Xuan, Dimitri Mawet, Jason Wang, Garreth Ruane, Daniel Echeverri, Nemanja Jovanovic, Clarissa Ó, et al. (June 2023). "On-sky speckle nulling through a single-mode fiber with the Keck Planet Imager and Characterizer". In: *Adaptive Optics for Extremely Large Telescopes 7th Edition*. ONERA. Avignon, France. DOI: 10.13009/AO4ELT7-2023-008. URL: https://hal.science/hal-04419970.

Y.X. wrote the code to perform speckle nulling with KPIC, tested the algorithm using the internal source, demonstrated it on-sky, performed the data analysis, and wrote the paper.

Xin, Yinzi, Nemanja Jovanovic, et al. (Oct. 2022). "Efficient Detection and Characterization of Exoplanets within the Diffraction Limit: Nulling with a Mode-selective Photonic Lantern". In: *The Astrophysical Journal* 938.2, 140, p. 140. DOI: 10.3847/1538-4357/ac9284. arXiv: 2209.07644 [astro-ph.IM].

Y.X. conceptualized the instrument, derived its operating principles, and performed simulations predicting its behavior and performance, and wrote the paper.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 The Study of Exoplanets

Though humans have long speculated about the existence of exoplanets — planets orbiting stars other than our own sun — the first evidence came only in 1992, with the discovery of two planets orbiting a pulsar (Wolszczan and Frail, 1992). This revelation was quickly followed by the 1995 discovery of another planet, the hot jupiter 51 Pegasi b, orbiting a Solar-type star (Mayor and Queloz, 1995). As shown in Fig. 1.1, over 5,000 exoplanets have since been found through a variety of methods (Akeson et al., 2013)[1], and the study of these objects holds promise for answering two of the greatest scientific questions of our time: how did we get here? And — are we alone?

Theories of planet formation are intricately linked to those of stellar formation and the properties of the circumstellar environment. When a star forms from the gravitational collapse of a molecular cloud, conservation of angular momentum leads to the emergence of a disk of leftover gas and dust that orbits the star. This initial protoplanetary disk then provides the materials from which planets are formed.

Two primary pathways have been proposed for planet formation. One is a bottom-up process (depicted in Fig. 1.2 with solid lines), in which small dust grains first coagulate into millimeter-size pebbles, which then clump together (e.g. through streaming instability (Youdin and Goodman, 2005)), then accrete more pebbles over time (Johansen et al., 2021) to eventually form a rocky core. While the smaller cores will remain as rocky planets, more massive cores may continue to accrete gas and dust until they become gas giants (Youdin and Zhu, 2025).

The second pathway is a top-down process that only applies to giant planets, in which an instability in the protoplanetary disk causes the local material to collapse into a giant planet, which then accretes additional material along its orbit (Mercer, Anthony and Stamatellos, Dimitris, 2020). This process is indicated in Fig. 1.2 as the dashed blue line. This disk instability formation pathway is expected to primarily produce gas giants far away from their host stars, where conditions for collapse are more favorable.

[1] NASA exoplanet archive: https://exoplanetarchive.ipac.caltech.edu/

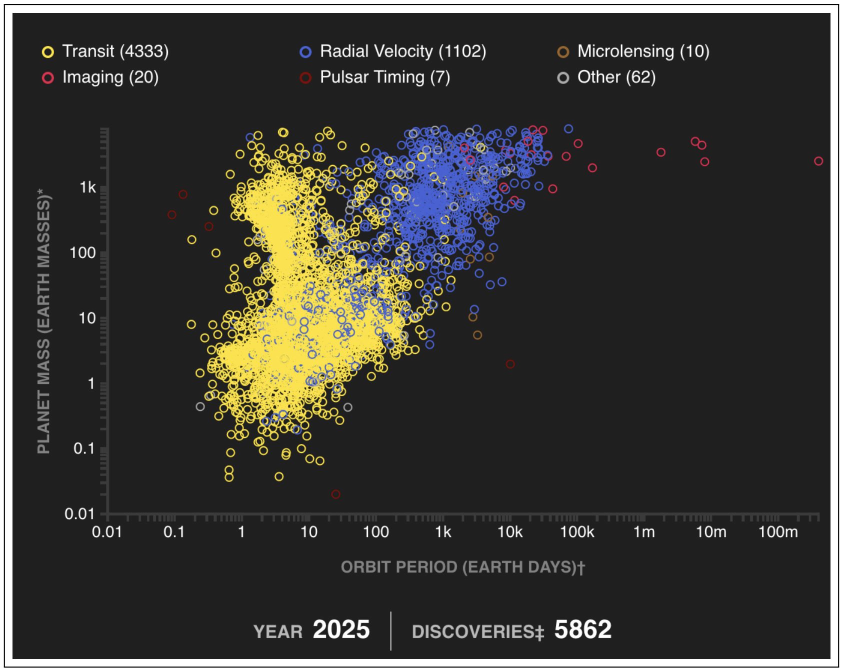

Figure 1.1: Known exoplanets as of March 2025. Since the first discovery of two exoplanets orbiting a pulsar in 1995, over 5000 exoplanets have been found by a variety of methods, spanning a wide range of parameters. Studying exoplanets can help answer questions about how our own solar system formed and inform us about the existence of life elsewhere in the universe. Figure from Akeson et al. (2013).

Testing these theories requires observational data. Observable quantities such as an exoplanet's mass, luminosity, orbital parameters, and molecular composition can probe the formation history of the exoplanet, and their distributions in aggregate can also provide insight about exoplanet demographics as a whole. One complication is that formed planets will continue to interact with their own disks, other planets in the system, and other gravitationally bound objects (such as a binary companion to their host star). As a result, they can migrate through their disk on timescales that depend on the specific process, and their current-day location may not be indicative of where they originally formed. Untangling the myriad of factors that contribute to the diversity of observed exoplanets is therefore a significant challenge.

Characterizing exoplanet atmospheres through spectroscopy is a particularly powerful approach for understanding exoplanet formation, as the chemical composition of an exoplanet is a strong tracer of its history (Mollière et al., 2022). For example, a giant planet that formed via disk instability is expected to have a composition similar to that of its host star, while planets that formed through core accretion may have a different composition that varies with formation location and the phases (gas

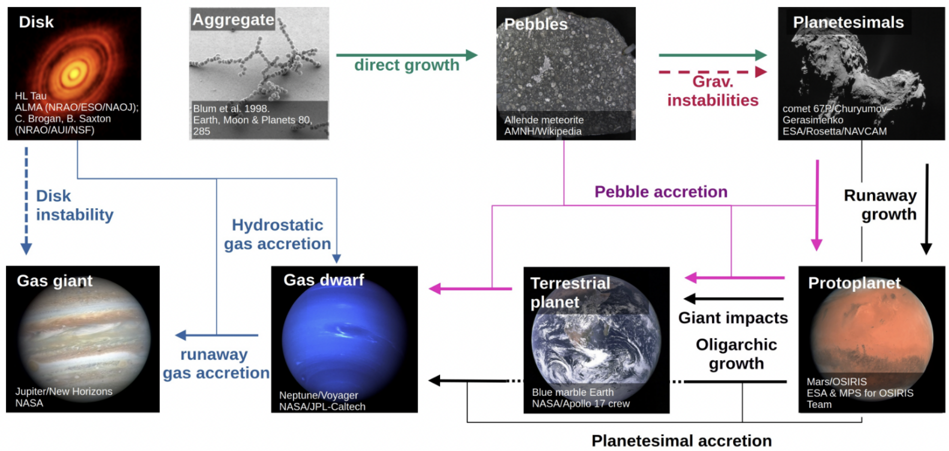

Figure 1.2: Routes to form planetary bodies. Mechanisms are indicated with an arrow and labeled. Gravitational instabilities (involving a collection of particles or gas) are indicated by dashed arrows. Accretion processes that rely on surface forces are in green, gravitational forces in black, gas accretion in blue, and pebble accretion (which involves both gas and gravity) is in magenta. The disk instability (DI) mechanism is depicted at the very left. Other processes fall under the core accretion (CA) umbrella. Figure and caption from Ormel (2024).

or solid) of the compounds there. If a planet migrates its formation, its composition will reflect the materials actually accreted during its migration, helping to distinguish between different models of its formation and migration history (Mollière et al., 2022).

Atmospheric characterization is also crucial for the search for biosignatures that may indicate extraterrestrial life, as biological processes can lead to chemical disequilibrium in a planet's atmosphere (see Fig. 1.3) (National Research Council, 2021). Currently, transit spectroscopy methods can probe the chemical composition of small rocky planets around M dwarfs, while direct spectroscopy can probe the composition of massive self-luminous gas giants. Neither of these planet types is promising for hosting life similar to Earth's — even rocky planets in the habitable zone around M dwarfs, where water is in its liquid phase, likely receive too much X-ray irradiation as a result of the M dwarfs' high stellar activity (National Research Council, 2021). Therefore, one of the most exciting directions in the future of exoplanet science is the direct characterization of true Earth analogues — rocky planets in the habitable zone around solar type stars, the only place we already know that life exists.

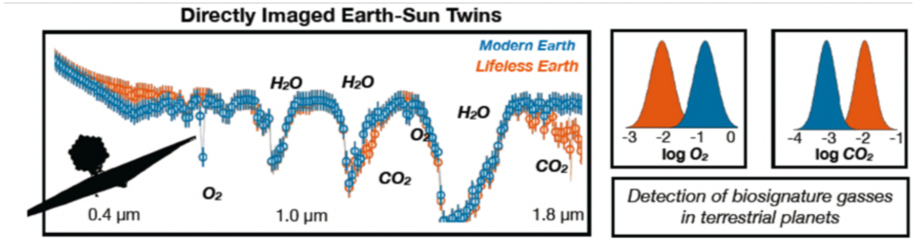

Figure 1.3: Simulated spectra of an Earth-like exoplanet with and without life. Characterizing the chemical abundances of molecules such as O_2 and CO_2 can detect the disequilibrium chemistry that can arise from the existence of life. Figure from the Decadal Survey for Astronomy and Astrophysics (National Research Council, 2021), courtesy of N. Batalha and the PICASO project (Batalha et al., 2019).

1.2 Overview of Observational Methods

Although the very first exoplanets were detected using pulsar timing variations, in which the impact of the exoplanet is seen as a deviation in the arrival times of radio pulses (Wolszczan and Frail, 1992), the majority of planets that have been found since then were detected using other techniques. This section first summarizes the four major *indirect* detection methods, which rely on measurements of the planet's host star to infer the presence of the planet. It then describes the method of directly detecting and measuring the light of exoplanets. Figure 1.4 shows a schematic illustrating each of these detection techniques.

Radial Velocity

The radial velocity approach measures the spectral lines of stars over time to search for periodic frequency oscillations — the Doppler redshift and blueshift of stellar motion induced by the gravitational tug of a planet. The semi-amplitude K_{RV} of this shift is given by

$$K_{RV} = \frac{28.4329\mathrm{ms}^{-1}}{\sqrt{1-e^2}} \frac{M_p \sin i}{M_\mathrm{J}} \left(\frac{M_p + M_*}{M_\odot}\right)^{-2/3} \left(\frac{P}{\mathrm{yr}}\right)^{-1/3}, \tag{1.1}$$

where e is the planet eccentricity, M_p the planet mass, i the planet inclination, M_* the stellar mass, P the planet period, and M_J and $M_\odot$ the masses of Jupiter and the sun respectively (Lovis, Fischer, et al., 2010). This method has been responsible for finding $\sim$ 1000 of known planets to-date (Akeson et al., 2013), and is able to constrain their orbital periods and eccentricities. Unfortunately, the semi-amplitude K_{RV} depends on both the mass of the planet and its orbital inclination, so while the quantity $M_p \sin(i)$ can be deduced, these two parameters cannot be disentangled

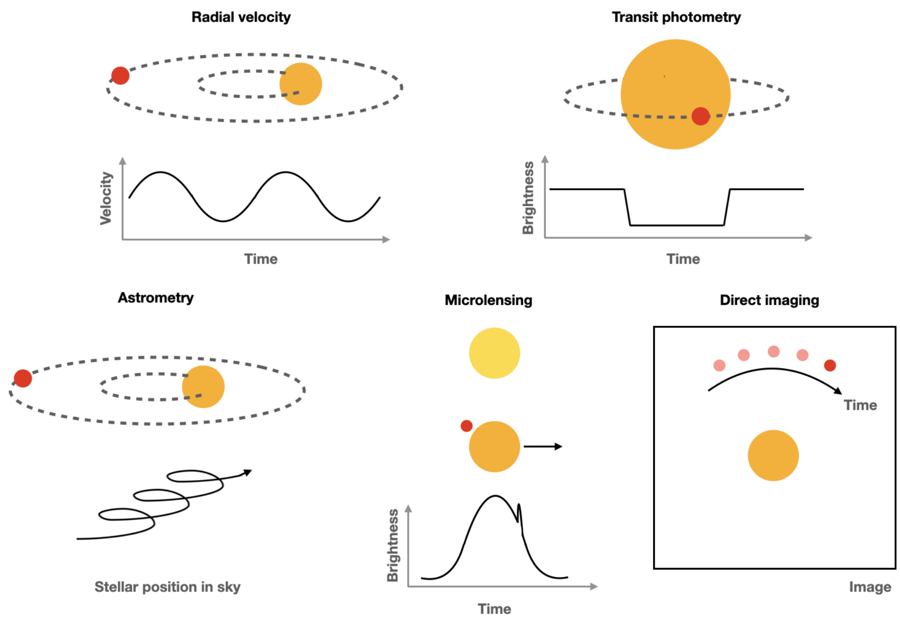

Figure 1.4: Schematics of various exoplanet detection techniques.

from each other. This method also does not probe the planet's atmosphere, and thus provides no direct measurement of its composition.

Transit Photometry

Transit photometry searches for periodic dips in a star system's brightness that result from a planet occulting part of its light as it passes in front of the star (the primary eclipse), as well as when the planet's contribution to the total intensity is blocked as it passes behind the star (the secondary eclipse). If the star is assumed to be uniformly bright, then transit depth ΔF (the fraction of light lost during primary transit) is then the relative areas of the two objects, given by

$$\Delta F = \left(\frac{R_p}{R_*}\right)^2, \tag{1.2}$$

where R_p and R_* are the planet and the stellar radius, respectively.

The transit method, such as used by the National Aeronautics and Space Administration (NASA) Kepler and Transiting Exoplanet Survey Satellite (TESS) missions, have been collectively responsible for detecting over 4000 planets (Akeson et al., 2013). Photometric measurements can be very precise, and this method has been successful in not only detecting planets and measuring their orbital periods, but also in constraining many other parameters (stellar radius, planet radius, orbital inclination) based on the shape of the transit lightcurve. The measurements can also be used to detect the presence of additional planets through variations in the transit timing, a technique known as transit-timing variations (Agol and Fabrycky, 2018).

Additionally, when the planet passes in front of the star, part of the starlight passes through the planet's atmosphere, allowing for the measurement of atomic and molecular absorption features; and when the planet passes behind the star, the light that is lost is then a measure of the planet's emission. These techniques are collectively known as transit spectroscopy, and are one way to measure the atmospheric composition of exoplanets (Seager, 2008).

Astrometry

Like the radial velocity method, the astrometry method relies on measuring a perturbation to the star induced by the planet, in this case deviations in the spatial position of the star relative to its proper motion. The astrometric deviation α in arcsec, induced by a planet, is given by

$$\alpha = (\frac{M_p}{M_*})(\frac{a_p}{\mathrm{AU}})(\frac{d}{\mathrm{pc}})^{-1}\mathrm{arcsec}, \tag{1.3}$$

where a_p is the semi-major axis of the planet, and d the distance to the system (Perryman et al., 2014).

The European Space Agency (ESA) GAIA mission (Gaia Collaboration et al., 2016) has measured the positions of over a billion stars with very high precision, and early data releases have already been used with great success to identify stars whose astrometric motion hints at potential planets, to be subsequently followed up by other methods to make detections. In the coming years, upcoming data releases from GAIA will provide exciting opportunities for additional follow-up observational campaigns.

Microlensing

Microlensing relies on capturing the chance alignment between a planet's host star and a distant background star. The host star, when passing in front of the background star, will act as a gravitational lens that distorts and magnifies the background star's light, and the presence of planets can be inferred from deviations in the lightcurve of this lensing event (Gaudi, 2012). So far, over 50 planets have been found through this method (Akeson et al., 2013), with more detections anticipated from the Roman

Space Telescope mission. However, unlike for the other planet detection methods, microlensing observations of the same system cannot be repeated, so any additional information about the planet would have to be obtained through other means.

Direct Observation

Direct observation, commonly referred to as direct imaging or high contrast imaging, measures the light from the planet itself, which has either been thermally emitted by the planet or reflected from its host star. The most simple and intuitive method is to pass the combined star and planet light through an imaging instrument and record the intensity, corresponding to the astrophysical intensity distribution convolved with the optical transfer function of the telescope. Planets whose separation from their star is greater than the telescope's diffraction limit ($1.22\lambda/D$ where λ is the wavelength of light and D the telescope diameter), and are bright enough to stand out against various sources of noise, will appear in the image as point sources distinct from their host star.

However, directly measuring exoplanet light is extremely difficult: the small angular separations between the star and the planet require large telescopes to resolve, and the flux ratios between a planet and its star ranges from 10^{-4} in the infrared for hot, young, massive planets to 10^{-10} in the optical for mature Earth-like planets (National Research Council, 2021). Photon noise from the star drowns out the planet signal in conventional imagers or spectrographs, so dedicated instruments are needed to remove the majority of the starlight before it reaches the detector. The challenge of reaching small angular scales with high sensitivity is why, to date, the handful of directly detected planets have all been massive (multiple M_J, or the mass of Jupiter), young and self-luminous from the heat of their formation, and widely separated from their host star (Traub and Oppenheimer, 2010).

Despite the limited number of planets that have been detected through this method, direct observations have already led to a greater understanding of this population of giant, widely-separated planets, as well as of brown dwarfs, a type of object larger than planets but smaller than stars, typically deuterium burning but not massive enough to fuse hydrogen. Notable directly imaged systems include the four HR8799 planets (Christian Marois, Bruce Macintosh, et al., 2008), two β Pictoris planets (Lagrange et al., 2009; Nowak et al., 2020), and 51 Eri b (Macintosh et al., 2015), which are among the first to have ever been imaged. A few dozen companions have been detected since, tending to appear around higher mass stars with ages of

10-100 Myr with an occurrence rate of less than 10% (Currie et al., 2023). The masses of these companions cannot be directly measured using imaging but can be inferred based on the age of the system, though they are sensitive to both the age that is assumed and the particular model used to predict the evolution of luminosity over time. Any constraints on the dynamical mass from other methods (primarily astrometry) thus give valuable insight into both the age of a system as well as the validity of various evolutionary models for the system.

A complication in the study of these objects is that the distinction between massive planets and brown dwarfs close to the mass boundary between them ($\sim 13M_J$) is not obvious. The preferred criterion for categorization is based on their formation pathway — planets are formed primarily through accretion while brown dwarfs are formed primarily through collapse — but this is not a directly observable quantity. Many studies have attempted to characterize the potential differences between the planet population and the brown dwarf population. Some approaches analyze the demographics of the objects and their orbital parameters, tentatively showing that the objects on the lower end of the mass spectrum ($2 - 15M_J$, ostensibly planets) are statistically distinct from the objects on the higher end of the mass spectrum ($15 - 75M_J$, ostensibly brown dwarfs), and they tend to appear closer to their host stars, on less eccentric orbits, and with spin axes that align with the spin axis of their star (Currie et al., 2023). These findings are consistent with the theory that the lower mass objects/planets primarily form in a disk around the host star while the higher mass objects/brown dwarfs form through the direct collapse of a molecular cloud.

As discussed in Section 1.1, atmospheric characterization is another approach to constrain the formation pathway of these objects, as objects that form through gravitational collapse are expected to have compositions similar to their host star, while objects that form through core accretion are expected to have the composition of the solid materials in the vicinity of their formation. High-resolution spectroscopy, in particular, enables the detection of distinct absorption lines corresponding to atomic and molecular species in the planet's atmosphere (Q. M. Konopacky et al., 2013). By measuring the abundance of compounds such as CO_2 and H_2O, the C/O and C/H ratios of several directly imaged objects have been constrained to have approximately solar compositions (around stars whose composition is expected to also be approximately solar) (J. W. Xuan et al., 2024). These results are consistent with gravitational collapse, but they may also be consistent with core accretion outside the snowlines of the relevant compounds (CO in this case), as those compounds would be in their solid state and therefore accreted, thus also resulting in a stellar composition. Additional measurements, such as that of the $^{12}C/^{13}C$ ratio, can help distinguish between these possible explanations, as CO inside the snowline is expected to be enriched in ^{13}C (Zhang et al., 2021). However, measuring $^{12}C/^{13}C$ requires higher signal-to-noise data, and has thus only been done for a few objects so far. The James Webb Space Telescope has also recently enabled access to cooler objects, in which NH_3 can be detected, the abundance of which can similarly constrain the atmospheric composition and be used to test models of the object's formation (Mâlin et al., 2025).

Moving forward, the direct characterization of fainter and closer-in planets is needed to understand the full diversity of the planet population and test formation theories for different systems across the exoplanet parameter space. Of particular interest is the bulk of the gas giant population, many of which have been detected using the radial velocity method but not spectroscopically characterized. Transit spectroscopy is inherently limited to exoplanets close to their host star, simply because the probability of transit decreases with orbital distance as approximately R_p/a_p, as, geometrically, a smaller range of possible inclinations will result in the star's occultation from our line of sight. Direct spectroscopy, on the other hand, is theoretically limited only by our telescopes and technology. Consequently, direct spectroscopy is perhaps the only way to characterize the atmosphere of not only gas giants, but true Earth-like planets in the habitable zones of sun-like stars, making the technology development needed to enable this one of the priorities recommended by the National Academy of Sciences 2020 Decadal Survey in Astronomy and Astrophysics (National Research Council, 2021).

1.3 Instrumentation for Direct Exoplanet Observations

This section provides an overview of the starlight-suppressing instruments needed to achieve the angular resolutions and sensitivities necessary for direct exoplanet characterization.

Starlight Suppression

The suppression of light from a central source has historical roots in the two fields of coronagraphy and nulling interferometry, and although these fields have developed their own tools, practices, and terminology, they ultimately share the same underlying principles.

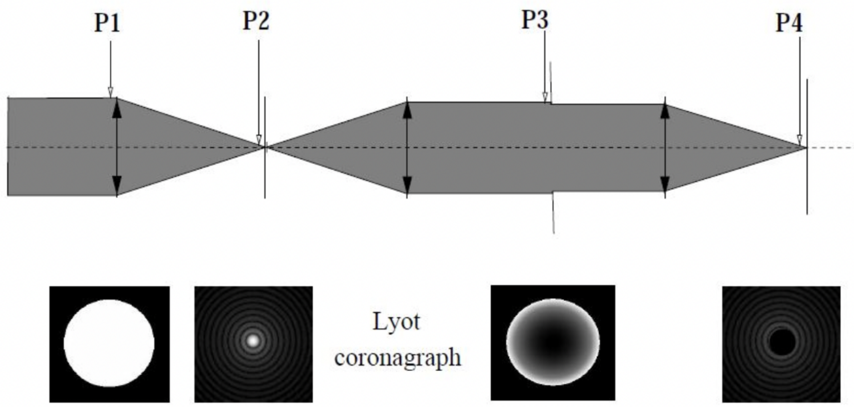

Figure 1.5: A schematic of a classical Lyot coronagraph. Plane P1 is the entrance pupil plane; plane P2 is the focal plane, where the occulting mask is; plane P3 is the second pupil plane, where the Lyot stop is; and plane P4 is the second focal plane, where the light from the on-axis source has been strongly attenuated. Figure and caption adapted from Olivier Guyon et al. (1999).

The first coronagraph was invented in 1931 by Bernard Lyot to observe the corona of the sun without the need for a natural eclipse (Lyot, 1939). The schematic of this classical Lyot coronagraph is shown in Figure 1.5. An intuitive explanation of the instrument is that when the on-axis light from the telescope is focused onto an occulting spot, it gets scattered into a ring in the following pupil plane. This ring of light is spatially filtered out with an undersized pupil mask (the Lyot stop), such that when the beam is refocused onto the detector, most of the light from the central source has been removed. Meanwhile, off-axis sources that do not land on the occulter will pass through the system mostly unchanged.

Coronagraphs in general are modeled and designed using Fourier optics. The propagation of light between successive pupil and focal planes in an imaging system can be approximated with Fraunhofer diffraction, which is valid when diffracted light is viewed in the far-field or, as in this case, a lens is used to focus the light. According to the Fraunhofer diffraction equation, the resulting complex amplitude $U(x, y, z)$ of an incoming wave of complex amplitude $A(x', y')$ (after being diffracted by the telescope aperture and focused by a lens) is given by

$$U(x, y, z) \approx \frac{e^{ikz} e^{ik(x^2+y^2)/2z}}{i\lambda z} \iint_{\text{Aperture}} A(x', y') e^{i\frac{k}{z}(x'x+y'y)} dx'dy'. \tag{1.4}$$

This transformation is proportional to the Fourier transform, which is thus one of the

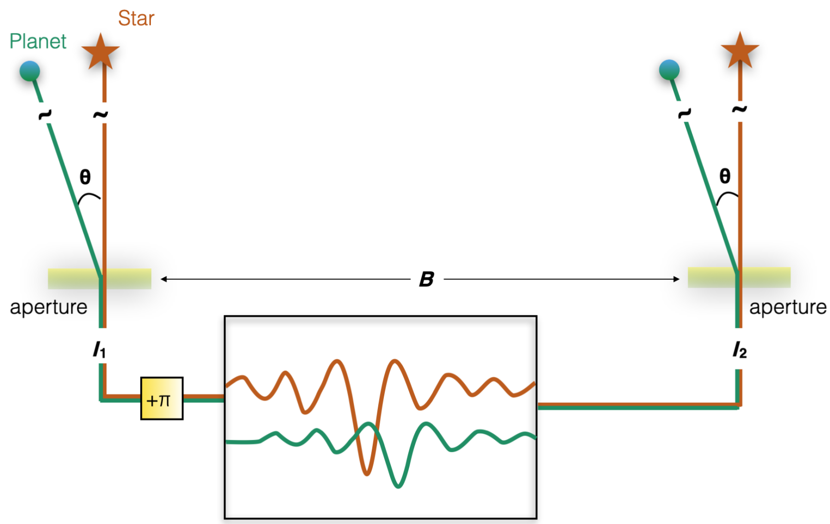

Figure 1.6: Illustration of the principle of nulling interferometry. The light of a star is collected by two apertures separated by a baseline B. A phase delay of π is introduced into one of the arms to produce a deep central minimum in the interference of the light beams. If a planet orbits the star at an angular separation of θ, its light enters the instrument off-axis introducing a further delay of $B \sin \theta$. As a consequence, the starlight is locally highly suppressed while the planet light is not. Figure and caption adapted from Lagadec et al. (2021).

primary tools for simulating, modeling, and designing coronagraphic instruments. Modern coronagraph designs that are more optimized than the classical Lyot coronagraph but remain based on a series of focal and pupil plane masks include the apodized pupil lyot coronagraph (Soummer, 2005), the four-quadrant phase mask coronagraph (D. Rouan et al., 2000), and the vortex coronagraph (D. Mawet et al., 2010).

Nulling interferometry, on the other hand, was born out of astronomical interferometry that combines the light of multiple telescopes. In 1978, Bracewell proposed that by combining the beams from two telescopes and shifting the phase of one of them by π, a transmission null would be created at the center (where the light deconstructively interferes), with an overall on-sky transmission pattern of sinusoidal fringes (see Fig. 1.6) (Bracewell, 1978). Variations of Bracewell nulling can be implemented on subapertures or subregions of single telescopes, resulting in instrument designs such as the Achromatic Interferometric Coronagraph (Gay, Rabbia, and Pierre Baudoz, 1997) or the Visible Nuller Coronagraph (Mennesson et al., 2003).

Theoretically, an instrument that manipulates the collected light can be described

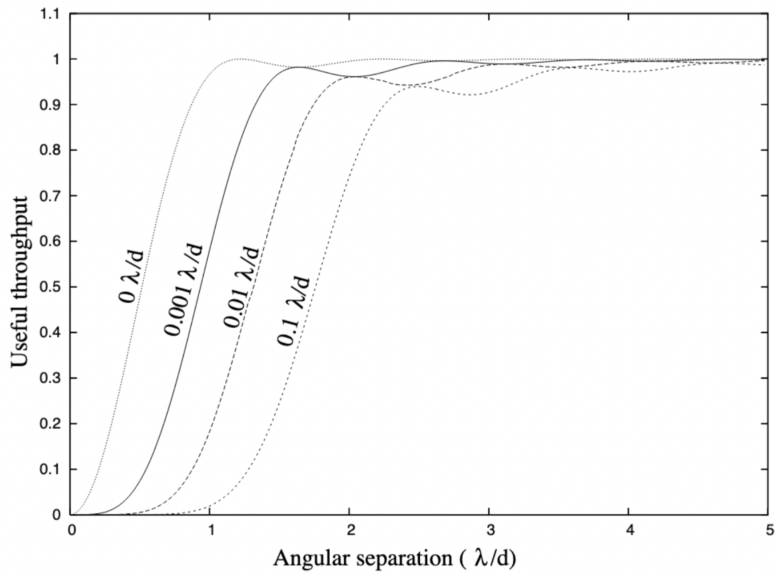

Figure 1.7: Upper limit on the off-axis throughput of a coronagraph for different stellar radii, assuming a stellar suppression level of 10^{-10}. Figure and caption adapted from O. Guyon et al. (2006a).

as a linear operator (a function of the instrument optics) acting upon the complex amplitude of the incoming wavefront. When all of the light in the system is accounted for (including light scattered away), then conservation of energy implies that the linear operator must be unitary. O. Guyon et al. (2006a) used this algebraic representation to explore the theoretical limits of stellar suppression, showing that for a given desired coronagraphic contrast (the level by which the stellar light is attenuated), a fundamental tradeoff exists between the achievable planet throughput and the angular size of the star. As shown in Fig. 1.7, a coronagraph that can reject the light of a larger star will also suffer from lower planet throughput, especially at close-in separations. O. Guyon et al. (2006a) proposes the idea of an optimal coronagraph, which uses a combination of beamsplitters to manifest the exact unitary matrix that reaches the theoretical limit for a given stellar size and planet separation. However, only in recent years have various instrument architectures and technologies (primarily based in photonics) shown promise for realizing these near-optimal unitary transformations.

Although nulling-type instruments are a subset of coronagraphs, for the rest of this work, ‘conventional coronagraphs’ will be used to refer to instruments designed to observe objects at separations $> 2\lambda/D$ from the host star, while ‘nullers’ will be used to refer to instruments designed to observe objects at separations at or less than $1\lambda/D$. While the design space itself is continuous and there is no clear boundary, this distinction is somewhat useful when discussing the classes of instruments that have been developed and implemented at real observatories.

1.4 Optical Waveguides

This section provides an overview of optical waveguides, one of the key technologies used in modern astronomical instrumentation (at optical and infrared wavelengths), and also one of the foundations for the instrumentation discussed in this work.

Light Propagation in a Medium

The following derivations for light propagation in waveguides are adapted from Paschotta (2022), *Step Index Fibers* (2025), and *Weakly Guiding Fibers* (2025).

The behavior of light propagating in a medium can be calculated using Maxwell's equations:

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial H}{\partial t}, \tag{1.5}$$

$$\nabla \times \mathbf{H} = \epsilon \frac{\partial E}{\partial t}. \tag{1.6}$$

In cylindrical coordinates, an electromagnetic wave propagating down a waveguide in the z direction takes the form of

$$\mathbf{E} = \vec{E}(r, \varphi) e^{i\beta z - i\omega t}, \tag{1.7}$$

$$\mathbf{H} = \vec{H}(r, \varphi) e^{i\beta z - i\omega t}. \tag{1.8}$$

The wave equations for each component of the electric-field $\vec{E}(r, \varphi)$ and magnetic field profile $\vec{H}(r, \varphi)$ are thus given by

$$(k^2 - \beta^2) E_r = i\beta \frac{\partial E_z}{\partial r} + i\omega\mu_0 \frac{1}{r} \frac{\partial H_z}{\partial \varphi}, \tag{1.9}$$

$$(k^2 - \beta^2) E_\varphi = i\beta \frac{1}{r} \frac{\partial E_z}{\partial \varphi} - i\omega\mu_0 \frac{\partial H_z}{\partial r}, \tag{1.10}$$

$$(k^2-\beta^2)H_r = i\beta\frac{\partial H_z}{\partial r} - i\omega\epsilon\frac{1}{r}\frac{\partial E_z}{\partial \varphi}, \tag{1.11}$$

$$(k^2-\beta^2)H_\varphi = i\beta\frac{1}{r}\frac{\partial H_z}{\partial \varphi} + i\omega\epsilon\frac{\partial E_z}{\partial r}, \tag{1.12}$$

$$\frac{\partial^2 E_z}{\partial r^2} + \frac{1}{r}\frac{\partial E_z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 E_z}{\partial \varphi^2} + (k^2-\beta^2)E_z = 0, \tag{1.13}$$

$$\frac{\partial^2 H_z}{\partial r^2} + \frac{1}{r}\frac{\partial H_z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 H_z}{\partial \varphi^2} + (k^2-\beta^2)H_z = 0. \tag{1.14}$$

Here, $k = 2\pi n/\lambda$ is the (potentially spatially varying) wavenumber, where n is the (potentially spatially varying) index of refraction of the medium, and λ the vacuum wavelength of the light. The solutions to these equations for a given waveguide geometry are the modes that can be supported by that waveguide.

Optical Fibers

A simple step-index optical fiber consists of a cylindrical core of a constant index of refraction n_{core} surrounded by a cladding, also of constant index of refraction n_{cl}. Typically, the difference in the indices of refraction is small ($(n_{\text{core}} - n_{\text{cl}})/n_{\text{core}} \ll 1$), such that the weakly guiding approximation can be used. In this case, the wave equations are significantly simplified and many of the modes become degenerate. The complex electric field profiles of the resulting supported modes (known as the linearly polarized modes, or LP modes) are the solutions to

$$\frac{\partial^2 E}{\partial r^2} + \frac{1}{r}\frac{\partial E}{\partial r} + \frac{1}{r^2}\frac{\partial^2 E}{\partial \varphi^2} + (k^2-\beta^2)E = 0. \tag{1.15}$$

Because the index of refraction within the material is constant and the geometry cylindrically symmetric, the solution will take the separable form of

$$E(r,\varphi) = F_{lm}(r)(a\sin(il\varphi) + b\cos(il\varphi)). \tag{1.16}$$

Plugging this into Eq. 1.15, performing a separation of variables, and multiplying through by r^2 yields a radial equation of

$$r^2 F''_{lm}(r) + rF'_{lm}(r) + (r^2k^2 - l^2 - r^2\beta^2)F_{lm} = 0. \tag{1.17}$$

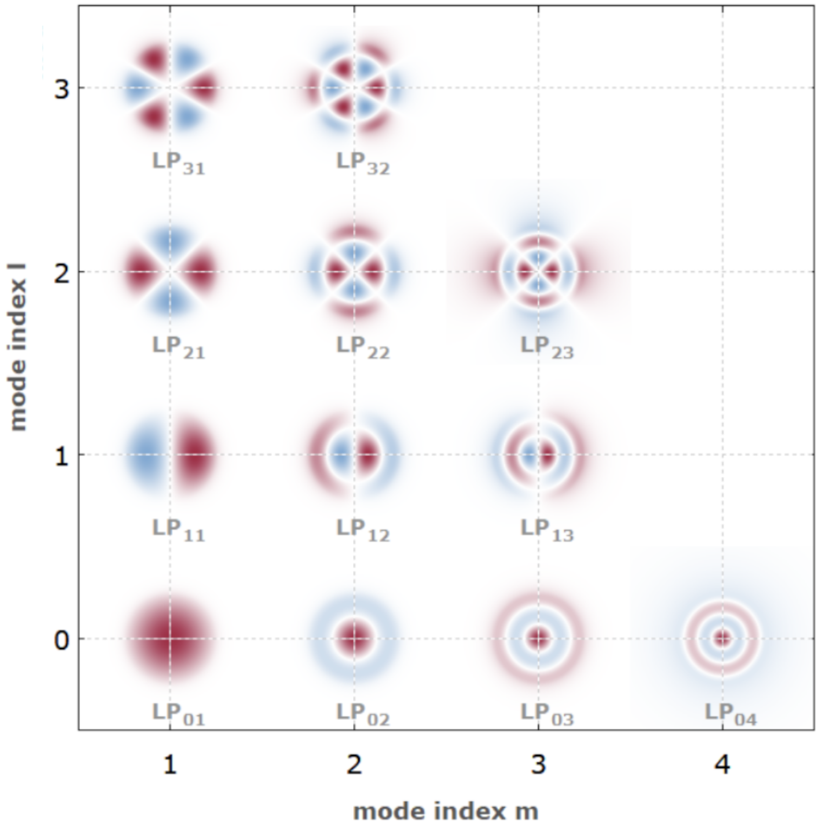

Figure 1.8: Select electric field amplitude profiles (LP modes) for the guided modes of a step-index fiber. Figure and caption adapted from Paschotta (2022).

The solutions to this differential equation are given by: Bessel functions $J_l(ur/r_{\text{core}})$ for the core (where $u = r_{\text{core}}\sqrt{n_{\text{core}}^2 k^2 - \beta^2}$), and modified Bessel functions $K_l(wr/r_{\text{core}})$ for the cladding, (where $w = r_{\text{core}}\sqrt{\beta^2 - n_{\text{cl}}^2 k^2}$). Meanwhile, the azimuthal component of the solution are the sines and cosines indexed by the mode order l. Figure 1.8 shows the solutions for the core, i.e. the LP modes, across a range of azimuthal mode index l and radial mode order m.

The boundary condition for continuity across the core-cladding interface gives

$$u^2 + w^2 = r_{\text{core}}^2(n_{\text{core}}^2 - n_{\text{cl}}^2)k^2. \tag{1.18}$$

The quantity $\sqrt{n_{\text{core}}^2 - n_{\text{cl}}^2}$ is often denoted as the numerical aperture, or NA, which is also related to the maximum angle of incidence of light that the fiber can accept by $\text{NA} = n_{\text{in}} \sin\theta_{\text{max}}$, where n_{in} is the index of refraction of medium through which the input beam is traveling.

Step-index fibers can be characterized by a dimensionless quantity known as the V number, or $V = (2\pi/\lambda) r_{\text{core}}\text{NA}$, which determines the number of modes that the fiber can support. At V numbers of less than $V_c \approx 2.405$, a fiber can support only the fundamental mode (LP 01), while fibers with higher V numbers can support a larger number of modes (which can be estimated as approximately $V^2/4$). For a given fiber, the cutoff wavelength λ_c below which the fiber is no longer single-moded can also be calculated as $\lambda_c = \frac{2\pi}{V_c} r_{\text{core}}\text{NA}$.

A Gaussian profile is often used to approximate the LP 01 mode of a single-mode fiber (SMF):

$$\Psi(r) = \sqrt{\frac{2}{\pi a^2}} e^{-(r/a)^2}. \tag{1.19}$$

The quantity $2a$ is referred to as the mode field diameter (MFD), which for a given fiber varies with wavelength as approximately (Jeunhomme, 1989)

$$\text{MFD} = 2r_{\text{core}}\big(0.65 + 0.434(\frac{\lambda}{\lambda_c})^{3/2} + 0.01419(\frac{\lambda}{\lambda_c})^6\big). \tag{1.20}$$

SMFs and their modal filtering have found many applications in astronomical instrumentation as a way to route light across (potentially long) distances while maintaining modal purity, e.g. for beam combination over long distances or to maintain a stable line-spread profile for spectroscopy. Of particular relevance to this work are fiber nullers such as the Palomar Fiber Nuller (Haguenauer and E. Serabyn, 2006), which isolates two subpupils of a telescope and applies a π phase shift across one of them before injecting the beam into an SMF. For an on-axis source, the two beams destructively interfere such that no light couples into the fiber, but some of the light of an off-axis source would couple in depending on its location on-sky. The light in the fiber can then be routed to a detector or spectrograph to be analyzed. Figure 1.9 shows variations on the fiber nuller phase mask, in analogy with their discrete telescope counterparts. The Vortex Fiber Nuller design (Ruane, Ji Wang, et al., 2018) in particular is already implemented as part of the Keck Planet Imager and Characterizer (KPIC) instrument at the Keck II Telescope, with first detections demonstrated in 2024 (Echeverri, J. W. Xuan, et al., 2024) and a currently ongoing science campaign.

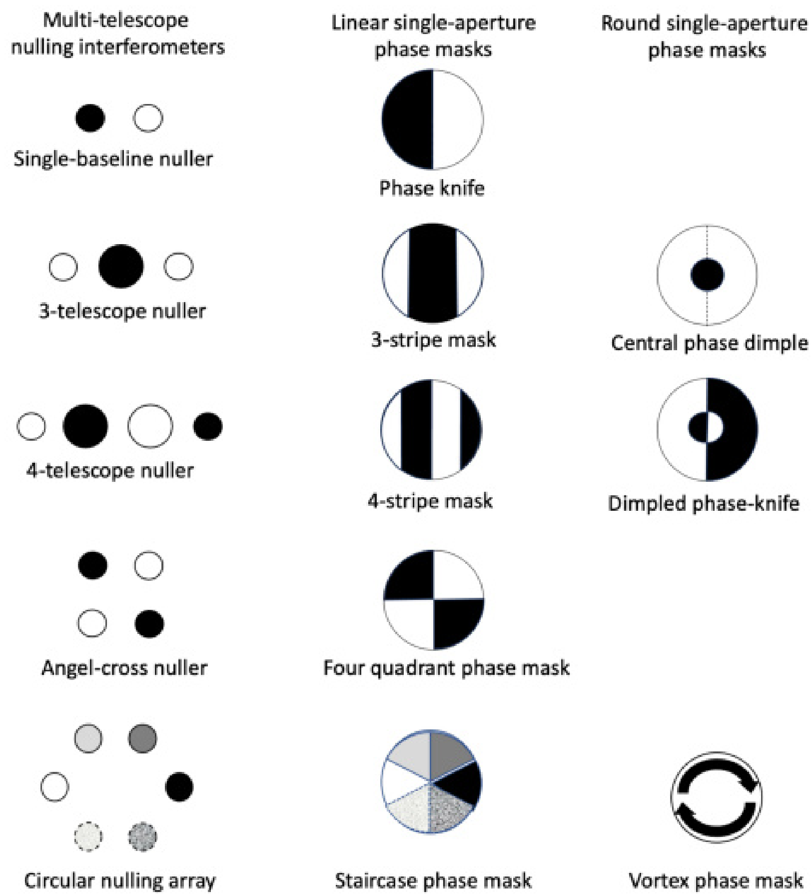

Figure 1.9: Comparison of multi-aperture nulling interferometer configurations (left column) with single-aperture nulling phase masks (central column). Phases of 0 and π are shown as white and black, respectively. Intermediate phases in the bottom row are shown in various shades of gray. Each small circle in the left column represents a separate telescope aperture. Central column: single-aperture “linear" nulling phase masks, arranged to show their topological correspondence to the separated-aperture nulling cases to their left. Right column: “Round" single-aperture phase masks (in which at least one of the phase transitions is circular) that also topologically correspond to the leftmost entries in their rows. Figure and caption adapted from Eugene Serabyn, Liewer, and Ruane (2024).

Photonic Lanterns

A waveguide somewhat more complex than optical fibers is the photonic lantern (PL), a device of key interest to this work. Intuitively, a photonic lantern is a bundle of SMF cores embedded in a shared cladding, the entirety of which has been stretched and tapered on one end (see Fig. 1.10a for a schematic of a 2-port lantern). This results in an adiabatic transition in the behavior of the waveguide as a bundle of discrete SMFs on one end into a multi-mode fiber on the other end, where the

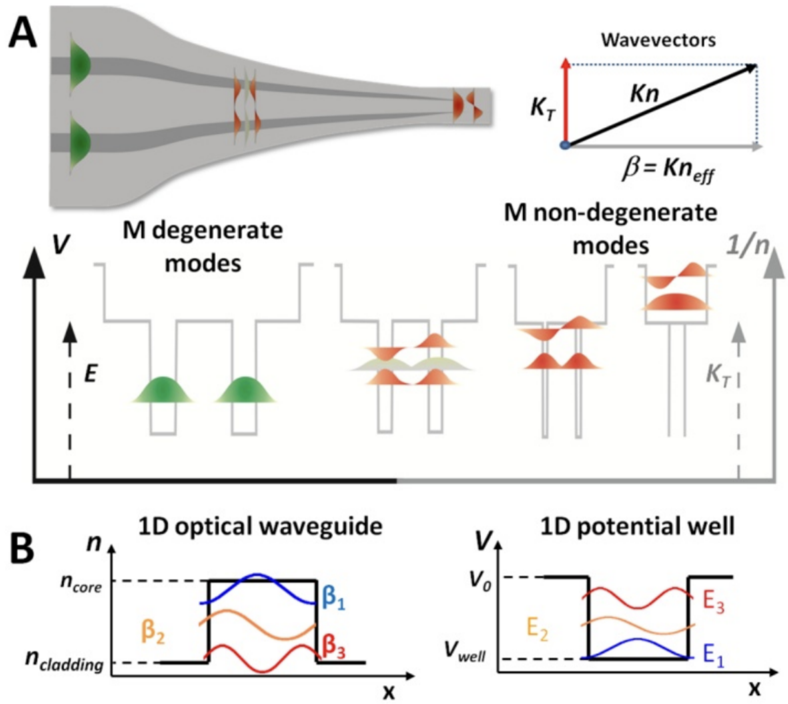

Figure 1.10: A) Schematics of a 2-port photonic lantern, and the Kronig-Penney model analogy for the transition between the single-mode and the multi-mode regime. B) Schematics of a one dimensional waveguide (left) and a one dimensional quantum well (right). Figure and caption adapted from Leon-Saval, Argyros, and Bland-Hawthorn (2013a).

original cladding of the bundle now forms the propagation medium.

Although there is no analytic solution for light propagation through a photonic lantern, an analogy for the behavior of the supported modes can be found in the Kronig-Penney model for particles in a potential well; as shown in Fig. 1.10b, the modes supported in a waveguide are mathematically analogous to the states supported by a quantum potential well, with the quantity $1/n$ (where n is the index of refraction) corresponding to the potential V, and the transverse wavevector k_T of the waveguide mode (see Fig. 1.10a for the definition of k_T relative to β and k) corresponding to the energy E of the electron state (Leon-Saval, Argyros, and Bland-Hawthorn, 2013a). By analogy, the single-mode end corresponds to an array of separate potential wells each supporting one electron state, which are gradually squeezed together such that the states become less localized, until the final states on the multi-mode end are the states supported by the broader potential well of the cladding material. The effective behavior is that when light enters the array of SMFs

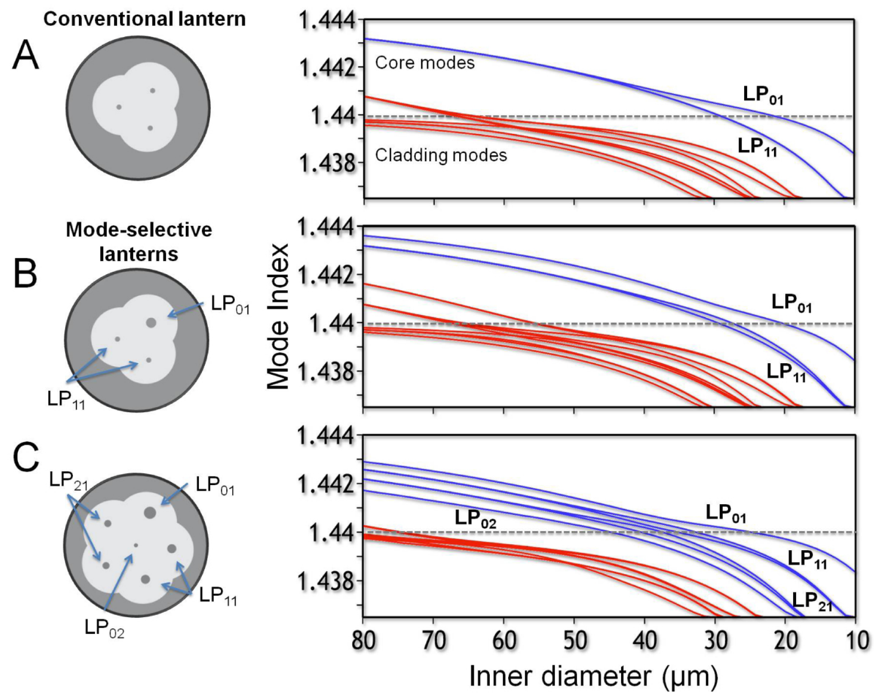

Figure 1.11: Modal analysis of A) a conventional 3-SMF photonic lantern, B) a mode-selective 3-SMF photonic lantern, and C) a mode-selective 6-SMF photonic lantern. (Left A, B, and C) Schematics of FMF end of the modeled photonic lanterns showing the different cores sizes corresponding to the similar/dissimilar fibers. Figure and caption adapted from Leon-Saval, Fontaine, et al. (2014).

on one end, the photonic lantern multiplexes them together into a multi-mode output. When used in reverse, light that enters the multi-mode end gets demultiplexed into separate SMFs that each correspond to a specific mode profile.

When fiber cores with different properties are used to construct the lantern (as shown in Fig. 1.11), the degeneracies in the transition region are broken, the modes retain distinct values of β without mixing, and each SMF thus maps to a specific LP mode; this type of lantern is called a mode-selective photonic lantern (MSPL) (Leon-Saval, Fontaine, et al., 2014). Hybrid lanterns, in which certain modes are isolated (such as the LP 01 mode) but the rest are mixed are also possible.

Several astronomical applications of photonic lanterns have been proposed and demonstrated, such as for focal-plane wavefront sensing (Lin, Fitzgerald, et al., 2022a; Norris, Wei, et al., 2020), interferometric imaging (Kim, Fitzgerald, Lin, Sallum, et al., 2024), and spectroastrometry (Kim, Fitzgerald, Lin, Xin, et al., 2024). The majority of the rest of this thesis (Chapters 2-5) focuses on using a mode-selective photonic lantern for nulling applications. Chapters 6 and 7 concern other advances in high-contrast imaging instrumentation and data analysis that can also be applied to photonic lantern nulling.

Chapter 2

EFFICIENT DETECTION AND CHARACTERIZATION OF EXOPLANETS WITHIN THE DIFFRACTION LIMIT: NULLING WITH A MODE-SELECTIVE PHOTONIC LANTERN

Xin, Yinzi et al. (Oct. 2022). "Efficient Detection and Characterization of Exoplanets within the Diffraction Limit: Nulling with a Mode-selective Photonic Lantern". In: *The Astrophysical Journal* 938.2, 140, p. 140. DOI: 10.3847/1538-4357/ac9284. arXiv: 2209.07644 [astro-ph.IM].

This chapter is a reproduction of my first paper on using photonic lanterns for nulling, which introduced the concept behind the instrument, explained the theory, and presented simulations that characterized the properties of the photonic lantern nuller (PLN) with respect to various design parameters (such as mode count) and external factors (such as wavefront aberrations and tip-tilt jitter). I showed that relative to designs such as the Vortex Fiber Nuller, the PLN achieved higher off-axis throughput and also provided enough spatial information to localize a hypothetical companion, albeit with a 180 degree degeneracy due to symmetries in the instrument. The PLN then became a significant part of the rest of my thesis work, with several future chapters dedicated to laboratory and on-sky demonstrations.

Abstract

Coronagraphs allow for faint off-axis exoplanets to be observed, but are limited to angular separations greater than a few beam widths. Accessing closer-in separations would greatly increase the expected number of detectable planets, which scales inversely with the inner working angle. The vortex fiber nuller (VFN) is an instrument concept designed to characterize exoplanets within a single beam width. It requires few optical elements and is compatible with many coronagraph designs as a complementary characterization tool. However, the peak throughput for planet light is limited to about 20%, and the measurement places poor constraints on the planet location and flux ratio. We propose to augment the VFN design by replacing its single-mode fiber with a six-port mode-selective photonic lantern, retaining the original functionality while providing several additional ports that reject starlight but couple planet light. We show that the photonic lantern can also be used as a

nuller without a vortex. We present monochromatic simulations characterizing the response of the photonic lantern nuller (PLN) to astrophysical signals and wavefront errors, and show that combining exoplanet flux from the nulled ports significantly increases the overall throughput of the instrument. We show using synthetically generated data that the PLN detects exoplanets more effectively than the VFN. Furthermore, with the PLN, the exoplanet can be partially localized, and its flux ratio constrained. The PLN has the potential to be a powerful characterization tool complementary to traditional coronagraphs in future high-contrast instruments.

2.1 Introduction

While conventional coronagraphs dramatically reduce the photon noise from the star, they are practically limited to angular separations greater than a few λ/D (the size of a resolution element, where λ is the wavelength and D the telescope diameter). The ability to access closer-in exoplanets would greatly increase the expected yield of detectable planets, since yield scales approximately inversely with the inner working angle (IWA), with yield $\propto$ IWA $^{-0.98}$ (Stark et al., 2015). Additionally, planets observable with coronagraphy in the visible and near-infrared regime may fall within the inaccessible inner working angle at longer wavelengths, where features of key biosignatures such as carbon monoxide and methane exist. Gaining access to closer separations at those longer wavelengths will thus enable better characterization of planets detected.

Meanwhile, techniques such as nonredundant masking interferometry (P. G. Tuthill, Monnier, and Danchi, 2000) or cross-aperture nulling interferometry (Bracewell, 1978; E. Serabyn et al., 2019) can access very small angular separations. However, these approaches result in lower efficiency than coronagraphy since only a small portion of the aperture is used. The Vortex Fiber Nuller (VFN) is an instrument concept that straddles the space between the two approaches, with a smaller IWA than coronagraphs but more efficient at routing the planet light to a diffraction-limited spectrograph than single-baseline cross-aperture interferometry (Ruane, Ji Wang, et al., 2018). This technique is capable of characterizing exoplanets within 1 λ/D, requires few optical elements, and is compatible with many coronagraph designs as a complementary characterization tool.

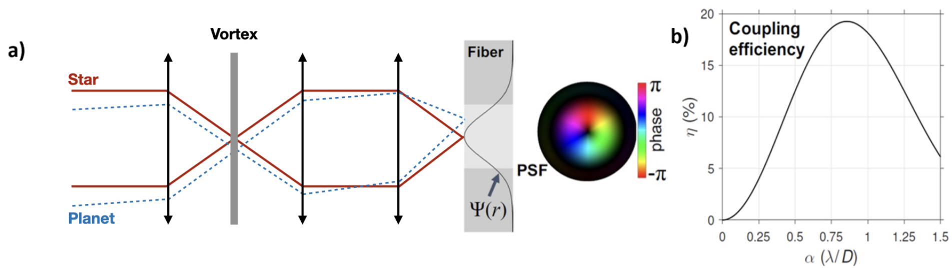

Figure 2.1: (a) Schematic of a focal-plane VFN with a single-mode fiber. The beam is focused onto a vortex mask, which imparts a different phase pattern on the star and planet point-spread-functions. The beam is then collimated and refocused onto a single-mode fiber. The on-axis star light rejected while the planet light gets partially coupled. (b) Coupling efficiency, η, or throughput, of a planet as a function of its angular separation from the star.

2.2 Photonic Lantern Nuller Concept

Vortex Fiber Nulling

The Vortex Fiber Nuller is an instrument concept that enables spectroscopy of exoplanets within 1 λ/D, using a vortex mask to generate a vortex phase pattern on the incoming beam (Ruane, Ji Wang, et al., 2018). Figure 2.1(a) shows that when the beam is on-axis (such as light from a star), the resulting pattern is orthogonal to the fundamental mode of a single-mode fiber (SMF) and does not couple to it. This result can be demonstrated by calculating the coupling efficiency of a field $f(r, \theta)$ with the SMF mode $\psi_{01}(r)$:

$$\int \psi_{01}(r) f(r, \theta) dA. \tag{2.1}$$

For the field created by a vortex, the integral is separable, and the polar term is given by

$$\int_0^{2\pi} \exp(il\theta) d\theta, \tag{2.2}$$

where l is an integer that denotes the vortex charge. This integral evaluates to 0 for $l \neq 0$, reflecting that the vortex field is orthogonal to the SMF mode.

However, as shown in Fig. 2.1(b), off-axis planet light from $\sim 0.5\lambda/D$ to $\sim 1.3\lambda/D$ can couple in, with a peak throughput of 19% at 0.9 λ/D. The coupled planet light can thus be directed to a spectrograph for immediate characterization, while the starlight is rejected. A focal-plane VFN is explored in this work, but Ruane, Echeverri, et al. (2019) showed that the vortex can also be placed in the pupil plane, resulting in a pupil-plane VFN that operates on the same principle of rejecting on-axis starlight with an imprinted vortex.

The range of angular separations probed by the VFN is smaller than the inner working angle of all classical coronagraphs, and is a region known to harbor potentially habitable exoplanets detected via radial velocity (RV) and transit methods. Additional advantages of the VFN compared to classical coronagraphs include its relative insensitivity to telescope aperture shape, polarization aberrations, and many wavefront aberration modes (Ruane, Ji Wang, et al., 2018). Since its conceptual development, the VFN concept has been tested in the lab, achieving azimuthally averaged peak coupling of 16% (close to the theoretical limit) and starlight suppression of 6×10^{-5}, which can be attributed to the minor wavefront errors in the system (Monochromatic; Broadband, Echeverri, Ruane, Nemanja Jovanovic, Dimitri Mawet, et al., 2019; Echeverri, Ruane, Nemanja Jovanovic, Hayama, et al., 2019).

While the original VFN design is already compelling, it has several drawbacks. The planet throughput is relatively low, with a theoretical limit of $\sim$ 20%, depending on the configuration. The measurement from a VFN also lacks spatial information — since the coupling map is circularly symmetric, there is no way to determine from the data the position angle of the planet, information that is (in the absence of other measurements) necessary for constraining the orbital parameters of the planet. Since there is only one flux measurement and the coupling into the SMF varies with the radial separation of the planet, there is also a degeneracy between the planet flux and its separation. Here, we present an augmentation to the VFN that enhances throughput and provides additional constraints on the orbit and flux of the planet, while retaining the functionality of the VFN concept. This new design relies on a device called the mode-selective photonic lantern.

Mode-Selective Photonic Lanterns

A photonic lantern is a photonic mode converter that adiabatically interfaces between a multi-mode port and several single-mode ports, where the distribution of flux in the single-mode outputs is related to the power in each mode at the multi-mode input (Leon-Saval, Argyros, and Bland-Hawthorn, 2013b). Photonic lanterns have been proposed for use in astrophysics for spectrometer coupling (Lin, Nemanja

Jovanovic, and Fitzgerald, 2021) and for focal-plane wavefront sensing, allowing for the measurement of the input wavefront while maintaining single-mode fiber outputs suited for injection into spectrographs for spectral characterization (N. Jovanovic, Schwab, et al., 2016; Corrigan et al., 2018; Norris, Wei, et al., 2020). Each mode at the few-mode fiber (FMF) face of the lantern is mapped to a SMF output, such that light coupling to a given mode at the FMF side will result in flux in the corresponding SMF core. The device is bi-directional, so light injected into one of the SMF ports will propagate into the mode corresponding to that port at the FMF face.

While standard photonic lanterns have similar cores and are not designed with a particular mode structure in mind, mode-selective photonic lanterns (MSPL, Leon-Saval, Fontaine, et al., 2014) utilize dissimilar cores that enable ports to be mapped into LP modes. A partially mode-selective photonic lantern has one port corresponding to the LP 01 mode, while the rest of the ports exhibit an unspecified structure. In a fully mode-selective photonic lantern, all ports correspond to LP modes. Figure 2.2 shows a schematic of a six-port MSPL based on the design from Leon-Saval, Fontaine, et al. (2014), where each port corresponds to one of the first six LP modes.

To synergize the action of the VFN with symmetry properties of the LP modes, we propose to replace the single-mode fiber of the original VFN with a MSPL, resulting in a Photonic Lantern Nuller (PLN) instrument concept that improves upon the original design.

VFN with a Mode-Selective Photonic Lantern

The PLN replaces the single-mode fiber of the VFN by a MSPL as described in Section 2.2. Specifically, the light after the vortex mask is focused onto the FMF face of the MSPL and propagates through to the single-mode outputs. Each output port can then be coupled into individual SMFs and routed to photodetectors or spectrographs. The port corresponding to the LP 01 mode provides the same response as the VFN, where on-axis light is nulled while off-axis light can couple. Additionally, if we label the LP mode azimuthal order by m' analogously to the Zernike polynomials, i.e. positive m' indicating an azimuthal component of $\cos(m'\theta)$ and negative m' indicating $\sin(m'\theta)$, then, a photonic lantern port combined with an optical vortex with azimuthal charge l, will result in an on-axis null *except* when $l \pm m' = 0$. This result can be derived by extending Equation 2.1 to an arbitrary fiber mode $\psi_{n'm'}$, and separating out the polar integral:

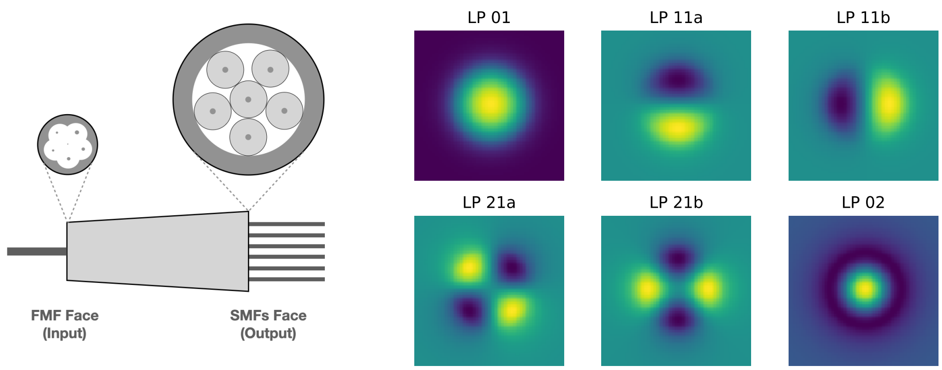

Figure 2.2: Left: Schematics of a six-port mode-selective photonic lantern spatial-multiplexer fiber system. Each LP mode at the few mode fiber (FMF) face is mapped to one of the six single-mode ports of the SMF face, such that light with an LP mode shape at the FMF side will result in flux in the corresponding SMF core. The device is bi-directional, so light injected into one of the SMF ports will propagate into the LP mode corresponding to that port at the FMF face. Right: The field amplitudes of the first six LP modes, corresponding to the ideal modes of six-port MSPL.

$$\begin{aligned} &\int_0^{2\pi} \exp(il\theta)\cos(m'\theta)d\theta, \quad m' \geq 0, \quad \text{or} \\ &\int_0^{2\pi} \exp(il\theta)\sin(m'\theta)d\theta, \quad m' < 0. \end{aligned} \tag{2.3}$$

Recalling the exponential trigonometric identities $\cos(x) = (e^{ix} + e^{-ix})/2$ and $\sin(x) = (e^{ix} - e^{-ix})/(2i)$, we find that these overlap integrals evaluate to 0 for $l \pm m' \neq 0$. Thus, on-axis nulls are created in multiple ports, from which planet spectra can be extracted. Additionally, the existence of ports with $m' \neq 0$ allows for a nuller configuration with no vortex at all, as the overlap integrals for the LP11ab and LP21ab ports evaluate to zero when $l = 0$. This means that the photonic lantern can be used by itself as a nuller, as contemporaneously presented in P. Tuthill (2022).

To demonstrate these properties, we simulate the PLN configurations using HCIPy (Por, S. Y. Haffert, et al., 2018). Our optical propagation model propagates the desired input wavefront through a circular pupil (with λ/D chosen to equal 1), and then into a focal plane. For the configuration without a vortex, this becomes the final focal-plane electric field. For the configurations with a vortex, either a charge 1 or 2 vortex is applied in the focal plane. As with the VFN, a vortex with charge higher than 2 results in lower peak throughput and larger IWA, so we do not focus on them in this work.

The square of the overlap integral of the focal-plane electric field distribution with each LP mode gives the relative intensity coupled into the corresponding port. We explore using an MSPL with six LP modes and a V number equal to 4.71. Our simulations assume perfect mode shapes as well as perfect transitions, free from cross-coupling and losses. Characterizing the impact of these real-world imperfections, from realistic designs as well as from fabrication errors, is left for future work.

Given wavelength, the optimal coupling into the lantern depends on the the mode field diameter (MFD) of the lantern modes and the focal ratio $F\#$ (Ruane, Echeverri, et al., 2019). While the real MFDs of photonic lanterns are tunable within a small range (Leon-Saval, private communication), in practice, the coupling in a real system will be optimized by changing the focal ratio. However, since our simulations already set $\lambda/D = 1$ and $F\# = 1$, we optimize coupling by tuning the MFD (expressed in units of λ/D). Specifically, for each configuration (no vortex, $l = 1$, $l = 2$), we simulate a range of MFDs and find the value that maximizes the peak of the x-axis cross-section of the summed throughput of the nulled ports. Although Section 2.4 shows that summed throughput does not fully predict instrument performance, it is still a useful proxy for choosing the MFD, as optimizing directly for detection capability would require knowledge of the level and distribution of on-sky wavefront error, which is not predictable *a priori*.

From our simulations, we find that the optimal MFD is 2.8 λ/D for the no vortex and charge 1 cases, and 3.2 λ/D for the charge 2 case. We present the results of our simulations using these diameters. Figure 2.3 shows the ideal spatial coupling efficiency for a point source as a function of angular separation from the optical axis, or coupling map, for every port (top panels) along with the line profile along the horizontal axis (bottom panels). We also plot the total flux collected across all ports (dashed pink lines) as well as the total flux collected from only the nulled ports satisfying $l \pm m' \neq 0$ (solid black lines). The total nulled throughput curves demonstrate that the additional ports increase both the peak throughput as well as the field of view for which planet light couples.

While MSPLs with more than six ports can in theory be fabricated, manufacturing MSPLs with large numbers of modes remains a practical challenge because the adiabaticity of the lantern transition becomes more difficult to achieve as the number of

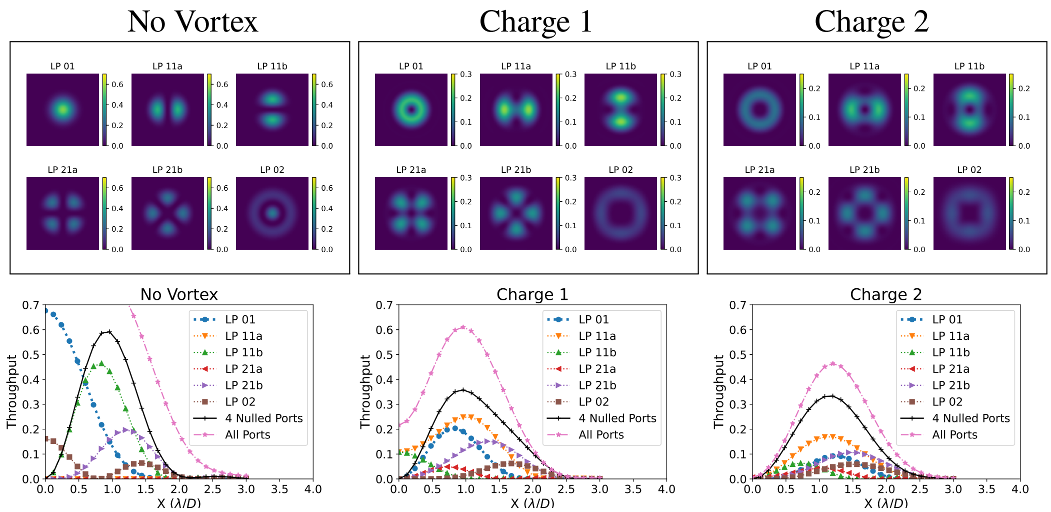

Figure 2.3: Coupling maps for each port with no vortex (top left), and a charge 1 (top middle) and charge 2 (top right) vortex. The maps span -3 λ/D to 3 λ/D in each direction. Bottom left: Throughput line profiles with no vortex. The four nulled ports satisfying $l \pm m' \neq 0$ are LP 11ab and LP 21ab. Bottom middle: Throughput line profiles with a charge 1 vortex. The four nulled ports satisfying $l \pm m' \neq 0$ are LP 01, LP 21ab, and LP02. Bottom right: Throughput line profiles for each port with a charge 2 vortex. The four nulled ports satisfying $l \pm m' \neq 0$ are LP 01, LP 11ab, and LP02. Although nulls in the LP 21ab ports are not guaranteed by symmetry, in this case, their central throughputs are spuriously low, and including them in the data analysis may provide some additional gains.

modes increases (Velázquez-Benítez et al., 2018). While larger port numbers may become available with the advancement of photonics technology, Figure 2.4 shows that increasing the total number of ports brings diminishing returns in throughput, especially at angular separations $< \lambda/D$. In addition, using fewer ports has the advantage that it requires fewer detector pixels, which are always at a cost premium. Considering these factors, and that MSPLs with more than six ports are not readily manufacturable with current photonics technology, we choose to focus our investigations on a PLN design with a six-port MSPL.

2.3 Sensitivity to Aberrations

Zernike Aberrations

One benefit of the original VFN was its insensitivity to many low order Zernike wavefront error modes. If the charge of the vortex is denoted by l, and the Zernike aberrations are denoted by $Z_n^m(r, \theta)$, where n is the radial order and m indicating the azimuthal structure, i.e. $\cos(m\theta)$ for positive m and $\sin(m\theta)$ for negative m, then

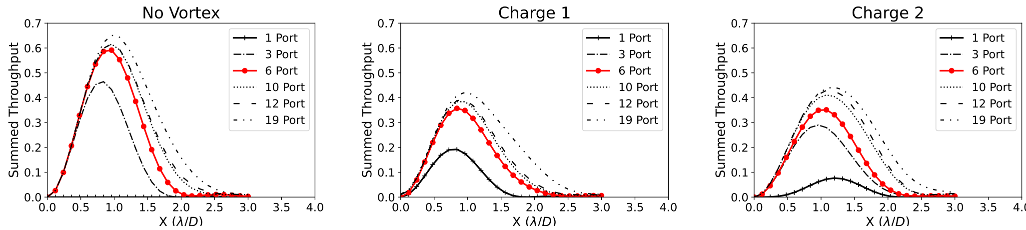

Figure 2.4: Line profiles for summed throughput of nulled ports for PLNs with no vortex (left), a charge 1 vortex (middle) and a charge 2 vortex (right), using MSPLs with varying numbers of output ports. As the number of ports increases, each additional port brings decreasing returns in additional throughput. The current limit of what can be practically manufactured is six ports. Thus, we choose to use a six-port MSPL in our PLN design, which balances the total throughput of the nulled ports with what is practically manufacturable. Note that a higher V number of 8.48 was necessary to generate up to 19 LP modes. Here, we wish to compare the effect of port number independently of V number effects, so fix the V number at 8.48 for all port numbers. Thus, due to the difference in V number, the line profiles shown in this analysis have slightly different shapes from those in Figure 2.3.

only aberrations that cancel out the vortex charge ($l \pm m = 0$) will couple. This can be demonstrated analogously to the case of LP modes, replacing the m' of a given port in Equation 2.3 with the m of a given Zernike mode.

The additional photonic lantern ports obey a similar principle, but the structure of the LP mode and the Zernike mode will interact, and the polar overlap integral is now given by

$$\begin{aligned}
&\int_0^{2\pi} \exp(il\theta)\cos(m'\theta)\cos(m\theta)d\theta, \quad m', m \geq 0, \quad \text{or} \\
&\int_0^{2\pi} \exp(il\theta)\cos(m'\theta)\sin(m\theta)d\theta, \quad m' \geq 0, m < 0, \quad \text{or} \\
&\int_0^{2\pi} \exp(il\theta)\sin(m'\theta)\cos(m\theta)d\theta, \quad m' < 0, m \geq 0, \quad \text{or} \\
&\int_0^{2\pi} \exp(il\theta)\sin(m'\theta)\sin(m\theta)d\theta, \quad m', m < 0.
\end{aligned} \tag{2.4}$$

Thus, for each port, only aberrations satisfying $l \pm (m' + m) = 0$ will couple (to first order). Figure 2.5 shows the simulated stellar coupling, η_s, as a function of the input amplitude of the first ten Zernike aberrations. In this work, we compute coupling normalized to the summed intensity of the beam, such that the stellar coupling is equivalent to the null-depth. The fact that the LP 01 port is sensitive primarily to tip, tilt, and coma (for charge 1) and astigmatism followed by second-order responses to tip and tilt (for charge 2) is consistent with theoretical predictions as well as the numerical simulations presented in Ruane, Echeverri, et al. (2019). The results for the other ports show that, as predicted by the azimuthal order conditions, each port is only sensitive to a few specific lower-order aberrations satisfying $l \pm (m' + m) = 0$. For example, the LP 21ab ports with a charge 1 vortex and the LP 11ab ports with a charge 2 vortex are all insensitive to defocus ($m = 0$) and astigmatism ($m = \pm 2$). The LP 02 ports have the same azimuthal order as the corresponding LP 01 ports, and thus reject the same low-order aberrations.

Tip-tilt Jitter

Ruane, Echeverri, et al. (2019) predicted that for ground-based observatories, tip-tilt jitter (evolving much faster than the typical exposure times) will likely be a significant contribution to degradation of the VFN's null-depth. We thus present simulations of average null-depth achieved (η_s) as a function of the standard deviation of tip-tilt jitter (σ_{tt}). For each data point, 100 independent realizations of tip-tilt are generated, with amplitude drawn from a normal distribution with standard deviation σ_{tt} and position angle drawn uniformly between 0 and 2π. The 100 frames are then averaged to calculate an averaged η_s. The results are presented in Figure 2.6. For example, to achieve a null depth of 10^{-3} in the LP11ab ports of the no vortex PLN, the standard deviation of tip-tilt jitter must be smaller than $\sim 0.1\lambda/D$. To achieve a null depth of 10^{-3} in the LP01 port of the charge 1 and charge 2 configurations, the standard deviation of tip-tilt jitter must be smaller than $\sim 0.1\lambda/D$ and $\sim 0.3\lambda/D$, respectively. For context, the Keck Planet Imager and Characterizer (KPIC) instrument at the Keck II telescope, a fiber injection unit for high resolution spectroscopy that currently has an VFN mode as well as the capability to test a future PLN on-sky, typically achieves on-sky jitter standard deviations of 6-7 mas, corresponding to 0.14 waves at 2.2 μm (Delorme et al., 2021a).

KPIC Atmospheric Residuals

We also simulate the performance of the PLN under WFE conditions measured by the pyramid wavefront sensor (PyWFS) of KPIC. The atmospheric seeing the night the data was taken was 0.6 arcsec, and the wavefront sensor achieved residuals of 150 nm RMS. It should be noted that the PyWFS does not see all of the errors in the optical system, as recent on-sky demonstrations of the VFN on KPIC (Echeverri et al, in prep) do not achieve the level of starlight suppression predicted by these residuals alone. Specifically, in the real KPIC instrument, there is additional tip-tilt error downstream of the PyWFS that is not captured in these simulations. Thus, these simulations should be interpreted as an optimistic limit, while the real performance will be impacted by additional errors invisible to the PyWFS.

For our simulation, we take 590 frames of measured wavefront error, expressed in the form of reconstructed Zernike coefficients. From each frame of coefficients, we generate a pupil plane WFE map. As an intermediate diagnostic, we calculate the focal-plane image PSF averaged over these frames, compared it to an ideal PSF with no WFE in Figure 2.7.

For our simulation, we propagate an on-axis beam with that WFE through our PLN models to calculate the output null depths. We also propagate off-axis beams with each frame of WFE (at $0.84\lambda/D$ for no vortex and charge 1 configurations and $1.3\lambda/D$ for charge 2 configuration). Figure 2.8 shows the mean coupling over all the frames. In the nulled ports of the PLN, the mean off-axis planet coupling over these frames (where it is expected based on the coupling maps) remains significantly higher than the stellar coupling in the presence of this WFE.

2.4 Simulation of Exoplanet Characterization

In this section, we demonstrate the exoplanet detection and characterization capabilities of a PLN and compare it to those of the VFN.

Synthetic Data Generation

We consider the outputs of the instrument to be the intensity at the single simulated wavelength in each port. In reality, the light in each port can be fed in to a spectrograph, and spectral analysis can be used to increase detectability by orders of magnitude (Ji Wang et al., 2017). However, we neglect spectral information in this preliminary demonstration of the PLN performance relative to the VFN, and leave exploring the combination of a broadband PLN and spectral analysis to future work.

We assume that the integration time of an observation is significantly longer than the coherence time of atmospheric residuals, such that fluctuations in wavefront error will average out to the null depth. Consequently, we assume that the primary contribution to non-static noise is photon noise.

The following process was used to generate the synthetic data. We first average the 590 intensity frames from the simulation of KPIC PyWFS residuals in Section

2.3 to obtain the average null depth. To generate realizations of photon noise, we calculate the stellar photon rate entering the instrument:

$$\mathrm{PR} = f_0 \times 10^{-m/2.5} \times A \times \Delta\lambda \times \eta_t, \tag{2.5}$$

where $f_0 = 9.56\times10^9$ photons m^{-2} s^{-1} μm^{-1} is the zero point number corresponding to the photon flux per unit wavelength of a magnitude zero star in H band, m is the stellar magnitude, A the telescope area, $\Delta\lambda$ the bandwidth, and η_t the throughput of the telescope before reaching the PLN instrument. We choose the stellar magnitude to be $m = 5$ and use the Keck telescope area ($A = 76$ m^2). We assume a bandwidth of $\Delta\lambda = 0.15\mu$m and upstream telescope throughput of $\eta_t = 0.06$, a typical value for Keck.

For each port of the PLN, we multiply PR by its null depth to calculate the photon rate per port. We then multiply that photon rate by the assumed exposure time of 60 s to obtain the counts per exposure. We add normally-distributed noise with a variance equal to the number of counts, an approximation for Poisson-distributed photon noise that is valid at our high photon count rates. We assume that each dataset corresponds to 5 hours of integration time, and thus generate 300 exposures per dataset. We generate a total of 1000 such datasets for analysis.

We also generate off-axis point-spread-functions (PSFs) that can be injected as astrophysical signal. The off-axis PSFs do not include WFE, since the simulations show that, at the WFE amplitudes of interest in our work, the planet coupling at separations of interest is not significantly impacted. In order to create data with an injected companion, the off-axis PSF at the desired separation is scaled appropriately based on the desired flux ratio, then added to each exposure of the simulated intensity of the on-axis source.

Detection

In this section, we characterize the detectability of planets, comparing the performance of the VFN and the PLN. For each dataset generated in Section 2.4, we first take the mean of the 300 exposures and subtract off the nominal on-axis signal with no WFE. We then perform detection testing on the resulting data, using a total energy test statistic:

$$\epsilon = \sum_i y_i^2, \tag{2.6}$$

where i is the port index of the PLN and y_i the signal in the port. The test statistic ϵ is calculated from the data and compared to a threshold ξ, which is chosen to provide a desired false-alarm rate. A detection is claimed if $\epsilon \geq \xi$, and a lack of detection is claimed otherwise.

There are four possible outcomes when comparing the test statistic calculated from a dataset to the value of the test statistic set as the detection threshold. The first is a true positive, in that a real companion in the data is detected; the fraction of real companions detected is the true positive rate (TPR). A second possible outcome is that a real companion is *not* detected, occurring at a rate of $1 - \mathrm{TPR}$. A third outcome is that there is no companion in the data, but the detection test incorrectly claims a detection. The rate at which this occurs is the false positive rate (FPR). The fourth and last outcome is that there is no companion, and a detection is correctly not claimed, occurring at a rate of $1 - \mathrm{FPR}$.

Choosing a threshold for the test statistic is a balancing act between the TPR and FPR: as the threshold is decreased, detecting real companions becomes more likely, but false detections also become more likely. This dependency can be characterized by examining the possible values of the test statistic and calculating the TPR and FPR *if* that value were the detection threshold. Plotting the TPR as a function of the FPR results in a receiver operating characteristic (ROC) curve, which characterizes the performance of a detection scheme and can be used in the determination of flux ratio detection limits.

Figure 2.9 shows ROC curves from the distribution of ϵ over the 1000 datasets. The VFN corresponds to the case where only the LP 01 port is used, while with the PLN, all four nulled ports are used. The simulations show that for both charges, the inclusion of the other nulled ports of the PLN provides detection gains relative to the VFN. For a given rate of false positives, the PLN can achieve a higher true positive rate than the VFN. At close in separations $\leq 1\lambda/D$, the charge 1 PLN achieves the best performance. At separations greater than $\approx 1.25\lambda/D$, the charge 2 PLN starts to perform better. Despite having higher throughput, the photonic lantern without a vortex does not outperform both the charge 1 and the charge 2 PLNs at any separation, emphasizing that the distribution of flux relative to the achievable null-depths matters more than sheer throughput. However, the no vortex PLN has the advantage of not requiring an additional optic in a pupil or focal plane, and can thus be realized with a simpler optical system. Additionally, the relative performance of the different configurations will ultimately depend on the distribution of WFE, as the ports in each configuration are sensitive to different subsets of modes.

Model-Fitting

Data from the VFN consists of only one measurement that contains no information on position angle and cannot discriminate between the effects of flux ratio and separation. Unlike the VFN, the spatial structures of the PLN modes allows for the retrieval of the planet's location, albeit with degeneracy in the position angle as a result of their symmetry.

To illustrate this capability, we attempt to fit models to one of the simulated datasets of the charge 2 VFN from Section 2.4, where a planet with a flux ratio of 2×10^{-6} is injected at ($X = 1.25\ \lambda/D$, $Y = 0\ \lambda/D$). We believe that a configuration that slightly breaks the symmetry would be a better strategy for localization than any of the configurations presented here. Determining how to do this effectively would be part of future work. For this work, our primary aim was to show that this localization capability exists in this architecture, so we choose to focus on just one configuration.

First, we assume that the average null-depth can be estimated, such as by observing a reference star. This assumes telescope conditions are reasonably stable between observations of the reference and target stars, as the accuracy of the null-depth estimation will be impacted by quasi-static aberrations as well as differential alignment onto the vortex or lantern centers, which would lead to differences between the reference and target observations.

The estimated null-depth is subtracted from the average of the measurement frames. This step is necessary to debias the data, since if only the nominal on-axis signal (without any wavefront error) is subtracted, the WFE that sets the null-depth will contribute to the apparent flux of the planet. We then fit a model to the data through Chi-squared (χ^2) minimization, using only data from the LP 01 port for the VFN, and data from all six ports for the PLN.

The three model parameters for a planet are its location coordinates (X, Y) and its flux ratio (FR). We first generate a grid of parameter values, choosing X to span from 0 λ/D to 3 λ/D and Y to span from -3 λ/D to 3 λ/D. This spans the spatial half-plane, which is enough for our purposes, as the symmetry of the modes means the position angle can at best be localized with a 180 degeneracy. The flux ratios are chosen to range logarithmically from 10^{-7} to 10^{-5}.

A planet corresponding to each set of parameters from the grid is simulated with the instrument model. The χ^2 of the difference between the model and the data is calculated using $\chi^2 = \sum_i (y_i - x_i)^2/\sigma_i^2$, where y_i is the measured data in port i, x_i is the model, and σ_i is the standard deviation of the noise across the 300 frames. The probability distribution is then calculated by taking $\mathrm{P}(X, Y, \mathrm{FR}) \propto \exp\{-\chi^2/2\}$, and normalizing such that the total probability over the entire explored parameter space is 1.

Figure 2.10 depicts the three spatial cross-sections of the resulting probability distributions for the charge 2 VFN and PLN, corresponding to the flux ratio values from the grid closest to the injected value of 2×10^{-6}. The parameter set in the grid closest to that of the injected planet is marked with an orange star. Also shown is the probability distribution of the flux ratio, marginalized over the spatial dimensions. As expected, it is largely unconstrained by the VFN, which cannot distinguish between the competing effects of flux ratio and separation. However, with the spatial information provided by the PLN, the retrieved probability distribution of the flux ratio peaks at the correct value of 2×10^{-6}. Given the best fit flux ratio using PLN, fitting a Gaussian curve to the y-axis cross-section of the spatial probability distribution reveals that the position angle can be localized to $\sim 1\ \lambda/D$ with the PLN, while it is completely unconstrained by the VFN. These simulation results show that compared to the VFN, the PLN can provide better constraints on the planet's location and flux ratio.

The response of the PLN to off-axis signal is not rotationally symmetric. We thus explore injecting and recovering a planet signal at varying position angles. Figure 2.11 shows that, given the correct flux ratio, the localization response varies as a function of position angle. At position angles other than 0 and $\pi/2$, additional solutions exist beyond the two guaranteed by the instrumental symmetry. However, an observing strategy that involves taking data with multiple rotations of the instrument relative to the sky will reduce the number of best fit position angle solutions to the fundamental two. Finding the most efficient observational strategy to best constrain the position angle given an unknown random initial orientation, and exploring the possibility of introducing slight asymmetries to break this degeneracy, are topics left for future work.

2.5 Conclusions

This work presents a proof-of-concept study of the Photonic Lantern Vortex Fiber Nuller. The advantage the MSPL offers over the SMF is two-fold. First, a photonic lantern, regardless of modal selectivity, accepts more input modes than the SMF, increasing the overall amount of light that can couple in. This improves the overall field of view and total planet coupling provided by the VFN. Second, the symmetries resulting from modal selectivity interact with the vortex field to create not just on-axis nulls, but also ports insensitive to low-order aberrations that do not meet a specific azimuthal order condition. Together, these properties of the PLN result in an instrument that rejects starlight while maintaining a substantial amount of planet light in the regions of interest. Additionally, while the PLN is meant for integration with spectrographs, motivated by the science that can be done in the spectral domain, the ports with different modal structures captures some spatial information, enabling planet localization that is not possible with the VFN. However, the instrumental symmetries that provide starlight and wavefront error rejection currently also cause degeneracies in the spatial information captured. Future work will explore whether introducing slight asymmetries into the instrument can lift the spatial degeneracies with minimal impact to the achievable null depth.

This work simulates the PLN's ideal behavior at a single wavelength. However, the modes of a realistic mode-selective photonic lantern will deviate from the ideal LP modes. Furthermore, its modes will actually vary with wavelength. Finite-difference beam propagation simulations are needed to simulate the behavior of a realistic photonic lantern design across different wavelengths, since its modes will no longer correspond to perfect LP modes, and there will be modal cross-coupling due to imperfections in the design as well as the fabrication process. Additional performance simulations will be conducted to characterize the impact of this non-ideal, wavelength dependent behavior on science results. This work includes simulating the PLN with synthetic planetary spectra and investigating methods to analyze the data, building upon current practices in exoplanet spectral analysis (J. J. Wang, Ruffio, et al., 2021). We will identify best practices to account for the wavelength dependent mode-structure and throughput and the optimal method for combining data from the different ports, including the possibility of obtaining concurrent stellar spectra in the non-nulled ports to be used for calibration and analysis. We will investigate if multiple sets of spectroscopic data can be used to cross-calibrate systematic errors. The single-mode outputs are ideal for downstream spectroscopy using photonic spectrographs (Gatkine, Veilleux, and Dagenais, 2019).

We will thus investigate strategies for optimal integration of PLN with an on-chip photonic spectrograph on each of the single-mode outputs (nulled or otherwise) to measure the spectra of the planet/companion and star, as well for cross-calibration.

Future work also includes verifying the behavior of a PLN in the lab — both the characterization of the photonic lantern device itself, and after integration with a vortex. We intend to characterize the PLN with different levels of wavefront error, as well as investigate the possibility of performing wavefront control to achieve better nulls, potentially compensating for defects such as residual optical surface error or even non-ideal photonic lantern modes. If the laboratory characterization validates the performance of the PLN, an on-sky demonstration will be attempted.

This work on the PLN also naturally ties in to several related topics, such as the development of wavefront sensing algorithms through photonic lanterns (Lin, Fitzgerald, et al., 2022a; Norris, Wei, et al., 2020), or the leveraging of the photonic lantern design paradigm to push towards the theoretical limits of optical signal separation.

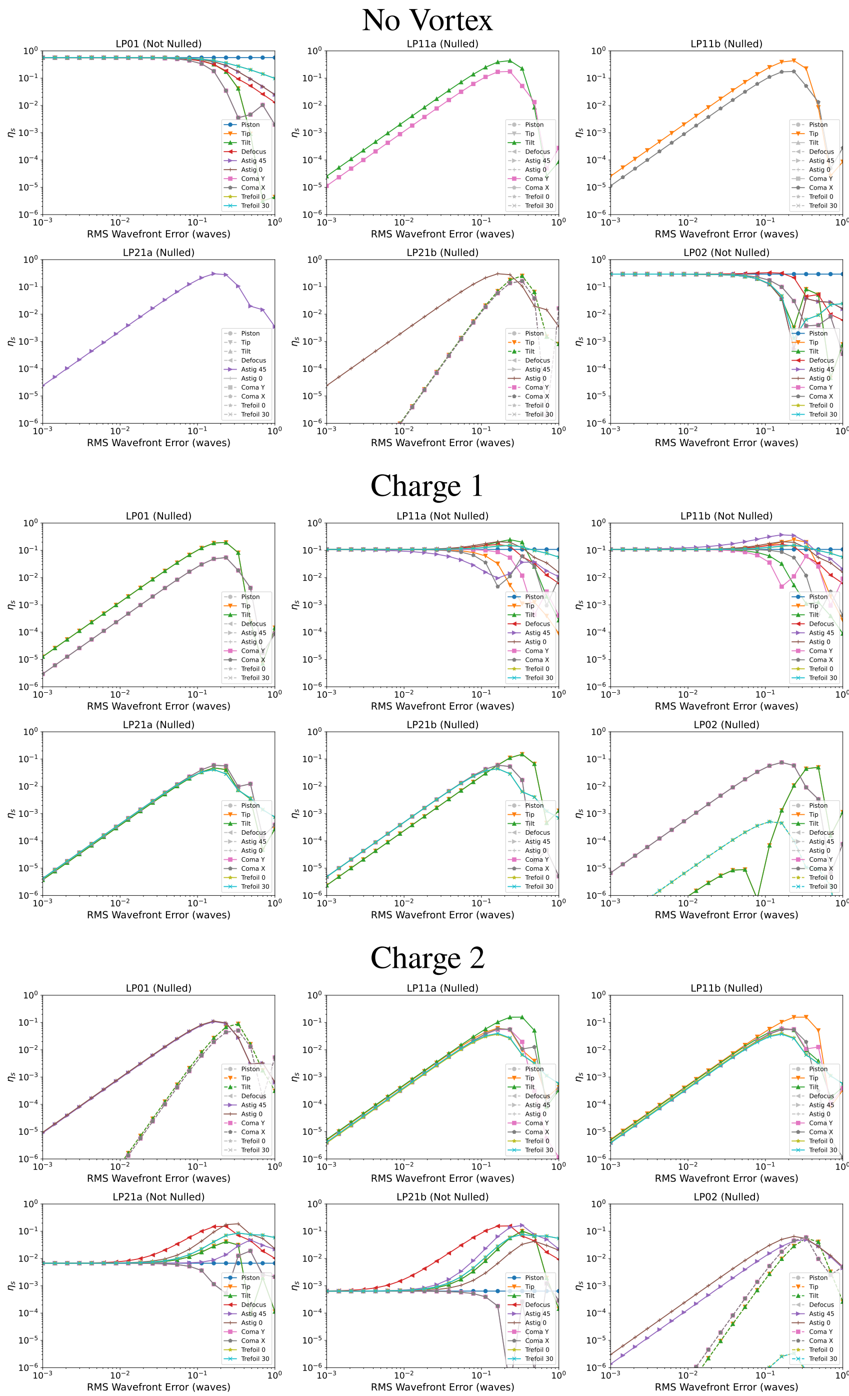

Figure 2.5: Stellar coupling rate as a function of individual Zernike polynomial amplitude, with no vortex (top), a charge 1 vortex (middle), and a charge 2 vortex (bottom). For the nulled ports, solid lines indicate modes predicted to couple (those satisfying $l \pm (m'+m) = 0$), while dashed lines indicate modes that are not predicted to couple (to first order, though higher-order coupling effects can be seen). Values of η_s falling below 10^{-6} are likely numerical noise, and are not shown. Lines that fall entirely below 10^{-6} are light grey in the legend.

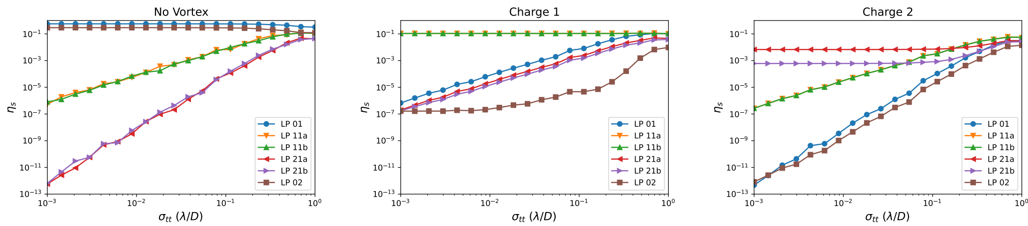

Figure 2.6: Left: Stellar coupling rates as a function of tip-tilt jitter, random-uniformly distributed in position angle, with no vortex (left), a charge 1 vortex (middle), and a charge 2 vortex (right). The standard deviation of the per-frame tip-tilt amplitude is given by σ_{tt}, with position angle drawn uniformly between 0 and 2π.

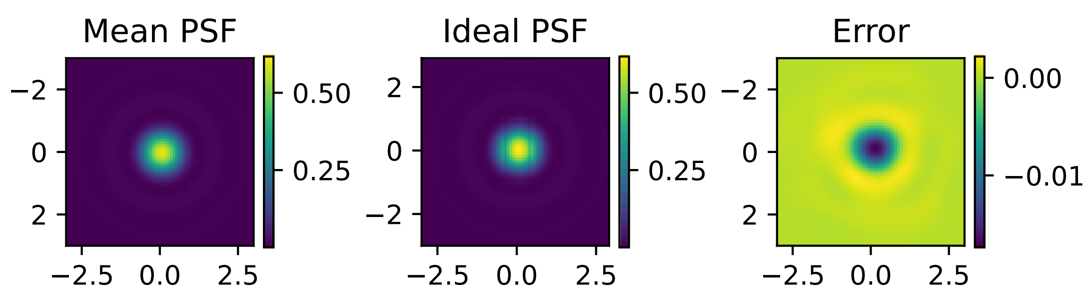

Figure 2.7: Left: Mean focal-plane PSF in the presence of WFE as measured by the KPIC PyWFS. Middle: Unaberrated focal-plane PSF. Right: Difference between the aberrated and ideal PSFs. Reminder that these are **not** simulations of the Keck PSF, but of the measured wavefront error residuals propagated through a system with an ideal circular aperture.

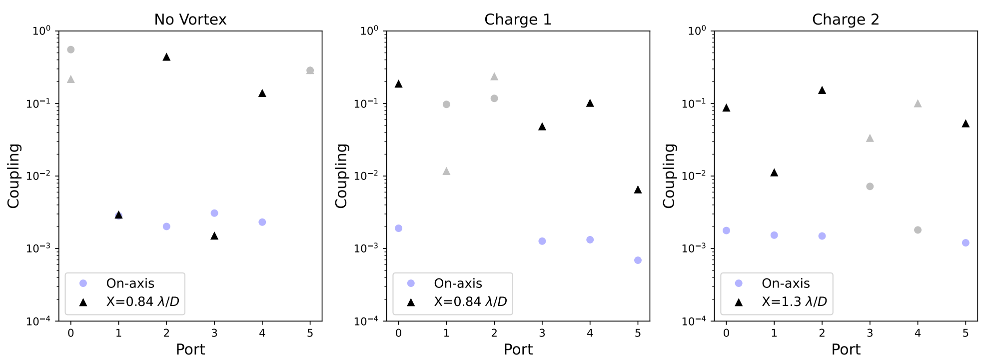

Figure 2.8: Mean coupling calculated over 590 frames of WFE residuals from the KPIC PyWFS, for the no vortex (left), charge 1 (middle), and charge 2 (right) configurations. The ports on the bottom axis are (from left to right): LP01, LP11a, LP11b, LP21a, LP21b, and LP02. Coupling values for ports that are not considered nulled are depicted in light grey. Off-axis planet coupling (where it is expected based on the coupling maps) remains higher than the stellar coupling in the presence of these WFE realizations.

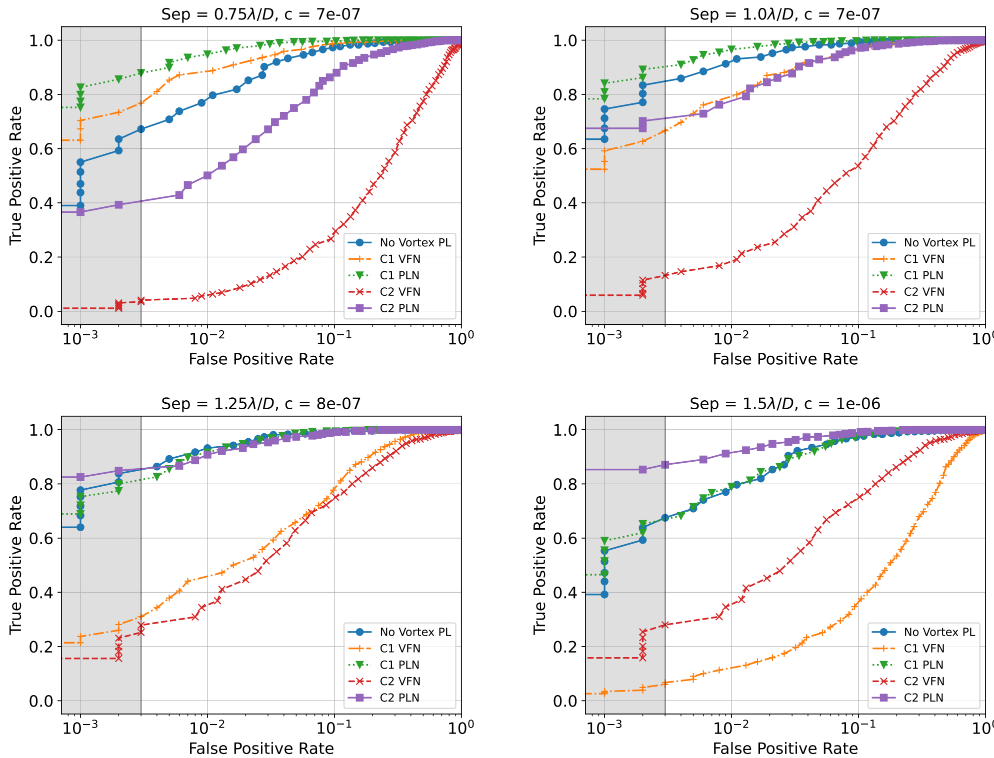

Figure 2.9: Example ROC curves at different separations in the presence of with photon noise, assuming wavefront error averages to a baseline null depth. For both vortex charges, the inclusion of other ports of the PLN provides detection gains relative to the VFN. The grey areas indicate false positive rates which are not well sampled as they involve fewer than 3 datasets with false detections.

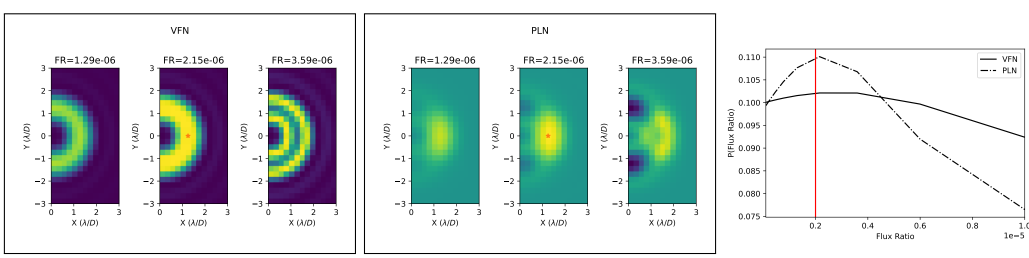

Figure 2.10: Left: Select spatial probability distribution cross-sections, using a charge 2 VFN. The three panels are plotted on the same color scale. Middle: Select spatial probability distribution cross-sections, using a charge 2 PLN. The three panels are plotted on the same color scale. The parameters closest to that of the injected planet are marked with orange stars. Right: Probability distributions of the flux ratio, marginalized over the spatial dimensions. The flux ratio of the injected planet is marked by the red line. The model-fitting shows that the PLN can provide better constraints on planet model parameters compared to the VFN.

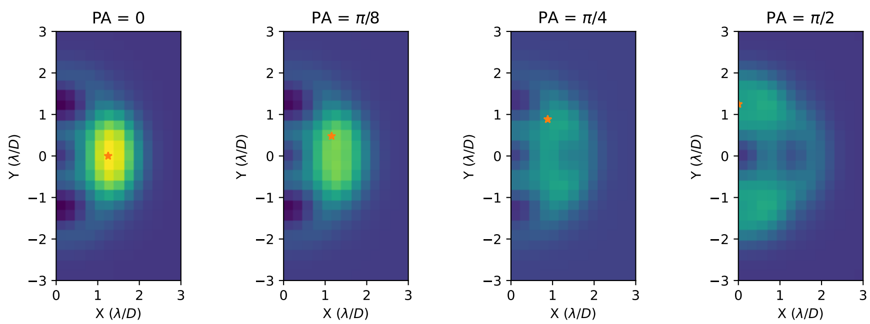

Figure 2.11: Spatial probability distributions given the correctly identified flux ratio of 2.15×10^{-6} (the panels are plotted on the same color scale). Planets at a separation of $1.25\,\lambda/D$ are injected at a variety of injected position angles (marked by the orange stars). At position angles other than 0 and $\pi/2$, additional solutions exist beyond the two guaranteed by the instrumental symmetry. However, an observing strategy that involves taking data with multiple rotations of the instrument relative to the sky will reduce the number of best fit position angle solutions to the fundamental two.

Chapter 3

LABORATORY DEMONSTRATION OF A PHOTONIC LANTERN NULLER IN MONOCHROMATIC AND BROADBAND LIGHT

Xin, Yinzi et al. (2024). "Laboratory demonstration of a Photonic Lantern Nuller in monochromatic and broadband light". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 10.2, p. 025001. DOI: 10.1117/1.JATIS.10.2.025001. URL: https://doi.org/10.1117/1.JATIS.10.2.025001.

This chapter is a reproduction of my paper which details the first laboratory demonstration of the PLN. In this work, I characterized a mode-selective photonic lantern sent to us by collaborators at the University of Sydney and the University of Central Florida, then demonstrated its use as a nuller on the Polychromatic Reflective Testbed (PoRT) at the Caltech Exoplanet Technology Lab. I measured the modes of the lantern, which are not ideal LP modes due to imperfections that result in 'cross-coupling' between the ports. I then characterized the nulling behavior of the lantern on PoRT and compared it to simulations using the measured modes, and showed that the experiment agrees with the prediction. The null-depths achieved in this work are on the order of 10^{-2}, with future work in Chapter 4 further improving these null-depths using wavefront control.

Abstract

Photonic lantern nulling (PLN) is a method for enabling the detection and characterization of close-in exoplanets by exploiting the symmetries of the ports of a mode-selective photonic lantern (MSPL) to cancel out starlight. A six-port MSPL provides four ports where on-axis starlight is suppressed, while off-axis planet light is coupled with efficiencies that vary as a function of the planet's spatial position. We characterize the properties of a six-port MSPL in the laboratory and perform the first testbed demonstration of the PLN in monochromatic light (1569 nm) and in broadband light (1450 nm to 1625 nm), each using two orthogonal polarizations. We compare the measured spatial throughput maps with those predicted by simulations using the lantern's modes. We find that the morphologies of the measured throughput maps are reproduced by the simulations, though the real lantern is lossy

and has lower throughputs overall. The measured ratios of on-axis stellar leakage to peak off-axis throughput are around 10^{-2}, likely limited by testbed wavefront errors. These null-depths are already sufficient for observing young gas giants at the diffraction limit using ground-based observatories. Future work includes using wavefront control to further improve the nulls, as well as testing and validating the PLN on-sky.

3.1 Motivation

Previous conceptual work on the PLN assumed a perfect, lossless MSPL (Xin, Nemanja Jovanovic, et al., 2022). However, a real MSPL will not have the ideal mode shapes corresponding to perfect LP modes, and the device itself will have some additional throughput loss. In this work, we characterize the properties of a real photonic lantern including manufacturing imperfections (described in Section 3.2). We then integrate it into a testbed to demonstrate the PLN in the lab (described in Section 3.3).

3.2 Lantern Characterization

A picture of the MSPL (optimized to have six ports at 1550 nm) is shown in Fig. 3.1a. The lantern is the stiff silver portion in the top right, with the MM end facing towards the right. Each SMF output of the lantern is connected to one of the white fiber pigtails.

We first characterize the properties of this lantern on its own by taking microscope images of the MM interface, measuring the throughput through each port, and using an interferometer to reconstruct the mode shapes corresponding to each port.

Microscope Imaging of Lantern Interface

We use a Dino-Lite Edge 3.0 microscope to image the multimode interface of the lantern, which is shown in Fig. 3.1b. The residual fiber cores are arranged in a pentagonal pattern, visible towards the bottom left as small black circles. The residual fiber cores are embedded in the multi-mode core (dark brown), which is surrounded by fiber-doped glass cladding (light brown ring). Surrounding the cladding is a silica substrate (outermost dark brown ring), followed by the glue (rough tan material) that attaches the lantern to the connector. The microscope image verifies the expected pentagonal arrangement of the single cores for the MSPL design presented in Leon-Saval, Fontaine, et al. (2014), also depicted in schematic form in Fig. 3.1c. The observed MM core diameter and distance between

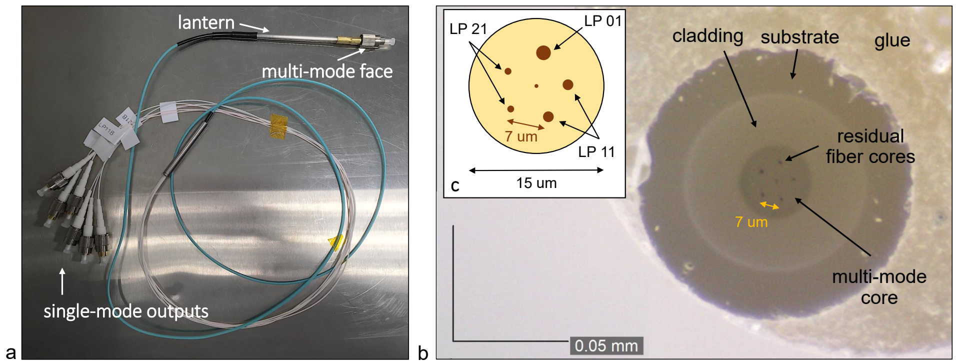

Figure 3.1: a) A picture of a six-port MSPL. The lantern is the stiff silver portion in the top right, with the MM end facing towards the right. Each SMF output of the lantern is connected to one of the white fiber pigtails. b) A microscope image of the MM face taken with a Dino-Lite Edge 3.0. The residual fiber cores are arranged in a pentagonal pattern, visible towards the bottom left as small black circles. The residual fiber cores are embedded in the multi-mode core (dark brown), which is surrounded by fiber-doped glass cladding (light brown ring). Surrounding the cladding is a silica substrate (outermost dark brown ring), followed by the glue (rough tan material) that attaches the lantern to the connector. c) A schematic of the design of the MM face for comparison. The observed MM core diameter and distance between adjacent residual SMF cores are both consistent with the design values of 15 μm and 7 μm respectively.

adjacent residual SMF cores are both consistent with the design values of 15 μm and 7 μm respectively.

End-to-end Throughput Measurements

Next, we measured the throughput of each of the lantern's ports from the single-mode inputs to the MM face using a power meter (Thorlabs S122C), with a laser diode (Thorlabs KLS1550) as the light source. We first took a background measurement with the light turned off. Next, we took a measurement of the power coming out of the fiber directly connected to the laser. Then, we connected the light source fiber to one of the SMF pigtails of the photonic lantern and measured the power coming out of the MM face. After background subtracting the measurements, we take the ratio of the power coming out of the MM face to the power coming out of the source fiber

Table 3.1: The average and standard deviation of five throughput measurements for each lantern port. Note that these throughput measurements are of the entire lantern assembly, including any connectorization losses or losses through the SMF pigtails, such as Fresnel loss and propagation loss.

Port	Throughput	Standard Deviation
LP 01	0.886	0.013
LP 11a	0.890	0.013
LP 11b	0.778	0.017
LP 21a	0.583	0.009
LP 21b	0.614	0.022
LP 02	0.617	0.014

to be the throughput of that port. We repeat this process five times for each port, and report the mean and standard deviation of those five measurements in Table 3.1. These throughput measurements are of the entire lantern assembly and include any connectorization losses (e.g. at the interfaces to the LC connectors), splice losses (e.g. between the lantern and the SMF pigtails), and losses from Fresnel reflection and propagation through the pigtails — and are consistent with losses expected from the assembly manufacturing. Note that in comparison, a typical SMF patch cable of equivalent length would have throughput > 95%.

Characterization of Modes

We use a technique called off-axis holography (OAH) to measure the complex electric field corresponding to each of the lantern's ports. A detailed discussion of the principles of OAH can be found in Cuche, Marquet, and Depeursinge (2000). In summary, a broad reference beam is interfered with an image of the lantern mode, creating fringes across the image. The fringes create sidelobes in Fourier space, and by filtering the Fourier-transformed signal, the electric field of the mode can be reconstructed. The principles and process of OAH is also similar to that of the self-coherent camera (P. Baudoz et al., 2006a).

A picture of our optical setup for OAH is shown in Fig. 3.2. For this experiment, we use a polarized tunable narrow linewidth laser (Thorlabs TLX2), set to a wavelength of 1568.772 nm with a linewidth of 10 kHz, which equates to $\sim 10^{-16}$ m. This results in a coherence length of $\sim 10^4$ m. The light is split by a 50:50 polarization-maintaining (PM) splitter, which sends half the light to the imaging arm of the interferometer, and the other half to the reference beam arm. The light in the imaging arm goes through a polarization controller, then to one of the single-mode inputs of the MSPL. The light coming out of the MM face of the lantern goes through a lens that collimates the beam, then to a lens that focuses the image onto an InGaAs camera (First Light C-RED 2), which has a pixel pitch of 15 μm.

The light in the reference beam path passes through a PM fiber coil (to better match path lengths between the two arms, even though the long coherence length does not necessitate this), then through a lens that forms a diverging beam large enough to cover the entire lantern mode image. This reference beam interferes with the lantern mode image, creating fringes that allow us to retrieve the complex mode field using Fourier analysis. Ideally, the reference beam should be collimated into a flat wavefront; however, we have simply made it large enough that its phase is slowly varying over the extent of the mode image, so does not significantly impact the reconstructions.

The visibility of the fringes is highest when the polarization state of the two beams are matched. To match the polarization between the two arms, a calibration polarizer is first inserted into the reference arm. The polarizer is then fixed to the angle that cancels the flux of the reference beam on the detector. Then, the polarizer is moved into the lantern arm, and the polarization controller set to minimize the flux that goes through the polarizer and onto the detector. This process aligns the polarization states of both arms to each other. The calibration polarizer is then removed from the beam for the rest of the experiment.

See Cuche, Marquet, and Depeursinge (2000) for the principles of using Fourier analysis to retrieve complex amplitudes from off-axis holography (OAH). For our work, an example hologram of the LP 21b port (after dark subtraction and centering) is shown in Fig. 3.3a, and the same hologram zoomed in to the center (such that the fringes are visible) is shown in Fig. 3.3b. Note that the fringes of interest resulting from interference between the two beams are the fine horizontal ones. The faint, wide, vertical stripes are not from the interference of the two beams, but rather the structure of the reference beam itself, as shown in Fig. 3.3c. We apply a 2D Fourier transform to the (full-sized) hologram to obtain the Fourier space signal (Fouriergram) shown in Fig. 3.3d. We then isolate the top right lobe of the Fouriergram, shown in Fig. 3.3e.

The cropped lobe in Fig. 3.3e is centered to minimize the amount of tip-tilt signal in the final reconstructed mode, shown in Fig. 3.3f. To obtain this final reconstructed mode, we first apply a broad Gaussian window (with a σ of 27 pixels) to the cropped

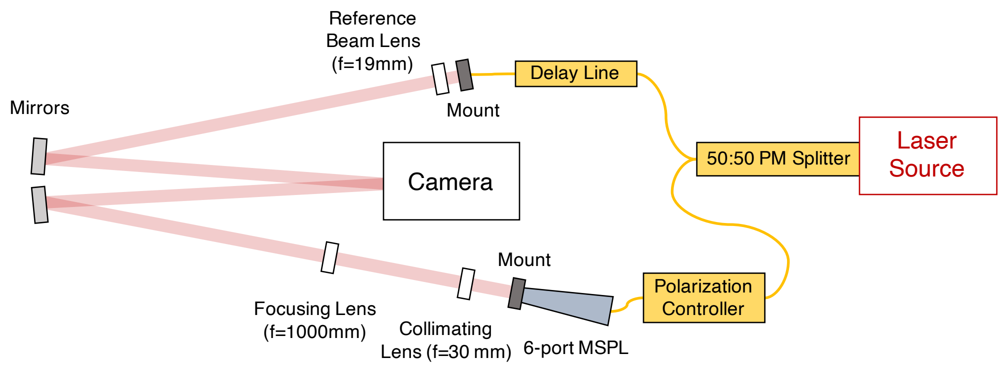

Figure 3.2: The optical setup for OAH measurements of the lantern modes. The light from the laser is split by a 50:50 polarization-maintaining splitter. Half the light is sent to the imaging arm, through a polarization controller, then to one of the single-mode inputs of the MSPL. The light coming out of the MM face of the lantern goes through a lens that collimates the beam, then to a lens that focuses the image onto an infrared camera. The other half of the light is sent to the reference beam path, where it passes through a PM fiber coil delay line, then through a lens that creates a diverging beam large enough to cover the entire lantern mode image. The polarization controller is used to match the polarizations of the two beams. The reference beam interferes with the lantern mode at the detector, creating fringes that allow us to retrieve the complex mode field using Fourier analysis.

Fouriergram to filter out edge effects, then apply an inverse Fourier transform. We then divide by the square-root of the reference beam intensity to remove the impact of its non-uniformity. Finally, we normalize the mode such that its summed intensity is 1. Because the arms of the interferometer are long, and the path length difference between the reference arm and the lantern arm fluctuates widely (due to vibrations and other bench instabilities), we are unable to constrain the global phase of the modes (i.e, the uniform phase term, or the phase piston as expressed in the focal plane). Fortunately, we do not need to know the global phase to predict the coupled intensities, since they are not impacted by the global phases of each mode.

We apply this same reconstruction process to the other 5 modes of the lantern. We also obtain and analyze three separate datasets taken across multiple days and confirm that the measurements are qualitatively stable. After matching the global phases between the three different measurements, we take their mean and re-normalize each mode to a total intensity of 1 to obtain the final reconstructions, shown in Figure 3.4a. In Figure 3.4b, we plot the dot-products between the measured modes, which show that they are orthogonal as predicted (the median of the dot-product magnitudes between two different modes is 0.011, which is commensurate with our

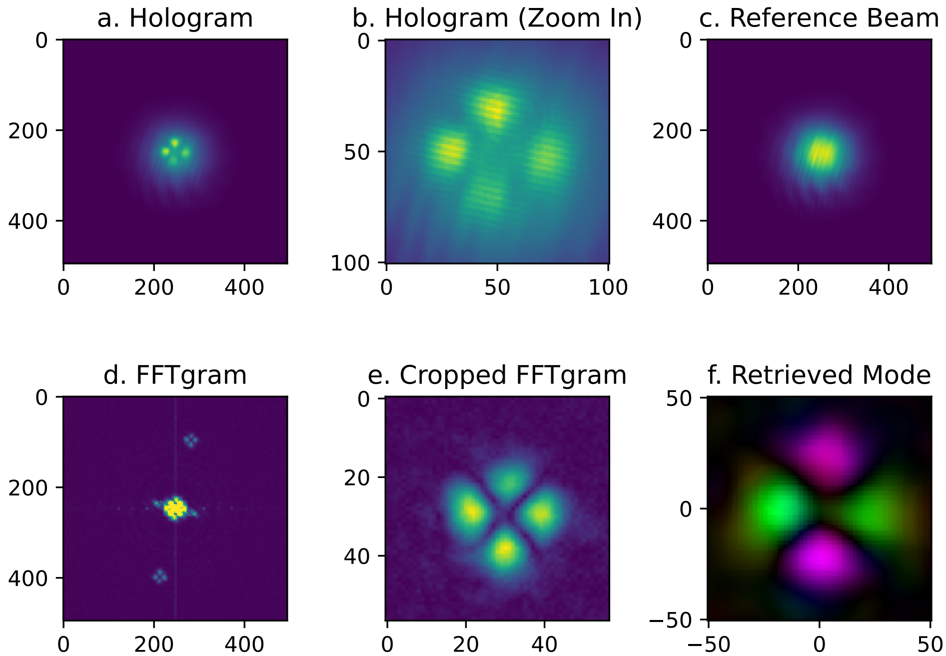

Figure 3.3: a) The measured hologram of the LP 21b port, centered and background subtracted. b) The same hologram as part (a), but zoomed in to the center to show the fringes. Note that the fringes of interest resulting from interference between the two beams are the fine horizontal ones. c) The reference beam intensity, plotted on the same spatial scale as part (a), showing that the faint, wide, vertical stripes are not from the interference of the two beams, but rather the structure of the reference beam itself. d) The 2D Fourier transform of the hologram in part (a). e) The Fourier-space signal in part (d), cropped to the top right lobe. This lobe is centered to minimize the tip-tilt signal in the final reconstructed mode. f) The final reconstructed mode, obtained by first applying a Gaussian window with a σ of 27 pixels to part (d) to filter out edge effects, then a 2D inverse Fourier transform. The signal is then divided by the square root of the reference beam intensity to remove its impact, then normalized to a total intensity of 1. The amplitude is indicated by brightness, and the phase indicated by hue.

measurement uncertainty for the mode shapes themselves).

In Section 3.3, the throughput maps simulated using these measured modes are compared to the actual throughput maps obtained on the testbed.

3.3 Photonic Lantern Nuller Demonstration

After characterizing the properties of the MSPL, we integrated it into the Polychromatic Reflective Testbed (PoRT) (Echeverri, Ruane, Benjamin Calvin, et al., 2020)

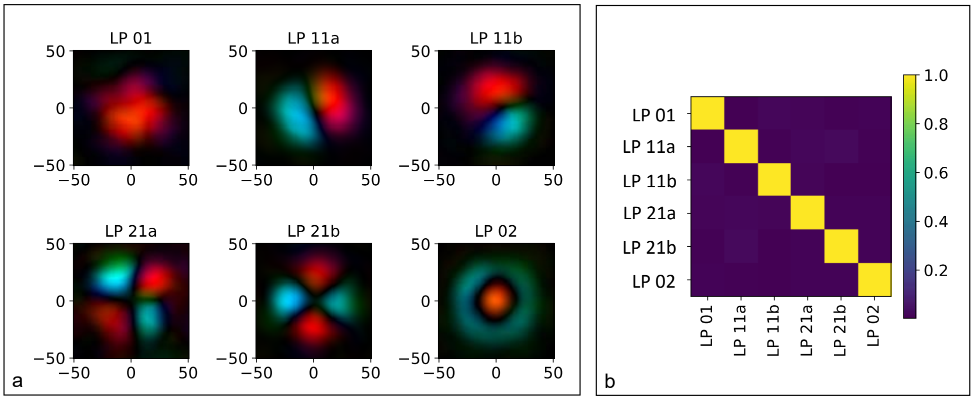

Figure 3.4: a) The measured modes corresponding to each port of the MSPL, obtained using OAH. The amplitude of each mode is indicated by brightness, and the phase indicated by hue. Each mode has its global piston phase term removed, and has been normalized to a total intensity of one. The axes correspond to pixels on the camera, centered about zero. b) The dot-products between the measured modes, which show that they are orthogonal as predicted. The median of the dot-product magnitudes between two different modes is 0.011, which is commensurate with our measurement uncertainty for the mode shapes themselves.

at Caltech to demonstrate using it as a nuller.

Experimental Setup

A diagram of the PoRT testbed is shown in Fig. 3.5. A light source is fed into the bench with a single-mode fiber mounted to the source stage. The light is collimated by an off-axis parabola (OAP) mirror. The collimated light is filtered by a baffle before reflecting off of a 12×12 Boston Micromachines deformable mirror (DM). Then, a set of relay OAPs magnifies the beam. In the resulting collimated beam is a mask mount that can be used to insert a pupil plane mask (such as a vortex); however, we leave it empty for this work. The beam then passes through an adjustable-size iris, which we use to control the $F\#$ of the system. The iris aperture diameter (D) can range from 1-15 mm, which, given the injection focal length (f) of 54.4 mm, can provide focal numbers ($F\# = f/D$) ranging from approximately 3.6 to 55.

The beam is then focused by the last OAP onto the injection stage, which holds both a SMF and the 6-port MSPL. The injection stage can move in translation in all three axes, which allows light to be injected into either the SMF or the MSPL, and can also be used to scan the face of either optic (the fiber or the lantern) across the focused beam. To measure the coupled flux through the SMF, the output end of the SMF is routed to an InGaAs (Femto OE-200-IN2) photodiode. To measure the coupled flux through one of the lantern ports, the corresponding SMF pigtail is routed to the Femto photodiode. This setup can only measure one port at a time, so the measurements for the different ports are made sequentially.

To normalize the data, we use a retractable stage to insert a power meter (Thorlabs S122C) into the beam just before the injection mount to measure the incident flux. After calibrating the readings to that of the Femto photodiode, these beam flux measurements can be used to normalize the coupled flux measurements, in order to obtain throughput measurements.

To best compare our results to the predictions from simulation, we wish to inject a beam with a flat wavefront into the photonic lantern. To flatten the wavefront, we use the SMF as a wavefront calibrator, since coupling into an SMF is maximized when the wavefront is flat. This is most easily seen by expressing the coupling as proportional to the square of the overlap integral of the electric field with the SMF mode (approximated as a Gaussian), expressed in the pupil plane with radial coordinate ρ and angular coordinate ϕ:

$$\eta \propto \left| \int_0^D e^{-\left(\frac{\pi D_f \rho}{2}\right)^2} e^{i\Phi(\rho,\phi)} dA \right|^2. \tag{3.1}$$

Here, D is the diameter of the aperture (assumed to be circular), D_f is the mode-field diameter of the SMF, and Φ is the wavefront error. The coupling is thus maximized when Φ is uniform over the entire aperture (i.e. there is no wavefront error), as any deviations from uniformity will reduce the value of the integral.

We first optimize coupling into the SMF by adjusting the X, Y, and Z directions of the injection stage. We then optimize the iris size, achieving maximum injection with an $F\#$ of 3.8. Then, we tune the twelve lowest-order Zernike modes of the DM map. We obtain a peak throughput of 74.1%, whereas the theoretical coupling (calculated as the overlap integral of the ideal SMF mode with the focal-plane electric field given a perfectly flat wavefront) is 82.8%. We have not accounted for Fresnel loss at the face of the fiber and propagation loss through the fiber, which may explain

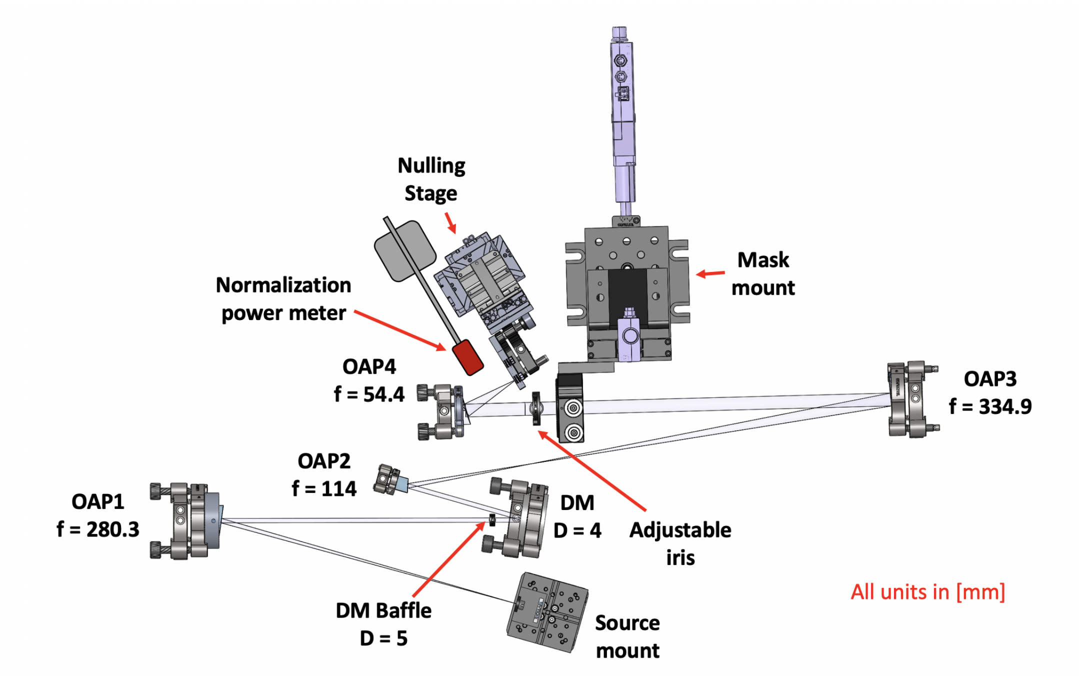

Figure 3.5: A light source is fed into the bench with a SMF mounted to the source stage. The light is collimated by an off-axis parabola (OAP) mirror. The collimated light is filtered by a baffle before reflecting off of a 12×12 Boston Micromachines deformable mirror (DM). Then, a set of relay OAPs magnifies the beam. We leave the mask mount empty for this work. The beam then passes through an adjustable-size iris, which we use to control the $F\#$ of the system. The beam is then focused by the last OAP onto the injection stage, which holds both a single-mode fiber and the 6-port MSPL. A Femto OE-200 photodiode is used to measure the coupled flux. A Thorlabs S122C power meter on a retractable stage can be inserted into the beam just before the injection mount to measure the incident flux, which can be used to normalize the coupled flux measurements for throughput measurements.

part of the discrepancy. The remaining losses may be a result of uncalibrated higher order wavefront errors.

Next, we keep the DM map that optimizes injection into the SMF (which implies minimal wavefront aberration), but translate the injection stage to the location of the photonic lantern. We then tune the iris size to set the $F\#$ into the lantern, a procedure that we discuss below.

There are several metrics one could use to determine the optimal $F\#$. The signal in a given port is the sum of both the stellar throughput at the center (η_s) and the off-axis throughput at the planet location (η_p). Broadly, we wish to maximize the signal-to-noise ratio (i.e., the planet light relative to photon noise from stellar leakage), which scales as $\eta_p/\sqrt{\eta_s}$. While η_s changes as a function of $F\#$, wavefront control can be used to further improve the null — without significantly impacting the $F\#$ that maximizes the off-axis throughput for each port. Although implementing wavefront control with a PLN is left for future work, because it would give us additional control over η_s, we choose to optimize iris size for planet throughput instead. We expect the exoplanet light to mostly couple into one of the LP 11 ports (since they tend to have higher throughput overall), so the peak off-axis throughput ($\eta_{p_{\text{peak}}}$) through either the LP11a or the LP11b port is a simple proxy for exoplanet throughput. In our work, we maximize through the LP 11a mode (though the LP 11b mode would be an equally valid choice), resulting in a $F\#$ of 6.2.

In theory, the truly optimal $F\#$ depends on the planet location, and what the right metric is depends on how well that location is known. In practice, the $F\#$ of an instrument will be fixed to a certain value regardless of the target being observed, and most reasonable optimization metrics targeting close-in planets will result in similar $F\#$'s.

Key Results - Monochromatic

In Figure 3.6a, we present the PLN throughput maps measured with 1568.772 nm light from the TLX2 tunable narrow linewidth laser, injected into the PoRT testbed with a PM fiber. The peak off-axis throughput of each port is reported in Table 3.2.

We also simulate throughput maps based on the mode profiles we reconstructed in Section 3.2 (note that the OAH data was taken using the same wavelength as the monochromatic PoRT measurements, but that the relative polarization between the OAH measurements and the PoRT measurements is unknown). We plot the simulated maps in Fig. 3.6b on the same color-scale as the PoRT measurements. We use a manual image alignment procedure to set the field of view, sampling, and rotation angle of the simulations to achieve the best match (across all six-ports) to the measured throughput maps.

The simulation with OAH modes assumes that the lantern is flux-preserving — that whatever light gets coupled into a given port is maintained through the lantern. The real lantern assembly is lossy, so has lower throughputs than in simulation. Qualitatively, however, the simulated throughput maps and the measured throughput maps have similar morphologies, showing that our measurements of the PLN's behavior largely agree with the model.

To better resolve the central null, we repeat the scans, but with finer spatial sampling by a factor of 5. The results are shown in Fig. 3.6c. Because the maps are asymmetric due to manufacturing imperfections, and the apparent centers of the modes slightly offset from each other, it is ambiguous which XY coordinate, even in simulation, should be designated as the theoretical axial center to which the star would be aligned. One approach, which we use for this work, is to sum up the central throughput maps of the four nulled ports (shown in Fig. 3.6d), and then take the location of minimum summed throughput as the center. The stellar leakages (η_s) at the identified center for each port are reported in Table 3.2.

In Figure 3.7, we plot select cross-sections of the throughput maps shown in Figure 3.6, and compare them against the throughput maps of an ideal MSPL with perfect LP modes. We show that the imperfect mode shapes cause different throughput profiles from an ideal MSPL, but that the measured profiles agree with the simulations using the modes obtained from OAH, except with additional losses from propagating through the lantern. Note that, while we do not see this phenomenon in our measured throughput maps, it is possible for the planet throughput at a given spatial location to be higher in one particular port than predicted with perfect LP modes. This is simply a result of the imperfect mode shapes, which distribute planet light differently amongst the ports, i.e., for a given planet position, the lantern imperfections can cause higher throughput in one port at the expense of throughput in other ports, relative to perfect LP modes.

In Table 3.2, we also report the values of $\eta_s/\eta_{p_{\mathrm{peak}}}$. Because the overall loss of each port cancels out in this ratio, it allows for a direct comparison between the PoRT measurements and the simulations using OAH modes. The quantity $\eta_s/\eta_{p_{\mathrm{peak}}}$ is often referred to as the 'null-depth' in discrete-aperture nulling interferometry contexts, though it is also called the 'raw contrast' in fiber-fed spectroscopy contexts (which is different from how raw contrast is typically used in coronagraphy contexts). To avoid confusion over terminology, we will refer to it in this work explicitly as $\eta_s/\eta_{p_{\mathrm{peak}}}$.

We then repeat the experiment using orthogonally polarized light by injecting the laser into the bench with a 90° polarization rotating fiber (see Appendix 3.3 for the throughput maps, as well as a comparison between the maps obtained for the two polarizations). A summary of the important measurements is presented in Table 3.2.

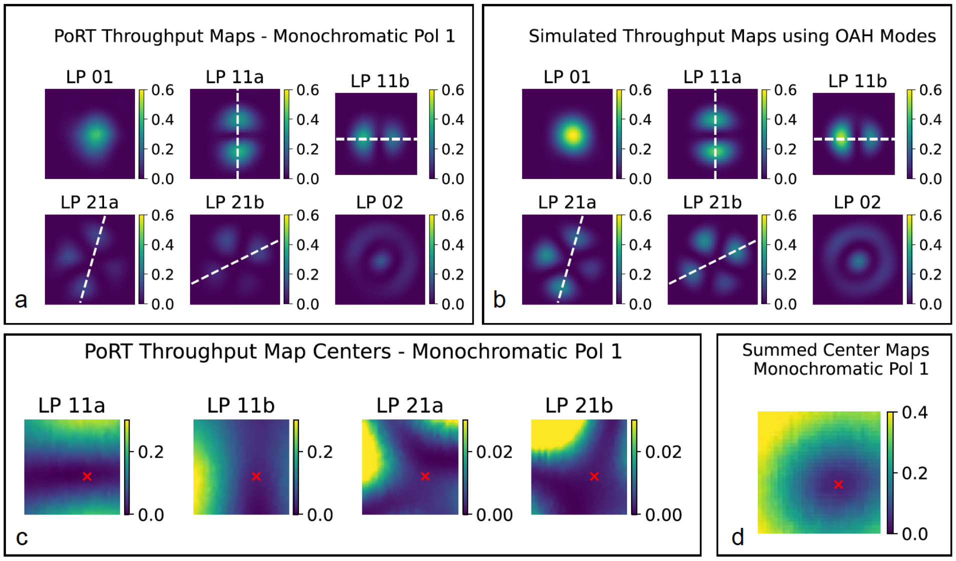

Figure 3.6: a) Monochromatic PLN throughput maps measured with 1568.772 nm light from the TLX2 tunable narrow linewidth laser injected into the PoRT testbed with a PM fiber. White dashed lines indicate cross-sections plotted in Figure 3.7. b) Simulated throughput maps based on the mode profiles reconstructed using OAH at the same wavelength, assuming that the lantern is flux-preserving. White dashed lines indicate cross-sections plotted in Figure 3.7. c) Monochromatic PoRT throughput maps of the nulled ports with fine spatial sampling of the center. The red crosses indicate the axial center of the lantern, identified using the map in part (d). d) The summed throughput of the four maps in part (c). The location of minimum summed throughput is taken to be the lantern center, where η_s is measured.

Key Results - Broadband

We also repeat the same experiment using a broadband Super-Luminescent Diode (SLD) light source with wavelength coverage from 1450 nm to 1625 nm (Thorlabs S5FC1550P-A2). The throughput maps for both polarizations are presented in the Appendix, along with a comparison between them. The peak and coaxial throughput values, as well as the ratio, determined using the same methodology as with monochromatic light, are reported in Table 3.3. The broadband null-depths are not significantly different from the monochromatic ones (with a difference varying from a few percent to a factor of 2), indicating that the nulls are fairly achromatic. This is as expected: the MSPL is designed to be mode-selective across a wide wavelength range, and thus the (vortex-free) PLN is intrinsically achromatic, one of its major advantages compared to other nulling techniques.

We did not perform OAH using the SLD, as the bandwidth of light results in a much

Table 3.2: Key monochromatic PLN metrics from a) simulations using the modes measured with OAH assuming that the lantern is flux-preserving, b) the PoRT testbed with the laser routed through a PM fiber (polarization 1), and c) the PoRT testbed with the laser routed through a 90° polarization rotating fiber (polarization 2). $\eta_{p_{\text{peak}}}$ refers to the maximum throughput of each map, or what the throughput of a planet at the location of maximum coupling would be. For the nulled ports, η_s refers to the throughput at the coaxial center of the lantern, corresponding to the stellar leakage. For the nulled ports, we also report the ratio $\eta_s/\eta_{p_{\text{peak}}}$, which is unaffected by the overall throughput of a port, and thus allows for a direct comparison between the PoRT measurements and the simulations using OAH modes.

	LP 01	**LP 11a**	**LP 11b**	**LP 21a**	**LP 21b**	**LP 02**
$\eta_{p_{\text{peak}}}$ (Sim.)	0.610	0.463	0.542	0.285	0.288	0.232
$\eta_{p_{\text{peak}}}$ (Pol. 1)	0.421	0.385	0.384	0.138	0.130	0.160
$\eta_{p_{\text{peak}}}$ (Pol. 2)	0.461	0.396	0.379	0.146	0.140	0.171
η_s (Sim.)	N/A	7.32×10^{-4}	3.65×10^{-3}	5.91×10^{-3}	9.05×10^{-4}	N/A
η_s (Pol. 1)	N/A	4.05×10^{-3}	2.64×10^{-2}	2.85×10^{-3}	1.77×10^{-3}	N/A
η_s (Pol. 2)	N/A	6.53×10^{-3}	3.00×10^{-2}	5.12×10^{-4}	1.07×10^{-3}	N/A
$\eta_s/\eta_{p_{\text{peak}}}$ (Sim.)	N/A	1.59×10^{-3}	6.69×10^{-3}	2.05×10^{-2}	3.18×10^{-3}	N/A
$\eta_s/\eta_{p_{\text{peak}}}$ (Pol. 1)	N/A	1.05×10^{-2}	6.88×10^{-2}	2.07×10^{-2}	1.37×10^{-2}	N/A
$\eta_s/\eta_{p_{\text{peak}}}$ (Pol. 2)	N/A	1.69×10^{-2}	7.82×10^{-2}	3.72×10^{-3}	8.24×10^{-3}	N/A

Table 3.3: Key broadband PLN metrics from a) the PoRT testbed with the SLD routed through a PM fiber (polarization 1), and b) the PoRT testbed with the SLD routed through a 90° polarization rotating fiber (polarization 2). $\eta_{p_{\text{peak}}}$ refers to the maximum throughput of each map, or what the throughput of a planet at the location of maximum coupling would be. For the nulled ports, η_s refers to the throughput at the coaxial center of the lantern, corresponding to the stellar leakage. For the nulled ports, we also report the ratio $\eta_s/\eta_{p_{\text{peak}}}$.

	LP 01	**LP 11a**	**LP 11b**	**LP 21a**	**LP 21b**	**LP 02**
$\eta_{p_{\text{peak}}}$ (Pol. 1)	0.438	0.404	0.374	0.132	0.124	0.177
$\eta_{p_{\text{peak}}}$ (Pol. 2)	0.451	0.387	0.386	0.133	0.129	0.205
η_s (Pol. 1)	N/A	7.98×10^{-3}	3.46×10^{-2}	2.55×10^{-3}	1.26×10^{-3}	N/A
η_s (Pol. 2)	N/A	8.54×10^{-3}	3.24×10^{-2}	2.52×10^{-3}	1.34×10^{-3}	N/A
$\eta_s/\eta_{p_{\text{peak}}}$ (Pol. 1)	N/A	2.06×10^{-2}	8.97×10^{-2}	1.92×10^{-2}	9.79×10^{-3}	N/A
$\eta_s/\eta_{p_{\text{peak}}}$ (Pol. 2)	N/A	2.21×10^{-2}	8.40×10^{-2}	1.89×10^{-2}	1.04×10^{-2}	N/A

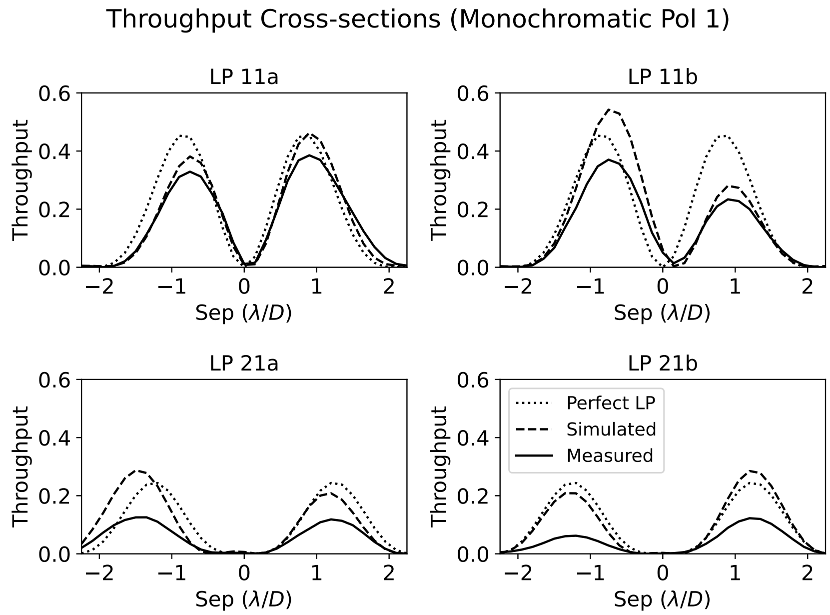

Figure 3.7: Simulated and measured throughput cross-sections (as indicated in Figure 3.6) for the nulled ports of the PLN. For comparison, the ideal throughput cross-sections with perfect LP modes is also plotted. The measured PLN throughput diverges from the ideal PLN because of imperfect mode shapes. However, considering the overall throughput losses through the lantern reported in Table 3.1, the shapes of the measured throughput maps agree well with the predictions simulated using the measured modes. The shown measurements are with the monochromatic laser in the first polarization state, but the other measured profiles lie closely to the same curve.

smaller coherence length ($\sim$ 10 microns), making matching path lengths practically infeasible.

Additional Throughput Maps and Polarization Comparison

We present the throughput maps measured using monochromatic light in the second polarization state (with the 90° polarization rotating fiber) in Fig. 3.8, using broadband light in the first polarization state (with the PM fiber) in Fig. 3.9, and using broadband light in the second polarization state in Fig. 3.10.

In Fig. 3.11, we plot the difference in the finely-sampled central throughput maps between the two orthogonal polarization states (the maps for polarization 1 are subtracted from the maps for polarization 2).

When observing the key measurements in Table 3.2 and 3.3, we find that the differences between the two polarization states are slight for most of the ports.

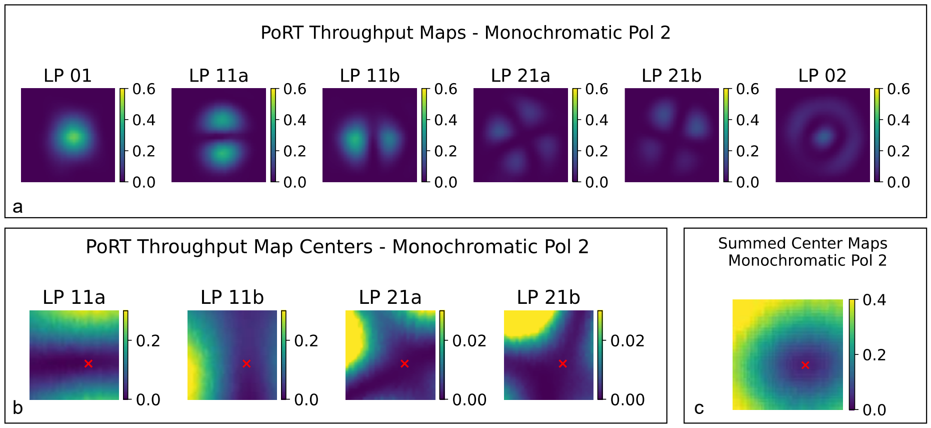

Figure 3.8: Monochromatic PLN throughput maps measured with 1568.772 nm light from the TLX2 tunable narrow linewidth laser injected into the PoRT testbed with a 90° polarization rotating fiber, such that the polarization is orthogonal to that of Fig. 3.6. a) Throughput maps of all ports across the PLN field of view. b) Throughput maps of the nulled ports with fine spatial sampling of the center (note that the LP 21 maps are on a different color scale from the LP 11 maps). The red crosses indicate the axial center of the lantern, identified using the map in part (c). c) The summed throughput of the four maps in part (b). The location of minimum summed throughput is taken to be the lantern center, where η_s is measured.

However, the monochromatic null-depth of the LP 21a port shows a stark difference, with the second polarization state exhibiting a null-depth that is almost an order of magnitude deeper than the first. This result suggests that the lantern mode shapes might depend on the polarization at some level. The limitations that these polarized differences impose on the null that wavefront sensing and control can achieve should be examined as part of future work.

One caveat is that the throughput maps presented were measured sequentially, not simultaneously. However, the measurements for polarization state 1 and for polarization state 2 were obtained about several hours apart, over which the testbench is relatively stable. We confirmed this stability by measuring the finely sampled LP21b map in both polarization states, one immediately after the other, and confirming that the difference map is consistent (in both features and magnitude) with the difference map obtained using the original data taken hours apart.

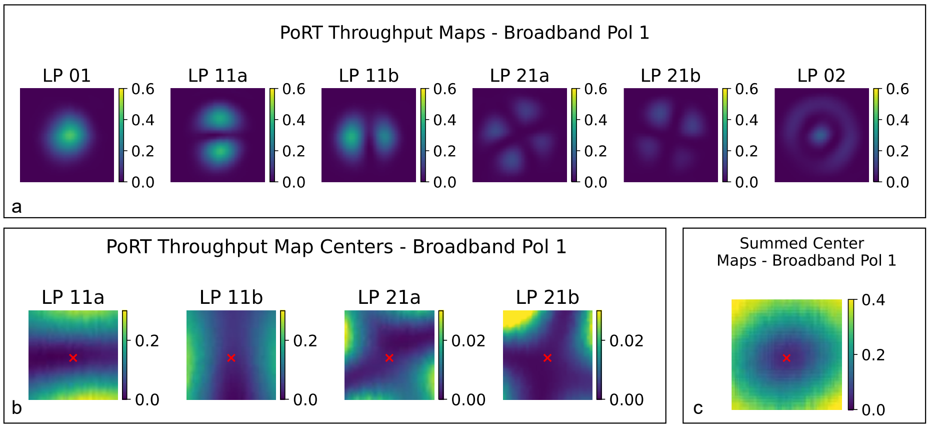

Figure 3.9: Laboratory PLN results measured on PoRT, using an SLD light source from 1450 nm to 1625 nm injected with a PM fiber. a) Throughput maps of all ports across the PLN field of view. b) Throughput maps of the nulled ports with fine spatial sampling of the center (note that the LP 21 maps are on a different color scale from the LP 11 maps). The red crosses indicate the axial center of the lantern, identified using the map in part (c). c) The summed throughput of the four maps in part (b). The location of minimum summed throughput is taken to be the lantern center, where η_s is measured.

3.4 Discussion and Future Work

In Table 3.2, we also include the values for $\eta_{p_{\text{peak}}}$, η_s, and $\eta_s/\eta_{p_{\text{peak}}}$ obtained from the monochromatic simulation using OAH modes. The simulated value of $\eta_s/\eta_{p_{\text{peak}}}$ assumes no wavefront error, so is indicative of the limit imposed by the modal impurities of the lantern. For the nulled port measurements, $\eta_s/\eta_{p_{\text{peak}}}$ is generally higher on the testbed than in simulation. The difference suggests that the wavefront of the PoRT testbed, optimized for injection through a single-mode fiber, is still not perfectly flat. This is likely due to a combination of limited calibration precision, uncalibrated higher order wavefront modes, and testbed drifts between the time of SMF calibration and the time the PLN measurements were taken. However, a perfectly flat wavefront is not necessarily the optimal wavefront for a real PLN with modal impurities, as the deformable mirror can be used to partially compensate for the modal impurities of the lantern, and improve the null even beyond that predicted by a flat wavefront. For example, with monochromatic light in the second polarization, we measure an η_s that is an order of magnitude than predicted by

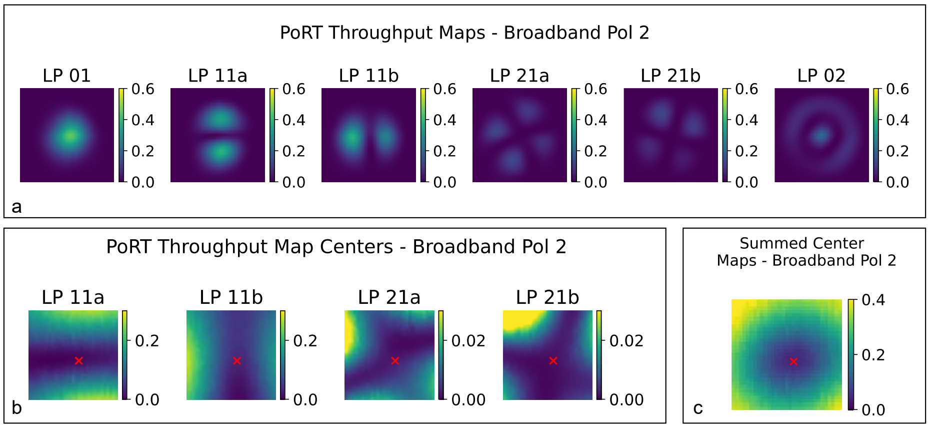

Figure 3.10: Laboratory PLN results measured on PoRT, using an SLD light source from 1450 nm to 1625 nm injected with a 90° polarization rotating fiber. a) Throughput maps of all ports across the PLN field of view. b) Throughput maps of the nulled ports with fine spatial sampling of the center (note that the LP 21 maps are on a different color scale from the LP 11 maps). The red crosses indicate the axial center of the lantern, identified using the map in part (c). c) The summed throughput of the four maps in part (b). The location of minimum summed throughput is taken to be the lantern center, where η_s is measured.

simulation. This suggests that the wavefront error in the system is interfering with the lantern mode in a way that deepens the stellar null. This also suggests that the difference in mode shape between the two polarizations is causing a difference in η_s on the order of $\sim 10^{-3}$. Exploring using wavefront sensing and control schemes with the PLN, including in the presence of polarization differences, is left for future work. Although the MSPL does not lend itself to linear wavefront control (Lin, Fitzgerald, et al., 2022b), it is a good candidate for data driven control such as Implicit Electric Field Conjugation (S. Y. Haffert, Males, et al., 2023a) because of the low dimensionality from its limited number of ports.

The $\eta_s/\eta_{p_{\text{peak}}}$ values of $\sim 10^{-2}$ currently achieved by the PLN are approximately the same as the null-depth the VFN achieves on sky, limited by the adaptive optics residuals of the Keck II telescope (Echeverri, J. Xuan, et al., 2023). Under these conditions, the VFN was able to tentatively detect a companion with flux ratio of 1/400 (Echeverri, J. W. Xuan, et al., 2024). Given that the PLN is expected to have even higher throughput, it could be used — as is — on ground-based telescopes, with

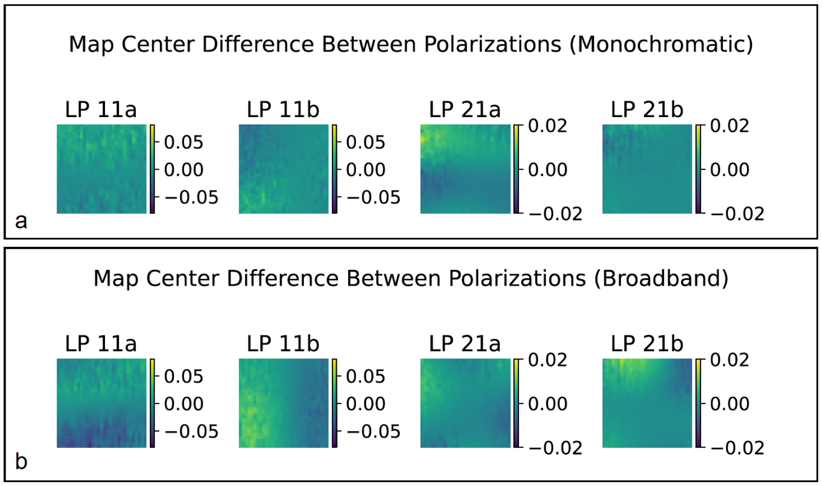

Figure 3.11: Difference in the finely-sampled central throughput maps between two orthogonal polarization states, with throughput maps for polarization 1 subtracted from the maps for polarization 2, using a) monochromatic light and b) broadband light. The monochromatic null-depth of the LP 21a port with the second polarization state is almost an order of magnitude deeper than the first, suggesting that the lantern mode shapes do depend on the polarization at some level. The limitations that these polarized differences impose on the null that wavefront sensing and control can achieve should be examined as part of future work.

signal-to-noise expected to exceed that of the VFN's. Additionally, the PLN has the potential to partially constrain the planet's location, a capability the VFN lacks (Xin, Nemanja Jovanovic, et al., 2022), as well as the potential for image reconstruction similar to what has been explored for a non-mode-selective photonic lantern in Kim, Fitzgerald, Lin, Sallum, et al. (2024). Thus, future work includes testing and validating the PLN on-sky, as it can already be used to obtain scientifically useful observations of close-in giant exoplanets.

Like the VFN, the PLN's null-depths on-sky at Keck II would be expected to be limited by the residuals of the AO system. However, given a better adaptive optics system (or a space environment) the modal impurity or 'cross-talk' of the lantern will become the dominant limiting factor. Wavefront control with one DM can only partially compensate for the cross-talk, as it would only modulate phase and not amplitude in the pupil plane. Wavefront control with two DMs is still chromatic, and thus expected to achieve nulls with smaller bandwidths than nulls resulting purely from the lantern's mode-selectivity. Therefore, more work should be invested in creating MSPLs with lower levels of cross-talk, as this would result in deeper, naturally broadband nulls that would relieve additional demand on the wavefront control system (Nemanja Jovanovic et al., 2023).

Lastly, although the lantern we use in this work is optimized for 1550 nm, silica-based mode-selective lanterns can be designed to operate at different wavelengths, ranging from the visible spectrum up to about 2 um. However, other fiber technologies (Price et al., 2006) would be needed to access wavelengths outside of this range.

3.5 Conclusion

In this work, we characterize the properties of a MSPL optimized around 1550nm, and perform the first laboratory demonstration of a PLN. We measure the throughput maps for each port (in two polarizations for both monochromatic and broadband light) and calculate the null-depths of the nulled ports, which are around the 10^{-2} level. We find that the mode shapes measured using off-axis holography can be used in simulations to model and predict the behavior of a real PLN. Future work involves using wavefront sensing and control to further improve the null-depths, as well efforts towards improving the modal purity of the lanterns themselves. In the meantime, the photonic lantern nuller already achieves broadband nulls suitable for observing young gas giants at the diffraction limit using ground-based observatories, a capability that should be tested on-sky.

Chapter 4

IMPLICIT ELECTRIC FIELD CONJUGATION WITH THE PHOTONIC LANTERN NULLER

Xin, Yinzi et al. (2025). "Implicit electric field conjugation with a photonic lantern nuller". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 11.2, p. 025004. DOI: `10.1117/1.JATIS.11.2.025004`. URL: `https://doi.org/10.1117/1.JATIS.11.2.025004`.

This chapter is a reproduction of my paper on adapting a wavefront control technique originally developed to create focal plane dark zones with coronagraphs to instead improve the nulls of a PLN. The implicit Electric Field Conjugation (iEFC) algorithm uses a deformable mirror to probe the existing starlight and apply a correction, minimizing it over successive iterations. In this work, I showed that using iEFC with a PLN can reduce null-depths from the 10^{-2} level to the 10^{-4} level for three of the four ports simultaneously, while the null of last port remains largely stagnant. I use simulations to investigate the source of this behavior, finding that it can be explained by the lantern cross-coupling manifesting partially as amplitude aberration in the focal plane, where the deformable mirror can only apply phase corrections. I find that solutions that simultaneously null all ports are physically possible (the coupling integral in this case allows for phase modulations to compensate for amplitude aberrations, which is not possible with coronagraphs), but that the pair-wise probing approach used in iEFC does not result in the identification of such solutions. Modifying the probing approach to enable more comprehensive sensing of potential solutions through mode-sorting instruments, such as the PLN, presents an interesting topic for future work.

Abstract

The Photonic Lantern Nuller (PLN) is an instrument concept designed to characterize exoplanets within a single beam-width from its host star. The PLN leverages the spatial symmetry of a mode-selective photonic lantern (MSPL) to create nulled ports, which cancel out on-axis starlight but allow off-axis exoplanet light to couple. The null-depths are limited by wavefront aberrations in the system as well as by imperfections in the lantern. We show that the implicit electric field conjugation

algorithm can be used to reduce the stellar coupling through the PLN by orders of magnitude while maintaining the majority of the off-axis light, leading to deeper null depths ($\sim 10^{-4}$) and thus higher sensitivity to potential planet signals. We discuss a theory for the tradeoff we observed between the different ports, where iEFC improves the nulls of some ports at the expense of others, and show that targeting one port alone can lead to deeper starlight rejection through that port than when targeting all ports at once. We also observe different levels of stability depending on the port and discuss the implications for practically implementing this technique for science observations.

4.1 Implicit Electric Field Conjugation

In this work, we present the results of using wavefront control — specifically the implicit electric field conjugation algorithm (iEFC) (S. Y. Haffert, Males, et al., 2023a) — to deepen the central nulls of the PLN.

The implicit electric field conjugation algorithm for active suppression of starlight is described in S. Y. Haffert, Males, et al. (2023a). We present a simplified overview of it here.

The stellar electric field can be modulated by applying probes on the deformable mirror (DM), and the electric field is linearly related to the difference between an image with some probe and the image with the same probe but with opposite sign. Minimizing the measurement δ — a series of such 'differenced' images — thus also minimizes the electric field. We can empirically calibrate the influence of the DM on δ by applying a mode on the DM and encoding the change in δ that it produces into a response matrix:

$$\delta = S\alpha, \tag{4.1}$$

where α is the DM command and S the calibrated response matrix. The basis set of modes whose influence is sequentially measured and encoded into S are referred to as the "calibration modes."

The iEFC solution for the desired DM input is given by

$$\alpha = \arg\min_{\alpha} |\delta + S\alpha|^2 + \lambda|\alpha|^2 = -(S^T S + \lambda I)^{-1} S^T \delta = -C\delta. \tag{4.2}$$

The parameter λ can be set to penalize large DM solutions, and the control matrix $C = (S^T S + \lambda I)^{-1} S^T$ is computed ahead of time, after calibrations are complete.

Unlike the alternative electric field conjugation (EFC) algorithm (Give'on et al., 2007), which is model-based, iEFC is data-driven. It has the advantage of not being limited by model fidelity but also the disadvantage of requiring testbed calibration time. We choose to use iEFC, however, because it is relatively more advantageous for the PLN than for conventional coronagraphs, as the PLN has a much more limited field-of-view (FOV) and thus requires significantly fewer modes to be calibrated than for a coronagraphic dark hole. The number of calibrated modes required scales linearly with the area of the FOV, which scales quadratically with FOV radius, so the iEFC calibration overhead for the PLN (with a FOV radius of $\sim 2\lambda/D$) compared to a typical coronagraphic dark hole (extending to $\sim 10\lambda/D$) is smaller by a factor of $2^2/10^2$, or about 0.04.

For this work, we use the implementation of iEFC from the lina package (Milani, Derby, and Douglas, n.d.).

4.2 Experimental Setup

A detailed schematic of the front-end of Polychromatic Reflective Testbed (PoRT) can be found in Xin, Echeverri, et al. (2024), along with the PLN coupling maps measured using the testbed without wavefront control — i.e. the DM merely flattened by maximizing coupling through a single-mode fiber. The results obtained in Xin, Echeverri, et al. (2024) were obtained using a photodiode with only one input, and therefore had to be measured sequentially. However, for wavefront control, it is convenient to be able to measure the coupling through all the relevant ports simultaneously.

In Figure 4.1a, we present an updated simplified diagram of PoRT, which now includes a back-end that images the outputs of the nulled ports onto a camera, allowing the fluxes coupled into each port to be simultaneously measured using photometry. As in Xin, Echeverri, et al. (2024), we use a Thorlabs TLX2 laser set to a wavelength of 1568.772 nm injected into the bench with a polarization maintaining fiber, and use a tunable iris to set the $F\#$ of the beam being injected into the lantern to maximize the throughput into the LP 11 ports, resulting in $F\# = 6.2$. Note that one polarization of light is used for this experiment. Xin, Echeverri, et al. (2024) showed that the lantern exhibited polarized differences at the level of a few 10^{-3} (in coupled intensity) between two orthogonal polarizations of light. With

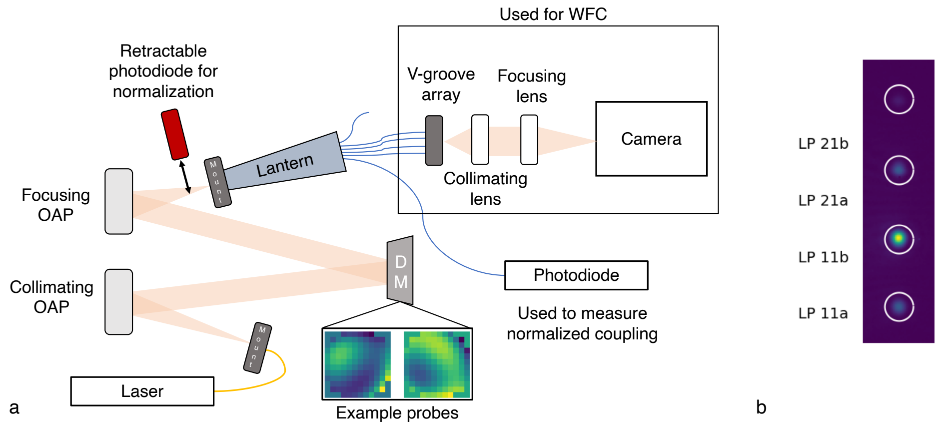

Figure 4.1: a) A simplified diagram of the experimental setup. A monochromatic 1568.772 nm laser is injected into the bench, and the beam is collimated. A 12×12 deformable mirror can be used to manipulate the wavefront of the beam before it is focused onto the lantern ; the inset shows example DM probe modes p_{1_+} and p_{2_+}. The SMF outputs of the lantern can then be routed to either a V-groove array to be imaged onto the camera, or to the photodiode. The photodiode is calibrated to a photometer that can slide into the beam just before the lantern, and thus provides measurements normalized to the incoming beam. While performing wavefront control, all four nulled ports are routed to the V-groove, with the non-nulled ports disconnected in order to not saturate the camera. After performing wavefront control, we sequentially route the nulled outputs to the coupling photodiode to obtain coupling maps for each port. The coupling photodiode is calibrated to a photodiode that can be inserted before the lantern in order to normalize the coupling maps to the incident light on the lantern, providing normalized measurements of η. b) The image on the camera cropped to the region of interest. The white circles indicate the photometric apertures used to measure coupling through the lantern's nulled ports.

conventional coronagraphs, these polarized differences would limit the achievable contrast due to different DM solutions being needed for each polarization (Ashcraft et al., 2025). However, the overlap integral of the PLN allows for more nulling degrees of freedom than conventional coronagraphs with a focal-plane detector (see 4.4). This potentially enables joint solutions that null both polarizations simultaneously despite the differences in mode shape, a study that is left for future work.

Figure 4.1b shows an example camera image, cropped to the region containing the lantern outputs and with the dark frame (the camera image when the light source is turned off) subtracted. The overlaid circles indicate the apertures used for photometry, where each intensity measurement is the sum of the counts contained within the defined circle. When the incoming beam is aligned to the 'center' of the lantern (which we choose as the location of minimum summed coupling through the nulled ports), these intensity measurements correspond to the stellar coupling through each port of the lantern. These intensity measurements are proportional to η_s, or the fraction of the incoming starlight coupled into each port. For closed-loop wavefront control, all four nulled ports are routed to the V-groove, and we work directly with the intensity measurements made on the camera (i.e. the sum of the counts within each defined aperture). We leave the non-nulled ports disconnected in order to not saturate the camera. After performing wavefront control, we sequentially route the nulled outputs to the coupling photodiode to obtain coupling maps for each port. The coupling photodiode is calibrated to a photodiode that can be inserted before the lantern in order to normalize the coupling maps to the incident light on the lantern, providing normalized measurements of η.

4.3 Implementation and Results

Example targeting all nulled ports

We define the following two sets of probes, where Z_n is the nth Noll-ordered Zernike mode defined across the 12×12 DM actuator grid:

$$p_{1_\pm} = \pm(Z_5 + Z_6 + Z_7 + Z_8)/\sqrt{4} \tag{4.3}$$

$$p_{2_\pm} = \pm(Z_5 - Z_6 + Z_7 - Z_8)/\sqrt{4}. \tag{4.4}$$

The probe mode shapes of p_{1_+} and p_{2_+} are shown in the inset of Fig. 4.1a. This choice of probe is motivated by the fact that the LP 11 modes are primarily sensitive to Coma (Z = 5 and 6) while the LP 21 modes are primarily sensitive to Astigmatism (Z = 7 and 8). In fact, simulations show that the dominant modes of S when sensed using *completely random* probes that modulate the whole DM, as obtained through a singular-value decomposition (SVD), are indeed Coma and Astigmatism first, followed by higher-order irregular modes. We find that in the presence of noise, however, using p_1 and p_2 as defined results in much better signal and iEFC performance than using more complex probes.

Using a probe amplitude of 0.02 and a calibration mode (i.e. the basis vectors for representing DM inputs) amplitude of 0.01 (both in DM control units that range

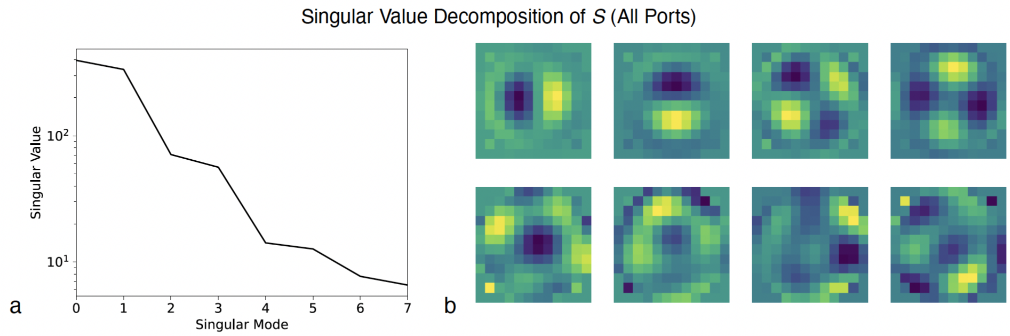

Figure 4.2: a) The singular values of the response matrix S when measuring all four nulled ports. b) The corresponding singular modes in DM space, arranged in order of descending singular value.

from 0 to 1, mapping to the control voltage range of the DM, which has a 3.5 μm stroke), we then calibrate the response matrix S across the set of Zernikes modes from Z_4 to Z_{30} (a total of 27 modes). The singular values and singular modes of S in DM space (arranged in order of descending singular value) are displayed in Fig. 4.2, showing the dominance of Coma and Astigmatism in the instrumental response. From S, we calculate the control matrix C using λ equal to 1/10 of the maximum diagonal value of $S^T S$.

Figure 4.3a shows an example closed-loop iEFC run, which significantly deepens the nulls in three of the ports, while the last null ends up approximately the same as it started. Normalized coupling maps are obtained using the photodiode, both with the original DM map and with the DM map after running iEFC, and are presented in Figure 4.4. The metrics of stellar coupling, planet coupling, and null depth before and after iEFC are presented in Table 4.1, and select cross-sections of the relative signal-to-noise (S/N) ratio, given by $\eta_p/\sqrt{\eta_s}$, before and after iEFC are also presented in Figure 4.5.

This experiment shows a typical tradeoff that we observe with iEFC with the PLN, where certain nulls become degraded to deepen others if it reduces the stellar coupling overall. In this case, the LP 11 nulls have likely reached the limit imposed by the stability of the testbed, as measurements of the null with the loop open show significant fluctuations on the timescale of an iEFC measurement (Fig. 4.3c). Meanwhile, the LP 21 ports have not reached the limit imposed by testbed stability. This can be demonstrated by running iEFC on them individually as in Section 4.3, which results in deeper nulls in those ports than when running iEFC on all four

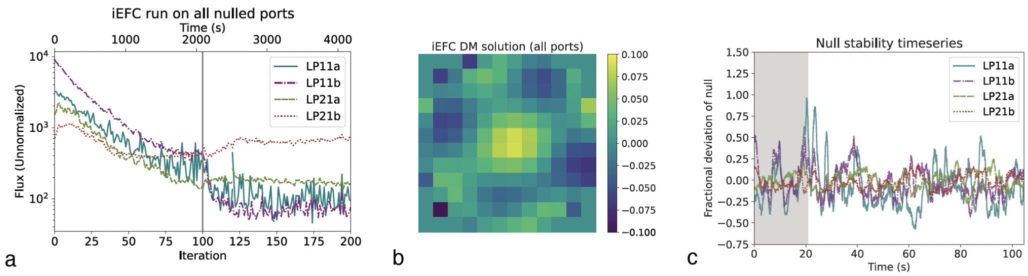

Figure 4.3: a) An example iEFC run targeting all four nulled ports. The vertical gray line indicates a change in camera integration time along with a control matrix recalibration. The final null depths from this run are compared to the initial null depths in Table 4.1. b) The DM solution found by iEFC targeting all four nulled ports at once. The root-mean-square (RMS) of the DM control values is 0.033, corresponding to approximately 120 nm of stroke. c) A timeseries of the fractional fluctuations in the null obtained with iEFC, spanning the timescale of one iEFC iteration. The gray shaded region corresponds to the timescale of one photometric measurement, and each iEFC iteration takes five photometric measurements: four for the probes and one without any probe. The LP 11a and LP 11b nulls exhibit significant fluctuations over the timescale of an iEFC iteration, showing that these nulls are likely limited by testbed instability and not the lantern itself.

Table 4.1: The values of $\eta_{p_{peak}}$ (the peak planet coupling, or the maximum coupling value for each port) and η_s (the stellar coupling) for the four nulled ports before and after performing iEFC, targeting all nulled ports at once. Also shown is the null depth ($\eta_s/\eta_{p_{peak}}$) for each port before and after performing iEFC.

	LP 11a	**LP 11b**	**LP 21a**	**LP 21b**
$\eta_{p_{\text{peak}}}$ (Before)	0.394	0.358	0.131	0.112
$\eta_{p_{\text{peak}}}$ (After)	0.246	0.224	0.108	0.0715
η_s (Before)	5.34×10^{-3}	1.65×10^{-2}	3.70×10^{-3}	1.31×10^{-3}
η_s (After)	5.80×10^{-5}	8.10×10^{-5}	2.35×10^{-4}	1.15×10^{-3}
$\eta_s/\eta_{p_{\text{peak}}}$ (Before)	1.35×10^{-2}	4.62×10^{-2}	2.83×10^{-2}	1.16×10^{-2}
$\eta_s/\eta_{p_{\text{peak}}}$ (After)	2.36×10^{-4}	3.62×10^{-4}	2.18×10^{-3}	1.61×10^{-2}

ports at once. In Section 4.4, we discuss simulations that provide more insight into this observed tradeoff, as well as the predicted behavior of iEFC under various conditions that are not realizable on the actual testbed.

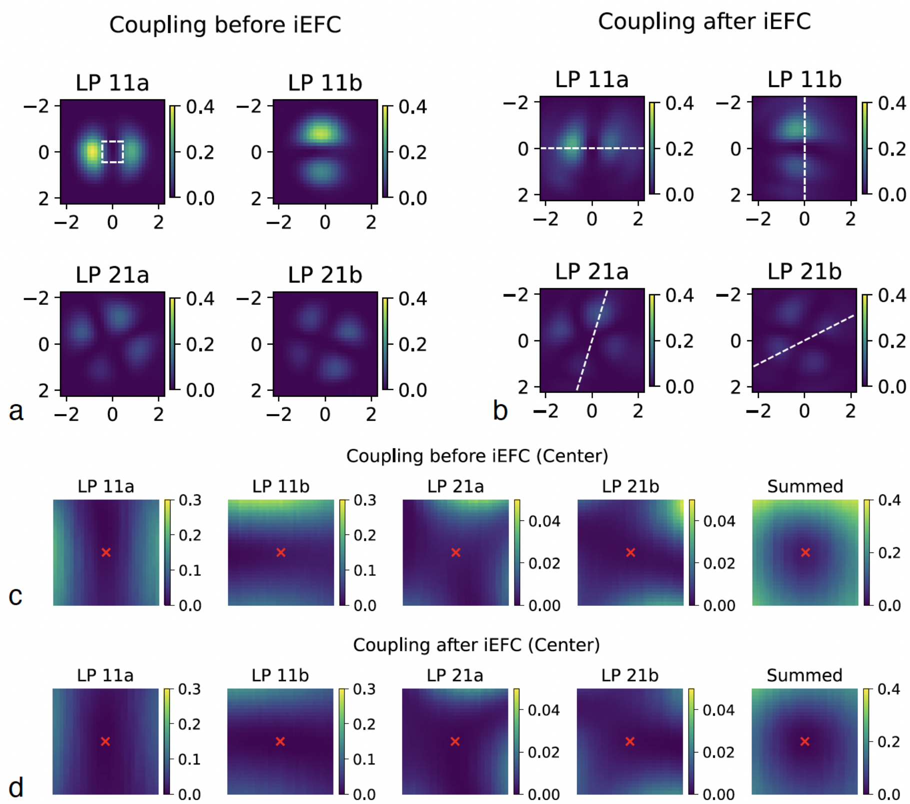

Figure 4.4: a) The normalized coupling maps through the nulled ports with the original DM surface, with spatial extent in units of λ/D. The white dashed box in the first panel indicates the finely-sampled region shown in (c) and (d) b) The normalized coupling maps using the DM solution found with iEFC, with spatial extent in units of λ/D. The dashed white lines indicate the locations of the S/N ratio cross-sections shown in Fig. 4.5. c) Finely-sampled coupling maps of the lantern center with the original DM surface, spanning 1/5 of the spatial extent in part (a). d) Finely-sampled coupling maps of the lantern center using the DM solution found with iEFC, spanning 1/5 of the spatial extent in part (a). The red crosses indicate the location where the beam is aligned for the camera measurements, and also where η_s is measured. We observe that iEFC spatially redistributes the coupling values through the lantern, lowering the coupling in the middle of the field of view and causing a diffuse extension of the coupling distribution at larger separations.

Example targeting a single port

For these runs targeting a single port, LP 21b, we use a different set of probes that better captures the effect of the control modes on this port. Simulations show that while the most dominant mode of S in this case is Astigmatism, as expected, the second dominant mode has a complex and irregular shape (somewhat resembling the eventual simulated DM solution), and thus sensing it well requires probes comprised

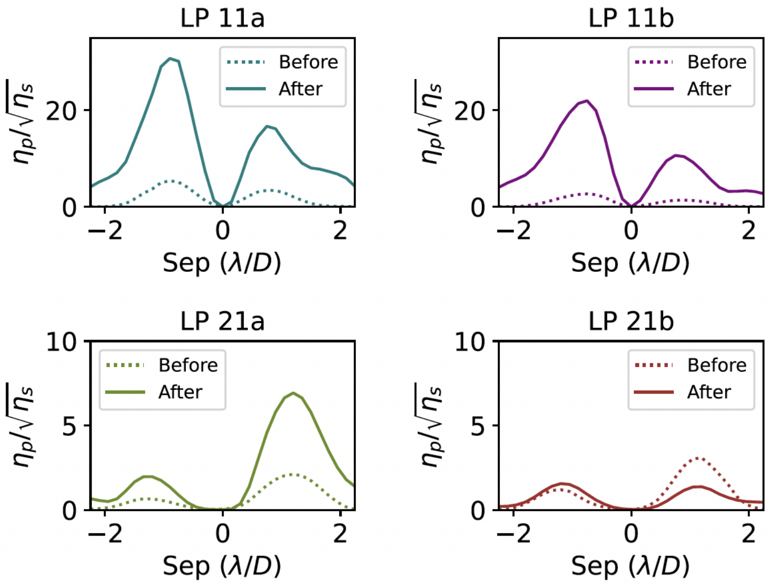

Figure 4.5: A comparison of the relative S/N ratio $\eta_p/\sqrt{\eta_s}$ along the cross-sections indicated in Fig. 4.4b, before and after performing iEFC. In this case, wavefront control significantly improves the S/N ratio for three of ports while slightly degrading it in one. A theory for the observed tradeoff between ports is discussed in Section 4.4.

of many Zernikes:

$$p_{3_\pm} \propto \pm \sum_{n=4}^{30} Z_n \tag{4.5}$$

$$p_{4_\pm} \propto \pm \sum_{n=4}^{30} (-1)^n Z_n. \tag{4.6}$$

In this case, $p_{3_\pm}$ and $p_{4_\pm}$ are each normalized to have an RMS of 1 before being multiplied by the desired probe amplitude in DM control units.

Figure 4.6 shows an example iEFC run targeting just the LP 21b port (i.e. the corresponding control matrix C is calculated using a response matrix S that includes only measurements from the LP 21b port). A probe amplitude of 0.02 is used, and a final η_s of 7.42×10^{-5} is achieved, significantly deeper than reached during the experiment targeting all four ports at once. For this experiment, the final η_s is extracted from camera photometry and converted to a normalized coupling value based on the calibrated initial measurement. This is because we observed the

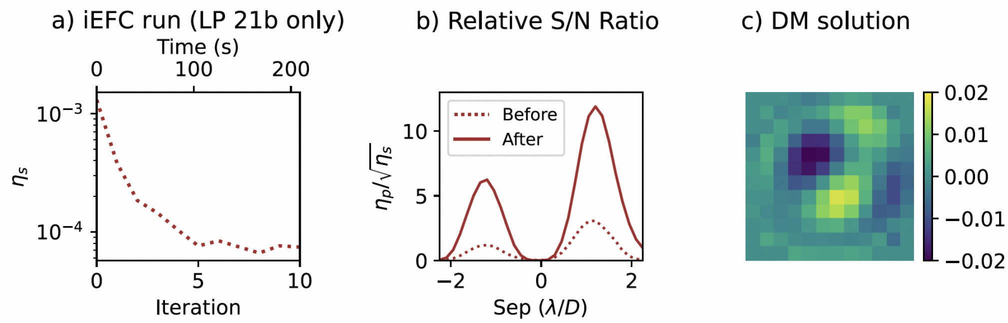

Figure 4.6: a) An example iEFC run targeting only the LP 21b port, showing significant improvement in the null. The final $\eta_s = 7.42 \times 10^{-5}$ is extracted from camera photometry and converted to a normalized coupling value based on the calibrated initial measurement. b) The relative S/N ratio ($\eta_p/\sqrt{\eta_s}$) along the cross-section indicated in Fig. 4.4b, before and after running iEFC. The peak coupling $\eta_{p_{peak}}$ is 0.110, giving a null depth $\eta_s/\eta_{p_{peak}}$ of 6.73×10^{-4}. c) The DM solution found by iEFC targeting only the LP 21b port. The RMS of the control values is 0.0064, corresponding to approximately 22 nm of stroke, five times smaller than the RMS of the control values for the solution targeting all four ports.

null degrade significantly in the time it took to measure a coupling map with the photodiode (approximately 30 minutes). This drift in the null also limits our ability to push to deeper nulls by relinearizing the response matrix, since the null also significantly degrades during the calibration process. A model-based algorithm (Give'on et al., 2007) or an algorithm that recalibrates on-the-fly may be able to avoid this problem.

However, the peak planet coupling $\eta_{p_{peak}}$ is not significantly impacted by these drifts, so we measure it using the photodiode, obtaining $\eta_{p_{peak}} = 0.110$. The retained planet throughput shows that the iEFC solution is less aggressive when targeting only one port, and ultimately achieves a null depth $\eta_s/\eta_{p_{peak}}$ of 6.73×10^{-4}. The S/N ratio cross-sections before and after iEFC are presented in Figure 4.6b, showing that the limitation observed in Fig. 4.3 is not due to this port in particular, but is a consequence of targeting all four ports at once, a phenomenon discussed further in Section 4.4.

4.4 Discussion

We conducted several simulations to further explore and corroborate the behavior observed on the testbed, elucidating some notable properties of performing iEFC with a PLN. Simulations were conducted with both a realistic lantern model (using the measured modes of the real lantern) as well as with a 'perfect' lantern model with ideal LP modes. The first set of simulations and their key takeaways are as follows:

1. Simulating iEFC through a realistic lantern model corroborates the behavior observed on the testbed, where the LP 11ab and LP 21a ports improve at the expense of the LP 21b port. However, both LP 11ab ports can now reach an η_s of 2×10^{-5} simultaneously (even deeper if traded-off against each other by relinearizing at different times), likely because there is no instability or noise in the simulation. Meanwhile, the LP 21a port bottoms out at a shallower $\eta_s \approx 2 \times 10^{-3}$, perhaps a tradeoff related to the deeper LP 11 nulls.

2. Simulating iEFC through a perfect lantern model with static phase-only aberrations in the pupil shows all four ports reaching extremely deep contrasts ($\eta_s \sim 10^{-9}$ to 10^{-12} in the absence of other noise, depending on the exact aberration). The static phase aberration in this case has a peak-to-valley (PTV) value of 0.2 radians.

3. Simulating iEFC through a perfect lantern model with static phase and amplitude aberrations (0.1 radians PTV each) in the pupil shows that the ports are now limited to shallower nulls ($\eta_s \sim 10^{-5}$ to 10^{-7} in the absence of other noise, with an observed tradeoff amongst the ports that depends on the exact aberration).

Back-propagating the measured modes and the ideal LP modes into the pupil plane and taking their difference shows that the cross-coupling of the lantern from manufacturing imperfections appears as both phase and amplitude deviations in the pupil plane. These simulations thus suggest that the cross-coupling manifesting partially as amplitude aberration in the pupil plane is what prevents all four ports from being nulled.

However, the typical coronagraphic solution of adding a second DM to correct for amplitude errors is not a necessary or even particularly effective solution in this case. The observed behavior does not seem to be a fundamental limitation of physics, but rather a unique property of mode-sorters such as the PLN and how iEFC senses the electric field through the ports. For example, simulating model-based EFC with a realistic lantern model given perfect knowledge of the electric field shows that one

DM is, in fact, capable of compensating for cross-coupling down to $\eta_s < 10^{-11}$ and beyond. Therefore, one DM has enough *control* authority to compensate for both phase and amplitude errors in the pupil, e.g. by manipulating phase to balance out amplitude asymmetries such that the overlap integral is still zero.

The limitation, therefore, lies with the sensing: as mentioned in Section 4.3, an SVD of the iEFC response matrix S shows that the most dominant modes are approximately Coma X&Y (primarily impacting the LP 11 ports), followed by 0° and 45° Astigmatism (primarily impacting the LP 21 ports), followed by weaker irregular modes. Fully compensating for amplitude asymmetries would require higher spatial frequency modes — as is observed in the DM solution obtained with EFC given perfect knowledge of the electric field; however, these modes are not well-sensed by the current pairwise probing approach, which mainly captures the effect of the most dominant low order modes.

While improvements in the lantern manufacturing process may reduce cross-coupling to the degree that this is no longer a problem at the contrasts required for ground-based astronomy, these sensing considerations will likely remain relevant at the higher contrasts required for space telescope instrumentation. Therefore, future work includes exploring alternative methods for sensing the electric field, and building response matrices in a way that addresses the unique physical characteristics of mode-sorters, where the outputs are not pixels directly sensing the local intensity but rather ports whose coupling is determined by an overlap integral. This work would be applicable to the PLN and as well as other designs such as the Single-mode Complex Amplitude Refinement (SCAR) coronagraph (Por and S. Y. Haffert, 2020; S. Y. Haffert, Por, et al., 2020).

Meanwhile, the experiment targeting the LP 21b port alone is limited by a rapid drift in the null. This behavior differs from those of the LP 11 nulls when targeting all four ports at once, which exhibit large fluctuations but very little degradation over time. Based on theoretical sensitivities to aberrations (Xin, Nemanja Jovanovic, et al., 2022), we believe this is because the LP 11 nulls are primarily affected by zero-mean tip-tilt fluctuations, while the LP 21 nulls are primarily affected by higher order modes (as can be inferred from the DM solution in Fig. 4.6c) that drift over time.

The combined behavior of instrument sensitivity and the aberrations' spatial and temporal characteristics, the layout of the adaptive optics (AO) system, and whether the planet location is already known — these factors all have implications for the use of iEFC for on-sky observations. For example, if the planet location is already known, then it is more advantageous to target one or two ports, maximizing sensitivity at the planet's location. However, for a survey, or for a planet whose location is not well-constrained, it would be necessary to target either both LP 11 ports, both LP 21 ports, or all four ports at once in order to obtain more coverage of the sky.

The choice of which ports to target also depends on whether the AO system is set up to perform iEFC in real time during the observation. Some systems, such as the Keck Planet Imager and Characterizer (KPIC) (Delorme et al., 2021b), have a dedicated second-stage DM that is not seen by the primary AO wavefront sensor. This means that the DM can be changed during the course of the observation without impacting the AO loop, as is done by KPIC to perform speckle nulling through the single-mode fiber (Xin, J. W. Xuan, et al., 2023), a control architecture that would allow the iEFC loop to run independently on sky. However, many instruments do not have this extra DM, so any changes to the DM made by the iEFC loop have to be accounted for with an 'offset' to the primary AO wavefront sensor. This process adds significant complexity, and though it has been demonstrated on sky with a Shack-Hartmann wavefront sensor (Potier et al., 2022), it is more difficult with Pyramid wavefront sensors, with which it has not yet been reliably implemented on-sky. Although any combination of ports could be targeted on sky using systems with a second-stage DM, since iEFC can be run in real time, for instruments without a second-stage DM, a more viable strategy would be to run iEFC during daytime calibrations, then observe in open loop. This method relies on the null being relatively stable, or primarily sensitive to aberrations that fluctuate about a zero mean in time, which can inform the choice of ports used (i.e. the LP 11 ports on the PoRT testbed, though this would depend on environmental conditions). Although there is no primary AO loop for space-based telescopes, similar interactions may arise with upstream low-order wavefront sensors (Pourcelot et al., 2023), and additional effort is needed to ensure that both the low-order and focal-plane loops can run concurrently.

An additional consideration for ground-based telescopes is whether to observe using pupil-tracking or field-tracking mode. Running iEFC actively on-sky would benefit from a more stable pupil-to-instrument relationship, but would require a more complicated data analysis approach to fit the rotationally-varying planet signal (Goyal et al. in prep). Depending on various factors (such as the speed of sky rotation, the rate of quasi-static speckle evolution, and the planet brightness and separation), the rotating planet signal (after post-processing with a modified version of Angular Differential Imaging (Christian Marois, Lafrenière, et al., 2006)) may or may not provide an overall gain in detection sensitivity compared to the static planet signal from field-tracking mode. Meanwhile, if iEFC is used to calibrate the PLN during the day to compensate for static bench and lantern imperfections without seeing the pupil, then either mode can be used depending on which provides the best post-processed performance, which is currently an open question. Additional work and empirical experience is needed to identify the best approach for scientific observations, including considerations of robustness and ease of use.

Additionally, this demonstration uses a laser due to the limitations of the photometric backend, but broadband nulls would be necessary for spectroscopic applications. Laboratory results presented in Xin, Echeverri, et al. (2024) indicate that the nulls of the PLN are naturally broadband at the $\eta_s \sim 10^{-3}$ level, and multi-wavelength control scales more advantageously for the PLN than for coronagraphs, as there are fewer degrees of freedom per wavelength that must be controlled by the DM. However, future work is needed to either demonstrate that monochromatic control can achieve a null that is sufficiently spectrally wide enough for science, or to demonstrate full multi-wavelength control.

4.5 Conclusion

In this work, we showed that implicit electric field conjugation, a focal-plane wavefront control method developed for coronagraphs, can also be used to improve the null depths of the PLN. We find that there is a tradeoff amongst the ports, where deepening some of the nulls can occur at the expense of the others, but overall, the null depths can be improved by one to two orders of magnitude for up to three of the four ports at once. We conduct simulations that corroborate and provide additional insight into the behavior observed on the testbed, as well as reveal an electric field sensing problem for mode-sorters that will be explored in future work. Future work also includes chromatically dispersing the outputs of the PLN, performing iEFC in broadband light, and characterizing the null depths achievable across the 10-20% bandwidths applicable to high-resolution spectroscopy science. Although the symmetries of the mode-selective photonic lantern prohibit it from being used as a linear wavefront sensor (Lin, Fitzgerald, et al., 2022a), it may be possible to design a hybrid lantern with some asymmetric ports and some symmetric ports, allowing for nulling and wavefront sensing simultaneously without the need for DM probes. Additionally, preparations for on-sky demonstrations of the PLN are also currently underway.

Chapter 5

ON-SKY DEMONSTRATION OF THE PHOTONIC LANTERN NULLER AT THE SUBARU TELESCOPE

In 2024, I had the opportunity to install the mode-selective photonic lantern on the Subaru Coronagraphic Extreme Adaptive Optics Instrument (SCExAO) at the Subaru Observatory in Hawaii, which has a spectrally dispersed backend with a resolving power of approximately $R = \lambda/\Delta\lambda = 500$. The first attempt at an on-sky demonstration in September 2024 was unsuccessful due to the difficulty in aligning the lantern on-sky in the presence of atmospheric turbulence, identifying the need for a more robust alignment procedure. A second opportunity to go on-sky in March 2025 was lost due to a motor failure that caused the injection carriage (which controls the $F\#$ of the beam) to become stuck very far from the correct position. This chapter describes the third attempt at an on-sky demonstration using a new calibration strategy: fixing the position of the lantern relative to the central pixel of the PSF monitoring camera during daytime calibrations and using a tip-tilt loop on-sky to drive the PSF to the same pixel. The new calibration strategy was unsuccessful due to an extreme misalignment between the lantern injection arm and the PSF monitoring camera, which occurred between the daytime calibrations and the on-sky observations and continued to drift even after the observation. However, the conditions were sufficiently good on-sky that the position of the lantern could be approximately recovered in real time. I took on-sky data of a bright star at three different lantern positions, then used this data to calculate null-depths for two of the ports, as well as to place constraints on the detection sensitivity to a hypothetical binary companion observed using reference differential imaging. This chapter thus presents the results of the first successful on-sky demonstration of the PLN.

5.1 The Subaru Coronagraphic Extreme Adaptive Optics Instrument

The Subaru Coronagraphic Extreme Adaptive Optics Instrument (SCExAO) is an extremely complex instrument with multiple modules operating at multiple wavelengths. For simplicity, only the elements relevant to this work will be described and shown in the schematic in Fig. 5.1. For a detailed description of the instrument, see Lozi et al. (2020).

SCExAO receives light from AO3k, the primary adaptive optics (AO) system of the

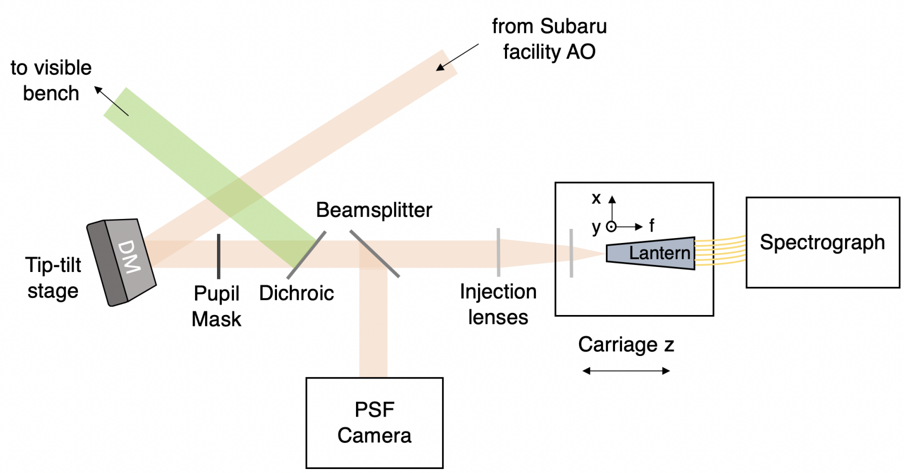

Figure 5.1: A simplified schematic of the elements of SCExAO relevant to photonic lanterns work. SCExAO receives light from AO3k, the primary adaptive optics system of the Subaru Telescope. The visible wavelength light is sent to a separate bench, which contains the AO wavefront sensor. The near-infrared light is sent to the photonics injection unit as well as a PSF monitoring camera. The injection unit consists of a two-lens injection system (in which the carriage position for one of the lenses can be adjusted to control the $F\#$ of the beam), and stage to which lanterns or fibers can be mounted, which can also be translated in XY (in the plane perpendicular to the beam) as well as F (to control the lateral position relative to the focus of the beam). The outputs of the lantern are dispersed onto a spectrograph. There is no dedicated tip-tilt mirror; however, coarse alignment can be obtained by adjusting the stage on which the DM is mounted, and faster control can be obtained by applying tip-tilt on the DM itself.

Subaru Telescope. SCExAO is also equipped with its own 2000-actuator deformable mirror and wavefront sensor, and typically runs a second AO loop for finer wavefront correction. For work with the PLN, near-infrared light is sent to the photonics injection unit as well as a PSF monitoring camera. The injection unit consists of a two-lens injection system (in which the carriage position for one of the lenses can be adjusted to control the $F\#$ of the beam), and stage to which lanterns or fibers can be mounted, which can also be translated in XY (in the plane perpendicular to the beam) as well as F (to control the lateral position relative to the focus of the beam). The outputs of the lantern are dispersed onto a spectrograph. There is no dedicated tip-tilt mirror; however, coarse alignment can be obtained by adjusting the stage on which the DM is mounted, and faster control can be obtained by applying tip-tilt on the DM itself, although precautions have to be taken as the same DM is also applying the AO correction.

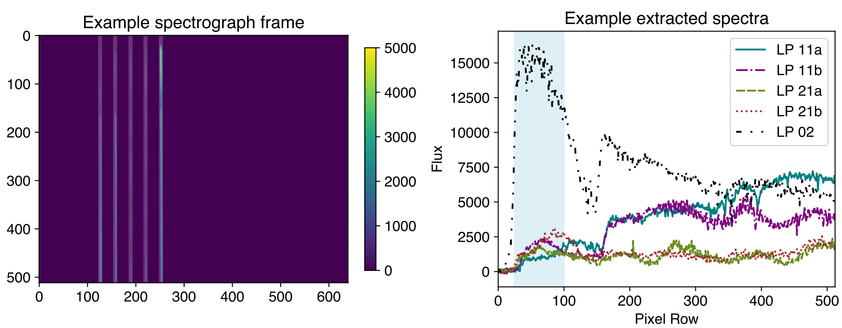

Figure 5.2: Left) An example frame obtained from the detector, showing the spectra from (from left to right) the LP 11a, LP 11b, LP 21a, LP 21b, and LP 02 ports. The shaded regions indicate the traces used for spectral extraction. Right) The extracted spectra from the traces indicated on the left, obtained by summing the counts in each row. The blue shading indicates the wavelength region summed over for computing coupling maps. Smaller row numbers correspond to longer wavelengths.

After the light has been injected into the lantern, the outputs of the lantern are routed through a V-groove array and spectrally dispersed onto a camera to create parallel spectral traces. An example frame on the spectrograph detector is shown in Fig. 5.2a, along with the traces used for spectral extraction. Example spectra for the five ports, extracted by summing across each row, are shown in Fig. 5.2b. The resolving power of this spectrograph is $R \approx 500$.

To prevent saturation of the detector, the trace of the LP 01 port, which is very bright when the lantern is centered, is placed off to the side such that it lands off the detector when working with the other five ports. The stage holding the camera can be translated relative to the spectral traces to view the LP 01 trace when desired, e.g. to aid with alignment.

5.2 Daytime Calibration and Testing

We performed calibrations during the daytime to align the lantern to the beam, using the default flattened DM setting of SCExAO. We optimized the injection $F\#$ to maximize the peak coupling into the LP 11a port, resulting in an $F\#$ of about 6, consistent with previous laboratory experiments presented in Xin, Echeverri, et al. (2024).

The wavelengths corresponding to the pixels are calibrated by inserting a series of narrowband filters in front of the light source and measuring the resulting peak intensity on the detector. The wavelength as a function of pixel row for each port is

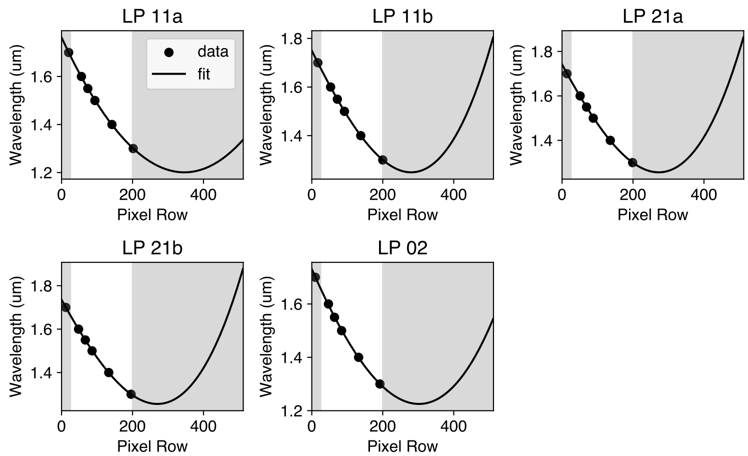

Figure 5.3: Calibration of detector pixel row to wavelength using narrowband filters. A cubic polynomial is fit to the data for each port. The overall flux is very low below the 25th row, and the inversion in the fitted polynomial results in the wavelength calibration only being valid up to approximately the 200th row. The rest of this work thus only uses data between the 25th and 200th rows, or between approximately 1.30 μm and 1.65 μm. The gray shaded regions indicate the regions excluded from the rest of this work.

shown in Fig. 5.3, along with the best fit cubic polynomials used to map pixel row number to wavelength in μm. The overall transmission is very low below the 25th row, and because of the inversion in the fitted polynomial, the wavelength calibration is only valid up to the 200th row. The rest of this work thus only uses data between the 25th and 200th rows, or between approximately 1.30 μm and 1.65 μm.

The lantern was designed to operate at 1.55 μm. To calculate coupling maps in the wavelength region of interest, we sum the flux between the 25th and 100th row for each port (corresponding to the wavelengths between approximately 1.47 μm and 1.67 μm). There are two methods for translating the beam relative to the lantern: by translating the injection stage itself or by applying a tip-tilt to the incoming beam. Tip-tilt applied to the incoming beam can be seen as a shift on the PSF monitoring camera, which is calibrated to milliarcseconds (mas), while the injection stage translation is not calibrated to physical units. However, the range of tip-tilt we can apply to the beam is limited to about 50 mas from the center, so it cannot capture the full extent of the coupling maps. We therefore use the following approach to calibrate the lantern position. We first fix the PSF on the PSF monitoring camera to the central pixel and scan the injection stage in X and Y. We then move the stage to the position of minimum summed coupling (across the four nulled ports), then fix

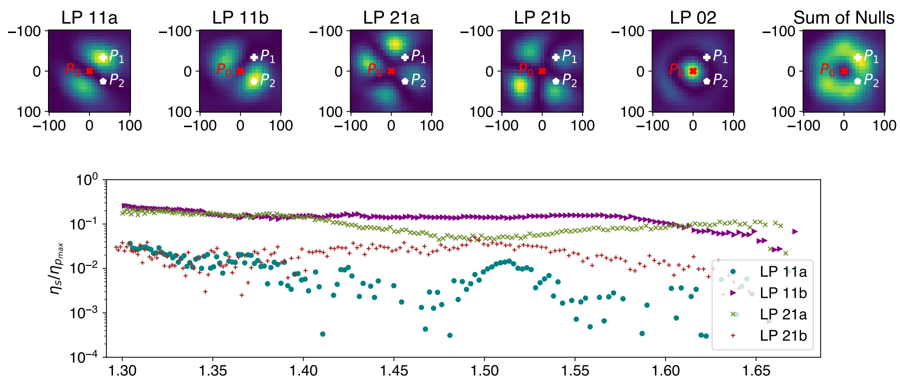

Figure 5.4: Top) The laboratory coupling maps corresponding to the five measured ports, with spatial scale in mas. The right-most panel is the summed coupling across the first four (nulled) ports. Three points of interest are marked, corresponding to intended locations at which on-sky data was later taken. P_0 indicates the center, the location of minimum summed coupling, where stellar coupling is measured. P_1 indicates the location of maximum coupling for the LP 11a port. P_2 indicates the location of maximum coupling for the LP 11b port. Bottom) Wavelength-calibrated null depths (stellar coupling divided by peak coupling) for each port.

the position of the injection stage and scan the PSF across a 100×100 mas window. We then obtain another scan using the injection stage, centered about the null. We fit the two sets of data to each other, finding the best-fit magnification and rotation angle, to calibrate units of lantern translation to mas.

The full coupling maps obtained by translating the lantern, with spatial scale calibrated to mas, are shown in Fig. 5.4a. Three points of interest are marked, corresponding to intended locations at which on-sky data was later taken (see Section 5.3). P_0 is the center of the lantern to which the star would be aligned, and where the stellar coupling (η_s) is measured. P_1 and P_2 mark where the peak off-axis coupling ($\eta_{p_{peak}}$) occurs for the LP 11a and LP 11b ports respectively.

Figure 5.4b shows the computed null-depths ($\eta_s/\eta_{p_{peak}}$) for each port. In this case, $\eta_{p_{peak}}$ for each wavelength is measured at the location of peak summed coupling for that port, identified using the coupling maps from Fig. 5.4 — i.e., for each port, the spatial location of this measurement is consistent across wavelengths rather than being independently calculated for each wavelength.

Because multiple submodules of SCExAO are used in any given engineering night and the hardware configuration is different for each, the instrument cannot be left in the calibrated configuration. We save the settings for the lantern alignment in order to restore it later on-sky. However, as detailed in Section 5.3, precisely recovering the alignment of the lantern poses significant difficulties, and future work is needed to identify sources of misalignment between the daytime and on-sky configurations, as well as to develop more robust strategies for recovering the alignment on-sky.

5.3 On-sky Observations

Misalignment and Attempted Recovery

During an engineering night on March 11th, 2025, we observed HD144206, a B-type star with an H-band magnitude of 4.923. We found that the flux through the lantern was higher when the AO loop was open than when it was closed, indicating the presence of a severe misalignment. We used manual gradient ascent to reposition the lantern such that the beam fell within its field-of-view (FOV), then conducted a scan to obtain the coupling maps shown on the left of Fig. 5.5. Because of the relatively good atmospheric conditions and AO correction, the scan revealed a central minimum in the sum of flux across the nulled ports. We aligned the lantern to this position, which we indicate as $\hat{P}_0$, our estimate of the location of P_0, with $\hat{P}_1$ and $\hat{P}_2$ indicating our estimates of P_1 and P_2 respectively.

The relative position of the injection stage between the daytime position and the onsky position, converted to units of mas, is plotted on the right of Fig. 5.5, showing that the misalignment caused the beam to initially land entirely out of the lantern's FOV. Additionally, data taken the afternoon after the observation shows that the alignment continued to drift in the same direction even after the observation. It is unclear what exactly caused this misalignment, but possibilities include mechanical settling (due to the instrument being craned just prior to calibrations), or a potential shift in an optic upstream. It is possible that this extreme of an misalignment will not typically occur on a regular basis; however, characterizing the typical level of misalignment will be an important part of future work.

On-sky measurements

After the lantern alignment was approximately recovered, we took data at the three points $\hat{P}_0$, $\hat{P}_1$, and $\hat{P}_2$. The median across all frames of the spectra obtained at each of these points is presented in Fig. 5.6. At the estimated center $\hat{P}_0$, LP 02 (the non-nulled port) has the highest coupling, as expected. Meanwhile, LP 11a has the highest coupling at $\hat{P}_1$, and LP 11b has the highest coupling at $\hat{P}_2$, as expected.

From these flux measurements, we calculate the onsky null-depths for the LP 11a and LP 11b ports by dividing the median of their stellar coupling $\eta_{\hat{P}_0}$ by the median

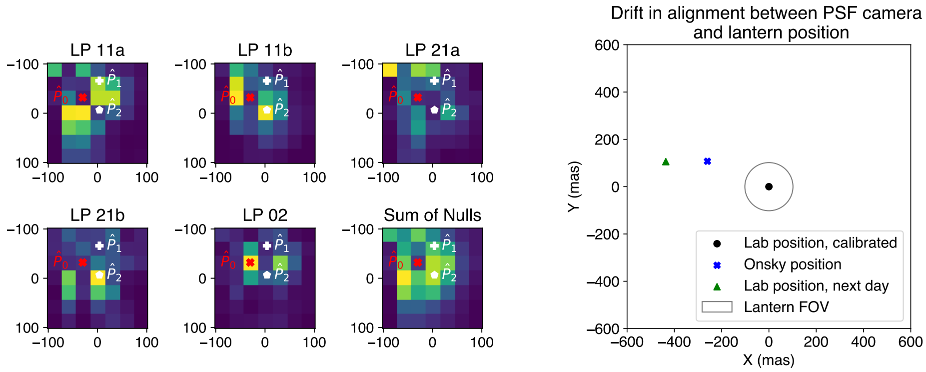

Figure 5.5: Left) The on-sky coupling maps corresponding to the five measured ports, with spatial scale in mas. Additional on-sky data was taken at the three locations $\hat{P}_0$, $\hat{P}_1$, $\hat{P}_2$. These are the estimated positions of the points indicated in Fig. 5.4, with some alignment uncertainty due the coarse sampling and rough quality of the on-sky coupling maps. Right) The measured drift in the alignment between the lantern position and the central pixel on the PSF monitoring camera. The calibrated daytime position is at the center, marked with the black dot. The range of the scans presented on the left as well as in Fig. 5.4 (approximately the FOV of the lantern) is indicated by the gray circle. The position of the lantern obtained during the on-sky observation is indicated with the blue square, while the position of the lantern measured the afternoon after the observation is indicated with the green triangle, showing a significant shift in the alignment between calibration and observation, which continues in the same direction even after the observation.

of their peak coupling ($\eta_{\hat{P}_1}$ and $\eta_{\hat{P}_2}$ respectively). These null-depths are plotted across wavelength in Fig. 5.7. We also plot the $\eta_{\hat{P}_0}$ measured in the LP 21b port divided by its $\eta_{\hat{P}_2}$ coupling. Although this is not a null-depth (since peak coupling does not occur for the LP 21b port at $\hat{P}_2$), as seen in Fig. 5.6, this port receives a significant amount of flux for an object at this position and thus can provide scientifically valuable information about the object.

Rather than first taking the median across the $\eta_{\hat{P}_0}$ frames, we can also divide the flux by the appropriate median of $\eta_{\hat{P}_1}$ or $\eta_{\hat{P}_2}$ and plot the distribution. The bottom row of Fig. 5.7 shows these histograms at the lantern design wavelength of 1.550 μm. A potential avenue for future work is analyzing these distributions using the null self-calibration (NSC) method, which can increase sensitivity by fitting a statistical model to the distribution to separately constrain the contributions to the null-depth from the underlying astrophysical signal and the atmospheric or instrumental fluctuations (Hanot et al., 2011; Martinod et al., 2021).

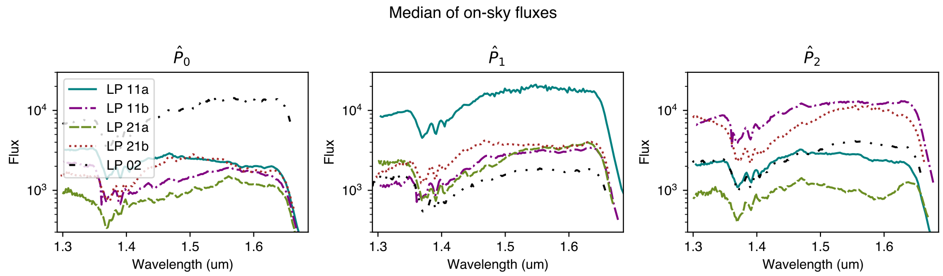

Figure 5.6: The median over time of the fluxes measured at points $\hat{P}_0$, $\hat{P}_1$, $\hat{P}_2$. At the estimated lantern center $\hat{P}_0$, the non-nulled port LP 02 has the highest coupling, as expected. Meanwhile, LP 11a has the highest coupling at $\hat{P}_1$, and LP 11b has the highest coupling at $\hat{P}_2$, as expected.

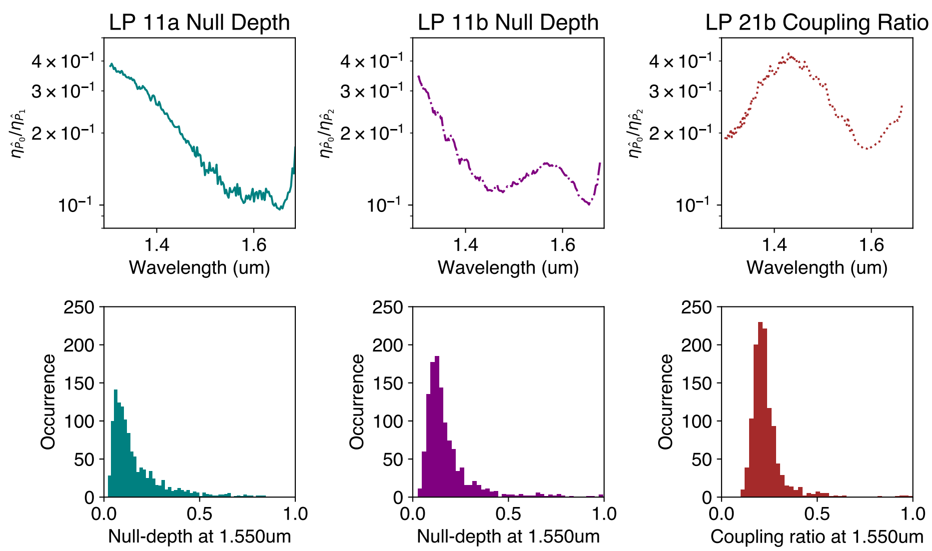

Figure 5.7:

Estimated detection sensitivity

Lastly, we estimate the detection sensitivity to a binary companion located at $\hat{P}_1$. Our data at $\hat{P}_0$ was taken in two separate sequences, one before the data at $\hat{P}_1$ and $\hat{P}_2$, and one after. We choose to use the first sequence for our analysis, as the second sequence occurred during the end of the night when the sun was rising and the conditions were degrading. We split this first sequence of frames obtained at $\hat{P}_0$ into two halves of 400 frames (over approximately 2 minutes) each. We use the first half as the science data to which we can choose to inject a fake companion (by scaling the data measured at $\hat{P}_1$ by a desired flux ratio). The second half we use as a reference dataset, as if we had conducted a reference differential imaging observation of a point-source star. This order of frame allocation will be referred to as Dataset 1. We also perform the same analysis allocating the first half as the reference dataset and injecting a fake companion (or not) to the second half to serve as the science data. This order of frame allocation will be referred to as Dataset 2. Although SCExAO operates in pupil-tracking mode (there is a pupil mask within the instrument rotationally matched to the telescope pupil), for this work, we assume that the sky rotation is negligible such that the companion remains at $\hat{P}_1$ the entire time, a reasonable assumption for our data which spans the order of minutes. For longer observations with noticeable sky rotation, a more sophisticated algorithm would be needed to take advantage of the additional angular 2diversity and fit for the location of the companion (Goyal et al. in prep).

The four mock science datasets are thus Dataset 1 with a fake companion, Dataset 1 without a fake companion, Dataset 2 with a fake companion, and Dataset 2 without a fake companion, where the fake companion, when injected, has a flux ratio of 1/10. We first perform PSF calibration by subtracting a normalized median estimate of the reference PSF from the science data for each wavelength independently, as to make the PSF subtraction process agnostic to potential differences in spectra between the science target star and the reference star. The resulting data are shown in Fig. 5.8. In the case of Dataset 1, the difference in the scale of the PSF-subtracted data between the companion and no-companion case is visible by eye, while in Dataset 2, the signal of the companion is comparable to that of the post-subtraction residuals.

To estimate the detectability of the companion, we fit the PSF-subtracted data D to forward-models of the planet PSF M, in this case the spatially-varying coupling data obtained during daytime calibrations propagated through the same normalized median subtraction as the science data. For simplicity, we collapse our individual wavelength solutions for each port into the mean across all of them. At each spatial location i, for each spectral channel j, we find the companion flux $\bar{f}_{ij}$ that maximizes the log-likelihood computed across all ports k:

$$\bar{f}_{ij} = \arg\max_f \{-\frac{1}{2}\sum_k \frac{(D_{jk} - fM_{ijk})^2}{\sigma_{jk}^2}\}. \tag{5.1}$$

Here, $\bar{f}_{ij}$ is the relative scaling factor between the companion signal in the data and the flux as measured with the lab source. We compute it across a logarithmically-spaced grid spanning from 10^{-2} to 1. Its actual value is not physically meaningful as each were taken with different camera settings, and neither are calibrated to

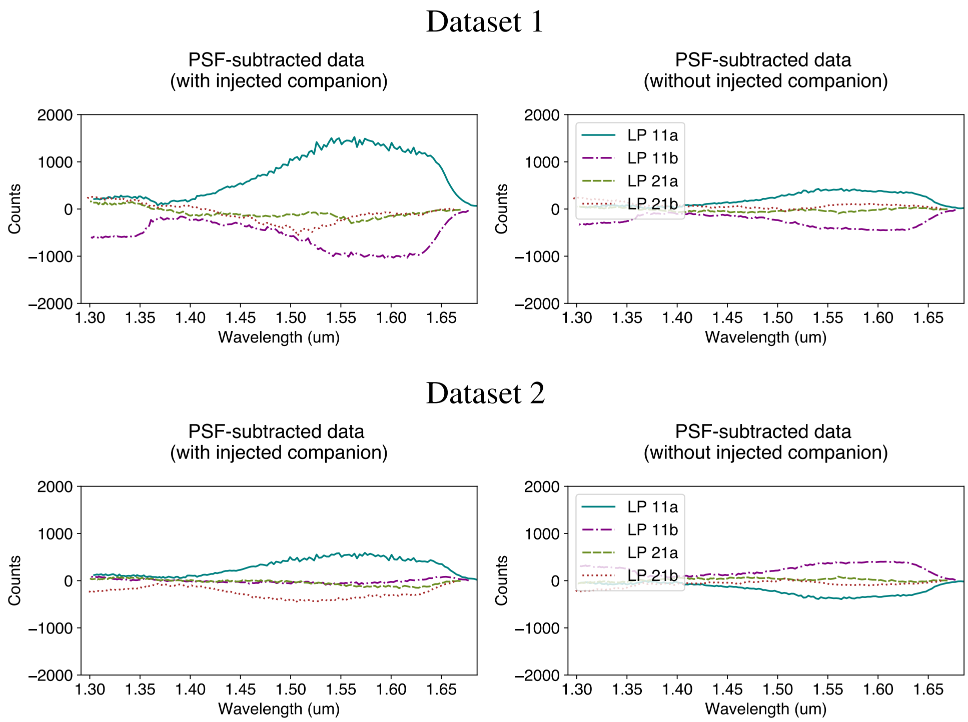

Figure 5.8: Left) Data after PSF subtraction for Dataset 1 (top) and Dataset 2 (bottom) with a fake companion injected with a flux ratio of 0.1. Right) Data after PSF subtraction for Dataset 1 (top) and Dataset 2 (bottom) without any injected fake companions. In the case of Dataset 1, the difference in the scale of the PSF-subtracted data between the companion and no-companion case is visible by eye, while in Dataset 2, the signal of the companion is comparable to that of the post-subtraction residuals.

absolute flux units. We leave $\bar{f}_{ij}$ in uncalibrated units for this work; however, it would be possible to calibrate these values to astrophysical units, such as by observing photometric standard stars.

We then assign the sum of log-likelihoods across wavelengths as the log-likelihood for that spatial position:

$$\ln \mathcal{L}_i = -\frac{1}{2} \sum_j \sum_k \frac{(D_{jk} - \bar{f}_{ij} M_{ijk})^2}{\sigma_{jk}^2}. \tag{5.2}$$

Lastly, we estimate $\ln \mathcal{L}_0$ as the log-likelihood of the data being entirely noise,

$$\ln \mathcal{L}_0 = -\frac{1}{2} \sum_j \sum_k \frac{D_{jk}^2}{\sigma_{jk}^2} \tag{5.3}$$

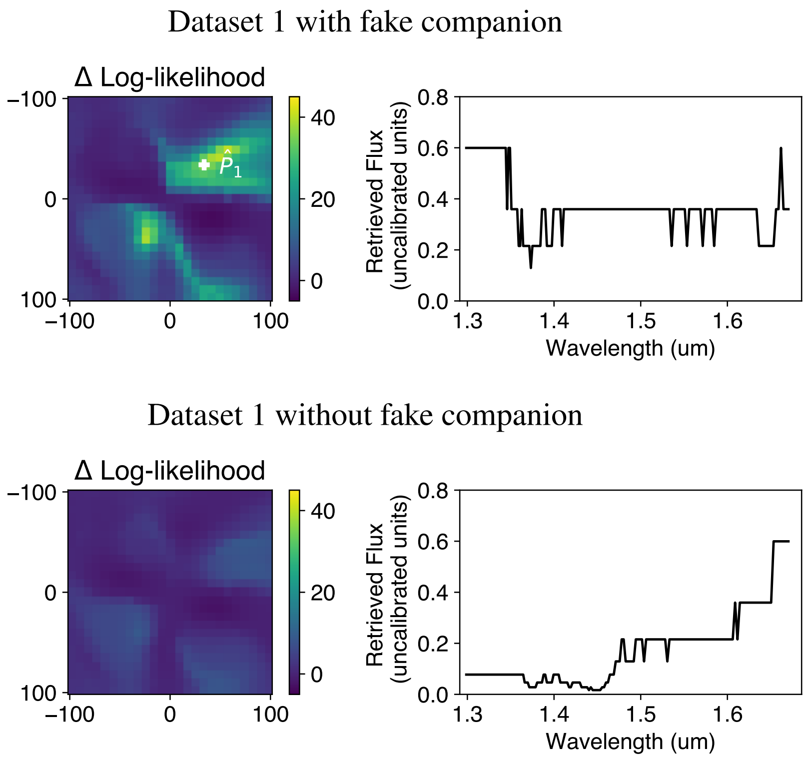

Figure 5.9: Detection results for Dataset 1. Left) The difference in log-likelihoods $\ln \mathcal{L}_i - \ln \mathcal{L}_0$ between a fitted model with a companion at each location and a model containing only noise, for a dataset that has a fake companion injected (top) and for one without a fake companion (bottom). Right) The respective best-fit flux (in uncalibrated units) as a function of wavelength at the best-fit spatial location. This data shows a strong detection of the binary companion when it exists, with both $\ln \mathcal{L}_i - \ln \mathcal{L}_0$ and the retrieved flux significantly higher than in the case without an injected companion, as well as an accurate localization of the companion.

and plot the difference in log-likelihood $\ln \mathcal{L}_i - \ln \mathcal{L}_0$ for all positions i in Fig. 5.9 (for Dataset 1) and Fig. 5.10 (for Dataset 2). This quantity $\ln \mathcal{L}_i - \ln \mathcal{L}_0$ is a measure of the relative likelihood of the model with a companion in that given position, relative to the likelihood of there being no companion, in logarithmic scale. Also plotted are the retrieved fluxes (in uncalibrated units) for each of the mock science datasets.

We find that for Dataset 1, we obtain a strong detection of the binary companion when it exists, with both $\ln \mathcal{L}_i - \ln \mathcal{L}_0$ and the retrieved flux significantly higher than in the case without an injected companion, as well as an accurate localization of the companion. Dataset 2, on the other hand, shows a present but weaker detection of the binary companion when it exists. The map of $\ln \mathcal{L}_i - \ln \mathcal{L}_0$ clearly favors the existence of a binary companion, but the retrieved flux is not significantly above

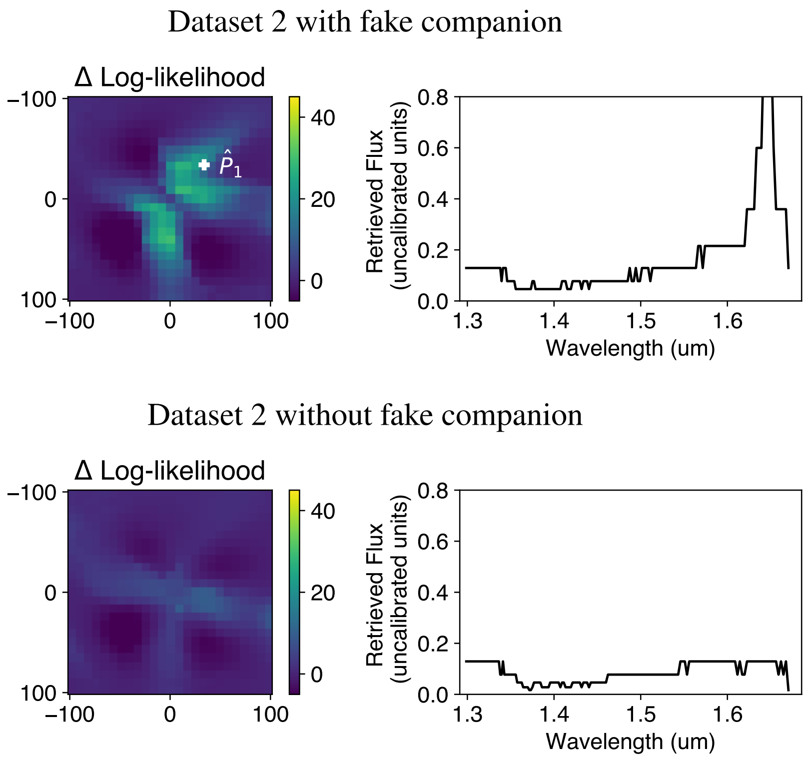

Figure 5.10: Detection results for Dataset 2. Left) The difference in log-likelihoods $\ln \mathcal{L}_i - \ln \mathcal{L}_0$ between a fitted model with a companion at each location and a model containing only noise, for a dataset that has a fake companion injected (top) and for one without a fake companion (bottom). Right) The respective best-fit flux (in uncalibrated units) as a function of wavelength at the best-fit spatial location. This data shows a present but weaker detection of the binary companion when it exists. The map of $\ln \mathcal{L}_i - \ln \mathcal{L}_0$ clearly favors the existence of a binary companion, but the retrieved flux is not significantly above the retrieved flux in the case of no injected companion. The localization is also weaker than with Dataset 1.

the retrieved flux in the case of no injected companion. The localization is also weaker than with Dataset 1, though still informative. The difference between these two orders of frame allocation indicates the presence of time-correlated systematic noise that has not been averaged out, which is expected as the data only spans about 4-minutes. Additionally, we observe the presence of non-Gaussianity in our noise, as seen in the histograms in Fig. 5.7, and as manifested as a bias towards positive values in the maps of $\ln \mathcal{L}_i - \ln \mathcal{L}_0$ for the noise-only datasets (which would be symmetric about zero in the case of Gaussian, zero-mean noise). While we have implemented one of most simple detection methods as an initial estimate of the detectability of a hypothetical 1:10 binary companion, future work includes using more sophisticated post-processing algorithms to improve the sensitivity, such as PCA-based PSF subtraction methods (which typically whitens the noise) (Soummer, Laurent Pueyo, and Larkin, 2012a; Laurent Pueyo, 2016a), calibration of tip-tilt using centroids obtained from the PSF monitoring camera, and statistical null self-calibration using the distribution of null fluctuations (Hanot et al., 2011). Additionally, Bayesian inference tools can be used to calculate the Bayes factors between the various models in order to more rigorously characterize the sensitivity.

5.4 Conclusion

In this work, we present the first on-sky demonstration of the PLN on the SCExAO instrument, which injects an AO-corrected beam into the PLN and disperses the outputs onto a spectrograph with $R \approx 500$. We first explain the daytime calibration procedure and show results from laboratory testing. We then present the results of on-sky engineering: first our attempt to correct for a severe misalignment between daytime calibrations and on-sky observations, then our measurements of on-sky spectra at three different points (the null, the location of peak coupling for the LP 11a port, and the location of peak coupling for the LP 11b port). From this data, we calculate null-depths for the LP 11 ports to be approximately 10^{-1}. We then use our on-sky data to generate a mock binary dataset and a reference star dataset in order to constrain the detectability of a hypothetical binary companion, showing that a companion with a flux ratio of 0.1 can be detected with 2 minutes of science exposures and 2 minutes of reference star exposures. However, the level of significance and retrieved companion flux depends on which subset of frames are used as the science data as opposed to the reference data, indicating the presence of time-correlated systematic noise that has not been averaged out. Future work includes diagnosing sources of misalignment between daytime calibrations and on-sky observations and testing more robust on-sky alignment procedures. It also includes observing binary systems on-sky to detect and characterize companions, as well as using more sophisticated post-processing algorithms to improve the detection sensitivity and Bayesian tools to better characterize it. Additionally, dispersing the light onto a spectrograph with higher resolving power ($R >> 4000$) in the future would enable high-resolution techniques based on spectrally resolving atomic and molecular absorption lines, which can improve detection sensitivity by a factor of ~ 100 in flux ratio and also enable atmospheric characterization of the companion by detecting the absorption signatures of particular compounds.

Chapter 6

ENHANCING FIBER-FED SPECTROSCOPY WITH SPECKLE NULLING ON THE KECK PLANET IMAGER AND CHARACTERIZER

Xin, Yinzi et al. (July 2023). "On-sky speckle nulling through a single-mode fiber with the Keck Planet Imager and Characterizer". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 9, 035001, p. 035001. DOI: 10.1117/1.JATIS.9.3.035001. arXiv: 2307.11893 [astro-ph.IM].

While the bulk of my thesis focuses on developing the PLN, I was also an active part of the instrument team for the Keck Planet Imager and Characterizer (KPIC) instrument, a fiber-fed spectrograph routing AO-corrected light from the Keck II telescope, through a single-mode fiber, to NIRSPEC, Keck's high-resolution spectrograph with $R = 35000$. KPIC has both a direct spectroscopy mode (where the fiber is centered on the companion) and a Vortex Fiber Nuller mode, so it is a useful platform and benchmark for exploring the possibilities of a PLN-fed high-resolution spectrograph. I helped to characterize the upgraded version of KPIC in 2022 before it was deployed, with calibrating the instrument for observations, and with conducting the observations themselves. I also used the instrument as a platform for demonstrating wavefront sensing and control techniques, and this chapter reproduces my paper on using speckle nulling through the single-mode fiber (in direct spectroscopy mode). Speckle nulling is a variation on probed focal-plane wavefront control, similar to the iEFC algorithm used in Chapter 4, so this work also provides an estimate of how well iEFC with a PLN might work if implemented on-sky with similar hardware (spectrograph resolving power, detector characteristics, etc.). For this work, I used the deformable mirror of KPIC to destructively interfere starlight leaking through the fiber with itself to suppress it during science observations, reducing the contamination caused by stellar noise. This technique can consistently achieve suppression factors of 2.5 to 3 on-sky and led to the most sensitive detection limit achieved by KPIC at this spatial separation (J. J. Wang, Dimitri Mawet, et al., 2024).

Abstract

The Keck Planet Imager and Characterizer (KPIC) is an instrument at the Keck II telescope that enables high-resolution spectroscopy of directly imaged exoplanets and substellar companions. KPIC uses single-mode fibers to couple the adaptive optics system to Keck's near-infrared spectrometer (NIRSPEC). However, KPIC's sensitivity at small separations is limited by the leakage of stellar light into the fiber. Speckle nulling uses a deformable mirror to destructively interfere starlight with itself, a technique typically used to reduce stellar signal on a focal-plane imaging detector. We present the first on-sky demonstration of speckle nulling through an optical fiber with KPIC, using NIRSPEC to collect exposures that measure speckle phase for quasi-real-time wavefront control while also serving as science data. We repeat iterations of measurement and correction, each using at least 5 exposures. We show a decrease in the on-sky leaked starlight by a factor of 2.6 to 2.8 in the targeted spectral order, at a spatial separation of 2.0 λ/D in K-band. This corresponds to an estimated factor of 2.6 to 2.8 decrease in the required exposure time to reach a given SNR, relative to conventional KPIC observations. The performance of speckle nulling is limited by instability in the speckle phase: when the loop is opened, the null-depth degrades by a factor of 2 on the timescale of a single phase measurement, which would limit the suppression that can be achieved. Future work includes exploring gradient-descent methods, which may be faster and thereby able to achieve deeper nulls. In the meantime, the speckle nulling algorithm demonstrated in this work can be used to decrease stellar leakage and improve the signal-to-noise of science observations.

6.1 Introduction

The Keck Planet Imager and Characterizer (KPIC) is a dedicated instrument for the high-resolution spectroscopy of directly imaged companions, combining Keck II's adaptive optics system with its high-dispersion near-infrared spectrometer (NIRSPEC) (Delorme et al., 2021b) using single-mode fibers. It is part of a family of instruments that leverages adaptive optics in combination with high-resolution spectroscopy, including HiRISE (Otten et al., 2021), RISTRETTO (Lovis, Blind, et al., 2022), REACH (Kotani et al., 2020), and VIS-X (Sebastiaan Y. Haffert et al., 2021). Although the combination of high-contrast imaging and high-dispersion coronagraphy is currently unable to access rocky planets in the habitable zone, the use of this technique on the Extremely Large Telescopes has the potential to reach this regime (Snellen et al., 2015).

KPIC operates by coupling the planet's light into a single-mode fiber and routing it to NIRSPEC. Because the star is typically orders of magnitude brighter than the planet and the planet is close to the star (< 1"), the signal on the spectrograph is often dominated by stellar light leaking into the fiber. Although cross-correlating the observed spectrum with an atmospheric model of the planet can help disentangle the star and planet signals (Quinn M. Konopacky et al., 2013; Schwarz et al., 2016), the leaked starlight nevertheless contributes a significant amount of photon noise as well as systematic errors, such as through increased telluric signals and fringing. This noise limits the sensitivity of the instrument at close separations less than 1 arcsecond, though (as a rule of thumb) KPIC becomes limited by thermal background at separations greater than ~ 1 arcsecond, depending on the target magnitude.

To motivate speckle nulling, we examine the impact of stellar leakage and throughput on the integration time required to reach a given signal-to-noise ratio (SNR) on a planet's spectra. We denote the fractional planet throughput (co-axial with the SMF) as η_p and the fractional off-axis stellar throughput (or leakage) as η_s. According to Ruane, Ji Wang, et al. (2018), if the measurement is photon-noise limited (e.g. the noise scales as the square-root of photon count), the integration time τ required to reach a given SNR scales as

$$\tau \propto \frac{\eta_s}{\eta_p^2}. \tag{6.1}$$

If the measurement is limited by systematic errors (e.g. the noise scales linearly with photon count), then the relative integration time to reach a given SNR scales as

$$\tau \propto \frac{\eta_s}{\eta_p}. \tag{6.2}$$

In either case, it is advantageous to decrease η_s with wavefront control, while keeping η_p as high as possible. For wavefront control, we use a Boston Micromachines deformable mirror (DM) with 1000 actuators, which was added as part of Phase II of KPIC (N. Jovanovic, B. Calvin, et al., 2020). In this work, we explore using speckle nulling with a single-mode-fiber-fed spectrograph, adapting an algorithm originally used for suppressing starlight (speckles) on a focal-plane imaging detector (Bottom et al., 2016). Speckle nulling aims to reduce η_s using the DM, and because the DM typically only applies a small perturbation to the wavefront, its impact on η_p is expected to be very minor. This is unlike conventional coronagraphs, which can achieve very good starlight suppression and thus very low η_s, but may result in a sizable decrease in η_p, especially at close separations (O. Guyon et al., 2006a).

In practice, there are other considerations that impact the relative integration time, such as different fractional overheads for each per-frame integration time, or if any frames taken for probing or calibration need to be discarded. In our case, our integration overheads while speckle nulling are only slightly higher than that of conventional KPIC observations (20% versus 10-15%), and, as discussed in Section 6.3, all of the frames taken in our implementation may be used for science.

We derive the equations for speckle nulling through a fiber in Section 6.2, and present laboratory and on-sky demonstrations with KPIC in Section 6.3.

6.2 Methods

The equations describing traditional speckle nulling at a focal-plane are given in Bottom et al. (2016). We adapt those equations for speckle nulling through a fiber here.

We define the complex amplitude of the scalar electric field (“electric field" hereafter) contributed by the speckle to the focal plane as $E_{s_0}(x, y)$ and the focal plane electric field induced by the DM as $E_{\mathrm{DM}_0}(x, y)$. The coupling of the focal plane electric field into a single-mode fiber can be described as an overlap integral between the field and the mode of the fiber centered at its physical location, denoted by (the real-valued) $\Psi(x, y)$.

We can define the (potentially time-varying) contribution of the speckle to the electric field through the fiber as

$$E_s = a_s e^{i\phi_s} = \int dxdy \Psi(x, y) E_{s_0}(x, y), \tag{6.3}$$

and the contribution of the DM through the fiber as

$$E_{\mathrm{DM}} = a_{\mathrm{DM}} e^{i\phi_{\mathrm{DM}}} = \int dxdy \Psi(x, y) E_{\mathrm{DM}_0}(x, y). \tag{6.4}$$

The electric field through the fiber can thus be described as

$$E_{\mathrm{fib}} = E_s + E_{\mathrm{DM}} = a_s e^{i\phi_s} + a_{\mathrm{DM}} e^{i\phi_{\mathrm{DM}}}. \tag{6.5}$$

The intensity (I_{fib}) measured at the output of the fiber is

$$I_{\text{fib}} = |E_{\text{fib}}|^2 = a_s^2 + a_{\text{DM}}^2 + 2a_s a_{\text{DM}} \cos(\phi_s - \phi_{\text{DM}}). \tag{6.6}$$

Assigning any flux through the fiber as speckle flux to be nulled, the speckle amplitude can be calculated as $a_s = \sqrt{I_{\text{fib}}}$. The speckle phase can be determined by applying DM probes. By taking measurements with $\phi_{\text{DM}} = [0, \pi/2, \pi, 3\pi/2]$, we obtain the following probe measurements (assuming no noise):

$$I_1 = a_s^2 + a_{\text{DM}}^2 + 2a_s a_{\text{DM}} \cos(\phi_s) \tag{6.7}$$

$$I_2 = a_s^2 + a_{\text{DM}}^2 + 2a_s a_{\text{DM}} \sin(\phi_s) \tag{6.8}$$

$$I_3 = a_s^2 + a_{\text{DM}}^2 - 2a_s a_{\text{DM}} \cos(\phi_s) \tag{6.9}$$

$$I_4 = a_s^2 + a_{\text{DM}}^2 - 2a_s a_{\text{DM}} \sin(\phi_s). \tag{6.10}$$

The speckle phase can be estimated as

$$\phi_s = \tan^{-1}\left[\frac{I_2 - I_4}{I_1 - I_3}\right]. \tag{6.11}$$

Any incoherent flux would appear in all probe measurements and thus subtract out in the phase calculation. The DM can then be used to apply $a_{\text{DM}} = a_s$ and $\phi_{\text{DM}} = \phi_s + \pi$ such that $I_{\text{fib}} = 0$. Theoretically, there are many ways to achieve the desired a_{DM} and ϕ_{DM} through the fiber. We choose to use a sinusoid on the DM, which applies a speckle with a point-source-like extent in the focal plane. We calibrate the sinusoid's spatial frequency and direction to maximize influence through the fiber, and apply a_{DM} and ϕ_{DM} as the sinusoid's amplitude and phase as respectively.

This derivation assumes that the light is monochromatic. In our application, we target a narrow wavelength band (chosen to be one echelle order of the NIRSPEC spectrograph, which spans approximately 45 nm or a $\Delta\lambda/\lambda$ of 0.02) that corresponds to a spread in the focal plane of $\sim 0.05\ \lambda/D$. Since this is much smaller than the spatial extent of a single-sinusoid speckle ($\sim 1\ \lambda/D$), we expect the effect of chromaticity to be small. In Section 6.3, we show empirically that the monochromatic assumption is indeed valid, as the wavelength band we work with is narrow compared to the null created by the sinusoid. Extending speckle nulling into broadband may be possible by calibrating multiple sinusoids that target different wavelengths, a topic left for future work.

6.3 Results

Laboratory Test

A simplified diagram of KPIC is shown in Figure 6.1a. Light from the telescope is reflected by a deformable mirror, which is used to change the phase of the wavefront. The light then propagates through to a tip-tilt mirror, which is used to put the star's point-spread-function (PSF) on a specified pixel on the tracking camera which receives J- and H-band light from the dichroic. The K- and L-band light is sent to a focusing lens, which injects it into the SMF that then routes it to NIRSPEC.

We first tested the algorithm using Keck's internal broadband source to characterize the suppression achievable in the lab, in the absence of atmosphere. We calibrated the DM sinusoid spatial frequency and direction to maximize influence into a fiber separated by 98 mas (the predicted separation of HD 206893c at the time of testing) from the star. On KPIC, 98 mas corresponds to 2.2 λ/D in K-band (the science band seen by NIRSPEC, with a central wavelength of 2.2 μm). We also calibrated the amplitude of the sinusoid in DM units (from 0 to 1) to a_{DM} in square-root-of-raw-contrast units.

We then perform the speckle nulling sequence, first taking the four probe frames to calculate the phase, then taking an unprobed frame to measure the raw contrast and converting it to DM units. We apply a DM correction with a gain on the amplitude (0.5 for in-lab tests, though we found that 0.25 works better on-sky) and a phase of $\phi_s + \pi$. We repeat this procedure until the null stops improving. Here, we do not apply any other perturbations to the DM and just use it to null the underlying static speckle. However, in future work, it may be possible to simulate on-sky conditions by injecting atmospheric turbulence or phase drift on the DM.

Figure 6.1b shows the nulling sequence, using the raw contrast (the ratio of leaked off-axis starlight to the fractional co-axial throughput) in NIRSPEC order that spans 1.99 to 2.04 μm as the metric. The ratio of initial raw contrast to the mean raw contrast of the last three iterations is 27.5. Meanwhile, because the DM-induced perturbation to the wavefront is indeed very small, the change in co-axial throughput is actually below our ability to measure, since Keck's internal light source is variable by 10-20%. Towards the end of the nulling sequence, increasing the exposure time did not result in a deeper null, indicating that the null depth was limited by systematic effects. We did not explore the source of this limitation, since this suppression ratio was already much higher than we expected to achieve on-sky (where the performance is expected to be limited by instrumental phase drift or atmospheric

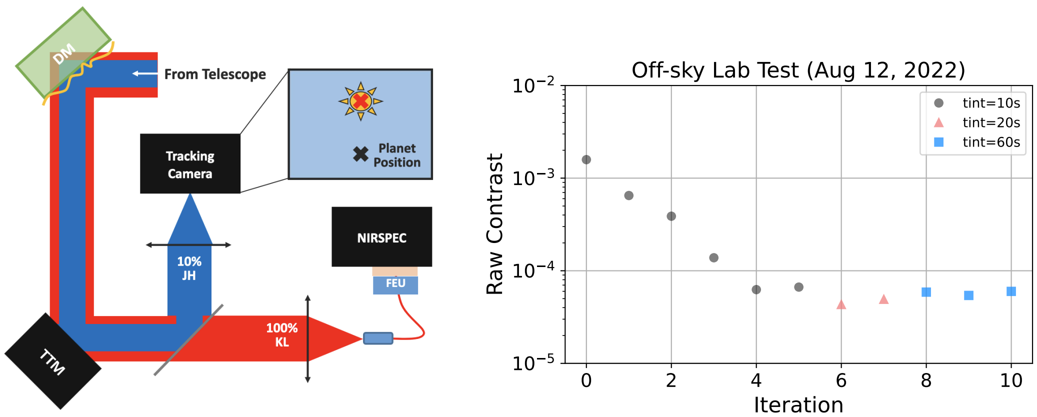

Figure 6.1: a) Diagram of KPIC. Light from the telescope is reflected off of a deformable mirror, which is used to change the shape of the wavefront. The light then propagates to a tip-tilt mirror, which is used to put the star's PSF on a specified pixel on the tracking camera that receives J and H band light from the dichroic. The K and L band light is sent to the fiber injection unit, which routes it to NIRSPEC. The inlay depicts the alignment procedure on the tracking camera: the star is positioned on the red cross such that the fiber position (indicated by the black cross) coincides with the predicted planet position. b) Raw contrast from 1.99 to 2.04 μm at each iteration during a laboratory test of speckle nulling. The ratio of initial raw contrast to the mean raw contrast of the last three iterations is 27.5. Further increasing the exposure time did not result in a deeper null, indicating that the null depth was limited by systematic effects.

turbulence). However, we speculate that this limit may be due to incoherent light in the instrument, introduced, for example, by optical elements such as dichroics, whose internal reflections can result in ghosts, or light that loses optical coherence with the main beam if the optical path difference between the two beams is greater than the coherence length. Here, the coherence length of the (combined K and L band) beam in question is on the order of ten microns.

On-sky Engineering

We tested speckle nulling on-sky on October 11, 2022 on HD 206893, a star with a K-band magnitude of 5.593 (Cutri et al., 2003). We placed the fiber at a separation of 91 mas (2.0 λ/D in K-band), which was the predicted separation of the companion HD 206893c on that night (J. J. Wang, Kulikauskas, and Blunt, 2021; Hinkley et al., 2023). The on-sky nulling sequence, targeting the raw contrast from 2.29 to 2.34 μm (where many useful CO absorption lines lie), is shown in Figure 6.2a. The frames were obtained using an exposure time of 59.0 seconds, with the exception of

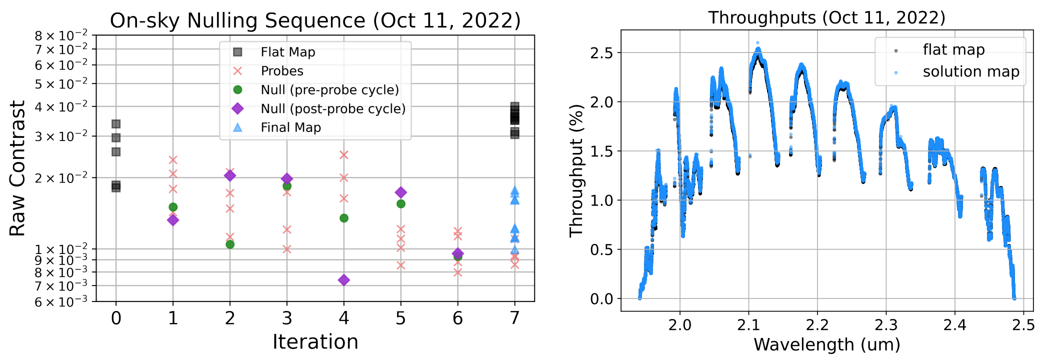

Figure 6.2: On-sky speckle nulling results from October 11, 2022. a) Raw contrast from 2.29 to 2.34 μm at each iteration of speckle nulling. Black squares correspond to the original flat map, which is the DM map that maximizes co-axial throughput through the fiber, calibrated using the internal source before the observing night begins. Pink crosses correspond to probes used to determine the speckle's phase, as described by Equations 7-11. We take one measurement of the intermediate null immediately after applying a solution map (shown in green) and a second measurement (with the same solution map) just after the next probe cycle but before applying the next map (shown in purple). In the absence of noise or drift, we expect these measurements to be the same as the DM state is identical, so the difference between them gives us a measure of the variability. Blue triangles correspond to the final solution map at the end of seven iterations. There is significant variability in measurements with the same instrument state, but by comparing the average of the blue triangles and the average of the black squares in iteration 7, we find that speckle nulling improves the raw contrast by a factor of 2.6 relative to using the flat map. b) Comparison of co-axial throughputs with and without speckle nulling, showing that speckle nulling does not decrease the throughput η_p.

the probe measurements for iterations 6 and 7 and the raw contrast measurements for iterations 5 and 6, which were obtained with an exposure time of 119.5 seconds (to test if increasing exposure time would improve the null). There is not enough data to fully compare the behavior of the iterations using 2 minute exposures to those using 1 minute exposures, but it tentatively seems that using the longer exposure time decreases the spread in probe measurements but does not lead to a deeper null. There is a trend towards deeper raw contrasts in the beginning, but the null becomes limited around 1×10^{-2}.

We observed significant variability in the measurements given the exact same instrument state. As a result, we took certain measures to characterize the level of variability, and to make sure that our comparisons and conclusions about our implementation are as valid as possible.

First, we took two raw contrast measurements of each intermediate null to measure the level of variability given the same instrument state. We take one measurement immediately after applying a solution map and a second measurement (with the same solution map) just after the next probe cycle but before applying the next map. These two separate measurements of intermediate nulls before and after the probe cycle are shown in Figure 6.2a in green and purple respectively.

Second, the blue triangle measurements in Figure 6.2a were all taken with the same DM map, and also taken as close to each other in time as possible. Although iteration 6 had a deeper null, iteration 7 is a fairer comparison to the final flat map measurements, since they were taken closer together in time, so the difference between them is less impacted by the variability.

Additionally, to check that the apparent gain from speckle nulling is not due to variability, after the last iteration, we flipped back and forth between the final DM map solution from speckle nulling and the original flat map — i.e. we took three frames using the solution map, followed by three frames with the original flat map, then repeated this twice for a total of 9 frames with each map. Comparing the mean raw contrasts across this set of measurements, we find that speckle nulling improved the null-depth by a factor of 2.6 relative to the flat map. Additionally, the raw contrasts during the probe measurements are not noticeably higher than the measurements with the DM solutions, and are still deeper than the raw contrast without speckle nulling at all. Thus, even the probe measurements can be used as science data, and no exposures have to be excluded from analysis.

Lastly, after collecting data comparing the final solution map with the flat map, we reset the DM to our final solution map. Then, we opened the loop and took short exposures to characterize the timescale of the null degradation. The open loop raw contrast over time is shown in Figure 6.3. This sequence shows that the null quickly degrades — in the time it takes to make a phase measurement using 1-minute exposures plus overheads (indicated with blue shading), the raw contrast increased by a factor of two. This degradation limits our ability to measure the drifting phase quickly enough to correct it, which is likely limiting our null-depth. This drift in phase may be due to the rotation of the pupil plane in our observation mode, as we fix the companion location to a specific point on the tracking camera and allow the pupil to rotate to maintain that position throughout the night. Unfortunately, in our case, the loop speed is limited by the photon rate, so the exposure time cannot be further

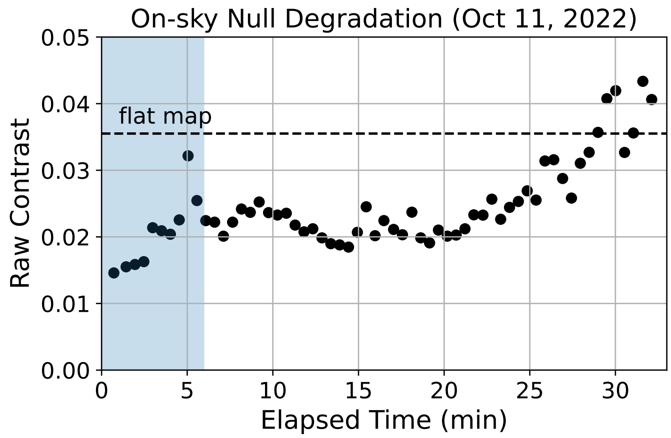

Figure 6.3: Raw contrast measurements over time when the loop is opened after the speckle nulling sequence. In the time it takes to make a phase measurement using 1 minute exposures plus overheads (indicated with blue shading), the raw contrast increased by a factor of two. This degradation limits our ability to measure the drifting phase quickly enough to correct it, which is likely limiting our null-depth. The dashed line indicates the raw contrast with the flat map (average of the black squares from iteration 7 from Figure 6.2), which is commensurate with the raw contrast that the null eventually degrades to after 30 minutes.

decreased. Future work may involve trying gradient-descent-based approaches to nulling, or keeping a running estimate of the speckle phase. These methods may be faster at combating phase drift, since an update can be made using fewer probes. Future instruments with higher photon sensitivity may also be able to run with lower integration times and therefore at faster rates.

In Figure 6.2b, we compare the fractional co-axial planet throughput obtained with the flat map with the throughput obtained with the final speckle nulling DM solution, and show that speckle nulling does not decrease the throughput η_p. In Figure 6.4a, we plot the average spectrum from the last 9 frames with the flat map and the average spectrum from the last 9 frames with the final speckle nulling DM solution. We also plot the ratio of the two across wavelength to show the spectral shape of the null. The null ratios across the different orders are summarized in Table 6.1. Interestingly, the null is not deepest in the targeted order from 2.29 to 2.34 μm. There is no conclusive explanation for this behavior, but one possible interpretation is that by the time these frames were taken, the underlying phase had already drifted in a way that caused the null to be deeper in an order different from the targeted one. The null is also rather broad, spanning several orders adjacent to the targeted one, and

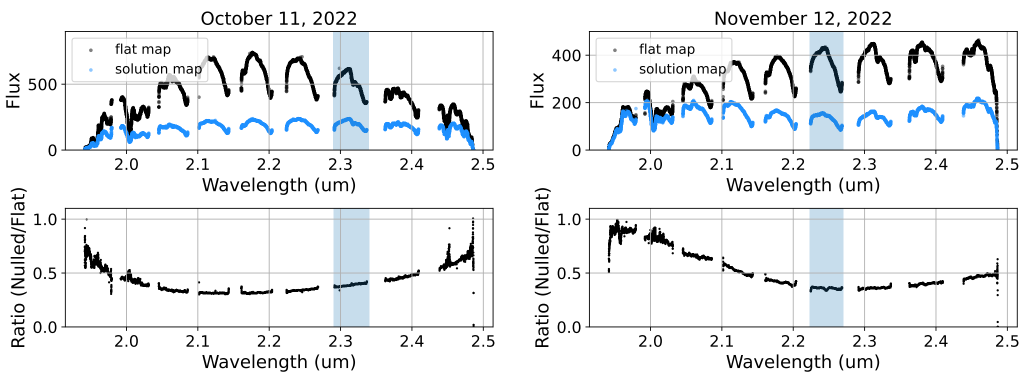

Figure 6.4: Comparison of speckle spectra obtained with and without speckle nulling. a) Data from October 11, 2022. On top, the average spectrum from the last 9 frames with the final speckle nulling DM solution is plotted in blue, and the average spectrum from the last 9 frames with the flat map is plotted in black. Below it, the ratio of the two shows the spectral shape of the null. The wavelength range targeted by speckle nulling (the range over which flux was summed to calculate the raw contrast) is indicated with light blue shading. In the targeted order from 2.29 to 2.34 μm, speckle nulling achieved a suppression ratio of 2.6. b) Data from November 12, 2022. On top, a frame at the end of a speckle nulling iteration is plotted in blue, and the spectrum obtained immediately afterwards with the flat map is plotted in black. Below it, the ratio of the two shows the spectral shape of the null. The wavelength range targeted by speckle nulling (the range over which flux was summed to calculate the raw contrast) is indicated with light blue shading. In the targeted order from 2.22 to 2.27 μm, speckle nulling achieved a suppression ratio of 2.8.

the starlight does not increase in any of the orders as a result of speckle nulling, so data across the entire observed spectrum benefits from using the technique.

Science Observation

We used speckle nulling for a science observation of the same target on November 12, 2022. This time, we targeted the order from 2.22 to 2.27 μm, which has slightly higher flux on the detector, which we believed might provide slightly better signal for speckle nulling. We prioritized obtaining science-suitable exposures (including the probe measurements), so kept the loop closed without stopping to obtain diagnostic engineering data. We only flipped back to the flat map once, immediately after a frame at the null. The comparison of the fluxes from two frames is shown in the right panel of Figure 6.4. The null ratios across the different orders are summarized in Table 6.1.

Table 6.1: Decrease in stellar coupling achieved through on-sky speckle nulling. The observed star is HD 206893, a star with a K-band bagnitude of 5.593. The fiber was placed 91 mas from the star at the predicted location of HD 206893c. The table lists the mean null ratio for each spectral order. Values for the order targeted by speckle nulling are in bold.

Wavelength (μm)	Null Ratio (Oct. 11)	Null Ratio (Nov. 12)
1.94-1.98	1.8	1.1
1.99-2.03	2.3	1.2
2.05-2.09	2.9	1.5
2.10-2.14	3.2	1.9
2.16-2.20	3.1	2.4
2.22-2.27	2.9	**2.8**
2.29-2.34	**2.6**	2.8
2.36-2.41	2.2	2.5
2.44-2.49	1.7	2.1

On this night, speckle nulling achieved a suppression ratio of 2.8 relative to using the flat map. This shows that the gain achieved with speckle nulling is repeatable over different nights.

6.4 Conclusion

We demonstrate the successful on-sky application of speckle nulling through an optical fiber, using a science spectrograph to simultaneously collect science data and measure speckle phase. We achieve a gain in stellar suppression of about 2.6 to 2.8, and show that this gain is repeatable over different nights. This suppression is achieved with minimal impact to the planet throughput (any change is well below what we can measure). Thus, using speckle nulling is expected to decrease the required integration time to reach a desired SNR by a factor or 2.6 to 2.8. The performance of speckle nulling on-sky is likely limited by the drift in phase on the timescale of a phase measurement, likely caused by instabilities in the instrument. Future work could involve exploring gradient-descent algorithms to null starlight, which may be faster and therefore better able to combat the drifting phase. In the meantime, the speckle nulling algorithm demonstrated in this work can be used to decrease stellar leakage and improve the signal-to-noise of science observations.

Chapter 7

CORONAGRAPHIC DATA POST-PROCESSING USING PROJECTIONS ON INSTRUMENTAL MODES

Xin, Yinzi et al. (Mar. 2024). "Coronagraphic Data Post-processing Using Projections on Instrumental Modes". In: *The Astrophysical Journal* 963.2, 96, p. 96. DOI: 10.3847/1538-4357/ad1879. arXiv: 2401.04269 [astro-ph.IM].

This chapter reproduces my paper on modeling a coronagraph instrument to find a subset of the data that is robust to wavefront aberrations, such that projecting the data onto the subspace results in higher signal-to-noise and stronger detections. This work models the Roman Space Telescope Hybrid Lyot Coronagraph; however, while the PLN has much fewer output degrees of freedom, the underlying principles are the same, and a similar concept can be applied, (e.g. by weighing the signal in linear combinations of ports by their relative sensitivity to aberrations). An important finding of this work is that the robust subspace is dependent on the instrument design as well as the level of dynamic wavefront aberrations relative to the static portion of the electric field. Although simple in premise, treating the case of large dynamic aberrations required an expansion of the term quadratic in wavefront error that was very atypical for and not yet explicitly documented within the field. I simulated mock detections with synthetic Roman Hybrid Lyot Coronagraph data to show that using such a projection increased detection sensitivity by about 28% with aberrations in the linear regime, and a factor of 2 with aberrations in the quadratic regime. While a projection is a simple way to implement information about the instrument in the data analysis, future implementations will likely include this model as a statistical prior along with other sources of information. This project has followed me to many places: I began it as a Master's student at MIT, met a key collaborator on a visit to the Observatory Côte d'Azur, continued it as a visiting researcher at the Space Telescope Science Institute, and finished it while a graduate student at Caltech.

Abstract

Directly observing exoplanets with coronagraphs is impeded by the presence of speckles from aberrations in the optical path, which can be mitigated in hardware with wave front control, as well as in post-processing. This work explores using an instrument model in post-processing to separate astrophysical signals from residual aberrations in coronagraphic data. The effect of wave front error (WFE) on the coronagraphic intensity consists of a linear contribution and a quadratic contribution. When either of the terms is much larger than the other, the instrument response can be approximated by a transfer matrix mapping WFE to detector plane intensity. From this transfer matrix, a useful projection onto instrumental modes that removes the dominant error modes can be derived. We apply this approach to synthetically generated Roman Space Telescope hybrid Lyot coronagraph data to extract “robust observables," which can be used instead of raw data for applications such as detection testing. The projection improves planet flux ratio detection limits by about 28% in the linear regime and by over a factor of 2 in the quadratic regime, illustrating that robust observables can increase sensitivity to astrophysical signals and improve the scientific yield from coronagraphic data. While this approach does not require additional information such as observations of reference stars or modulations of a deformable mirror, it can and should be combined with these other techniques, acting as a model-informed prior in an overall post-processing strategy.

7.1 Introduction

Specialist high-contrast techniques are required to directly observe faint astrophysical objects near brighter objects, such as exoplanets, brown dwarfs, or circumstellar disks orbiting much brighter central stars. High contrast observations are essential for answering scientific questions involving binary and planetary system population statistics, planet and disk formation and evolution, planetary atmospheres, and planet habitability and the search for biosignatures (Traub and Oppenheimer, 2010). Measuring these exoplanet signals is difficult because they often lie at small angular separations from their host star and can be many orders of magnitude fainter. One major obstacle for high contrast observations is photon noise from the light of the central star. Practical matters such as detector saturation aside, if the star is orders of magnitude brighter than its companion, the photon noise associated with the outer lobes of star’s point-spread-function (PSF) can overwhelm any signal from the companion, even if the on-axis star’s signal is perfectly known. As a result, instruments to directly suppress starlight, such as coronagraphs and nullers, are important in increasing the photon signal-to-noise (SNR) of faint companions.

Another important source of noise is wavefront error (WFE), which distorts the signal of the on-axis point source. Sources of wavefront error include atmospheric turbulence, imperfections in the optics, or thermo-mechanical changes in the telescope or instrument. At an instant in time, a perturbation to the wavefront scatters energy from the core of the PSF into speckles throughout the image that can resemble off-axis sources. If the magnitude of the WFE electric field is smaller than that of the underlying electric field from the PSF, then the speckles are symmetric about zero in detector plane intensity and average out over time. When the wavefront error is the larger term, as in the case of uncorrected atmospheric turbulence, the speckles are predominantly positive, increasing rather than decreasing the intensity over most of the focal plane, and averaging out to a halo that can obscure off-axis signals. Scattered starlight at larger spatial separations increases the photon noise at those locations in the detector plane, which can also dominate over signals from faint companions.

The goal of high contrast instruments is to separate the signal of the on-axis star from off-axis sources. Coronagraphs are passive optical elements that spatially filter the light to suppress the signal of an on-axis star, reducing its associated photon noise while letting through off-axis signals (O. Guyon et al., 2006b). Adaptive optics (AO; Tyson, 2000) and focal plane wavefront control (Groff et al., 2015) actively correct for wavefront error to reduce their impact. However, even with suppression from coronagraphs or nullers, the sensitivity to faint astrophysical signals is still limited by residual starlight and its associated photon noise.

Post-processing techniques can use additional available information to further mitigate the effects of WFE and increase sensitivity to real astrophysical signals (see Cantalloube et al. (2022) for a discussion of the state-of-the-art of high contrast post-processing in the context of a direct imaging data challenge). For example, angular differential imaging (ADI) exploits observations at different roll angles, taking advantage of azimuthal averaging of the wavefront error (Christian Marois, Lafreniere, et al., 2006; Flasseur et al., 2018). Other methods rely on performing principal component analysis (PCA) on reference observations of a calibration star similar to the host star, but without astrophysical companions, to calibrate out residual static or quasi-static starlight (Lafrenière et al., 2007; Soummer, Laurent Pueyo, and Larkin, 2012b; Laurent Pueyo, 2016b). Additional sources of information on residual WFE include telemetry from wavefront sensing and control (WFSC) systems such as wavefront sensor residuals (Vogt et al., 2011) or focal plane electric field estimates (Pogorelyuk, Kasdin, and Rowley, 2019), data from a self-coherent camera (P. Baudoz et al., 2006b), and data at different wavelengths as exploited in spectral deconvolution (Sparks and Ford, 2002).

This work shows that the modeled or measured instrument sensitivity to wavefront error can be included as an additional source of information in the post-processing of coronagraphic data, information that, in theory, can be combined with the other techniques discussed. This work examines an approach that uses this physical optics model to construct a projection removing the dominant error modes in the appropriate wavefront error regime, and finds that this can improve sensitivity to faint companions by up to and over a factor of 2.

7.2 Coronagraphic Signals

Data Formation

The model used in this work assumes the light through the instrument is monochromatic. With a discrete representation of the optical planes of an instrument, a coronagraph can be modeled as a linear operator $\mathbf{C}$, a constant 2D matrix transforming the electric field vector at the pupil plane, $\boldsymbol{E_s}$, into the electric field vector at the detector plane, $\boldsymbol{E}_{\mathbf{det}}$. If $\boldsymbol{E}_{s\mathbf{0}}$ is the electric field vector of the central source (star) at the pupil plane in the absence of aberrations, and $\boldsymbol{\Delta E_s}$ a vector of small perturbations to that electric field, representing wavefront aberrations (which can be variable in time), then the electric field vector at the detector plane, assuming that the star is the only source of light, is

$$\boldsymbol{E}_{\mathbf{det}} = \mathbf{C}\boldsymbol{E}_s = \mathbf{C}\boldsymbol{E}_{s\mathbf{0}} + \mathbf{C}\boldsymbol{\Delta E}_s(t). \tag{7.1}$$

The intensity measured is the element-wise norm squared of the detector plane electric field (here, $\overline{x}$ indicates the element-wise complex conjugate of x and $\circ$ indicates the element-wise product):

$$\begin{aligned}\boldsymbol{I}_s &= |\boldsymbol{E}_{\mathbf{det}}|^2 \\ &= |\mathbf{C}\boldsymbol{E}_{s\mathbf{0}}|^2 + 2\,\mathrm{Re}\left\{\overline{(\mathbf{C}\boldsymbol{E}_{s\mathbf{0}})} \circ \mathbf{C}\boldsymbol{\Delta E}_s(t)\right\} + |\mathbf{C}\boldsymbol{\Delta E}_s(t)|^2.\end{aligned} \tag{7.2}$$

The vector of the pupil plane electric field of a binary companion is given by

$$\boldsymbol{E}_{\boldsymbol{p}} = \sqrt{c}\boldsymbol{E}_{s\mathbf{0}}e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} + \sqrt{c}\boldsymbol{\Delta E}_s(t)e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} = \boldsymbol{E}_{\boldsymbol{p}\mathbf{0}} + \boldsymbol{\Delta E}_{\boldsymbol{p}}(t), \tag{7.3}$$

where c is the flux ratio between the planet and the star, $\boldsymbol{k}$ is the pupil plane wave vector indicating the companion's location, and $\boldsymbol{x}$ is the pupil plane coordinate vector. Namely, the planet's pupil-plane electric field is the star's electric field, but tilted and scaled by the square root of the flux ratio. The detector plane intensity for the planet can be expressed as

$$\boldsymbol{I_p} = |\mathbf{C}\boldsymbol{E_{p0}}|^2 + 2\,\mathrm{Re}\left\{\overline{(\mathbf{C}\boldsymbol{E_{p0}})} \circ \mathbf{C}\boldsymbol{\Delta E_p}(t)\right\} + |\mathbf{C}\boldsymbol{\Delta E_p}(t)|^2. \tag{7.4}$$

The total intensity on the detector plane from the star and the planet is the sum of Eqs. 7.2 and 7.4. However, we can make two simplifying assumptions. The first assumption is that the flux of the planet is small relative to the flux of the star, such that $c \ll 1$. The second assumption is that the magnitude of the wavefront error is small relative to the total magnitude of the electric field, namely $\boldsymbol{\Delta E_s}(t) \ll \boldsymbol{E_{s0}}$, which implies $\boldsymbol{\Delta E_p}(t) \ll \boldsymbol{E_{p0}}$. This is true if we are both in the "small phase regime" (when there is much less than one wave of wavefront error) and the fractional amplitude error is much less than one. These assumptions imply that the last two terms of Eq. 7.4 are small relative to the other terms, so we can approximate the total intensity as

$$\boldsymbol{I_{\mathrm{tot}}} \approx |\mathbf{C}\boldsymbol{E_{s0}}|^2 + 2\,\mathrm{Re}\left\{\overline{(\mathbf{C}\boldsymbol{E_{s0}})} \circ \mathbf{C}\boldsymbol{\Delta E_s}(t)\right\} + |\mathbf{C}\boldsymbol{\Delta E_s}(t)|^2 + |\mathbf{C}\boldsymbol{E_{p0}}|^2. \tag{7.5}$$

The first term $|\mathbf{C}\boldsymbol{E_{s0}}|^2$ is the residual starlight not blocked by the coronagraph in the case of no aberrations. The second term $2\,\mathrm{Re}\left\{\overline{(\mathbf{C}\boldsymbol{E_{s0}})} \circ \mathbf{C}\boldsymbol{\Delta E_s}\right\}$ is linear in the wavefront aberration and corresponds to the interference of the aberration, propagated to the focal plane, with the underlying residual starlight from the coronagraph, analogous to speckle pinning (Bloemhof, 2002; Perrin et al., 2003). The third term $|\mathbf{C}\boldsymbol{\Delta E_s}|^2$ is the quadratic term, corresponding to the norm squared of the wavefront error propagated to the focal plane. The last term $|\mathbf{C}\boldsymbol{E_{p0}}|^2$ is the nominal off-axis signal of interest.

Whether the effect of wavefront errors at some location in the detector plane are dominated by the linear term or the quadratic term depends on the attenuation of starlight by the coronagraph and the level of the propagated wavefront error at that location. If the propagated wavefront aberrations are smaller in complex amplitude than the residual starlight after the coronagraph with no aberrations, the linear term is dominant. When a coronagraph is not used, this corresponds to the speckle pinning regime, in which the aberrations primarily interfere with the wings of the telescope's PSF (Bloemhof, 2002). The same phenomenon occurs with a coronagraph; however, as the amplitude of the PSF wings are reduced by the coronagraph, the range of WFE over which this occurs is much more limited. Otherwise, if the propagated wavefront aberrations have relatively larger magnitudes, the quadratic term is dominant. For a given location in the focal plane, the local point of transition between the linear and quadratic regimes occurs when $|2\,\mathrm{Re}\left\{\overline{(\mathbf{C}\boldsymbol{E}_{s\mathbf{0}})} \circ \mathbf{C}\boldsymbol{\Delta E}_s\right\}| = |\mathbf{C}\boldsymbol{\Delta E}_s|^2$, or roughly when $|2\mathbf{C}\boldsymbol{E}_{s\mathbf{0}}| = |\mathbf{C}\boldsymbol{\Delta E}_s|$.

This point of transition is different for each pixel, and also depends on the coronagraph design as well as the "nominal" wavefront (whether it is flat, as is typical for ground-based coronagraphs, or the wavefront corresponding to a dark hole, as is planned for space-based coronagraphs). For this work, we use as an example the Hybrid Lyot Coronagraph (HLC) of the Coronagraph Instrument of the Roman Space Telescope. With the dark hole presented in Section 7.5, which has an average raw contrast (residual stellar intensity divided by unocculted peak intensity) of 5.6×10^{-9}, the point at which $|\mathbf{C}\boldsymbol{\Delta E}_s| > |2\mathbf{C}\boldsymbol{E}_{s\mathbf{0}}|$ for 50% of the pixels in the dark hole region occurs at roughly 0.1 waves root-mean-square (RMS) of phase error, on average. This means that wavefront errors less than 0.1 waves RMS will primarily be in the linear regime, while wavefront errors larger than 0.1 waves RMS will primarily be in the quadratic regime, although this is somewhat dependent on the form of the wavefront's spatial power spectral density (PSD) that we use in Section 7.5.

In this work, robust observables are only formulated for WFE that is predominantly linear or predominantly quadratic throughout the entire focal plane. However, it may be possible to obtain robust observables for when both terms have comparable contributions, a topic that is left for future work.

Linear Regime

From Equation 7.5, if we then assume that the linear error term is dominant, then we can drop the quadratic contribution such that the detector plane intensity is approximately

$$\boldsymbol{I}_{\mathbf{tot},\boldsymbol{l}} \approx |\mathbf{C}\boldsymbol{E}_{s\mathbf{0}}|^2 + 2\,\mathrm{Re}\left\{\overline{(\mathbf{C}\boldsymbol{E}_{s\mathbf{0}})} \circ \mathbf{C}\boldsymbol{\Delta E}_s(t)\right\} + |\mathbf{C}\boldsymbol{E}_{p\mathbf{0}}|^2. \tag{7.6}$$

The contribution of the wavefront error to the intensity can be expressed as a linear transformation $\mathbf{A_l}$ acting on the wavefront error:

$$\boldsymbol{I}_{\textbf{tot},\boldsymbol{l}} \approx |\mathbf{C}\boldsymbol{E}_{\boldsymbol{s0}}|^2 + \mathrm{A}_{\mathrm{l}}\boldsymbol{\varDelta E}_{\boldsymbol{s}}(t) + |\mathbf{C}\boldsymbol{E}_{\boldsymbol{p0}}|^2. \tag{7.7}$$

The transfer matrix $\mathbf{A_l}$ can be calculated semi-analytically from the coronagraph operator and the unaberrated electric field, as derived from Equation 7.6:

$$\mathrm{A}_{1kj} = \frac{\partial I_k}{\partial \varDelta E_{s_j}} = 2\,\mathrm{Re}\left\{(\sum_i \mathrm{C}_{ki} E_{s0_i})^* \mathrm{C}_{kj}\right\}. \tag{7.8}$$

The indices i and j label the input basis vectors used to represent the wavefront error, and the index k labels the detector pixel.

It is desirable to reduce the term dependent on WFE, $\mathbf{A_l}\boldsymbol{\varDelta E_s}(t)$, relative to the terms containing astrophysical signals of interest. This can be achieved by left-multiplying the measured intensities by a matrix $\mathbf{K_l}$, that projects out the dominant modes of $\mathbf{A_l}$. The following Section 7.3 describes the process of calculating $\mathbf{A_l}$ and finding from it an appropriate $\mathbf{K_l}$. The observables obtained using projection matrix $\mathbf{K_l}$ are given by

$$\boldsymbol{O_l} = \mathrm{K}_{\mathrm{l}}\boldsymbol{I}_{\textbf{tot},\boldsymbol{l}}. \tag{7.9}$$

When the wavefront errors are in the linear regime, this projection is expected to suppress the contribution of wavefront errors to the measured data. As long as the measurements retain most of the astrophysical signal, then the projection will boost its SNR.

Quadratic Regime

In the quadratic-dominated regime, we can drop the linear contribution in Eq. 7.5, such that the detector plane intensity is approximately

$$\boldsymbol{I}_{\textbf{tot},\boldsymbol{q}} \approx |\mathbf{C}\boldsymbol{E}_{\boldsymbol{s0}}|^2 + |\mathbf{C}\boldsymbol{\varDelta E}_{\boldsymbol{s}}(t)|^2 + |\mathbf{C}\boldsymbol{E}_{\boldsymbol{p0}}|^2. \tag{7.10}$$

In a discrete numerical model, the contribution of the quadratic term to each pixel labeled k in the detector plane can be expressed as

$$|\mathbf{C}\boldsymbol{\varDelta E}_{\boldsymbol{s}}|_k^2 = (\sum_m \mathrm{C}_{km}\varDelta E_{s_m})^*(\sum_n \mathrm{C}_{kn}\varDelta E_{s_n}) = \sum_i \sum_j \varDelta E_{s_i}^* \mathrm{M}_{kij}\varDelta E_{s_j}. \tag{7.11}$$

The indices m, n, i, and j label the input basis vectors used to represent the wavefront error (expressed here in terms of perturbation to the complex electric field), and the index k labels the detector pixel. The quantity $\hat{\mathbf{M}}$ with elements M_{kij} is a 3-tensor containing the second order partial derivative matrix (Hessian) of each pixel intensity with respect to the wavefront error, and relates each pairwise combination of pupil basis vectors to its effect on each detector plane pixel k. Each entry can be calculated semi-analytically from the coronagraph operator using the following formula derived from Equation 7.11:

$$\mathrm{M}_{kij} = \frac{\partial^2 I_k}{\partial \mathit{\Delta} E^*_{s_i} \partial \mathit{\Delta} E_{s_j}} = \mathrm{C}_{ki}\mathrm{C}^*_{kj} + \mathrm{C}_{kj}\mathrm{C}^*_{ki}. \tag{7.12}$$

Assuming there are N_{pix} pixels of interest on the detector, and N basis vectors are used to represent the wavefront error, then, through a remapping, the 3-tensor $\hat{\mathbf{M}}$ of size $(N_{\text{pix}} \times N \times N)$ can be expanded into a matrix acting on the space of all *pairwise combinations* of pupil basis vectors. Since Hessians are symmetric because partial derivatives commute ($\mathrm{M}_{kij} = \mathrm{M}_{kji}$), the ordering of each pair of segments does not matter, and the derivatives corresponding to the same pair of original basis vectors can be consolidated into the same entry. This results in an input vector space of size $\binom{N+1}{2}$, or the number of pairwise combinations of pupil basis vectors.

The 3-tensor $\hat{\mathbf{M}}$ can thus be represented as a $(N_{\text{pix}} \times \binom{N+1}{2})$ matrix $\mathbf{A_q}$ of second derivatives, acting on a vector $\boldsymbol{\beta}$ of perturbations defined for each pairwise combination $\Delta E_{s_i} \Delta E_{s_j}$. This results in the following expression for the quadratic term:

$$|\mathbf{C}\boldsymbol{\mathit{\Delta} E_s}|^2 = \mathbf{A_q}\boldsymbol{\beta}. \tag{7.13}$$

The projection is similar to the linear case: the detector intensities can be left-multiplied by a matrix $\mathbf{K_q}$ that projects out the dominant quadratic error modes of $\mathbf{M}$. The observables with the appropriate projection $\mathbf{K_q}$ are given by

$$\boldsymbol{O_q} = \mathbf{K_q}\boldsymbol{I}_{\textbf{tot},\boldsymbol{q}}. \tag{7.14}$$

7.3 Response Matrices and Robust Observables

Calculating the Response Matrix

This section details the numerical calculation of instrument response matrices and the projection matrices. In this work, the response matrix is calculated with the wavefront aberrations represented in the Zernike basis. In this basis, ΔE_{Z_n} is the coefficient of the aberration induced by the n^{th} Noll ordered Zernike polynomial (Noll, 1976), and N is the total number of polynomials chosen to construct the response matrix:

$$\boldsymbol{\Delta E_s} = \begin{pmatrix} \Delta E_{Z_1} \\ \dots \\ \Delta E_{Z_N} \end{pmatrix}. \tag{7.15}$$

We define N_{pix} as the total number of detector pixels of the optical model and N_{basis} as the number of Zernike modes to include. The coronagraph operator **C** is the $N_{\text{pix}} \times N_{\text{basis}}$ matrix that, when applied to a vector of Zernike coefficients, gives the perturbation they induce in the focal plane electric field. This operator is typically either already part of the optical model, or obtainable by propagating Zernike modes through the optical model and using finite differences to populate its columns. Given the operator **C** and the initial unaberrated focal plane electric field, we can calculate both $\mathbf{A_l}$ and $\mathbf{A_q}$ using Equations 7.8 and 7.12. Note that the term $(\sum_k \mathrm{C}_{kj} E_{s0_j})$ in Equation 7.8 is simply the initial unaberrated focal plane electric field at pixel k. For more complicated models without simple analytical solutions (such as those that include distortion), automatic differentiation, in which arbitrary exact derivatives can be computed without finite differences, may be useful (Pope et al., 2021).

The linear transfer matrix poses no computational problems, as its size is $N_{\text{pix}} \times N_{\text{basis}}$. However, for the quadratic transfer matrix $\mathbf{A_q}$, the size of the input dimension quickly becomes computationally burdensome for high N_{basis}. For the example system shown in Section 7.5, a N_{basis} of 528 results in a $\mathbf{A_q}$ matrix of width $\binom{N_{\text{basis}}+1}{2} = 139,656$ (the number of pairwise combinations of pupil basis vectors), and length 5,476 (the number of detector pixels of the model). This $\mathbf{A_q}$ matrix, when represented as (non-complex) doubles, is over 6GB in size. As explained in Section 7.3, the calculation of the projection matrix involves a singular value decomposition (SVD) of the response matrix. Since calculating the SVD of a matrix of this size is too computationally expensive, we restrict our quadratic transfer matrix to only include the first $N_{\text{redu}} = 100$ Zernikes, which results in an $\mathbf{A_q}$ with a width of only $\binom{N_{\text{redu}}+1}{2} = 5050$. This model is valid only in a smaller area closer to the central star — namely within $\sim 5\,\lambda/D$, where λ is the wavelength and D the telescope diameter. However, in Appendix 7.6, we explore using an approximation of the quadratic transfer matrix that can extend the area of applicability while circumventing impractical computational costs.

Calculating the Projection Matrix

Once a response matrix $\mathbf{A}$ is obtained, a singular value decomposition of $\mathbf{A} = \mathbf{USV}^T$ is performed, revealing its singular modes and corresponding singular values. Then a choice of the number of modes to project out (N_m) is made. The remaining $N_{\text{pix}} - N_m$ modes are kept in the post-processing projection $\mathbf{K}$. Accordingly, $\mathbf{K}$ is the subset of $\mathbf{U}$ that contains the $m + 1^{th}$ and higher left singular modes of $\mathbf{A}$. A pseudo-code summary of the process to find $\mathbf{K}$ is given in Algorithm 1. The optimal N_m depends on the signal of interest. For the point-source companion signals explored in this work, N_m is chosen as the cutoff that results in the best detection limit at the separation of interest.

Algorithm 1 Calculate projection matrix $\mathbf{K}$

Input: Transfer matrix $\mathbf{A}$
Input: Cutoff mode (number of modes to project out), N_m
Input: Indices of detector plane pixels in region of interest, idx
Output: Projection matrix $\mathbf{K}$

$U, S, V^T \leftarrow \text{svd}(\mathbf{A})$

$\mathbf{K} \leftarrow \text{transpose}(U(\text{idx}, N_m + 1 : \text{end}))$

If the linear and quadratic projections are used in the appropriate regimes to increase SNR, they could, for example, allow for a binary signal detection with a deeper flux ratio than using the raw intensity data. Detection tests can be performed on both projected and unprojected data to quantify this effect.

7.4 Detection Testing

Detections are typically claimed from a statistical hypothesis test (see e.g. Kasdin and Braems, 2006; Jensen-Clem et al., 2017; Ceau et al., 2019). A test statistic T is calculated from the data and compared to a threshold ξ. A detection is claimed if $T \geq \xi$, and a lack of a detection is claimed otherwise. The fraction of real companions detected is the true positive rate (TPR). A false positive occurs if there is no companion in the data, but the detection test incorrectly claims a detection. The rate at which this occurs is the false positive rate (FPR).

As the detection threshold ξ is decreased, detecting real companions becomes more likely, but false detections also become more likely (Jensen-Clem et al., 2017). Varying the threshold and plotting the TPR as a function of the FPR results in a receiver operating characteristic (ROC) curve, an example of which is shown in Figure 7.6. ROC curves characterize the performance of a detection scheme and are used in the determination of flux ratio detection limits.

This work uses a simple Delta Reduced Chi-squared ($\Delta\chi_r^2$) statistic, or the difference in the reduced χ^2 of the data assuming it contains only noise, and the reduced χ^2 of the data assuming it contains noise and the companion signal. The formula for calculating this test statistic from the data is given by Equation 7.16 (the bars indicate vector norm, the divisions are element-wise):

$$\Delta\chi_r^2 = \frac{1}{\nu}\left(\left|\frac{\mathbf{y}}{\boldsymbol{\sigma}}\right|^2 - \left|\frac{\mathbf{y}-\mathbf{x}}{\boldsymbol{\sigma}}\right|^2\right). \tag{7.16}$$

In this formula, $\mathbf{y}$ is the data, which is the synthetically generated realizations of $\boldsymbol{I}_{\mathbf{tot}}$, with or without a planet. Meanwhile, $\mathbf{x}$ is the unaberrated model of the planet signal $\boldsymbol{I}_{\boldsymbol{p}\mathbf{0}} = |\mathbf{C}\boldsymbol{E}_{\boldsymbol{p}\mathbf{0}}|^2$ (assuming that it is known, such as through a maximum-likelihood-estimation). The estimated uncertainty of the data is denoted by $\boldsymbol{\sigma}$, and ν is the degrees of freedom (the number of data elements minus the number of free parameters; a binary system's three free parameters are the flux ratio, separation, and position angle). This use of this test statistic is motivated by an assumption that the noise is Gaussian and uncorrelated, under which this quantity is related to the relative log-probabilities of the data containing both the planet signal and noise, versus containing only noise. The noise being uncorrelated and Gaussian is generically not the case. However, the effects of the correlation and non-Gaussianity of the injected noise on the resulting test statistic distributions is properly simulated and captured by the Monte Carlo methods used in this work.

7.5 Example: Nancy Grace Roman Space Telescope Hybrid Lyot Coronagraph

In this section, the use of robust observables with the Hybrid Lyot Coronagraph of the Roman Space Telescope is analyzed through simulation. However, this approach could also be applied to other coronagraphs, as long as the exposure times are short enough such that wavefront error has not been averaged out. The optical model of CGI is shown in Figure 7.1 (Kasdin, Bailey, et al., 2020). The optical elements corresponding to the HLC mode (the relevant mode for this work) are depicted in the top panel.

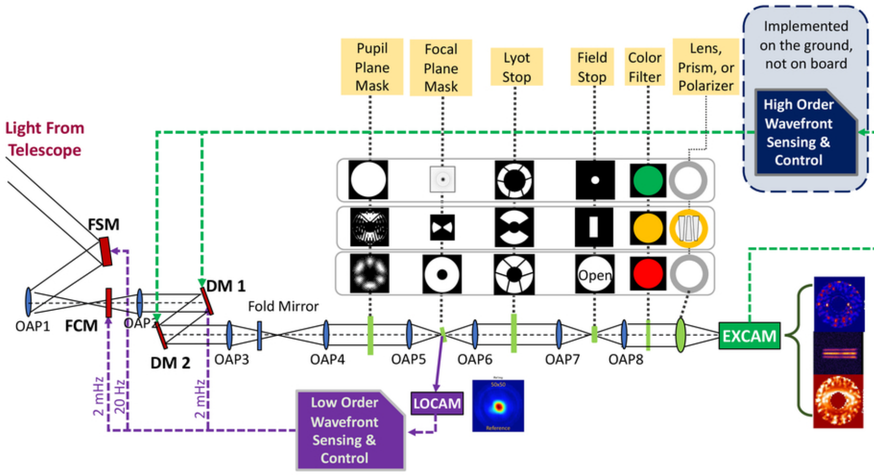

Figure 7.1: The CGI optical train and wavefront sensing and control architecture. The optical elements of the HLC mode of interest are depicted in the top panel. Before an observation, the high order wavefront sensing and control loop is performed on a bright reference star to generate a 'dark hole' (an area where starlight is suppressed). Then, the DM shapes are fixed, and the telescope slews to the target star for the observation. During the observation, wavefront errors accrue as a result of instrumental disturbances and drifts, the effects of which this work aims to mitigate in post-processing. Figure from Kasdin, Bailey, et al. (2020).

Optical Model

The HLC operates around a dark hole state, which is obtained using focal plane wavefront control (with deformable mirrors) to measure and minimize the electric field in the detector plane. Such focal plane wavefront control significantly suppresses the amount of starlight in the dark hole, and allows for much deeper raw contrasts than with just a flattened wavefront. Before an observation, the dark hole is generated using high order wavefront sensing and control loop on a bright reference star. Then, the DM shapes are fixed, and the telescope slews to the target star for the observation. During the observation, wavefront errors accrue as a result of instrumental disturbances and drifts. This work aims to mitigate the effects of those wavefront errors in post-processing. Note that as a result of the dark-hole generation, the nominal electric field $\boldsymbol{E}_{s0}$ is not a flat wavefront, but the pupil plane electric-field obtained at the end of the dark-hole generation sequence.

A Lightweight Space Coronagraph Simulator (LSCS) [1] derived from the HLC model in the Fast Linearized Coronagraph Optimizer (FALCO; Riggs et al., 2018) toolbox

[1] https://github.com/leonidprinceton/LightweightSpaceCoronagraphSimulator

is used for the following simulations. The LSCS relies on the HLC numerical model and focal-plane wavefront control algorithm included in FALCO to first generate the initial dark hole electric field. The numerical model in FALCO is also used to calculate **C** from the finite-difference sensitivities of the focal plane electric field to pupil plane phase error expressed in the Zernike basis (we have made the assumption that the matrix transformation is approximately linear in phase, valid when the phase error is much less than a wave). Although we use finite-differences to calculate **C**, one could also construct it by multiplying together all the matrix transformations of the optical model. These simulations are conducted at a single wavelength of 546 nm.

The average raw contrast of the initial dark hole is 5.6×10^{-9}. The LSCS model takes in Zernike coefficients for phase aberrations, calculates their effect on the focal plane electric field, and adds them to the initial dark hole electric field to obtain the focal plane electric field in the presence of wavefront errors. The intensity can be calculated as the norm-square of the focal plane electric field. Detector and photon noise are not simulated. Since the default LSCS models only the first 136 Zernikes, FALCO is first used to extend the LSCS model to 528 Zernikes in order for the entire dark hole to be sampled.

This results in using 528 Zernikes to sample the entire dark hole, or a N_{basis} of 528. The LSCS models a detector that is 74×74 pixels, with 3 pixels per λ/D, for a total pixel number of $N_{\text{pix}} = 5476$. The number of pixels defined to be in the dark hole is $N_{\text{DH}} = 2608$. This model does not consider the effects of amplitude errors, and only analyzes phase errors, which, from end-to-end modeling of Roman CGI, are expected to be the dominant form of dynamic aberrations (J. E. Krist et al., 2023). However, for a system where dynamic amplitude errors are comparable to dynamic phase errors, both should be included.

Response Matrices

The Zernike coefficient drift values from the Observing Scenario simulations (OS 9; J. Krist, 2020), based on physical modeling of the telescope, indicate that the WFE expected on Roman will fall within the linear regime of this dark hole. However, the level of wavefront error may end up being higher than currently expected. Additionally, on ground based telescopes, wavefront error from adaptive optics residuals is typically in the quadratic regime. Therefore, for illustrative purposes, both a linearly-dominated noise model and quadratically-dominated noise model are examined.

The matrices $\mathbf{A_l}$ and $\mathbf{A_q}$ are calculated according to Section 7.3. The linear matrix includes all Zernikes present in the optical model, and thus has an input dimension of $N_{\text{basis}} = 528$. The quadratic matrix includes only the first 100 Zernikes, and thus has an input dimension of $\binom{N_{\text{redu}}+1}{2} = 5,050$. The relevant dimensions of the objects used in this analysis are listed in Table 7.1. Note that the cutoff number N_m is a variable to optimized over.

Table 7.1: Quantities and Dimensions for Analysis of the Roman Space Telescope HLC

Quantity	Description	Dimension (Dependency)	Dimension (Value)
$\boldsymbol{E}_{\text{pup}}$	Vector of electric in pupil plane	N_{basis}	528
$\boldsymbol{E}_{\text{det}}$	Vector of detector plane electric field	N_{pix}	$5,476$
$\boldsymbol{I}_{\text{det}}$	Vector of detector plane intensity	N_{pix}	$5,476$
$\boldsymbol{E}_{\text{DH}}$	Vector of detector plane electric field in dark hole	N_{DH}	$2,608$
$\boldsymbol{I}_{\text{DH}}$	Vector of detector plane intensity in dark hole	N_{DH}	$2,608$
$\mathbf{A_l}$	Linear-regime instrument response matrix	$N_{\text{pix}} \times N_{\text{basis}}$	$5,476 \times 528$
$\mathbf{U_l}$	Left singular matrix of $\mathbf{A_l}$	$N_{\text{pix}} \times N_{\text{pix}}$	$5,476 \times 5,476$
$\mathbf{S_l}$	Singular value matrix of $\mathbf{A_l}$	$N_{\text{pix}} \times N_{\text{basis}}$	$5,476 \times 528$
$\mathbf{V_l}$	Right singular matrix of $\mathbf{A_l}$	$N_{\text{basis}} \times N_{\text{basis}}$	528×528
$\mathbf{K_l}$	Linear-regime projection matrix	$(N_{\text{pix}} - N_m) \times N_{\text{DH}}$	$(N_{\text{pix}} - N_m) \times 2,608$
$\boldsymbol{O_l}$	Vector of linear-regime observables	$(N_{\text{pix}} - N_m)$	$(N_{\text{pix}} - N_m)$
$\mathbf{A_q}$	Quadratic-regime instrument response matrix	$N_{\text{pix}} \times \binom{N_{\text{redu}}+1}{2}$	$5,476 \times 5,050$
$\mathbf{U_q}$	Left singular matrix of $\mathbf{A_q}$	$N_{\text{pix}} \times N_{\text{pix}}$	$5,476 \times 5,476$
$\mathbf{S_q}$	Singular value matrix of $\mathbf{A_q}$	$N_{\text{pix}} \times \binom{N_{\text{redu}}+1}{2}$	$5,476 \times 5,050$
$\mathbf{V_q}$	Right singular matrix of $\mathbf{A_q}$	$\binom{N_{\text{redu}}+1}{2} \times \binom{N_{\text{redu}}+1}{2}$	$5,050 \times 5,050$
$\mathbf{K_q}$	Quadratic-regime projection matrix	$(N_{\text{pix}} - N_m) \times N_{\text{DH}}$	$(N_{\text{pix}} - N_m) \times 2,608$
$\boldsymbol{O_q}$	Vector of quadratic-regime observables	$(N_{\text{pix}} - N_m)$	$(N_{\text{pix}} - N_m)$

Projection Matrices

According to Algorithm 1, a singular value decomposition of $\mathbf{A} = \mathbf{USV}^T$ is performed for each transfer matrix, revealing their singular modes and corresponding singular values. The singular values of the transfer matrices are shown in the top of Figure 7.2. The first 10 singular modes of each transfer matrix as represented in the detector plane intensity basis (with pixels not in the dark hole masked) are plotted in Figure 7.3.

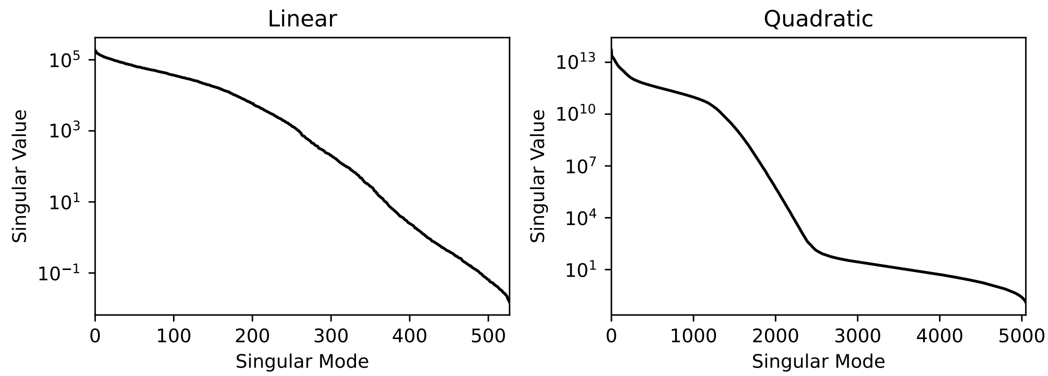

Figure 7.2: The singular values of $\mathbf{A_l}$ (left) and $\mathbf{A_q}$ (right). Note that the transfer matrices are rectangular and have $N_{\text{pix}} = 5476$ total singular modes, but the singular values beyond the size of the input dimension are all 0.

From both the linear and the quadratic transfer matrix, model-based projection matrices with a range of cutoff modes are calculated according to Algorithm 1. To rule out the effect of dimensionality alone on the dataset, random projection matrices of the same size are also generated. This is done by taking the SVD of a matrix the same size as the $\mathbf{A}$ matrices, but populated with values drawn uniformly from -1 to 1, and then removing the same number of dominant modes as is done with $\mathbf{A}$. These matrices are applied to synthetically generated data to quantify their effect on the detectability of binary companion signals.

Synthetic Data Analysis

Synthetic Data Generation

FALCO is used to generate a library of off-axis PSFs corresponding to the dark hole state, which can be injected as binary companions. These off-axis PSFs do not incorporate any WFE that is added on top of the dark hole state. However, the effect of WFE on the off-axis signal is expected to be much, much smaller than its effect

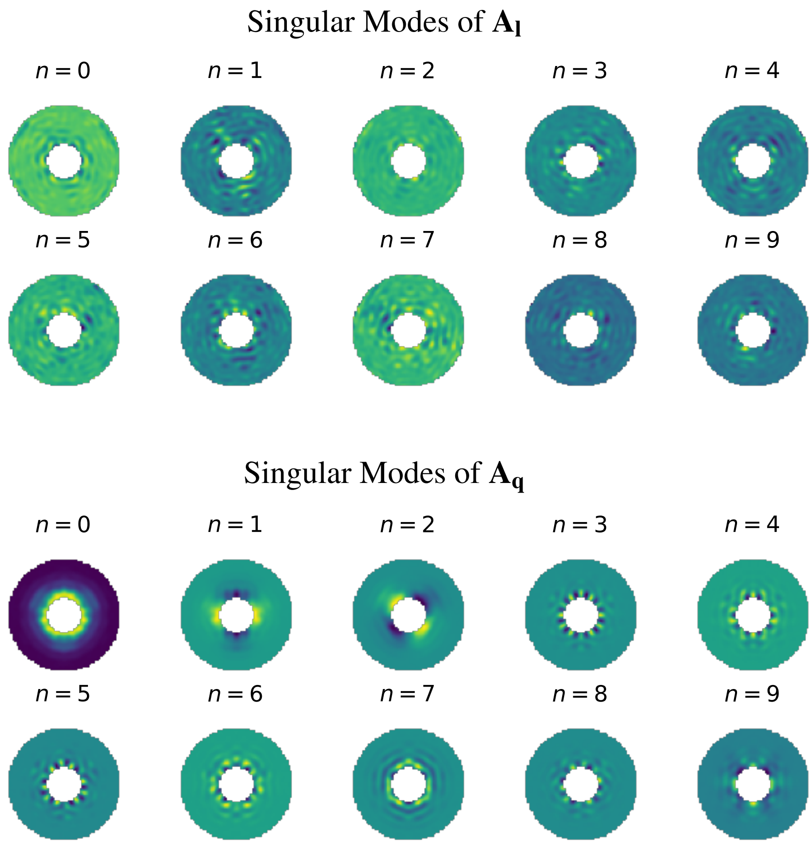

Figure 7.3: Top: The first 10 singular modes of $\mathbf{A_l}$ as represented in the the detector plane intensity basis (linear scale). Bottom: The first 10 singular modes of $\mathbf{A_q}$ as represented in the the detector plane intensity basis (linear scale). The HLC design is nearly circularly symmetric, broken only by the six secondary mirror struts (which can also be seen in the Lyot stop). Because the quadratic transfer matrix depends only on the coronagraph operator $\mathbf{C}$, its singular modes exhibit cosine and sine-like azimuthal behavior associated with circularly symmetric operators. However, the linear transfer matrix depends on both $\mathbf{C}$ as well as on the focal-plane electric field at the end of dark hole creation, which is random and not circularly symmetric. Thus, its singular modes show no such symmetry structures. These singular modes correspond to the intensity patterns most likely to be attributed to wavefront error. Meanwhile, the companion's intensity pattern (the PSF at its location) overlaps very little with these dominant modes, so its signal is mostly retained when the dominant modes are projected out.

on the on-axis stellar signal, so not modeling the effects of wavefront error on the off-axis signal should have a negligible impact on the data.

The optical system is first initialized in the dark hole state. Two noise models are considered, one in the linear regime, and one in the quadratic regime. Each dataset thus consists of 20 instantaneous frames of independent noise realizations. For each frame, the spatial PSD given in Equation 7.17 is used to generate the wavefront error.

$$PSD(n_z) = a{n_z}^b \tag{7.17}$$

In this equation, n_z is Noll-ordered index of the Zernike coefficient. The normalization parameter a is chosen to be 10 nm for the linear regime, and 130 nm for the quadratic regime. The power law exponent b is chosen to be -2. These PSDs correspond to an average wavefront error (calculated over 100 realizations) of about 7 nm (0.013 waves) RMS for the linear regime data, and about 110 nm (0.2 waves) RMS for the quadratic regime data. As discussed in Section 7.2, the linear-quadratic transition occurs at approximately 0.1 waves RMS. Although 110 nm RMS of dynamic wavefront error is unrealistically high for the Roman HLC, we include this regime for demonstration purposes, as this level of WFE would be relevant on ground-based telescopes.

The resulting 528 Zernike coefficients are propagated through the LSCS to calculate the resulting dark hole intensities. In order to create data with an injected companion planet, the off-axis PSF at the desired separation is scaled by the companion's flux ratio, and then added to the dark hole intensity. The separation of the injected companion is set to be $6.5\,\lambda/D$ in the linear case (which is the middle of the dark hole) and $4.0\,\lambda/D$ in the quadratic case (since the model is only valid within $\sim 5\,\lambda/D$). The position angles of both are set to be 0. Frames without the injected companion are used for the control case. Figure 7.4 shows example data frames: the initial dark hole, example frames with the aberrations from both noise models applied, and the same frames with injected companion signals. The flux ratio of the companion is 2×10^{-7} for frame with linear-regime errors and 5×10^{-6} for the frame with quadratic-regime errors. These flux ratios correspond to particularly bright planets chosen to be visible by eye.

It is worthwhile to examine how well the response matrices calculated in Section 7.5 can reconstruct the intensity errors present in the synthetic data. Figure 7.5 compares the intensity error resulting from WFE as calculated from the optical model with the intensity error calculated by multiplying the WFE by the appropriate response matrix, for example frames in both the linear and the quadratic regimes.

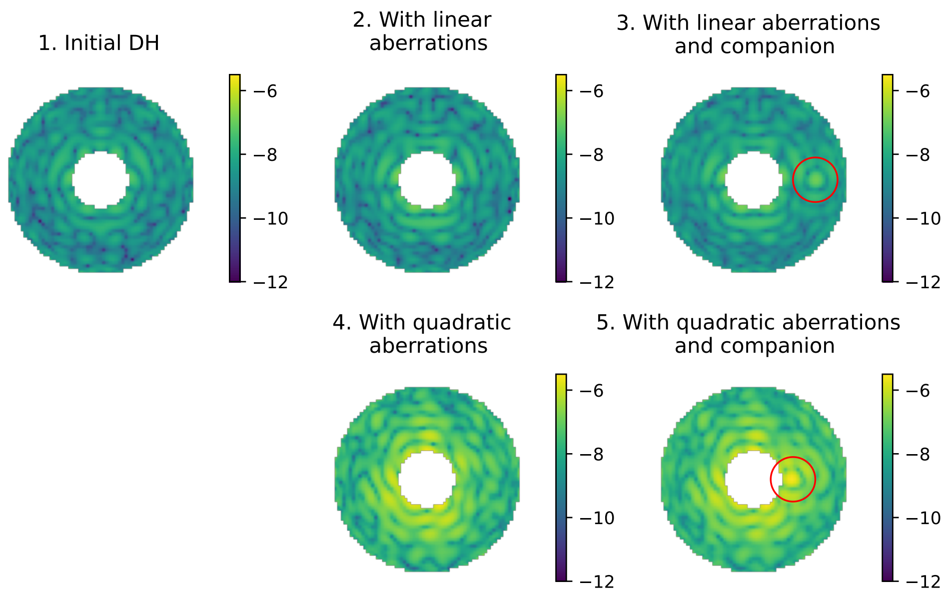

Figure 7.4: 1) Initial dark hole intensity achieved using electric field conjugation with the HLC. 2) A single snapshot with linear-regime wavefront aberrations. 3) The same snapshot with an injected companion with a flux ratio of 2×10^{-7} at $6.5\,\lambda/D$ (indicated with red circle). 4) A single snapshot with quadratic-regime wavefront aberrations. 5) The same snapshot with an injected companion with a flux ratio of 5×10^{-6} at $4\,\lambda/D$ (indicated with red circle). All intensities are shown in $\log_{10}$ of raw contrast.

In both regimes, the response matrices largely reproduce the spatial structure of the intensity error from the optical model.

Processing Synthetic Data

The quantity $|\mathbf{C}\boldsymbol{E}_{s0}|^2$ is the initial dark hole intensity without any extra WFE applied (as determined from the data at the end of the dark-hole digging sequence on the reference star, for example). This nominal signal is first subtracted from each frame. Then, the pixels within the defined dark hole are gathered into the vector $\boldsymbol{\Delta I}_{\mathrm{DH}}$. The data is left-multiplied by the appropriate $\mathbf{K}$ matrix to obtain the observables $\boldsymbol{O} = \mathbf{K}\boldsymbol{\Delta I}_{\mathrm{DH}}$. The data is also left-multiplied by the random matrix of the same size as $\mathbf{K}$ to obtain data whose dimension has been reduced randomly. For each case, the average of the data over the twenty frames is used as the final measurement, while the standard deviation of the frames is used as the measurement uncertainty. Note that the process outlined does not rely on reference stars or dithering by deformable mirrors, and can be used even on observations for which reference observations or

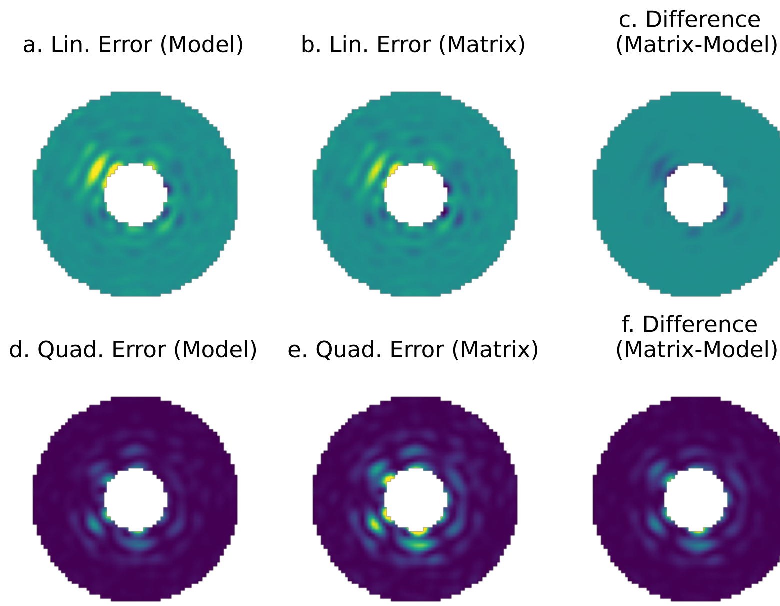

Figure 7.5: a. Example linear-regime intensity error from the optical model. b. Corresponding linear-regime intensity error reconstructed by response matrix $\mathbf{A_l}$, plotted on the same scale as (a). c. The difference between the response matrix prediction and the optical model prediction, plotted on the same scale as (a) and (b). d. Example quadratic-regime intensity error from the optical model. e. Corresponding quadratic-regime intensity error reconstructed by the response matrix $\mathbf{A_q}$, plotted on the same scale as (d). f. The difference between the response matrix prediction and the optical model prediction, plotted on the same scale as (d) and (e). Slight differences arise because the model includes both the linear and quadratic error terms while the matrix predictions only include one or the other, i.e. the linear matrix prediction neglects the contribution of the quadratic term and the quadratic matrix prediction neglects the contribution of the linear term (as well as the influence of any Zernikes past the first 100). While the linear matrix prediction is biased low near the peaks and the quadratic matrix prediction biased high overall, our method depends only on how well the *spatial structure* of the errors are reproduced. A relevant metric for characterizing the spatial overlap is the normalized inner product between the optical model prediction and the transfer matrix prediction, where a value of 1 indicates perfect spatial overlap and a value of 0 indicates perfect spatial orthogonality. In this case, the normalized inner product is 0.936 for the linear regime example and 0.985 for the quadratic regime example, sufficient for providing a quantifiable improvement in detection sensitivity.

wavefront diversity is unavailable.

Flux Ratio Detection Limits

Detection tests are applied to these measured intensities and observables in order to characterize the detectability of a companion with these measurements. Detection limits are determined using the Monte Carlo method. One thousand random datasets are generated for each noise model with a given flux ratio. Each dataset is processed as raw intensity data, and with each projection matrix with a different cutoff mode, and the $\Delta\chi_r^2$ values are calculated for each case. Figure 7.6 shows example histograms of the resulting $\Delta\chi_r^2$ values for a $c = 5.4 \times 10^{-7}$ at $4.0\,\lambda/D$ companion with the quadratic noise model, as well as the corresponding ROC curves, for the projection matrix with cutoff mode $N_m = 70$ (which, as shown in Figure 7.7, is the optimal cutoff at this spatial separation). The ROC curve shows that while using the robust observables results in a FPR = 0.01 and TPR = 0.9 detection of the injected companion, both the raw intensity and the randomly dimensionally reduced data remain very far from detectability.

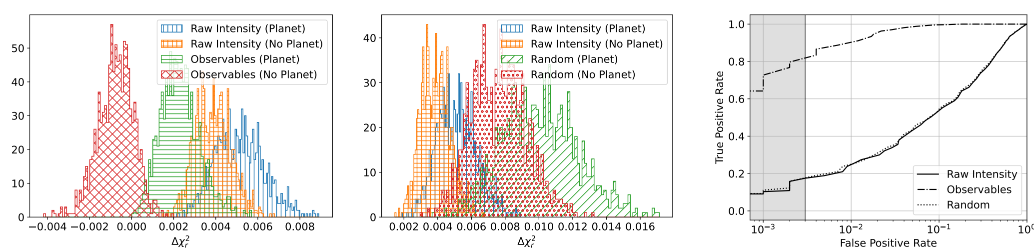

Figure 7.6: Detection test results for the quadratic regime noise model. The companion planet considered has flux ratio of 5.4×10^{-7} and is located at $4.0\,\lambda/D$. Left: Histograms from using raw intensities compared to those from using quadratic robust observables with the optimal cutoff of $N_m = 70$. The histograms using raw intensity overlap significantly, making it difficult to distinguish between a model with a planet and a model without one, while the histograms using the robust observables are further separated and more distinguishable. Middle: Histograms for using raw intensities and a random projection matrix of the same size as the instrumentally-motivated projection. Both sets of histograms overlap significantly, and the random projection does not improve the distinguishability of the two models. Right: ROC curves corresponding to the histograms. Grey area indicates false positive rates which are not well sampled as they involve less than 3 datasets with false detections. The ROC curve shows that while using the robust observables results in a FPR = 0.01 and TPR = 0.9 detection of the injected planet, both the raw intensity and the randomly dimensionally reduced data remain very far from detectability.

This process is repeated for a range of flux ratios (to a precision of two significant figures). The resulting FPR = 0.01 and TPR = 0.9 detection limits for both regimes, as a function of cutoff mode N_m, are shown in Figure 7.7. Note that these flux ratio detection limits are not based on any statistical assumptions or extrapolations, but rather real FPRs and TPRs calculated by analyzing one thousand synthetically generated datasets, with injected companions of the given flux ratios and separations. The results show that with the linear-regime noise model, the robust observables increases the detectability of a companion at $6.5\,\lambda/D$ by 28%. With the quadratic-regime noise model, using robust observables increases the detectability of a companion at $4.0\,\lambda/D$ by over a factor of two, and the improvement is not particularly sensitive to N_m beyond the first few modes. For the linear regime, this approach can also easily be extended to companions throughout the entire dark hole, though significant computation would be required to optimize N_m at all separations. For the quadratic regime, our model is only valid within $\sim 5\,\lambda/D$, though Appendix 7.6 discusses a method that can be used to extend the spatial coverage without incurring impractical computational costs.

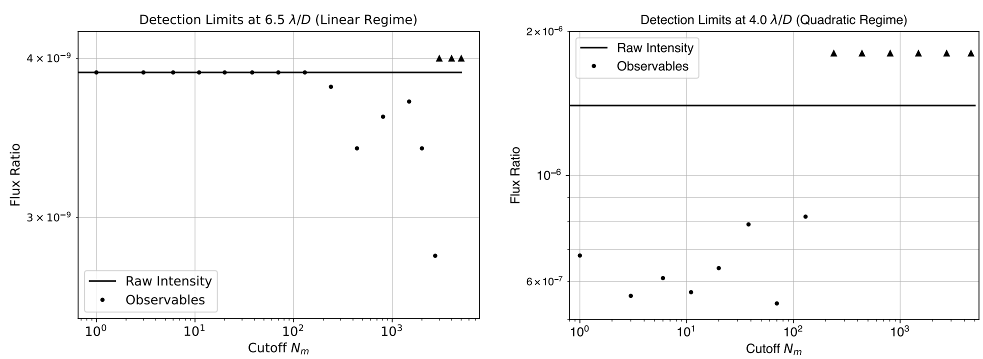

Figure 7.7: Flux ratio detection limits (FPR = 0.01, TPR = 0.90) for a binary companion (to two significant figures) as a function of cutoff mode. Upward triangles indicate that a projection matrix with the specified cutoff mode performs worse than using the raw intensity, which occurs when the modes the majority of the planet signal overlaps with have also been projected out. Left: Linear regime with a companion at 6.5 λ/D. The optimal cutoff mode is 2,727, which results in a detection limit of 2.8×10^{-9}. Unshowable in log-log scale is the detection limit with $N_m = 0$, which, with observables, is 3.9×10^{-9}. This is, as expected from the fact that no error modes are removed, equal to the raw intensity detection limit. Right: Left: Quadratic regime with a companion at $4.0\,\lambda/D$. The optimal cutoff mode is 70, which results in a detection limit of 5.4×10^{-7}. Unshowable in log-log scale is the detection limit with $N_m = 0$, which, with observables, is 1.4×10^{-6}. This is, as expected from the fact that no error modes are removed, equal to the raw intensity detection limit.

7.6 Quadratic Model Approximation and Extension

As explained in Section 7.3, the calculation of the projection matrix involves a singular value decomposition (SVD) of the response matrix, but a quadratic response matrix that includes all 528 Zernikes needed to span the dark hole would have a size of $5,476 \times 139,656$. Since calculating the SVD of a matrix of this size is too computationally burdensome, we explore an approximation of the quadratic response that models only the impact of norm-squared of each input basis vector while neglecting the effects of the pairwise combinations. Namely, we use an approximate response matrix $\mathbf{A'_q}$ with elements:

$$\mathrm{A'_{q}}_{kj} = \mathrm{C}^*_{kj}\mathrm{C}_{kj}. \tag{7.18}$$

The index j labels the input basis vector and the index k labels the detector pixel. The size of $\mathbf{A'_q}$ scales linearly with the number of Zernike models, and in our case would be of size $5,476 \times 528$, which is easily decomposable.

Note that $\mathbf{A'_q}$ cannot be used to accurately reproduce quadratic-regime intensity error. However, $\mathbf{A'_q}$ is nevertheless useful for identifying a subspace robust to quadratic regime wavefront errors, leading to increased signal-to-noise. We can observe this by comparing the detection test results with and without using the approximation for a model with 100 Zernikes. We calculate the approximation $\mathbf{A'_q}$ using Equation 7.18, and use the original $\mathbf{A_q}$ from Section 7.5. Detection tests on quadratic-regime synthetic data, similar to the one from Section 7.5 are performed, using projection matrices derived from both $\mathbf{A_q}$ and $\mathbf{A'_q}$. The resulting flux ratio detection limits as a function of cutoff mode are shown in Figure 7.8.

The full matrix achieves the best results with a cutoff mode of 70, leading to a detection limit of 5.4×10^{-7} while the approximate matrix achieves the best results with a cutoff mode of 2,727, also leading to a detection limit of 5.4×10^{-7}. These results show that the approximation performs as well as the full model.

To understand why this is the case, we analyze the subspaces spanned by the identified optimal projection matrices. We define $\mathbf{P}$ as the projection onto the dominant modes of $\mathbf{A_q}$, $\mathbf{P'}$ as the projection onto the dominant modes of $\mathbf{A'_q}$, and $\mathbf{P_r}$ as a random projection matrix the same shape as $\mathbf{P'}$. We also define $\mathbf{K}$ as the projection onto the remaining modes (the robust subspace) of the full model, $\mathbf{K'}$ as the projection onto the robust subspace of the approximate model, and $\mathbf{K_r}$ as a random projection matrix the same shape as $\mathbf{K'}$. We then calculate the subspace

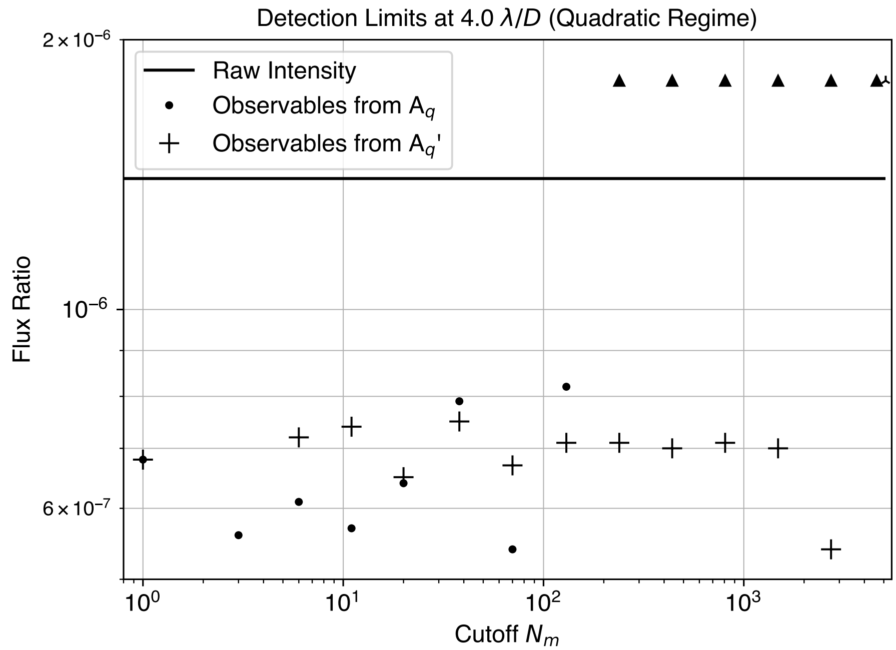

Figure 7.8: Quadratic regime flux ratio detection limits (FPR = 0.01, TPR = 0.90), to two significant figures, as a function of cutoff mode, for a companion at 4.0 λ/D. Only the first 100 Zernikes are used in the model used to calculated the full and approximate quadratic transfer matrices, but WFE up to 538 Zernikes are included in the synthetic data. Upward triangles or spikes indicate that a projection matrix with the specified cutoff mode performs worse than using the raw intensity, which occurs when the modes the majority of the planet signal overlaps with have also been projected out. The full matrix achieves the best results with a cutoff mode of 70, leading to a detection limit of 5.4×10^{-7} while the approximate matrix achieves the best results with a cutoff mode of 2,727, also leading to a detection limit of 5.4×10^{-7}. Unshowable in log-log scale is the detection limit with $N_m = 0$, which, with observables, is 1.4×10^{-6}. This is, as expected from the fact that no error modes are removed, the same as the raw intensity detection limit of 1.4×10^{-6}. These results indicate that the approximation performs as well as the full model.

angles (Jordan, 1875) between each of these projection matrices and **P** using the function scipy.linalg.subspace_angles. These subspace angles provide an indication of how much the subspace spanned by each of these projection matrices overlaps with the subspace spanned by the dominant modes identified by the full model. The results are shown in Figure 7.9.

The number of principle angles with value 0 is the dimension of overlap between the subspaces. As expected, the subspace angles between **P** and itself are all 0, meaning it overlaps completely with itself. Also as expected, the angles between **K** and **P**

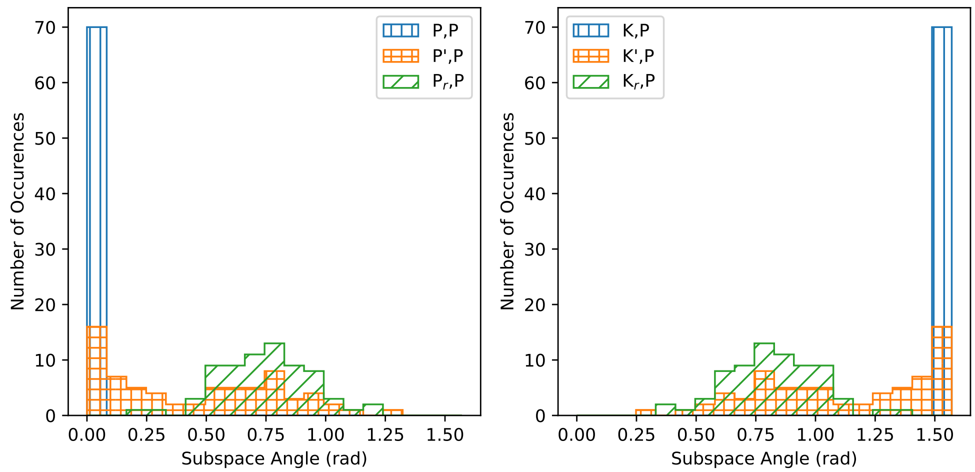

Figure 7.9: The subspace angles between various projection matrices (onto dominant modes on the right, onto a robust subspace on the left) and $\mathbf{P}$, the projection onto the dominant error modes determined from the full quadratic model. The number of principle angles with value 0 is the dimension of overlap between the subspaces. Angles with value $\pi/2$ indicate overlap with the subspace orthogonal to $\mathbf{P}$. $\mathbf{P}'$ (the space of dominant modes derived from the approximate model) overlaps with $\mathbf{P}$ (the space of dominant modes derived from the full model) significantly more than random. Crucially, $\mathbf{K}'$ (the robust subspace from the approximate model) overlaps with the subspace orthogonal to $\mathbf{P}$'s significantly more than random, which is why data projected onto this subspace is still robust to wavefront error.

are all $\pi/2$, as $\mathbf{K}$ is orthogonal to $\mathbf{P}$. Both of the random matrices have a random distribution of angles with $\mathbf{P}$ centered around $\pi/4$. Meanwhile, $\mathbf{P}'$ (the space of dominant modes derived from the approximate model) overlaps with $\mathbf{P}$ (the space of dominant modes derived from the full model) significantly more than random. Crucially, $\mathbf{K}'$ (the robust subspace from the approximate model) overlaps with $\mathbf{P}$ significantly less than random, and with the subspace orthogonal to $\mathbf{P}$'s significantly more than random, which is why data projected onto this subspace is still robust to wavefront error. This result shows why the approximate model, despite poorly predicting the detector intensity response, is nevertheless useful for identifying a subspace that overlaps significantly with the robust subspace of the full model.

We can thus use this approximation with all 528 Zernikes in our model to analyze spatial separations beyond the $\sim 5\ \lambda/D$ spanned by the first 100 Zernikes. To demonstrate this, we build $\mathbf{A}'_{\mathbf{q}_{528}}$ according to Equation 7.18, and perform detection tests at a separation of $6.5\ \lambda/D$. The results are shown in Figure 7.10.

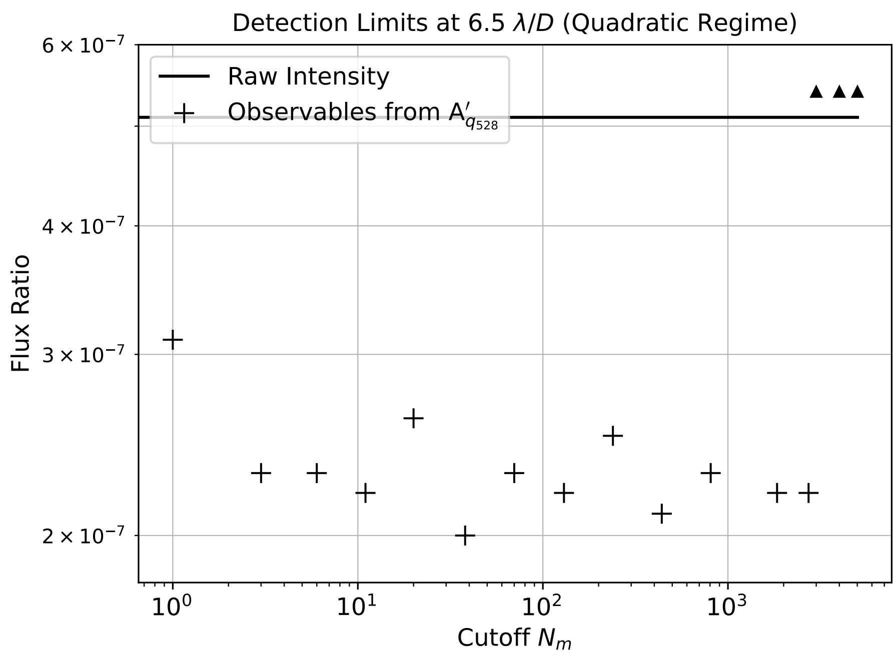

Figure 7.10: Quadratic regime flux ratio detection limits (FPR = 0.01, TPR = 0.90), to two significant figures, as a function of cutoff mode, for a companion at 6.5 λ/D. Upward triangles indicate that a projection matrix with the specified cutoff mode performs worse than using the raw intensity, which occurs when the modes the majority of the planet signal overlaps with have also been projected out. The optimal cutoff mode is 38, which results in a detection limit of 2.0×10^{-7}. Unshowable in log-log scale is the detection limit with $N_m = 0$, which, with observables, is 4.9×10^{-7}. This is, as expected from the fact that no error modes are removed, close to the raw intensity detection limit of 5.1×10^{-7}.

Our tests show that the approximation $\mathbf{A'_{q528}}$ can successfully increase signal-to-noise at spatial separations beyond the original regime of validity of $\mathbf{A_q}$. Thus, even though the input dimension of the quadratic model scales cumbersomely with the number of basis vectors, an approximation considering only norm-squared terms can still be used to find observables that are robust to quadratic wavefront error, and thus provide detection gains at farther spatial separations of interest.

7.7 Discussion

Temporally Correlated WFE and Compatibility with Other Post-processing Techniques

This work aims to characterize the effect of using robust observables in isolation. Thus, only noise models in which the WFE is uncorrelated in time are examined, since additional post-processing techniques are typically used to handle time-correlated data. Robust observables are compatible with these other post-processing techniques, and can serve as an instrument-motivated prior in the overall post-processing strategy. For example, random errors can first be reduced by projecting the data into a subspace that is robust to wavefront error. Then, reference observations along with PCA-based methods such as KLIP (Soummer, Laurent Pueyo, and Larkin, 2012b) can be used to calibrate static and quasi-static errors and de-correlate the frames in time. This is similar to the calibration approach used in non-redundant aperture masking (NRM) interferometry or kernel-phase interferometry, in which data is projected onto closure-phases or kernel-phases respectively, which are then calibrated based on reference observations (Martinache, 2010; Ireland, 2013; Pope et al., 2021). A more sophisticated approach would be to formulate post-processing as a statistical inference problem, where a least-squares fit with the reference frames makes one up term in the cost function, and a prior over the instrumental modes (e.g. weighted by the singular value spectrum) makes up another term.

Ygouf et al. (2016) shows that for the time-varying wavefront error expected on the Roman Space Telescope HLC, classical PSF subtraction with a reference observation increases the contrast gain by a factor of a few to about ten, depending on the scenario. Future work includes investigating how much overall post-processing gain can be achieved when robust observables and calibration strategies are combined, and which hybrid strategies maximize the sensitivity that can be obtained with all available information.

PSD Engineering

The robust observables derived in this work are agnostic to the actual temporal or spatial PSD of the static and dynamical wavefront errors, and are intended to be applied when these PSDs are not well-known or imperfectly characterized. As of today, this is the case for all ground-based instruments (as predictions of the influence of the atmosphere are quite imperfect), and space-based missions (as HST and JWST observatory level key metrics for requirements are expressed in terms of encircled energy, not contrast). However, it has been proposed that for future space telescope coronagraphs, the telescope WFE PSD must comply with stringent requirements in order to facilitate exoplanet detection (Nemati et al., 2020).

For instance, the PASTIS approach (Leboulleux et al., 2018; Laginja et al., 2019) considers the effects of the quadratic response on the *average* intensity contrast over the entire dark hole (or region of interest), calculating which modes the coronagraph is most sensitive to in order to determine stability tolerances for the segments accordingly. Calculating robust observables for post-processing is akin to doing PASTIS backwards, where the modes the coronagraph is most sensitive to are calculated in order to project them out of the data. For such telescopes, that have PSDs engineered based on the instrument response, the additional gain from using robust observables will depend on how well the error modes are suppressed in hardware, as well as the timescales at which power in those modes leaks through. To some extent, robust observables will remain applicable to such future telescopes and instruments in the spatial and temporal sub-spaces in which they do not meet their requirements.

Model Accuracy

In this analysis, the model used to generate the instrument response matrices is exactly the same model that is used to generate the synthetic data. In a real observations, the instrument model will not exactly match the behavior of the actual instrument, and one future avenue to explore is how well a model must match the instrument in order for robust observables to work on real data. This technique's robustness can be investigated by first calculating the response matrices using one model, then changing the parameters of the model (e.g. the coronagraphic mask size and displacement, the DM alignment, the detector pixel scale) before generating synthetic data, and examining how well the robust observables work in the presence of model mismatch.

For instruments equipped with wavefront modulating devices such as deformable mirrors, however, the instrument response matrix may also be calculated experimentally. If a perturbation within the linear regime is applied, the difference in measured intensity can be directly registered into the appropriate column of the linear response matrix. The technique for experimentally building the quadratic response matrices is equivalent to the approach used for PASTIS (Laginja et al., 2019), with the difference that the measurements are not averaged over a dark hole, but rather maintained for every pixel. Additionally, some wavefront and control schemes such as implicit electric-field conjugation (S. Y. Haffert, Males, et al., 2023b) already involve an empirical measurement of the instrument response, which can be used to derive linear-regime robust observables without having to set aside additional calibration time. Experimentally building instrument response matrices circumvents the need to have a well-matched numerical model, and allows for the response matrices to capture effects in the real instrument.

7.8 Conclusions

A coronagraph model with linear and quadratic contributions of wavefront error to detector plane intensity is developed, and when either term is dominant, the coronagraph response can be approximated by a transfer matrix. A useful projection can be found from this transfer matrix that removes the dominant error modes, resulting in observables that are more robust to WFE in the regime of interest. These robust observables are extracted from synthetically generated data with the Hybrid Lyot Coronagraph of the Roman Space Telescope in both the linear and quadratic regimes. The performance of the robust observables is compared to that of the raw intensity data through calculations of their respective binary companion flux ratio detection limits. In these examples, using the robust observables significantly increases the sensitivity to the signal of a binary companion. A projection onto a robust subspace can in theory be combined with other families of post-processing algorithms. Hybrid post-processing approaches would incorporate information on the instrument response alongside the other available information (such as angular diversity, spectral diversity, reference observations, or WFC telemetry) to fully maximize the sensitivity to astrophysical signals in coronagraphic data; however, the approach outlined in this work can be applied to observational data and result in post-processing gains even if such additional information is unavailable.

Chapter 8

SUMMARY AND PERSPECTIVES

This thesis covers advances in instrumentation for directly detecting and characterizing exoplanets at close-in separations, within the working angles accessible to conventional coronagraphs. A major focus of this work is the Photonic Lantern Nuller, which uses a mode-demultiplexing waveguide called a mode-selective photonic lantern to destructively interfere on-axis starlight while allowing off-axis planet light through. This thesis covers the development of the Photonic Lantern Nuller from the original concept, through a comprehensive laboratory characterization, to an on-sky demonstration. It also presents other advances in high contrast instrumentation, demonstrated with other instruments but with potential applications to the PLN.

8.1 Summary

The first few chapters in this thesis cover the development of the Photonic Lantern Nuller instrument. Chapter 2 reproduces the paper in which the concept, theory, and simulated behavior of the PLN with various parameters under different conditions. It shows that the PLN augments the overall planet throughput relative to the Vortex Fiber Nuller, a predecessor design, reaching 60% throughput compared to the VFN's 10-18%. It presents simulations that predict the anticipated null degradation due to sources of wavefront error, such as Zernike aberrations, fast tip-tilt jitter, or atmospheric residuals reconstructed from wavefront sensor data from the Keck II telescope. The PLN overall has similar wavefront error rejection properties to the VFN as a result of its symmetries, and also adds the ability to partially localize the location of the planet.

Chapter 3 details the laboratory characterization of a mode-selective photonic lantern. The throughput of the lantern ports were measured, and the mode-profiles characterized using off-axis holography, which interferes a reference beam with each mode of the lantern to reconstruct its electric-field profile. The lantern was then integrated with the Polychromatic Reflective Testbed and demonstrated as a nuller. The experiments measured null-depths (stellar coupling divided by peak planet coupling) of approximately 10^{-2}, as well as spatial coupling profiles that match the predictions from simulation. Measurements were made using both a monochromatic light source as well as a broadband light source, as well as with two orthogonal polarizations of light, to characterize the impact of bandwidth and polarization on the achieved null-depth, showing them both to be minor at this contrast level.

Chapter 4 details the use of the implicit Electric Field Conjugation algorithm to deepen the null-depths of the PLN to the 10^{-4} level, although only achieving this with three out of four ports simultaneously. I use simulations to explain the origin of this behavior, identifying an electric-field sensing problem unique to mode-sorting nullers that forms an intriguing topic for future work.

Chapter 5 presents an on-sky demonstration of the PLN at the Subaru Telescope with the Subaru Coronagraphic Extreme Adaptive Optics instrument. Despite significant challenges in instrument alignment, spectrally-dipsersed on-sky null-depths were successfully measured for the LP 11a and LP 11b ports to be about 10^{-1}, with statistical distributions that suggest the limitation is mostly due to fast-changing AO residuals. Further analysis of the on-sky data suggests that a 1:10 binary companion should be detectable with a few minutes of data each on the target system and on a point-source reference star.

While the bulk of my thesis focuses on developing the PLN, I was also an active part of the instrument team for the Keck Planet Imager and Characterizer (KPIC). KPIC is a useful platform and benchmark for a potential PLN-fed high-resolution spectrograph, and I leveraged my time as part of the instrument team to demonstrate wavefront sensing and control techniques, informing their potential on-sky feasibility with the PLN. Chapter 6 thus presents work on speckle nulling with KPIC, where starlight is actively destructively interfered with itself in order to reduce contamination in the science data, improving the suppression by a factor of 2.6-2.8 on-sky. Speckle nulling is algorithmically similar to iEFC and has similar requirements, so this work also serves as an estimate of how well iEFC with a PLN might work, if implemented on-sky with similar hardware (spectrograph resolving power, detector characteristics, etc.).

Chapter 7 reproduces my paper on using a model of a coronagraphic instrument to identify a subspace in the data that is less sensitive to wavefront aberrations. I show that projecting the data onto this robust subspace can improve detection sensitivity by a factor of 28% to a factor of 2, depending on the wavefront error regime. This work uses a simulation of the Roman Space Telescope Hybrid Lyot Coronagraph instrument. However, the concept of using instrumental responses to wavefront error

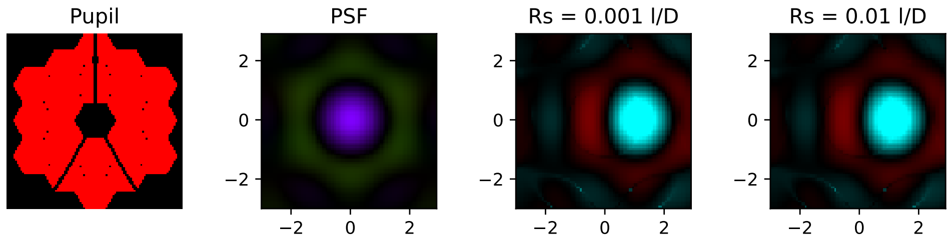

Figure 8.1: From left to right: 1) The pupil aperture for the James Webb Space Telescope. 2) The corresponding PSF in the focal plane. 3) A waveguide mode optimized for a planet at $X = 1.0\lambda/D$ and $R_s = 0.001\lambda/D$, achieving $\eta_s = 5.89 \times 10^{-7}$ and $\eta_p = 0.993$. 4) A waveguide mode optimized for a planet at $X = 1.0\lambda/D$ and $R_s = 0.01\lambda/D$, achieving $\eta_s = 4.40 \times 10^{-5}$ and $\eta_p = 0.993$.

in post-processing can be extended to instruments like the PLN, such as by weighing the signal in linear combinations ports by their relative sensitivity to aberrations.

8.2 Perspectives

Optimal Waveguide-based Nullers

Although the PLN is a compelling instrument that can access $1\lambda/D$ scales with high planet throughput, its design came from leveraging the symmetries of a mode-selective photonic lantern to cancel out starlight, without explicitly considering planet throughput in the design process. However, one can imagine a different waveguide that both nulls starlight and preserves more of the planet light, perhaps concentrating all of the planet light in one port to maximize sensitivity in the presence of photon noise. For simplicity, consider a waveguide that supports one mode, optimized for a planet at one spatial location. A mode that minimizes the relative integration time η_s/η_p^2 would depend on the pupil aperture function, the stellar radius R_s, and the planet location, and can be calculated using numerical optimizers.

Two example mode solutions for the James Webb Space Telescope pupil with a planet at $X = 1.0\lambda/D$ (for $R_s = 0.001\lambda/D$ and $R_s = 0.01\lambda/D$ respectively), calculated using the SLSQP algorithm from the scipy.optimize.minimize module (Virtanen et al., 2020) are shown in Fig. 8.1. This JWST pupil is also one of the primary pupil configurations being explored for the Habitable Worlds Observatory, which (as discussed in Section 8.2) would benefit from an infrared nulling instrument operating at the $\sim 1\lambda/D$ regime, in order to perform follow-up infrared spectroscopy of exoplanets discovered at optical wavelengths.

These solutions achieved $\eta_s = 5.89 \times 10^{-7}$ and $\eta_p = 0.993$ for $R_s = 0.001\lambda/D$ and $\eta_s = 4.40 \times 10^{-5}$ and $\eta_p = 0.993$ for $R_s = 0.01\lambda/D$. However, the solution returned by the optimizer depends on the initial guess, so more work would be needed to find globally-optimal solutions (i.e. by starting with an semi-analytic initialization using a framework similar to that in O. Guyon et al. (2006a)). Another complication is the manufacturability of the mode and the multiplexibility of the modes into a device like a photonic lantern, as numerically optimal solutions may require unrealistic manufacturing techniques or tolerances. Therefore, it will likely be worthwhile to include manufacturing constraints in the optimization process as well.

Simultaneous Nulling and Wavefront Sensing

This thesis focused on using a mode-selective photonic lantern to perform nulling, where four of the ports are nulled, and the other two ports are not-nulled. While the non-nulled ports can be useful for photometric and spectral calibration of the star or for monitoring overall atmospheric and AO conditions, they do not contribute meaningful information about the wavefront, as only asymmetric mixed modes are sensitive to both the shape and sign of wavefront aberrations (Lin, Fitzgerald, et al., 2022b).

A possible extension of the work in this thesis would be to use hybrid photonic lanterns to simultaneously perform nulling and wavefront sensing. For example, the first three ports (LP 01 and LP 11ab) could be made mode-selective, providing a throughput monitoring port along with two nulling ports, while the remaining ones are left mixed for wavefront sensing. The advantage of this configuration is that there is no non-common path aberration between the sensing ports and the nulling ports, whose relationship should remain very stable over time. The wavefront sensing ports could be used in closed-loop with a deformable mirror, and/or used in post-processing to calibrate the intensity in the nulled ports to remove the contribution known to come from wavefront error. This technique falls under the umbrella of Coherent Differential Imaging (CDI). In general, CDI methods seek to sense and remove the coherent portion of the light, which can only be due to the star. Designing instruments that enable and harness the potential of CDI is a particularly interesting avenue for future work.

Probing the Exoplanet Parameter Space with Nulling

The PLN and other close inner working angle nullers can serve different roles in the broader context of exoplanet science, depending on where and how they are used. The range of exoplanets that can be directly observed with a coronagraph or nuller design depends primarily on the observing wavelength λ, the size of the telescope aperture D (or the length of the baseline between apertures B), and the contrast achievable given the stability of the optical wavefront. Generally, high contrast is harder to achieve and maintain at closer separations to the star, so the smaller the planet's separation is in units of λ/D, the harder it will be to detect. Therefore, the exoplanets that an instrument can target will depend on the telescope it is on. As shown in Fig. 8.2, the different environments of ground- and space-based telescopes, as well as the different limitations they face, mean that they are ultimately better suited for probing different regions of the exoplanet parameter space.

On the 6 – 10m class telescopes on the ground, conventional coronagraphs are typically used to observe companions farther than $\sim$ 100 mas from their host star at near-infrared wavelengths, reaching flux ratio sensitivies down to 10^{-6} – 10^{-7}. Meanwhile, interferometer arrays such as VLTI/GRAVITY (S. Lacour et al., 2020) and CHARA (Brummelaar et al., 2005), as well as newer nulling instruments on 10m class telescopes (Echeverri, J. W. Xuan, et al., 2024; Norris, Cvetojevic, et al., 2020) — including the PLN — are just beginning to directly access the 1-5 AU parameter space for nearby stars, with significant efforts dedicated to improving performance and sensitivity in the near future.

Future generation instruments on the Extremely Large Telescopes (ELTs) will have very large diameters ($\sim$ 30m), but will ultimately remain limited in sensitivity due to the presence of the Earth's atmosphere. A compelling science case for coronagraphy on the ELTs is access to the rocky planets in the habitable zone of M dwarfs, which have less extreme flux ratios (on the order of 10^{-6}) relative to their solar-type counterparts, but are also closer in (National Research Council, 2021). Nulling instruments on the ELTs will be able to directly access the even closer-in population of Hot Jupiters (HJs), providing direct spectral measurements of known transiting HJs (as well as detecting similar objects that do not transit).

Space telescopes, on the other hand, are not limited by the Earth's atmosphere, but telescopes with larger aperture diameters are much harder to build and launch. Relative to the ELTs, space telescopes are therefore better suited for observing fainter but farther away exoplanets, such as Jovian-like planets (flux ratios of $\sim 10^{-8}$) with the Roman Space Telescope and eventually Earth-like planets (flux ratios of $\sim 10^{-10}$) with the Habitable Worlds Observatory (HWO). However, the requirements

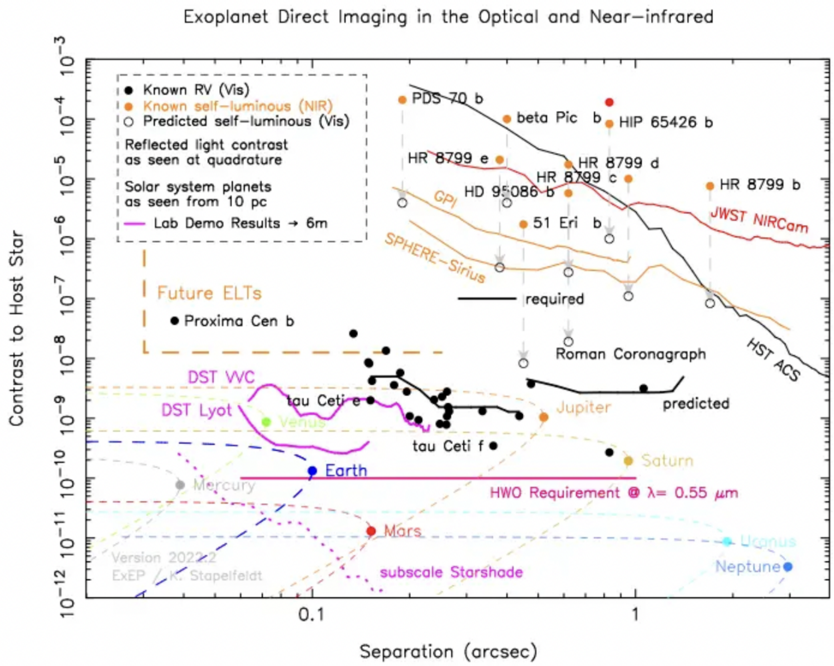

Figure 8.2: Flux ratio (contrast) versus apparent angular separation. The filled orange circles indicate direct imaging detections of young, self-luminous planets imaged in the near-infrared by ground-based telescopes. The orange curves show measured performance of ground-based coronagraphs. The GPI curve shows typical performance, while the SPHERE curve shows the best achieved performance to-date on Sirius. Achieved performance with Hubble Space Telescope (HST) and James Webb Space Telescope (JWST) are also shown. The predicted and required performance at 565 nm for (two configurations of) the Roman Coronagraph instrument is shown as solid black curves. State of the art lab coronagraph lab demonstrations in the Decadal Survey Testbed (DST) are shown as magenta curves. The notional performance goal for the Habitable Worlds Observatory is shown as a red horizontal line. For consistency, known self-luminous planets discovered in the near-infrared are shown with vertical arrows pointing to the predicted contrast ratios at visible wavelengths. Figure and caption adapted from *Exoplanet Program: Technology Overview* (2024), courtesy of Karl Stapelfeldt and NASA/JPL/Caltech.

for observing Earth analogues are different for instruments operating at optical wavelengths as opposed to those operating at near-infrared wavelengths. Plotting the distribution of habitable zone distances of a provisional HWO target list (Mamajek and Stapelfeldt, 2024) converted to λ/D units for $\lambda = 500$nm and $\lambda = 1500$nm (Fig. 8.3) shows that most of the habitable zones accessible by a coronagraph at optical wavelengths will no longer be accessible in the infrared. For these planets, a nulling mode operating at smaller separations ($\sim 1\lambda/D$) could perform

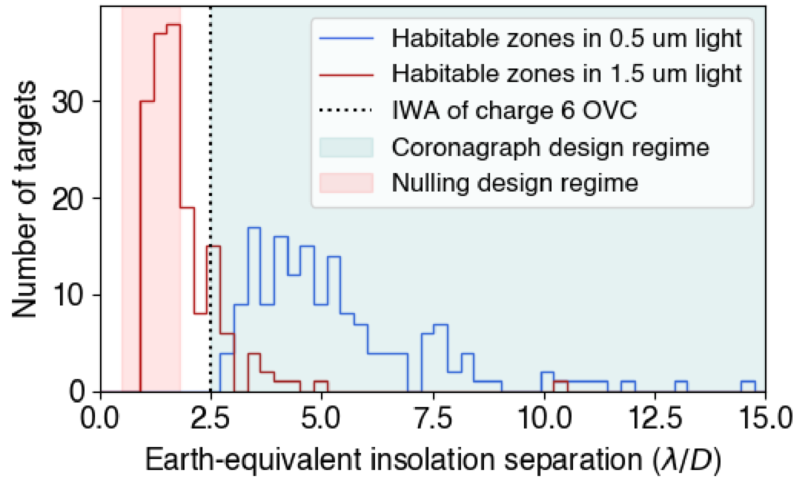

Figure 8.3: The distribution of Earth-equivalent insolation separations (a proxy for habitable zone location) of a tentative HWO target list in units of λ/D. Most habitable zones that are observable with a coronagraph at optical wavelengths ($\sim$ 0.5 um) are no longer accessible at infrared wavelengths ($\sim$ 1.5 um). Nulling interferometers designed for closer-in separations are needed to spectrally characterize these planets at infrared wavelengths.

spectroscopic follow-up in the infrared, where molecules like CO_2 and CH_4 have the strongest features. Their abundances, interpreted in conjunction with those of O_2 and H_2O, are needed to provide constraints on disequilibrium chemistry in the planet's atmosphere and the potential existence of life. It is unclear which nulling architecture or design will eventually best meet the demands of the HWO mission; however, the invention, characterization, and demonstration of the PLN presented in this thesis expands the space of possibilities, contributing valuable relevant knowledge in photonics-based instrument design, wavefront sensing and control with mode-sorters, calibration and operational procedures, and more.

BIBLIOGRAPHY

Agol, Eric and Daniel C. Fabrycky (2018). "Transit-Timing and Duration Variations for the Discovery and Characterization of Exoplanets". In: *Handbook of Exoplanets*. Ed. by Hans J. Deeg and Juan Antonio Belmonte, 7, p. 7. DOI: 10.1007/978-3-319-55333-7_7.

Akeson, R. L. et al. (Aug. 2013). "The NASA Exoplanet Archive: Data and Tools for Exoplanet Research". In: *Publications of the Astronomical Society of the Pacific* 125.930, p. 989. DOI: 10.1086/672273. arXiv: 1307.2944 [astro-ph.IM].

Ashcraft, Jaren N. et al. (Mar. 2025). "Polarization aberrations in next-generation Giant Segmented Mirror Telescopes (GSMTs): II. Influence of segment-to-segment coating variations on high-contrast imaging and polarimetry". In: *Astronomy and Astrophysics* 695, A28, A28. DOI: 10.1051/0004-6361/202452671. arXiv: 2501.03897 [astro-ph.IM].

Batalha, Natasha E. et al. (June 2019). "Exoplanet Reflected-light Spectroscopy with PICASO". In: *Astrophysical Journal* 878.1, 70, p. 70. DOI: 10.3847/1538-4357/ab1b51. arXiv: 1904.09355 [astro-ph.EP].

Baudoz, P. et al. (Jan. 2006a). "The Self-Coherent Camera: a new tool for planet detection". In: *IAU Colloq. 200: Direct Imaging of Exoplanets: Science & Techniques*. Ed. by C. Aime and F. Vakili, pp. 553–558. DOI: 10.1017/S174392130600994X.

– (Jan. 2006b). "The Self-Coherent Camera: a new tool for planet detection". In: *IAU Colloq. 200: Direct Imaging of Exoplanets: Science & Techniques*. Ed. by C. Aime and F. Vakili, pp. 553–558. DOI: 10.1017/S174392130600994X.

Bloemhof, E. E. (Dec. 2002). "Suppression of Speckle Noise by Speckle Pinning in Adaptive Optics". In: *The Astrophysical Journal* 582.1, pp. L59–L62. DOI: 10.1086/346100. URL: https://doi.org/10.1086%2F346100.

Bottom, Michael et al. (2016). "Speckle nulling wavefront control for Palomar and Keck". In: *Adaptive Optics Systems V*. Ed. by Enrico Marchetti, Laird M. Close, and Jean-Pierre Véran. Vol. 9909. International Society for Optics and Photonics. SPIE, p. 990955. DOI: 10.1117/12.2233025. URL: https://doi.org/10.1117/12.2233025.

Bracewell, R. N. (Aug. 1978). "Detecting nonsolar planets by spinning infrared interferometer". In: *Nature* 274.5673, pp. 780–781. DOI: 10.1038/274780a0.

Brummelaar, Theo A ten et al. (2005). "First results from the CHARA array. II. A description of the instrument". In: *The Astrophysical Journal* 628.1, p. 453.

Cantalloube, F. et al. (Aug. 2022). "Exoplanet imaging data challenge, phase II: characterization of exoplanet signals in high-contrast images". In: *Adaptive Optics Systems VIII*. Ed. by Laura Schreiber, Dirk Schmidt, and Elise Vernet.

Vol. 12185. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1218505, p. 1218505. DOI: 10.1117/12.2627968. arXiv: 2209.08120 [astro-ph.IM].

Ceau, Alban et al. (Aug. 2019). "Kernel-phase Detection Limits : Hypothesis Testing and the Example of JWST NIRISS Full Pupil Images". In: *arXiv e-prints*, arXiv:1908.03130, arXiv:1908.03130. arXiv: 1908.03130 [astro-ph.IM].

Corrigan, Mark K. et al. (2018). "Demonstration of a photonic lantern low order wavefront sensor using an adaptive optics testbed". In: *Adaptive Optics Systems VI*. Ed. by Laird M. Close, Laura Schreiber, and Dirk Schmidt. Vol. 10703. International Society for Optics and Photonics. SPIE, pp. 1313–1320. DOI: 10.1117/12.2311336. URL: https://doi.org/10.1117/12.2311336.

Cuche, Etienne, Pierre Marquet, and Christian Depeursinge (Aug. 2000). "Spatial filtering for zero-order and twin-image elimination in digital off-axis holography". In: *Appl. Opt.* 39.23, pp. 4070–4075. DOI: 10.1364/AO.39.004070. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-39-23-4070.

Currie, T. et al. (July 2023). "Direct Imaging and Spectroscopy of Extrasolar Planets". In: *Protostars and Planets VII*. Ed. by S. Inutsuka et al. Vol. 534. Astronomical Society of the Pacific Conference Series, p. 799. DOI: 10.48550/arXiv.2205.05696. arXiv: 2205.05696 [astro-ph.EP].

Cutri, R. M. et al. (June 2003). "VizieR Online Data Catalog: 2MASS All-Sky Catalog of Point Sources (Cutri+ 2003)". In: *VizieR Online Data Catalog*, II/246, pp. II/246.

Delorme, Jacques-Robert et al. (July 2021a). "Keck Planet Imager and Characterizer: a dedicated single-mode fiber injection unit for high-resolution exoplanet spectroscopy". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 7, 035006, p. 035006. DOI: 10.1117/1.JATIS.7.3.035006. arXiv: 2107.12556 [astro-ph.IM].

– (July 2021b). "Keck Planet Imager and Characterizer: a dedicated single-mode fiber injection unit for high-resolution exoplanet spectroscopy". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 7, 035006, p. 035006. DOI: 10.1117/1.JATIS.7.3.035006. arXiv: 2107.12556 [astro-ph.IM].

Echeverri, Daniel, Garreth Ruane, Benjamin Calvin, et al. (Dec. 2020). "Detecting and characterizing close-in exoplanets with vortex fiber nulling". In: *Optical and Infrared Interferometry and Imaging VII*. Ed. by Peter G. Tuthill, Antoine Mérand, and Stephanie Sallum. Vol. 11446. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1144619, p. 1144619. DOI: 10.1117/12.2563142. arXiv: 2012.04239 [astro-ph.IM].

Echeverri, Daniel, Garreth Ruane, Nemanja Jovanovic, Thomas Hayama, et al. (2019). "The vortex fiber nulling mode of the Keck Planet Imager and Characterizer (KPIC)". In: *Techniques and Instrumentation for Detection of Exoplanets IX*. Ed. by Stuart B. Shaklan. Vol. 11117. International Society for Optics and

Photonics. SPIE, pp. 264–274. DOI: 10.1117/12.2528529. URL: https://doi.org/10.1117/12.2528529.

Echeverri, Daniel, Garreth Ruane, Nemanja Jovanovic, Dimitri Mawet, et al. (May 2019). "Vortex fiber nulling for exoplanet observations I Experimental demonstration in monochromatic light". In: *Optics Letters* 44.9, p. 2204. DOI: 10.1364/OL.44.002204. arXiv: 1811.02083 [astro-ph.IM].

Echeverri, Daniel, Jerry Xuan, et al. (2023). "Vortex fiber nulling for exoplanet observations: implementation and first light". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 9.3, p. 035002. DOI: 10.1117/1.JATIS.9.3.035002. URL: https://doi.org/10.1117/1.JATIS.9.3.035002.

Echeverri, Daniel, Jerry W. Xuan, et al. (Mar. 2024). "Vortex Fiber Nulling for Exoplanet Observations: First Direct Detection of M Dwarf Companions around HIP 21543, HIP 94666, and HIP 50319". In: *arXiv e-prints*, arXiv:2403.17295, arXiv:2403.17295. DOI: 10.48550/arXiv.2403.17295. arXiv: 2403.17295 [astro-ph.EP].

Exoplanet Program: Technology Overview (2024). https://exoplanets.nasa.gov/exep/technology/technology-overview/. [Accessed 27-03-2025].

Flasseur, Olivier et al. (Apr. 2018). "Exoplanet detection in angular differential imaging by statistical learning of the nonstationary patch covariances. The PACO algorithm". In: *Astronomy and Astrophysics* 618. DOI: 10.1051/0004-6361/201832745.

Gaia Collaboration et al. (Nov. 2016). "The Gaia mission". In: *Astronomy and Astrophysics* 595, A1, A1. DOI: 10.1051/0004-6361/201629272. arXiv: 1609.04153 [astro-ph.IM].

Gatkine, Pradip, Sylvain Veilleux, and Mario Dagenais (2019). "Astrophotonic spectrographs". In: *Applied Sciences* 9.2, p. 290.

Gaudi, B. Scott (Sept. 2012). "Microlensing Surveys for Exoplanets". In: *Annual Review of Astronomy and Astrophysics* 50, pp. 411–453. DOI: 10.1146/annurev-astro-081811-125518.

Gay, Jean, Yves Rabbia, and Pierre Baudoz (1997). "Le coronographe interférentiel achromatique". In: *Comptes Rendus de l'Académie des Sciences - Series IIB - Mechanics-Physics-Chemistry-Astronomy* 325.1, pp. 51–56. ISSN: 1251-8069. DOI: https://doi.org/10.1016/S1251-8069(97)83264-2. URL: https://www.sciencedirect.com/science/article/pii/S1251806997832642.

Give'on, Amir et al. (2007). "Broadband wavefront correction algorithm for high-contrast imaging systems". In: *Astronomical Adaptive Optics Systems and Applications III*. Ed. by Robert K. Tyson and Michael Lloyd-Hart. Vol. 6691. International Society for Optics and Photonics. SPIE, 66910A. DOI: 10.1117/12.733122. URL: https://doi.org/10.1117/12.733122.

Groff, Tyler D. et al. (2015). “Methods and limitations of focal plane sensing, estimation, and control in high-contrast imaging”. In: *Journal of Astronomical Telescopes, Instruments, and Systems* 2.1, pp. 1–15. DOI: 10.1117/1.JATIS.2.1.011009. URL: https://doi.org/10.1117/1.JATIS.2.1.011009.

Guyon, O. et al. (Nov. 2006a). “Theoretical Limits on Extrasolar Terrestrial Planet Detection with Coronagraphs”. In: *The Astrophysical Journal Supplement Series* 167.1, pp. 81–99. DOI: 10.1086/507630. URL: https://doi.org/10.1086%5C%2F507630.

– (Nov. 2006b). “Theoretical Limits on Extrasolar Terrestrial Planet Detection with Coronagraphs”. In: *The Astrophysical Journal Supplement Series* 167.1, pp. 81–99. ISSN: 1538-4365. DOI: 10.1086/507630. URL: http://dx.doi.org/10.1086/507630.

Guyon, Olivier et al. (Oct. 1999). “The Nulling Stellar Coronagraph: Laboratory Tests and Performance Evaluation”. In: *Publications of the Astronomical Society of the Pacific* 111.764, p. 1321. DOI: 10.1086/316445. URL: https://dx.doi.org/10.1086/316445.

Haffert, S. Y., J. R. Males, et al. (May 2023a). “Implicit electric field conjugation: Data-driven focal plane control”. In: *Astronomy & Astrophysics* 673, A28, A28. DOI: 10.1051/0004-6361/202244960. arXiv: 2303.13719 [astro-ph.IM].

– (May 2023b). “Implicit electric field conjugation: Data-driven focal plane control”. In: *Astronomy and Astrophysics* 673, A28, A28. DOI: 10.1051/0004-6361/202244960. arXiv: 2303.13719 [astro-ph.IM].

Haffert, S. Y., E. H. Por, et al. (Mar. 2020). “The Single-mode Complex Amplitude Refinement (SCAR) coronagraph. II. Lab verification, and toward the characterization of Proxima b”. In: *Astronomy and Astrophysics* 635, A56, A56. DOI: 10.1051/0004-6361/201731615.

Haffert, Sebastiaan Y. et al. (Sept. 2021). “The Visible Integral-field Spectrograph eXtreme (VIS-X): a high-resolution spectrograph for MagAO-X”. In: *Techniques and Instrumentation for Detection of Exoplanets X*. Ed. by Stuart B. Shaklan and Garreth J. Ruane. SPIE. DOI: 10.1117/12.2594875. URL: https://doi.org/10.1117%5C%2F12.2594875.

Haguenauer, P. and E. Serabyn (2006). “Deep nulling of laser light with a single-mode-fiber beam combiner”. In: *Appl. Opt.* 45.12, pp. 2749–2754. DOI: 10.1364/AO.45.002749.

Hanot, C. et al. (Mar. 2011). “Improving Interferometric Null Depth Measurements using Statistical Distributions: Theory and First Results with the Palomar Fiber Nuller”. In: *Astrophysical Journal* 729.2, 110, p. 110. DOI: 10.1088/0004-637X/729/2/110. arXiv: 1103.4719 [astro-ph.IM].

Hinkley, S. et al. (Mar. 2023). "Direct discovery of the inner exoplanet in the HD 206893 system. Evidence for deuterium burning in a planetary-mass companion". In: *Astronomy and Astrophysics* 671, L5, p. L5. DOI: 10.1051/0004-6361/202244727. arXiv: 2208.04867 [astro-ph.EP].

Ireland, M. J. (Aug. 2013). "Phase errors in diffraction-limited imaging: contrast limits for sparse aperture masking". In: *Monthly Notices of the Royal Astronomical Society* 433.2, pp. 1718–1728. DOI: 10.1093/mnras/stt859. arXiv: 1301.6205 [astro-ph.IM].

Jensen-Clem, Rebecca et al. (Dec. 2017). "A New Standard for Assessing the Performance of High Contrast Imaging Systems". In: *The Astronomical Journal* 155.1, p. 19. DOI: 10.3847/1538-3881/aa97e4. URL: https://doi.org/10.3847%2F1538-3881%2Faa97e4.

Jeunhomme, L. B. (1989). *Single-Mode Fiber Optics: Principles and Applications*. New York: Marcel Dekker.

Johansen, Anders et al. (Feb. 2021). "A pebble accretion model for the formation of the terrestrial planets in the Solar System". In: *Science Advances* 7.8, eabc0444. DOI: 10.1126/sciadv.abc0444. arXiv: 2102.08611 [astro-ph.EP].

Jordan, Camille (1875). "Essai sur la géométrie à *n* dimensions". fr. In: *Bulletin de la Société Mathématique de France* 3, pp. 103–174. DOI: 10.24033/bsmf.90. URL: http://www.numdam.org/articles/10.24033/bsmf.90/.

Jovanovic, N., B. Calvin, et al. (Dec. 2020). "Enhanced high-dispersion coronagraphy with KPIC phase II: design, assembly and status of sub-modules". In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 11447. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 114474U, 114474U. DOI: 10.1117/12.2563107. arXiv: 2012.06638 [astro-ph.IM].

Jovanovic, N., C. Schwab, et al. (Dec. 2016). "Enhancing Stellar Spectroscopy with Extreme Adaptive Optics and Photonics". In: *Publications of the Astronomical Society of the Pacific* 128.970, p. 121001. DOI: 10.1088/1538-3873/128/970/121001. arXiv: 1609.06388 [astro-ph.IM].

Jovanovic, Nemanja et al. (2023). "2023 Astrophotonics Roadmap: pathways to realizing multi-functional integrated astrophotonic instruments". In: *Journal of Physics: Photonics*. URL: http://iopscience.iop.org/article/10.1088/2515-7647/ace869.

Kasdin, N. Jeremy, Vanessa P. Bailey, et al. (2020). "The Nancy Grace Roman Space Telescope Coronagraph Instrument (CGI) technology demonstration". In: *Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave*. Ed. by Makenzie Lystrup et al. Vol. 11443. International Society for Optics and Photonics. SPIE, pp. 300–313. URL: https://doi.org/10.1117/12.2562997.

Kasdin, N. Jeremy and Isabelle Braems (Aug. 2006). "Linear and Bayesian Planet Detection Algorithms for the Terrestrial Planet Finder". In: *The Astrophysical Journal* 646.2, pp. 1260–1274. DOI: 10.1086/505017. URL: https://doi.org/10.1086%2F505017.

Kim, Yoo Jung, Michael P. Fitzgerald, Jonathan Lin, Steph Sallum, et al. (Apr. 2024). "Coherent Imaging with Photonic Lanterns". In: *Astrophysical Journal* 964.2, 113, p. 113. DOI: 10.3847/1538-4357/ad245e. arXiv: 2402.08158 [astro-ph.IM].

Kim, Yoo Jung, Michael P. Fitzgerald, Jonathan Lin, Yinzi Xin, et al. (Oct. 2024). "On the potential of spectroastrometry with photonic lanterns". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 10, 045001, p. 045001. DOI: 10.1117/1.JATIS.10.4.045001. arXiv: 2409.09120 [astro-ph.IM].

Konopacky, Quinn M. et al. (2013). "Detection of Carbon Monoxide and Water Absorption Lines in an Exoplanet Atmosphere". In: *Science* 339.6126, pp. 1398–1401. DOI: 10.1126/science.1232003. eprint: https://www.science.org/doi/pdf/10.1126/science.1232003. URL: https://www.science.org/doi/abs/10.1126/science.1232003.

– (Mar. 2013). "Detection of carbon monoxide and water absorption lines in an exoplanet atmosphere". English (US). In: *Science* 339.6126, pp. 1398–1401. ISSN: 0036-8075. DOI: 10.1126/science.1232003.

Kotani, T. et al. (2020). "Extremely high-contrast, high spectral resolution spectrometer REACH for the Subaru Telescope". In: *Adaptive Optics Systems VII*. Ed. by Laura Schreiber, Dirk Schmidt, and Elise Vernet. Vol. 11448. International Society for Optics and Photonics. SPIE, p. 1144878. DOI: 10.1117/12.2561755. URL: https://doi.org/10.1117/12.2561755.

Krist, J. (2020). *WFIRST CGI OS9 Time Series Simulations (Hybrid Lyot Coronagraph, Band 1)*. Online; accessed 26 August 2020.

Krist, John E. et al. (2023). "End-to-end numerical modeling of the Roman Space Telescope coronagraph". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 9.4, p. 045002. DOI: 10.1117/1.JATIS.9.4.045002. URL: https://doi.org/10.1117/1.JATIS.9.4.045002.

Lacour, S. et al. (Dec. 2020). "The ExoGRAVITY project: using single mode interferometry to characterize exoplanets". In: *Optical and Infrared Interferometry and Imaging VII*. Ed. by Peter G. Tuthill, Antoine Mérand, and Stephanie Sallum. Vol. 11446. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 114460O, 114460O. DOI: 10.1117/12.2561667. arXiv: 2101.07098 [astro-ph.EP].

Lafrenière, David et al. (May 2007). "A New Algorithm for Point-Spread Function Subtraction in High-Contrast Imaging: A Demonstration with Angular Differential Imaging". In: *Astrophysical Journal* 660.1, pp. 770–780. DOI: 10.1086/513180. arXiv: astro-ph/0702697 [astro-ph].

Lagadec, T. et al. (2021). "The GLINT South testbed for nulling interferometry with photonics: Design and on-sky results at the Anglo-Australian Telescope". In: *Publications of the Astronomical Society of Australia* 38, e036. DOI: 10.1017/pasa.2021.29.

Laginja, Iva et al. (2019). "Wavefront error tolerancing for direct imaging of exo-Earths with a large segmented telescope in space". In: *Techniques and Instrumentation for Detection of Exoplanets IX*. Vol. 11117. International Society for Optics and Photonics. SPIE, pp. 382–396. DOI: 10.1117/12.2530300. URL: https://doi.org/10.1117/12.2530300.

Lagrange, A-M et al. (2009). "A probable giant planet imaged in the β Pictoris disk-VLT/NaCo deep L'-band imaging". In: *Astronomy & Astrophysics* 493.2, pp. L21–L25.

Leboulleux, Lucie et al. (2018). *Pair-based Analytical model for Segmented Telescopes Imaging from Space (PASTIS) for sensitivity analysis*. arXiv: 1807.00870 [astro-ph.IM].

Leon-Saval, Sergio G., Alexander Argyros, and Joss Bland-Hawthorn (Dec. 2013a). "Photonic lanterns". In: *Nanophotonics* 2, pp. 429–440. DOI: 10.1515/nanoph-2013-0035. arXiv: 1503.03269 [physics.optics].

– (Dec. 2013b). "Photonic lanterns". In: *Nanophotonics* 2, pp. 429–440. DOI: 10.1515/nanoph-2013-0035. arXiv: 1503.03269 [physics.optics].

Leon-Saval, Sergio G., Nicolas K. Fontaine, et al. (Jan. 2014). "Mode-selective photonic lanterns for space-division multiplexing". In: *Opt. Express* 22.1, pp. 1036–1044. DOI: 10.1364/OE.22.001036. URL: http://opg.optica.org/oe/abstract.cfm?URI=oe-22-1-1036.

Lin, Jonathan, Michael P. Fitzgerald, et al. (Oct. 2022a). "Focal-plane wavefront sensing with photonic lanterns: theoretical framework". In: *Journal of the Optical Society of America B Optical Physics* 39.10, p. 2643. DOI: 10.1364/JOSAB.466227. arXiv: 2208.10563 [astro-ph.IM].

– (Oct. 2022b). "Focal-plane wavefront sensing with photonic lanterns: theoretical framework". In: *Journal of the Optical Society of America B Optical Physics* 39.10, p. 2643. DOI: 10.1364/JOSAB.466227. arXiv: 2208.10563 [astro-ph.IM].

Lin, Jonathan, Nemanja Jovanovic, and Michael P. Fitzgerald (July 2021). "Design considerations of photonic lanterns for diffraction-limited spectrometry". In: *Journal of the Optical Society of America B Optical Physics* 38.7, A51. DOI: 10.1364/JOSAB.423664. arXiv: 2106.10990 [astro-ph.IM].

Lovis, Christophe, Nicolas Blind, et al. (Aug. 2022). "RISTRETTO: high-resolution spectroscopy at the diffraction limit of the VLT". In: *Ground-based and Airborne Instrumentation for Astronomy IX*. Ed. by Christopher J. Evans, Julia J. Bryant, and Kentaro Motohara. Vol. 12184. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 121841Q, 121841Q. DOI: 10.1117/12.2627923. arXiv: 2208.14838 [astro-ph.EP].

Lovis, Christophe, Debra Fischer, et al. (2010). "Radial velocity techniques for exoplanets". In: *Exoplanets*, pp. 27–53.

Lozi, Julien et al. (Dec. 2020). "Status of the SCExAO instrument: recent technology upgrades and path to a system-level demonstrator for PSI". In: *Adaptive Optics Systems VII*. Ed. by Laura Schreiber, Dirk Schmidt, and Elise Vernet. Vol. 11448. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 114480N, 114480N. DOI: 10.1117/12.2562832.

Lyot, Bernard (1939). "The study of the solar corona and prominences without eclipses (George Darwin Lecture, 1939)". In: *Monthly Notices of the Royal Astronomical Society, Vol. 99, p. 580* 99, p. 580.

Macintosh, B et al. (2015). "Discovery and spectroscopy of the young jovian planet 51 Eri b with the Gemini Planet Imager". In: *Science* 350.6256, pp. 64–67.

Mâlin, Mathilde et al. (Jan. 2025). "First unambiguous detection of ammonia in the atmosphere of a planetary mass companion with JWST/MIRI coronagraphs". In: *Astronomy and Astrophysics* 693, A315, A315. DOI: 10.1051/0004-6361/202452695. arXiv: 2501.00104 [astro-ph.EP].

Mamajek, Eric and Karl Stapelfeldt (Apr. 2024). "Provisional NASA ExEP Mission Target Star List for the Habitable Worlds Observatory". In: *AAS/Division for Extreme Solar Systems Abstracts*. Vol. 56. AAS/Division for Extreme Solar Systems Abstracts, 628.17, p. 628.17.

Marois, Christian, David Lafreniere, et al. (Apr. 2006). "Angular Differential Imaging: A Powerful High-Contrast Imaging Technique". In: *The Astrophysical Journal* 641.1, pp. 556–564. DOI: 10.1086/500401. URL: https://doi.org/10.1086%2F500401.

Marois, Christian, David Lafrenière, et al. (Apr. 2006). "Angular Differential Imaging: A Powerful High-Contrast Imaging Technique". In: *The Astrophysical Journal* 641.1, pp. 556–564. DOI: 10.1086/500401. arXiv: astro-ph/0512335 [astro-ph].

Marois, Christian, Bruce Macintosh, et al. (2008). "Direct imaging of multiple planets orbiting the star HR 8799". In: *science* 322.5906, pp. 1348–1352.

Martinache, Frantz (Nov. 2010). "Kernel Phase in Fizeau Interferometry". In: *Astrophysical Journal* 724.1, pp. 464–469. DOI: 10.1088/0004-637X/724/1/464. arXiv: 1009.3933 [astro-ph.IM].

Martinod, Marc-Antoine et al. (Jan. 2021). "Scalable photonic-based nulling interferometry with the dispersed multi-baseline GLINT instrument". In: *Nature Communications* 12, 2465, p. 2465. DOI: 10.1038/s41467-021-22769-x.

Mawet, D. et al. (Jan. 2010). "The Vector Vortex Coronagraph: Laboratory Results and First Light at Palomar Observatory". In: *Astrophysical Journal* 709.1, pp. 53–57. DOI: 10.1088/0004-637X/709/1/53. arXiv: 0912.2287 [astro-ph.IM].

Mayor, Michel and Didier Queloz (Nov. 1995). "A Jupiter-mass companion to a solar-type star". In: *Nature* 378.6555, pp. 355–359. DOI: 10.1038/378355a0.

Mennesson, Bertrand P. et al. (2003). "Optical Planet Discoverer: how to turn a 1.5-m telescope into a powerful exo-planetary systems imager". In: *High-Contrast Imaging for Exo-Planet Detection*. Ed. by Alfred B. Schultz and Richard G. Lyon. Vol. 4860. International Society for Optics and Photonics. SPIE, pp. 32–44. DOI: 10.1117/12.457646. URL: https://doi.org/10.1117/12.457646.

Mercer, Anthony and Stamatellos, Dimitris (2020). "Planet formation around M dwarfs via disc instability - Fragmentation conditions and protoplanet properties". In: *Astronomy and Astrophysics* 633, A116. DOI: 10.1051/0004-6361/201936954. URL: https://doi.org/10.1051/0004-6361/201936954.

Milani, Kian, Kevin Derby, and Ewan Douglas (n.d.). *lina (v0.1)*. https://doi.org/10.5281/zenodo.11195112.

Mollière, Paul et al. (July 2022). "Interpreting the Atmospheric Composition of Exoplanets: Sensitivity to Planet Formation Assumptions". In: *The Astrophysical Journal* 934.1, p. 74. DOI: 10.3847/1538-4357/ac6a56. URL: https://dx.doi.org/10.3847/1538-4357/ac6a56.

National Research Council (2021). *Pathways to Discovery in Astronomy and Astrophysics for the 2020s*. Washington, DC: The National Academies Press.

Nemati, Bijan et al. (2020). "Method for deriving optical telescope performance specifications for Earth-detecting coronagraphs". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 6.3, pp. 1–42. DOI: 10.1117/1.JATIS.6.3.039002. URL: https://doi.org/10.1117/1.JATIS.6.3.039002.

Noll, Robert J. (Mar. 1976). "Zernike polynomials and atmospheric turbulence$*$". In: *J. Opt. Soc. Am.* 66.3, pp. 207–211. DOI: 10.1364/JOSA.66.000207. URL: http://www.osapublishing.org/abstract.cfm?URI=josa-66-3-207.

Norris, Barnaby R. M., Nick Cvetojevic, et al. (Jan. 2020). "First on-sky demonstration of an integrated-photonic nulling interferometer: the GLINT instrument". In: *Monthly Notices of the Royal Astronomical Society* 491.3, pp. 4180–4193. DOI: 10.1093/mnras/stz3277. arXiv: 1911.09808 [astro-ph.IM].

Norris, Barnaby R. M., Jin Wei, et al. (Oct. 2020). "An all-photonic focal-plane wavefront sensor". In: *Nature Communications* 11, 5335, p. 5335. DOI: 10.1038/s41467-020-19117-w. arXiv: 2003.05158 [astro-ph.IM].

Nowak, Mathias et al. (2020). "Direct confirmation of the radial-velocity planet β Pictoris c". In: *Astronomy & Astrophysics* 642, p. L2.

Ormel, Chris (Oct. 2024). "Planet Formation Mechanisms". In: *arXiv e-prints*, arXiv:2410.14430, arXiv:2410.14430. DOI: 10.48550/arXiv.2410.14430. arXiv: 2410.14430 [astro-ph.EP].

Otten, G. P. P. L. et al. (Feb. 2021). "Direct characterization of young giant exoplanets at high spectral resolution by coupling SPHERE and CRIRES+". In: *Astronomy and Astrophysics* 646, A150, A150. DOI: 10.1051/0004-6361/202038517. arXiv: 2009.01841 [astro-ph.IM].

Paschotta, Rudiger (Jan. 2022). *LP modes*. URL: https://www.rp-photonics.com/lp_modes.html.

Perrin, Marshall D. et al. (Oct. 2003). "The Structure of High Strehl Ratio Point-Spread Functions". In: *Astrophysical Journal* 596.1, pp. 702–712. DOI: 10.1086/377689. arXiv: astro-ph/0306468 [astro-ph].

Perryman, Michael et al. (Nov. 2014). "ASTROMETRIC EXOPLANET DETECTION WITH GAIA". In: *The Astrophysical Journal* 797.1, p. 14. DOI: 10.1088/0004-637X/797/1/14. URL: https://dx.doi.org/10.1088/0004-637X/797/1/14.

Pogorelyuk, Leonid, N. Jeremy Kasdin, and Clarence W. Rowley (Aug. 2019). "Reduced Order Estimation of the Speckle Electric Field History for Space-based Coronagraphs". In: *The Astrophysical Journal* 881.2, p. 126. DOI: 10.3847/1538-4357/ab2ecf. URL: https://doi.org/10.3847%2F1538-4357%2Fab2ecf.

Pope, Benjamin J. S. et al. (Jan. 2021). "Kernel Phase and Coronagraphy with Automatic Differentiation". In: *Astrophysical Journal* 907.1, 40, p. 40. DOI: 10.3847/1538-4357/abcb00. arXiv: 2011.09780 [astro-ph.IM].

Por, E. H. and S. Y. Haffert (Mar. 2020). "The Single-mode Complex Amplitude Refinement (SCAR) coronagraph. I. Concept, theory, and design". In: *Astronomy and Astrophysics* 635, A55, A55. DOI: 10.1051/0004-6361/201731616.

Por, E. H., S. Y. Haffert, et al. (2018). "High Contrast Imaging for Python (HCIPy): an open-source adaptive optics and coronagraph simulator". In: *Adaptive Optics Systems VI*. Vol. 10703. Proc. SPIE. DOI: 10.1117/12.2314407. URL: https://doi.org/10.1117/12.2314407.

Potier, A. et al. (Sept. 2022). "Increasing the raw contrast of VLT/SPHERE with the dark hole technique. II. On-sky wavefront correction and coherent differential imaging". In: *Astronomy & Astrophysics* 665, A136, A136. DOI: 10.1051/0004-6361/202244185. arXiv: 2208.11244 [astro-ph.IM].

Pourcelot, R. et al. (Apr. 2023). "Low-order wavefront control using a Zernike sensor through Lyot coronagraphs for exoplanet imaging. II. Concurrent operation with stroke minimization". In: *Astronomy and Astrophysics* 672, A73, A73. DOI: 10.1051/0004-6361/202244857. arXiv: 2301.03242 [astro-ph.IM].

Price, Jonathan H. V. et al. (2006). "Non-silica microstructured optical fibers for mid-IR supercontinuum generation from 2 µm – 5 µm". In: *Fiber Lasers III: Technology, Systems, and Applications*. Ed. by Andrew J. W. Brown et al. Vol. 6102. International Society for Optics and Photonics. SPIE, 61020A. DOI: 10.1117/12.660942. URL: https://doi.org/10.1117/12.660942.

Pueyo, Laurent (June 2016a). "Detection and Characterization of Exoplanets using Projections on Karhunen Loeve Eigenimages: Forward Modeling". In: *Astrophysical Journal* 824.2, 117, p. 117. DOI: 10.3847/0004-637X/824/2/117. arXiv: 1604.06097 [astro-ph.IM].

– (June 2016b). "Detection and Characterization of Exoplanets using Projections on Karhunen Loeve Eigenimages: Forward Modeling". In: *Astrophysical Journal* 824.2, 117, p. 117. DOI: 10.3847/0004-637X/824/2/117. arXiv: 1604.06097 [astro-ph.IM].

Riggs, A J Eldorado et al. (2018). "Fast linearized coronagraph optimizer (FALCO) I: a software toolbox for rapid coronagraphic design and wavefront correction". In: *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*. Ed. by Makenzie Lystrup et al. Vol. 10698. International Society for Optics and Photonics. SPIE, pp. 878–888. DOI: 10.1117/12.2313812. URL: https://doi.org/10.1117/12.2313812.

Rouan, D. et al. (Nov. 2000). "The Four-Quadrant Phase-Mask Coronagraph. I. Principle". In: *Publications of the Astronomical Society of the Pacific* 112.777, p. 1479. DOI: 10.1086/317707. URL: https://dx.doi.org/10.1086/317707.

Ruane, Garreth, Daniel Echeverri, et al. (2019). "Vortex fiber nulling for exoplanet observations: conceptual design, theoretical performance, and initial scientific yield predictions". In: *Techniques and Instrumentation for Detection of Exoplanets IX*. Ed. by Stuart B. Shaklan. Vol. 11117. International Society for Optics and Photonics. SPIE, pp. 366–381. DOI: 10.1117/12.2528555. URL: https://doi.org/10.1117/12.2528555.

Ruane, Garreth, Ji Wang, et al. (Nov. 2018). "Efficient Spectroscopy of Exoplanets at Small Angular Separations with Vortex Fiber Nulling". In: *The Astrophysical Journal* 867.2, p. 143. DOI: 10.3847/1538-4357/aae262. URL: https://dx.doi.org/10.3847/1538-4357/aae262.

Schwarz, Henriette et al. (Sept. 2016). "The slow spin of the young substellar companion GQ Lupi b and its orbital configuration". In: *Astronomy & Astrophysics* 593, A74, A74. DOI: 10.1051/0004-6361/201628908. arXiv: 1607.00012 [astro-ph.EP].

Seager, S. (2008). "Exoplanet Transit Spectroscopy and Photometry". In: *Strategies of Life Detection*. Ed. by Oliver Botta et al. Boston, MA: Springer US, pp. 345–354. ISBN: 978-0-387-77516-6. DOI: 10.1007/978-0-387-77516-6_25. URL: https://doi.org/10.1007/978-0-387-77516-6_25.

Serabyn, E. et al. (Oct. 2019). "Nulling at short wavelengths: theoretical performance constraints and a demonstration of faint companion detection inside the diffraction limit with a rotating-baseline interferometer". In: *Monthly Notices of the Royal Astronomical Society* 489.1, pp. 1291–1303. DOI: 10.1093/mnras/stz2163. arXiv: 1908.05977 [astro-ph.IM].

Serabyn, Eugene, Kurt Liewer, and Garreth Ruane (May 2024). “Geometric-phase-based phase-knife mask for stellar nulling and coronagraphy”. In: *Opt. Express* 32.11, pp. 19924–19934. DOI: 10.1364/OE.520548. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-32-11-19924.

Snellen, I. et al. (Apr. 2015). “Combining high-dispersion spectroscopy with high contrast imaging: Probing rocky planets around our nearest neighbors”. In: *Astronomy and Astrophysics* 576, A59, A59. DOI: 10.1051/0004-6361/201425018. arXiv: 1503.01136 [astro-ph.EP].

Soummer, Rémi (Jan. 2005). “Apodized Pupil Lyot Coronagraphs for Arbitrary Telescope Apertures”. In: *Astrophysical Journal Letters* 618.2, pp. L161–L164. DOI: 10.1086/427923. arXiv: astro-ph/0412221 [astro-ph].

Soummer, Rémi, Laurent Pueyo, and James Larkin (Aug. 2012a). “Detection and Characterization of Exoplanets and Disks Using Projections on Karhunen-Loève Eigenimages”. In: *Astrophysical Journal Letters* 755.2, L28, p. L28. DOI: 10.1088/2041-8205/755/2/L28. arXiv: 1207.4197 [astro-ph.IM].

– (Aug. 2012b). “Detection and Characterization of Exoplanets and Disks Using Projections on Karhunen-Loève Eigenimages”. In: *Astrophysical Journal, Letters* 755.2, L28, p. L28. DOI: 10.1088/2041-8205/755/2/L28. arXiv: 1207.4197 [astro-ph.IM].

Sparks, William B. and Holland C. Ford (Oct. 2002). “Imaging Spectroscopy for Extrasolar Planet Detection”. In: *Astrophysical Journal* 578.1, pp. 543–564. DOI: 10.1086/342401. arXiv: astro-ph/0209078 [astro-ph].

Stark, Christopher C. et al. (Aug. 2015). “Lower Limits on Aperture Size for an ExoEarth Detecting Coronagraphic Mission”. In: *The Astrophysical Journal* 808.2, 149, p. 149. DOI: 10.1088/0004-637X/808/2/149. arXiv: 1506.01723 [astro-ph.SR].

Step Index Fibers (2025). https://www.fiberoptics4sale.com/blogs/wave-optics/step-index-fibers. [Accessed 27-03-2025].

Traub, W. A. and B. R. Oppenheimer (2010). “Direct Imaging of Exoplanets”. In: *Exoplanets*. Ed. by S. Seager. University of Arizona Press, pp. 111–156.

Tuthill, Peter (2022). “Nulling interferometry: high contrast science for single large apertures”. In: *Adaptive Optics Systems VIII*. Ed. by Laura Schreiber, Dirk Schmidt, and Elise Vernet. Vol. 12185. International Society for Optics and Photonics. SPIE, 121858P. DOI: 10.1117/12.2627887. URL: https://doi.org/10.1117/12.2627887.

Tuthill, Peter G., John D. Monnier, and William C. Danchi (2000). “Aperture masking interferometry on the Keck I Telescope: new results from the diffraction limit”. In: *Interferometry in Optical Astronomy*. Ed. by Pierre J. Lena and Andreas Quirrenbach. Vol. 4006. International Society for Optics and Photonics. SPIE, pp. 491–498. DOI: 10.1117/12.390244. URL: https://doi.org/10.1117/12.390244.

Tyson, Robert K. (2000). *Introduction to Adaptive Optics*. Bellingham: The Society of Photo-optical Instrumetation Engineers.

Velázquez-Benítez, Amado M. et al. (June 2018). "Scaling photonic lanterns for space-division multiplexing". In: *Scientific Reports* 8.1, p. 8897. ISSN: 2045-2322. DOI: 10.1038/s41598-018-27072-2. URL: https://doi.org/10.1038/s41598-018-27072-2.

Virtanen, Pauli et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17, pp. 261–272. DOI: https://doi.org/10.1038/s41592-019-0686-2.

Vogt, Frédéric P. A. et al. (Dec. 2011). "Coronagraphic Low-Order Wavefront Sensor: Postprocessing Sensitivity Enhancer for High-Performance Coronagraphs". In: *Publications of the Astronomical Society of the Pacific* 123.910, pp. 1434–1441. ISSN: 1538-3873. DOI: 10.1086/663723. URL: http://dx.doi.org/10.1086/663723.

Wang, Jason J., Matas Kulikauskas, and Sarah Blunt (Jan. 2021). *whereistheplanet: Predicting positions of directly imaged companions*. Astrophysics Source Code Library, record ascl:2101.003. ascl: 2101.003.

Wang, Jason J., Dimitri Mawet, et al. (June 2024). "The high-contrast performance of the Keck Planet Imager and Characterizer". In: *arXiv e-prints*, arXiv:2406.15028, arXiv:2406.15028. DOI: 10.48550/arXiv.2406.15028. arXiv: 2406.15028 [astro-ph.IM].

Wang, Jason J., Jean-Baptiste Ruffio, et al. (Oct. 2021). "Detection and Bulk Properties of the HR 8799 Planets with High-resolution Spectroscopy". In: *Astronomical Journal* 162.4, 148, p. 148. DOI: 10.3847/1538-3881/ac1349. arXiv: 2107.06949 [astro-ph.EP].

Wang, Ji et al. (Apr. 2017). "Observing Exoplanets with High Dispersion Coronagraphy. I. The Scientific Potential of Current and Next-generation Large Ground and Space Telescopes". In: *Astronomical Journal* 153.4, 183, p. 183. DOI: 10.3847/1538-3881/aa6474. arXiv: 1703.00582 [astro-ph.EP].

Weakly Guiding Fibers (2025). https://www.fiberoptics4sale.com/blogs/wave-optics/weakly-guiding-fibers. [Accessed 27-03-2025].

Wolszczan, A. and D. A. Frail (Jan. 1992). "A planetary system around the millisecond pulsar PSR1257 + 12". In: *Nature* 355.6356, pp. 145–147. DOI: 10.1038/355145a0.

Xin, Yinzi, Daniel Echeverri, et al. (2024). "Laboratory demonstration of a Photonic Lantern Nuller in monochromatic and broadband light". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 10.2, p. 025001. DOI: 10.1117/1.JATIS.10.2.025001. URL: https://doi.org/10.1117/1.JATIS.10.2.025001.

Xin, Yinzi, Nemanja Jovanovic, et al. (Oct. 2022). "Efficient Detection and Characterization of Exoplanets within the Diffraction Limit: Nulling with a Mode-selective Photonic Lantern". In: *The Astrophysical Journal* 938.2, 140, p. 140. DOI: 10.3847/1538-4357/ac9284. arXiv: 2209.07644 [astro-ph.IM].

Xin, Yinzi, Jerry W. Xuan, et al. (July 2023). "On-sky speckle nulling through a single-mode fiber with the Keck Planet Imager and Characterizer". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 9, 035001, p. 035001. DOI: 10.1117/1.JATIS.9.3.035001. arXiv: 2307.11893 [astro-ph.IM].

Xuan, Jerry W. et al. (July 2024). "Are These Planets or Brown Dwarfs? Broadly Solar Compositions from High-resolution Atmospheric Retrievals of ~10–30 M $_{Jup}$ Companions". In: *Astrophysical Journal* 970.1, 71, p. 71. DOI: 10.3847/1538-4357/ad4796. arXiv: 2405.13128 [astro-ph.EP].

Ygouf, Marie et al. (July 2016). "Data processing and algorithm development for the WFIRST coronagraph: comparison of RDI and ADI strategies and impact of spatial sampling on post-processing". In: *Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave*. Ed. by Howard A. MacEwen et al. Vol. 9904. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 99045M, p. 99045M. DOI: 10.1117/12.2231581.

Youdin, Andrew N. and Jeremy Goodman (Feb. 2005). "Streaming Instabilities in Protoplanetary Disks". In: *The Astrophysical Journal* 620.1, p. 459. DOI: 10.1086/426895. URL: https://dx.doi.org/10.1086/426895.

Youdin, Andrew N. and Zhaohuan Zhu (Jan. 2025). "Formation of Giant Planets". In: *arXiv e-prints*, arXiv:2501.13214, arXiv:2501.13214. DOI: 10.48550/arXiv.2501.13214. arXiv: 2501.13214 [astro-ph.EP].

Zhang, Yapeng et al. (July 2021). "The ^{13}CO-rich atmosphere of a young accreting super-Jupiter". In: *Nature* 595.7867, pp. 370–372. DOI: 10.1038/s41586-021-03616-x. arXiv: 2107.06297 [astro-ph.EP].