

Correcting for Quasi-static Wavefront Error Drifts in High-contrast and Wide-field Imaging Telescopes

SUSAN F. REDMOND

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
MECHANICAL AND AEROSPACE ENGINEERING

ADVISER: EGEMEN KOLEMEN, WILLIAM C. JONES, N. JEREMY KASDIN

SEPTEMBER 2023

© COPYRIGHT BY SUSAN F. REDMOND, 2023. ALL RIGHTS RESERVED.

ABSTRACT

The study of both Exoplanets and Dark Matter provide valuable information on how we humans came to exist and whether we are alone in the universe. Directly imaging planets requires long integration times (hours) with a coronagraphic instrument due to the limited number of photons. The wavefront must be stable over a small field-of-view on the same time scale, which is often difficult in space due to time-varying wavefront errors from thermal gradients and other mechanical instabilities. In order to directly observe a habitable exoplanet, a series of masks are used to redistribute and block the starlight in the region where the planet might be; these masks are referred to as coronagraphs. Wavefront errors from misalignments and manufacturing defects of the optical surfaces must be corrected via deformable mirrors in order to achieve the required level of starlight suppression. Dark Matter mapping via cluster weak lensing, like exoplanet imaging, requires a very stable wavefront over hour-long integration times but also requires a large field-of-view. Dark Matter acts as a gravitational lens and the Dark Matter in a foreground galaxy cluster will distort the light travelling from the background galaxy cluster. The level of distortion can be used to infer the mass of Dark Matter present in the foreground galaxy cluster. When performing weak lensing measurements from a balloon-borne telescope, there are thermal and structural changes throughout the integration time which can introduce quasi-static wavefront error drifts that introduce errors in the Dark Matter measurement.

The instruments required to directly image an exoplanet versus to measure Dark Matter via weak lensing are quite different but both suffer from the effects of quasi-static wavefront error drifts. I present solutions for exoplanet imaging that utilize active optics and focal plane wavefront sensing and control techniques. I also provide a study to inform the design of an active-optics system for a weak-lensing application.

For directly imaging exoplanets, an overview of a Dark Zone Maintenance algorithm is provided which combines an Extended Kalman Filter as an estimator and Electric Field Conjugation as a controller. Also provided are laboratory demonstrations that show the algorithm maintaining the dark zone presence of various wavefront error drifts. Deformable Mirrors are used to inject wavefront error drifts both for monochromatic and broadband (10% bandwidth at 635 nm) experiments. Low-photon cases are provided where the images obtained from the testbed are processed prior to being passed to the estimator to mimic the pixel sampling, photon rate, and detector noise expected on the Roman Space Telescope. I also demonstrate detection of fake planets injected into the testbed data using various

post-processing techniques. Experiments are performed on the High-contrast imager for Complex Aperture Telescopes at the Space Telescope Science Institute, and both the In Air Coronagraph Testbed and Occulting Mask Coronagraph testbed at the Jet Propulsion Laboratory.

For weak lensing, I perform a study of the stability of the image quality of the 2023 Super-pressure Balloon-borne Imaging Telescope flight results. These results are then used to inform the design for a new 1.3 m optical-near-UV balloon-borne telescope (GigaBIT). GigaBIT has a 0.5° field of view over which the wavefront must be stable on timescales of an hour as the instrument moves in azimuth, elevation, and yaw to remain locked on a target. The Super-pressure Balloon-borne Imaging Telescope was designed as a pathfinder experiment and we use the thermal, mechanical, and pointing behaviour to determine a sufficient active optics design for the next generation instrument.

Contents

ABSTRACT	iii
1 INTRODUCTION	1
1.1 Motivation	1
1.2 Scientific approach	5
1.3 The field today	10
1.4 Quasi-static wavefront error drifts	15
1.5 Thesis overview	16
2 BACKGROUND	19
2.1 Fourier and Fresnel optics	20
2.2 Coronagraphs	24
2.3 Wavefront sensing and control	28
2.4 Direct imaging post processing methods	37
2.5 Weak lensing measurement error ⁹⁵	40
2.6 Summary	42
3 GENERATING THE DARK ZONE USING FOCAL PLANE BROADBAND ESTIMATION	43
3.1 Motivation	44
3.2 Broadband estimator algorithm	45
3.3 Laboratory results	52
3.4 Conclusions and future work	59
4 QUASI-STATIC DRIFT CORRECTION ON A CORONAGRAPH TESTBED	61
4.1 Chapter overview	62
4.2 Introduction	63
4.3 Dark zone maintenance	64
4.4 HiCAT laboratory results	78
4.5 IACT laboratory results	90

4.6	Conclusions	92
5	LOW SIGNAL-TO-NOISE DARK ZONE MAINTENANCE	99
5.1	Chapter overview	100
5.2	Dark zone maintenance algorithm adjustments	102
5.3	HiCAT results	107
5.4	Conclusions and future work	110
6	POST-PROCESSING OF DARK ZONE MAINTENANCE DATA	112
6.1	Chapter overview	112
6.2	Planet injection	113
6.3	Signal-to-noise ratio calculations	115
6.4	Post processing using Karhunen-Loève image projection	116
6.5	Angular differential imaging with DZM	118
6.6	HiCAT experimental results using ADI	119
6.7	DZM with real-time incoherent light estimation	123
6.8	HiCAT experimental results using CDI and IAI	131
6.9	HiCAT experimental results using augmented EKF and RDI	135
6.10	Planet detection	137
6.11	Conclusions and future work	139
7	DEVELOPMENT OF THE GIGABIT INSTRUMENT SPECIFICATIONS	141
7.1	Chapter overview	141
7.2	GigaBIT technical specifications	142
7.3	Overview of the GigaBIT pathfinder: the Super-pressure Balloon-borne Imaging Telescope (SuperBIT)	144
7.4	SuperBIT 2023 flight results	146
7.5	GigaBIT instrument specifications	173
7.6	Conclusions and future work	181
8	SUMMARY AND FUTURE WORK	183
	APPENDIX A HiCAT HARDWARE AND AMBIENT ENVIRONMENT	188
	A.1 High Contrast Imager for Complex Aperture Telescopes (HiCAT)	189
	A.2 HiCAT hardware state	191
	APPENDIX B CHOOSING AN OPTIMAL DITHER FOR BMC RANDOM WALK DRIFT	196
	REFERENCES	221

List of Figures

1.1	Sample planet spectra. ⁶⁹	3
1.2	Evolution of the universe.	4
1.3	Basic high-contrast optical layout.	6
1.4	Cluster strong lensing.	7
1.5	Effect of weak lensing	8
1.6	SuperBIT-Hubble FOV comparison.	9
1.7	The Bullet Cluster.	9
1.8	Atmospheric transmission at different altitudes.	13
1.9	Predicted GigaBIT performance with current and upcoming space telescopes.	15
2.1	Fourier optics derivation plane definitions.	20
2.2	Simplified coronagraph optical layout.	24
2.3	RST shaped pupil coronagraph. ¹⁷	26
2.4	Lyot coronagraph masks.	26
2.5	PSF evolution with coronagraph components.	27
2.6	Pupil plane WFS.	29
2.7	Deformable mirror technologies.	30
2.8	Simple high-contrast optical layout with DMs.	32
2.9	Focal plane wavefront estimation and control.	33
2.10	Sample high contrast image and post processed image with planet.	37
2.11	Effect of PSF on an image.	41
3.1	Pairwise probe example.	51
3.2	Cube mode estimate comparison.	53
3.3	Contrast versus radius and time for cube mode experiment.	55
3.4	Broadband estimator estimates.	56
3.5	Broadband image-estimate comparison.	56
3.6	Broadband mode contrast vs radius for select iterations.	57
3.7	Broadband dark zone generation comparison.	58

4.1	Dark zone maintenance control loop.	66
4.2	Dark zone dig vs dark zone maintenance.	67
4.3	DZM iteration 26 open-loop image and estimate.	77
4.4	Dither effect on raw contrast.	79
4.5	Dither vs raw contrast fit.	80
4.6	BMC monochromatic drift DZM experiment.	82
4.7	BMC drift multi-wavelength DZM experiment.	84
4.8	Multi-wavelength DZM temporally averaged image and associated histogram.	86
4.9	IrisAO segmented aperture PTT drift DZM experiment.	88
4.10	BMC and IrisAO drift DZM experiment.	90
4.11	IACT multi-wavelength DZM results.	92
4.12	LUVOIR-HiCAT histogram comparison.	95
5.1	Low SNR DZM flow chart.	103
5.2	Low SNR conversion.	107
5.3	Low SNR DZM contrast vs time.	109
5.4	Low SNR DZM coadded image.	110
6.1	DZM planet injection flowchart.	114
6.2	Planet injection demonstration.	114
6.3	Mask for noise image.	116
6.4	ADI demonstration.	118
6.5	ADI using KLIP on HiCAT.	120
6.6	Contrast vs time for low SNR ADI HiCAT experiment.	122
6.7	ADI using KLIP for 4×10^{-8} planet.	122
6.8	Simulated contrast vs time for nominal EKF with bright planet.	124
6.9	Simulation results showing planet in coherent estimate for nominal DZM.	125
6.10	CDI with DZM demonstration.	130
6.11	CDI with a bright planet using HiCAT low SNR DZM data.	132
6.12	CDI-IAI comparison of SNR vs frames used for a bright planet	133
6.13	CDI and IAI for 8×10^{-8} planet in HiCAT DZM experiment.	134
6.14	HiCAT low SNR CDI dim planet SNR vs frames used.	135
6.15	SNR vs KLIP mode using RDI for HiCAT low SNR bright planet experiment.	136
6.16	RDI KLIP result for HiCAT low SNR experiment with dim planet.	136
6.17	KLIP planet detection results.	138
7.1	HST ACS FOV compared to a single GigaBIT detector.	144
7.2	SuperBIT gondola.	146
7.3	SuperBIT telescope.	147

7.4	SuperBIT 2023 Flight Path	148
7.5	FSC Centroid Position with FGS off.	149
7.6	Rolling extrema demonstration for FSC ₁ X centroid data.	150
7.7	Required tip-tilt throw vs time for FGS off data.	151
7.8	Heater breakdown for OTA.	152
7.9	Diurnal cycle of OTA control thermistors.	153
7.10	Reference OTA thermistor data.	154
7.11	Primary mirror diurnal thermal cycle.	155
7.12	Mean nightly temperature of primary mirror reference thermistors based on clocking.	156
7.13	SuperBIT source extraction examples for PLCKG287 target.	157
7.14	PSFEx point source selection.	159
7.15	Focal plane metapixel breakdown.	159
7.16	2D Gaussian fit of PSFEx model.	161
7.17	2D Gaussian fit PSF residual.	162
7.18	Focal plane metapixel validation.	163
7.19	Mean standard deviation of PSF characteristics within a metapixel across the focal plane.	164
7.20	Ellipticity correlation plots for Abell 3526.	166
7.21	FSC PSF statistics.	167
7.22	FSC-based PSF statistics for first column metapixels.	168
7.23	Diurnal PSF statistics.	170
7.24	Telescope elevation PSF statistics.	172
7.25	GigaBIT telescope design.	175
7.26	GigaBIT tip-tilt requirements vs distance from focal plane.	179
7.27	Tipt-tilt characteristics for a 16-bit controller.	180
A.1	HiCAT Layout ¹²⁷	190
A.2	Effect of humidity drift on HiCAT on raw contrast.	193
A.3	IrisAO closed-loop drift contrast-humidity correlation.	195
B.1	Estimation error vs dither.	199
B.2	Effect of varying dither on control response.	200
B.3	Choosing an optimal dither for BMC random walk drifts.	202

FOR SAM.

Acknowledgments

Growing up in Newfoundland and Labrador, I have a deep appreciation of community and the notion that “it takes a village”; my PhD experience reflects that on multiple levels. Most people are lucky to have on good mentor throughout their PhD, I had five!

Jeremy, thank you for taking me on and allowing me to expand my research in my own way. You have always been supportive of my different projects and provided important insight when I needed it, even while being remote. You adjusted your advising style to be what made me the most productive which is what great mentors do. I have been extremely privileged to work with you over the past five years and I look forward to our continued collaborations. Also, to Jeremy’s group especially He, Mia, Anthony, Leonid, Jessica, and Mike: thank you for welcoming me and answering all of my high-contrast and hardware related questions.

Egemen, thank you for taking me in! You already had a large group when I joined and you have been supportive of all of my research despite it not being related to your groups work. You have been a great career mentor and gave me a great group of students to bounce problems off of. Thanks to Ricardo, Joe, Brian, Andy, Daniel, Dario, Josiah, and (especially) Rory for helping me with my obscure python problems and letting me crash your group meetings to learn about Fusion (boo ELMs!).

Bill, I think it is safe to say I would not be at Princeton without you. Thank you for your unwavering support and flexibility. I have now worked with you in four countries on two projects in a variety of high stress and high consequence scenarios and every time I have been grateful to have you in our corner advocating for the team and the project. To the Jones group as well as the entire SPIDER-II and SuperBIT teams: thank you for being such a great group of humans regardless of the amount of stress or sleep deprivation. There aren’t many people I would rather be stuck in a rural, foreign place with for months while either getting cooked or frozen in the process of trying to build a world-class balloon telescope. Stevie, thank you for all your help getting here and being an amazing role model. Joseph, Vy, Suren, Sho, Jared, and Johanna, you were the best night shift anyone could ask for (even if you left the heat on too high). Steve, thank you for having infinite patience and being a great adventure buddy on every campaign. Corwin, you will always be the SPIDER-

II MVP in my books for finding the reaction wheel mounting joints. Sasha and Ricardo, thank you for making an extended pit stop in McMurdo and for picking up the pieces of SPIDER-II when it landed! Jason, team gondola would not have been the same without you. Steven Li, thank you for your creativity and optimism and being able to turn out a Solidworks model for a replacement part in under an hour (and being willing to dremel anything). Simon, thanks for being up for any outdoor activity regardless of the time and Elle thank you for being a great friend and tolerating my 100 questions on how to properly light seal the inserts. Jeff, thanks for being a great PI and getting SPIDER-II a launch opportunity! SuperBIT launched on April 17th and was operational for over a month; that kind of accomplishment is not possible without the entire team being on their A game for over three months (assembly and flight). Big shoutout to David for the autopilot feature that helped keep us all (mostly) sane during the flight. Also thank you to Barth, Richard, Emaad, Ajay, Spencer, and Ellen for helping with the infinite amount of arts and crafts as well as your day jobs! As for the SPIDER-II people also on SuperBIT, you already got your appreciation.

A balloon flight requires a lot more than the science team and I have been fortunate to work with a great set of launch, telemetry, and recovery teams throughout my PhD. To Randall, Hugo, and Robert: the SPIDER-II and SuperBIT launches would not have happened without your expertise and advocacy. Big thank you to the recovery teams which included Rose, Scott, Alec, Daniel, Zac, Leonardo, and Jesse, we would not have our data without you; thanks for braving the cold, carbon fibre dust, and Guanacos! Additional thanks to all the wonderful people I met in McMurdo that helped SPIDER-II happen and were just great people to get to know. Kaija, Dan, Lacey, Ryan, Branden, Toby, and Lester, you all made the McMurdo experience richer.

Rémi, Laurent, and Marshall thank you for adopting me into the Makidon lab and being so supportive. You all went above and beyond the required 'grant testbed access' and have spent many hours helping me debug and enhance my code as well as helping me plan my career. I would not be where I am today without your expertise and insight. To all of the past and present Makidon lab members including Iva, Scott, Jules, Matt, Meji, Ananya, Raphaël, and Bry, thank you for being a great remote team and a wealth of knowledge. Special thanks to Emiel for your patience and insight which has greatly helped my research progress and evolve.

My PhD would have been a lot less enjoyable without the incredible set of people I am fortunate to call friends. Meghan and Amlan, I would not have passed generals without our weekly study sessions. Thank you for making that process a little more tolerable and being great friends. Alec, thanks for always being up for an adventure or a coffee break, regardless of what time zone I am in. To the Princeton climbing crew, thanks for keeping me sane and being amazing adventure buddies. Annie, Oak, Kelsey, Jack, Isabel, Eric, Tom, and

Valentin: you are more family than friend at this point and I know you're always in my corner. Jiarong and Link, you were there day-in and day-out throughout my PhD. Thanks for always being there and for being great COVID buddies. My friends from home especially Erin, Myfanwy, Philippa, and Claire, you have seen me through it all and make sure to let me know when I start sounding like a mainlander. To the barn girls, you have a way of hyping me up while keeping me grounded and for that I will always be grateful. Devin, thanks for always reminding me to take the trash out. Robyn, Matt C., and Matt M.: thanks for being an amazing team whether or not I am in the boat.

Last but not least, thank you to my incredible family that has supported me through all four degrees. To my parents that always encouraged me to explore in all aspects of my life (even if they don't really understand what I do) and to Thomas and Jacquelyn that have seen me through everything, I would not have made it here without you. To my extended family abroad especially Julie, Juanita, Phillip, Aunt Elizabeth, Uncle David, and Steven, thanks for often being my 'home away from home'. Also thanks to my extended family at home that make coming home so special.

Flipbook artwork provided by the wonderful Kailey Whitman.

This work was supported in part by the National Aeronautics and Space Administration under Grant 80NSSC19K0120 issued through the Strategic Astrophysics Technology/Technology Demonstration for Exoplanet Missions Program (SAT-TDEM; PI: R. Soummer). E.H.P. contributed to this work and is supported by the NASA Hubble Fellowship grant #HST-HF2-51467.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. Part of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). SuperBIT is funded in part through NASA grants NNX16AF65G and 80NSSC22K0365.

This dissertation carries T#3456 in the records of the Department of Mechanical and Aerospace Engineering.

1

Introduction

1.1 MOTIVATION

1.1.1 ARE WE ALONE?

“Are we alone in the Universe?” is a question that has intrigued humanity from the beginning. Seeing as we have not yet developed a functional Millennium Falcon, we must investigate the potential for life on other planets using imaging. The study of exoplanets, or

planets that orbit other stars, is a growing field. To determine if an exoplanet could sustain life, we use the spectra of the light from the exoplanet. The absorption lines in the spectra of the exoplanet indicate the molecules present in the exoplanet's atmosphere as shown in Fig. 1.1. Molecules like water and oxygen are used as biosignatures for potential life.¹¹² To date, liquid water has not been found on an exoplanet. One of the main reasons for this is that the planets we can obtain spectra for with the technology available are mainly large, self-luminous planets (like an exo-Jupiter) with large orbits around their host star. Planets that would have liquid water, a sufficient condition for a habitable exoplanet, are expected to be small, rocky planets that orbit relatively close to their host star (like Earth).¹¹² Another important consideration is the type of star, as the system must be old enough such that life could have evolved but not so old that the star is in decline and flaring. Schwieterman et al. 2018¹¹² considers planets orbiting F, G, K, M stars (10^9 – 10^{12} years old) as potential targets. Sample spectra for different types of exoplanets are shown in Fig. 1.1.

The bottom middle panel of Fig. 1.1 highlights another reason why studying small rocky planets is so interesting. To date, the methods used to detect exoplanets have been biased to exoplanets with certain characteristics (large mass, low orbital radius). In order to understand how planetary systems form and the roles of different types of planets within a solar system, we need a more comprehensive dataset of the exoplanets that exist. This will allow us to determine how unique our solar system is, even if we cannot find an exo-Earth.

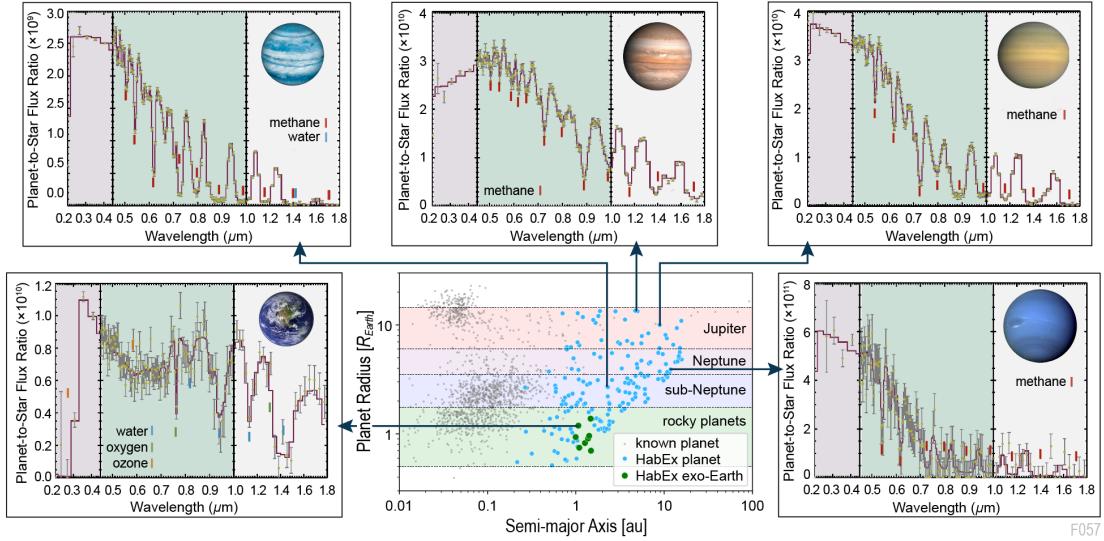


Figure 1.1: Sample planet spectra for targets from the HabEx mission study by Mennesson et al. 2020.⁶⁹ In the bottom middle panel, the cyan and green points are planets that would be discovered by HabEX and the grey points are previously discovered planets.

1.1.2 HOW DID WE GET HERE?⁹⁵

Cosmology is the study of the large scale evolution of the Universe. The Λ Cold Dark Matter (Λ CDM) is the most widely accepted model and assumes the Universe contains radiation, matter, and dark energy. The Friedmann equations⁸³ (Eqn 1.1), describe the temporal evolution of the Universe and its geometry as a function of the component densities (Ω_i):

$$H^2 = H_0^2 (\Omega_R a^{-4} + \Omega_M a^{-3} + \Omega_k a^{-2} + \Omega_\Lambda) \quad (1.1)$$

where the size of the universe is given by the scale factor a , and the expansion rate by the Hubble parameter $H = \dot{a}/a$ and its current value $H_0 = 67.36 \pm 0.54 \text{ km s}^{-1} \text{ Mpc}^{-1}$.⁸⁶ Since there is currently no theoretical understanding of why these parameters have their

specific values in our Universe, the model parameters must be determined observationally. Ω_M describes the mass of the Universe which includes both baryonic and dark matter. Baryonic matter is relatively easy to study as we can see it. Dark matter has mass but does not emit or interact with anything emitting around it. If we develop instruments that can determine the amount of dark matter in the Universe, we can improve the accuracy of the Λ CDM model. Figure 1.2 is a visual depiction of the Λ CDM model combined with inflation theory. It starts with quantum fluctuations which lead to the commonly phrased ‘Big Bang’ after which the Universe rapidly expands (inflation). As the plasma cools, the photons decouple giving us the Cosmic Microwave Background (CMB). There is then a dark period before the first stars begin to form from Hydrogen and Helium. As the stars begin to die, they spew the new elements that they cooked in their cores into the Universe. Eventually, enough of these new elements exist to create planets and stars, which begin to group in gravitationally-bound systems we dub ‘galaxies’.

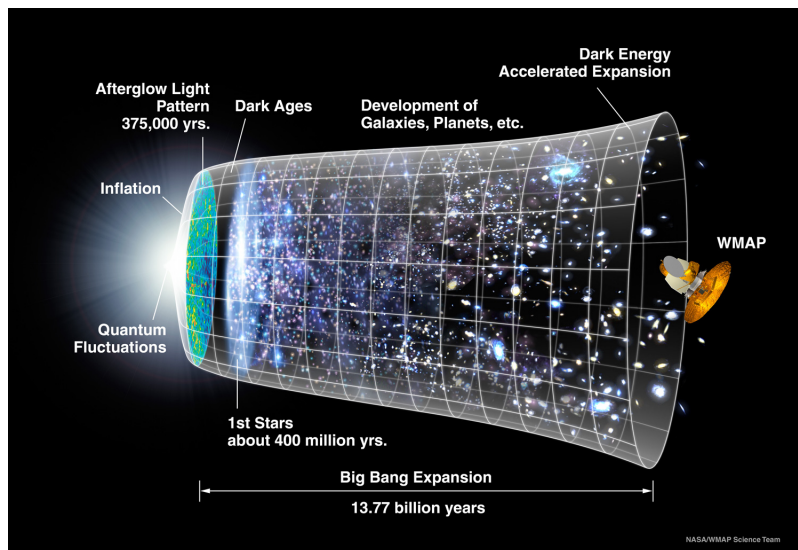


Figure 1.2: Visual depiction of the evolution of the Universe using based on the Λ CDM model and inflation theory.⁴

1.2 SCIENTIFIC APPROACH

This section is taken from Redmond et al. 2022⁹⁶ and Redmond 2018.⁹⁵

1.2.1 DIRECT IMAGING OF EXOPLANETS

Since stars are many orders of magnitude brighter than the planets they host, the lobes of the star's Point Spread Function (PSF) will hide a planet even if the optical system would otherwise be able to resolve the planet. The lobes of the star's PSF must be suppressed in a way that maximizes the throughput of planet light; optical systems that accomplish this are referred to as coronagraphs.⁵¹ There are many types of coronagraphs but most contain one mask in the focal plane to block the main stellar PSF and one mask in the pupil plane to deal with diffraction effects of the focal plane mask, as shown in Fig 1.3. The pupil plane masks can be located before, after, or on either side of the focal plane mask. These masks create a 'dark zone' (DZ) or 'dark hole' (DH) in the focal plane of the coronagraph, near the star where a planet is observable. Though the coronagraph drastically reduces the starlight in the DZ, there will always be residual starlight due to manufacturing defects of the optical components and this residual starlight often has structure. We refer to these structures as 'speckles'^{94,85} and there are many research areas that focus on distinguishing between speckles and planets.^{129,117}

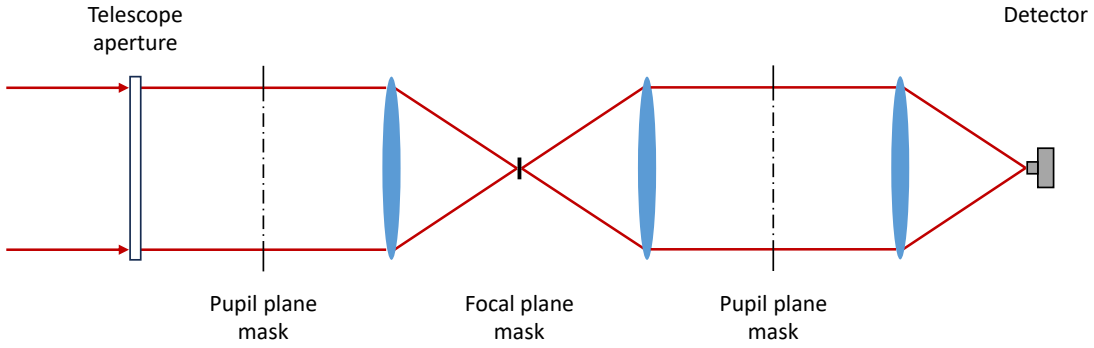


Figure 1.3: Basic high-contrast optical layout.

In order to directly observe a rocky Earth-like planet, the starlight must be suppressed by a factor of 10^{10} .⁴⁶ For this level of suppression in a space-based high-contrast system, deformable mirrors (DMs) are required to correct for the small manufacturing defects and misalignments in the optical system.^{124,121,62} The magnitude of suppression is commonly referred to as the contrast, which we calculate here via:

$$\text{contrast} = \frac{I_{\text{coron}}}{\max(I_{\text{direct}})} \quad (1.2)$$

where I_{coron} is the intensity of the coronagraph image at the planet location and I_{direct} is the intensity of the nominal image of the star without the focal plane mask (FPM) in place.

While there exists more sophisticated and statistically rigorous definitions for contrast in the context of point source detection,⁴¹ here I adopt a simple metric as we are mostly concerned about relative maintenance around a nominal level of starlight extinction. When discussing contrast, a ‘high’ contrast (10^{-10}) is considered good and a ‘low’ contrast (10^{-5}) is considered bad. The characterization of exoplanet atmospheres using direct imaging spectroscopy also requires high-contrast over a wide wavelength range.

Even with the coronagraph and DMs, a large number of exposures are required to accumulate enough photons to characterize dim exoplanets and these long observation times put strict stability requirements on the PSF. Often, one can observe a planet that is dimmer than the static contrast obtained by the instrument, if the image is stabilized at a level that is a small fraction of the raw contrast so that one can calibrate that noise a posteriori. ¹⁴⁵

1.2.2 WEAK LENSING AS A COSMOLOGICAL PROBE ⁹⁵

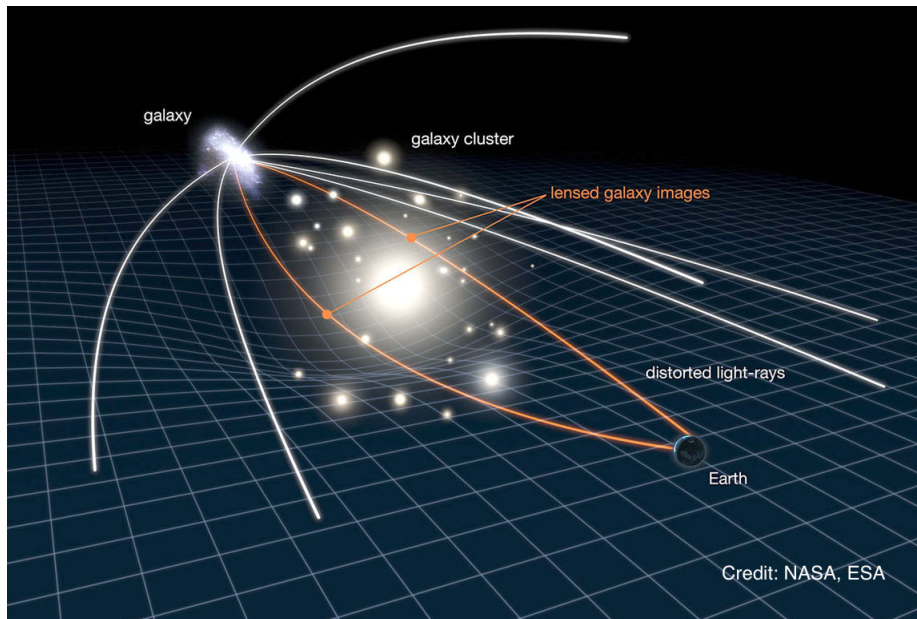


Figure 1.4: Here we see cluster strong lensing shown by the orange lines; the foreground cluster acts as a lens and creates a double image of the background galaxy. ¹⁴

Cluster weak gravitational lensing can be used as a method to constrain the matter density Ω_M described in Eqn. 1.1. This type of weak lensing uses galaxy clusters at different redshifts; a cluster far, far away as the source and a nearby galaxy cluster as the lens. The mass of the near galaxy cluster acts as a gravitational lens and bends the light rays as shown by

the white lines in Fig. 1.4. Strong lensing occurs when the lens cluster is sufficiently massive that it creates multiple images of select source galaxies as shown by the orange lines in Fig. 1.4. Weak lensing is based on the same mechanism but produces a milder distortion of the source galaxies as shown in Fig. 1.5 (illustrative purposes only). The resulting image(s) depends on the not-necessarily-symmetric geometry of the whole system. Strong lensing is rarer due to the mass required to produce it as well as the excellent alignment required between the source (background) and lens (foreground) clusters; note that a perfectly symmetric and aligned system will create a single Einstein ring. Weak lensing is thus a more favourable observational probe due to its abundance.

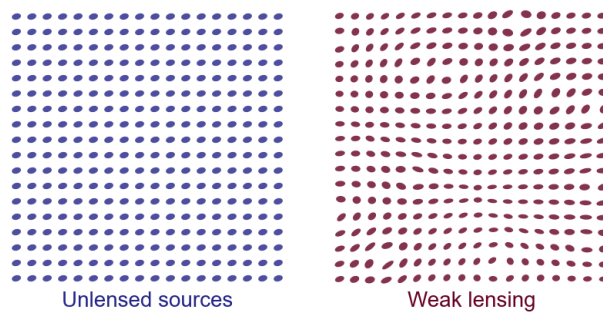


Figure 1.5: Example of the effect of weak lensing.⁵⁹

In reality the distortion induced by weak lensing is on the order of 1%.¹⁴² For this reason any weak lensing instrument must have a wide field of view (FOV) and be capable of deep imaging to capture as many sources as possible. With this in mind, the SuperBIT telescope has been developed to have a 22.5×15.1 arcminute field of view with a throughput above 80% for the optical and near-ultra-violet (NUV) wavelengths. For comparison, Fig. 1.6 shows an image of the Eagle Nebula taken during a SuperBIT engineering flight in 2016 with the Hubble Space Telescope’s Wide Field Camera 1 FOV overlaid in blue.²⁷

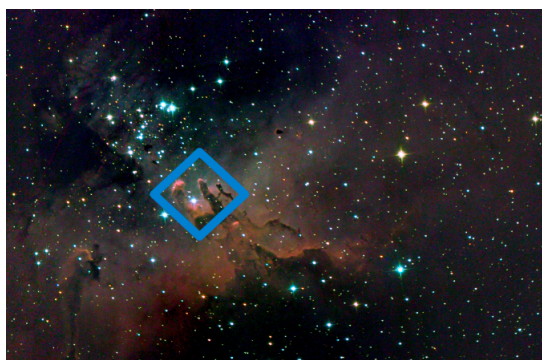


Figure 1.6: False colour SuperBIT image (2016 test flight) of the Eagle Nebula with the Hubble FOV overlaid in blue.³¹

It is important to note that the light rays are affected by both dark matter and baryonic matter, which means that weak lensing does not require any assumptions regarding the relationship between the two. This is important as there are often scenarios where the location of the luminous matter does not coincide with the location of the majority of the mass of a galaxy. Such an example is the Bullet Cluster shown in Fig. 1.7 where the pink highlights luminous matter (mapped using X-rays) and the blue indicates dark matter (mapped using gravitational lensing). Note this image is in false-colour.

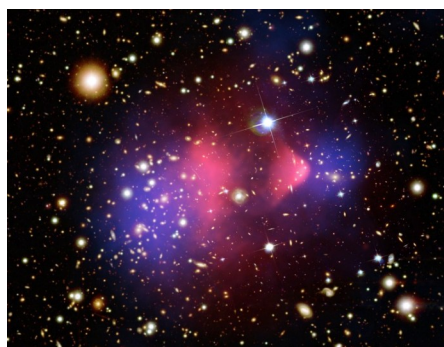


Figure 1.7: Composite image of the Bullet Cluster.⁷⁸ X-ray emission is in pink and dark matter is in blue. The Bullet Cluster was once two clusters as is evident by the two pink and blue clouds. We can see that during the merger, the luminous matter becomes entangled but the dark matter does not self-interact and stays as two distinct entities.

Weak lensing represents a minimally biased method of measuring the total matter as well as the matter distribution within clusters, regardless of their dynamical state. Large samples of unbiased cluster mass measurements are necessary for improving the constraint on cluster mass-observable relationships. These mass-observable relationships are necessary for leveraging large cluster surveys to constrain other cosmological parameters.⁸⁶ In addition, measurements of the dark matter distribution (from weak lensing) in combination with independent tracers of baryonic matter within a sample of galaxy clusters provides an opportunity to constrain interactions within the dark sector and between the dark and light sectors in a model independent way.

1.3 THE FIELD TODAY

1.3.1 SPACE-BASED HIGH-CONTRAST INSTRUMENTS

Before examining high-contrast imaging, it is important to acknowledge the exoplanet detection techniques that have been used to advance the field of exoplanet science for the past three decades. To date, the majority of exoplanets have been detected using techniques such as Radial Velocity¹⁴³ (RV) or transits.^{108,143} Radial velocity uses the slight wobble in the host-star's position as it orbits the star-planet centre of mass to detect the planet and infer its mass and orbital radius. RV is thus biased towards more massive planets with small orbital radii. Transits look for periodic dimming of the host-star as the planet passes in front of it in order to make a detection. Directly imaging exoplanets provides much more information about planetary systems and targets a different class of exoplanet but benefits extensively from RV and transit detection results.

The Nancy Grace Roman Space Telescope (RST) is expected to be the first NASA space

mission to use DMs in space for creating high-contrast dark zones to directly image exoplanets. The RST Coronagraphic Instrument (CGI) is a technology demonstration for future high-contrast space missions. The required contrast for CGI is 10^{-7} but it is predicted to perform at a contrast of 10^{-9} .⁴⁴ CGI has two 48x48 AOA Xnetic DMs and three coronagraph options. Two of the coronagraphs are Shaped Pupil Coronagraphs (SPC) and the third is a Hybrid Lyot Coronagraph (HLC); these will be further discussed in Ch. 2. Ongoing thermo-mechanical modelling, as part of the mission development cycle, suggests that the image quality necessary to detect temperate giant exoplanets can be achieved by

1. pointing at a bright reference star to dig the dark zone, i.e., find the DM command that corrects the static wavefront error (WFE) induced by the optical surfaces etc.,
2. performing a slew to point at the science target at which point the DM command is kept constant (Krist et al. 2020⁴⁹)
3. rolling the telescope four times while on the science target
4. slewing back to the reference star to re-dig the dark zone that has degraded due to the various drifts in the system.

This reference star–target star–reference star cycle would repeat three times for each target if RST is operating in HLC mode. The telescope rolls are to enable the use of angular differential imaging^{111,29,63} (ADI) in post-processing. As explained in Sec. 2.4, the residual starlight speckles remain in the same position on the focal plane when the telescope is rolled but the planet moves.

The 2020 decadal survey recommended exoplanet science as a primary objective for the Large Strategic Mission “A Future IR/Optical/UV Telescope Optimized for Observing

Habitable Exoplanets and General Astrophysics”;⁷³ this mission is now referred to as the Habitable Worlds Observatory (HWO).³⁹ HWO will build on the JWST segmented aperture technology as well as the RST coronagraph technology. The instrument will orbit L2 and be designed to accommodate robotic service missions. The target contrast is 10^{-10} with the goal to be able to look for biosignatures in the atmospheres of potentially habitable exoplanets. The mission concept study of the Large UV/Optical/IR Surveyor (LUVOIR)⁹³ will be extensively referenced in the HWO design.

1.3.2 BALLOON-BORNE TELESCOPES

On Earth, optical telescope capabilities and image quality are limited by atmospheric seeing. Moving out into the far-UV or far-IR, the atmosphere is mainly opaque reducing the science goals achievable by ground-based telescopes. For these reasons, space telescopes were developed to improve data quality and probe wavelengths which are out of reach for ground-based instruments. Unfortunately, space telescopes are very costly and have very long lead times.⁸² Balloon-borne telescopes offer a unique compromise in which data quality losses and costs are significantly reduced. They fly at altitudes ranging from 30 to 45 km in the stratosphere which provides a space-like atmospheric environment but maintains gravity loading as seen by ground-based instruments. Figure 1.8 shows the transmission gain for balloon instruments between Near-Ultraviolet (NUV) and Near-Infrared (NIR) wavelengths when compared to ground-based instruments either at sea level or the Hawaii Mauna Kea site. At the end of the flight, the payload returns to the ground via a parachute. Often the majority of the instrument can be re-used for the next flight. This enables closed-loop engineering of the components as well as relatively short turnaround times between

upgrades.

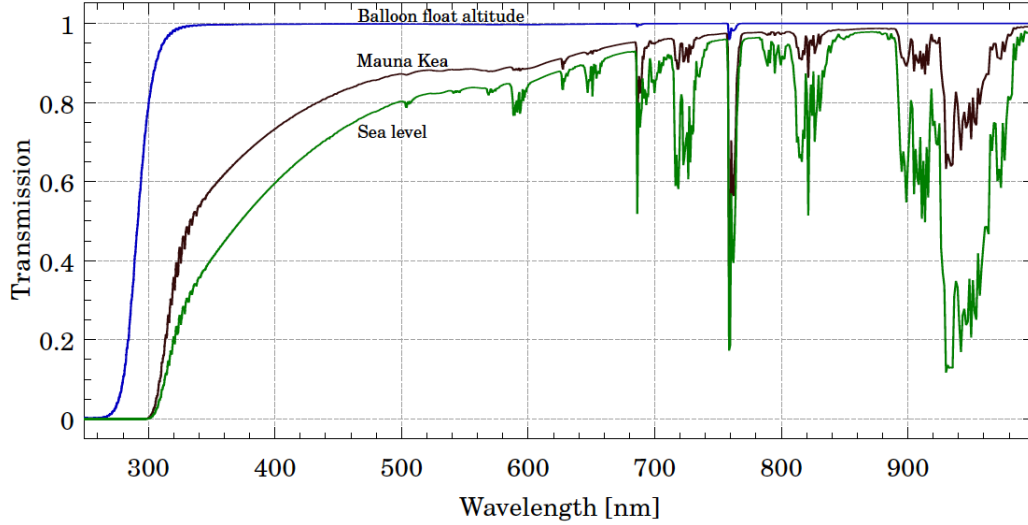


Figure 1.8: MODTRAN4 transmission calculation altitude comparison for the Near-Infrared (NIR, 900nm) and Near-Ultraviolet (NUV, 300 nm).⁷⁹

There are a number of launch sites for balloon-borne instruments depending on the science goals and wavelength range. For longer wavelength instruments such as SPIDER-II or the Galactic / Extragalactic ULDB Spectroscopic Terahertz Observatory (GUSTO), the preferred launch location is the Long Duration Balloon Base near the McMurdo station in Antarctica. Launching during the austral summer from McMurdo provides continuous solar power and the stratospheric winds generally keep the payload over land which facilitates recovery of the payload when the flight terminates. For telescopes targeting shorter wavelengths, such as the Planetary Imaging Concept Testbed Using a Recoverable Experiment-Coronagraph (PICTURE-C) or the Super-pressure Balloon-borne Imaging Telescope (SuperBIT), the preferred launch locations are Fort Sumner, NM or Wanaka, NZ. Fort Sumner only provides one–two night flights due to the flight path and the fact that they

use zero-pressure balloons. When launching from Wanaka, NZ, a Super Pressure Balloon (SPB) is used. This is a novel concept that has been in development by NASA over the last two decades. SPB are completely sealed and maintain a much more consistent volume during day-night cycles which minimizes the altitude variations. The current Super Pressure Balloon (SPB) flight duration record is 54 days²⁶ but that is expected to increase as developments continue. As with satellites, mass and power consumption have strict limits which results in unique design challenges for the optical design.

SuperBIT is the current state-of-the-art diffraction-limited near-infrared–near-UV wide-field balloon-borne observatory. It has completed four one-night test flights and recently completed a 49 night science flight in the spring of 2023. During the science flight it operated semi-autonomously and demonstrated 50 milliarcsecond (mas) pointing stability over 30 minute periods. This both demonstrated the benefit of turning SuperBIT into a standalone observatory as well as complete its goal as a pathfinder instrument for the Gigapixel Balloon-borne Imaging Telescope (GigaBIT). GigaBIT will be a 1.3 m version of SuperBIT and aims to launch in the late 2020s. An important note for both SuperBIT and GigaBIT is that, similar to ground-based telescopes, science targets are only available at certain times as they rise and set. Target availability also depends on the current longitude of the balloon as the payload circumvents the southern hemisphere. This means that a target might be imaged in certain wavelengths on one night and other wavelengths on another night.

Figure 1.9 shows the predicted GigaBIT performance for a 1 hr exposure as a function of wavelength. There are two other wide-field imaging instruments shown in Fig. 1.9 that compliment the science goals of SuperBIT and GigaBIT. As shown in Fig. 1.9, none of the current or upcoming instruments have NUV capabilities. Also, space telescopes are

well known for being over-subscribed with only 17.7% of HST Cycle 27 proposals being awarded. The existence of a balloon platform with similar capabilities to space telescope that can upgrade technology on a yearly basis is very advantageous to the scientific community.

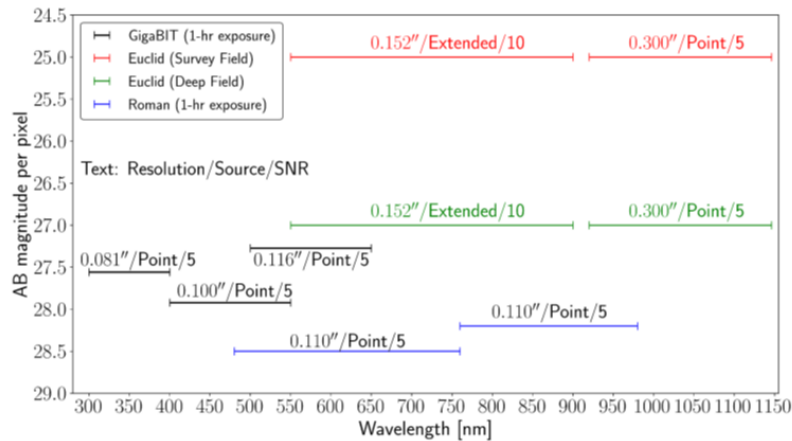


Figure 1.9: Comparison of the GigaBIT imaging resolution and sensitivity with current and upcoming space telescopes.⁸⁰ This figure is adapted from Fig. 2-10 in Spergel et al. 2015.¹³¹

1.4 QUASI-STATIC WAVEFRONT ERROR DRIFTS

As previously stated in Sec. 1.2.1, long integration times are required for directly observing exoplanets. Currently, multiple solutions are available to correct high-spatial frequency, static wavefront errors (WFE)¹³³ as well as low spatial frequency, fast WFE.^{118,119} Slow or quasi-static instabilities in the system also degrade the contrast over an observation and can produce speckles that can be falsely identified as planets.¹³⁰

HWO, envisioned as a segmented aperture telescope, presents specific stability challenges beyond what will be demonstrated with RST.^{70,93,22} Since each segment has its own mounting structure, thermal effects introduce modes with higher spatial frequencies than

when compared to traditional monolithic primary mirrors.^{110,57,54} In addition, each segment has both controllable and un-controllable degrees of freedom (DOF). The former can be adjusted using actuators mounted on a hexapod structure akin to the one used for the James Webb Space Telescope²⁰ (JWST). The actuators must be repeatable and stable in order to avoid introducing WFE into the optical system. Non-actuated segment edges also need to remain in place at levels commensurate with picometers of wavefront at the entrance of the coronagraph.

In order to perform accurate weak lensing measurements, the wavefront for SuperBIT and GigaBIT must be stable on timescales of an hour and they must be repeatable. As mentioned in Sec. 1.3.2, the instrument might observe a target at different times after sunset and telescope elevations on different days. In order to be able to combine the data from different observations, the systematics due to thermal variations and gravity sag must be minimal. Due to the wide FOV, this is a slightly different controls problem than for high-contrast imaging. For SuperBIT, since it only has a 0.5 m primary mirror and more mass available to stiffen the telescope, these slow thermal and gravity gradients can be corrected for by using a temperature controlled baffle and passive structural components. For GigaBIT, this will likely not be sufficient and active optics will be required either in the form of a back-end deformable mirror or a deformable primary mirror.

1.5 THESIS OVERVIEW

This thesis focuses on developing solutions to correct for quasi-static drifts in high-contrast and wide-field imaging applications. For high-contrast imaging, two focal plane wavefront estimation and control algorithms are presented which provide alternate approaches to (1)

generating the dark zone and (2) maintaining it in the presence of different drifts. Experimental results are provided which use the High Contrast Imager for Complex Aperture Telescopes (HiCAT) testbed at the Space Telescope Science Institute (STScI); an overview of the HiCAT testbed is provided in App. A. Post-processing techniques are investigated to ensure that the algorithm actually facilitates detection of exoplanets and does not just maintain the dark zone contrast. For wide-field imaging, the results from the 2023 SuperBIT flight are used to quantify the effects of the thermal and gravity related wavefront error on the PSF. These results are then used to inform the active optics design for GigaBIT. The breakdown for this thesis is as follows:

Chapter 2: An overview of the concepts used throughout this thesis including theoretical foundations and important instrument information.

Chapter 3: Description and hardware demonstration of a DZ generating focal plane wavefront estimator that uses broadband images to estimate the electric field at select wavelengths within the band from Redmond et al. 2021.⁹⁷

Chapter 4: Derivation and hardware demonstration of a Dark Zone Maintenance (DZM) algorithm that corrects for high-order drifts. Hardware results are using the High-contrast Imager for Complex Aperture Telescopes (HiCAT) testbed and the In-Air Coronagraph Testbed (IACT). This chapter is taken from Redmond et al. 2022.⁹⁶

Chapter 5: Adjustment of the DZM algorithm to replicate the low SNR environment predicted for space-based coronagraphic instruments as well as hardware demonstrations; taken from Redmond et al. 2022.⁹⁹

Chapter 6: Direct-imaging post-processing techniques and how they work with the DZM algorithm.

Chapter 7: Overview of SuperBIT focal plane wavefront stability using the 2023 flight results. These results are then used to inform design decisions for GigaBIT.

2

Background

In this chapter I discuss all of the building blocks required for this research. This includes the theory of the optical systems, how high-contrast systems work, and the critical hardware components.

2.1 FOURIER AND FRESNEL OPTICS

Fourier optics is an elegant approach for modeling the far-field effects of optical components on a given electric field. Additional material on this derivation can be found in Hecht 2017³⁸ and Goodman 2005.³⁵ The Huygens-Fresnel principle states that the wavefront of an electrical field at any point is the superposition of secondary spherical emitters along the primary wave. We can use the Rayleigh-Sommerfeld integral to describe the superposition. I define the propagation planes in Fig. 2.1.

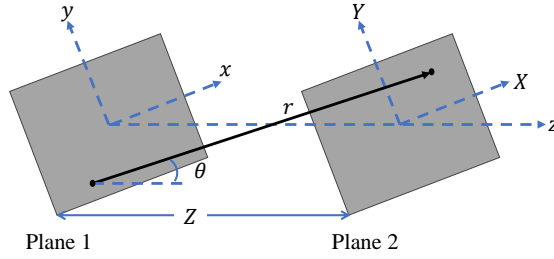


Figure 2.1: Propagation planes for Fourier optics derivation.

If we take $z = 0$ to be the location of the emission (Plane 1 in Fig. 2.1) then we can express the cumulative result at some distance away Z as

$$E(X, Y, Z) = \frac{1}{i\lambda} \iint_{-\infty}^{\infty} \varepsilon(x, y, 0) \frac{e^{ikr}}{r} \cos \theta dx dy \quad (2.1)$$

where $k = 2\pi/\lambda$ and $r = \sqrt{(X-x)^2 + (Y-y)^2 + Z^2}$. If we then let

$$\rho^2 = (X-x)^2 + (Y-y)^2 \quad (2.2)$$

we get

$$r = z\sqrt{1 + \frac{\rho^2}{Z^2}}. \quad (2.3)$$

We assume $Z \gg \rho$ (paraxial approximation) and let $r \approx Z$ for the amplitude $\frac{1}{r}$ term.

For the r in the phase term (e^{ikr}) we need to be more rigorous as k can be quite large and thus the error in the r approximation must be much smaller for the phase term than for the amplitude term. For the phase term, we instead Taylor expand r to get

$$r = Z \left(1 + \frac{\rho^2}{2z^2} - \frac{1\rho^4}{8z^4} \right). \quad (2.4)$$

Plugging r into the e^{ikr} term from Eqn. 2.1 we get

$$e^{ikr} = e^{ikZ} e^{0.5ik\rho^2/Z} e^{-ik\rho^4/(8z^3)}. \quad (2.5)$$

Looking at the last term we can see that we can neglect it if $\frac{k\rho^4}{z^3} \ll 2\pi$; since we took $Z \gg \rho$ we will assume that to be true. Note that an additional reduction from the $Z \gg \rho$ assumption is that $\cos \theta \approx Z/\rho \approx 1$.

Plugging the above simplifications into Eqn. 2.1 and rearranging we get

$$E(X, Y, Z) = \frac{e^{ikZ}}{i\lambda Z} \iint_{-\infty}^{\infty} \varepsilon(x, y, 0) e^{0.5ik\rho^2/Z} dx dy \quad (2.6)$$

From here, we sub Eqn 2.2 back in and expand the integrand to get

$$e^{\frac{ikp^2}{Z}} = e^{\frac{ik(X^2+Y^2)}{2Z}} e^{\frac{ik(x^2+y^2)}{2Z}} e^{-\frac{ik(xX+yY)}{Z}}. \quad (2.7)$$

We can plug this in and rearrange to get

$$E(X, Y, Z) = \frac{e^{ikZ} e^{\frac{ik(X^2+Y^2)}{2Z}}}{i\lambda Z} \iint_{-\infty}^{\infty} \varepsilon(x, y) e^{\frac{ik(x^2+y^2)}{2Z}} e^{-\frac{ik(xX+yY)}{Z}} dx dy \quad (2.8)$$

which is the Fresnel Equation. Note that this can be expressed in terms of a Fourier transform

$$E(X, Y, Z) = \frac{e^{ikZ}}{i\lambda Z} e^{\frac{ik(X^2+Y^2)}{2Z}} \mathcal{F}\{\varepsilon(x, y) e^{\frac{ik(x^2+y^2)}{2Z}}\} \left(\frac{X}{\lambda Z}, \frac{Y}{\lambda Z}\right). \quad (2.9)$$

We can take this one step further and look at the Fresnel Equation in the case that $\frac{Z}{\lambda} > \frac{x^2+y^2}{\lambda^2}$ where $x^2 + y^2$ can be thought of as the aperture in the case that Plane 1 in Fig. 2.1 is partially opaque. Using this assumption, we can drop the red term in Eqn 2.9 as it will be approximately unity which yields the Fraunhofer diffraction equation:

$$E(X, Y, Z) = \frac{e^{ikZ}}{i\lambda Z} e^{\frac{ik(X^2+Y^2)}{2Z}} \mathcal{F}\{\varepsilon(x, y)\} \left(\frac{X}{\lambda Z}, \frac{Y}{\lambda Z}\right). \quad (2.10)$$

From the Fraunhofer diffraction equation, we can see that the blue term is the equation of a plane wave and the purple term is an additional quadratic phase factor on top of the Fourier transform. For space-based astronomy, we assume that the incoming electric field is a plane wave so we can drop the blue term. As for the purple term, it is just a phase shift so

for our application it can also be neglected. From here we can see that in the ‘far field’ the electric field is the Fourier transform of the incoming field,

$$E(X, Y, Z) \propto \mathcal{F}\{\varepsilon(x, y)\}_{\left(\frac{x}{\lambda Z}, \frac{y}{\lambda Z}\right)}. \quad (2.11)$$

We will now take the above result and express it in a form that represents an optical system. As described in Hect 2017,³⁸ a lens transforms incoming light to the far-field diffraction pattern expressed by Eqn. 2.11. We also introduce the definition of an aperture function

$$\mathcal{A}(x, y) = \mathcal{A}_0(x, y)e^{i\varphi(x, y)} \quad (2.12)$$

which defines the apodization of the aperture and assumes a plane wave. Equation 2.11 then becomes

$$E(X, Y) = \mathcal{F}\{\mathcal{A}(x, y)\} \quad (2.13)$$

showing that the electric field at the focal plane of a telescope is the Fourier transform ($\mathcal{F}\{\}$) of the aperture using the aforementioned assumptions. Note that the inverse of Eqn. 2.13 is also true where the electric field in the pupil plane is the inverse Fourier transform of the electric field in the far-field (focal plane).

Fourier optics is a clean method for propagating the electric field from the aperture of a telescope to the detector plane that captures diffraction effects. We now apply this to the setup for a simple coronagraph with a focal plane mask (\mathcal{M}_{FPM}) and a pupil plane mask

(M_{PPM}) as shown in Fig. 2.2. In blue we have lenses that focus and collimate the light in the appropriate way for each of the coronagraph components. The focal length³⁸ of the lenses is denoted as f ; we can see that the focal plane mask (FPM) is located at the focus of the lens and the pupil plane mask is located at a conjugate pupil plane to the aperture (a distance f from the lens). The masks are denoted as M_i which is equivalent to \mathcal{A}_0 in Eqn. 2.12. We see that the electric field at the science camera (E_f) is defined as

$$E_f(X, Y) = \mathcal{F} \{ M_{PPM} \mathcal{F}^{-1} \{ M_{FPM} \mathcal{F} \{ M_A E_p \} \} \} \quad (2.14)$$

where E_p is the incoming plane wave and \mathcal{F}^{-1} is the inverse Fourier transform. The masks in Eqn. 2.14 are further discussed in Sec. 2.2.

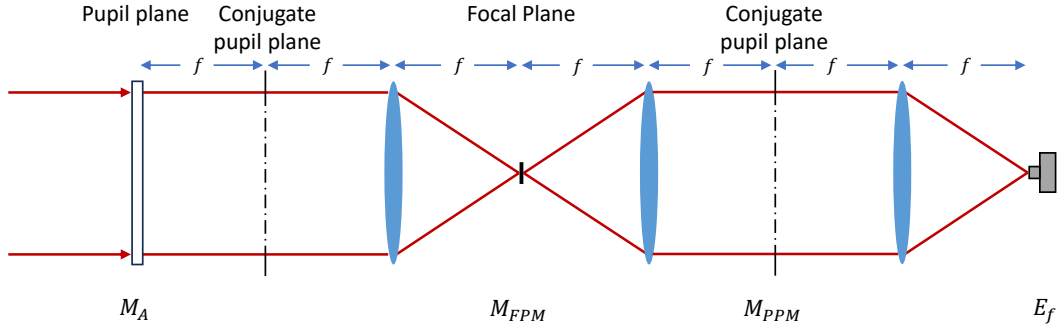


Figure 2.2: Simplified coronagraph optical layout assuming flat DMs. M_A is the aperture, M_{FPM} is the focal plane mask, M_{PPM} is the pupil plane mask, and E_f is the electric field at the focal plane. The focal length of the lenses is denoted as f and the PPM stop is in a conjugate pupil plane to the aperture.

2.2 CORONAGRAPHS

There are a number of different coronagraphs in development today. I will focus on those that will be flown on RST¹⁰¹ and that are used on HiCAT^{74,127} (App. A). The two main

architectures are a Shaped Pupil Coronagraph (SPC)⁴⁵ and a Lyot coronagraph.^{61,128} A SPC first modifies the shape of the PSF in the focal plane by changing the shape of the entrance pupil using a pupil plane mask; in the example diagram provided in Fig. 2.2 this pupil plane mask goes at the first conjugate pupil plane after the aperture (between M_A and the first lens). The pupil plane mask generates a high contrast zone and then the main stellar PSF is rejected via a focal plane mask (M_{FPM} in Fig. 2.2) to reduce the required dynamic range of the detector. The SPC can often generate a higher contrast region⁴⁵ and is robust with respect to observatory jitter¹⁰¹ but those used on RST are not azimuthally symmetric. In order to directly image an exoplanet with a SPC on RST, either the planet location needs to be known a priori or the telescope must perform multiple rolls to locate it. Figure 2.3 shows the aperture (left panel) and one of the Shaped Pupil pupil plane masks for RST (middle panel); this figure is taken from Cady et al. 2015.¹⁷ The aperture (left panel) is circular with obscurations due to the secondary mirror and support struts. The dark zone consists of two wedge-shaped regions separated by 180° as shown in the right panel. Note that the focal plane mask is not provided in this figure to demonstrate the PSF created by the shaped pupil. RST will fly multiple SPCs that create slightly different dark zones to target different science cases.

Unlike a SPC, a Lyot coronagraph first rejects the main stellar PSF in the focal plane and then re-collimates the beam and passes it through a ‘Lyot stop’; a simple optical layout for a Lyot coronagraph is shown in Fig. 2.2. The Lyot stop reduces the diffraction effects induced by the the FPM and a high contrast region is generated in the focal plane. Note that a Lyot stop can be paired with many coronagraphs and the SPC on RST will use a Lyot stop with each of the two SPC masks. In Fig. 2.4 I show the Lyot coronagraph masks used

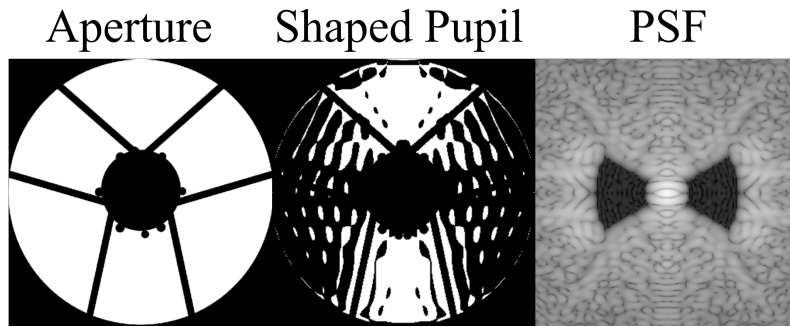


Figure 2.3: RST aperture (left panel) and one of the shaped pupil masks (middle panel). The resulting PSF due to the combination of these masks is shown in the right panel. A wedge-shaped dark zone is created on either side of the star. These figures are taken from Cady et al. 2015.¹⁷

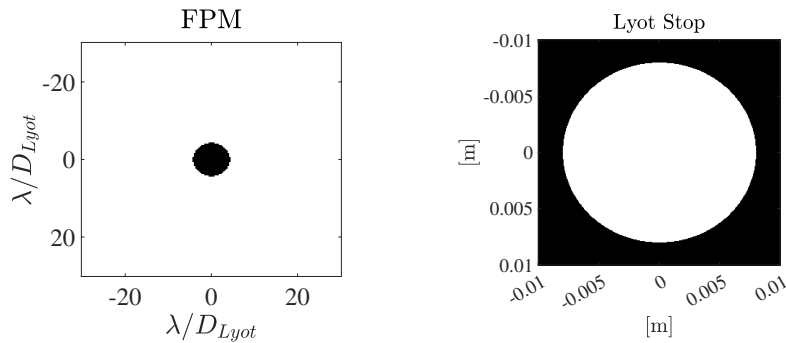


Figure 2.4: HiCAT Lyot coronagraph masks. In the left panel is the FPM; in reality HiCAT uses a reflective FPM that has a hole in the middle that allows the main lobe of the stellar PSF to pass through and reflects the rest of the light. In the right panel is the Lyot stop which is located in the pupil plane (M_{PPM} in Fig. 2.2).

on HiCAT. Note that the FPM dimensions are λ/D_{Lyot} where λ is the wavelength used and D_{Lyot} is the diameter of the Lyot stop. For a space telescope, D is generally defined as the diameter of the aperture of the telescope.

In Fig. 2.5 I show the effect of each component of the HiCAT high-contrast system. All figures are in log contrast units and have the same colour limits. In the left panel we see the un-occulted PSF with no masks in place other than a circular aperture; note that the peak contrast is one which is expected based on Eqn. 1.2. The second panel shows what

the PSF looks like when the FPM is in place. We get a small dark zone where the main lobe of the star used to be but there are a lot of bright speckles around it. In the third panel, both coronagraph components are in place and the target dark zone is shown in black. The Lyot stop pushes the speckles outside of the region of interest and reduces the contrast in the dark zone to $\sim 10^{-5}$. The third panel is a visual representation of Eqn. 2.14. In the rightmost panel, the DMs are activated and set to a shape to reduce the residual light in the dark zone; determining the optimal DM shape is discussed in Sec. 2.3.3. The DMs allow for a dark zone contrast of $\sim 10^{-8}$ on HiCAT.

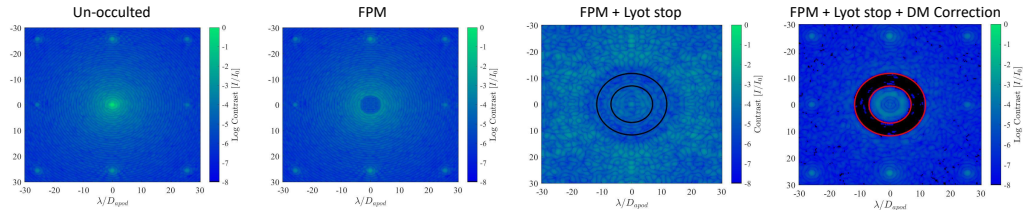


Figure 2.5: Evolution of the HiCAT PSF as components of the coronagraph are added. In the left panel we have the un-occulted or ‘direct’ PSF. The peak contrast in the direct image is one by the definition of contrast provided in Eqn. 1.2. Next we put the FPM in place and we get a small region with a higher contrast. When we install the FPM and the Lyot stop (third panel) we get a larger high-contrast region with a mean dark zone contrast on the order of 10^{-5} – 10^{-4} ; note a sample dark zone is outlined by the black annulus. In the far right panel, a correction is applied to the DMs to suppress the residual starlight in the dark zone and we obtain a mean dark zone contrast on the order of 10^{-8} .

Two important parameters of any coronagraph are the inner working angle (IWA) and outer working angle (OWA). These parameters are the angular separation between the centre of the stellar PSF and extremes of the largest possible dark zone. We often do not try to generate a dark zone that hits both extreme limits of the coronagraph as it results in a overall degradation of the dark zone. In Fig 2.5 the dark zone has an IWA of $6 \lambda/D_{Lyot}$ and an OWA of $11 \lambda/D_{Lyot}$; the DZ IWA is much larger than the minimum IWA set by the coronagraph. The IWA and OWA set the exoplanet orbital radii for which we can still directly image the exoplanet. If the angular separation between the star and the planet is not in

the range of the working angles of the coronagraph, it will be outside the dark zone and we will not be able to image it. The IWA of a coronagraph is often the topic of discussion as exo-Earths are predicted to have a small orbital radius around their host stars so future space-based high-contrast instruments will need a coronagraph with a small IWA to directly image an exo-Earth.³³

2.3 WAVEFRONT SENSING AND CONTROL

2.3.1 ACTIVE OPTICS AND WAVEFRONT SENSING

Many ground-based high-contrast imagers use one detector for science and another detector for determining the WFE they must correct. When using a separate detector to determine the WFE, the resulting WFE measurement contains non-common path aberrations (NCPA). Since the light hits different optical surfaces en route to the WFE detector vs the science detector, the aberrations present in the images will be slightly different. At lower contrasts or for very large WFE, NCPA do not limit the instrument performance as they are generally very small. Space-based high contrast imagers, however, are very sensitive to NCPA and must use Focal Plane Wavefront Sensing and Control (FPWSC) to correct for the quasi-static WFE. Note that this assumes there are DMs in the back-end optics, which is not the case for the Hubble Space Telescope or the James Webb Space Telescope, each of which depends on modelling and post-processing to remove drift effects.^{24,84}

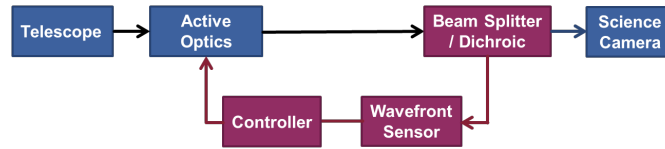


Figure 2.6: Schematic for a typical pupil plane wavefront estimation and control loop. The science instrument is shown in blue and the wavefront control system is shown in burgundy. A beam splitter or dichroic is used to send a portion of the light to the wavefront sensor. Due to the different optics between the DMs (active optics) and the science camera vs the wavefront sensor, NCPA errors are present in the estimate of the electric field at the science camera.

2.3.2 DEFORMABLE MIRRORS

In high-contrast imaging, the choice of which deformable mirror is used has a large impact on the achievable dark zone size, contrast, and stability. The maximum size of the dark zone is determined by a combination of the number of deformable mirrors and the number of actuators per deformable mirror. We can loosely think of the maximum dark zone size in terms of the concept of controllability from modern controls theory; the more actuators available, the more pixels on the detector we can control. From a flight mission perspective, we also want a high actuator density to minimize the size, and thus mass, of the DMs and other optical components.

The DMs also have an impact on the achievable contrast level. When operating at high contrasts, the change in the DM correction iteration-to-iteration is very small. This means that the better the resolution of the DM actuators, the higher the achievable contrast. In order to directly image an exoplanet, the dark zone must be stable for hours. Some DMs have the resolution to generate a high contrast but then degrade over time due to thermal sensitivities or electro-mechanical instabilities. This characteristic is difficult to measure for a DM as instabilities at the picometer level can affect the dark zone contrast.

There are currently two main types of deformable mirrors (DMs): (1) continuous facesheet

and (2) segmented, as shown in Fig. 2.7. Continuous facesheet DMs are used for wavefront error correction and segmented DMs are used in laboratories as a proxy for a segmented primary mirror on an on-sky telescope.

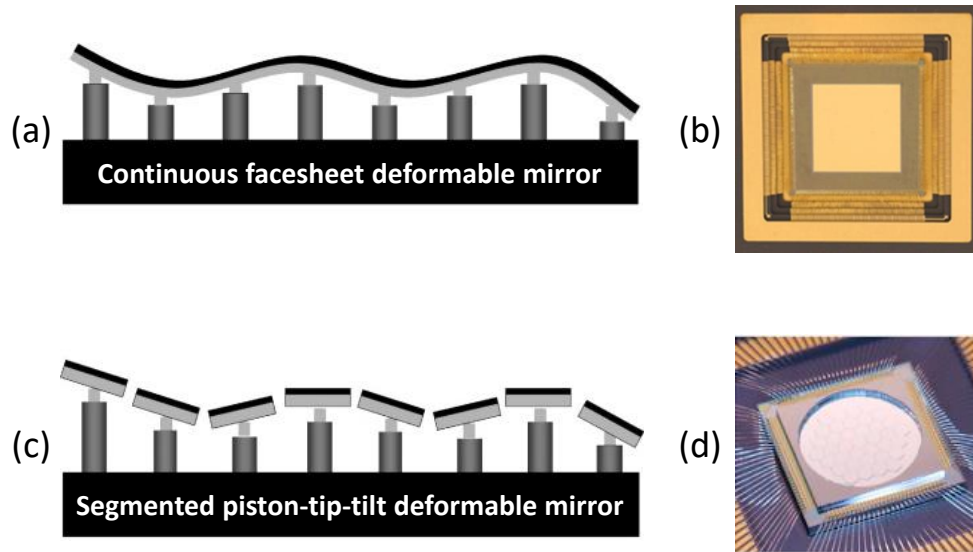


Figure 2.7: Types and examples of deformable mirrors. Panels (a) and (c)²⁵ show section views for a continuous facesheet DM and segmented DM with piston-tip-tilt control respectively. Panel (b) is the BMC continuous facesheet DM used on HiCAT. Panel (d) is the IrisAO PT111 segmented DM that is used on HiCAT; note the segments are hexagons like other on-sky segmented aperture telescopes.

Continuous facesheet DMs have a flexible reflective top surface with piston actuators that deform the facesheet as shown in Fig. 2.7a.²⁵ There are two main actuator architectures that have been developed for space-based high-contrast imaging which have different advantages. Boston Micromachines Corporation (BMC) produces high-density DMs that use Microelectromechanical systems, or MEMS-based, actuators as shown in Fig. 2.7b. BMC DMs are used on the Princeton, HiCAT, and IACT testbeds. AOA Xinetics manufactures DMs that use lead magnesium niobate (PMN) electroceramic actuators; this is the type of DM that RST CGI will fly. One downside to the AOA Xinetec DMs is that

they have been shown to drift at the 10^{-10} contrast level when attempting to hold a particular shape for multiple hours.¹¹⁴ This drift is not expected to negatively affect RST as the operating contrast for RST is 10^{-8} – 10^{-9} but will be something to consider for the HWO mission. The PMN strain is also known to change with temperature so the RST DMs have active thermal control to 10 mK.⁸⁷ Relevant properties of the BMC DMs used on HiCAT and the AOA Xintec DMs to fly on RST are provided in Tab. 2.1.

The Iris AO PT111 segmented deformable mirror installed on HiCAT is shown in Fig. 2.7d. It has 37 flat hexagonal segments (each 1.212 mm flat-to-flat) separated by a $10\ \mu\text{m}$ segment gap that can be controlled in piston, tip, and tilt. The Iris AO DM is used on HiCAT as a proxy for a segmented primary mirror to recreate the effects of the segment gaps and inter-segment wavefront errors that are seen on segmented telescopes. Relevant properties of the Iris AO PT111 are provided in Tab. 2.1.

Table 2.1: DM properties for DMs used throughout this thesis. The DM resolution is limited by the controller; a 16-bit controller is available for the BMC DMs which would improve the resolution to 22 pm. Note the Iris AO number of actuators is listed as 3×37 as there are 37 segments and each segment has three actuators to allow for piston, tip, and tilt control.

Manufacturer	Inscribed diameter [mm]	Number of Actuators	Actuator stroke [μm]	Controller resolution [bits]	DM resolution [pm]
BMC	9.9	1,156	1.5	14	91
AOA Xintec	48.0	2,304	0.5	16	7.5
Iris AO	7.0	3×37	5	14	305

2.3.3 GENERATING A HIGH-CONTRAST DARK ZONE

Figure 2.8 shows a simplified version of the optical layout for the High Contrast Imager for Complex Aperture Telescopes (HiCAT) at the Space Telescope Science Institute. A detailed overview of HiCAT is provided in App. A. HiCAT consists of a Lyot coronagraph (discussed in Sec. 2.2) and two DMs between the aperture and the focal plane mask. If there is only one DM in the optical chain, the dark zone must be ‘one sided’ in order to still correct for the phase and amplitude errors.¹⁴ With two DMs we can correct phase and amplitude for a symmetric dark zone. As shown in Fig. 2.8, DM1 is located at a conjugate pupil plane to the aperture which enables it to correct for the amplitude errors. DM2 is some distance d downstream from DM1 where there is mixing between the phase and amplitude errors; this enables correction of the phase error.^{116,92} Note that in a high contrast system, the DMs are located between the aperture and the first mask of the coronagraph.

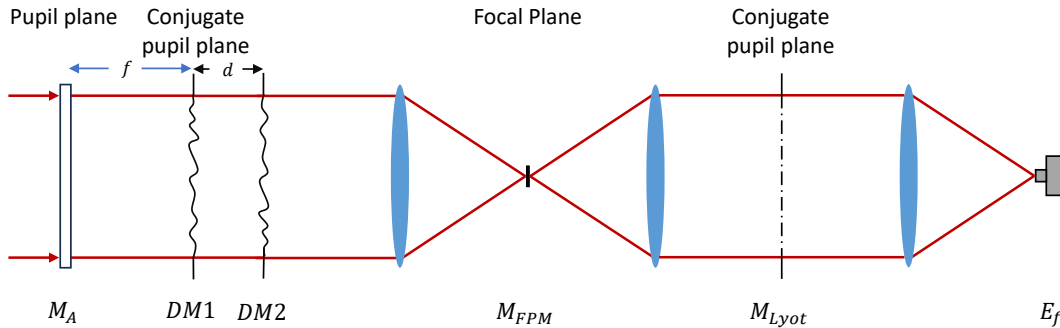


Figure 2.8: Simple high-contrast optical layout with DMs. Note that DM1 is located at a conjugate pupil plane to the aperture (M_A) and DM2 is not. Putting DM2 outside a conjugate pupil plane leverages the Talbot effect¹³⁷ and allows correction of both the phase and amplitude errors of the electric field.¹¹⁶ Both DMs are located before the coronagraph.

Figure 2.9 shows the flowchart for a typical focal plane wavefront estimation and control system; note that k denotes the iteration of the control loop. The optical system is outlined

in blue and the wavefront estimation and control system is outlined in red. The light passes through the telescope optics which introduce small static aberrations due to surface imperfections and misalignments; this produces an aberrated electric field E_{ab} . The intensity (I^k) of the closed-loop electric field (E_f^k) is detected at the science camera after interacting with the deformable mirrors and the set of masks that make up the coronagraph. The measured closed-loop intensity I^k is used to estimate of the electric field at the camera (focal plane), \hat{E}^k . The estimate is used to find the change in the DM command Δu_{opt}^k that will suppress the electric field in the dark zone region. Probe commands (δu_{probe}^k) are applied to the DMs when acquiring the images for the estimator but removed for science images. The nature of the probe commands is dependent on the estimator used.

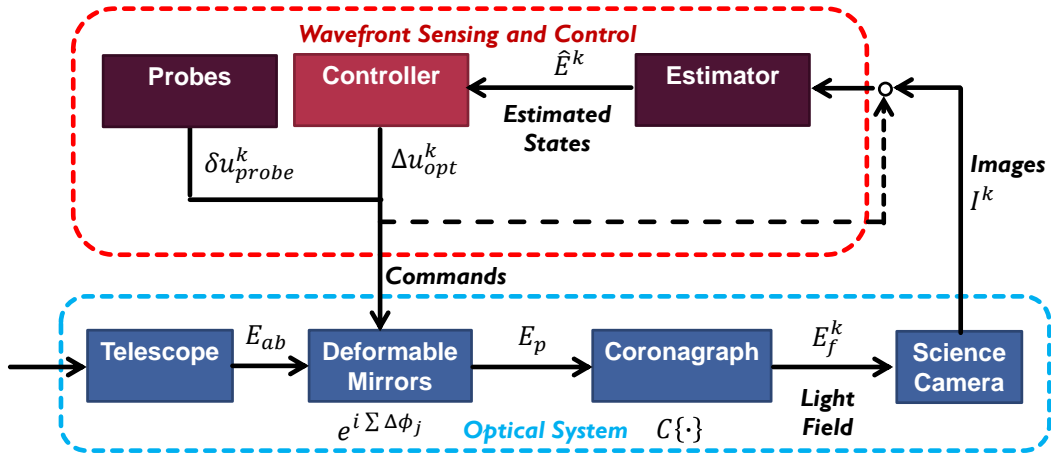


Figure 2.9: Typical focal plane wavefront estimation and control system. The optical system is outlined in blue and the wavefront sensing and control system is outlined in red. Note that $e^{i \sum \Delta \phi_j}$ is the effect of the DMs on the electric field and $C\{\cdot\}$ is the coronagraph operator as described in Sun 2019.¹³³ The light passes through the optical system and the images are passed to the estimator. The estimator produces an estimate of the electric field at the focal plane which is passed to the controller. The controller generates a DM command to suppress the electric field at the focal plane which is then applied to the DMs.

There are multiple focal plane wavefront estimators and controllers in use today that

generate a high-contrast region. A summary of the common methods is provided in Groff et al. 2015.³⁷ I will summarize the pair-wise probe (PWP) estimator as well as the electric field conjugation (EFC) and stroke minimization (SM) controllers. All of these techniques depend on having a model of the optical system including a Jacobian that describes the effect of each DM actuator on the electric field each pixel in the dark zone. The Jacobian is calculated as

$$\mathbf{G} = \begin{bmatrix} \Re \left(\frac{\partial E_0}{\partial u} \right) \\ \Im \left(\frac{\partial E_0}{\partial u} \right) \\ \vdots \\ \Re \left(\frac{\partial E_n}{\partial u} \right) \\ \Im \left(\frac{\partial E_n}{\partial u} \right) \end{bmatrix} \quad (2.15)$$

where E_i is the electric field for pixel i and u is the DM command. Note that the Jacobian is a $2n \times m$ matrix where n is the number of pixels in the dark zone and m is the total number of DM actuators. The factor of two is due to the fact that we split the real and imaginary components of the Jacobian to deal with a purely real matrix.

The PWP estimator operates under the assumption that we are in the linear regime of the DMs and the probe command is relatively small. In this case the electric field E at the

focal plane in the presence of a small DM command can be expressed as

$$E = E_0 + Gu \quad (2.16)$$

where E_0 is the initial electric field (with the real and imaginary components split), G is the jacobian, and u is the command applied to the DM. As implied by the name, the PWP estimator uses pairs of images taken with a positive and negative version of the same DM command (I_+, I_-). When we take the difference of these images and equate it to $\Delta|E|^2$ from Eqn. 2.16 we can set up an observer as

$$y = Hx \quad (2.17)$$

where y is the difference of the probe images, H is the observation matrix, and the state x is the electric field. We can then generate an estimate the electric field via $\hat{x} = H^\dagger y$ where \dagger is the pseudo-inverse. Note that we need a minimum of two sets of probe images to obtain an estimate of the electric field as described in Sun 2019.¹³³

The stroke minimization⁹² (SM) controller generates the smallest DM command possible that still suppresses the electric field sufficiently to meet an indicated contrast constraint. SM is formulated as a cost function augmented with a Lagrange multiplier μ ; it performs a line search to find the μ that produces an acceptable contrast in the dark zone as described in Pueyo et al. 2009.⁹² Note that if a very high contrast is requested when in the early stages of generating the dark zone the SM controller can diverge but as long as the contrast requests are increased at a reasonable rate, the SM controller is very robust especially when initially generating the dark zone.

The electric field conjugation (EFC) controller is designed to minimize the electric field subject to constraints on the DM stroke. If we look at Eqn. 2.16 and set the desired E_0 to be zero, we have $E = Gu$ which we could solve for u . In reality, we cannot do this and instead minimize the error between Gu and the change in the electric field (∂E) induced by a DM command u via a least-squares approach. We implement Tikhonov regularization to ensure the inversion of the Jacobian is possible, which produces

$$u = (G^T G + \alpha \mathcal{I})^{-1} G^T \partial E, \quad (2.18)$$

where α is the Tikhonov regularization parameter and \mathcal{I} is an $m \times m$ identity matrix. Note that α must be tuned so an excessively large DM command is not requested. If EFC is used to generate the dark zone, α often has to start out large and gradually be decreased as we reach higher contrasts so to avoid the control loop diverging.

An extension of EFC for broadband applications is discussed in Give'on et al. 2007.³⁴ Here the Jacobian and the electric field estimate contain the information for l wavelengths within the band

$$\mathbf{G}_L = \begin{bmatrix} \mathbf{G}(\lambda_1) \\ \mathbf{G}(\lambda_2) \\ \vdots \\ \mathbf{G}(\lambda_l) \end{bmatrix}, \quad E_L = \begin{bmatrix} E(\lambda_1) \\ E(\lambda_2) \\ \vdots \\ E(\lambda_l) \end{bmatrix}$$

where $G(\lambda_j)$ and $E(\lambda_j)$ are the Jacobian and electric field for a single wavelength. The least-squares solution $\left((G_L^T G_L + \alpha \mathcal{I})^{-1} G_L^T \partial E_L \right)$ is then the minimum over all wavelengths and

thus optimally suppresses the electric field across the band.

2.4 DIRECT IMAGING POST PROCESSING METHODS

Even with a coronagraph and DMs, direct imaging of exoplanets benefits from data post processing. The goal when post processing direct imaging data is to find and remove the residual starlight in the dark zone. This does two things: (1) it removes bright speckles that could be falsely identified as planets and (2) it allows for imaging of planets below the dark zone contrast. This is shown by the sample image shown in the left panel of Fig. 2.10. A planet is located in the white circle but the dark zone is still speckle limited. In this case, the planet contrast is 4×10^{-8} and the mean dark zone contrast is 7.8×10^{-8} ; we are able to extract this planet using post processing as shown by the right panel of Fig. 2.10. Note that the sample image is in units of log contrast and the post processed image is in units of contrast.

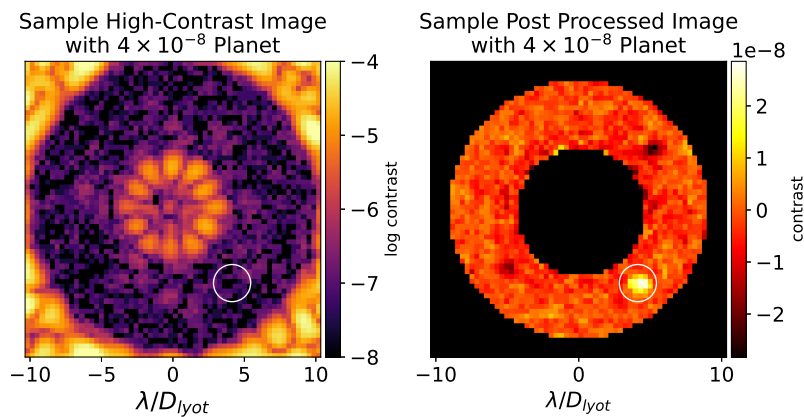


Figure 2.10: Sample high-contrast image (left panel) with a planet located in the white circle. The planet is not visible by eye but can be extracted using post processing as shown by the right panel. Note that the left image is in units of log contrast and the right is in units of contrast.

There are four main techniques to extract an exoplanet from direct imaging data:

1. Angular Differential Imaging (ADI)^{111,29,63}
2. Reference-star Differential Imaging (RDI)^{52,125}
3. Coherent Differential Imaging (CDI)^{30,15}
4. Spectral Differential Imaging (SDI)¹²²

ADI is the main technique planned to be implemented on RST where the telescope is rolled about the boresight every N hours while on a target star (it is also used in conjunction with RDI on JWST¹⁹). This moves the planet in the dark zone but, in theory, the residual starlight speckle pattern should stay the same as that depends on the optics. This allows identification of the planet as the speckles will get subtracted out since they are stationary on the focal plane. While this technique greatly facilitates post-processing, it is also quite costly. There is the time it takes to roll the telescope as well as the time it takes for the vibration modes to settle. The settling time for a smaller instrument like RST is reasonable but for a large instrument like JWST or HWO, it is a major consideration. In theory, to roll JWST 26° (the angle used by RST ADI) and wait for the instrument to settle and acquire a guide star it would cost at least 15 mins.⁶ In reality, JWST rolls must be less than 14° to stay within the shadow of the sunshield.⁵

RDI uses a reference library of the speckle pattern in the dark zone from reference star data which does not contain a planet. It is the main method employed for HST high-contrast imaging²¹ and is also used for JWST.¹⁹ RDI does not require rolling the telescope and often instruments point at a bright reference star to generate the dark zone before slewing to the target star and thus have a library of reference images already. One

downside is the speckle pattern in the dark zone must be consistent over time in order for this technique to be effective. For instruments targeting lower contrasts such as HST or ground-based telescopes, the speckle pattern is repeatable. Relying on reference star data for higher contrasts becomes much more difficult as the optics must be much more stable (sub-nanometer).

CDI is a hardware implementation and a post-processing technique combined into one. It leverages the fact that the planet light is incoherent whereas the residual starlight speckles are coherent. This means the starlight speckles can be modulated by the DMs but only the wings of the planet PSF are modulated and the core of the planet PSF is largely unaffected. There are multiple ways to implement CDI, as described in Bottom et al.¹⁵ or Galicher et al. 2010,³⁰ but generally there is a reference beam of light that is combined with the science beam to create an interference pattern that allows extraction of the planet. This concept inspired one of the approaches that is discussed later in Ch. 6.

SDI is mainly used in conjunction with an Integral Field Spectrometer which provides the spectra for Nyquist-sampled points on the focal plane. Planets can be identified from star-induced speckles based on the wavelength-dependant behaviour of the spectra. As the wavelength varies, speckles move on the focal plane but the planet PSF just changes size. It is a common approach for ground-based high-contrast instruments^{140,47} and will likely be used by JWST.¹⁴⁴ SDI unfortunately does not work well for Earth-like planets as they are dim, small, and too close to the main lobe of the stellar PSF. Speckles near the main stellar PSF move sub-pixel distances as the wavelength varies making it very difficult to distinguish them from planets.

2.5 WEAK LENSING MEASUREMENT ERROR⁹⁵

Changing gears from high-contrast imaging, I now provide an overview of considerations when performing weak lensing measurements. Weak gravitational lensing is prone and sensitive to both statistical and systematic errors. One approach for mitigating statistical errors is to restrict the scale at which the lensing signal is measured to the scale at which the signal becomes comparable to the noise.¹⁴² This is due to the issues that arise at either extreme. At large scales there are fewer modes available to measure, and thus an increase in sample variance uncertainty. On the other hand, at small scales shape noise from the telescope beam begins to dominate, decreasing the ability to measure the lensing signal. Another common issue is accurately determining the redshift of the source galaxy. This is nominally done using photometric redshifts which use crude filters and depend on strong features such as the Balmer break. This can be problematic as the Lyman- α ($Ly-\alpha$) feature can also be redshifted to the point that it appears to be the Balmer break, which is known as “photo- z degeneracy.”¹⁴² A 1% change in the redshift of the source galaxy leads to a 2% change in the power spectrum hence the source redshift must be known to the sub-percent level.¹⁴² The uncertainty in the source galaxy redshift can be reduced by using narrower filter bands over a wider range such that both the Balmer and $Ly-\alpha$ features are visible or by comparing measurements with large redshift surveys.

Instrument imperfections also contribute to errors in the weak lensing measurement. Misaligning or over-constraining the optical surfaces can drastically affect the point spread function (PSF). The PSF describes the 2D intensity pattern that results when a point source is mapped to the image plane by the telescope optics. If the PSF is not isotropic it can ap-

appear as a weak lensing signal as shown in Fig. 2.11. Imperfections in the optical surfaces as well as the wavelength-dependent nature of coatings cause the PSF to change across the focal plane. Due to the finite number of stars in the sky, the PSF cannot be characterized on every point on the focal plane thus interpolation is required. This interpolation can be unconstrained so instrument models and heuristic methods are often used in parallel.¹⁴² By dithering the telescope between exposures, SuperBIT moves the image slightly on the focal plane, which further facilitates characterization of the systematics.

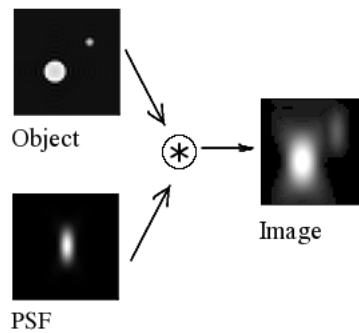


Figure 2.11: Effect of PSF on an object.¹¹³

In addition to changing across the focal plane, the PSF also changes with wavelength.¹⁴² Repeating observations in multiple bands using relatively narrow filters can aid in removing this systematic in the post-processing. The last important PSF parameter is its overall size. As the PSF size increases with respect to the galaxy size on the focal plane it dilutes the ellipticity measurement and increases the required observing time.¹⁴²

When the images are corrected for the estimated systematics it is also possible to perform null tests to check the correction success. One such test is to compute the correlation function between the PSF-corrected galaxies and the star ellipticities in the original image¹⁴². Correlation between the two is a good indicator that residual instrument systematics exist

in the image.

2.6 SUMMARY

We have now gone through the building blocks required for this thesis. Fourier and Fresnel optics provide elegant ways to propagate the electric field from the entrance of a telescope, through the high-contrast imaging components, to the focal plane. The main hardware components of a high-contrast instrument are: (1) the coronagraph and (2) the deformable mirrors. The coronagraph suppresses the starlight by a factor of $\sim 10^5$ and determines the shape and limits of the dark zone where we can directly image exoplanets. The DMs correct for the manufacturing and alignment defects in the optics. The number, stroke, and resolution of the actuators determines how large and deep of a dark zone we can generate. Even with a coronagraph and DMs, post-processing methods are important for distinguishing between residual starlight speckles and planets. Post-processing methods can also enable direct imaging of planets below the contrast in the dark zone. When considering weak gravitational lensing, the shear measurement is very sensitive to the accuracy of the redshift measurement as well as the PSF shape and stability. I will use these building blocks in the following chapters in the development of focal plane wavefront and control algorithms for high-contrast imaging and the investigation of an active optics system for a wide-field balloon-borne telescope.

3

Generating the Dark Zone using Focal Plane Broadband Estimation

This chapter is taken from Redmond et al. 2021⁹⁷ and builds on simulation results from Pogorelyuk et al. 2020.⁹⁰

3.1 MOTIVATION

One main advantage of directly imaging exoplanets is the ability to do spectroscopy on the exoplanet^{141, 23} to determine its atmospheric composition. To do spectroscopy on directly observed exoplanets, the starlight must be suppressed over a broad band of wavelengths.¹³⁹ In order to suppress the starlight within a certain band, an estimate of the electric field over the band must be obtained which can then be used to determine a DM command to reduce the light in the dark zone. There are currently two accepted approaches to estimating the electric field within a broadband dark zone: (1) use an integral field spectrometer (IFS)¹³⁵ and (2) use multiple narrowband filters.³⁶ For both the IFS and narrowband filter approaches, any of the monochromatic electric field estimators developed, such as pairwise-probing or Kalman filters³⁷, can be used.

In Chapter 4, the narrowband filter approach is exclusively used. Using narrowband filters (monochromatic images) to estimate the electric field at the filter wavelengths is the most common approach to obtain an electric field estimate across a broad band. The quality of the broadband dark zone is dependent on how many filters are used. Using narrowband filters for broadband estimation is relatively simple and does not introduce strict alignment requirements other than the planarity of the filters.³⁶ The narrowband filter approach does require a factor of l more images than the IFS approach where l is the number of filters. From a spacecraft design perspective, the use of narrowband filters for a DZM algorithm would be undesirable due to the number of filter wheel moves that would be required during the lifespan of the spacecraft. If the exposure time is 39 s (RST exposure time in Observing Scenario 11⁵⁰) and it is desired for the spacecraft to have a lifespan of 5

years, that is up to 5 million filter wheel moves.

In this chapter, I provide preliminary results for a broadband estimation scheme that does not require narrowband filters. To demonstrate this approach, I focus on using it to generate the dark zone rather than maintaining it in the presence of a drifting electric field. I provide preliminary HiCAT results for a novel estimation scheme that uses broadband images to estimate the electric field at discrete sub-bands. I show that, on HiCAT, at contrasts of $\sim 7 \times 10^{-6}$ it performs as well as the narrowband filter approach. At the time of these experiments, the broadband performance of HiCAT is limited by the laser source and the accuracy of the Jacobians. Simulations demonstrating the performance of this broadband estimator at a contrast of 10^{-9} with an RST model are provided in Pogorelyuk et al. 2020⁹⁰ (Fig. 3).

3.2 BROADBAND ESTIMATOR ALGORITHM

The Broadband estimator is an extension of the monochromatic Pairwise-probe estimator³⁷ commonly used to generate a dark zone (summarized in Sec. 2.3.3). DM probe commands are developed to modulate the electric field at the focal plane and images are taken using the positive and negative versions of those DM commands. By taking the difference of the positive and negative probed images we can estimate the electric field. The main difference for the Broadband estimator is that the electric field for multiple wavelengths is estimated simultaneously. To do this, we must assemble a broadband Jacobian containing information for all l wavelengths of interest. The monochromatic Jacobians are calculated

as

$$\mathbf{G}_i = \sqrt{t_{exp}} \begin{bmatrix} \Re \left(\frac{\partial E_0}{\partial u} \right) \\ \Im \left(\frac{\partial E_0}{\partial u} \right) \\ \vdots \\ \Re \left(\frac{\partial E_n}{\partial u} \right) \\ \Im \left(\frac{\partial E_n}{\partial u} \right) \end{bmatrix} \quad (3.1)$$

by taking the partial derivative of the electric field (E) at each pixel ($0-n$) in the dark zone with respect to each DM actuator for wavelength i ; note that u is the vector of DM actuators. The estimator operates in units of counts so the Jacobian must be multiplied by the square root of the exposure time (t_{exp}) every iteration. To have a purely real Jacobian, the real and imaginary components of the partial derivatives are stacked in a pixel-wise manner.

The monochromatic Jacobians are merged to produce a broadband Jacobian defined as

$$\mathbf{G}_{bb} = \begin{bmatrix} \mathbf{G}_0(0, :) \\ \mathbf{G}_0(1, :) \\ \mathbf{G}_1(0, :) \\ \mathbf{G}_1(1, :) \\ \vdots \\ \mathbf{G}_{l-1}(0, :) \\ \mathbf{G}_{l-1}(1, :) \\ \vdots \\ \mathbf{G}_{l-1}(n-2, :) \\ \mathbf{G}_{l-1}(n-1, :) \end{bmatrix} \quad (3.2)$$

where l is the number of wavelengths. This groups the information for each pixel since $\mathbf{G}_i(k, :)$ and $\mathbf{G}_i(k+1, :)$ are the real and imaginary components for the same pixel. In general, assuming DM commands are small, the monochromatic intensity at the focal plane after a DM command is applied can be approximated as

$$I_i = \mathbf{B}_i |E_i + \mathbf{G}_i u|^2 \quad (3.3)$$

where E is the electric field prior to the command being applied, I is the intensity at the focal plane after the command is applied, i is the wavelength, and u is the DM command. In the monochromatic case, \mathbf{B}_i is a matrix that sums the squared real and imaginary components electric field. Note that the electric field vector is split in a similar manner to the

Jacobians as shown by

$$E_i = \begin{bmatrix} \Re(E_{p=0}) \\ \Im(E_{p=0}) \\ \vdots \\ \Re(E_{p=n}) \\ \Im(E_{p=n}) \end{bmatrix} \quad (3.4)$$

where p is the pixel. Using the small command assumption, if we subtract the images at wavelength i obtained from a positive ($I_{i,j+}$) and negative ($I_{i,j-}$) probe command u_j we then get

$$I_{i,j+} - I_{i,j-} = \mathbf{B}_i |E_i + \mathbf{G}_i u_j|^2 - \mathbf{B}_i |E_i - \mathbf{G}_i u_j|^2 = 4\mathbf{B}_i \text{diag}(\mathbf{G}_i u_j) E_i \quad (3.5)$$

$$E_i = \mathbf{H}_i (I_{i,j+} - I_{i,j-}) \quad (3.6)$$

which is the monochromatic pairwise-probe estimator where $\mathbf{H} = (4\mathbf{B}_i \text{diag}(\mathbf{G} u_j))^\dagger$. Note that to make \mathbf{H} well conditioned, multiple sets of probed images must be obtained. Even with multiple sets of probes, a regularization parameter is often required to perform the psuedo-inverse. When this is expanded to create the Broadband estimator, broadband images are taken and \mathbf{B} now sums the wavelength components of $\mathbf{G}_{bb} u_j$ for each pixel as well as

the real and imaginary components. This produces the broadband observation matrix

$$\mathbf{H} = \begin{bmatrix} 4\mathbf{B} \cdot \text{diag}(\mathbf{G}_{bb}\mathbf{u}_0) \\ \vdots \\ 4\mathbf{B} \cdot \text{diag}(\mathbf{G}_{bb}\mathbf{u}_s) \end{bmatrix}^+ \quad (3.7)$$

where s is the number of probes. The observation matrix is then used to determine the electric field at the selected wavelengths,

$$\begin{bmatrix} \Re(E_{0,0}) \\ \Im(E_{0,0}) \\ \Re(E_{0,1}) \\ \Im(E_{0,1}) \\ \vdots \\ \Re(E_{0,l}) \\ \Im(E_{0,l}) \\ \Re(E_{1,0}) \\ \Im(E_{1,0}) \\ \vdots \\ \Re(E_{n,l}) \\ \Im(E_{n,l}) \end{bmatrix} = \mathbf{H} \begin{bmatrix} I_{bb,0+} - I_{bb,0-} \\ \vdots \\ I_{bb,s+} - I_{bb,s-} \end{bmatrix} \quad (3.8)$$

where n is the number of pixels, l is the number of wavelengths, and all the images are broadband images. To check the accuracy of the estimate, the broadband intensity estimate is calculated and compared to the un-probed broadband image. The estimator operates

in units of counts but we prefer the error in units of contrast so the broadband intensity estimate (\hat{I}_{bb}) is calculated as

$$\hat{I}_{bb} = \sum_{i=1}^l \frac{\Re(E_{\lambda_i})^2 + \Im(E_{\lambda_i})^2}{\max I_{d_{bb}}} \quad (3.9)$$

where $\max I_{d_{bb}}$ is the peak intensity of the un-masked broadband PSF as described in Eqn. (1.2).

The estimate error is then expressed as

$$\varepsilon_{bb} = |I_{bb} - \hat{I}_{bb}| \quad (3.10)$$

and will be discussed more in Section 3.3.

The DM probe commands are an important aspect of the Broadband estimator. For this chapter, ‘ordinary’ probes are used which introduce a uniform phase across the dark zone, determined by θ , and are calculated via

$$\mathbf{u}_j = (\mathbf{G}_i^T \mathbf{G}_i + \alpha_{probe} \mathbf{I})^{-1} \mathbf{G}_i^T \begin{bmatrix} \cos \theta_j \\ \sin \theta_j \\ \vdots \end{bmatrix} \quad (3.11)$$

where α_{probe} is the regularization parameter and \cos/\sin alternate for all n pixels in the dark zone. For a given set of s probes, θ sweeps from $0 - \pi$ in equal steps. A sample broadband probe for $\theta_0 = 0$ is shown in Fig. 3.1 for iteration 18 of a focal plane wavefront control (FPWC) experiment. The amplitude of the probe is updated each iteration based on the contrast; a higher contrast requires a smaller probe amplitude. It is important to note that the probes are wavelength dependent. Currently, the probes are calculated using the Ja-

cobians for the centre wavelength of the band. Further investigation into improving the probe effectiveness is in progress.

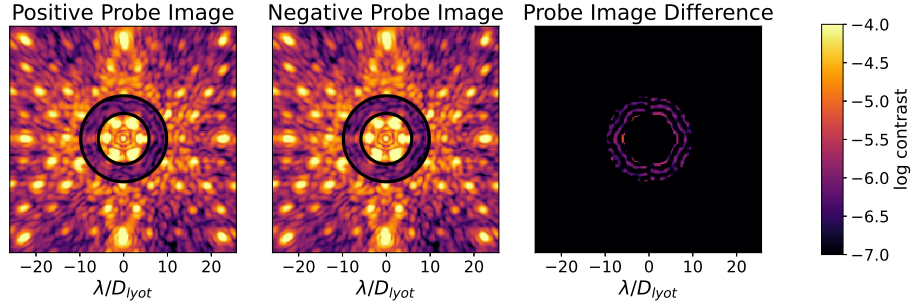


Figure 3.1: Example probe for the Broadband estimator ($\theta = 0$). The difference between the positive and negative probe images cannot be seen by eye but the phase introduced is easily seen in the probe difference image in the right panel. The probe images use the 6% filter as is evident by the smearing of the satellite spots towards the edge of the frame.

Once an estimate is obtained for each of the desired wavelengths, Electric Field Conjugation (EFC)³⁴ is used to determine the optimal DM command. As discussed in Pogorelyuk et al. 2020⁹⁰, this improves the numerical conditioning of the overall algorithm. The DM command based on a single wavelength estimate is expressed as

$$\Delta u_i = (\mathbf{G}_i^T \mathbf{G}_i + \alpha \mathbf{I})^{-1} \mathbf{G}_i^T \hat{\mathbf{E}}_i \quad (3.12)$$

where α is the Tikhonov regularization parameter that limits the maximum allowable DM command. For broadband control, the command is taken as the mean of the EFC output for each wavelength,

$$\Delta u = \frac{1}{l} \sum_{i=0}^{i=l} \Delta u_i. \quad (3.13)$$

For future experiments we plan on implementing the broadband EFC algorithm discussed in Give'on et al. 2007³⁴ and summarized in Sec. 2.3.3.

3.3 LABORATORY RESULTS

Using a dark zone size of $5.8 - 9.8 \lambda / D_{lyot}$ on HiCAT, there are six very bright lobes near the IWA. These lobes can be difficult to suppress when they are on the border of the dark zone. To improve suppression of the bright inner lobes, the dark zone is initially dug to a contrast of 10^{-6} using an oversized region with an IWA of $3.8 \lambda / D_{lyot}$. The DM command obtained by the oversized dark zone run is referred to as the 'resume command' and is used as a starting point for the 'real' experiments. In some cases the resume command is obtained for only the 640 nm band which can cause issues for the broadband estimator as the starting contrast will not be uniform across the wavelength band; this will be demonstrated in Section 3.3.1.

3.3.1 CUBE MODE

As an intermediate step between monochromatic and true broadband estimation, the 'cube mode' of the Broadband estimator was developed. In cube mode, images are taken using 10 nm bandpass filters and added together (in intensity units, not contrast) to create a stand-in broadband image. The estimator only has access to the broadband image but the monochromatic images can be used to check the estimate quality. Cube mode would never be used for on-sky operations.

To demonstrate cube mode, Fig. 3.2 and 3.3 show results from an experiment run using the 640 and 660 nm filters. The 640 and 660 nm filters are chosen as they provide a

20 nm band that span a 30 nm band (5%) and should provide a good reference point for when the 6% filter is used. For this experiment, the estimating wavelengths are also 640 and 660 nm. The initial DM command is a 640 nm resume command (as described above). Six probed images are taken at each wavelength and added together to obtain $I_{bb,+}$, $I_{bb,-}$ (to be used in Eqn. (3.8)) for each of the s probes. After the monochromatic electric field estimates are obtained using the probed images, the broadband estimate error is calculated via Eqn. (3.10).

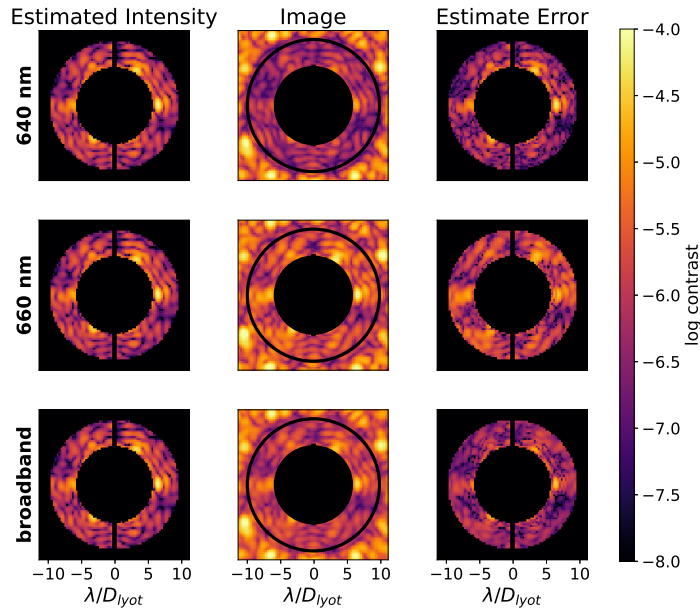


Figure 3.2: Comparison of the estimates and images for iteration 0 in the cube mode experiment. The bottom row contains broadband data which is calculated via Eqn. (3.14). The left column contains the estimates where the top two rows are calculated by the estimator. The middle column contains the images; the bottom broadband image is the type of image that is provided to the estimator. The center of the images is masked to highlight the features in the dark zone near the IWA. Here we can see the effect of using the resume DM command as the 640 nm dark zone starts off with a much higher contrast. Since all estimated wavelengths are assumed to contribute equally to the broadband intensity, the electric field is over-estimated at 640 nm and under-estimated at 660 nm. The right column contains the estimate error described by Equation (3.10) where we can see that the 660 nm estimate error is much larger than the 640 nm estimate error, again, due to the use of the resume command.

As shown in Fig. 3.2, the Broadband estimator has moderate success at estimating the electric field at the selected wavelengths. The left column shows the estimates at each wavelength, the center column shows the images at each wavelength, and the right column shows the error between the left and center columns. Note that the contrast for the first two rows is expressed as $I_{m_i} / \max I_{d_i}$, where i is the wavelength, but the bottom row contrast is expressed as

$$\text{contrast}_{bb} = \sum_{i=1}^l \frac{I_{m_i}}{\max I_{d_{bb}}} \quad (3.14)$$

as the peak intensity of the stand-in direct broadband image ($I_{d_{bb}}$) is approximately twice that of the individual wavelengths. The effect of using the 640 nm resume command is evident in the middle column as the initial 640 nm dark zone is much more pronounced. The estimator assumes all wavelengths to be equal contributors to the broadband intensity and thus over-estimates the 640 nm electric field and under-estimates the 660 nm electric field.

Figure 3.3 shows the contrast plots for the cube experiment. Note that the 3.3a is a worst case scenario as it is taken along the $y = 0$ line where the residual spoke pattern seen in Fig. 3.2 is most prominent. Figure 3.3b shows the broadband and monochromatic mean dark zone contrast vs. time. Note that the mean dark zone contrast refers to the spatially averaged contrast within the $5.8 - 9.8 \lambda / D_{\text{lyot}}$ annulus. Here we can see that the 640 nm resume command causes the initial 640 nm contrast (purple dashes) to be a factor of four better and the steady state contrast to be a factor of two better than the 660 nm contrast (blue dots). The large contrast discrepancy between the two wavelengths is likely what

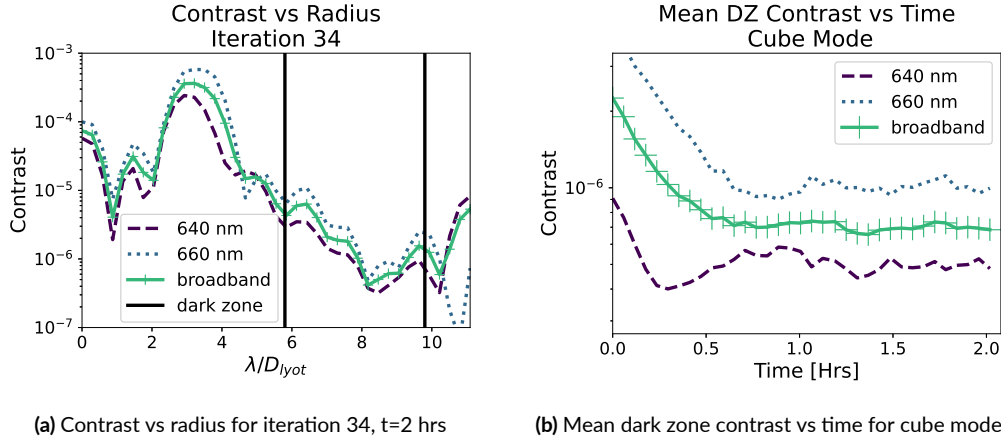


Figure 3.3: Contrast plots for cube mode experiment. Figure 3.3a shows the worst case scenario contrast vs. radius with an IWA broadband contrast of $\sim 3.5 \times 10^{-6}$ and an OWA broadband contrast of $\sim 2 \times 10^{-6}$. Figure 3.3b shows the mean dark zone contrast at each wavelength as well as the broadband mean dark zone contrast vs. time. The 640 nm contrast begins at a factor of four better contrast than the 660 nm due to the resume command used. By iteration 34 (2 hrs), the monochromatic contrasts are within a factor of two. The steady state broadband contrast is 6.9×10^{-7} .

causes the 640 nm contrast increase after t=0.25 hrs.

3.3.2 BROADBAND MODE

Broadband mode is the approach that would be used on-sky where a single broad filter is used to acquire the broadband probe images. On HiCAT we use a 6% filter centered at 640 nm. The broadband probe images are then plugged into Eqn. (3.8) to estimate the electric field at the chosen wavelengths. For this experiment, increments of 10 nm were chosen from 620 – 670 nm as the estimation wavelengths to match the filters available to the narrowband filter approach. Also, the transmission of the filter is $>90\%$ from 616–662 nm and we wanted to ensure the edges of the band were covered. In future experiments we will likely not estimate 670 nm when using the 6% filter. For this experiment, 30 probes were used which is five probes per estimated wavelength. Figure 3.4 shows the product of the

Broadband estimator for iteration 0. The right broadband panel is computed separately after. As expected the estimates have similar features but are not identical and differ at the 10^{-7} level.

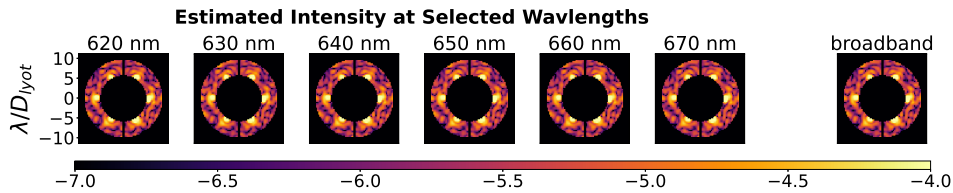


Figure 3.4: Broadband estimator estimate breakdown by wavelength for iteration 0.

Figure 3.5 shows the comparison of the broadband estimate for iteration 20 and the broadband image. This experiment does not use the resume command and the six bright lobes inside the IWA are very evident. The estimator captures the hot spots at the IWA but does not capture the full spokes as they extend into the dark zone.

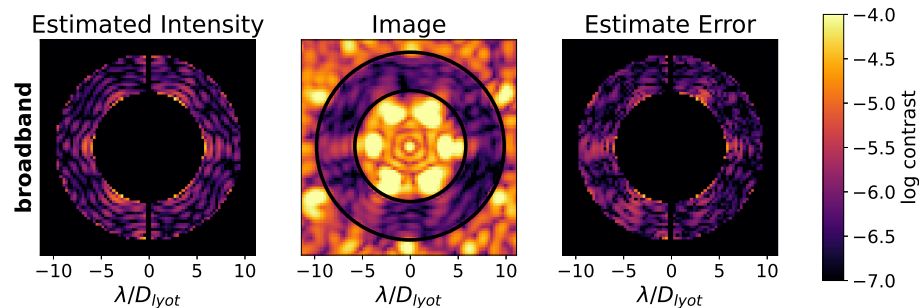


Figure 3.5: Broadband estimate comparison to the broadband image for iteration 20.

Figure 3.6 looks at the contrast vs. radius and how that varies with time. The first iteration is shown by the dashed blue lines and the third last iteration is shown by the green dots. The majority of the correction occurs in the first five iterations at which point the algorithm fails to continue removing the six-fold spoke pattern. As in Section 3.3.1, the slice

in Fig. 3.6 is along $y = 0$ which is a worst case scenario since it follows one of the spokes. It could be that we need to run a short experiment with the Broadband estimator using the 6% filter and a smaller IWA to kill the bright lobes and the spoke pattern but it is our hope that we can achieve this by improving the estimator instead.

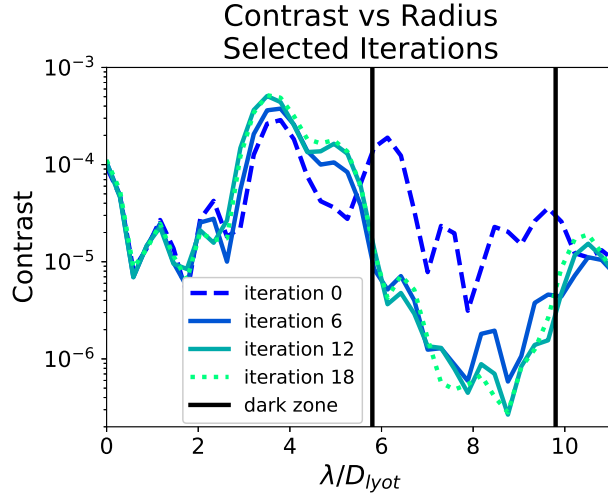


Figure 3.6: Broadband contrast vs radius for select iterations. Iteration 0 is shown by the dark blue dashes and iteration 18 is shown by the green dots. The dark zone limits are shown by the black vertical lines. This slice is taken through the center of the dark zone where $y = 0$ and is a worst case scenario since it contains the residual spoke pattern. The bulk of the correction occurs in the first five iterations at which point the contrast begins to asymptote.

3.3.3 MODE COMPARISON

When performing these experiments in 2021, HiCAT had not yet been optimized for broadband performance so it is not quite fair to judge the absolute dark zone digging performance of the Broadband estimator. In Fig. 3.7 we compare different methods of generating a dark zone on HiCAT. It should be noted that the cube mode (green solid curve) and monochromatic pairwise approach (purple dashes) utilize the resume command. The cube mode and monochromatic pairwise method are both only estimating 640 and

660 nm as they tend to have stability issues when incorporating more wavelengths. The stability issues can be addressed for the most part by choosing very conservative controller parameters but that results in very slow dark zone generation. Broadband mode as shown by the blue dots in Fig. 3.7 is slightly more robust and can have slightly more aggressive controller parameters. Though the monochromatic pairwise method achieves the highest mean dark zone broadband contrast (6.3×10^{-7}), the slopes of all of the curves are approximately equivalent after they pass 2×10^{-6} . This is a good sign as it indicates the Broadband estimator performance is not significantly worse than its monochromatic partner. It should also be noted that the monochromatic pairwise experiment used four probes per wavelength estimated, the cube mode experiment used three probes per wavelength estimated, and the broadband mode used five probes per wavelength estimated.

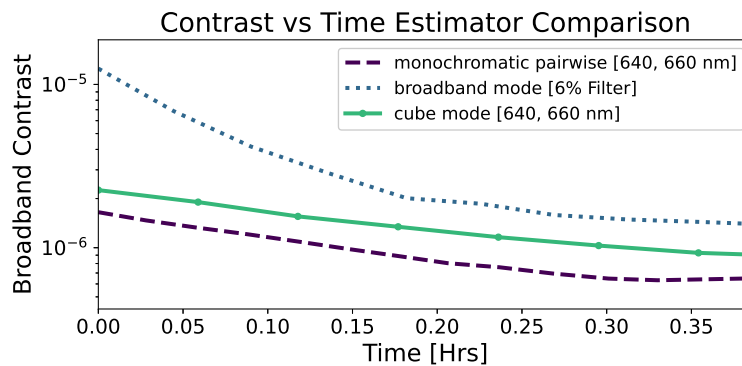


Figure 3.7: Comparison of methods to generate a broadband dark zone on HiCAT. The Broadband estimator approaches are shown by the blue dotted line and green solid line representing broadband mode and cube mode respectively. The monochromatic pairwise approach is shown by the purple dashes. The monochromatic pairwise method achieves the highest contrast but all curves have approximately the same slope once they hit 2×10^{-6} .

3.4 CONCLUSIONS AND FUTURE WORK

Generating a dark zone across a wide wavelength band is an important capability for a high contrast imaging system. In order to minimize hardware requirements, I investigate the potential of using broadband images to estimate the electric field at discrete wavelengths within the band. I demonstrate focal plane wavefront estimation and control using broadband images where the mean dark zone contrast is increased by an order of magnitude (from 10^{-5} to 10^{-6}) across a 6% band centered at 640 nm. Using images taken with the 6% filter we estimate the electric field at 10 nm increments from 620–670 nm.

Though I have demonstrated the potential of the Broadband estimator, there are many areas of improvement. Pogorelyuk et al. 2020⁹⁰ use an RST model to show that this estimator can operate at a contrast of 10^{-9} so the HiCAT results are far from any fundamental limit. A common issue encountered throughout the experiments presented in this chapter is the accuracy of the Jacobians. Extensive model matching has only been done for the 640 nm Jacobian causing the contrast at 640 nm to improve much more quickly than the other wavelengths. Performing model matching at other wavelengths is expected to improve the performance of both the Broadband estimator and the monochromatic pairwise estimator. Moving forward a main point of interest is the probes used for the Broadband estimator. In Pogorelyuk et al. 2020⁹⁰ random probes were used but these were found to perform worse in the HiCAT simulator and on the HiCAT testbed. The goal is to induce chromatic effects that probe the entire dark zone. For this reason using two-dimensional Sinc-shaped probe commands on DM2 might be a better option than ordinary probes (Eqn. 3.11) on DM1. This is because the Fourier transform of a Sinc function ($\sin(x)/x$)

is a rectangular function and thus using a Sinc-shaped DM command (in a pupil plane) creates a uniform change in the electric field at the focal plane (see Sec. 2.1 for supporting background). By putting the command on DM₂, which is not at a conjugate pupil plane on HiCAT, the chromatic effects are exasperated. Lastly, the chromaticity of the testbed has not been considered. Adding spectral weights to the estimated electric fields may also improve the Broadband estimator performance.

4

Quasi-static Drift Correction on a Coronagraph Testbed

This chapter is taken from Redmond et al. 2022.⁹⁶

4.1 CHAPTER OVERVIEW

Due to the limited number of photons, directly imaging planets requires long integration times with a coronagraphic instrument. The wavefront must be stable on the same time scale, which is often difficult in space due to time-varying wavefront errors from thermal gradients and other mechanical instabilities. In this chapter, I discuss a laboratory demonstration of a photon-efficient dark zone maintenance (DZM) algorithm in the presence of representative wavefront error drifts. The DZM algorithm allows simultaneous estimation and control while obtaining science images and removes the necessity of slewing to a reference star to re-generate the dark zone mid-observation of a target. The experiments are performed on the High-contrast imager for Complex Aperture Telescopes (HiCAT) at the Space Telescope Science Institute (STScI). The testbed contains an IrisAO segmented primary surrogate, a pair of continuous Boston Micromachine (BMC) kilo deformable mirrors (DMs), and a Lyot coronagraph. Both types of DMs are used to inject synthetic high-order wavefront aberration drifts into the system, possibly similar to those that would occur on telescope optics in a space observatory, which are then corrected by the BMC DMs via the DZM algorithm. In the presence of BMC, IrisAO, and all DM wavefront error drift, I demonstrate maintenance of the dark zone contrast ($5.8\text{--}9.8 \lambda/D_{\text{Lyot}}$) at monochromatic levels of 8.5×10^{-8} , 2.5×10^{-8} , and 5.9×10^{-8} respectively. In addition I show multi-wavelength maintenance at a contrast of 7.0×10^{-7} over a 3% band centred at 650 nm (BMC drift). This work demonstrates the potential of adaptive wavefront maintenance methods for future exoplanet imaging missions and our demonstration significantly advances their readiness.

4.2 INTRODUCTION

For our HiCAT tests of the DZM algorithm, I manually inject drifts into the optical system that loosely represent the quasi-static drifts expected on space telescopes, such as RST and HWO, using a random walk model. I make these drifts large enough so they are larger than the HiCAT environmental noise. The drifting DM can be either the two BMC DMs or the IrisAO segmented aperture (or all three). I study the DZM performance at multiple wavelengths in order to quantify chromatic effects. A summary of the experiment configurations is shown in Tab. 4.1. Our successful tests demonstrate that the monochromatic DZM algorithm can operate at least at the 10^{-8} level of contrast (current HiCAT contrast limit), which can potentially increase science observing time by eliminating the need for a reference star and relax stability requirements for large aperture segmented telescopes.

Table 4.1: Configurations for the HiCAT experiments included in this chapter and the relevant sections which provide an in depth description.

Section	Wavelength [nm]	Drifting DM	IrisAO Installed? [Y/N]
4.4.2	640	BMC	N
4.4.3	640, 660	BMC	Y
4.4.4	640	IrisAO	Y
4.4.5	640	BMC & IrisAO	Y

Section 4.3 outlines the DZM algorithm and its ability to correct for different drifts at multiple wavelengths. Next, I cover the dither calibration and experimental results from relevant experiments on HiCAT in Sec. 4.4. Notes on the HiCAT hardware are provided in Appendix A and optimal dither selection for the BMC drift is provided in Appendix B. Lastly, Sec. 4.6 provides an overview of the DZM results so far and outlines the plan for

future work.

4.3 DARK ZONE MAINTENANCE

4.3.1 DARK ZONE CREATION VS DARK ZONE MAINTENANCE

The goal of the DZM is to maintain high contrast in the dark zone, implying that we must first create, or ‘dig’, the dark zone. HiCAT uses a pair-wise probe estimator³⁷ combined with the stroke-minimization^{92,37} controller as discussed in Soummer et al. 2019¹²⁶ and summarized in Sec. 2.3.3. The dark zone generation algorithm is referred to as the PPSM (Pairwise Probe Stroke Minimization) algorithm. The PPSM algorithm uses the stroke-minimization controller as the wavefront changes drastically during the generation of the dark zone and the dynamic tuning of the regularization parameter is advantageous.^{37,134} Since the wavefront is relatively stable during DZM, we can use EFC as the controller with a constant regularization parameter¹³⁴; EFC will be further discussed in Sec. 4.3.2. Though the PPSM algorithm is effective for generating the dark zone, it is less advantageous for maintaining the dark zone. PPSM requires at least two sets of ‘probe’ images (four images total) to obtain an estimate of the electric field. The number of required images drives the iteration time and limits the rate of quasi-static drifts PPSM can correct for. Ideally, the DZM algorithm is provided 10–20 images prior to the controller turning on to allow the extended Kalman filter (EKF) to converge and then requires a single image per iteration; note that the DZM experiments are 1000+ iterations. In addition, the PPSM ‘probe’ images are taken at a lower contrast making them less useful for science analysis. The dithering technique used in the DZM algorithm to calculate the estimate of the open loop electric field can also be used in post-processing to locate a planet⁸⁹ making it an attractive ap-

proach for maintaining a high-contrast dark zone.

For all results in this chapter, the control region is an annulus from $5.8\text{--}9.8 \lambda / D_{\text{lyot}}$ where D_{lyot} is the diameter of the Lyot stop. The mean contrast is the spatially averaged contrast (Eqn. (1.2)) in the dark zone,

$$\text{mean contrast} = \mu = \frac{1}{n} \sum_{i,j} \frac{I_{\text{coron}}(i,j)}{\max(I_{\text{direct}})}, \quad (4.1)$$

where n is the number of pixels in the dark zone and i, j are the indices of the pixels in the dark zone. I leave demonstrations with larger or smaller dark zones for future investigations.

The high-level Dark Zone Maintenance (DZM) algorithm is shown in Fig. 4.1. An aberrated electric field, E_{ab} , arrives at the DMs and is partially corrected to create the closed-loop (CL) electric field at the science camera for iteration k , E_f^k . The camera measures the intensity, $z^k = I^k = |E_f^k|^2$, which is used to estimate the open-loop (OL) electric field, $x^k = E_{OL}^k$, using an EKF. Once the estimate is obtained, the Electric Field Conjugation (EFC) algorithm is used to determine the next DM command. A small random dither command is added to the EFC command to increase the phase diversity at the focal plane which improves the estimate. When using continuous DMs to introduce wavefront instabilities, a small drift command is added to the DM command when the image is taken; these drift commands are unknown by the estimator.

Figure 4.2 shows how the PPSM and DZM algorithms work together. In the left panel of Fig. 4.2, the mean contrast as a function of time using the PPSM algorithm (in a high-flux regime with no drifts added other than the natural drifts of the testbed) is indicated in

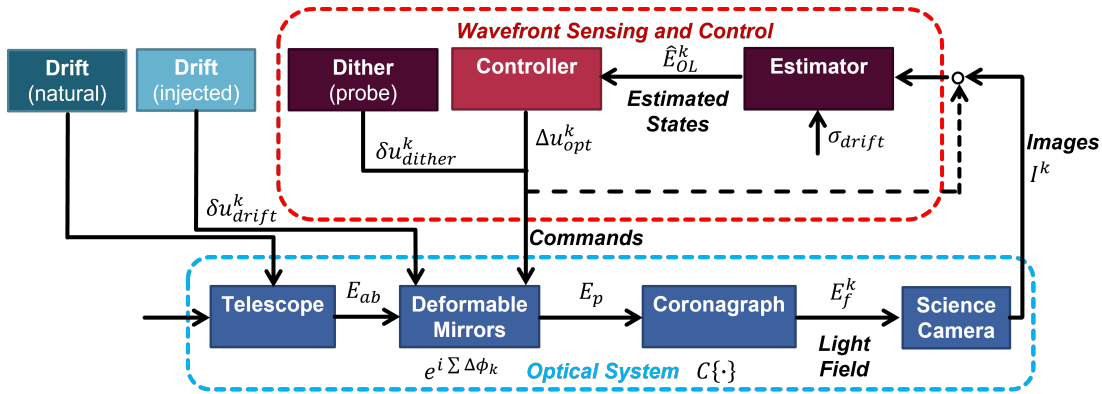


Figure 4.1: The light passes through the telescope optics which introduce small static aberrations due to surface imperfections and misalignments. In addition, there are transient aberrations or natural drifts, ∂E_{ab} , which come from thermal drifts and pointing jitter; these transient aberrations are approximated by δu^k_{drift} in our experiments using the available DMs. The intensity of the closed-loop electric field E^k_f is detected at the science camera after interacting with the deformable mirrors and the set of masks that make up the coronagraph. The measured closed-loop intensity, $z^k = I^k$, is used to estimate the open-loop electric field at the camera (focal plane), $\hat{x}^k = \hat{E}^k_{OL}$. The estimate is used to find the change in the DM command Δu^k_{opt} that will correct the open-loop field in the dark zone region. Before being applied, a random dither (δu^k_{dither}) and drift (δu^k_{drift}) are added to the DM command.⁶⁸

black diamonds. Once the mean contrast stops improving (after ~ 0.8 hrs for Figure 4.2), the digging process is terminated and the state of the correction (\hat{x}_0, u_0) is used as the initial condition for the Dark Zone Maintenance (DZM) algorithm. Results from such a DZM experiment are shown in Fig. 4.2 in magenta crosses, mitigating a simulated OL contrast drift where each BMC DM actuator is performing a random walk as described by Eqn. 4.4, with $\sigma^2_{drift} = 100 \text{ pm}^2/\text{iter}$. The open-loop contrast associated with this WFE drift is shown in cyan in the left panel of Fig. 4.2. Due to the short timescale used for Fig. 4.2, the open-loop contrast has not had time to significantly drift which is why the magenta and cyan curves are still rather similar with respect to the mean DZ contrast. The final open-loop image (no DZM, BMC drift) is shown in the right panel of Fig. 4.2 where there are now a number of low contrast (10^{-5}) speckles. Note that the DZM algorithm is estimat-

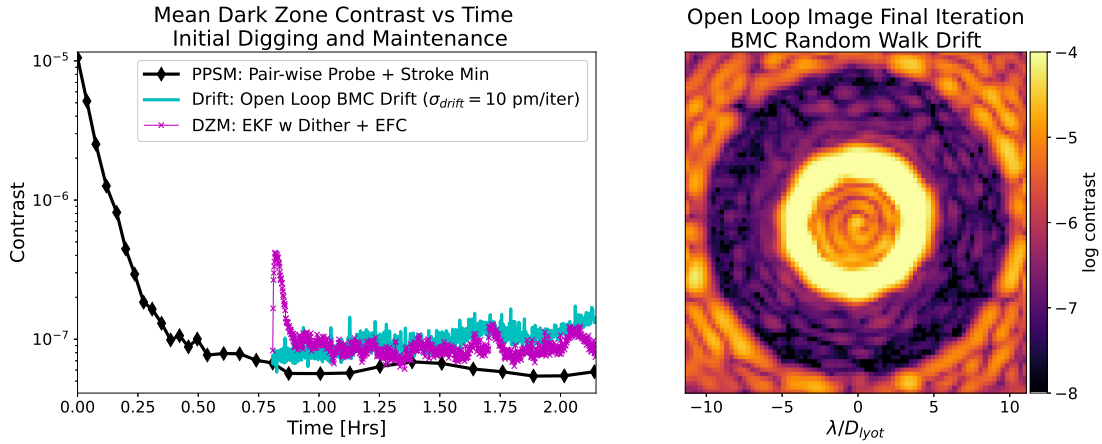


Figure 4.2: In the left panel I show HiCAT lab data comparing a dark zone dig experiment and a dark zone maintenance experiment. The PPSM algorithm is used to dig the dark zone (black diamonds); no drift is injected during the dark zone dig. At 0.8 hrs I overlay a sample DZM experiment; all curves have a contrast of $\sim 8 \times 10^{-8}$ at 0.8 hrs. For the DZM experiment (magenta crosses), an EKF combined with DM dithering is used to estimate the electric field, and EFC is used to correct for the injected drift. The injected drift (solid cyan) is characterized by the variance of the random walk of each actuator, $\sigma_{drift}^2 = 100 \text{ pm}^2/\text{iter}$, and the dither is characterized by the standard deviation of the random command added to the EFC control, $\sigma_{dither} = 200 \text{ pm}$. The DM command applied to obtain the cyan curve is described by Eqn. (4.4) and will be further discussed in Sec. 4.3.2. After 1.3 hrs, the open-loop contrast has degraded to 1.4×10^{-7} while the closed loop contrast is 7.6×10^{-8} . The right panel shows the final image of the dark zone for the open loop curve (solid cyan) in the left panel. The dark zone has degraded and many low-contrast speckles are present that can hide exoplanets.

ing the electric field of the solid cyan curve, represented by the intensity image in the right panel of Fig. 4.2.

4.3.2 DARK ZONE MAINTENANCE ALGORITHM

In this section I discuss the math behind the DZM algorithm. Standard controls notation is used to describe the EKF so Table 4.2 is provided to highlight the conversion between traditional WFSC notation (used in Fig. 4.1) and the EKF notation. Note in both notations, a $\hat{\cdot}$ indicates that the variable is an estimate, not the true value or measurement.

Table 4.2: Conversion table between WFSC notation and EKF notation.

Parameter	WFSC Notation	EKF Notation
Open loop electric field	E_{OL}	x
Estimate of open loop electric field	\hat{E}_{OL}	\hat{x}
Closed loop electric field	E_f	x_{CL}
Measured closed loop intensity	I	z

STATE SPACE MODEL AND SYNTHETIC DRIFTS

For the dynamical system described in Fig. 4.1, the state is the open-loop drifting electric field in the dark zone (x), the measurement is the closed-loop intensity at the science camera in the dark zone (z), and the control is the BMC DM command (u). Note that the IrisAO has 37 segments, each with piston-tip-tilt (PTT) degrees of freedom, and is only used for drift injection, not control. There are $n = 2, 228$ pixels in the dark zone and $m = 2 \times 952$ actuators are available for the control due to the two kilo-DMs. The electric field is a complex number so the real and imaginary components are split to facilitate working with real matrices; as a result, the state contains $2n = 4, 456$ entries. The DZM algorithm uses an Extended Kalman Filter (EKF) to estimate the open-loop electric field at the desired wavelengths and an EFC controller to determine the DM command to correct for the drift. The EKF operates in a pixel-wise manner so the real and imaginary components for each pixel are stacked for both the state variables and the Jacobians,

$$G = \sqrt{t_{exp}} \begin{bmatrix} \Re \left(\frac{\partial x_0}{\partial u} \right) \\ \Im \left(\frac{\partial x_0}{\partial u} \right) \\ \vdots \\ \Re \left(\frac{\partial x_{n-1}}{\partial u} \right) \\ \Im \left(\frac{\partial x_{n-1}}{\partial u} \right) \end{bmatrix}, \quad (4.2)$$

$$x = \begin{bmatrix} \Re(x_0) \\ \Im(x_0) \\ \vdots \end{bmatrix}, \quad (4.3)$$

where the pixel index in the dark zone ranges from 0 to $(n - 1)$ and G is the BMC Jacobian. Note that G is the total BMC Jacobian where the Jacobians for BMC DM₁ and BMC DM₂ are stacked horizontally ($G = [G_{DM1} \ G_{DM2}]$). The EKF estimator operates in units of counts so the Jacobians are in units of $[(\text{counts/s})^{1/2}/\text{nm}]$ and must be multiplied by the square-root of the exposure time ($\sqrt{t_{exp}}$) each iteration.

In our testbed experiments, the drift is injected as a random walk of each DM actuator and can be expressed as

$$u_{drift}^{k+1} = u_{drift}^k + \mathcal{N}(0, \sigma_{drift}^2 \mathcal{I}), \quad (4.4)$$

where k is the iteration, \mathcal{I} is the identity matrix (as we assume there is no cross-talk between

the actuators), and σ_{drift} is the standard deviation of the normal distribution, \mathcal{N} . For the BMC DMs, σ_{drift} is a single number; for the IrisAO, σ_{drift} is broken up into $\sigma_P, \sigma_T, \sigma_t$ since each segment has piston (P), tip (T), and tilt (t) capabilities (PTT). When using the IrisAO to inject drift, each degree of freedom drifts independently and is not constrained to be from the same distribution.

The nonlinear state space model is described as

$$x^{k+1} = f(x^k) + w^k = x^k + w^k, \quad (4.5)$$

$$x_{CL}^k = x^k + Gu^k, \quad (4.6)$$

$$z^{k+1} = h(x^k, u^k) + n^k = x_{CLRe}^k \circ x_{CLRe}^k + x_{CLIm}^k \circ x_{CLIm}^k + n^k, \quad (4.7)$$

where \circ is the Hadamard operator, u^k is the DM command for the current iteration, x_{CL}^k is the closed-loop electric field, and w^k, n^k are the process and observation noise terms. As with the BMC Jacobians, the DM command u is a stacked vector containing the command for both BMC DMs ($u = [u_{DM1} \ u_{DM2}]^T$). Since we are injecting drift using the DMs, the process noise term at each iteration is

$$w^k = G(u_{drift}^{k+1} - u_{drift}^k) + G_P(u_P^{k+1} - u_P^k) + G_T(u_T^{k+1} - u_T^k) + G_t(u_t^{k+1} - u_t^k), \quad (4.8)$$

where G_{PTt} are the IrisAO Jacobians for piston, tip, tilt and u_{PTt} are the IrisAO piston, tip, tilt drift commands. Note that for this chapter, in all BMC drift cases both BMC DMs are used for drift injection simultaneously. For a space telescope, a model of the predicted drifts would be used to determine the process noise.

ESTIMATOR: EXTENDED KALMAN FILTER

The Extended Kalman Filter is a nonlinear estimator with memory solving a maximum likelihood estimation (MLE) problem.¹²⁰ The EKF takes into account the previous time step biases in addition to the current estimate residual.¹³³ The first step in the EKF is to linearize the nonlinear model around the current state to produce

$$F^k = \left. \frac{\partial f}{\partial x^k} \right|_{x^k = \hat{x}^{k-1}}, \quad (4.9)$$

$$H^k = \left. \frac{\partial b}{\partial x^k} \right|_{x^k = \hat{x}^{k|k-1}}, \quad (4.10)$$

where H^k is the observation matrix and $F^k = \mathcal{I}$ is the state transition matrix. Since our state is the open-loop electric field there are no dynamics in the system and the time dependent component is contained in the process noise term.

Next we calculate the estimate (\hat{x}) and variance (P) for the current time step given the estimate from the previous time step:

$$\hat{x}^{k|k-1} = F^k \hat{x}^{k-1}, \quad (4.11)$$

$$P^{k|k-1} = F^k P^{k-1} (F^k)^T + Q^k, \quad (4.12)$$

where Q^k is the process noise. For the initial time step, we have the electric field estimate from the PPSM algorithm and initialize $P^{0-1} = 0$. The process noise must contain the total predicted drift in the system. Due to the additive property of random matrices, the covariance of the sum is the sum of the covariances.¹³⁶ In the presence of a random walk

drift of each BMC DM actuator, the BMC process noise is expressed as

$$Q_{BMC}^k = GG^T\sigma_{drift}^2 \quad (4.13)$$

as discussed in Pogorelyuk et al. (2020).⁹⁰ Here, σ_{drift} is the standard deviation of the random drift command applied to each BMC actuator as shown in Eqn. 4.4. The formulation is similar for the IrisAO except we must account for all three degrees of freedom,

$$Q_{iris}^k = (G_p G_p^T \sigma_p^2 + G_T G_T^T \sigma_T^2 + G_t G_t^T \sigma_t^2), \quad (4.14)$$

where G_p , G_T , and G_t are the piston, tip, and tilt Jacobians for the IrisAO. To get the total process covariance matrix we add the contributions from each DM to get

$$Q^k = Q_{BMC}^k + Q_{iris}^k. \quad (4.15)$$

Using the HiCAT simulator,^{72,81,127} we can predict how accurate Q needs to be in order for the system to be stable. Short simulations (100 iterations) with σ_{drift} errors of 50% show no effect on the DZM performance. Longer testbed runs (200 iterations) with σ_{drift} errors of 5% have corroborated the simulation results for small errors at a variety of drifts ($0.1 \leq \sigma_{drift} \leq 0.9 \text{ nm}/\sqrt{\text{iter}}$). The DZM stability with respect to the IrisAO drift error has not yet been studied but a similar result is expected. For this chapter, experiments have perfect knowledge of σ_{drift} , $\sigma_{P,T,t}$ as a detailed look into DZM stability is outside the scope.

Up next we calculate the Kalman gain which is used to determine the electric field esti-

mate and covariance for the current iteration

$$K^k = P^{k|k-1}(H^k)^T (H^k P^{k|k-1}(H^k)^T + R^k)^{-1}, \quad (4.16)$$

$$\hat{x}^k = \hat{x}^{k|k-1} + K^k [z^k - h(\hat{x}^{k|k-1}, u^k)], \quad (4.17)$$

$$P^k = P^{k|k-1} - K^k H^k P^{k|k-1}, \quad (4.18)$$

where R^k is a square matrix that approximates the covariance of measurement noise and z^k is the measurement. It contains the nonlinear measurement estimate of the state (open-loop intensity) on the diagonal: $R^k = \text{diag}(\hat{y}^{k|k-1}) = \text{diag}(h(\hat{x}^{k|k-1}, u^k))$. Note that the EKF does not have access to the entire command applied to the BMC DMs, only the EFC command and dither:

$$u^k = \Delta u_{opt}^k + \delta u_{dither}^k \quad (4.19)$$

$$\delta u_{dither}^k = \mathcal{N}(0, \sigma_{dither}^2 \mathcal{I}) \quad (4.20)$$

where Δu_{opt}^k is the optimal DM command determined by EFC and δu_{dither}^k is the dither command. Currently on HiCAT we apply a unique dither command to each BMC DM, but dithering two DMs is not strictly necessary for the DZM algorithm to work. Riggs et al. (2016)¹⁰² and Pogorelyuk et al. (2019)⁸⁸ discuss the risk of using an EKF without any probe images as it can cause the estimate to converge to the wrong value. The dither command acts as a small probe which increases the phase diversity of the electric field between iterations to improve the estimate while not significantly degrading the contrast. Since the dither only minimally degrades the contrast, the DZM algorithm produces more science images in a given amount of time at the target contrast than the pair-wise probe es-

timator. The dither is not required to be random and optimal dither probes will be investigated in the future. Also, as discussed in Pogorelyuk et al.⁹⁰, this DZM algorithm improves the post-processing results for various techniques including Angular Differential Imaging (ADI) and Electric Field Order Reduction (EFOR) when compared to the open-loop approach.

CONTROLLER: ELECTRIC FIELD CONJUGATION

Once the estimate of the open-loop electric field is obtained, EFC³⁴ is used to determine the optimal DM command to minimize the closed-loop field present via

$$\Delta u_{opt}^{k+1} = -\beta (G^T G + \alpha \mathcal{I})^{-1} G^T \hat{x}^k, \quad (4.21)$$

where α is the Tikhonov regularization parameter to avoid excessively large DM commands and β is the control gain. The EKF operates by storing information from the previous estimates and needs to accumulate a certain amount of data before the estimate will converge (App. B). The control gain (β) is often set to zero for the first 10–20 iterations, depending on the speed of the drift, to allow the estimator to converge prior to applying an EFC command (this will be discussed further in Sec. 4.4.2). Another important note is that the total control command update⁸⁸ changes slightly when the open-loop electric field is assumed to drift and is used as the state variable. In a traditional EFC use case,¹³³ where the closed-loop electric field is the state variable and the open-loop electric field is assumed to be constant, the DM command is updated via $u_{control}^{k+1} = u_{control}^k + \Delta u_{opt}^{k+1}$. In DZM, once the EFC correc-

tion is obtained, the DM command applied to the BMC mirrors follows as

$$\mathbf{u}_{control}^{k+1} = \mathbf{u}_0 + \Delta \mathbf{u}_{opt}^{k+1} + \delta \mathbf{u}_{drift}^{k+1}, \quad (4.22)$$

$$\mathbf{u}_{tot}^{k+1} = \mathbf{u}_{control}^{k+1} + \mathbf{u}_{drift}^{k+1}, \quad (4.23)$$

where \mathbf{u}_0 is the DM command obtained from the final iteration of PPSM, $\mathbf{u}_{control}^{k+1}$ is the total closed-loop DM command, and \mathbf{u}_{tot}^{k+1} is the total command sent to the DMs including the drift. For all experiments discussed in this chapter, both BMC DMs are used for control.

MULTI-WAVELENGTH DZM

Performing DZM at multiple wavelengths requires only a minor adjustment to the algorithm. For each iteration, an image is taken at each of the l wavelengths of interest using the narrowband filters described in App. A. The same DM command is used for all wavelengths. The EKF is run on each wavelength, individually producing an estimate of the electric field at each wavelength ($\hat{x}^k(\lambda_i)$). Note that the Jacobians (G), observation matrix (H^k), and covariance matrices (Q^k, P^k, R^k) are now all wavelength dependant. An EFC command is then calculated for each wavelength using the EKF output and the average command is used for control as shown by

$$\Delta \mathbf{u}_{opt}^{k+1} = \frac{1}{l} \sum_{i=1}^l \mathbf{u}_{opt}^{k+1}(\lambda_i), \quad (4.24)$$

where $\mathbf{u}_{opt}^{k+1}(\lambda_i)$ is the EFC output for λ_i . If desired, weights can be applied to prioritize certain wavelengths. Note that I plan on moving to the standard broadband EFC formulation described in Give'on et al. ³⁴ and Sec. 2.3.3 for future experiments.

ESTIMATE ERROR

The HiCAT hardware experiments take an open-loop image every k_{ol} iterations to track the open-loop contrast and the error in the estimate of the state variable (\hat{x}). Note that the state variable error must use the intensity error as a proxy as we cannot measure the electric field. The error in the state estimate for iteration k is described by

$$\varepsilon_{OL}^k = |I_{OL}^k - \hat{I}_{OL}^k|, \quad (4.25)$$

where I_{OL}^k is the open-loop image and \hat{I}_{OL}^k is the estimate of the open-loop intensity. The open-loop intensity estimate and BMC DM command used to acquire the open-loop image are calculated as

$$\hat{I}_{OL}^k = \hat{x}_{\Re}^k \circ \hat{x}_{\Re}^k + \hat{x}_{\Im}^k \circ \hat{x}_{\Im}^k, \quad (4.26)$$

$$u_{OL}^k = u_0 + u_{drift}^k, \quad (4.27)$$

where \hat{x} is calculated in Eqn. (4.17) and u_0 is the initial DM command obtained from the PPSM algorithm as discussed in Sec. 4.3.1. Note that due to the way the state vector is assembled, it follows that $\hat{x}_{\Re}^k = \hat{x}^k[0, 2, \dots, 2n - 2]$ and $\hat{x}_{\Im}^k = \hat{x}^k[1, 3, \dots, 2n - 1]$.

Taking periodic open-loop images provides us with a more accurate indication of the estimator performance as it allows us to directly compare the estimate of the state to a measurement as shown by Fig. 4.3. In an on-sky application, the closed-loop estimate error would be used as a proxy for the error in the state estimate and open-loop images would not be acquired. For Fig. 4.2, $k_{ol} = 1$ but for most experiments I use $k_{ol} = 5-25$ to remove

the initial contrast degradation shown by the spike in the magenta crosses curve in Fig. 4.2. Figure 4.3 shows the open-loop intensity image and estimate for iteration 26 of the DZM experiment described in Sec. 4.4.4. Here, iteration 26 is where $\beta = 1$ and the controller is turned on. The image in Fig. 4.3 (left panel) is taken prior to the EFC correction. The mean DZ intensity estimate error is reduced from 3.37×10^{-8} to 1.16×10^{-8} by iteration 26. As shown by the right panel of Fig. 4.3, the main structures in the image have been captured by the estimate including the hot spots near the inner working angle (IWA) and the speckle structure.

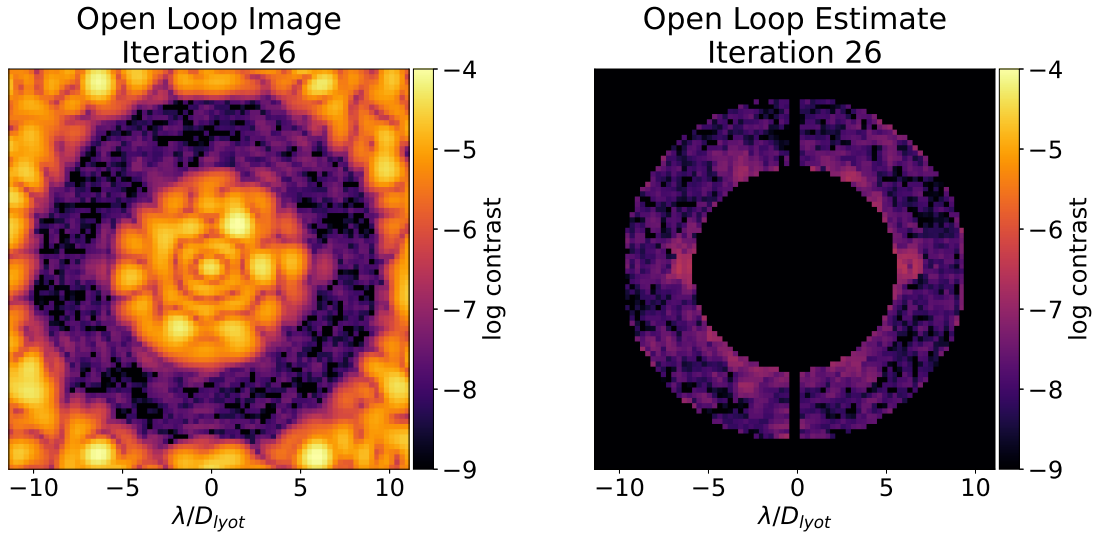


Figure 4.3: Log contrast open-loop image and estimate for iteration 26 of the HiCAT DZM experiment described in Sec. 4.4.4 where the IrisAO DM segments are drifting in piston, tip, and tilt. The controller is off for the first 25 iterations to allow for the estimator to converge. By iteration 26, the estimate captures the magnitude and structure of the open-loop image. Note that the DM is drifting while the controller is off.

4.4 HiCAT LABORATORY RESULTS

In this section I discuss the experimental results obtained on the HiCAT testbed at STScI. I first describe the dither effect on the raw contrast and then walk through four DZM experiments. Supplemental information related to HiCAT-specific hardware and procedures is provided in Appendix A. A more detailed discussion of dither tuning for BMC random-walk drifts is included in Appendix B.

4.4.1 OPTIMIZING THE DITHER AMPLITUDE

Adding a random dither to a DM command degrades the contrast as shown in Fig. 4.4. Here the dotted lines are produced using the HiCAT simulator⁸¹ and the solid lines are produced using the HiCAT testbed. The DM command, u_0 , producing a high-contrast dark zone ($\sim 6 \times 10^{-8}$) is used at the start and then a random realization of the dither normal distribution is added to it for each iteration. Note that this random command does not accumulate and, as a result, the contrast degradation is controlled (via the amplitude of the dither) and approximately constant for a given dither magnitude. As shown by the green square, blue diamond, and purple dotted lines in Fig. 4.4 the simulation and lab data agree very well. For the red triangle line, where the dither amplitude is the smallest, the initial contrast on the testbed was too low (‘bad’) for a significant change in the contrast to be induced by the dither.

We can see in Fig. 4.4 that the contrast degrades more for a larger dither magnitude; we are in the small aberration regime where the contrast scales with the square of the wavefront error. The dither–contrast relationship can be modelled using a second-order polynomial by averaging the mean DZ contrast values over the 50 iterations. For $\sigma_{dither} = 0$, there

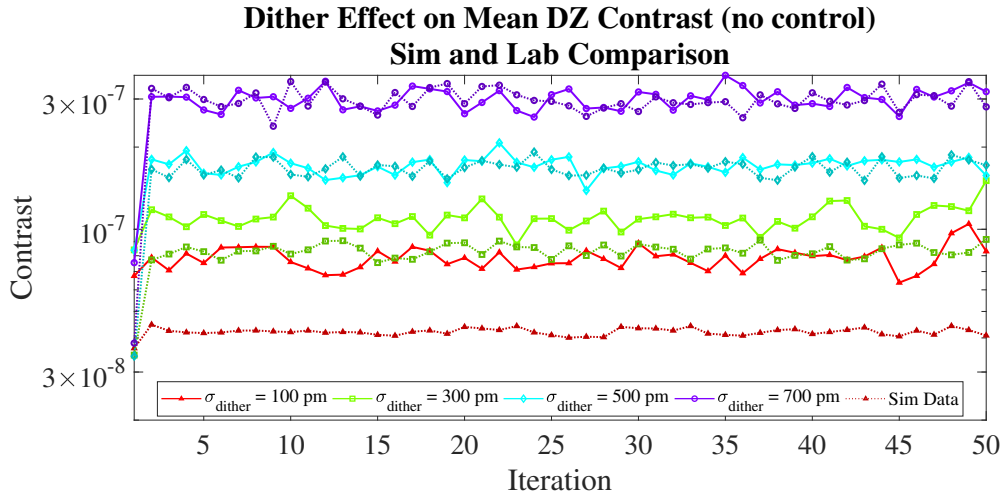


Figure 4.4: Comparing simulated and hardware data for the effect of varying DM dither standard deviations (σ_{dither} [pm]) on the mean contrast in the dark zone. Solid lines are hardware data and the dashed lines are simulated data. For the red triangle hardware curve the initial contrast was not sufficiently high to see the effect of the dither.

is no dither-induced contrast change. To fit the model, the contrast for $\sigma_{dither} = 0$ is taken to be the average initial contrast. The data and models for simulated and hardware experiments are shown in Fig. 4.5 where the data points include error bars to show the standard deviation of the contrast in the dark zone region over the 50 images. As shown in the left two panels of Fig. 4.5, the models are well matched for both the simulated and lab data sets. In the right panel of Fig. 4.5 we see that the hardware and the simulation models agree well at larger dithers but deviate at the smaller dithers. At small dithers, the contrast change is commensurate with the raw testbed contrast and thus there is no dither induced modulation as shown by the red triangle curves in in Fig. 4.4.

The optimal dither for a DZM run is dependent on the magnitude and speed of the drift and the starting contrast (final PPSM contrast). For lower (worse) contrasts, a larger dither is required to induce enough phase diversity for the estimator to converge. Similarly for larger drifts, a larger dither is required to ensure the majority of the phase diversity is

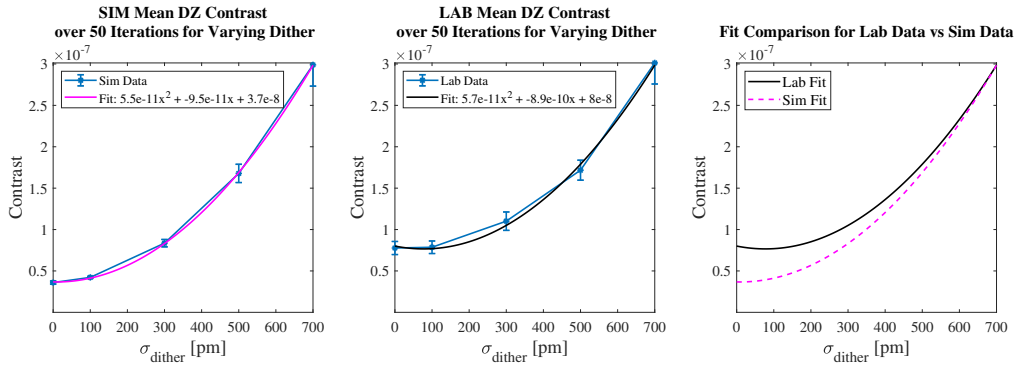


Figure 4.5: Second order model of the mean dark-zone contrast with respect to the dither for the HiCAT simulator and hardware. The second-order fit produces different equations for the lab and sim data due to the contrast floor on the testbed (right panel). Data points for the simulated and hardware experiments include error bars to show the standard deviation of the contrast in the dark zone region over the 50 images; note that the error bars increase with dither magnitude which is expected. At smaller dithers, the hardware does not reliably produce high enough contrasts to capture the behaviour.

coming from the dither and not the drift. In the experiments where I use the BMC DMs, I inject mid-to-high spatial frequency continuous wavefront drifts (defined in Eqn. 4.4) and can determine the optimal drift–dither relationship as covered in Appendix B. This is because the spatial-frequency content of the drift and dither is similar. However, when synthetic segment level drifts with the IrisAO are added, this relationship becomes harder to characterize. For those experiments, the dither is tuned using the smallest drift that produces a noticeable change in the open-loop electric field in the span of 30 mins. An initial dither is selected as the dither value that produces the final contrast from the PPSM algorithm via the left panel in Fig. 4.5. If the estimator performs poorly, the dither is increased until it is stable. From there, if the drift changes, the dither is scaled with it. Our numerical model of HiCAT is sufficiently precise, see Appendix B, that the dither can often be tuned in simulation, only requiring minor adjustments when I transfer to the testbed. This ad-hoc dither tuning for segment-level drifts has proved effective so far in our experiments,

however I am further investigating more efficient approaches.

4.4.2 DZM WITH BMC RANDOM WALK DRIFT WITHOUT SEGMENTED APERTURE

I first demonstrate DZM of high-spatial frequency drifts, injected using both of HiCATs Boston Micromachine DMs. In an observatory, such drifts can occur because of optical misalignments from thermal effects or electronic drifts in actuated mechanisms. As discussed in Sec. 4.3, I create these synthetic drifts by applying BMC DM actuator commands that follow a random walk with $\sigma_{drift}^2 = 100 \text{ pm}^2/\text{iter}$. The dither is chosen, using the method in Appendix B, to be $\sigma_{dither} = 200 \text{ pm}$. A single exposure is taken for each image making the total exposure time for each iteration 0.3 s; due to testbed overhead, iterations are $\sim 5.6 \text{ s}$. By the end of the experiment, the BMC DM drift command for each DM has an RMS of 0.63 nm and a PV of 2.20 nm.

Figure 4.6 shows the closed-loop mean contrast of the dark-zone region in magenta for the duration of the experiment (6.25 hrs, 4000 iterations). Note there is a sharp peak in the closed-loop contrast in the beginning of the experiment. Here, the control gain (β) is one for all iterations and a very aggressive (small) Tikhonov parameter (α) is used. This combination of β and α values in Eqn. 4.21 causes the contrast to temporarily degrade while the estimator is converging. In the following sections, this effect is avoided by setting $\beta = 0$ for the first 10–20 iterations and choosing a more conservative α . Once the estimator is converged, the closed-loop contrast is stable around the initial value. The dotted black line in Fig. 4.6 is the mean of the magenta curve (closed-loop contrast) and the dashed black line is the standard deviation of the magenta curve values. To track the drift at the focal plane, six images were taken throughout the experiment with the open-loop DM command

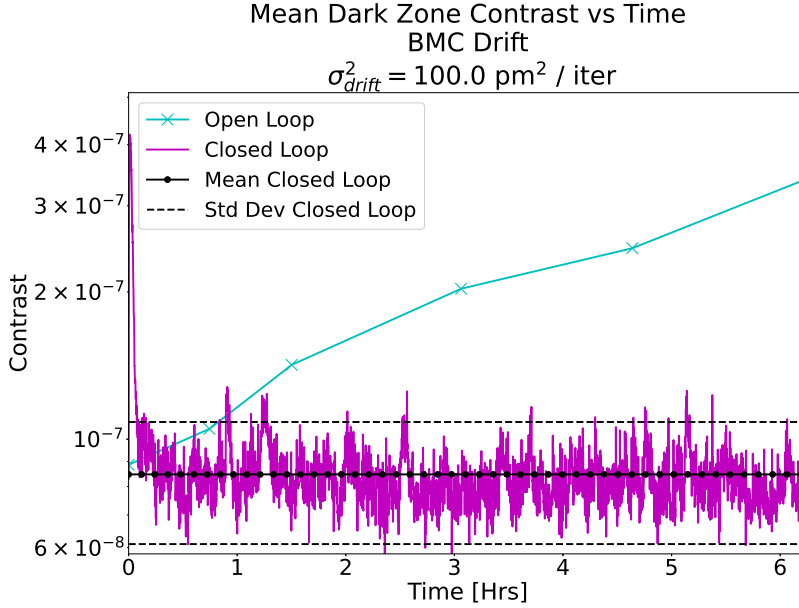


Figure 4.6: DZM experiment for BMC random walk drift and no segmented aperture. The spatially averaged contrast of the dark zone is shown in magenta with its mean and standard deviation shown by the dotted and dashed black lines, respectively. The blue crosses are the mean contrast in the dark zone for the open-loop DM command. The open-loop mean contrast degrades by a factor of four while the closed-loop mean contrast remains at the initial value of 8.5×10^{-8} within a standard deviation of 2.4×10^{-8} .

as shown by the cyan crosses in Fig. 4.6. The cyan crosses represent what an observation period would look like if a constant DM command was used for all science images and a drift was present in the system. The DM command used for the open-loop images is solely the initial DM command obtained from PPSM and the accumulated drift command as described by Eqns. 4.4 and 4.27. Over the course of the experiment, the open-loop mean contrast degrades by a factor of four while the closed-loop mean contrast remains at the initial value of 8.5×10^{-8} within a standard deviation of 2.4×10^{-8} . Using this data with the post-processing method discussed in Pogorelyuk et al.⁸⁸, we can further reduce the contrast to the level of the standard deviation.

The Extended Kalman Filter is a nonlinear estimator that linearizes about each time step.

Since the DM command, and thus the estimate, are continuously accumulating, this can lead to large linearization errors. A way around this is to shift the DM command and electric field estimate periodically throughout the experiment so that

$$\hat{E}_0 = \hat{E}^{k_r} \quad (4.28)$$

$$u_0 = u^{k_r}, \quad (4.29)$$

where k_r is the estimator reset iteration. From Fig. 4.6 it is shown that the linearization errors are not sufficiently large to cause issues when the open-loop contrast is four times the closed-loop contrast. When I start pushing the limits of the DZM algorithm to determine the largest correctable drift or when using DZM with ADI (to be discussed in Ch. 6), I may need to consider performing estimator resets.

4.4.3 BMC DRIFT AT MULTIPLE WAVELENGTHS

Building upon previous BMC drift DZM results, I perform a similar DZM experiment with the BMC DMs performing a random walk drift while using a broadband laser source. The slight algorithmic changes required for multi-wavelength DZM are described in Sec. 4.3.2. For this experiment the segmented aperture is installed but is not used for drift injection. Note that the HiCAT testbed is not optimized for broadband performance, limited to a contrast of 3×10^{-7} even when monochromatic FPWSC is performed using the broadband source and a single narrowband filter. Due to limitations when using the broadband laser source I have not yet attempted true broadband DZM. For this experiment, I use the 640 nm and 660 nm filters which result in a 3% band spanning 30 nm. At each wavelength,

two exposures are averaged for each image resulting in a net exposure time of 0.1 s per wavelength and a total iteration time of 30 s.

Operating at 10^{-7} for this experiment, I use larger dither and drift values of $\sigma_{dither} = 300$ pm and $\sigma_{drift}^2 = 900$ pm²/iter respectively. The control gain β is zero for the first 10 iterations and then set to one. There is no spike in contrast at the beginning of the experiment so 10 iterations is long enough for the estimator to converge, yet short enough so the open-loop electric field does not significantly drift during that time. The final BMC drift command for each DM has an RMS of 0.93 nm and a PV of 6.15 nm.

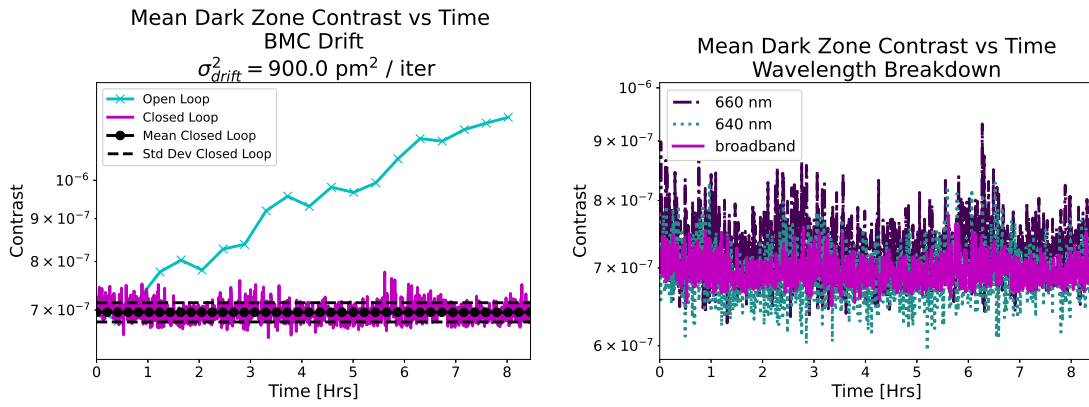


Figure 4.7: Mean DZ contrast as a function of time for multi-wavelength DZ maintenance in the presence of BMC random walk drift. The left plot shows the open- and closed-loop broadband contrasts, with the cyan crosses and magenta curves respectively, for the duration of the experiment. The mean closed-loop broadband contrast is maintained at 7.0×10^{-7} (black dots) with a standard deviation of 1.8×10^{-8} (black dashes). The open-loop broadband contrast degrades by a factor of 1.72 to 1.2×10^{-6} . The right plot shows the closed-loop contrast vs time for the wavelengths controlled. The 660 nm contrast (mauve dashes) is slightly worse than the 640 nm contrast (teal dots) but they are within 5% of each other. For reference, the broadband contrast is provided in the right plot in magenta.

Figure 4.7 shows the contrast vs. time plots for multi-wavelength dark zone maintenance where the BMC DMs are drifting. As shown in the left panel of Fig. 4.7, a mean broadband dark zone contrast of 7.0×10^{-7} is maintained for 8.5 hrs with a standard deviation of 1.8×10^{-8} . Certainly, operating at 10^{-7} is part of what makes the small standard deviation

possible as the other drifts in the system (such as temperature, humidity, and air turbulence) have less of an effect at this level. The open-loop broadband contrast increases by a factor of 1.72 to 1.19×10^{-6} by the end of the experiment. Figure 4.7 (right panel) shows the wavelength breakdown for the closed-loop mean dark-zone contrast during the experiment. Here we can see that the 640 nm contrast is slightly better than the 660 nm contrast but they are within 5% of each other. Overall, these results illustrate what I expect from previous theoretical and numerical work: multi-wavelength DZM operates at the same level of performance as monochromatic DZM and the ultimate contrast obtained is commensurate with a given instruments/testbeds static chromatic behavior. Hardware upgrades to improve the broadband contrast on HiCAT are in progress. Due to potential logistical upgrade delays as well as the range of experiments already performed, it was decided that waiting for full 10^{-8} broadband contrast DZM results was unnecessary for this chapter.

I also analyze the spatial contrast distribution in the dark zone. To do so, I average over all 1000 images from our multi-wavelength experiment to mitigate for the intrinsic noise in our short exposures. I obtain the temporally averaged broadband image and distribution shown in Fig. 4.8. There is a clear residual of the six-fold spoke pattern created by the IrisAO segmented aperture inside the dark zone; this defect is mainly seen when the broadband laser source is used and is a limitation for the contrast. In the right panel of Fig. 4.8 I provide the histograms for the temporally averaged images for each wavelength as well as the combined broadband case; note that a log scale is used for the x-axis. The characteristics of these histograms are provided in Table 4.3 which shows an equal level of DZM across the band.

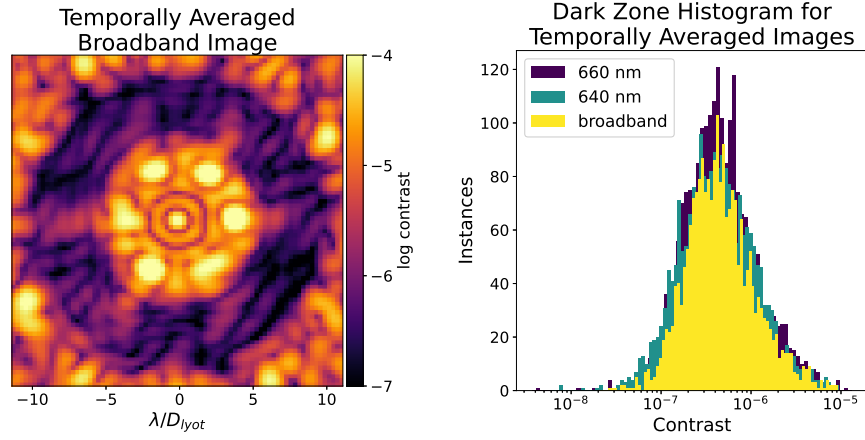


Figure 4.8: Temporally averaged broadband image and associated histograms. The left panel shows the temporally averaged broadband image for all 1000 iterations where D_{Lyot} is the diameter of the Lyot stop. There is a distinct residual of the six-fold spoke pattern induced by the IrisAO which is the main source of the lower contrast speckles and positively skewed distributions. This residual pattern is primarily present on HiCAT when the broadband laser source is used and I are conducting hardware upgrades to mitigate this phenomenon. The right panel shows the histograms for the 660 nm, 640 nm, and broadband temporally averaged images. Note that this histogram has a log scale on the x-axis. All histograms are very similar, showing that all wavelengths are being equally maintained in the broadband DZM algorithm.

Table 4.3: Temporally averaged image statistics for DZM with a drift on the BMC DMs. Note that the DZ contrast oscillates around the mean but does not have a significant trend (as shown in Fig. 4.7); thus, the initial contrast (pre-drift) is on par with the mean (provided in column two). Also, the standard deviations provided in this table are for the spatial distributions unlike the previous values quoted which were for the temporal distributions.

Wavelength	Mean Contrast in DZ	Std Dev of Contrast in DZ	Skew of Contrast Distribution
660 nm	7.2×10^{-7}	9.9×10^{-7}	4.62
640 nm	6.9×10^{-7}	9.1×10^{-7}	4.11
broadband	7.1×10^{-7}	9.1×10^{-7}	4.13

4.4.4 DZM WITH IRISAO PISTON TIP TILT DRIFT

The IrisAO (Appendix A.2.2) PTT drift is aimed at emulating the drift environment of future large segmented telescopes in space such as LUVOIR.⁹³ I present a monochromatic experiment where the piston, tip, and tilt of each IrisAO segment drift independently. For this experiment, the drift in each DOF has approximately the same weight with $\sigma_p^2 = 25 \text{ pm}^2/\text{iter}$ and $\sigma_{Tt}^2 = 25 \text{ nrad}^2/\text{iter}$. Note that the flat-to-flat distance of each IrisAO segment is 1.212 mm thus a 5 nrad tip or tilt causes a $\pm 3 \text{ pm}$ displacement of the segment edge. A pinhole was temporarily placed after the laser launch to try and filter some of the fiber effects (Appendix A.2.2). This reduced the throughput by a factor of five thus requiring larger exposure times. To achieve the desired signal-to-noise ratio (SNR), each image consists of 16 exposures with the total exposure time being 0.48 s and the iteration time being $\sim 30 \text{ s}$. Note, a relatively high SNR was chosen for this experiment. The final IrisAO drift command has a surface RMS of 0.19 nm and a surface PV of 1.06 nm. Under this configuration, the contrast after Dark Zone digging is 2.5×10^{-8} and the magnitude of the drifts are commensurate with estimates of primary mirror stability for the Roman Space Telescope.⁴⁹ The dither used is $\sigma_{dither} = 100 \text{ pm}$ and was chosen as outlined in Sec. 4.4.1.

Figure 4.9 shows the mean dark-zone contrast evolution during the experiment. The experiment runs overnight for 9.5 hrs maintaining a closed-loop contrast of 2.5×10^{-8} (black dots) within a standard deviation of 5.5×10^{-9} (black dashes). The open-loop contrast drifts by a factor of 4.75 to 1.2×10^{-7} (cyan crosses) during the experiment. The control gain β is set to zero for the first 25 iterations and then set to one for the duration of the experiment. As shown by the first 10 mins of the plot in Fig. 4.9, there is no large overshoot as the estimator converges and the open-loop electric field does not significantly drift while

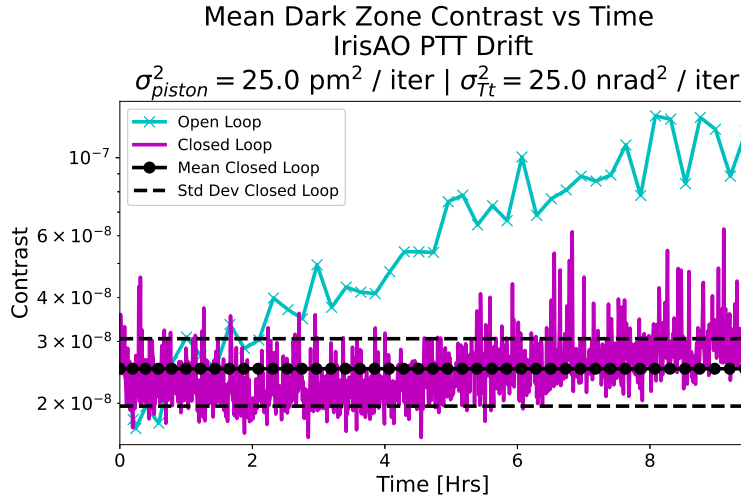


Figure 4.9: DZM with an IrisAO PTT drift. The mean contrast of the dark zone is shown in magenta with its mean and standard deviation shown by the dashed and dotted black lines respectively. The cyan crosses mark the mean contrast in the dark zone for the open-loop DM command. The open-loop mean contrast degrades by a factor of 4.75 while the closed-loop mean contrast remains at the initial value of 2.5×10^{-8} within a standard deviation of 5.5×10^{-9} .

the controller is off. The closed-loop contrast (magenta curve in Fig. 4.9) is extremely stable for the first 4.5 hours and then begins to deviate slightly as the open-loop contrast becomes significantly larger.

In this experiment we are operating at starlight suppression levels three times fainter than in Sec. 4.4.2. We are now observing some environmental effects that the DZM algorithm can only partially correct as it is only tuned for our synthetic drifts. If I split the experiment into two blocks (0–4.5 hrs and 4.5–9.5 hrs) we see that the mean contrast increases by 20% from 2.3×10^{-8} to 2.7×10^{-8} after 4.5 hrs and the standard deviation increases by 54% from 3.8×10^{-9} to 5.9×10^{-9} . This increase in contrast is likely due to a combination of an overly-cautious Tikhonov parameter (α) when determining the EFC correction and a humidity drift (Appendix A.2.4). In future experiments I will loosen the EFC Tikhonov parameter to allow for a more aggressive DM correction. As mentioned in Sec. 4.3.2, errors

in the drift rate have not been proven to have a negative effect on the performance provided the model is correct. The reason we see a drift leakage here is predominately because the humidity changes induce a low-order drift (tip-tilt) that is very different to the DM actuator random-walk drift model provided to the EKF. In the future, low-order drifts will be corrected using the low-order wavefront control loop to be implemented on HiCAT.

4.4.5 DZM FOR SIMULTANEOUS BMC AND IRISAO DRIFT

Finally, I study the effect of two different simultaneous drifts with both BMC DMs drifting ($\sigma_{drift}^2 = 100 \text{ pm}^2/\text{iter}$) as well as the IrisAO ($\sigma_p^2 = 0.64 \text{ pm}^2/\text{iter}$, $\sigma_{Ti}^2 = 0.64 \text{ nrad}^2/\text{iter}$). The dither used is $\sigma_{dither} = 150 \text{ pm}$ and since the drift is predicted to be relatively fast, the control gain β is only set to zero for the first five iterations. This is likely too few iterations as we see a small spike in the closed-loop contrast in Fig. 4.10 when the controller turns on while the estimator is still converging. Unlike in Sec. 4.4.4, the number of exposures is capped at three for this experiment. With three exposures per image, the total exposure time is 0.024 s and the total iteration time is 15 s. It should be noted that the pinhole is removed for this experiment which aides in reducing the exposure time. The final BMC drift command for each DM has an RMS of 0.40 nm and a PV of 2.86 nm. For the IrisAO, the final drift command has an RMS of 0.04 nm and a PV of 0.28 nm. The final command is near the resolution of the IrisAO and in the future, larger IrisAO drifts should be used so that the covariance matrices more accurately represent the drift.

As shown by the contrast vs time plot in Fig. 4.10, the open-loop contrast drifts by a factor of 9.25 reaching 5.5×10^{-7} after 1500 iterations or 7 hrs. The spatially averaged closed-loop contrast has a mean of 5.9×10^{-8} and a standard deviation of 1.1×10^{-8} over the

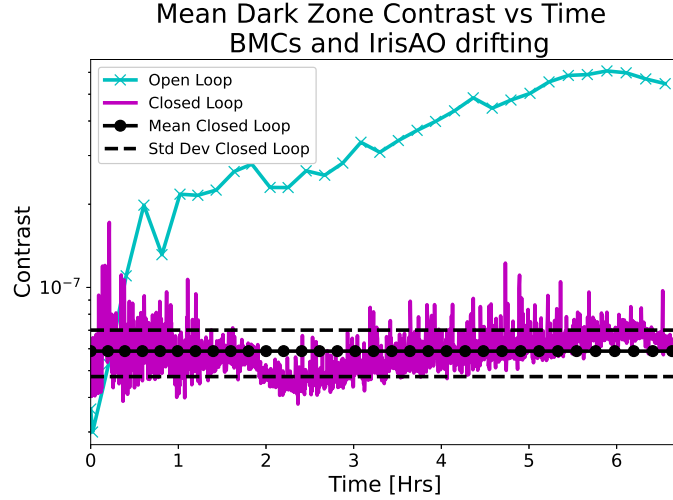


Figure 4.10: DZM with BMC and IrisAO PTT drifts. The mean contrast of the dark zone is shown in magenta with its mean and standard deviation shown by the dotted and dashed black lines respectively. The cyan crosses mark the mean contrast in the dark zone for the open-loop DM command. The open-loop mean contrast degrades by a factor of 9.25 while the closed-loop mean contrast remains at the initial value of 5.9×10^{-8} within a standard deviation of 1.1×10^{-8} .

entire experiment. This performance is slightly worse than the experiment in Sec. 4.4.4 for a number of reasons. When applying the final DM command from the PPSM algorithm, HiCAT reliably resumes at an average of 6×10^{-8} but with significant variance around that operating point. In Sec. 4.4.4, HiCAT resumes from the PPSM at a deep contrast of 2×10^{-8} and thus the mean contrast is deeper. Moreover, I suspect the Tikhonov parameter is too conservative for this experiment which allows some of the synthetic drift to leak into the closed loop contrast.

4.5 IACT LABORATORY RESULTS

The In-Air Coronagraph Testbed (IACT)^{9,109} at the Jet Propulsion Laboratory uses a single 2k BMC DM in the pupil plane combined with a vector vortex coronagraph to create a one-sided dark zone. The broadband light source is a NKT SuperK Extreme laser coupled

to a photonic crystal fiber. ¹⁰⁹

Here I present the results of testing the DZM algorithm on IACT. This experiment has a very similar setup to that described in Ch. 4.4.3 it is just executed on a different testbed at a higher contrast. I use a 10% band centred at 635 nm and sample the band using three wavelengths. The single DM on IACT performs a random walk drift at a rate of $1.1025 \times 10^{-10} \text{ V}^2/\text{iteration}$ and a dither of $8 \times 10^{-5} \text{ V}$ is used. IACT requires much longer exposure times and each iteration is approximately 225 s.

The left panel of Fig. 4.11 shows the mean dark zone contrast throughout the experiment for the open loop (dashed crosses) and closed-loop cases (solid). Each wavelength is plotted individually to show that the electric field is suppressed uniformly across the band. We can see that the initial contrast for the open-loop curves is higher than the closed-loop curves for all wavelengths; this is the dither slightly degrading the contrast. By the end of the 21.8 hr experiment, the open-loop contrast has degraded (on average) by a factor of 2.8. The right panel of Fig. 4.11 shows the final dark zone averaged over the three wavelengths sampled.

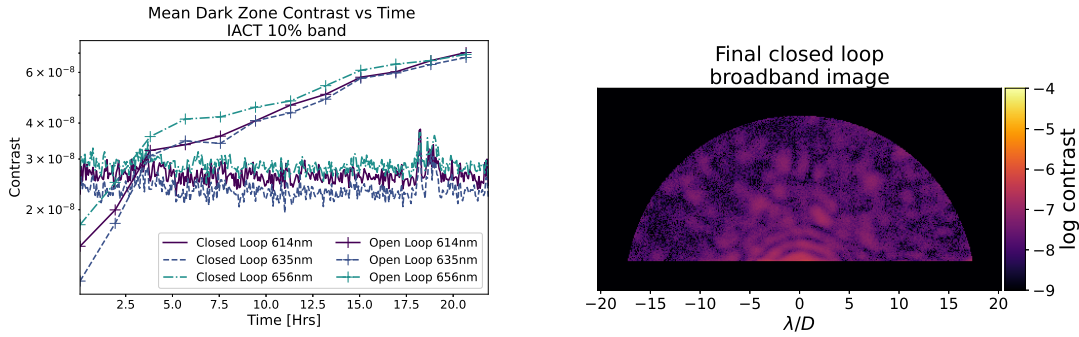


Figure 4.11: Dark zone maintenance over a 10% band centred at 635 nm on the IACT testbed at JPL. Three sub-bands were imaged to uniformly suppress the electric field over the 10% band as shown by the left panel. The open-loop contrast is shown by the crosses and degrades by a factor of two during the experiment. The closed loop contrast remains constant and the light is suppressed uniformly across the different wavelengths. The discrepancy between the initial open-loop and closed-loop contrast is due to the dither added to the deformable mirrors to aid the estimator; this is a known effect and is small enough that it does not negatively impact the science. This experiment was 350 iterations and took over 20 hrs. The right panel shows the final closed-loop image averaged over the three sub-bands.

The DZM experiment run on IACT shows that the algorithm is robust for different types of coronagraphs as well as different overall testbed configurations. This is extremely important as there are multiple coronagraphs on RST and there are many options being considered for HWO. Having an algorithm that is insensitive to the hardware allows for more applications and allows for development when the hardware state is not yet finalized (HWO).

4.6 CONCLUSIONS

4.6.1 DISCUSSION

Future space missions will need a reliable method for dealing with quasi-static high-order WFE drifts. In this chapter I demonstrate the DZM capabilities in a variety of situations as summarized in Table 4.4. These results are very promising and have the potential to

greatly facilitate the design and operation of future high-contrast space missions. I start by maintaining a contrast of 8.5×10^{-8} with a monolithic aperture in the presence of a BMC DM random walk drift ($\sigma_{drift}^2 = 100 \text{ pm}^2/\text{iter}$) as described in Sec. 4.4.2. In order to do spectroscopy on directly observed exoplanets, the contrast must be consistent over a wide wavelength range. HiCAT is still developing its broadband capabilities so we are restricted to multi-wavelength control at contrasts lower than 10^{-7} . In Sec. 4.4.3 I use DZM to correct for BMC random walk drift at a rate of $\sigma_{drift}^2 = 900 \text{ pm}^2/\text{iter}$ and maintain the contrast of a 3% band at 7.0×10^{-7} for 8.5 hrs. I also provide the wavelength-dependant statistics for the temporally averaged dark zone demonstrating that the closed-loop contrast is uniformly maintained. Large segmented aperture telescopes introduce new challenges that HiCAT aims to address. Using the IrisAO segmented aperture DM, in Sec. 4.4.4 I demonstrate the ability to maintain a contrast of 2.5×10^{-8} in the presence of a PTT random walk drift with a rate of $\sigma_p^2 = 25 \text{ pm}^2/\text{iter}$ and $\sigma_{Tt}^2 = 25 \text{ nrad}^2/\text{iter}$. Space telescopes will have multiple drift sources with different temporal and spatial frequencies. To represent this effect, Sec. 4.4.5 shows the results for an experiment where the BMC DMs drift at a rate of $\sigma_{drift}^2 = 100 \text{ pm}^2/\text{iter}$ and the IrisAO drifts at a rate of $\sigma_p^2 = 0.64 \text{ pm}^2/\text{iter}$, $\sigma_{Tt}^2 = 0.64 \text{ nrad}^2/\text{iter}$. The closed-loop contrast is maintained at 5.9×10^{-8} . In Table. 4.4 I summarize the experiments performed and translate the results into drift variances and RMS rates for the proposed LUVOIR-B architecture. The reported RMS rates in Table 4.4 are meant to provide a more intuitive representation of the 1/f drifts to the reader. To do so, I performed a linear fit to the cumulative RMS of our injected drifts as a function of time. In the case of our experiments, I found that we are indeed remaining in the linear regime of the RMS growth and that this linear approximation is overall representative of

the open-loop behavior on HiCAT.

Since our drifts are implemented as discrete increments occurring at each iteration, we first need to estimate the projected LUVOIR-B exposure times equivalent to our images. To do so I compare each experiment's dark zone speckle statistics with simulated images generated using the LUVOIR simulator in Juanola-Parramon et. al 2019.⁴² The `pastis`^{53,55} package is used to simulate the level of contrast commensurate with what we see in the lab and the `exoscene`¹⁴⁷ package is used to convert contrast into photons. I simulate using LUVOIR-A but the primary mirror is scaled down to 8 m. I assume a static wavefront error (RMS ranging from 0.26–1.22 nm) yielding a steady-state contrast commensurate with the one of each HiCAT experiment. A $m_V = 5$ star is used and I adjust the exposure time until the simulated speckle statistics (mean and variance) match the experimental ones. A sample set of images to demonstrate the final result of this process is provided in Fig. 4.12 for the experiment where the IrisAO is drifting (Sec. 4.4.4). The right two panels of Fig. 4.12 show the HiCAT image and the final simulated LUVOIR image. The left panel of Fig. 4.12 shows the histograms for the HiCAT experiment (hollow black) and the simulated LUVOIR image (solid blue); note that the histograms are normalized such that the sum of the bin areas is one and the histograms only consider the HiCAT dark zone. In order for the speckles statistics of the simulated LUVOIR image (right panel) to match the experimental HiCAT image (middle panel) the exposure time for the LUVOIR image must be 5.5 mins.

Note when converting to LUVOIR, I adjust the HiCAT tip/tilt drifts such that the segment edge displacement behaviour is the same. The same angle applied to a LUVOIR segment would produce a much larger WFE compared to the IrisAO. For example, 5 nrad

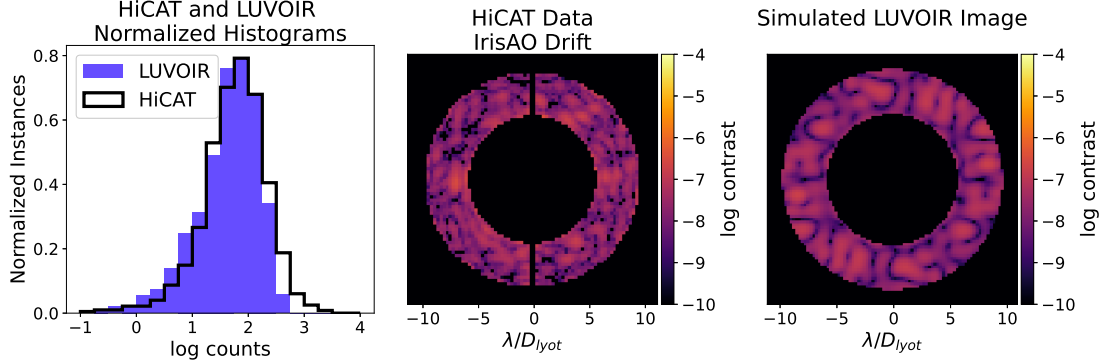


Figure 4.12: HiCAT and LUVOIR histograms and contrast images for the segmented aperture drift experiment (Sec. 4.4.4). The static WFE applied to the LUVOIR primary is scaled such that the mean dark zone contrast of the middle and right panel is the same. The exposure time for the LUVOIR image (right panel) is then adjusted until the histograms in the left panel have approximately the same mean and standard deviation; the solid blue histogram is for the LUVOIR image and the hollow black histogram is for the HiCAT image. For the segmented aperture drift experiment shown here, an exposure time of 0.48 s on HiCAT translates to an exposure time of 5.5 mins for LUVOIR.

tip on an IrisAO segment would produce a ± 3 pm displacement of the segment edges as the flat-to-flat distance of a segment is 1.212 mm and the segment rotates about the center. Considering a LUVOIR segment instead, with a flat-to-flat distance of 0.955 m¹², the edge displacement of the segment would be 2.38 nm; an order of magnitude larger. Since we are in the small angle regime, the conversion between edge displacement and segment angle is linear and $Var(rX) = r^2 Var(X) = r^2 \sigma_{Ti}^2$ where r is the ratio of the IrisAO and LUVOIR flat-to-flat distance.

For all monochromatic experiments, contrasts deeper than 10^{-7} , our experiments are much longer (99 to 600+ hours) than a typical LUVOIR-B sequence. This means that our experiments are operating in the “photon rich regime”, that is not representative of how DZM algorithms are envisioned to operate on future exoplanet imaging instruments.⁸⁸ As a result, the LUVOIR-B equivalent exposure time is unrealistically large, and the projected LUVOIR-B surface RMS drift rates are on the order of 0.1 picometer per minute; this still

places challenging constraints on stabilizing the observatory. Future work focusing on designing DZM experiments with photon statistics and exposures commensurate with actual operations on future mission will further the present proof of concept to actually establishing by how much continuous wavefront sensing and control can relax observatory level requirements.

Table 4.4: DZM experiment summary. The HiCAT drift variances are converted into predicted LUVOIR-B drift variances for the experiments with the IrisAO segmented aperture installed. These drift variances assume a Magnitude 5 star and the same contrast as in the corresponding HiCAT experiment. It is assumed that the exposure time is the limiting factor and there is no other overhead during the LUVOIR WFSC iterations

	Monochromatic BMC Drift	Multi-wavelength BMC Drift	Monochromatic IrisAO Drift	Monochromatic All DMs Drift
λ [nm]	638	640, 660	638	638
Initial μ_{DZ}	8.3×10^{-8}	7.2×10^{-7}	2.6×10^{-8}	5.4×10^{-8}
Final μ_{DZ}	8.9×10^{-8}	6.9×10^{-7}	3.0×10^{-8}	5.9×10^{-8}
HiCAT σ_{drift}^2 [pm ² /iter]	100	900	-	100
σ_p^2 [pm ² /iter]	-	-	25	0.64
σ_{Tt}^2 [nrad ² /iter]	-	-	25	0.64
Time [hrs]	6.24	8.46	9.45	6.72
LUVOIR σ_{drift}^2 [pm ² /min]	15.4	600.0	-	4.0
σ_p^2 [pm ² /min]	-	-	4.54	2.6×10^{-2}
$r^2 \sigma_{Tt}^2$ [nrad ² /min]	-	-	1.3×10^{-6}	1.6×10^{-9}
rms _{BMC} [pm/hr]	1.5	39.3	-	0.6
rms _{Iris} [pm/hr]	-	-	1.6	6.0×10^{-2}
Time [hrs]	433.3	25	98.5	660.8

A common theme shown in all the experiments presented in this chapter is that, on HiCAT, we are currently limited by the raw contrast of the testbed. I have shown that if we can achieve a certain contrast using the PPSM algorithm, we can maintain that contrast in the presence of a DM random walk drift. This will not always be true on HiCAT as the dither command I use is approaching the resolution of the 14-bit BMC DM driver. If we do not have good control over the magnitude of the dither, our estimate will likely suffer. In space or on a vacuum testbed, we will likely require a DM driver with a finer resolution

to maintain contrasts better than 10^{-8} . Also, at the deepest contrast (2×10^{-8}) we are near the environmental limit of the testbed, as discussed in Appendix A.2.4, and we see drifts associated with the humidity change leaking through. To determine the true limits of the DZM algorithm we need to move to a vacuum testbed such as the HCIT at JPL.

4.6.2 FUTURE WORK

The robust validation of DZM for future direct imaging space missions is a large task. This chapter provides a proof of concept using hardware experiments on the HiCAT testbed. It is important for the community to consider DZM as an option when designing missions such as HWO. The first step in furthering the study of the DZM algorithm will be to address the items causing sub-optimal performance discussed in this chapter. This will include improving the controller tuning and performing a humidity sensitivity study at a range of contrasts. The HiCAT broadband source has been upgraded since these experiments were run and I aim to re-produce these results at a wider wavelength band. For the broadband DZM, I also plan on investigating the broadband EFC method outlined in Give'on et al. 2007³⁴ to replace the average EFC method (Eqn. 4.24) currently used.

In addition to the minor upgrades, now that I have proved that DZM is possible in a range of drift scenarios, I can push the limits in each case to determine where it breaks down. This will include determining the maximum error in the Q matrix, the maximum correctable drift, and the effect of the exposure time. I will also study the dither-drift relationship for the IrisAO DM to determine the optimal dither when using the IrisAO DM to inject drift as was done with the BMC DMs in App. B. It is our goal to perform a subset of these experiments both on HiCAT and on a vacuum testbed to ensure that the limits we

find are not due to the air environment. I then aim to take the HiCAT DZM limitations and relate them to stability requirements for future space missions including RST and HWO. Chapters 5 and 6 continue the discussion of the DZM work and probe the effect of the SNR of the images as well as the effectiveness of various post-processing methods.

5

Low Signal-to-Noise Dark Zone

Maintenance

This chapter is taken from Redmond et al. 2022.⁹⁹ The experiments presented are from 2022 prior to the HiCAT hardware and software upgrades.

5.1 CHAPTER OVERVIEW

Directly imaging exoplanets requires long integration times when using a space-based coronagraphic instrument due to the small number of photons. Wavefront stability on the same timescale is of the utmost importance; a difficult feat in the presence of thermal and mechanical instabilities. In this chapter, I demonstrate that dark zone maintenance (DZM) functions in the low signal-to-noise (SNR) regime similar to that expected for the Roman Space Telescope (RST) and HWO recommended by the 2021 decadal survey. I develop low-photon experiments with tunable noise properties to provide a representative extrapolation. The experiments are performed on the High-contrast Imager for Complex Aperture Telescopes (HiCAT) at the Space Telescope Science Institute (STScI). High-order wavefront error drifts are injected using a pair of kilo-deformable mirrors (DMs). The drifts are corrected using the DMs via the DZM algorithm; note that the current limiting factor for the DZM results is the air environment. I show that DZM can maintain a contrast of 5.3×10^{-8} in the presence of DM random walk drift with a low SNR.

5.1.1 FUTURE SPACE TELESCOPE OBSERVING CONDITIONS

For the purpose of this Chapter, I will focus on the Roman Space Telescope (RST) as a case study for high contrast imaging in space. The current observing plan for RST involves slewing between bright reference stars and dimmer science target stars as described in Ch. 1.3.1. Starting on a bright reference star decreases the time required to find the DM command that corrects for the static wavefront error in the system and generates the dark zone. RST aims to have a mean dark-zone contrast between 10^{-8} – 10^{-9} (7) and thus must use focal plane wavefront sensing and control (FPWSC) when generating the dark zone

to avoid correcting for non-common path (NCP) aberrations⁸ introduced by pupil plane techniques. Once the dark zone is generated, RST slews to the target star and performs a series of observations at different roll angles for ~ 8 hrs while keeping the DM shape constant. It is predicted that after 8 hrs, the quasi-static high-order wavefront drifts will have degraded the dark zone^{8,124} requiring RST to return to the target star to regenerate the dark zone. These quasi-static drifts come from thermal gradients, mechanical relaxation of the structure, and the DMs themselves.

The RST Coronagraph Instrument (CGI) will be operating in the low signal regime with a photon rate around 10 milli-photons per second.⁷⁵ The individual coronagraphic exposures on RST will be relatively short (≤ 60 s⁵⁰) to avoid cosmic ray effects. Combining the low signal and short exposure times, we can see that RST CGI images will have a very low signal-to-noise ratio (SNR). The RST team chose an electron multiplying charge coupled device (EMCCD) detector due to its unique photon-counting mode which boosts the SNR by a factor of $\sqrt{2}$.¹⁰³ A comparison of the RST³ and HiCAT detector properties including dark current, read noise, and clock induced charge (CIC) is provided in Table 5.1. For reference, the detector design requirements for the Large Ultraviolet Optical Infrared (LUVOIR) Surveyor are also included.¹³⁸ The values for RST assume the electron multiplying (EM) gain is being used in photon counting mode.³ Note that HiCAT uses a complementary metal oxide semiconductor or CMOS detector which suffers from random telegraph noise (RTN) and not CIC.

To create an RST-like observing scenario on HiCAT, I start by injecting high-order wavefront error drifts to represent the quasi-static drifts seen during a target observation. In this chapter, I use the two BMC DMs to inject the drift, allowing each actuator to per-

Table 5.1: Detector properties for HiCAT, RST, and LUVOIR. Note that the LUVOIR properties are the required capabilities for a future-developed detector.

Instrument	Detector	Dark current [e-/hr]	Read noise [e-]	CIC [e-/pixel/frame]
HiCAT	ZWO ASI178MM	0	2.42	n/a
RST	T-e2v CCD311	0.76	120	0.01
LUVOIR	n/a	0.108	0	0.0013

form a random walk. I also scale the photon rate in the images to put HiCAT in the low signal regime like RST. Lastly, I apply a simple CCD detector noise model to the images and pass the adjusted images to the estimator. Manually injecting the noise allows us to take a short exposure (≤ 0.5 s) using a CMOS detector and make it look like a longer exposure (~ 60 s) from a CCD or EMCCD detector. In the future, it will also enable us to determine the minimum photon rate required in order for DZM to work. This low SNR conversion process is described in Sec. 5.2.1. The results of the low SNR DZM experiments are covered in Sec. 5.3 followed by the conclusions and plans for future work in Sec. 5.4.

5.2 DARK ZONE MAINTENANCE ALGORITHM ADJUSTMENTS

Low SNR DZM follows the same procedure described in Ch 4.3.2 with one additional step as shown by the light blue box in Fig. 5.1. The DM command (u_0) that corrects for the static WFE and generates the dark zone is determined using a pair-wise probe estimator and a stroke-minimization controller. This DM command is used as the starting point for DZM. With each DZM iteration, a dither command (δu_{dither}^k) is applied to the DMs which acts as a ‘probe’ and increases the phase diversity of the images between iterations to help the estimator converge. For the dither, a random command is applied to each actuator with

a standard deviation of σ_{dither} . I then take a single image and pass it through the low SNR conversion process (blue box in Fig. 5.1) which applies a scaling and injects noise to create a low SNR image with CCD noise properties (discussed in detail below). The resulting low SNR image is passed to the Extended Kalman Filter (EKF) which estimates the open-loop, drifting electric field (\hat{x}^{k+1}). The electric field conjugation (EFC) controller takes the estimate and determines the DM command (Δu_{opt}^{k+1}) required to correct for the open-loop electric field. Unknown to the control loop, the drift command (Δu_{drift}^{k+1}) is also applied to the DMs each iteration. The EKF is provided some statistical information on the drift (σ_{drift}) but not the exact drift command, as described in Ch. 4.3.2.

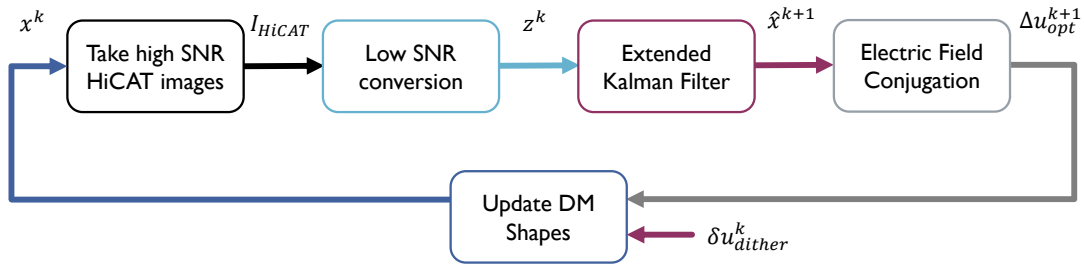


Figure 5.1: Simplified flow chart for low SNR DZM. The process is the same as that described in Redmond et al. 2022⁹⁶ with the addition of the light blue box which manipulates the HiCAT image (I_{HiCAT}) to produce a low SNR measurement (z^k) which gets passed to the EKF. The dither DM command (δu_{dither}^k) increases the phase diversity of the images between iterations to help the EKF converge. The EKF calculates the open-loop electric field estimate (\hat{x}^{k+1}) and passes it to the electric field conjugation (EFC) controller. The controller determines the required DM command (Δu_{opt}^{k+1}) to correct for the drifting electric field.

5.2.1 LOW SIGNAL-TO-NOISE CONVERSION

To emulate RST images using HiCAT, I first adjust the sampling of the focal-plane images. Without any binning, the sampling on HiCAT is 13.623 pixels per λ/D_{Lyot} , where D_{Lyot} is the diameter of the Lyot stop. Binning by a factor of six results in a sampling of

2.28 metapixels per λ/D_{Lyot} , or 1.75 metapixels per λ/D where D is the diameter of the HiCAT aperture, at 638 nm. This sampling is on par with the ~ 2.6 pixels per λ/D sampling expected for RST⁵⁰ and LUVOIR (where D is the aperture diameter of RST). The 6×6 binning is applied to the raw images prior to any other manipulation.

Next, I must adjust the photon rate in the image. The direct image of the target star in the RST CGI Observing Scenario (47 UMa, $V=5.04$, G1V, $D_{star}=0.9$ mas)⁵⁰ has a total photon rate of $\sim 18.5 \times 10^6$ ph/s. Note that this is the sum of the photon rate over the entire image. To get close to this value, I start by taking a direct image (I_{direct}) on HiCAT (without the Lyot stop in place) using the longest exposure time possible without saturating any pixels. This maximizes the signal-to-noise ratio (SNR) of the image as a longer exposure time increases the signal in the image and the exposure time is still much too short for dark current to accumulate. Note that I_{direct} is a $N \times 1$ vector where N is the number of metapixels in the image. I normalize the image by the exposure time (t_{direct}) and sum the count rate to get

$$dI_{direct} = \frac{I_{direct}}{t_{direct}}, \quad (5.1)$$

$$F_{HiCAT} = \sum_{i=0}^N dI_{direct_i}, \quad (5.2)$$

where dI_{direct} is a vector of the direct HiCAT image in counts/s, N is the number of metapixels in the image, and F_{HiCAT} is the total count rate in the image. I then choose a photon rate (F) that I want to scale the image to and determine the counts-to-photons conversion factor

(γ) for any type of image (direct or coronagraph) via

$$\gamma = \frac{F}{F_{HiCAT}}. \quad (5.3)$$

Note that γ is calculated once via the direct image and then used to convert the coronagraph images to photons/s. For this chapter, I choose $F = 15 \times 10^6$ ph/s which is intentionally slightly less than RST photon rate quoted above. HiCAT operates at a lower contrast than RST and will thus have more photons in the dark zone for a given direct image photon rate. Our goal for the low SNR DZM experiments is to use representative RST coronagraph images with the EKF, so it is more important that the photon rate in the dark zone matches RST than the direct image.

Once I have γ , I can scale the images and apply the desired noise characteristics. For the purpose of this chapter, I use a very basic noise model that only includes dark current and read noise, but the methodology is not restricted to this basic noise model. For each image (I_{HiCAT}), I first normalize by the exposure time (t_{HiCAT}),

$$dI_{HiCAT} = \frac{I_{HiCAT}}{t_{HiCAT}}, \quad (5.4)$$

to get dI_{HiCAT} in counts/s and then apply the scaling and noise model to obtain a low SNR image (I_{lsnr}) to use in the estimator:

$$I_{poisson} = P((\gamma dI_{HiCAT} + \mu_I)t), \quad (5.5)$$

$$I_{lsnr} = \text{round}(I_{poisson} + \mathcal{N}(0, \sigma_{rn}^2 \mathcal{I})). \quad (5.6)$$

Note that I_{HiCAT} , $I_{poisson}$, and I_{lsnr} are $N \times 1$ vectors where N is the number of metapixels in the full image (not just the dark zone). Here, $P()$ generates random variates selected from a Poisson distribution, μ_j is the dark current in e-/s, t is the RST-like exposure time in seconds, $I_{poisson}$ is an image (in units of photons) that contains Poisson noise and dark current only, σ_{rn} is the standard deviation of the read noise, and \mathcal{I} is the identity matrix with the same dimension as the image. The result is rounded since it is not physical to have a fraction of an electron. Note that the Jacobian and covariance matrices (for EFC and the EKF) need to be calculated with or scaled by t and I_{lsnr} as opposed to t_{HiCAT} and I_{HiCAT} .

In Fig. 5.2, I show the conversion from the initial HiCAT image in the left panel, I_{HiCAT} , to the measurement passed to the EKF in the right panel, $z^k = I_{lsnr}$. In the right panel we can clearly see the change in the sampling as well as the decrease in the SNR and overall signal. It should be noted that the peak intensity in the low SNR dark zone is on the order of 50 photons which is a factor of 10 higher than that expected for RST; this will be reduced for future experiments.

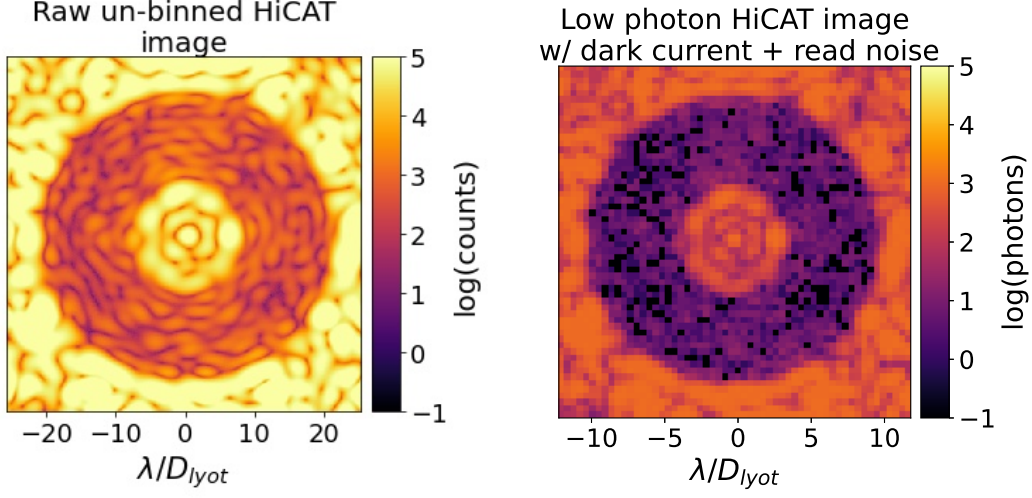


Figure 5.2: Conversion from high SNR HiCAT image (left) to the binned low SNR image for EKF (right). The left panel shows the log of the initial image (un-binned) obtained from HiCAT (I_{HiCAT}). There is definite structure to the speckles and the maximum intensity in the dark zone is on the order of 10^7 counts. In the right panel I have I_{lsnr} which is the result of passing I_{HiCAT} through the low SNR conversion process. There is still some structure to the speckles but it is drastically reduced. The exposure time used for the right panel is 39 s.

5.3 HiCAT RESULTS

In this section, I present the initial results of low SNR DZM on HiCAT; the parameters used for this experiment are summarized in Table 5.2. Both BMC DMs inject high-spatial-frequency drifts via a random walk of each actuator with $\sigma_{drift}^2 = 400 \text{ pm}^2/\text{iter}$ and the dither is chosen to be $\sigma_{dither} = 250 \text{ pm}$. The counts to photons conversion parameter is found to be $\gamma = 9.2 \times 10^{-7}$ when I choose $F = 15 \times 10^6 \text{ ph/s}$. The exposure time used to generate I_{lsnr} is chosen to be the same as the exposure time for 47 UMa used in Observing Scenario 11⁵⁰. This is a monochromatic experiment at 638 nm using an annular dark zone of extent $5.8\text{--}9.8 \lambda/D_{lyot}$. The mean dark-zone contrast is defined as the mean of the contrast values within the dark zone.

Table 5.2: Experiment parameters used. Note that t_{HiCAT} is the exposure time used to obtain the initial image and t is the exposure time used to determine the noise injected into the image as discussed in Sec. 5.2.1. For this experiment, t is taken to be the exposure time used in Observing Scenario 11⁵⁰ for for 47 UMa.

Parameter	Value	Units
σ_{diaber}	250	pm
σ_{drift}^2	400	pm ² /iter
F	15×10^6	photons/s
γ	0.92	
t_{HiCAT}	0.01	s
t	39	s
μ_I	0.005	e-/s
σ_m	2	e-

Figure 5.3 shows the mean closed-loop contrast of the dark-zone region in magenta for the duration of the experiment (3500 iterations). The iteration time is set to $t_{iter} = t = 39$ s for Fig. 5.3 as the exposure time on RST or HWO will be the lower limit on the iteration time. I use the high SNR HiCAT images (I_{HiCAT}) described in Sec. 5.2.1 to calculate the mean dark-zone contrast to ensure it is the true contrast and not artificially high due to noise. The open-loop contrast (cyan crosses) diverges to 1.1×10^{-6} by the final iteration at which point the BMC DM drift command for each DM has a root-mean-square of 1.26 nm and a peak-to-valley of 8.39 nm. The dotted and dashed black lines show the mean and standard deviation of the magenta curve (mean closed-loop dark-zone contrast) respectively for the duration of the experiment. The mean dark-zone contrast is held at 5.3×10^{-8} within a standard deviation of 6.4×10^{-9} .

As seen in the magenta curve in Fig. 5.3, the mean and standard deviation vary with time. From 0–12 hrs, the mean is relatively constant but the standard deviation is rather large. After 12 hrs, the standard deviation decreases but the mean has a positive slope. This

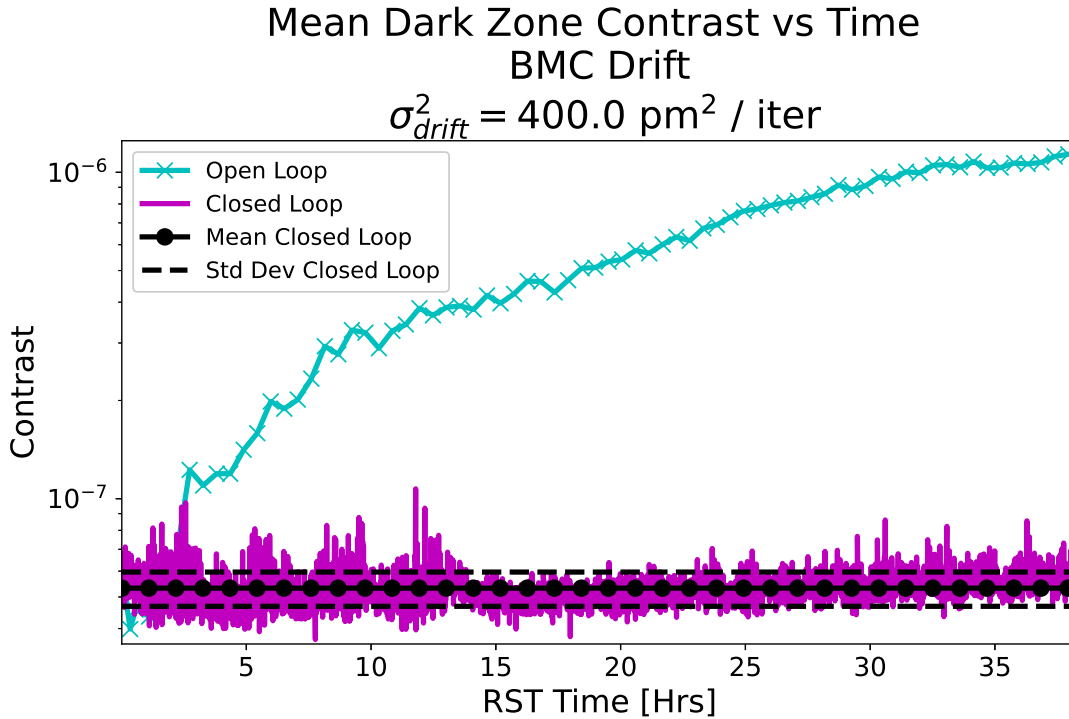


Figure 5.3: Mean dark-zone contrast against time for low SNR DZM experiment on HiCAT. Note that the mean dark-zone contrast plotted here is calculated using the high SNR I_{HiCAT} images and is thus the true contrast. The closed-loop contrast (solid magenta) is maintained at 5.3×10^{-8} (dotted black line) within a standard deviation of 6.4×10^{-9} (dashed black line) for the duration of the experiment. The open-loop contrast (cyan crosses) diverges to 1.1×10^{-6} by the final iteration. Note that the time axis for this plot uses assumes an iteration time of $t = 39 \text{ s}$.

is due to temperature and humidity instabilities in the lab due to issues with the HVAC system, which caused small low-order drifts (mostly tip-tilt). This tip-tilt drift is not modelled or captured by the EKF and thus slightly degrades the DZM performance. The humidity instability has been improved since these data were taken and, in addition, HiCAT now has an improved software architecture that allows faster control loops by one order of magnitude and running low order corrections in parallel with DZM as discussed in Soummer et al. 2022.¹²⁷ I expect significant gains in contrast as well as the inner working angle (IWA) in future experiments.

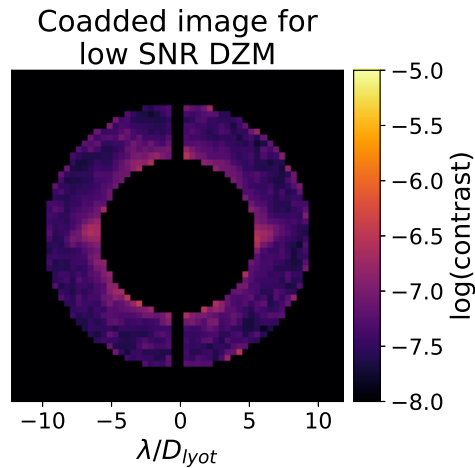


Figure 5.4: Coadded image for low SNR DZM experiment. I generate the noisy contrast image of the dark zone in each iteration using the I_{lsnr} direct and coronagraph images (see Eqn. 1.2). To produce this image I use simple coaddition¹⁴⁶ where I sum all dark-zone contrast images for the experiment and divide by the number of iterations. The SNR has increased in the coadded image when compared to the right panel of Fig. 5.2.

In Fig. 5.4, I show the coadded dark zone for the entire experiment. I use a basic coaddition approach¹⁴⁶ that simply averages the 3,500 contrast images from the experiment. As one would expect, the SNR drastically increases when the images are coadded. This was a good sanity check and Ch. 6 provides a thorough overview of post-processing techniques to extract a planet signal.

5.4 CONCLUSIONS AND FUTURE WORK

In this chapter, I experimentally demonstrate that the DZM algorithm can function in the low SNR regime using the HiCAT testbed. I outline the procedure of converting high SNR HiCAT images to low SNR images that emulate RST simulations and perform DZM using these low SNR images. With both BMC DMs drifting, the mean DZ contrast is maintained at 5.3×10^{-8} within a standard deviation of 6.4×10^{-9} . The open-loop

mean DZ contrast degrades over an order of magnitude to 1.1×10^{-6} over the course of the experiment. Note that our results are currently limited by the HiCAT air environment and not the DZM algorithm itself. These results further support that DZM would work on RST and if implemented, would increase science observing time since slews to a reference star are no longer needed. There are three main enhancements to implement to the low SNR DZM approach: (1) a more accurate noise model, (2) more accurate drifts, and (3) a larger dark zone. I aim to replicate these experiments with more realistic noise models using the `emccd_detect`⁷⁶ and `PhotonCount`⁷⁷ packages based on Nemati 2020.⁷⁵ The `emccd_detect.sim_sub_frame` function accepts a flux map which we create via

$$I_{flux_map} = P(\gamma d I_{HiCAT} t) / t, \quad (5.7)$$

where the HiCAT image is scaled, Poisson noise is added, and then the result is normalized by the exposure time. The output of `emccd_detect.sim_sub_frame` is an image with EMCCD noise in units of e- (including the bias) and scaled by the EMCCD gain. In the Observing Scenario 11 example script⁵⁹, if the EMCCD gain is greater than 1000, the photon counting mode of the EMCCD is used, otherwise the final image is simply normalized by the EMCCD gain. I will apply the same logic for our experiments. As for the drift injected, I aim to repeat or expand on the drift cases covered in Redmond et al. 2022⁹⁶ which includes the IrisAO DM drifting and all DMs drifting. Lastly, I will reduce the IWA of the dark zone; in Ch. 6 I provide low SNR results with an IWA of $4 \lambda / D_{Lyot}$ but ideally this will be further reduced to $3 \lambda / D_{Lyot}$ to match that of RST.

6

Post-processing of Dark Zone Maintenance

Data

6.1 CHAPTER OVERVIEW

In Ch. 4 and 5 I show how the DZM algorithm can hold the mean contrast in the dark zone constant while the open loop electric field drifts. While the mean dark zone contrast

is a useful metric, it does not tell us the most important information: does the DZM algorithm facilitate exoplanet detection? In this chapter I delve into post-processing techniques that extract planets from direct imaging data and why they are applicable (or not) to the DZM data.

It should be noted that the experiments presented in this chapter were performed after an upgrade of both the HiCAT testbed and software. The testbed light source and camera were updated and the software foundation was completely re-written. The software upgrades resulted in a drastic reduction in the iteration time for the DZM experiments. All experiments are monochromatic at 640 nm and use a $4-9 \lambda/D_{Lyot}$ dark zone.

6.2 PLANET INJECTION

The HiCAT testbed only has a laser source to mimic the target star around which an exoplanet could be orbiting. In order to test post-processing techniques to determine how DZM affects our ability to detect a planet, I need to inject a planet signal in software. I do this while running the experiment to ensure that the effect of the planet signal on the control loop is captured. To generate a fake planet, I take the direct image and scale it to be the desired planet contrast. I then shift the centre of the planet to be in the dark zone and add the result to the image obtained from the camera. The process of obtaining the science images and injecting the planet is shown in Fig. 6.1. If I am operating in low SNR mode, the planet injection happens prior to the low SNR conversion described in Fig. 5.1.

I use a variety of different planet contrasts to probe different limitations of the DZM algorithm. For example, it is possible that the EKF would struggle with very bright exoplanets as they will keep requesting DM commands to suppress the bright ‘speckle’ but since

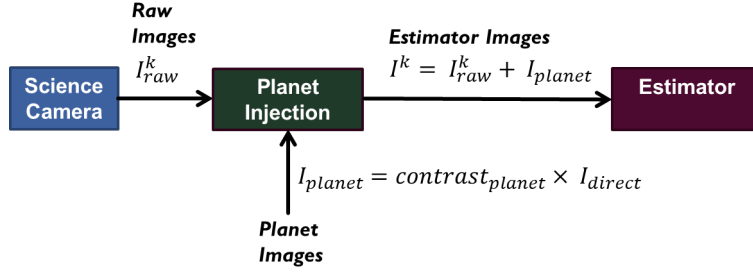


Figure 6.1: Flowchart for injecting a fake planet into the data acquired from the testbed during a DZM run.

the planet light is incoherent, the DMs will have little effect which will cause the loop to diverge. On the other hand, very dim planets will be difficult to detect under the contrast floor. Note that for this work, ‘bright’ and ‘dim’ planets are with respect to the contrast floor or mean dark zone contrast maintained by the DZM algorithm. I test planets slightly dimmer, slightly brighter, and an order of magnitude brighter than the mean DZ contrast (4×10^{-8} , 8×10^{-8} , 1×10^{-6}). I set the angular separation between the planet and the target star to be $6.62 \lambda / D_{lyot}$ and place the planet at an angle of 40° from the horizontal as shown in Fig. 6.2.

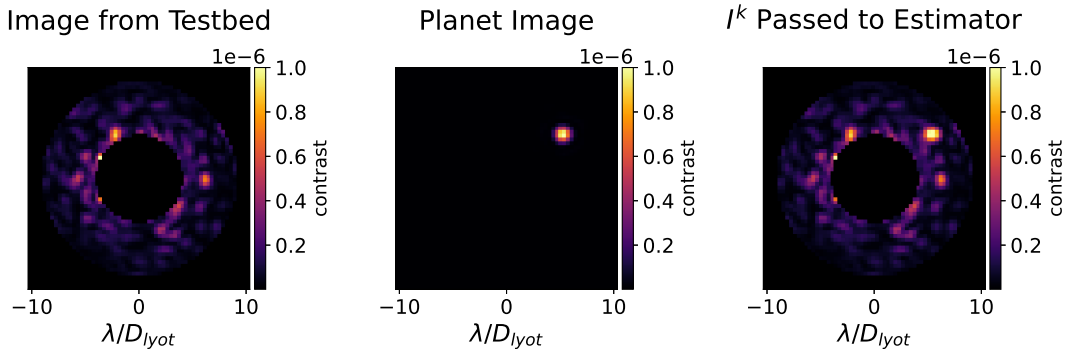


Figure 6.2: Demonstration of planet injection process using a very bright planet with a contrast of 10^{-6} which is placed $6.62 \lambda / D_{lyot}$ from the centre of the target star.

6.3 SIGNAL-TO-NOISE RATIO CALCULATIONS

One of the main ways to indicate planet detectability is via the signal to noise ratio (SNR). There are many rigorous ways to detect and determine the SNR of the planet as described in Marois et al. 2008⁶⁴, Mawet et al. 2014⁶⁶, and Bonse et al. 2023.¹³ In this work I adopt a simple approach to calculate the SNR of the planet. I assume we know the planet location and take the mean of the values of a $p_s \times p_s$ square around the centre of the planet to determine the planet signal. I create a noise image by masking the planet as well as any known artifacts and convolving it with a $p_s \times p_s$ square. The noise is then the standard deviation of the convolved noise image and the SNR follows as

$$s = \frac{1}{p_s^2} \sum_{i=1}^{p_s^2} I_{planet}[i] \quad (6.1)$$

$$b = 1_{p_s \times p_s} \frac{1}{p_s^2} \quad (6.2)$$

$$I_{noise} = (I \circ M) * b \quad (6.3)$$

$$n = \sqrt{\frac{\sum_{i=1}^N (I_{noise}[i] - \mu_{noise})^2}{N}} \quad (6.4)$$

$$SNR = \frac{s}{n}. \quad (6.5)$$

The pixels that are masked to create the noise image are either known bright speckles due to the optical elements or artifacts introduced by the post processing technique. For example, when using `pyklip` in ADI mode (further discussed in Sec. 6.4), there are two very low noise lobes that are created at an angle of $\pm \theta_{roll}/2$ from the planet in its initial roll location. This effect is a result of taking the autocorrelation of the planet signal and is de-

scribed in detail in Cantalloube et al. 2015.¹⁸ Including these lobes in the noise calculation would artificially increase the standard deviation of the noise in the image. A sample mask for planet square size of $p_s = 3$ is shown in Fig. 6.3 where the black pixels will be used for the noise image and the white pixels are discarded. Though it is not obvious here, the mask also discards the inner and outer ring of pixels in the dark zone as that is where we are most likely to get low contrast speckles bleeding into the dark zone.

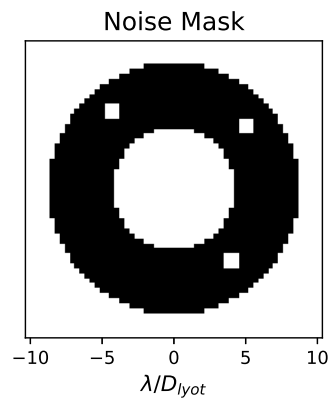


Figure 6.3: Mask for creating a noise image from a `pyklip` result. Black pixels are kept whereas white pixels are discarded. The square in the upper right quadrant is the planet mask and the other two squares are to mask the low noise lobes created by `pyklip`. Note that the mask needs to be rotated by θ_{roll} depending on if the initial roll angle is 0 or θ_{roll} as `pyklip` will always align the images to the first roll angle.

6.4 POST PROCESSING USING KARHUNEN-LOÈVE IMAGE PROJECTION

I start with one of the most commonly used algorithms in direct imaging: Karhunen-Loève Image Projection^{129,91} (KLIP). The basis of the KLIP algorithm is to perform a Principal Component Analysis (PCA) of a reference image library and project the PCA results onto the set of science images to extract the planet signal. KLIP can be used with most differential imaging techniques (as described in Sec. 2.4) and is implemented as a python package

called `pyklip`. When using `pyklip` with ADI the procedure is as follows:

1. Perform PCA of all the images from roll 1, keep first k modes
2. Project k PCA modes onto each image from roll 2 and sum the k projections for each image
3. Subtract each roll 2 image from the sum of the projections and sum the result for each image
4. Repeat steps 1–3 but switch roles of roll 1 & 2
5. Rotate roll 2 result to align with roll 1 and add

There is also an RDI method available with `pyklip`. As described in Sec. 2.4, RDI uses a library of reference images to disentangle residual starlight speckles from planet candidates. With RDI the procedure is very similar to ADI and is as follows:

1. Perform PCA of all the reference images, keep first k modes
2. Project k PCA modes onto each science image and sum the k projections for each image
3. Subtract each science image from the sum of the projections and sum the result for each image

Note for this analysis, I generate results for $k = [1, 5, 10, 20, 50]$.

6.5 ANGULAR DIFFERENTIAL IMAGING WITH DZM

Angular Differential Imaging (ADI) is a well developed post processing technique as described in Sec 2.4 for directly imaging exoplanets and here I pair it with DZM. ADI rotates the telescope around the bore-sight to move the planet on the detector. The speckles induced by the optics should rotate with the telescope which allows extraction of the planet signal. On HiCAT, instead of rotating the optical bench, I move where I inject the planet as shown in Fig. 6.4 using a roll angle of 90° . When running ADI experiments I also cycle the dither commands used by the EKF (Eqn. 4.20). I do this so that the speckles introduced by the dither commands are more consistent which aides with any type of differential imaging post processing algorithm. For these experiments, the roll angle is changed either halfway through, so $\delta u_{dither}^k = \delta u_{dither}^{k+n_{iter}/2}$, or every $n_{iter}/4$ iterations.

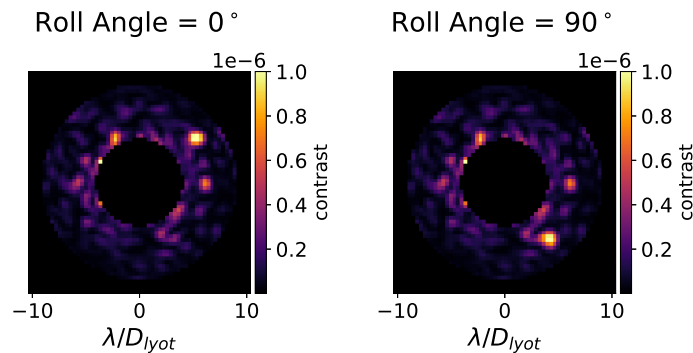


Figure 6.4: Demonstration of planet injection process for ADI. At the nominal roll angle for the telescope the planet is in the upper right quadrant of the dark zone (left panel) but when I roll by 90° the planet moves to the lower right quadrant of the dark zone (right panel).

6.6 HiCAT EXPERIMENTAL RESULTS USING ADI

6.6.1 LOW EXPOSURE TIME

For this experiment I injected a 5×10^{-8} planet into a dark zone with a mean DZ contrast of 7.85×10^{-8} and used a very short exposure time to reduce the overall signal. The full list of parameters used for this experiment are shown in Tab. 6.1. The experiment lasted a total of 1500 iterations and at iteration 750 I changed the roll angle and started repeating the dither commands.

Table 6.1: Experiment parameters used.

Parameter	Value	Units
σ_{dither}	250	pm
σ_{drift}^2	400	pm ² /iter
t_{HiCAT}	8×10^{-3}	s
$contrast_{planet}$	5×10^{-8}	-
θ_{planet}	6.62	λ/D_{Lyot}
θ_{roll}	90	degrees
n_{iter}	1500	-

Figure 6.5 shows the final result from `pyklip` using the data from this experiment. The noise mask from Fig. 6.3 was used when calculating the SNR. In the right panel of Fig. 6.5 we have the final image produced by `pyklip` using all of the data from this experiment. The two low noise lobes are obvious as well as the planet. The planet contrast is also approximately correct and the background has been drastically reduced. An image like the right panel is created for each point in the left panel to determine how the SNR of the planet evolves as the integration time increases. I start by only handing `pyklip` four frames and

then increase by 200 frames until all 1500 are used. This is to ensure that the noise is actually being reduced by increasing the integration time. If too much drift is leaking into the dark zone, the SNR vs integration time plot will eventually turn over and trend down. The SNR curve is plotted for each KLIP mode calculated. This is also a good sanity check as if too many modes are used, KLIP will start fitting the noise and the end result will be worse. The SNR is 9.61 when all exposures and 50 KLIP modes are used.

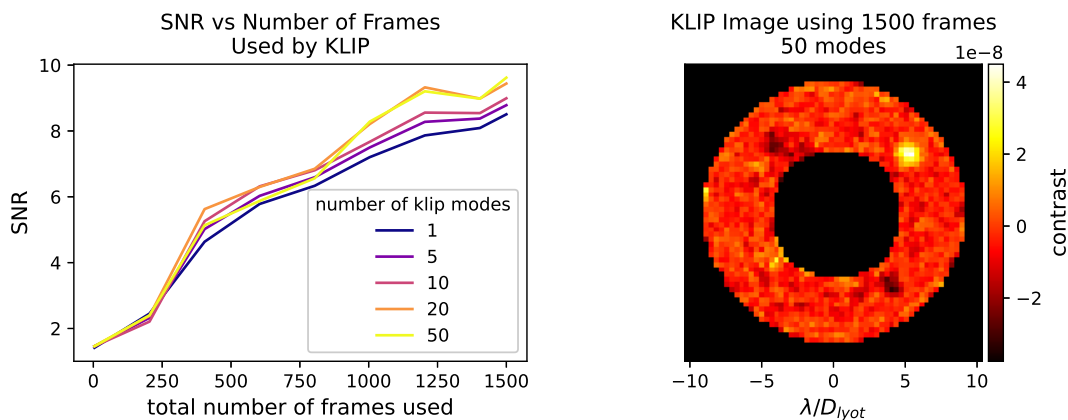


Figure 6.5: Left Panel: SNR as a function of number of frames provided to KLIP. This can be interpreted as SNR vs integration time. The SNR trends up as the number of frames increase and all modes have similar behaviour. For 50 modes, the planet has an SNR of 9.61 when all 1500 frames are used. Right Panel: Image produced by `pyklip` for the 50 mode case when all 1500 frames are used. The planet is clearly visible in the upper right quadrant as well as the low noise lobes on either side of it.

6.6.2 LOW SNR MODE

For this experiment I inject a 4×10^{-8} planet into a dark zone with a mean DZ contrast of 7.76×10^{-8} . The BMC DMs perform a random walk drift over 2200 iterations. The parameters for this experiment are summarized in Tab. 6.2 and the low photon procedure is described in Ch. 5. The exposure time simulated for the Low SNR mode is 39 s which results in a total simulated integration time of 23.8 hrs. For this experiment, I ‘roll’

the telescope four times (every 550 iterations) instead of two to match the RST Observing Scenario 11 procedure. The experiment summary plot is provided in Fig. 6.6 with the

Table 6.2: Experiment parameters used. Low photon method and parameters are described in Ch. 5.

Parameter	Value	Units
σ_{dither}	250	pm
σ_{drift}^2	400	pm ² /iter
t_{HiCAT}	0.1	s
t	39	s
F	1.5×10^7	photons/s
μ_I	0.005	e-/s
σ_{rn}	2	e-
contrast_{planet}	4×10^{-8}	-
θ_{planet}	6.62	λ/D_{Lyot}
θ_{roll}	90	degrees
n_{rolls}	4	-
n_{iter}	2200	-

mean dark zone contrast vs time in the left panel and the closed loop image in the right panel. The time axis for this plot is HiCAT time. As shown in the right panel of Fig. 6.6, the planet is not visible, even when using a logscale for the image.

As shown in the left panel of Fig. 6.7, when all 2200 frames are used the planet achieves a SNR of 9.58. The behaviour of all the modes is similar so I know I am not fitting the noise. An SNR of 9.58 is well above the detection threshold and one can also clearly see the two negative lobes created by the pyklip algorithm in the right panel of Fig. 6.7. This experiment demonstrates how DZM can be used in conjunction with ADI in the case of dim planets below the contrast floor.

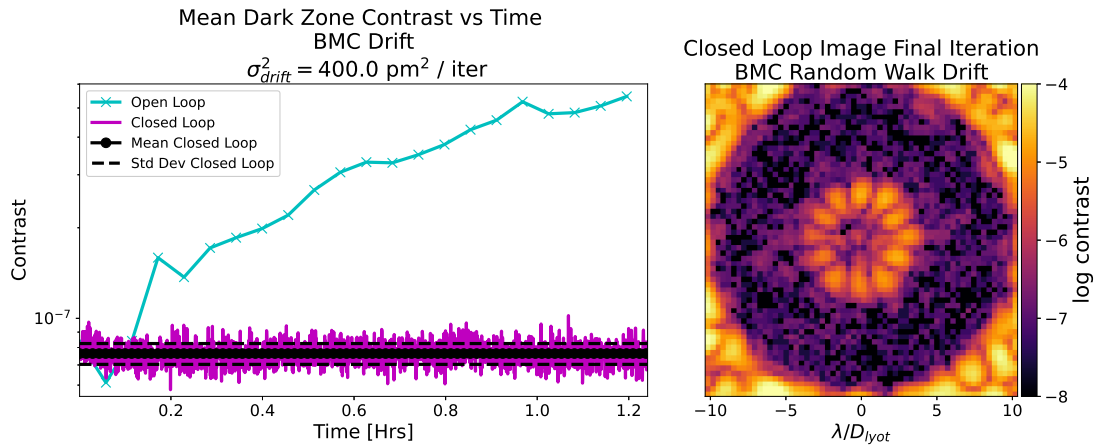


Figure 6.6: Left panel: Mean DZ contrast vs time for the low SNR case with a 4×10^{-8} planet. The cyan is the open-loop contrast, the magenta is the closed-loop contrast, and the black dots and dashes are the mean and standard deviation of the magenta curve respectively. Note that the time axis is the actual duration of the experiment and does not use the simulated 39 s exposure time. Right panel: Sample science image from the final iteration. The planet is not visible by eye in the dark zone.

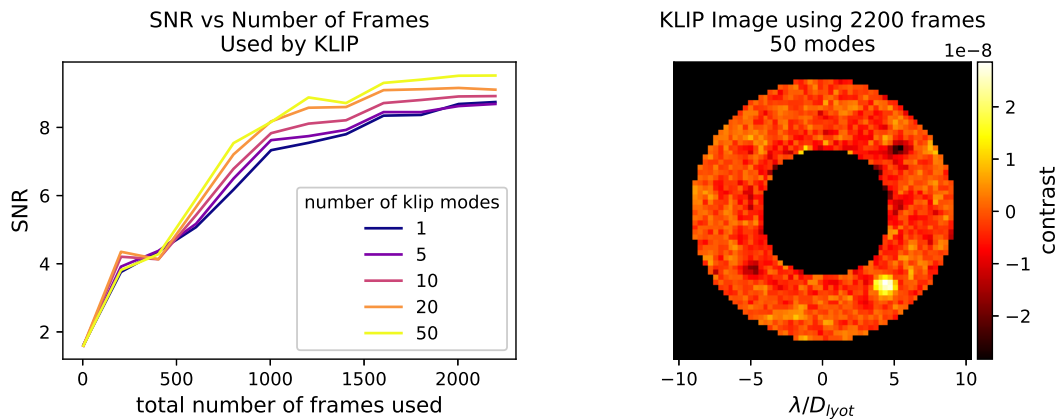


Figure 6.7: Left Panel: SNR as a function of number of frames provided to KLIP. This can be interpreted as SNR vs integration time. The SNR trends up as the number of frames increase and all modes have similar behaviour. For 50 modes, the planet has an SNR of 9.58 when all 2200 frames are used. Right Panel: Image produced by pyklip for the 50 mode case when all 2200 frames are used. The planet is clearly visible in the lower right quadrant as well as the low noise lobes on either side of it.

6.7 DZM WITH REAL-TIME INCOHERENT LIGHT ESTIMATION

6.7.1 MOTIVATION

One potential limitation of the DZM implementation described in Ch. 4 is in the case of bright planets or bright exozodiacal dust emission.¹⁰⁴ If the planet is very bright, the estimator will capture it and the DMs will try to suppress it. Since the planet light is incoherent, the DMs will not succeed which can cause the control loop to diverge. This is demonstrated in Fig. 6.8 which shows the contrast vs time for two **simulated** DZM runs where a bright planet is injected. The initial mean DZ contrast without the planet is 5×10^{-8} and the planet contrast is either 10^{-6} (magenta dots) or 10^{-5} (blue). The dashed black line shows where the roll angle of the planet changes. For both cases, one can see that the estimator takes longer than normal to converge; note that these peaks at the beginning could be reduced by turning on the controller later. For the 10^{-6} planet, we see the mean DZ contrast degrade over the 750 iterations. For the 10^{-5} planet, the system struggles when the roll angle changes. This is something that could be mitigated by disabling the controller for a period of time after a roll change but while the controller is off, drifts will leak in. If the controller has to be disabled every roll change, that can result in contrast degradation.

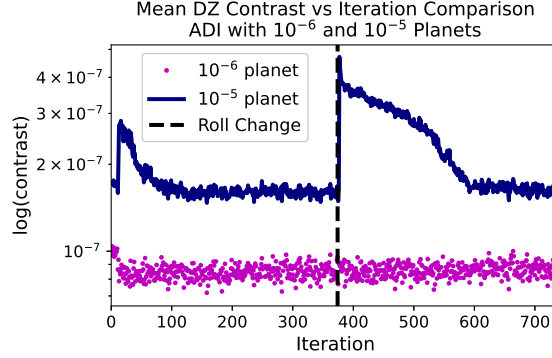


Figure 6.8: Simulated HiCAT contrast vs time data. The magenta dots curve has a 10^{-6} planet injected and the mean DZ contrast creeps up over the course of the experiment. The blue curve has a 10^{-5} planet injected and struggles to obtain a good estimate at the start as well as update the estimate for the new planet location when the roll angle changes at iteration 375 (black dashes).

The reason for the second peak in the blue curve in Fig. 6.8 can be seen in Fig. 6.9. The top panel shows the **simulated** image and the bottom panel shows the closed-loop intensity estimate. When the roll angle changes at iteration 375, we see the planet move from the top right quadrant to the bottom right quadrant in the image. In the estimate, it captures the new planet location quickly but holds on to the old planet location and we can see a bright speckle in both the top and bottom right quadrants of the dark zone. If we want to avoid the contrast degradation seen in Fig. 6.8, we would have to leave the controller of until the EKF had ‘forgotten’ about the initial roll location of the planet.

As described in Sec. 2.4, ADI also results in a loss of observing time due to the time associated with rolling the telescope. If a DZM approach can be implemented that removes the requirement of ADI for dim planets as well, that is very advantageous to future space-based coronagraphs. I investigate if I can use solely the data acquired from the DZM experiment to separate the coherent and incoherent light in the image. This allows for different post-processing techniques that do not require any additional hardware or slew time.

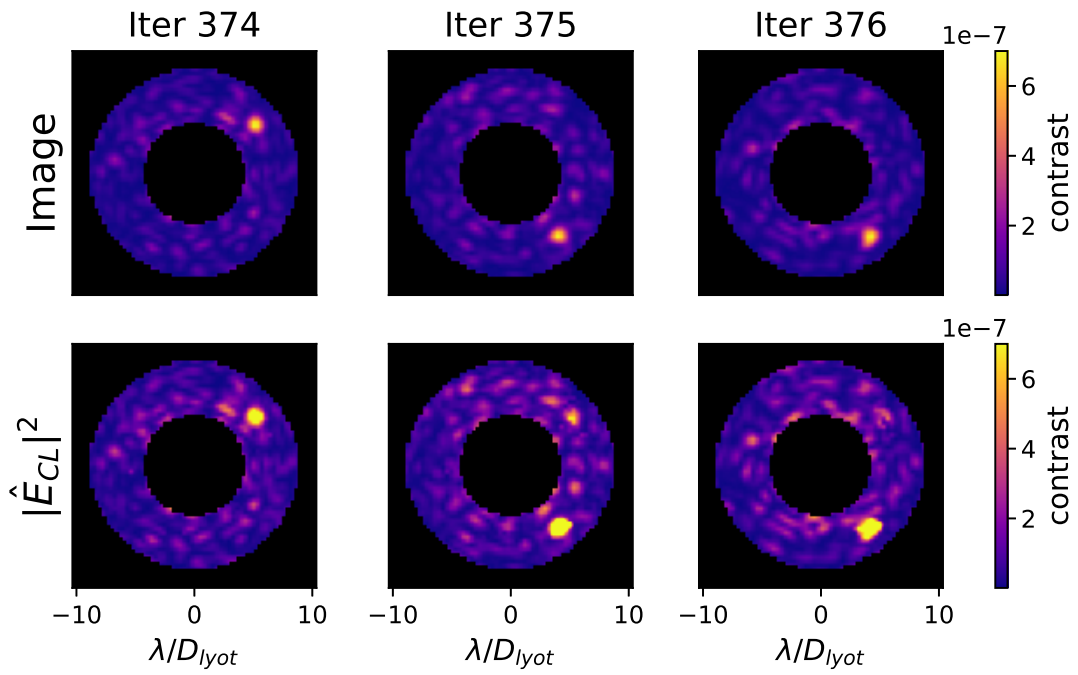


Figure 6.9: Simulated HiCAT Data. Top row: images for iteration 374–376 of the experiment shown by the blue curve in Fig. 6.8. The roll angle changes at iteration 375 and the planet moves from the top right quadrant to the bottom right quadrant. Bottom row: closed-loop intensity estimate showing how the EKF holds on to the initial roll location of the planet. A bright speckle can be seen in the top right quadrant of the estimate in iterations 375 and 376 as well as the new planet location in the bottom right quadrant.

6.7.2 AUGMENTED EXTENDED KALMAN FILTER

A way to ensure stability of the control loop in the case of bright planets, incoherent background, and exozodiacal dust emissions is to augment the EKF to have a third state which is the incoherent light in the image. This augmented EKF is described in detail for digging a dark zone in Riggs 2016; ^{too} I adapt it here for maintaining the dark zone. The response matrix (equivalent to the Jacobian G for the nominal EKF) and state vector become

$$\Gamma = \sqrt{t_{\text{exp}}} \begin{bmatrix} \Re\left(\frac{\partial E_{OL_0}}{\partial u}\right) \\ \Im\left(\frac{\partial E_{OL_0}}{\partial u}\right) \\ 0 \\ \vdots \\ \Re\left(\frac{\partial E_{OL_{n-1}}}{\partial u}\right) \\ \Im\left(\frac{\partial E_{OL_{n-1}}}{\partial u}\right) \\ 0 \end{bmatrix}, \quad (6.6)$$

$$x = \begin{bmatrix} \Re(E_{OL_0}) \\ \Im(E_{OL_0}) \\ I_{inco_0} \\ \vdots \\ \Re(E_{OL_{n-1}}) \\ \Im(E_{OL_{n-1}}) \\ I_{inco_{n-1}} \end{bmatrix} \quad (6.7)$$

where I_{inco} is the incoherent intensity in the image and E_{OL} is the open loop electric field as described in Tab. 4.2. The closed loop version of the state vector (Eqn. 4.6) can be expressed as

$$x_{CL}^k = x^k + \Gamma u^k \quad (6.8)$$

which then leads to the following for Eqn. 4.7

$$\mathbf{z}^{k+1} = E_{CLRe}^k \circ E_{CLRe}^k + E_{CLIm}^k \circ E_{CLIm}^k + I_{inco} + \mathbf{n}^k \quad (6.9)$$

where E_{CLRe} and E_{CLIm} are the real and imaginary parts of the closed loop electric field respectively as calculated in Eqn 6.8. Assuming BMC drift, the process noise is then

$$Q_{BMC}^k = \Gamma \Gamma^T \sigma_{drift}^2 \quad (6.10)$$

but the 0 rows of Q are set to a small number to allow the EKF to converge and be numerically stable. The last major change for the EKF is the H matrix. For each pixel the H matrix is augmented to be

$$H_i^k = \begin{bmatrix} 2\Re(\hat{E}_{CLi}^k) & 0 & 0 \\ 0 & 2\Im(\hat{E}_{CLi}^k) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.11)$$

where i indicates the pixel index in the dark zone. Note that for the EFC side of the control loop, it is still only handed the open loop electric field estimate; not the augmented state vector.

A very important aspect of any Kalman Filter is how the state vector is initialized. For the original EKF, the open loop electric field estimate can be initialized either as zero and it will converge in ~ 30 iterations or as the final estimate of the PWP experiment. For the

incoherent light there were three options considered:

$$\hat{I}_{inco}^0 = |\overline{z^0 - \hat{z}^0}| \quad (6.12)$$

$$\hat{I}_{inco}^0 = |z^0 - \hat{z}^0| \quad (6.13)$$

$$\hat{I}_{inco}^0 = |z^0 - z_{ref}| \quad (6.14)$$

where z^0 is the initial science image on the target, \hat{z}^0 is the initial estimate of the closed-loop intensity (generated from the reference star), and z_{ref} is the final image from generating the dark zone on the reference star. Each of these options is physical in its own way. For Eqn. 6.12, it assumes that the primary contribution to the incoherent light is a uniform background which should be the same order of magnitude as the error in the intensity estimate of the coherent light and the image. Equation 6.13 assumes the incoherent light has structure which matches the estimate error. Also note that since \hat{z}^0 is from the reference star, it will not contain the planet but z^0 will; taking the difference and keeping the structure potentially aids the EKF in locating the planet. The last option (Eqn. 6.14) is based on a similar idea in that we can take the difference of the final reference star image and the first target star image to start the estimator off with a potential planet location.

6.7.3 POST PROCESSING OPTIONS ENABLED BY THE AEKF

Another benefit of the AEKF is that it opens the door for more post-processing options. Rolling a space telescope is costly as it takes time both to execute the roll and allow the vibrational modes excited by the roll to settle. In Pogorelyuk et al. 2019^{88,89} the electric field order reduction (EFOR) method is discussed; this is a robust option that does not require

ADI but is very computationally expensive. There are three simpler post-processing options I consider to use with the AEKF data:

1. Simplified Coherent Differential Imaging (CDI)
2. Incoherent Accumulated Imaging (IAI)
3. Reference Differential imaging.

The first two are similar and leverage the estimates produced by the AEKF. I implement a simplified software-based version of CDI where I take the closed-loop coherent intensity estimate for each iteration and subtract it from the science image,

$$I_{pf}^k = z^k - |\hat{E}_{CL}^k|^2 \quad (6.15)$$

$$I_{pf} = \frac{1}{n_{itr}} \sum_{i=1}^{n_{itr}} I_{pf}^k \quad (6.16)$$

where I_{pf} is the ‘post processing’ image for the CDI approach (shown in Fig. 6.10). Since the planet is not in the coherent electric field estimate, this should reveal the planet in the science image. When all the estimate differences are obtained, they are summed to average out the noise.

The IAI approach is to take the mean of the incoherent estimates since the incoherent estimate has captured the planet,

$$I_{pi} = \frac{1}{n_{itr}} \sum_{i=1}^{n_{itr}} \hat{I}_{inco}^k \quad (6.17)$$

where I_{pi} is the post-processing image for the incoherent estimate approach.

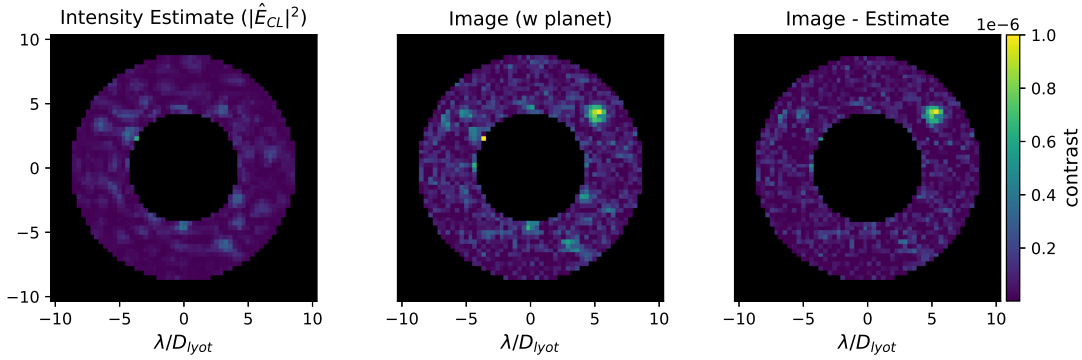


Figure 6.10: Visual representation of generating I_{pf}^k for a 10^{-6} planet in the low SNR mode.

The third approach goes back to using `pyklip` but in RDI mode instead of ADI mode. The RDI procedure is explained in Sec. 6.4; when combining this with the AEKF, the idea is to use the closed-loop coherent intensity estimates as the ‘reference images’ as they do not contain the planet and should capture the PSF structure. First I generate a PSF library which contains the closed-loop coherent intensity estimates and the science images. Using this PSF library, `pyklip` generates a correlation matrix for the entire dataset. Then, the science dataset is assembled which only contains the science images. Both the PSF library and the science dataset are passed to `pyklip` which performs the PCA procedure described in Sec. 6.4; note that the images are indexed and `pyklip` automatically removes any science images from the PSF library when it assembles the reference PSF dataset.

6.8 HiCAT EXPERIMENTAL RESULTS USING CDI AND IAI

Table 6.3: Experiment parameters used for all experiments in this section. Low photon method and parameters are described in Ch. 5.

Parameter	Value	Units
σ_{ditber}	250	pm
σ_{drift}^2	400	pm ² /iter
t_{HiCAT}	0.1	s
t	39	s
F	1.5×10^7	photons/s
μ_I	0.005	e-/s
σ_{rn}	2	e-
θ_{planet}	6.62	λ/D_{Lyot}
n_{iter}	1500	-

6.8.1 BRIGHT PLANET LOW SNR

For this experiment, I inject a 10^{-6} planet into a dark zone with a mean contrast of 9.8×10^{-8} . The images passed to the estimator are flux adjusted as per Tab. 6.3. The experiment runs for 1500 iterations; all of which are used in the post-processing. The results of the CDI and IAI post-processing techniques are shown in Fig. 6.11. The top row takes the difference between the coherent estimate and the science image (CDI) and the bottom row uses the incoherent estimate (IAI). As more frames are used in the stack, the background

smooths out; this is especially evident between 4–500 frames. It should also be noted that

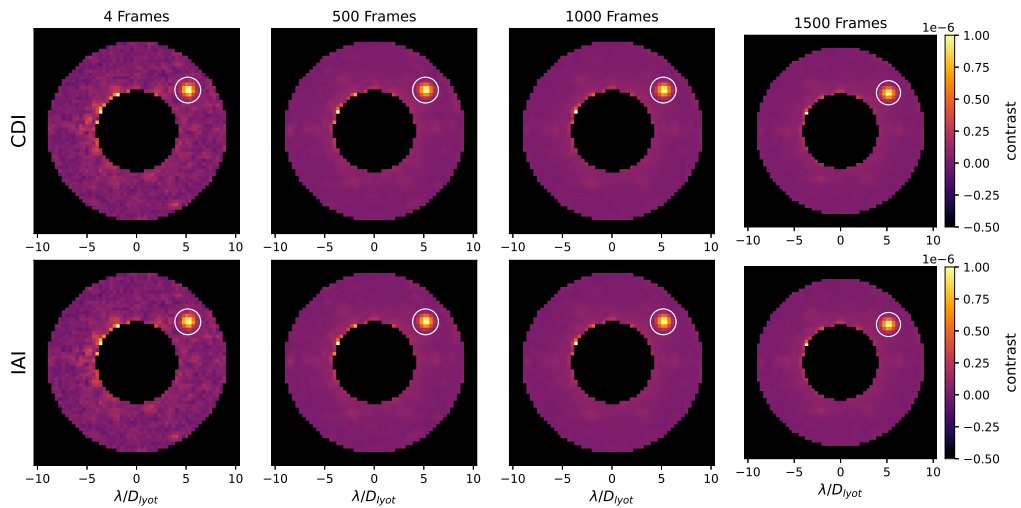


Figure 6.11: Here I show the evolution of I_{pf} (top row) and I_{pi} (bottom row) for a hardware experiment on HiCAT. Each column stacks a different number of frames. For this experiment, a 10^{-6} planet is injected and it is clearly visible even when only four frames are used. Between 4–500 frames the background is noticeably smoother as the noise averages out.

the two rows in Fig. 6.11 look very similar. I can quantify this by looking at the SNR of the planet as a function of how many frames are used as shown in Fig. 6.12. The CDI method is shown in orange and the incoherent estimate method is shown in black. Both methods end with an SNR of ~ 29 when all 1500 frames are used.

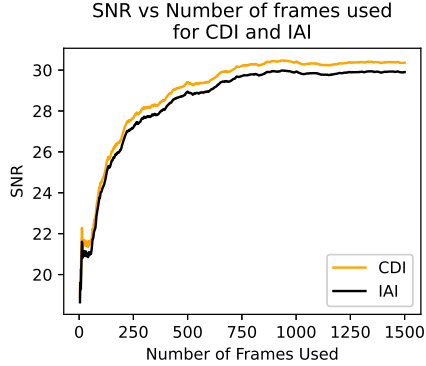


Figure 6.12: Comparison of SNR vs number of frames used for CDI (orange) and incoherent estimate (black) post-processing methods.

6.8.2 DIM PLANET LOW SNR

For this experiment, I inject a 8×10^{-8} planet into a dark zone with a mean contrast of 8.9×10^{-8} . The images passed to the estimator are flux adjusted as per Tab. 6.3. The experiment runs for 1500 iterations; all of which are used in the post-processing. The results of the CDI and IAI post-processing techniques are shown in Fig. 6.13. Using the AEKF with DZM is still a very novel concept and even showing that the incoherent state in the AEKF captures the dim planet (bottom row in Fig. 6.13) is a very promising result. As with Sec. 6.8.1, the more frames that are used in the stack, the more the background smooths out. Unfortunately for the dim planet case, there are a significant number of residual speckles on the same order of magnitude as the planet in the post-processing images. These residual speckles are erroneously captured by the incoherent intensity estimate instead of the coherent electric field estimate which is why the two rows look the same. This may be a fundamental limit of the AEKF as the incoherent estimate will tend to capture any light that is not being modulated by the DMs. If there are other sources of incoherent light in

the dark zone, they will be captured by \hat{I}_{inco} . If other incoherent light sources form a uniform background, it is much less of an issue than if it has structure.

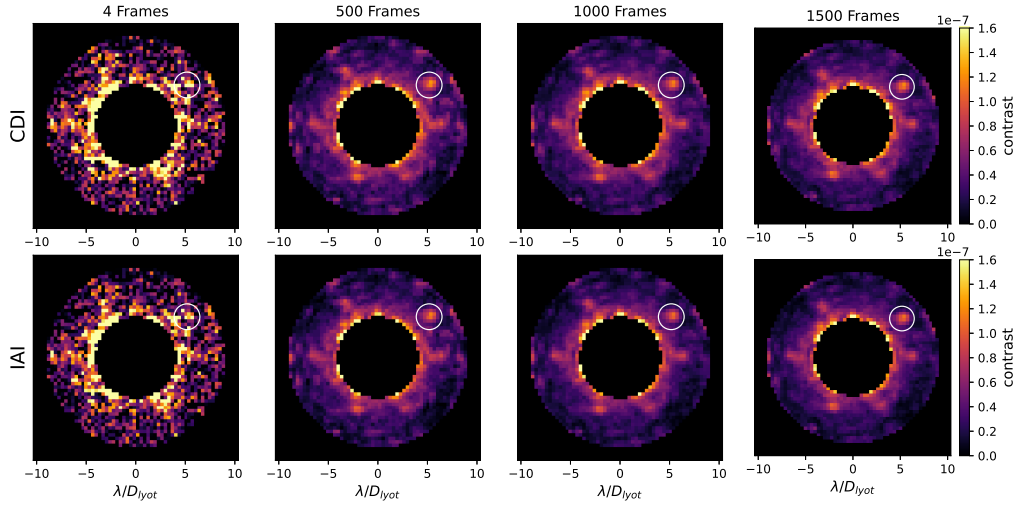


Figure 6.13: Here I show the evolution of I_{pf} (top row) and I_{pi} (bottom row) for a hardware experiment on HiCAT. Each column stacks a different number of frames. For this experiment, a 8×10^{-8} planet is injected (circled in white) and it is not visible initially but emerges as more frames are used. The background smooths out but does not decrease and there is obvious structure other than the planet remaining.

Figure 6.14 shows the comparison between CDI (orange) and IAI (black) of the SNR as a function of the number of frames used. As with Sec. 6.8.1, the two behave very similarly with an SNR of 4.9 when all 1500 frames are used. This experiment demonstrates the importance of planet detection methods as, if only the SNR is considered, this could be viewed as a good result that simply needs more frames. In reality, we can see that determining which blob in Fig. 6.13 is a planet vs a speckle is much more problematic than the other ADI or CDI/IAI cases I have considered. One way to determine between a speckle and a planet for CDI or IAI would be through target revisits. By returning to a target multiple times, we image the planet at different locations in along it's orbit. If the optics are sufficiently stable, the speckle pattern will stay the same and the planet will move.

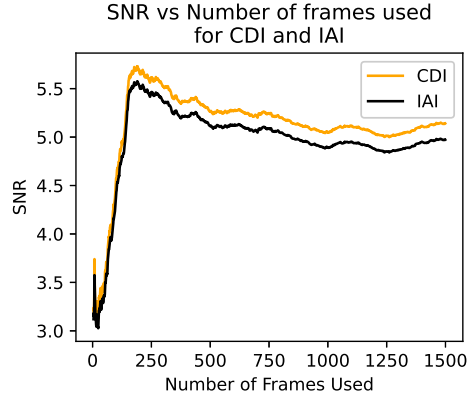


Figure 6.14: Comparison of SNR vs number of frames used for CDI (orange) and incoherent estimate (black) post-processing methods.

6.9 HiCAT EXPERIMENTAL RESULTS USING AUGMENTED EKF AND RDI

I now take the same experiments from Sec. 6.8 and put them through `pyklip` but using the RDI mode instead of ADI. For the 10^{-6} planet case, the result is good as expected with an SNR of 23.1 when all 1500 frames and 10 KLIP modes are used. It should be noted that this is a case where the number of KLIP modes has a significant impact as shown in Fig. 6.15. If one KLIP mode is used, the planet has an SNR of 32.9 but if 50 modes are used it has an SNR of 18.7. As the number of modes increases, KLIP begins to fit the residual structure as shown in the increase in speckles between the left and right panel of Fig. 6.15. This KLIP effect is not present when using the ADI mode with the DZM data.

Next I consider a 8×10^{-8} planet which is on par with the mean dark zone contrast. As shown in Fig. 6.16, KLIP RDI barely recovers the planet and it only has an SNR of 4.2 after 1500 frames. For planets near the contrast floor, the SNR still decreases as a function of the number of KLIP modes but by $<5\%$.

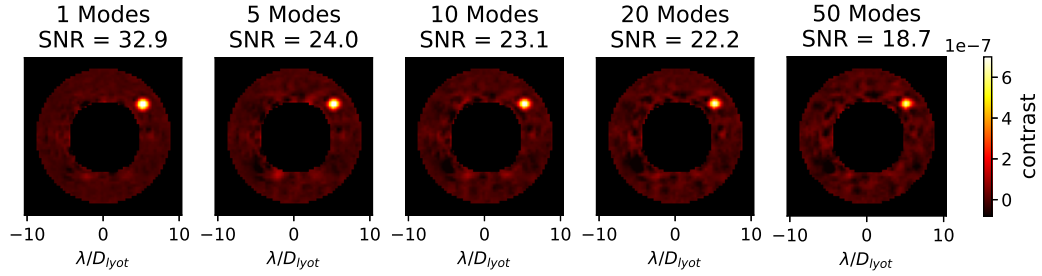


Figure 6.15: KLIP RDI result using 1500 frames for a 10^{-6} planet in a low SNR dark zone. The planet is clearly visible by eye for all KLIP modes. The SNR decreases as the number of KLIP modes increases because KLIP begins to fit the residual structure in the image which increases the noise.

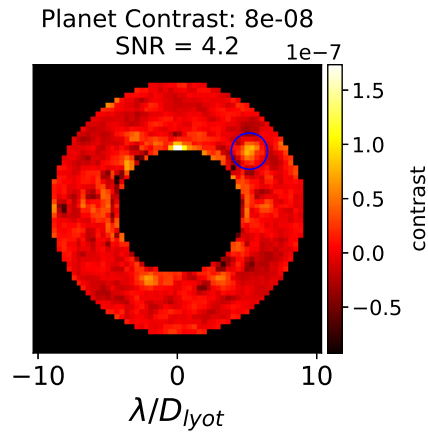


Figure 6.16: KLIP RDI result using 1500 frames and 10 KLIP modes for a 8×10^{-8} planet in a low SNR dark zone. The planet is circled in blue and is barely visible.

The `pyklip` results in RDI mode are on par with the CDI and IAI results in Sec. 6.8 as far as the planet SNR is concerned. When looking at the whole picture, `pyklip` outperforms the CDI and IAI approaches as it drastically reduces the speckle pattern; this is further discussed in Sec. 6.10.

6.10 PLANET DETECTION

Once a dataset from an experiment has been post-processed to produce a single low-noise image, the final step is to distinguish between speckles and planets in the image. There are a number of approaches, the most robust of which involve using a spectrograph to look at the spectra of each planet candidate in the image or by revisiting the target when the planet is at a different orbital position. A spectra-based approach is not possible with the data obtained from HiCAT so instead I use the `point_source_detection` function within `pyklip`. In order to use this function, I first generate a `pyklip` SNR map. This SNR map is generated in a similar way to the method described in Sec. 6.3 but is more robust and accurate; note that the SNR numbers quoted in this section are slightly different since I am using the `pyklip` function. The SNR map is passed to the `point_source_detection` function which performs an iterative process where it finds the highest SNR candidate, masks it, and then repeats until the highest SNR candidate is below the SNR threshold (which I choose to be 5). The output is a candidate list with the location and SNR of each candidate above the threshold.

Figure 6.17 shows sample detection results using each of the four post-processing techniques previously discussed. For all of the techniques, 1500 frames are used to generate the post-processed image and a 8×10^{-8} planet is injected into the dark zone. For the ADI experiment, the mean DZ contrast is 7.7×10^{-8} and for the RDI/CDI/IAI experiment (Sec. 6.8.2) the mean DZ contrast is 8.9×10^{-8} . The top two panels of Fig. 6.17 show the detection results for the KLIP-based post-processing methods and the bottom two show the simple stacking post-processing methods. The candidate with the highest SNR is cir-

cluded in black and the other candidates above the threshold are circled in red. In all cases, the planet is the candidate with the highest SNR. For CDI and IAI we see that the planet has a higher SNR than the RDI but with the large caveat that there are a number of other candidates above the SNR threshold detected.

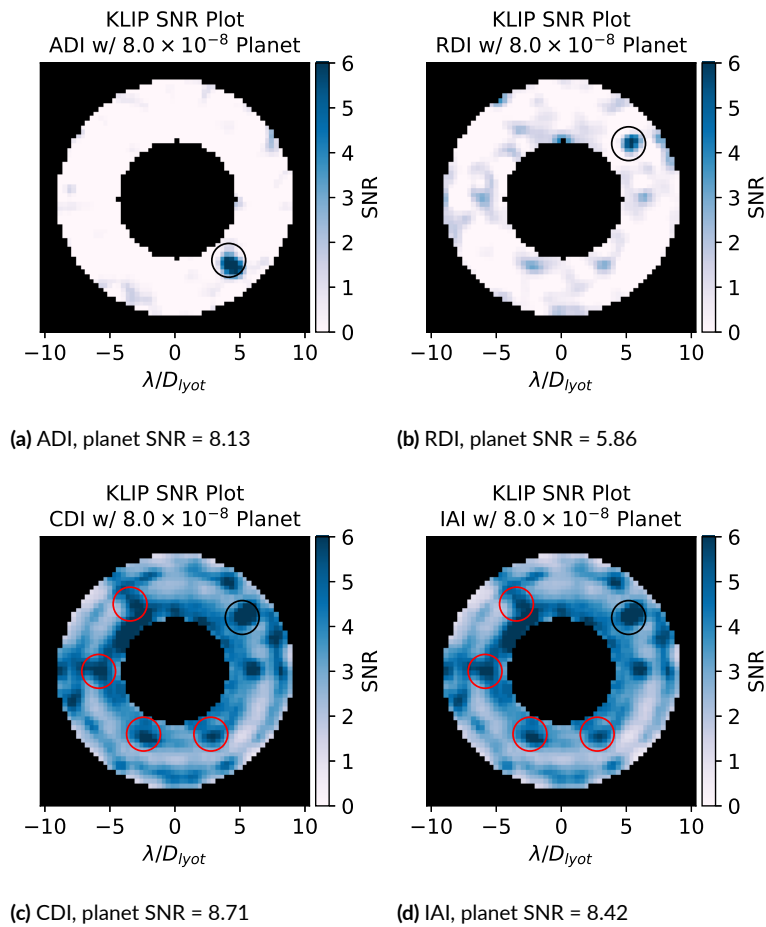


Figure 6.17: Planet detection results from KLIP using the four post-processing methods discussed. For all cases, an 8×10^{-8} planet is injected to a dark zone with a comparable mean DZ contrast and 1500 frames are used. The RDI, ADI, and IAI cases use data from the same experiment. The candidate with the highest SNR is circled in black and, in all cases, the planet has the highest SNR. In red are all other candidates returned by `pyklip` that have an SNR > 5.

The false candidates detected in the CDI and IAI cases have SNR values between 5.7–

7.5. This is obviously not ideal and future work on CDI and IAI should work to remove the speckles contributing to these false candidates. In a space mission, the targets are revisited and the planet is in a different orbital position during each revisit. This would help the CDI and IAI approaches distinguish between speckles and planets. It should also be noted that an SNR of five is a relatively low threshold for planet detection and a threshold of eight is more commonly used. Setting the threshold to eight would result in the rejection of all of the false candidates in this case.

6.11 CONCLUSIONS AND FUTURE WORK

This chapter demonstrates how we can augment the DZM EKF to estimate the incoherent light in the image and how DZM data can effectively be used with a number of different post-processing methods. Post-processing with DZM data is still in the early stages of development and there are a number of future research avenues. When using the nominal DZM algorithm with ADI, it would be worthwhile to investigate the shortest allowable dither-repeat cycle that is longer than the memory of the EKF. Currently, the dither commands start to repeat whenever the planet is ‘rolled’ but if the timescale was shorter it is possible that `pyklip` would perform better as the residual starlight speckle pattern would be more consistent.

Using the AEKF with DZM is a novel concept and there are many exciting paths forward. Repeating the dither commands could also potentially aid with the RDI approach for a similar reason so experiments should be run with the AEKF and dither-cycling. One could also run AEKF with planet-PSF-sized blocks instead of single pixels. The H matrix (Eqn. 6.11) would then be $3n_p \times 3n_p$ where n_p is the number of pixels in an expected planet

PSF. As the incoherent estimate finds the planet, the blocking of the dark zone could be dynamically updated to centre the block on the planet. It would also be interesting to look into the tuning of the Q matrix and have it dynamically update throughout the experiment. Often when using an EKF, the Q matrix is made to be more aggressive at the start and taper off. It is possible that by implementing a dynamic Q matrix, I could avoid accumulating the coherent estimate error in the incoherent estimate which would improve the CDI and IAI post-processing results. Since exozodiacal dust is predicted to be a potential issue for directly imaging exo-Earths¹⁰⁴, it would also be beneficial to test the AEKF with representative exozodiacal dust emission injected in to the images. Lastly, it would be good to look into using the AEKF with ADI and develop a methodology for re-initializing the incoherent estimate when a roll is performed. The AEKF is much more robust for dealing with rolls as it does not require re-initializing the coherent electric field estimate as that will not change. For this reason, it should be relatively trivial to avoid the peaks seen in Fig. 6.8 when the nominal DZM algorithm is used when a bright planet is present.

Yeab, this is fun.

James Castrission

7

Development of the GigaBIT Instrument Specifications

7.1 CHAPTER OVERVIEW

Prior to considering the GigaBIT hardware, I first review the technical specifications that will enable GigaBIT to meet its science goals. I then go through the SuperBIT results

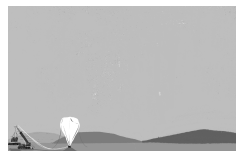
which provide valuable data on expected behaviour for GigaBIT. Lastly, I outline the GigaBIT instrument requirements that emerge from the combination of the technical specifications and the SuperBIT results.

7.2 GIGABIT TECHNICAL SPECIFICATIONS

GigaBIT aims to be a balloon-based observatory that provides diffraction-limited wide-field imaging capabilities in the optical to near-ultraviolet (NUV). These qualities are targeted to fill a gap in the collective observing capabilities of current and future telescopes for important science goals (Fig. 1.9 in Sec. 1.3.2). One of the main science goals is to study cluster weak gravitational lensing as described in Sec. 1.2.2.

In order to ensure the errors on the weak lensing measurement are inherent to the shape-measurement method, GigaBIT must be able to detect a minimum of 35 galaxies / arcmin² at a redshift greater than 0.3.^{40,65,48,80} In order to perform accurate shear measurements, it is also important to have a high resolving power to maximize the galaxy shape measurements. A resolution of 0.081 arcseconds in the NUV achieves this goal.⁸⁰ These two requirements combined with a finite amount of observing time feeds into the required collecting area (or primary mirror diameter) of GigaBIT which will be discussed in Sec. 7.5.

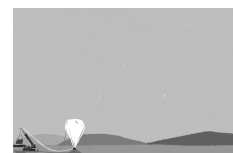
In order to take advantage of the stratospheric environment and complement the other available astronomical observatories, GigaBIT targets a wavelength range from 300–700 nm. Imaging in the NUV has multiple benefits. The first is that having NUV measurements allows for more accurate estimation of the galaxy redshifts as the Balmer break is typically around 400 nm.⁶⁰ Having accurate redshift estimation is crucial as discussed in Sec. 1.2.2 and allows separation of the foreground and background galaxies. Secondly, foreground



objects tend to emit less in the NUV (lower surface brightness) which enables more accurate and aggressive cuts of objects that can contaminate the weak lensing measurement. Lastly, this decision is also driven by the results presented in Shaaban et al. 2022.¹¹⁵ This study showed that when operating in the stratospheric environment with a wide-field diffraction limited telescope, the most efficient band to observe in to acquire the desired number of sources in the minimal amount of integration time is in the blue. This is different than for ground-based instruments which target the red end of the spectrum. Additional work on weak lensing measurements in the blue are provided in McCleary et al. 2023.⁶⁷ A bonus of imaging in the NUV is that no contemporary wide-field instrument has this capability making the GigaBIT measurements very complementary to other datasets.

The final piece of the puzzle is the size of the patch of sky that GigaBIT must observe per exposure. In order to accurately determine the mass distribution in a galaxy cluster, GigaBIT must be able to image the cluster, its environs, and its connection to the cosmic web. This is demonstrated by Fig. 7.1 which shows a simulation of a **single** GigaBIT detector with the HST ACS detector overlaid in yellow. In this example, HST ACS would be able to constrain the mass in the core of the cluster (dominated by strong lensing) but would have to perform multiple exposures to comment on the cluster as a whole. Note that Fig. 7.1 shows only one of several detectors that will tile the GigaBIT focal plane. A FOV of 0.25deg^2 allows for imaging many galaxy clusters of interest out to a large radii in a single exposures. The PSF must be stable to sub-resolution ($1/4$ of the FWHM) on timescales of 300 s and repeatable over days at the pixel level.

These science goals create very strict instrument requirements that will be discussed in



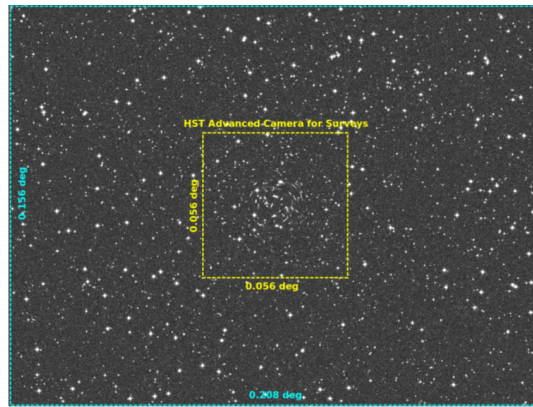


Figure 7.1: Comparison of a single GigaBIT detector and HST ACS FOV for a simulated galaxy cluster.⁶⁰ Note that the centre of the cluster is dominated by strong lensing. This is only one of several detectors that will tile the GigaBIT focal plane.

Sec. 7.5. As a stepping stone, the Super-pressure Balloon-borne Imaging Telescope (SuperBIT) was developed to bridge the technology gap and demonstrate the benefit of a balloon-based wide-field imaging observatory. The results from the SuperBIT 2023 flight directly feed into design decisions for GigaBIT and SuperBIT has been instrumental in the GigaBIT development.

7.3 OVERVIEW OF THE GIGABIT PATHFINDER: THE SUPER-PRESSURE BALLOON-BORNE IMAGING TELESCOPE (SUPERBIT)

7.3.1 INSTRUMENT

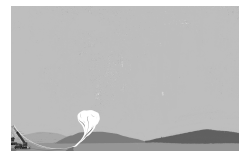
As previously discussed, SuperBIT was initially designed as a pathfinder mission for GigaBIT. SuperBIT is comprised of three nested frames that allow for yaw, roll, and pitch control of the telescope as shown in Fig. 7.2. Yaw control is achieved using a combination of a reaction wheel located in the bottom of the outer frame as well as a pivot motor at the top of the outer frame that connects to the flight train (to the balloon). The flight train is



a ladder structure which acts as a torsional spring. The middle frame performs the ± 3 deg roll control using two frameless motors (one at the bow, one at the stern). Roll control is important to enable tracking as the sky rotates during an observation. The inner frame moves the telescope in pitch using a stepper motor in line with a frameless motor. There are two orthogonal star cameras mounted to the telescope baffle used for coarse pointing of the telescope.

The design and implementation of the SuperBIT pointing system is described in Romualdez 2017¹⁰⁶ and test flight results are provided in Romualdez et al. 2018 and 2020.^{105,107} In order to point at a desired target, the right ascension and declination of the target are converted into a coarse yaw and pitch angle. SuperBIT slews to that position and then the star cameras take over. The two star cameras on the baffle first take full-frame images and use a pattern matching “lost-in-space” algorithm⁵⁶ to provide high-fidelity pointing information. In between lost-in-space solutions, the cameras select a bright star in their FOV and move in yaw/roll/pitch to keep the bright star in the same place on the detector. SuperBIT tracks a target either until it has completed the pre-selected set of images or until the target is outside the allowable roll/elevation range.

The telescope for SuperBIT (Fig 7.3) is a modified Dall-Kirkham design with a 0.5 m primary mirror. The primary mirror is made from a CLEARCERAM-Z² glass that is designed to have a very small coefficient of thermal expansion (CTE) at 273 K. The advantage of a low CTE is that as the mirror changes temperature due to the diurnal cycle, the WFE remains small. The secondary mirror and lenses are made of fused silica. Note that the lenses drastically reduce the throughput in the NUV. The primary mirror is actuated in tip/tilt and the secondary is actuated in tip/tilt/piston. In the back-end optics box there



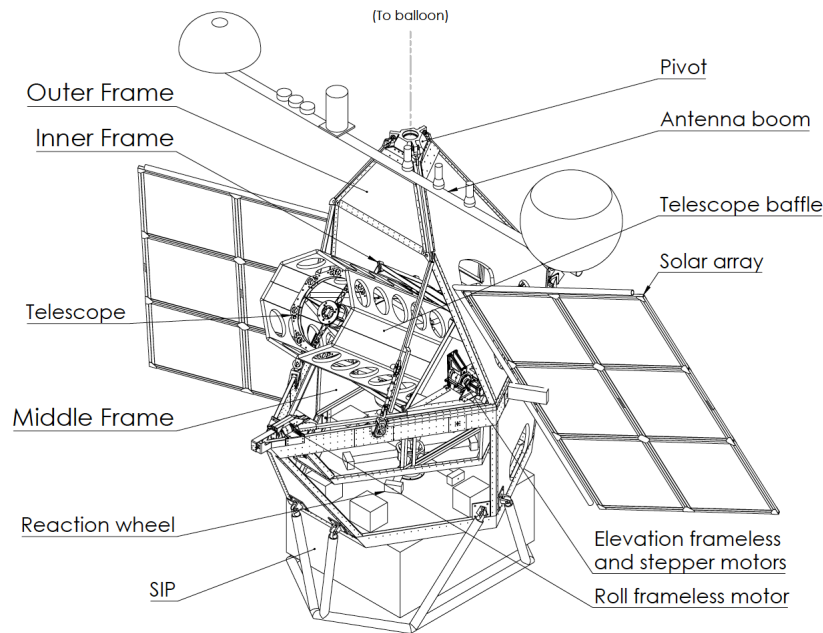


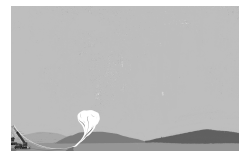
Figure 7.2: SuperBIT gondola. ¹⁰⁶ Three nested frames allow for yaw, roll, and pitch control.

is a high-speed piezo-actuated tip-tilt mirror (fine guidance mirror (FGM)) to correct for the vibrations in the structure. There are two focal plane star cameras (FSCs) in the optics box; one on either side of the focal plane. When pointing at a target, a star is placed on one of the two FSCs. The FSC as well as the rate gyroscopes are used to control the FGM at ~ 50 Hz, together forming the fine guidance sensor (FGS).

7.4 SUPERBIT 2023 FLIGHT RESULTS

7.4.1 2023 FLIGHT PROCEDURE AND GENERAL NOTES

SuperBIT launched from Wanaka, NZ on April 16, 2023. The full flight path is shown in the left panel of Figure 7.4 and the landing site is shown in the right panel. SuperBIT



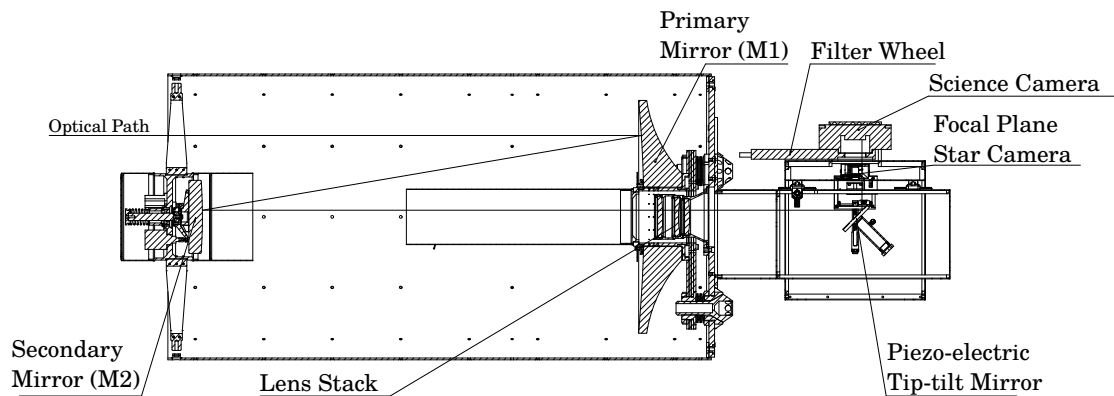
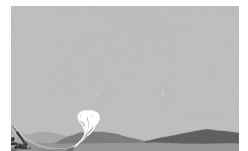


Figure 7.3: Cross section of the SuperBIT telescope.¹⁰⁷

had a functioning Starlink satellite dish for the first 15 days of flight; during this time we were able to download the flight data in real time. When Starlink stopped working, our downlinking abilities were drastically reduced and required us to physically recover the hard drives from the instrument where it landed in Argentina. The first three nights were used for alignment and calibration. This is the only time that the primary mirror was actuated. The primary mirror actuators were disabled on April 21, 2023 and were not turned on for the rest of the flight. The secondary mirror was occasionally checked for focus throughout the flight but otherwise remained static after the initial alignment procedure. There are three environmental considerations from the flight conditions that are important to mention. The first is that since SuperBIT is moving, a day-night cycle is slightly less than 24 hrs and varies depending on how fast it is moving. The next is that the stratospheric winds over the Andes were quite fast and SuperBIT was sometimes travelling up to 140 knots (260 km/hr). Even with the low air pressures (~ 5 torr) at float altitudes, traveling at that speed can affect pendulations and the coarse pointing ability of the telescope. Lastly, SuperBIT occasionally got to witness the Aurora Australis. While this was likely a beautiful



sight, the Aurora increases the diffuse background light in the images. This not only degrades the science images but also increases the required FSC exposure time for a given guide star. The longer the required FSC exposure time, the slower the FGS loop and this can impact the fine pointing stability.

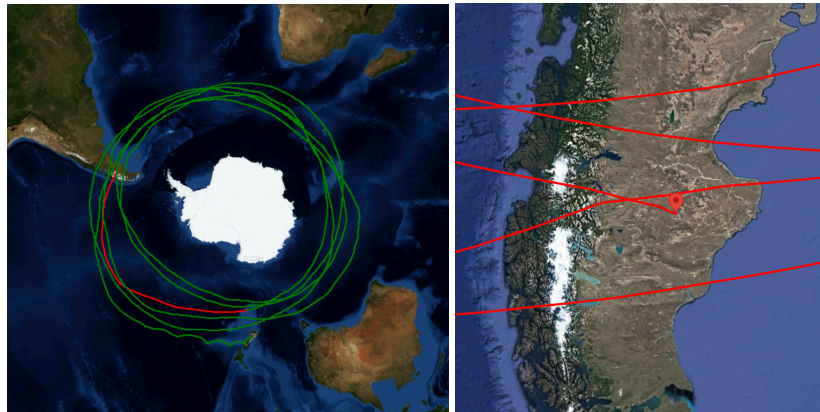
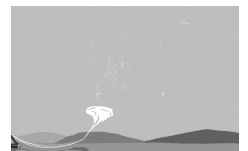


Figure 7.4: SuperBIT flight path and landing location during the 2023 flight.³²

7.4.2 COARSE POINTING RESIDUAL

As previously mentioned, SuperBIT has two FSCs, one on either side of the focal plane. During the 2023 flight, we took 12 mins of data where we simultaneously put bright stars on FSC₁ and FSC₂ and turned off the FGS to have a direct measurement of the coarse pointing system performance and the magnitude of the jitter that the FGS corrects for. Figure 7.5 shows the behaviour of the centroid of the star on each FSC. The top panel contains the X and Y position of the centroid for FSC₁ and the bottom panel contains the same for FSC₂. The behaviours of both centroids are very correlated which is a good sanity check. The centroid is a factor of two more stable in X than Y. Note that X is pitch (elevation) and Y is cross-pitch (approximately yaw). In cross-pitch, the reaction wheel instabili-



ties make it harder for the FGS to stabilize.

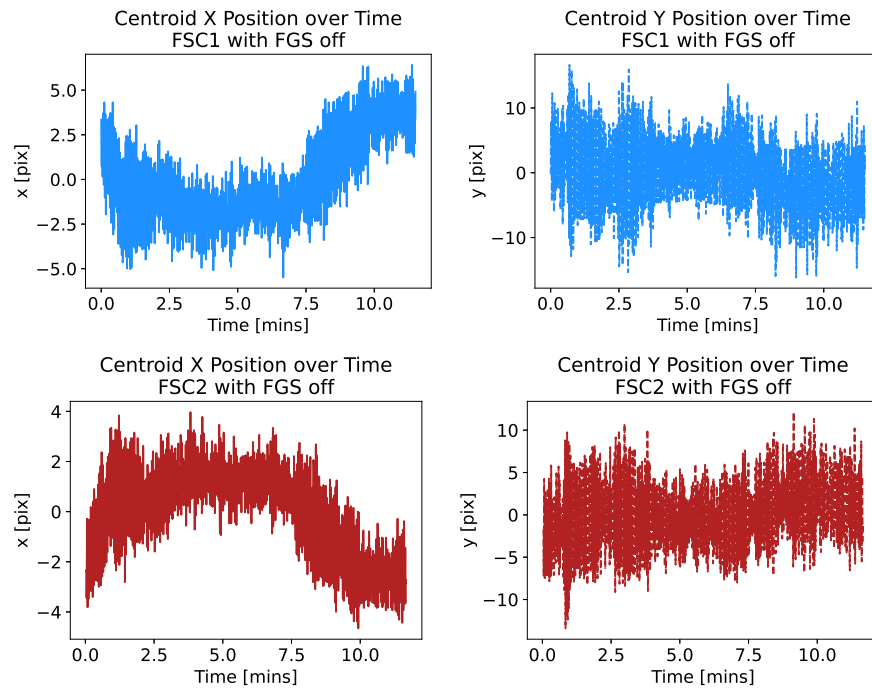
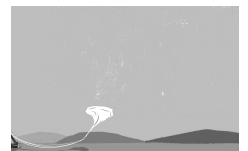


Figure 7.5: Stellar centroid vs time for both FSCs for time period when the FGS is off. The top row is FSC1 and the bottom row is FSC2. The coarse pointing system stabilizes in X a factor of two better than in Y.

In order to comment on the coarse pointing stability, I take the rolling minimum and maximum of each frame in Fig. 7.5; a sample is shown by the green curve in Fig. 7.6 for the X centroid position on FSC1. I use a 10 s window when calculating the rolling extrema which is significantly longer than the centroid sampling rate of 100 Hz. It should also be noted that even though the data sampling rate is 100 Hz, the frame rate of the FSC varies between 10–200 Hz depending on the required exposure time for a given guide star.



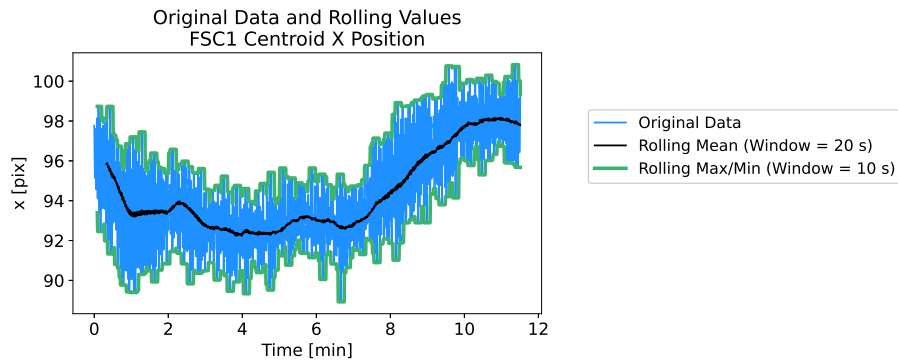
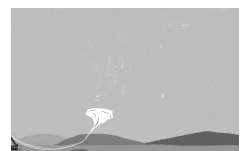


Figure 7.6: Demonstration of the rolling extrema calculation for the X centroid position for FSC1 (FGS off). The rolling extrema is shown in green, the original data is in blue, and a rolling mean is shown in black for reference. A 10 s window is used when calculating the rolling extrema which is significantly longer than the 100 Hz sampling rate.

Repeating process in Fig. 7.6 for all panels in Fig. 7.5 yields Fig. 7.7. In Fig. 7.7, I take the difference of the rolling maximum and rolling minimum to get the required throw of the tip-tilt mirror on the focal plane to correct for jitter at that point in time. I do this for the X (solid lines) and Y (dashed lines) centroid coordinates separately. FSC₁ is shown in the left panel and FSC₂ is shown in the right panel. The maximum required throw is dictated by the Y centroid position on FSC₁ and is ~ 32 pix. The angular throw of the SuperBIT tip-tilt is 2 mrad and it is located 0.169 m from the focal plane which translates to a 90 pix throw on the focal plane. The required throw to correct for the jitter (32 pix) is well within that throw and leaves a comfortable margin for the longer-timescale or static tip-tilt offsets required. This will be further discussed with respect to GigaBIT in Sec. 7.5.



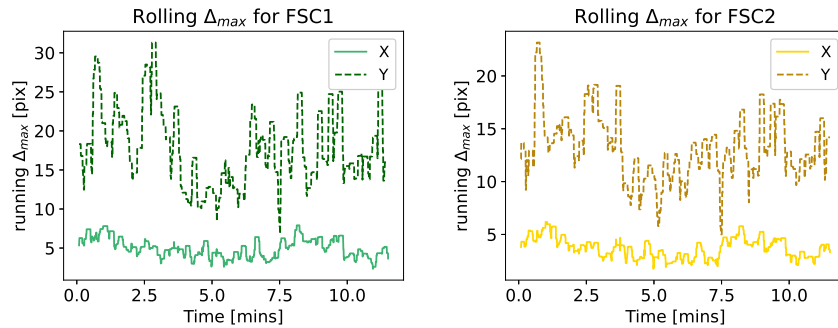
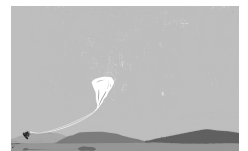


Figure 7.7: Rolling peak-to-valley jitter of the centroids on FSC1 (left panel) and FSC2 (right panel) during the 12 min window where the FGS is off. The X jitter is shown by the solid lines and the Y jitter is shown by the dashed lines. The peak Δ_{max} is 32 pix as shown by the Y curve in the left plot.

7.4.3 THERMAL PERFORMANCE

For all of the SuperBIT flights since 2017, I have been the thermal design and control lead. In this section I look into the performance of the OTA thermal control system. The Baffle of the telescope is actively controlled using the rowcolumn (RC) breakdown shown in Fig 7.8. The back of the telescope (R1) is where the science instrument is located and the front of the telescope (R3) is where the secondary mirror is located; note that R0 (not shown) is the back face of the OTA. The OTA is mounted to the inner-frame at R1C3 and R1C7 and the majority of the inner-frame electronics are located on top of the inner frame above R1C8, R1C1, and R1C2. The ‘heaters’ are lengths of Nichrome wire woven back and forth and adhered to the OTA using non-conductive tape. The OTA is covered in 50 mm thick polystyrene foam with an outer layer of aluminized mylar. The gap between R1C5 and the bottom panel of the inner frame is quite small and does not allow for the same amount of insulation as on the top.

In Fig. 7.9 I show the performance of the OTA thermal control system for a 48 hr time period. The diurnal cycle of the OTA is quite consistent so this is a representative dataset.



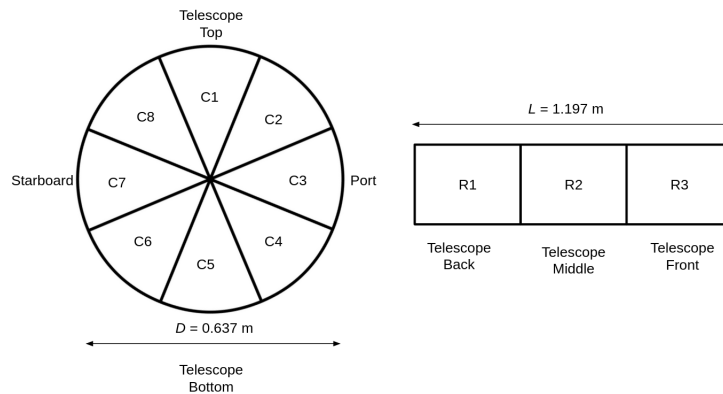


Figure 7.8: Heater breakdown naming convention for the OTA. Note that R0 is not shown here but it is the back plate of the OTA.

The grey panels show the nights (sun elevation $< -9^\circ$) and the white panels show the days; note that 48 hrs corresponds to slightly more than two sunsets due to the speed and direction (East–West) that SuperBIT orbits the southern hemisphere. From top to bottom in Fig 7.9 we go from R0 at the back of the telescope to R3 at the front of the telescope. The rows are controlled to different temperatures in a way that reduces power consumption while minimizing the thermal-induced WFE.

We wish to control the primary mirror to 0°C to minimize the coefficient of thermal expansion of the mirror so R0 (purple) and R1 (pink) are set slightly warmer to 1°C . We can see that R0 is actually only actively controlled during the day and is warmer than the set-point during the night due to the thermal mass of the back plate. A similar occurrence can be seen by top three R1 segments near the inner-frame. The power dissipated by the inner-frame leaches into the OTA and brings the top three R1 segments above the set point. It took a couple days for the diurnal trend of the OTA to become established at which point the telescope had been aligned. There was concern with respect to both power consumption and having to realign the telescope if we increased the R1 set point to have the seg-



ments actively controlled at all times during the night. The decision to keep the R0 and R1 set-point at 1 °C was further supported by the primary mirror thermal behaviour which will be discussed later in this chapter.

For R2 and R3 we can see that all segments are actively controlled for the majority of the night. Occasionally one or two segments are uncontrolled at the start of the night but it is a relatively small fraction of the night and those segments do not have as large of an impact on the thermal stability of the mirrors as R0 and R1. It is also relevant to note that, when the control system is active, the segments are stabilized to the set-point within a standard deviation of <math><0.1\text{ K}</math> on average.

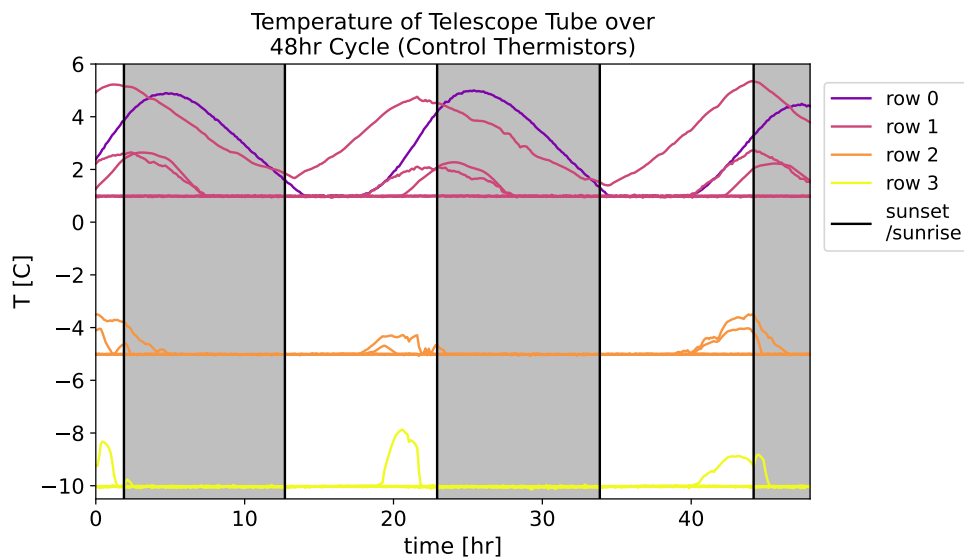
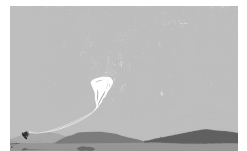


Figure 7.9: Diurnal cycle of OTA control thermistors. Row 0 is the back of the OTA and row 3 is by the secondary mirror. To conserve power, the row 2 and 3 set points are colder than row 0 and 1. The row 1 segments at the top of the OTA (near the inner frame electronics) are above the set-point for the majority of the night and thus are not actively controlled.

In addition to the thermistors used as feedback to the heaters, SuperBIT also has a number of passive thermistors. In Fig 7.10 I show the behaviour of the reference thermistors on



the OTA. There are three at R0 (purple) and one at R3 (yellow). The R0 thermistors oscillate by up to 5°C over the course of a day. The behaviour of the R0 reference thermistors is expected as it matches the behaviour of the control thermistor shown in Fig. 7.9 which is rarely actively controlled. The R3 reference thermistor behaves much better as it oscillates around the -10°C set-point with a standard deviation of 0.5°C .

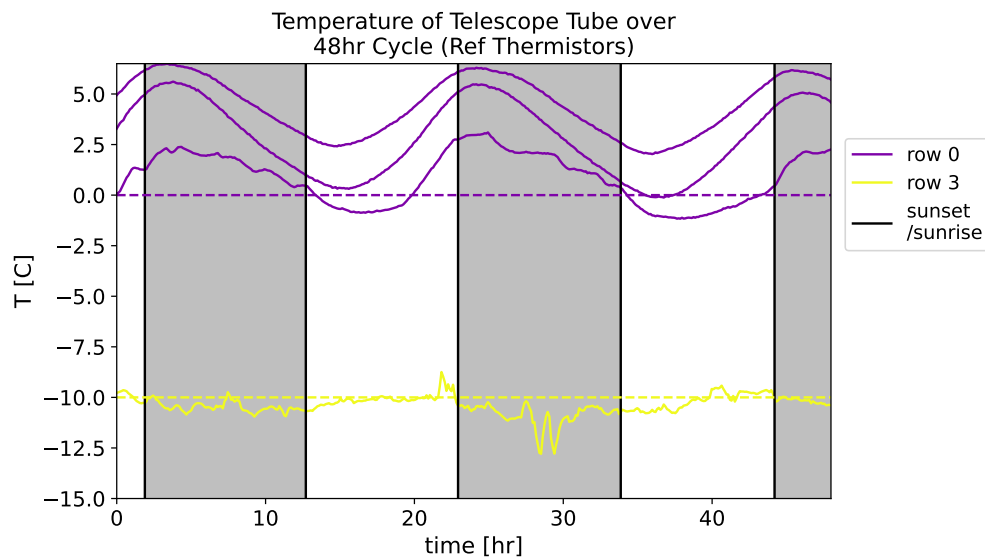


Figure 7.10: Behaviour of reference thermistors on the OTA over a 48 hr period. Purple indicates R0 and yellow indicates R3. The R0 thermistors oscillate by up to 5°C as that row is poorly controlled as shown by Fig. 7.9. The R3 reference thermistor is held at the set-point within a standard deviation of 0.5°C .

As mentioned above, one of the main reasons we control the temperature of the OTA is to stabilize the primary mirror temperature at 0°C . There are 13 thermistors epoxied to the back or edge of the primary mirror. The naming scheme is similar to that of the OTA and uses the same clocking layout with but R indicates the radial distance from the centre of the mirror. Fig. 7.11 shows the temperature of the primary thermistors over the same 48 hr time frame as the previous plots. Note that a cyclic colormap is used for this plot so the lighter colours are the indicate a clocking near the top of the mirror (near the inner frame



electronics) and the darker colours are at the bottom. We can see the same diurnal cycle seen by the R0 and R1 rows of the OTA in Fig. 7.9 and there is a gradient of up to 2°C across the mirror. A preliminary study in Redmond 2018⁹⁵ indicated it was desired to keep the temperature gradient across the primary mirror below 1.5°C but that conclusion was derived using the original telescope which was much more susceptible to thermal-induced WFE due to the method in which the primary mirror was constrained. Also, all of these thermistors are located on the back of the primary mirror so it is possible that the mirror face is more isothermal.

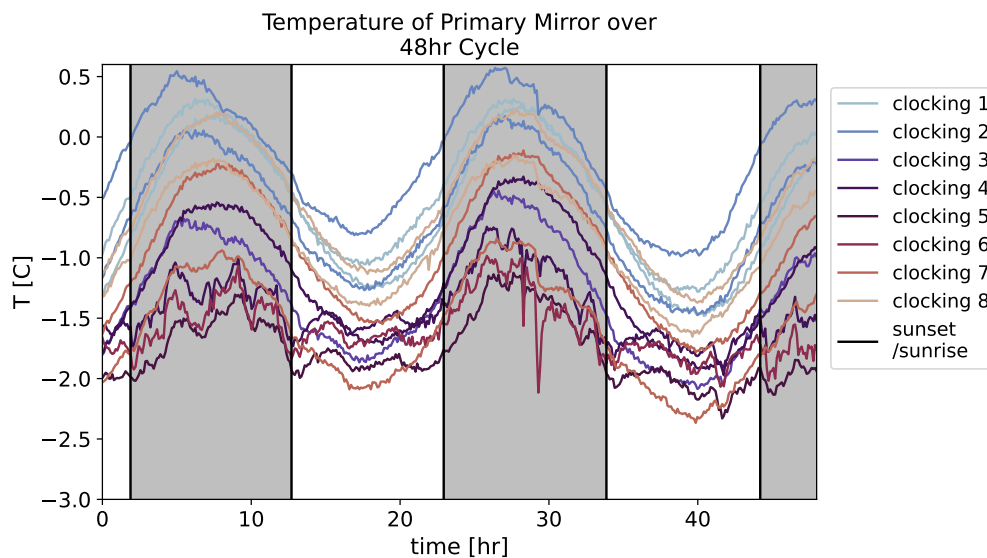


Figure 7.11: Temperature of the primary mirror reference thermistors over a 48 hr time period. A cyclic colormap is used so the lighter colours correspond to clocking locations towards the top of the telescope and the darker colours are towards the bottom of the telescope. The mirror is colder on the bottom and warmer on the top which is to be expected based on the behaviour of the OTA.

While Fig. 7.11 provides information on the diurnal nature of the primary mirror temperature, it is difficult to really see the clocking-dependant nature of the thermal behaviour. In Fig 7.12 I show the mean and peak-to-valley temperature over the course of the two



nights shown in Fig. 7.11 for each thermistor on the primary based on the clocking index. If a clocking has more than one dot, that is because there are two thermistors but they are at different radial segments. The same colormap is used for Fig 7.12 and 7.12. The left panel in Fig 7.12 shows that the primary mirror is colder on the bottom than the top. This is expected based on the poorly-controlled OTA R1 segments. In the right panel I show the peak-to-valley of the temperature for the two nights studied. The peak-to-valley of the temperature does not appear to depend on the clocking so the mirror seems to maintain the same gradient as it warms and cools.

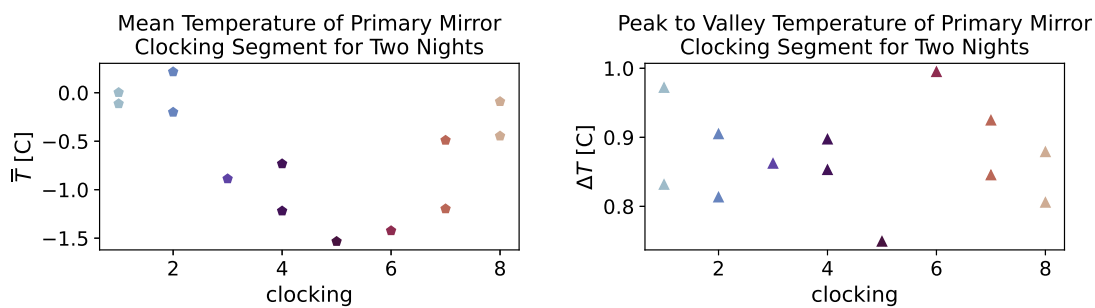
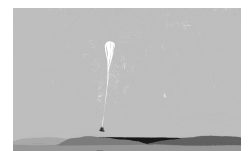


Figure 7.12: Clocking-dependent thermal behaviour of the primary mirror. In the left panel we have the mean temperature of each thermistor for the two nights shown in Fig. 7.11; note the same colormap is used for both figures. We see that the bottom of the mirror (C3–C6) is colder than the top of the mirror. In the right panel we have the peak-to-valley for the same dataset. There does not appear to be a correlation between the peak-to-valley temperature seen by a primary mirror segment based on the clocking.

This study of the thermal behaviour of the SuperBIT OTA provides us with valuable information which on the systematics to be wary of in the 2023 SuperBIT data, how to improve for the next SuperBIT flight, and insights for designing GigaBIT. There is still a wide range of work that can be done with the thermal performance data from the 2023 flight. The first step would be to perform a similar study over the entirety of the flight and look for longer timescale trends. Another important future study is to generate a thermo-



mechanical model of the new SuperBIT optics and use this data as an input to more accurately predict the WFE contribution of the diurnal OTA behaviour.

7.4.4 PSF STUDY SETUP

As previously discussed, both SuperBIT and GigaBIT need to have small, round, and stable PSFs in order to avoid introducing systematics into the weak-lensing measurement.

To study the PSF quality throughout the SuperBIT 2023 flight, I use the PSF Extractor¹⁰ (PSFEx) package to generate a model of the PSF at each location on the focal plane for each science image.

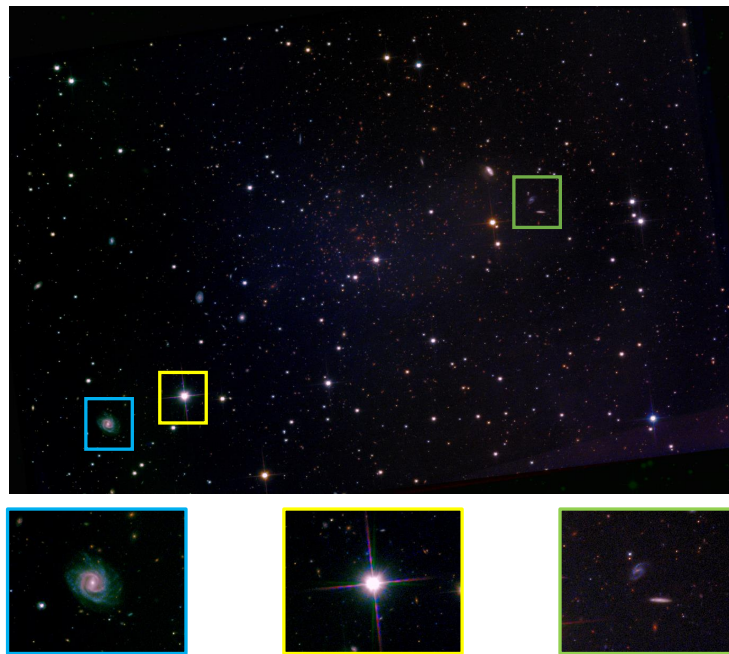


Figure 7.13: Source extraction examples for target PLCKG287. This is a color image created by aligning and stacking all of the exposures from the red, green, and blue bands. Note that non-linear scaling has been applied to the colour scaling. The bottom panels show examples of objects that should not be used for creating a PSF model. In blue there is a small spiral galaxy, in yellow there is a saturated star, and in green there is both a small spiral galaxy as well as an edge-on galaxy.



The images are first processed using SExtractor¹¹ which classifies the point sources in the image and generates an LDAC-format FITS catalog. This is an important step as the PSF models should only be built from un-saturated stars in the image. As discussed in Sec. 2.1, a star is a near-perfect point source and thus can be used to gain un-biased information regarding the optical system. An image of a galaxy is the convolution of the galaxy shape and the PSF and should not be used to generate a PSF model. The classification is a multi-step process which first removes saturated sources, then separates blended sources, and finally uses a neural net to distinguish between galaxies and stars. A sample of objects that should not be used when generating a PSF model are shown in Fig. 7.13. Moderately bright stars can be identified using by looking at the magnitude of the object and its half-light radius as stars will follow the stellar locus shown in Fig. 7.14.^{1,43} A classification example is shown in Fig. 10 of Bertin et al. 1996.¹¹ There will be some contamination of the point source selection due to dim galaxies that get included in the sample. This effect should be isotropic but will introduce some degree of bias.

Once the stars have been identified, PSFEx is handed the SExtractor catalog for that exposure and it generates a model of the PSF for each source using a linear combination of a user-selected vector basis; for this thesis the default pixel basis is used. With a sampling of the PSF across the focal plane, PSFEx then interpolates to provide a model of the PSF at each location on the focal plane. Note that the quality of the interpolation across the focal plane depends on the distribution and quantity of point sources in the image. For this reason, only the ‘science’ targets are considered as they are primarily galaxy clusters with an abundance of evenly distributed point sources. This is not true for ‘glamour’ targets such as the M83 Pinwheel Galaxy.



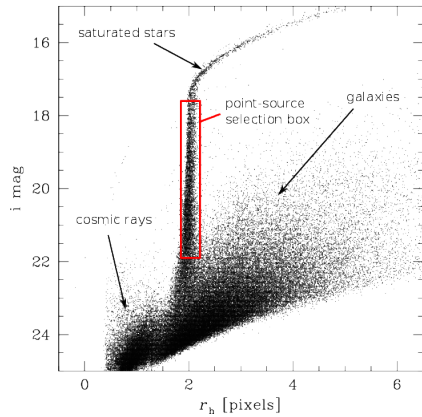
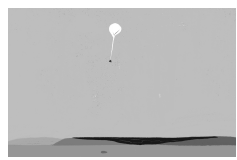


Figure 7.14: Point source selection box used by PSFEx based on the magnitude and half-light radius of an object identified by SExtractor.¹ The stellar locus described in Kaiser et al. 1995⁴³ is shown in the red box.

The output from PSFEx (a .psf file) enables the generation of a representative PSF at each pixel on the SuperBIT focal plane. For the purpose of this analysis, a PSF model for each of the 6.2×10^7 pixels is needless since the PSF changes relatively slowly across the focal plane. Discretizing the focal plane into 288 metapixels captures the location-dependent nature of the PSF while drastically expediting the analysis. The breakdown of the focal plane is shown in Fig. 7.15 where each square is a ‘metapixel’ representing a 500×400 pixel region of the detector.

1	2	3	4	...	21	22	23	24
25	26	27	28	...	45	46	47	48
...
241	242	243	244	...	261	262	263	264
265	266	267	268	...	285	286	287	288

Figure 7.15: The focal plane is nominally 6422×9600 pixels but is broken up into 288 metapixels (12 rows, 24 columns) to facilitate and expedite the analysis.

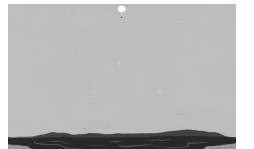


The PSF models generated by PSFEx are very non-gaussian which makes it difficult to generate statistics on the PSF quality. The main two parameters I consider for this study are the full width at half maximum (FWHM) of the PSF and the ellipticity. To measure these parameters across the focal plane, I load the PSFEx for a single exposure and fit a 2D Gaussian (using `astropy`) to the PSFEx generated PSF at $m = 288$ locations across the focal plane. The fitting function returns the maximum and minimum standard deviation of the 2D Gaussian as well as the angle Θ between the horizontal and the axis of the maximum standard deviation. Using these outputs the parameters of interest are calculated via

$$fwhm = \frac{1}{2} (\sigma_{max} + \sigma_{min}) \quad (7.1)$$

$$e = \frac{\sigma_{max} - \sigma_{min}}{\sigma_{max} + \sigma_{min}} \quad (7.2)$$

where e is the ellipticity. Note that the FWHM is in units of pixels and the ellipticity is dimensionless. Any image that has a FWHM < 1 pixel anywhere in the image is discarded as it is assumed to not have a sufficient density of point sources for PSFEx to interpolate the PSF accurately. A sample PSFEx model and associated 2D Gaussian fit are provided in Fig. 7.16; note that this sample is a **strategically chosen worst-case scenario** to demonstrate the accuracy of the ellipticity approximation.



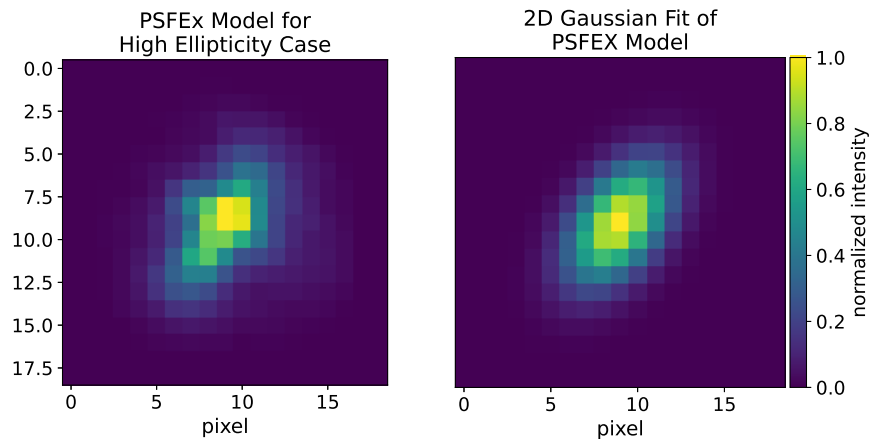
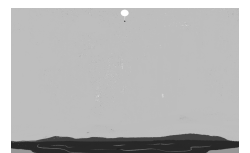


Figure 7.16: PSFEx generated SuperBIT PSF and associated 2D Gaussian fit for a **strategically chosen worst-case scenario** high-ellipticity PSF. The 2D Gaussian captures the structure sufficiently well to use the ellipticity values produced as a metric of the PSF quality. For this example, the 2D Gaussian fit reports the FWHM as 4.84 and the ellipticity as 0.26.

To summarize, the high-level procedure for using the PSFEx output to study the PSF quality is as follows:

1. Load the PSFEx output (.psf file) for a single blue-band 300 s exposure
2. Generate PSFEx models of the PSF at m locations across the focal plane
3. Fit a 2D Gaussian to each PSFEx model
4. Calculate FWHM and e from the 2D Gaussian fit
5. Discard target if $\text{FWHM} < 1$ at any of the m locations.

It should be noted that the SuperBIT PSF is not Gaussian, even for perfectly manufactured optics and a perfectly aligned telescope. To sanity check that the 2D Gaussian was producing an informative result, I looked at the residual between the PSFEx generated PSF and the 2D Gaussian. In Fig. 7.17, I take the average residual over all metapixels and all



exposures and compare that to the residual between an Airy disk and the 2D Gaussian fit of an Airy disk. The residuals are similar with a well in the middle and then a bright ring followed by another well. This indicates that the main contributor to the 2D Gaussian-PSFEx residual is from the non-Gaussian nature of the SuperBIT PSF. This residual causes an overall offset in the FWHM measurement but it is consistent across the focal plane and with respect to time so the 2D Gaussian fit results can still be used to draw conclusions on the PSF stability.

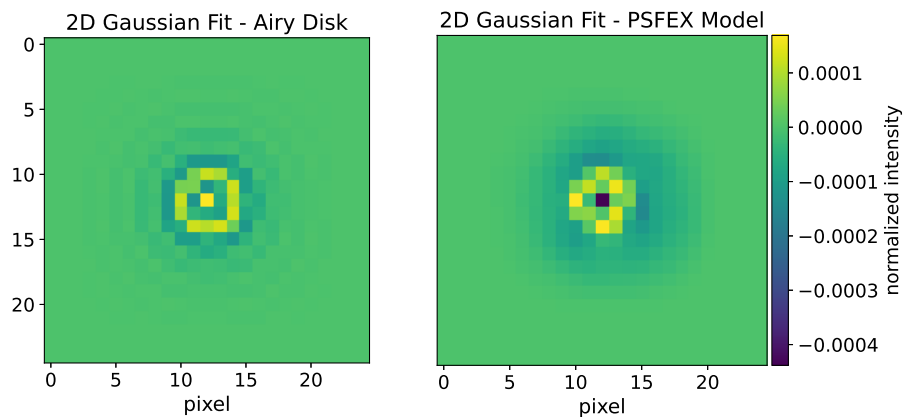
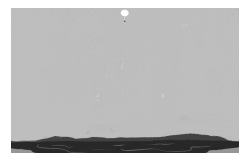


Figure 7.17: Residual comparison between a 2D Gaussian fit and an Airy disk vs a 2D Gaussian fit and a PSFEx generated SuperBIT PSF. The two have a similar structure showing that the dominant contribution to the residual is intrinsic to the non-Gaussian nature of the SuperBIT PSF.

Before moving forward with the analysis, it is also important to verify that the variation of the FWHM and ellipticity within the m metapixels is small. In order to check this, I divide each metapixel into a 3×3 grid. I then look at the statistics for each metapixel based on the 9 sub-metapixels. Colourmaps for the FWHM and ellipticity results for a single exposure are shown in Fig. 7.18. The left panels are the means of the sub-metapixels and the right panels are the standard deviations of the sub-metapixels for each metapixel. For both the FWHM and ellipticity, the standard deviation is correlated with the mean. This



is because the large FWHM or ellipticity values are also in areas of larger gradients which increases the variance.

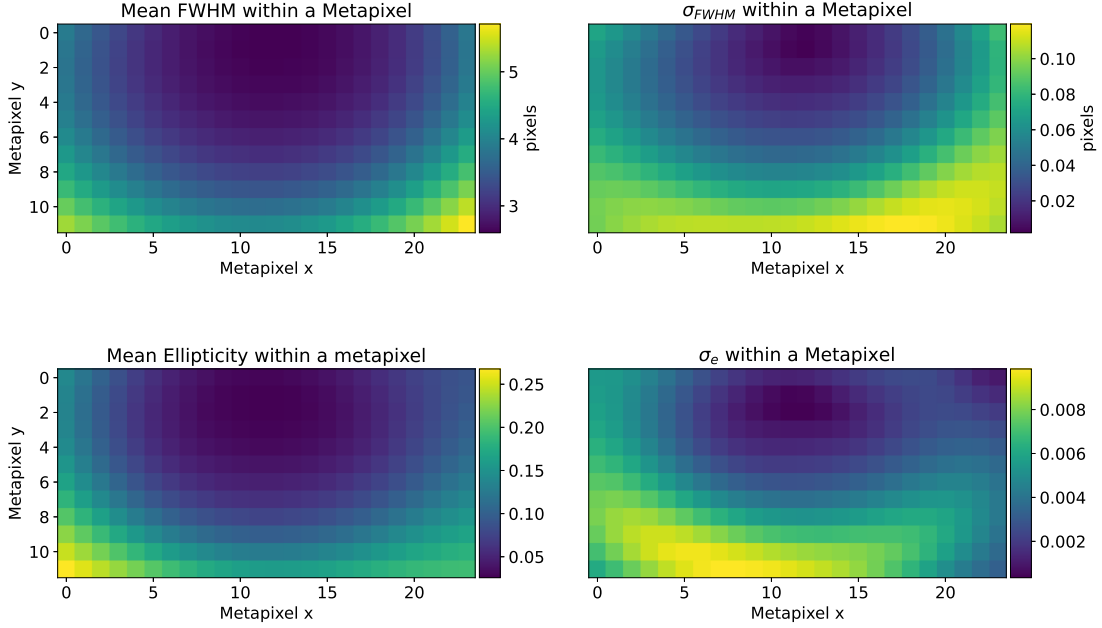


Figure 7.18: Validation of the metapixel breakdown using the PSFEx output for a single exposure. The top row is the FWHM statistics and the bottom row is the ellipticity statistics. The left panel is the mean value of the sub-metapixels for each metapixel. The right panel is the standard deviation of the 9 sub-metapixels for each metapixel. The standard deviations track the mean as locations with high mean values are also in regions with a larger gradient of values.

Note that Fig 7.18 is only for a single exposure but it is representative of the full dataset. This is shown by Fig. 7.19 which is the mean of the standard deviations in each metapixel for the entire flight. For the ellipticity this can be expressed as

$$\bar{\sigma}_{e_k} = \frac{1}{N} \sum_{j=1}^N \left(\sqrt{\frac{1}{9} \sum_{i=1}^9 (e_{i,j} - \mu_{e_{k,j}})^2} \right) \quad (7.3)$$

where k is the metapixel, N is the number of exposures, i is the sub-metapixel, and $\mu_{e_{k,j}}$ is the mean of the sub-metapixel ellipticity values for exposure k . As shown by Fig. 7.18 and

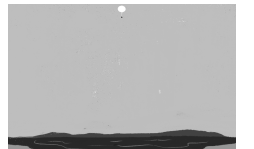


Fig. 7.19, the standard deviation within a metapixel for both the FWHM and ellipticity is $<10\%$ of the mean and is much less than the variations we see image-to-image so the chosen metapixel breakdown is determined to be acceptable for this application. For this reason, the error bars associated with the metapixel variation are not included in the following sections.

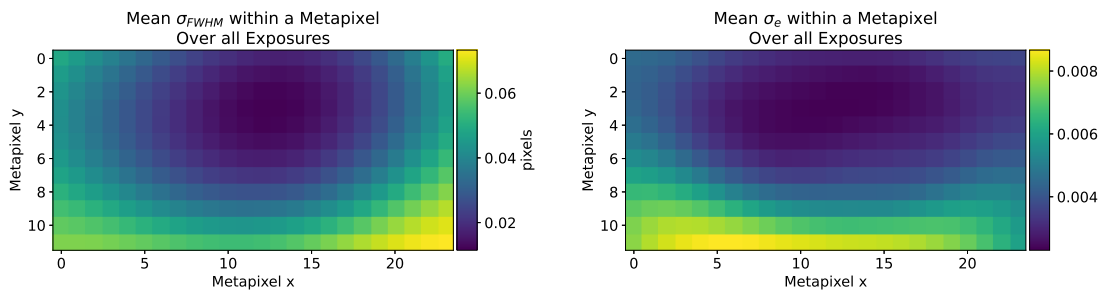
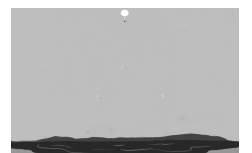


Figure 7.19: Mean standard deviation of FWHM (left) and Ellipticity (right) within a metapixel over all exposures as calculated by Eqn. 7.3.

Based on the flight condition variability discussed in Sec. 7.4.1 and the thermal performance discussed in Sec. 7.4.3, I investigate if the PSF qualities are related to the telescope elevation (or pitch) angle, the FSC used for the FGS, the sun elevation (or time since sunset), and the amount of time since the first exposure on a target. I determine if these variables influence the PSF through correlation matrices; if there is a trend between the correlation matrices I then investigate further. There are two types of correlation matrices, one for the PSF quality variables (FWHM and ellipticity), and the other for the observing variables (elevation, FSC, sun elevation, and time).

First let us calculate the correlation matrices for the PSF FWHM and ellipticity. For each exposure of a target, I generate a FWHM map and ellipticity map (like the left panels in Fig. 7.18) and calculate the Pearson’s correlation coefficient between each image taken for



that target for the last 35 days of the flight. This generates a correlation matrix R where each entry is calculated via

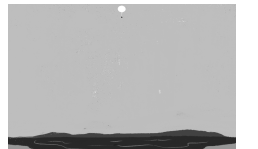
$$R_{i,j} = \left| \frac{\sum_{k=1}^m (M_k^i - \bar{M}^i) (M_k^j - \bar{M}^j)}{\sqrt{\sum_{k=1}^m (M_k^i - \bar{M}^i)^2} \sqrt{\sum_{k=1}^m (M_k^j - \bar{M}^j)^2}} \right| \quad (7.4)$$

where M^i is the ellipticity or FWHM map for exposure i of the target in question and m is the number of of metapixels. I take the absolute value since, for the purpose of this study, the sign of the correlation coefficient is not relevant. Note that the diagonal of R is ones and if the PSF did not change at all, the FWHM and ellipticity maps would be the same for every exposure and the entire matrix would be ones.

Next up we calculate the correlation matrices for the observation variables. These matrices are essentially one minus the relative change of the observation variable between exposures (telescope elevation, FSC used, sun elevation, and elapsed time). The relative change matrices are calculated via

$$R_{i,j}^c = 1 - \frac{|v_i - v_j|}{\Delta v_{max}} \quad (7.5)$$

where v_i is the value of the variable value when exposure i was taken of the target and Δv_{max} is the maximum possible change in that variable. One exception for Δv_{max} is the elapsed time R^c matrix where Δt_{max} is taken to be the time between the first and last exposure for the target, not for the entire flight. Here we can see that R and R^c should have similar behaviour where the diagonal is ones and the more similar the data provided, the higher the value. Figure 7.20 shows the R matrices for target Abell 3526 (Centaurus Cluster) and is



representative of the behaviour seen for all targets. Note that the data is sorted with respect to time before creating these matrices so it is expected for the elevation and sun elevation plots to change relatively smoothly. The ellipticity R matrix is in the left-most panel and has clear structure that relates to each of the R^c matrices provided. The general blocking lines up with time and the finer structure matches the FSC, elevation, and sun elevation. This plot supports my hypothesis that, when investigating the contributors to the PSF quality, the variables plotted in Fig. 7.20 are a good place to start. In the following sections I delve deeper into how the PSF changes as a function of the telescope elevation, FSC, and sun elevation.

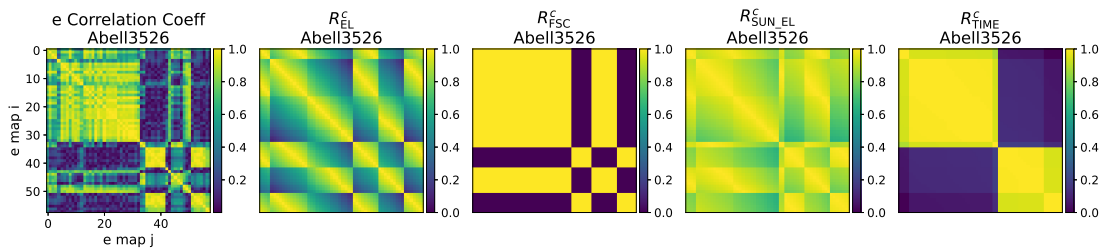
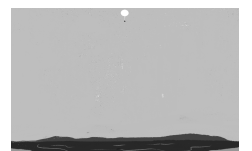


Figure 7.20: Correlation matrices for Abell 3526. The left panel shows the Pearson correlation coefficient for the ellipticity maps generated for each exposure of Abell 3526. The right panels show how the telescope elevation, sun elevation, FSC, and time differ between each Abell 3526 exposure. The structure in the ellipticity matrix is clearly correlated with a combination of the variables shown.

7.4.5 PSF AS A FUNCTION OF LOCATION ON THE FOCAL PLANE FOR FSC₁ TARGETS VS FSC₂ TARGETS

The first variable I investigate is the FSC used for feedback to the FGS. This can have implications on the PSF quality for three reasons:

1. The FSCs are located on opposite sides of the focal plane
2. The FSCs are different cameras and have different pixel sizes and frame rates



- The FSCs have different read noise characteristics.

Figure 7.21 shows the average FWHM and ellipticity over the focal plane for FSC1 (top row) vs FSC2 (bottom row). The exposures that contribute to Fig. 7.21 are restricted to be when the sun elevation is less than -50° (the middle of the night) and when the telescope elevation is between $37-42^\circ$. This should remove the other PSF systematics and allow us to see the effect of the FSCs. When these data cuts are applied, we get 18 exposures for FSC1 and 20 exposures for FSC2. As shown in Fig. 7.21, the results are quite similar for the two FSCs. FSC2 outperforms FSC1 with respect to the ellipticity but the structure and magnitude of the FWHM and ellipticity across the focal plane is largely the same.

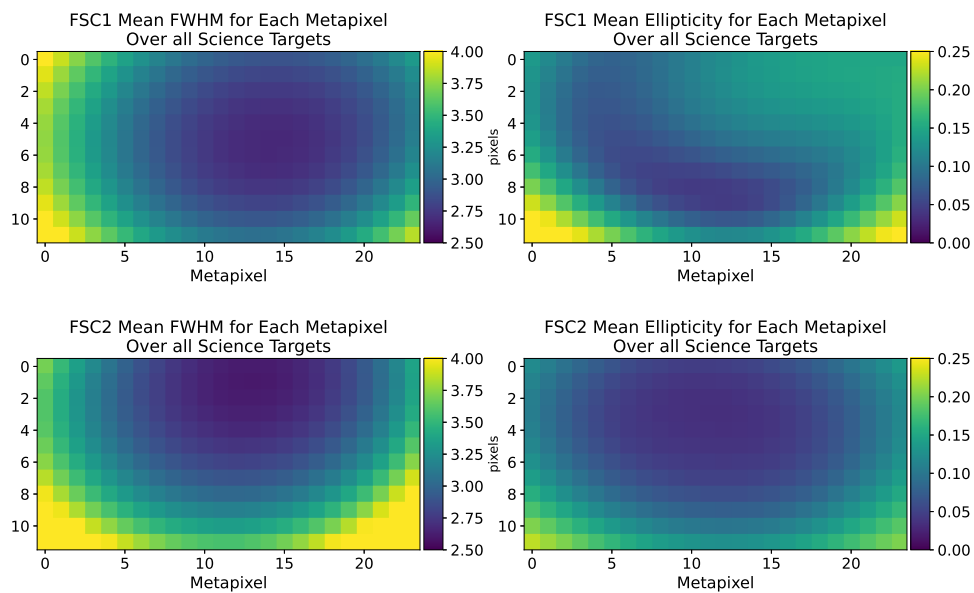


Figure 7.21: Mean FWHM (left column) and ellipticity (right column) maps for targets that used FSC1 (top row) for tracking vs FSC2 (bottom row). The dominant structure in these maps is due to the optics; due to SuperBIT's large FOV, the PSF degrades towards the edges of the focal plane. The effect of the optics is likely not symmetric due to the beam being slightly off centre on the focal plane. Outside of the dominant optics effect, we can see that the FSC1 and FSC2 ellipticity and FWHM maps have slightly different structure. FSC2 outperforms FSC1 when it comes to ellipticity but other than that they are comparable.



I further quantify the results from Fig 7.21 by looking at the 0th metapixel column in Fig. 7.22. The y-axis is the metapixel row and the x-axis is the FWHM (left panel) or ellipticity (right panel). The markers are the mean values (shown in Fig. 7.21) and the errorbars are the standard deviation of the values of the 42–44 exposures that contribute to Fig. 7.21. A reminder that the variance of the PSF within a metapixel was determined to be small (Sec. 7.4.4) and thus the errorbars in Fig. 7.22 are due to true variations of the PSF exposure-to-exposure. In the right panel we can see that the FSC2 images have a rounder PSF across all metapixel rows. In the left panel we see that the FWHM is better for FSC2 at the top of the focal plane and then there is a cross over where FSC1 becomes slightly better towards the bottom of the focal plane.

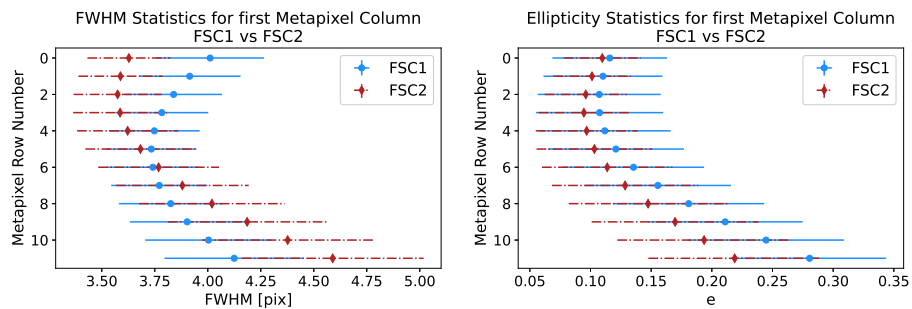


Figure 7.22: PSF statistics by FSC for first column of metapixels. Solid blue dots are FSC1 and red dashes are FSC2. 42 exposures are used for the FSC1 data and 44 exposures are used for the FSC2 data after cuts are made to restrict the telescope and sun elevations. In the left panel we have the FWHM which shows a slight trend for FSC2 to have a better PSF at the top of the focal plane and FSC1 to have a better PSF at the bottom of the focal plane. The ellipticity is shown in the right panel and FSC2 slightly outperforms FSC1 across the focal plane.

Figure 7.21 and 7.22 provide us with additional context with respect to the correlation plots shown in Fig. 7.20. We can see that the distributions of the FWHM and ellipticity across the focal plane are slightly different for FSC1 vs FSC2. We can also go the other way and use Fig. 7.20 to inform us on why the errorbars in Fig. 7.22 largely overlap. Though



I have shown that the FSC has a small contribution to how correlated the PSF quality is image-to-image, there are other contributors that I have not controlled for when making the data cuts. The main two residual factors are (1) the absolute time (discussed in Sec. 7.4.4) and (2) the FSC exposure time. The brighter the guide star placed on the FSC, the shorter the exposure time and the faster the control loop can run. The influence of the FSC exposure time is an important future research area especially as we make design decisions for GigaBIT.

7.4.6 INVESTIGATION OF DIURNAL PSF TRENDS

Since we see that there is a diurnal thermal trend of the primary mirror in Sec. 7.4.3, I want to investigate if we can see a similar trend in the PSF quality. This effect is studied by looking at the PSF quality as a function of the sun elevation. This study is also motivated by the similarities seen in Fig. 7.20 between the sun elevation and the PSF parameter correlation coefficient.

Figure 7.23 shows the FWHM and ellipticity vs sun elevation for a sample set of metapixels over the course of the flight. Sunset is on the left of the plots and sunrise is on the right. As with the FSC study, data cuts are made to reduce the overlap of systematics in the exposures used. All points use FSC₂ as the guide camera and the telescope elevation is between 37–42°. After these data cuts, there are a total of 115 exposures to consider. The PSF parameters are plotted individually for each exposure. The left panel shows the FWHM; as expected the center of the focal plane (green) has the smallest FWHM followed by the top-left corner (purple) with the worst-case being the bottom left (blue). This distribution is consistent across all sun elevations and is dominated by telescope optics as discussed in



Sec. 7.4.4. The bottom left metapixel is the only location on the focal plane that we see any plausible trend in the FWHM or ellipticity where it consistently improves over the course of the night. Figure 7.23 shows that the correlation between the FWHM and ellipticity Pearson's coefficients shown in Fig. 7.20 and the sun elevation is likely driven by the metapixels around the edge of the focal plane that are more sensitive to the thermal-induced WFE. This effect is quite small and not the main contributor to the PSF variations throughout the flight.

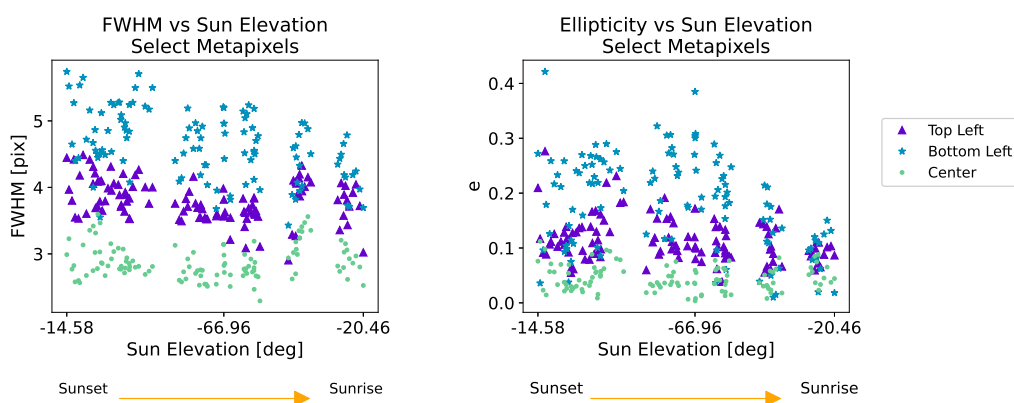


Figure 7.23: PSF metrics as a function of sun elevation. All 115 exposures that contribute to these plots use FSC2 as the tracking camera and have a telescope elevation between $37\text{--}42^\circ$. Sample points are taken from three locations on the focal plane: the top-left (purple), bottom-left (blue), and center (green) which are metapixels 1, 265, and 132 respectively. Sunset is on the left of each plot and sunrise is on the right. The left panel shows the FWHM; there is a slight trend in the blue points where the FWHM decreases as a function of time-since-sunset but it is not a large effect. The right panel shows the ellipticity and has similar behaviour to the FWHM. The top and centre of the focal plane are consistent throughout the night and the bottom corners of the focal plane improve slightly as the instrument thermalizes.

7.4.7 TELESCOPE ELEVATION

I now investigate if the telescope elevation has a measurable effect on the PSF quality. To isolate the elevation effect I only consider exposures where FSC2 is used for tracking and the sun elevation is $< -50^\circ$. I also use 5° telescope elevation bins to acquire statistics on the



elevation-dependent behaviour. With the data cuts described and the binning applied, I get between 10–67 exposures to use for each telescope bin. The breakdown is provided in Tab. 7.1; the upper two elevation bins have the least exposures which should be considered when looking at the plots that follow.

Bin Centre [deg]	# Exposures
25.5	67
30.5	67
35.5	47
40.5	46
45.5	27
50.5	10
55.5	20

Table 7.1: Number of exposures available for each telescope elevation bin when the sun elevation is restricted to be $< -50^\circ$ and only exposures tracking with FSC2 are used.

Figure 7.24 is a similar plot to the sun elevation plot but here I have binned the telescope elevation range and provided statistics. Each errorbar uses n exposures as listed in Tab. 7.1. Note that the 50.5° bin only has 10 exposures that contribute, the rest have >20 . The same metapixels are plotted as in Fig. 7.23 and we see the same metapixel-based trend where the centre metapixel (green dot) behaves the best, followed by the top left (purple triangle), and finally the bottom left (blue star). For the both the FWHM and the ellipticity, there is no significant trend with respect to the telescope elevation.



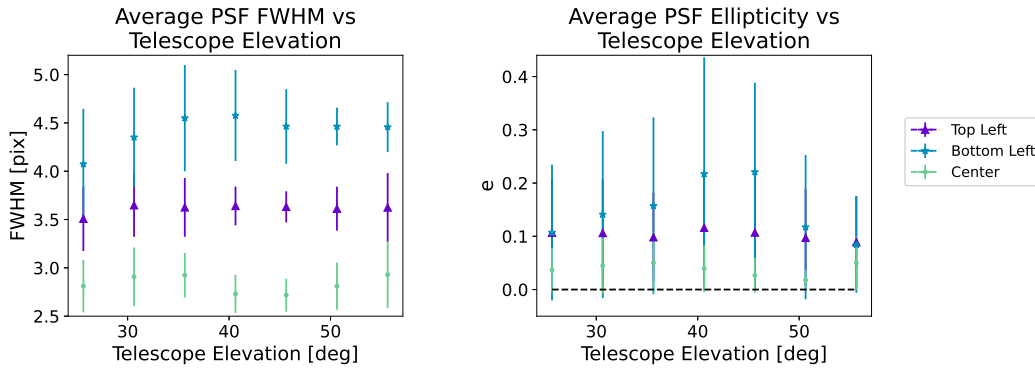


Figure 7.24: PSF qualities as a function of telescope elevation; for all exposures FSC2 is used for the FGS and the sun elevation is $<-50^\circ$. The telescope elevation is discretized into 5deg bins which is how the statistics are done. The left panel is the FWHM and the right panel is ellipticity. Three sample metapixels are provided. The center metapixel (green dot) has the best PSF quality and the smallest variance. The next best is the top left metapixel (purple triangles) and the worst is the bottom left metapixel (blue stars). This metapixel-based trend is due to the static optics error. There does not appear to be a significant elevation-based trend.

7.4.8 SUPERBIT RESULTS SUMMARY

In summary, we know based on the correlation plots (Fig 7.20) that the quality of the PSF is correlated with time, the FSC used for tracking, sun elevation, and telescope elevation. When we control for all other variables we only see a significant trend in the FWHM and ellipticity when we compare exposures that use FSC1 for the FGS vs FSC2. A reminder that the variance of the PSF within a metapixel was determined to be small (Sec. 7.4.4) and thus the trends seen in Fig. 7.22 are due to true variations of the PSF exposure-to-exposure. Even though there is a FSC-based trend, the variance suggests it is only marginally significant. More investigation is required to find the dominant cause of the PSF variation.



7.5 GIGABIT INSTRUMENT SPECIFICATIONS

7.5.1 OVERVIEW

From Sec. 7.2 we have a number of requirements that drive the instrument design. Assuming a circular aperture, the angular resolution of a telescope is calculated via $\theta = 1.22\lambda/D$ where λ is the wavelength and D is the diameter of the aperture. Given that the desired resolution in the NUV is 0.081 arcseconds, this yields a desired primary mirror diameter of 0.94 m at 300 nm. We choose a larger diameter of 1.34 m for the GigaBIT primary mirror for multiple reasons. As the resolution increases, the FWHM decreases and we become less susceptible to beam errors. Increasing the resolution also decreases the number of binary objects that get mis-classified as a single asymmetric object which affects the overall shear measurement; this benefit is essentially monotonic and increasing the resolution will always improve the shear measurement accuracy. Lastly, as shown in Fig. 4 of Miyazaki et al. 2007⁷¹, the peak of the galaxy size typically used in weak lensing environments is <0.5 arcseconds in environments with little atmospheric seeing.

Targeting the NUV introduces very strict requirements on the surface coatings for the optics. Many typical coatings have poor throughput in the UV as it is not a commonly used wavelength range. This also requires all optics to be reflective as transmissive optics, such as lenses, will also drastically reduce throughput. For a balloon-borne instrument, this is a difficult design challenge for two reasons: (1) there is very restricted space and mass allowance in the back-end optics and (2) the budget is much smaller than a space telescope and does not allow for high-quality off-axis-parabola (OAP) mirrors to adjust the beam convergence.

Due to the telescope elevation restrictions and the diurnal nature of balloon-based as-



tronomy, it is often true that it takes more than one night to acquire all of the required exposures in the required bands for a target in a single night as described in Sec. 7.4. Balloon-based telescopes operate similar to space-based telescopes where they have a list of targets and each target has a priority so the priority combined with the observable targets available at a given time determines which target is chosen. For each GigaBIT science target, approximately 25 five minute exposures are required to achieve the desired 35 galaxies per arcmin² (pre-cut) at $z > 0.3$.⁸⁰ This sets both the short timescale stability requirements and long timescale stability requirements for the optics. On short timescales, the jitter must be controlled below the telescope resolution to 20 mas for the duration of a single exposure (five minutes).⁸⁰ On longer timescales, the PSF must be repeatable night-to-night to ensure that the data for each target is cohesive.

Based on the SuperBIT 2023 results provided in Sec. 7.4 we see that there is no ‘smoking gun’ that drives the PSF quality variations throughout the flight.

7.5.2 CURRENT HIGH-LEVEL OPTO-MECHANICAL DESIGN

Figure 7.25 shows the current telescope design for GigaBIT which is a Three Mirror Anastigmatic (TMA) reflector. Only reflective optics are used to increase the throughput in the NUV and they are made of Silicon Carbide to reduce mass. The primary and secondary mirrors have an optimized honeycomb backing and a facesheet that is between 2–4 mm thick. In the current design, the fine guidance mirror is 98 mm in diameter and is located 1.06 m from the science camera; note that it is located at the throat of the beam.



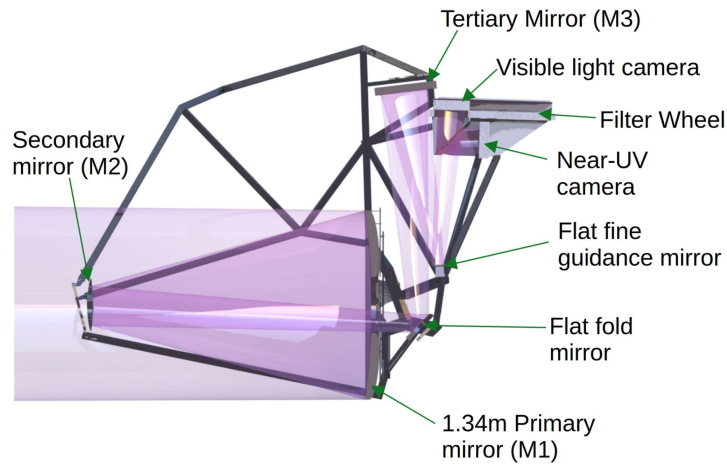


Figure 7.25: GigaBIT telescope design. The optical layout is by Thuy Vy Luu and the opto-mechanical design is by Steven Li.⁸⁰

7.5.3 THERMO-MECHANICAL IMPLICATIONS

Table 7.2 summarizes the relevant properties for the SuperBIT and GigaBIT primary and secondary mirrors. For GigaBIT, the mass for the 2 mm and 4 mm facesheet options are included. Note that the CTE for SiC is two orders of magnitude larger than for CLEARCERAM-Z but so is the conductivity. This is a commonly discussed trade-off between low CTE glasses and SiC. The glass will deform less for a given temperature gradient but the SiC will have less of a temperature gradient because the conductivity is higher. A metric for this trade-off is provided in the bottom row for each mirror and that is the CTE divided by the conductivity. For the primary mirror, it works out to be very similar but for the secondary mirror there is a order of magnitude improvement when moving from fused silica to SiC.



Table 7.2: Comparison of the SuperBIT and GigaBIT mirror thermal properties. Note that a 2 mm and 4 mm facesheet option is provided for GigaBIT.

		SuperBIT	GigaBIT	
			2 mm Facesheet	4 mm Facesheet
Primary	mass [kg]	19.1	19.7	26.7
	cp [J/(kg K)]	773	700	
	k [W/(m K)]	1.52	125	
	cte [m/(m K)]	2×10^{-8}	2.0×10^{-6}	
	cte / k [m/W]	1.3×10^{-8}	1.6×10^{-8}	
Secondary	mass [kg]	1.3	1.9	2.5
	cp [J/(kg K)]	736	700	
	k [W/(m K)]	1.38	125	
	cte [m/(m K)]	5×10^{-7}	2.0×10^{-6}	
	cte / k	3.6×10^{-7}	1.6×10^{-8}	

In Tab. 7.3 I compare the thermal mass of the main mirrors ($C_{tb} = mc_p$) for SuperBIT and the two GigaBIT options. Moving from SuperBIT to GigaBIT results in an increase in the thermal mass for all cases except for the 2 mm facesheet option for the primary mirror. For the secondary mirror, the thermal settling time will increase between 43–88%. This is not a large concern as the thermal settling time for the SuperBIT secondary mirror was never a driving factor. On the other hand, 25% increase in the thermal settling time could be detrimental to the science.

We look to the 2023 SuperBIT flight results to inform design of the GigaBIT instrument. Important results are provided in Sec. 7.4 on the effectiveness of the thermal control of the optics, mechanical support of the primary mirror, and coarse pointing statistics. In many ways, the SuperBIT results are a best case scenario for GigaBIT. In addition, the size of the optics and focal plane is larger for GigaBIT which makes it more difficult to obtain a flat wavefront across the entire focal plane. In Sec. 7.4 we saw that PSF quality in the



Table 7.3: Comparison of the thermal mass of the main SuperBIT and GigaBIT optics. Note that a negative percent decrease indicates an increase in the thermal mass when moving from SuperBIT to GigaBIT.

		Primary	Secondary
		C_{tb} [J/K]	
SuperBIT		14754.9	927.5
GigaBIT	2 mm Facesheet	13773.3	1323.7
	4 mm Facesheet	18720.0	1744.5

		Percent Decrease in C_{tb}	
GigaBIT vs SuperBIT	2 mm Facesheet	7	-43
	4 mm Facesheet	-27	-88
GigaBIT vs GigaBIT	2 mm vs 4 mm	-36	-32

corners of the SuperBIT focal plane is much more sensitive than in the centre of the focal plane; this effect will only be exasperated with GigaBIT. The effectiveness of the thermal control of the SuperBIT optics provides good information on the thermal-induced WFE for GigaBIT. The thermal control for the GigaBIT optics will have to perform better than SuperBIT with the increase in the thermal settling time.

7.5.4 ACTIVE OPTICS IMPLICATIONS

The active optics system for GigaBIT must serve two purposes: (1) correct for the high-speed low-order jitter and (2) correct for the slower time scale high order thermal and gravity induced WFE. This section aims to investigate the requirements for the location and design of the ‘fine guidance mirror’ shown in the back-end optics just before the science cameras in Fig. 7.25. Based on the combination of the results from Sec. 7.4.3 and Sec. 7.5.3, it is highly likely that GigaBIT will require a deformable mirror to correct for higher order WFE. For this reason I investigate deformable mirror options that will allow for higher



order correction and that can also handle the fine guidance requirements. DMs can run very fast (kHz) and take care of tip-tilt control as well as higher order modes though they do introduce a quilting error for lower order modes. It should be noted that the FGM in Fig. 7.25 is located at the throat of the beam to minimize the required diameter of the FGM (98 mm). Due to the limited space in the back-end optics, there is no relay to collimate the beam prior to the DM (as is done in most high-contrast imaging applications).

For this study, I look at the required maximum tip-tilt throw (Θ_{max}) and resolution ($\delta\Theta$) based on the distance between the focal plane and the FGM. The coarse pointing statistics provided in Sec. 7.4.2 inform how much residual pointing error must be corrected by the fine guidance system. The SuperBIT FGM has a factor of three on top of the expected required correction as margin; I adopt the same requirements for GigaBIT which results in a required maximum throw of 0.34 mm and resolution of 3.4 nm on the focal plane. I consider the ALPAO DMs as they have flight heritage with the PICTURE-C⁶⁸ mission and can meet the diameter requirements. The parameters for the two DMs in consideration are shown in Tab. 7.4. Note that ALPAO offers a large stroke upgrade that increases the stroke by a factor of 1.5. The settling time is an important characteristic as that determines how fast we can drive the DM. For both cases the setting time is on the order of a millisecond and will be a limiting factor for GigaBIT.



Mirror	Diameter [mm]	# Actuators across aperture	Stroke [um]	Resolution [um]	Settling Time [ms]
DM3228	93	64	8	0.08	0.5
DMX37	100	7	30	0.23	2

Table 7.4: ALPAO DM parameters assuming a 16-bit controller. For the large stroke upgrade, the stroke and resolution are increased by a factor of 1.5.

In Fig. 7.26 I provide the TT requirements based on the TT placement. In the left panel is the maximum throw and in the right panel is the resolution. The dashed and dotted horizontal lines show the capabilities of the DMs described in Tab. 7.4; any point that lies in a shaded area indicates a z_{TT} value for which the requirements are met.

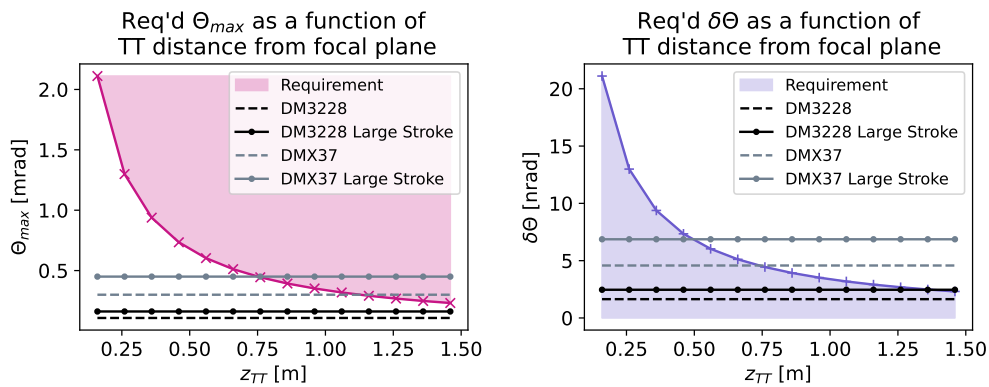
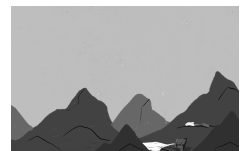


Figure 7.26: Required maximum throw (left panel) and resolution (right panel) for a GigaBIT tip-tilt mirror to have the same focal plane coverage as SuperBIT. As the mirror moves away from the focal plane, the Θ_{max} requirement relaxes while the $\delta\Theta$ requirement becomes tighter. In the left panel, the pink x's show the requirement and the dotted and dashed lines show the capabilities of two ALPAO DMs. ALPAO offers a large stroke option for all DMs that increases the stroke by a factor of 1.5; this is shown by the dotted lines. The two DMs in consideration are the DM3228 and the DMX37. The DM3228 is the desired option due to the large number of actuators. Unfortunately, even with the large stroke option the DM3228 does not meet the maximum throw requirements. The nominal DMX37 model could possibly work and the large stroke option is well within the requirements. Since I assume 16-bit electronics for all ALPAO options, the larger the stroke the lower the resolution. Due to this trade-off, there is no z_{TT} value in which both the maximum throw and resolution requirements are met.



Upon inspection of Fig. 7.26, it does not appear that any z_{TT} as a DM that meets both the maximum throw and resolution requirements. This is highlighted in the left panel of Fig. 7.27. The dashed line is the best possible resolution using 16-bit electronics for a given maximum throw. The purple line and shaded region is the GigaBIT requirement; the z_{TT} parameter is embedded in this plot where the right side corresponds to $z_{TT} = 0.25$ m and the left side corresponds to $z_{TT} = 1.45$ m. As we can see, for an f11 telescope, the dashed line is above the shaded region for all cases. If we increase the f-number of the telescope to f18 (longer focal length), the dashed line is within the required region for all cases.

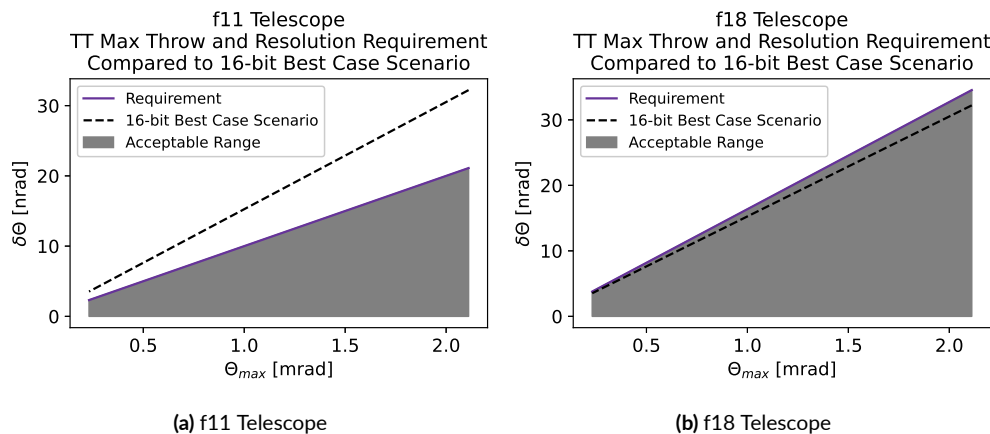
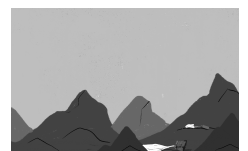


Figure 7.27: Achievable and required maximum throw and resolution of a tip-tilt for GigaBIT for an f11 and f18 telescope. In purple we have the required resolution vs resolution as a function of the distance between the tip-tilt and the focal plane; the right side of both plots corresponds to $z_{tt} = 1.45$ m and the left side corresponds to $z_{tt} = 0.25$ m from Fig. 7.26. Anything below the purple line in the grey shaded region is also acceptable. The dashed line shows the best-case resolution for a given maximum throw if 16-bit electronics are used. We can see that for an f11 telescope (left panel), it is not possible to meet the requirements with 16-bit electronics. If we increase the f-number to be f18, we can meet the requirement for z_{tt} placement of the tip-tilt between 0.25-1.45 m.

This study provides us with three important pieces of information:

1. The SuperBIT FGM cannot be used for GigaBIT
2. There is no commercially available standard deformable mirror that can meet both



the GigaBIT throw and resolution requirements

3. Increasing the f-number of the telescope can help reduce the amount of customization required in the development of an acceptable FGM.

7.6 CONCLUSIONS AND FUTURE WORK

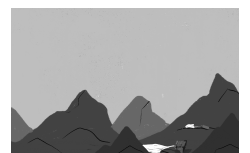
The SuperBIT 2023 flight terminated on May 25, 2023 and the data was not recovered until early June so the results provided in the section are very new. There are many additional paths this research can take and many aspects that need to be investigated prior to GigaBIT construction.

When data-mining the SuperBIT results, I found a number of interesting correlations but all of them appeared to be secondary effects. The data suggests that a host of different effects impact the PSF stability night-to-night, in particular the FSC in charge. The one parameter from the correlation plots that was not further investigated in this thesis was the absolute time. The absolute time of an exposure has an effect for two reasons: (1) the stratospheric conditions present that influence the coarse pointing stability and (2) the sky background. The coarse pointing stability of SuperBIT varies slightly day-to-day due to the varying conditions in the stratosphere as discussed in Sec. 7.4.1. The brighter the sky background, the longer the required FSC exposure time for a given star which slows down the FGS system. There were select nights during the SuperBIT 2023 flight where the instrument was in a region with active Aurora Australis which drastically increased the sky background. It would also be informative to make a correlation plot between the FWHM/ellipticity R matrices and the FSC exposure time for a given exposure. This would indicate FSC requirements for GigaBIT and potentially require us to update our guide star



magnitude requirements. In addition, it would be good to look at the fine pointing quality during each exposure. If the fine pointing is uniformly worse for a given exposure, we expect an increase in the PSF FWHM, if it is worse in one axis we expect to see an increase in the PSF ellipticity. Lastly, it would be beneficial to create a thermo-mechanical model for the 2023 SuperBIT optics and feed it the thermal data from flight to estimate the thermal-induced WFE.

As far as GigaBIT design work, we now have a sufficiently complete opto-mechanical design that we can start investigating thermo-mechanical induced WFE and create a constraint on the required thermal stability of the optics. The results of this study will also inform the requirements for the deformable mirror in the back-end optics. With respect to the DM, there are multiple open questions. The first step is to close the technology gap and create a device that has the throw, resolution, and settling time required by GigaBIT. Next we need to create a control system for the DM. The SuperBIT FGS assumes a linear relationship between the FSC star centroid X and Y location and tip-tilt of the FGM; this model is likely too simplistic for GigaBIT. If the DM is also doing high order corrections, we need a control algorithm that can handle the fact that it is not in a pupil plane and that there is mixing between the modes; this requires heavily depending on modelling and calibration. Due to the lack of space in the back-end of the telescope, this will either have to be a focal plane wavefront control algorithm or there will have to be collimating optics that can be inserted and removed for taking wavefront measurements. Two additional technologies to investigate for the GigaBIT FGM are the custom large-format DMs being developed out of the University of Santa Cruz¹⁶ as well as the curved DMs being developed out of NASA Goddard Space Flight Center.¹³²



8

Summary and Future Work

Conclusions and suggestions for future work are provided at the end of each chapter but will be summarized here for convenience.

Generating a dark zone across a wide wavelength band is an important capability for a high contrast imaging system. In order to minimize hardware requirements, I investigate the potential of using broadband images to estimate the electric field at discrete wavelengths

within the band in Ch. 3. I demonstrate focal plane wavefront estimation and control using broadband images where the mean dark zone contrast is increased by an order of magnitude (from 10^{-5} to 10^{-6}) across a 6% band centered at 640 nm. Using images taken with the 6% filter we estimate the electric field at 10 nm increments from 620–670 nm. These experiments were performed prior to the major HiCAT upgrades in 2022 and thus should be re-executed at higher contrasts now that the broadband laser source has been improved. It would also be beneficial to investigate alternate DM probes; using Sinc probes on DM₂ might be a better option than ordinary probes on DM₁ to induce chromatic effects in the dark zone. Lastly, the chromaticity of the testbed has not been considered in the estimator. Adding spectral weights to the estimated electric fields may also improve the Broadband estimator performance.

Maintaining the dark zone in the presence of quasi-static wavefront error drifts is crucial for directly imaging exoplanets. I provide proof-of-concept of a Dark Zone Maintenance (DZM) algorithm based on an extended Kalman filter (EKF) that can maintain the dark zone contrast in a variety of drift scenarios. In addition to the nominal DZM algorithm, I show that it works in the low SNR regime by taking high SNR images and degrading them prior to passing the measurement to the estimator. I also show that the nominal DZM algorithm works when planets are present at or below the mean DZ contrast. In the case of bright planets, I show that we can augment the EKF to estimate the incoherent planet light separately from the coherent electric field. I use a publicly available post processing package, `pyklip`, to show that planets can be successfully detected in ADI or RDI mode from DZM data. I also present alternate post processing methods (CDI, IAI) for the augmented DZM algorithm that use the coherent and incoherent estimates to extract the planet from

the dark zone.

Future work for the DZM algorithm involve making the experiments as similar to a space mission as possible. For the most part, monochromatic experiments have been performed. This is fine for the first stage of development but a push for more broadband experiments should be made. Next we should investigate alternate drifts to inject. JWST has WFE time series data available and correcting for that drift in the lab would be an excellent exercise. It is also import to use EMCCD detector noise models for the low SNR experiments since RST will use an EMCCD detector. The `emccd_detect`⁷⁶ and `PhotonCount`⁷⁷ packages based on Nemati 2020⁷⁵ will be useful tools for this work. With the EKF, investigation into the Q matrix would be very worthwhile. We will never know the exact drift present in a system so we need to determine the maximum error in the Q matrix before we see a degradation in performance. For the AEKF, it would be a useful exercise to look into dynamic updating of the Q matrix for the incoherent term to try to avoid having the incoherent estimate capture features of the residual starlight speckles. One could also run AEKF with planet-PSF-sized blocks instead of single pixels. The H matrix (Eqn. 6.11) would then be $3n_p \times 3n_p$ where n_p is the number of pixels in an expected planet PSF. As the incoherent estimate finds the planet, the blocking of the dark zone could be dynamically updated to centre the block on the planet.

Post-processing with DZM data is still in the early stages of development and there are a number of future research avenues. When using the nominal DZM algorithm with ADI, it would be worthwhile to investigate the shortest allowable dither-repeat cycle that is longer than the memory of the EKF. Repeating the dither commands could also potentially aid with the RDI approach for a similar reason so experiments should be run with the AEKF

and dither-cycling. Since ADI is a commonly used approach, it would be good to look into using the AEKF with ADI and develop a methodology for re-initializing the incoherent estimate when a roll is performed. Lastly, the CDI and IAI post processing methods are still in the early stages. Further investigation past the simple approach implemented in this thesis could increase the post processing methods available that do not require rolling the telescope.

The SuperBIT 2023 flight terminated on May 25, 2023 and the data was not recovered until early June so the results provided in the section are very new. There are many additional paths this research can take and many aspects that need to be investigated prior to GigaBIT construction. It would be informative to make a correlation plot between the FWHM/ellipticity R matrices and the FSC exposure time for a given exposure. This would indicate FSC requirements for GigaBIT and potentially require us to update our guide star magnitude requirements. In addition, it would be good to look at the fine pointing quality during each exposure. If the fine pointing is uniformly worse in for a given exposure, we expect an increase in the PSF FWHM, if it is worse in one axis we expect to see an increase in the PSF ellipticity. Lastly, it would be beneficial to create a thermo-mechanical model for the 2023 SuperBIT optics and feed it the thermal data from flight to estimate the thermal-induced WFE.

As far as GigaBIT design work, we now have a sufficiently complete opto-mechanical design that we can start investigating thermo-mechanical induced WFE and create a constraint on the required thermal stability of the optics. The results of this study will also inform the requirements for the deformable mirror in the back-end optics. We need to close the DM technology gap and create a device that has the throw, resolution, and set-

ting time required by GigaBIT. For the high-order DM control system, the algorithm will need to be designed to handle the fact that the DM is not in a pupil plane. It may be useful to investigate adding collimating optics that can be inserted and removed for taking pupil-plane wavefront measurements with the science camera rather than developing a focal plane wavefront estimation and control algorithm.



HiCAT Hardware and Ambient Environment

In this appendix we discuss the HiCAT hardware state and atmospheric environment. The HiCAT design is described in N'Diaye et al. 2013⁷⁴ and the current layout can be found in Soummer et al. 2018.¹²³ Due to frequent changes to the hardware configuration, we

provide a description of why and when certain components are installed/removed. Similarly, there are many aspects of the experiments shown in Sec. 3.3 that are unique to the HiCAT testbed. The iteration time and atmosphere control are determined by external factors that are not representative of another high contrast testbed, ground telescope, or space telescope. The effect of these topics are discussed to aid in the interpretation of our results and what our results mean for future high-contrast missions.

A.1 HIGH CONTRAST IMAGER FOR COMPLEX APERTURE TELESCOPES (HiCAT)

The layout of the High Contrast Imager for Complex Aperture Telescopes (HiCAT) is shown below in Fig. A.1.

HiCAT currently contains a classical Lyot coronagraph,⁶¹ two Boston Micromachine kilo (BMC) deformable mirrors, and an IrisAO PTT I I L segmented primary surrogate. For the experiments presented here, the IrisAO segmentation pattern is not apodized,¹²⁸ and instead compensated for using DM based wavefront control. The light source is a NKT Photonics SuperK EVO4 with a tunable filter; the tunable filter allows for the central wavelength and bandpass to be set. The simulator used to emulate the testbed (`catkit`⁸¹) is discussed in Fowler et. al. 2020²⁸ and Moriarty et. al. 2018.⁷²

HiCAT can achieve contrast levels in a $5.8\text{--}9.8 \lambda/D_{\text{Lyot}}$ dark zone of 2.5×10^{-8} (6.5×10^{-8} 90% of the time) using the monochromatic laser source and 6.3×10^{-7} over a 3% band centred at 650 nm using the broadband laser source.⁹⁸

HiCAT optical components and paths

Note - Not to scale, and does not depict physical arrangement of components on the table.
 For clarity, many optics are depicted as transmissive that are actually reflective.
 Blue and green: optics Red: detectors. Purple: mechanisms Pink: manual adjustment axes

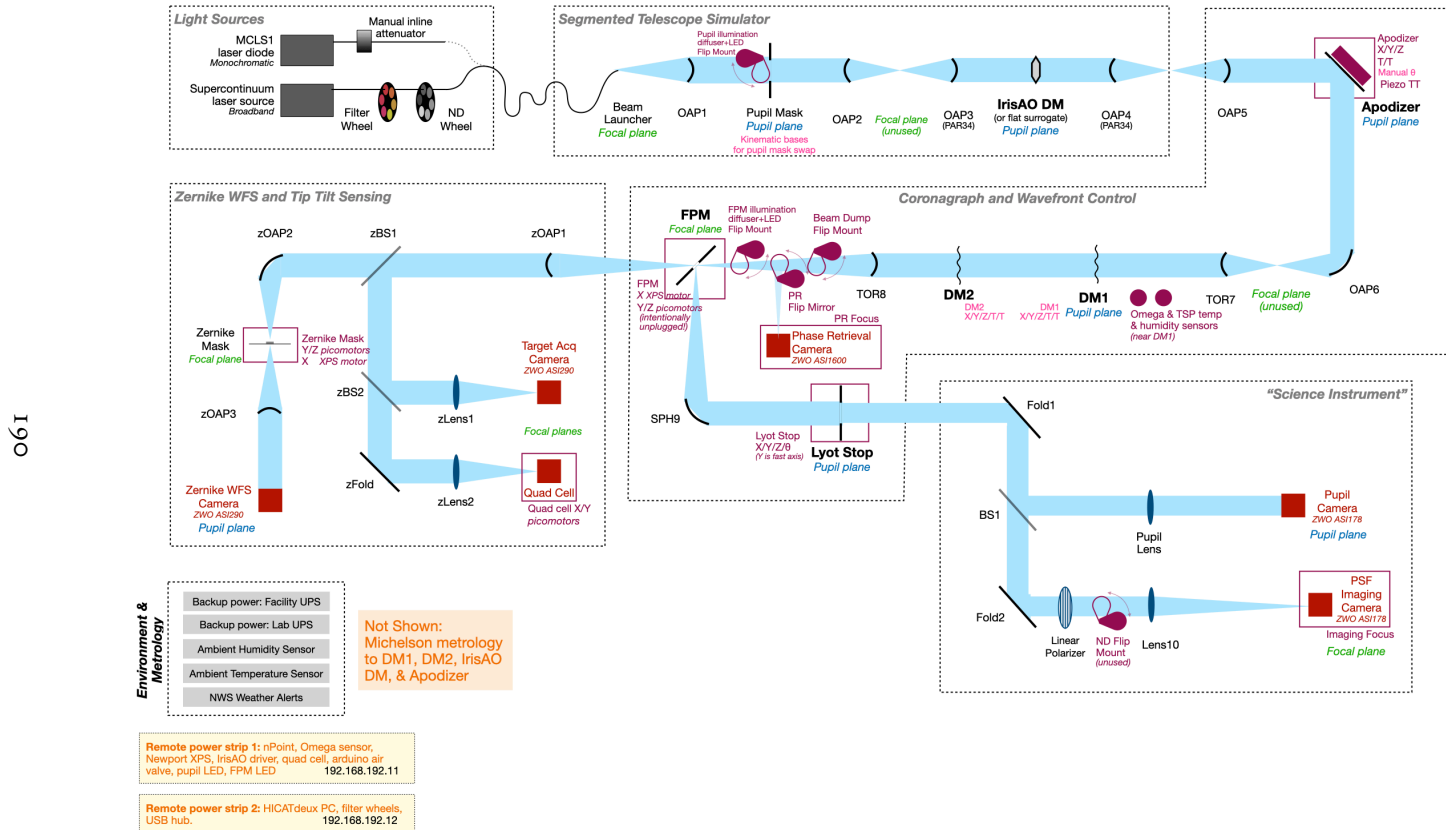


Figure A.1: Testbed layout for HiCAT.¹²⁷ The light leaves the laser source, is collimated, and then arrives at the IrisAO segmented deformable mirror. The light then hits the apodizer (if it is installed or a flat mirror if it is not) followed by each of the BMC DMs, one which is in the pupil plane and one which is out of the pupil plane. HiCAT uses a reflective focal plane mask (FPM) which, after focusing the light on the FPM, allows the main lobe of the PSF to pass through the middle and reflects the planet-searching region. The main lobe of the PSF travels to either the low order wavefront sensor or calibration detectors. The light reflected off of the FPM is collimated, passed through the Lyot stop in the pupil plane, and then focused onto the science camera.

A.2 HiCAT HARDWARE STATE

A.2.1 CH. 6 TESTBED STATE

In the latter half of 2022, HiCAT went through major software and hardware upgrades. The entire foundation of the software was rewritten to enable the parallelization of different wavefront control loops. An overview of these changes is provided in Soummer et al. 2022.¹²⁷ This upgrade drastically reduced the iteration time from ~ 15 s down to <1 s. In addition to the software upgrades, HiCAT also acquired a new laser source and science camera. The laser source is a NKT Photonics SuperK EVO₄ with a tunable filter. This means that the same laser source can be used for broadband and monochromatic experiments as both the central wavelength and the bandpass of the tunable filter can be set. The new science camera is the ZWO ASI₅₃₃MM which has lower read noise and a much higher dynamic range. The pixels are also slightly larger than the old science camera and the un-binned sampling is 8.699 pixels per λ/D_{Lyot} . For Ch. 6, the apodizer was installed (see Soummer et al. 2022¹²⁷ for figures). The goal of the apodizer is to reduce the diffraction effects of the non-circular segmented aperture, as well as the gaps between the segments, on the dark zone.

A.2.2 CH. 3–5 TESTBED STATE

For Ch. 3–5, the monochromatic laser source was a Thorlabs MCLS₁ diode laser at 638 nm and the broadband source was a Leukos SM-30-400 laser combined with a filter wheel. The filter wheel had five 10 nm bandpass filters available at 610, 620, 640, 660, and 670 nm as well as a 6% bandpass filter centered at 640 nm. The science detector for these chapters was

the ZWO ASI178MM and the un-binned sampling was 13.623 pixels per λ/D_{Lyot} . The apodizer was also not installed for any of the experiments shown in Ch. 3–5.

PINHOLE

In July 2021, a $3\ \mu\text{m}$ pinhole was temporarily placed after the laser launch to filter some of the fiber effects seen when using the broadband laser source. This reduced the throughput by a factor of five, thus requiring longer exposure times. It was removed shortly thereafter due to the lengthy experiment times. In this thesis, the pinhole was in place for Sec. 4.4.4 results but removed for all other experiments.

IRISAO SEGMENTED APERTURE

In December 2020, the IrisAO segmented aperture PTT I I I L was installed on HiCAT. The aperture has 37 hexagonal segments with piston, tip, and tilt degrees of freedom. Each segment is 1.212 mm from flat to flat of the hexagon. This thesis includes results with and without the IrisAO installed. The Results for Section 4.4.2 and Appendix B do not include the segmented aperture while results for Sec. 4.4.3–4.4.5 and Ch. 5–6 contain the segmented aperture.

A.2.3 DRIFT METRICS

On the HiCAT testbed, prior to 2022 (Chapters 3–5), the majority of the iteration time is due to file writing and thus the drift per iteration is a better metric when extrapolating to other experiments than the drift per minute. Both values will be provided for each experiment as the HiCAT environment is changing in real time. In each iteration there are

two guaranteed images taken, one image for the estimator and one image after the EFC correction, as well as two optional images, an open-loop image and a direct image without the FPM. The image after the EFC correction is not strictly necessary and could be removed to shorten iteration time. In addition to the drift variance, the final DM surface root mean square (RMS) and peak-to-valley (PV) are provided for the drifting DM.

A.2.4 HiCAT ENVIRONMENTAL DRIFT

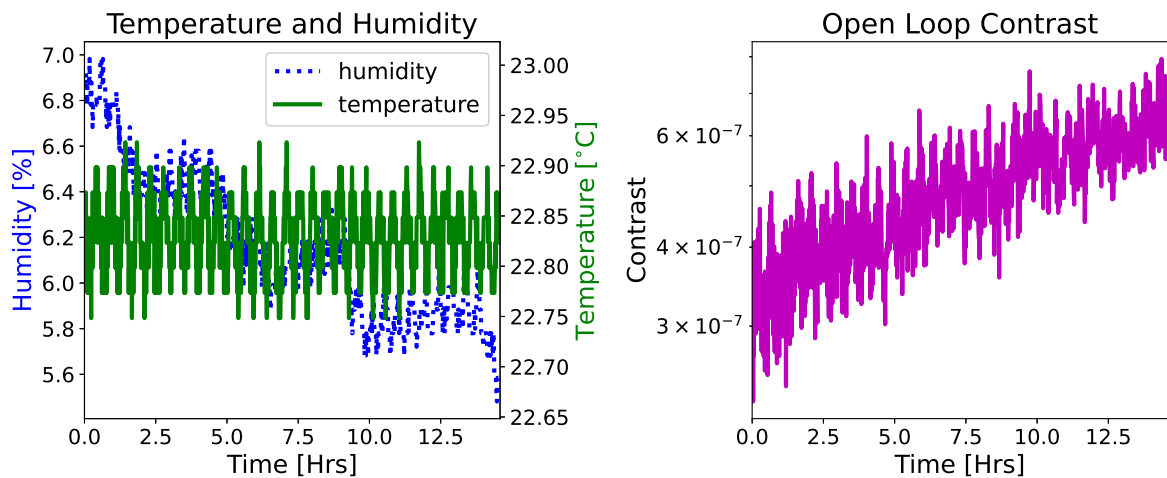


Figure A.2: HiCAT instabilities due to climate control issues in the summer 2020. A constant DM command is applied repeatedly and the thermal/humidity drifts cause a contrast degradation of a factor of two over the course of 14 hrs as shown by the magenta curve in the right panel. The main driver of this contrast degradation is the humidity drop from 6.85% to 5.55% as shown by the blue dotted curve in the left panel. The temperature is shown by the green curve in the left panel; it stays relatively constant oscillating by $\pm 0.09^{\circ}\text{C}$ about the mean of 22.83°C .

In the HiCAT lab, the main drivers of quasi-static WFE are temperature and humidity. This is captured by Fig. A.2 which shows the mean contrast on HiCAT degrading by a factor of two (magenta curve in the right panel) over the course of 14 hours while repeatedly applying a constant DM command. This factor of four is attributed to the humidity drop

from 6.85% to 5.55% shown by the blue dotted curve in the left panel. Note that there is a small sinusoidal oscillation with an amplitude of 0.1% on top of the dominant humidity drift as shown by the blue dotted curve in the left panel. This sinusoidal humidity drift correlates with high-order contrast variations seen in the magenta curve in the right panel which have an amplitude of 8×10^{-8} . The temperature is also oscillating by $\pm 0.09^\circ\text{C}$ (green curve, left panel), a factor of four worse than usual, and is likely contributing to the high order contrast oscillations. The mean of the temperature variations is relatively constant so it is not directly contributing to the bulk contrast change. We acquired these data during the summer of 2020 when there were issues with the temperature and humidity control in the lab, as shown by the blue humidity curve. As a result, the stability and performance shown in Fig. A.2 are an order of magnitude worse than what is possible in nominal HiCAT operations.

We potentially see this effect to a smaller degree in the results from Sec. 4.4.4. As shown in Fig. A.3 the mean DZ contrast (magenta) seems to track the humidity (blue dots) when it starts to increase at 4.5 hrs. The humidity drift is not captured by the process noise covariance matrix; thus, the EKF will not try to correct for it. The contrast-humidity correlation could be a coincidence, as the humidity change is very small, but the correlation of the high-order oscillations in Fig. A.2 is an indicator that humidity variations of 0.1% affect the contrast at the 10^{-8} level. Further investigation is warranted and will be performed as we have limited data on the contrast-humidity relationship on HiCAT.

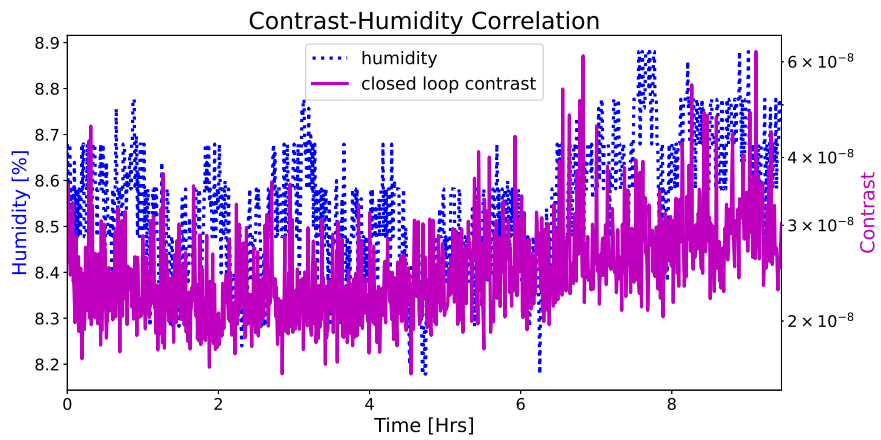


Figure A.3: Contrast-humidity correlation. The blue dotted line shows the humidity in the HiCAT enclosure during the experiment described in Sec. 4.4.4. The mean dark zone contrast shown in magenta starts trending up when the humidity increases at the 4.5 hr mark.

B

Choosing an Optimal Dither for BMC

Random Walk Drift

The dither magnitude has a large effect on the success of the DZM experiment. Since our initial injected drift (random walk of BMC DM actuators) is relatively simple, I perform an in-depth study of the correlation between the drift standard deviation and optimal dither

value. The optimal dither is considered to be the dither that minimizes the estimate convergence time, steady-state estimate error, and steady-state contrast. I build on the simulated results in Pogorelyuk et al. 2019⁸⁸ by performing HiCAT simulations⁸¹ and hardware experiments.

B.0.1 EFFECT OF DITHER ON ESTIMATOR CONVERGENCE (SIMULATIONS)

As described in Sec. 4.3, the dither increases the phase diversity of the electric field at the focal plane. One can think of the EKF as a pair-wise probe estimator with memory that acquires one probed image per iteration. The larger the dither magnitude, the larger the change between images and the better the estimator will perform. Note that we cannot increase the dither indefinitely as eventually the contrast will be limited by the dither and, even though the estimator will converge quickly, it is not the overall behaviour that is desired. We cannot assume we have access to the open-loop image; thus, the closed-loop estimate error must be used as a proxy as it is in the EKF. The closed-loop intensity estimate error for pixel i is calculated as

$$\varepsilon_{CLi} = |z_i^k - \hat{y}_i^k| = |z_i^k - b(\hat{x}^k, u^k)_i| \quad (\text{B.1})$$

where z is the measured closed-loop intensity, \hat{y}^k is the estimate of the closed-loop intensity as defined in Section 4.3, and k is the current iteration. The mean and standard deviation

of the closed-loop intensity estimate error for a given iteration are then

$$\mu_{err} = \frac{1}{n} \sum_{i=1}^n |z_i^k - \hat{y}_i^k| \quad (\text{B.2})$$

$$\sigma_{err} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left| |z_i^k - \hat{y}_i^k| - \mu_{err} \right|^2} \quad (\text{B.3})$$

where i is a given pixel and n is the number of pixels in the dark zone.

Figure B.1 shows the mean intensity estimate error in units of contrast for a simulated DZM run on HiCAT with a random walk DM drift of 0.01 nm/iter. As shown in Fig. B.1, the mean estimate error initially gets worse and then turns over as the estimator accumulates enough information to improve the estimate. It improves for a number of iterations and then asymptotes to a steady state mean estimate error caused by a discrepancies between the model and the actual system combined with the effectiveness of the dither. Estimator convergence is defined as the iteration where the absolute estimate error reaches the steady state absolute estimate error; for the orange solid curve in Fig. B.1 this would be iteration 98. Since we are looking at the absolute error and not the relative error, the estimate error will increase as the contrast degrades. Note that for these simulations and experiments, the control gain (β) is one for the entire experiment. This accentuates the effect of the poor initial estimate and allows us to identify the fastest converging dither.

The mean estimate error for the smallest dither in Fig. B.1 (red dot curve) takes longer to peak and longer to reach the same error level as the $\sigma_{dither} = 0.1$ nm case. The largest dither case (teal dot dashed curve) ‘converges’ the fastest but the steady state mean estimation error is very high and does not to improve. The EKF is performing well but the $\sigma_{dither} = 0.4$ nm case is at the dither limit, meaning the contrast remains constant at

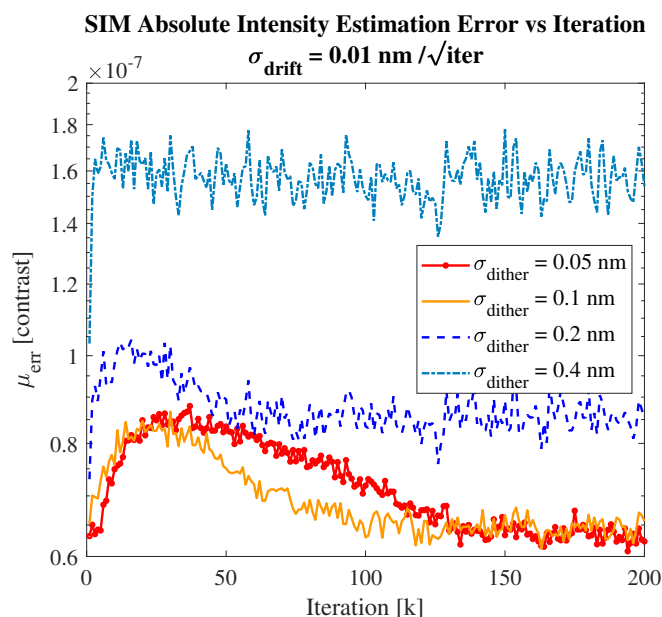


Figure B.1: Mean absolute error between the estimated closed-loop intensity and the measured closed-loop intensity at each pixel in the dark zone for 200 iterations during an active control sequence for various dither values, σ_{dither} [nm]. For all cases, $\sigma_{drift} = 0.01 \text{ nm}/\sqrt{\text{iter}}$.

the value determined in Section 4.4.1. Since this is a higher, constant contrast, the absolute error will scale appropriately and follow a similar trend. This dither limiting case is also shown in Fig. B.2 and will be discussed more in Section B.o.2. The two medium dithers (blue dashed and orange solid curves) have similar profiles but the smaller of the two asymptotes to a lower error. From this we can see that there is an ‘optimal’ dither if the goal is to minimize the mean estimate error.

B.o.2 CHOOSING AN OPTIMAL DITHER

Figure B.2 shows the mean contrast evolution for both simulation and hardware data for different dithers and a $\sigma_{drift} = 0.01 \text{ nm}/\sqrt{\text{iter}}$ (note one iteration is $< 10 \text{ s}$). The two plots are well matched for $\sigma_{dither} = 0.2, 0.4 \text{ nm}$. The lab result (right panel) for the solid orange

curve, $\sigma_{dither} = 0.1$ nm, initially behaves like the simulation result (left panel) for the red dot curve, $\sigma_{dither} = 0.05$ nm. These experiments are currently performed with no real-time tip-tilt correction so this is likely due to small lab instabilities that hinder the effect of the small dither making it harder for the estimator to converge. Once the estimate for the lab case peaks, around iteration 30, the contrast for $\sigma_{dither} = 0.1$ nm returns to the original contrast with a slope similar to the simulated $\sigma_{dither} = 0.1$ nm. In the future, when the real-time tip-tilt correction is incorporated into the DZM algorithm, the smaller dithers are expected to better match the simulations. Figure B.2 illustrates the combination of the described effects

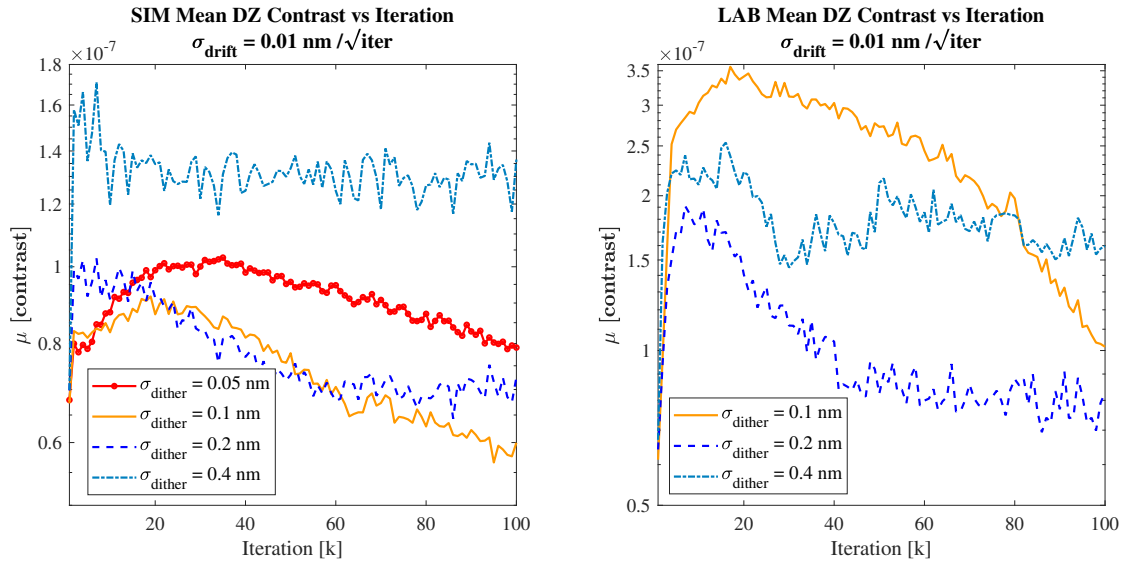


Figure B.2: HiCAT simulation and hardware comparison of the mean dark zone contrast response for $\sigma_{drift} = 0.01$ nm/ $\sqrt{\text{iter}}$ and varying dither. The contrast response as a function of dither follows a similar trend to the estimate error with the smallest and largest dithers performing worse than the dithers in between. The simulations match the hardware for $\sigma_{dither} = 0.2, 0.4$ nm but deviate at $\sigma_{dither} = 0.1$ nm.

in Fig. 4.5 and B.1. The contrast initially degrades as the estimate error is large and thus the optimal command provided by EFC is not cancelling out the electric field very well. As the estimate improves, so does the quality of the EFC command and the contrast trends

back to the original state. The steady state contrast is then determined by a combination of the dither contrast limit as shown in Fig. 4.5 and the estimate error shown in Fig. B.1. For larger dithers, the dither contrast limit dominates and for smaller dithers, the estimate error dominates. It is important to note that the red dot curve in Fig. B.2 does return to the original contrast after ~ 200 iterations. The teal dot dashed curve ($\sigma_{dither} = 0.4$ nm) stays relatively constant in both the simulation and on the hardware. The contrast is at the ‘dither limit’ as shown in Fig. 4.5 and the EFC cannot do much to improve it.

To choose an optimal dither there are three main parameters of interest: number of iterations required for the estimator to converge, steady-state mean estimate error, and steady-state contrast. The goal is to pick the dither that achieves an optimal point where all three are small. For a drift of 0.01 nm/iter the effect of varying dither values is shown in Fig. B.3. Note that the dashed red line representing the asymptotic contrast vs dither agrees with Fig. 4.5 for the larger dither values. It depends on how each of the parameters in Fig. B.3 are weighted, but the optimal dither lies in the yellow region highlighted where $0.1 \text{ nm} \leq \sigma_{dith} \leq 0.2 \text{ nm}$.

The behaviour of the DZM algorithm varies with a number of factors including starting contrast, observation noise, and drift. All plots shown in Sections 4.4.1 and B.o.2 vary the dither while keeping all other parameters constant. Repeating the process in Sections 4.4.1 and B.o.2 for a variety of σ_{drift} values shows that for a BMC DM actuator random walk drift, the optimal dither is $\sigma_{dither} \sim 15 \times \sigma_{drift}$. Note that we plan on repeating this process for drifts injected via the IrisAO DM during a more cohesive DZM parameter study.

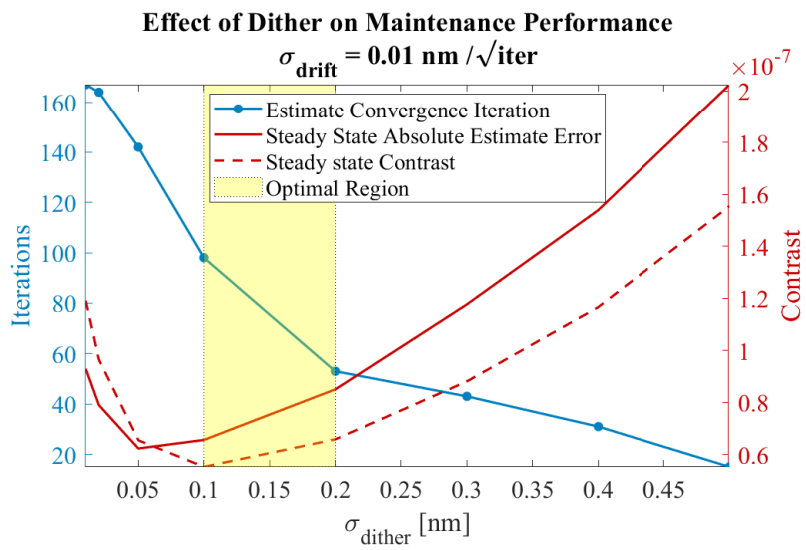



Figure B.3: Summary plot of Fig. B.1 and B.2. Simulated effect of dither on iterations until estimator converges (blue dot), steady state closed-loop contrast estimate error (solid red), and steady state contrast (dashed red). The yellow block shows the region in which the combination of these three parameters is optimal.

Bibliography

- [1] 5. How PSFEx works — PSFEx 3.18.2 documentation.
- [2] Low Expansion Glass - Ceramics.
- [3] Roman Space Telescope Spacecraft and Instrument Parameters.
- [4] Timeline of the Universe Image.
- [5] (2022a). Aperture Position Angle Special Requirements - JWST User Documentation.
- [6] (2022b). JWST Slew Times and Overheads - JWST User Documentation.
- [7] Akeson, R., Armus, L., Bachelet, E., Bailey, V., Bartusek, L., Bellini, A., Benford, D., Bennett, D., Bhattacharya, A., Bohlin, R., Boyer, M., Bozza, V., Bryden, G., Calchi Novati, S., Carpenter, K., Casertano, S., Choi, A., Content, D., Dayal, P., Dressler, A., Doré, O., Fall, S. M., Fan, X., Fang, X., Filippenko, A., Finkelstein, S., Foley, R., Furlanetto, S., Kalirai, J., Gaudi, B. S., Gilbert, K., Girard, J., Grady, K., Greene, J., Guhathakurta, P., Heinrich, C., Hemmati, S., Hendel, D., Henderson, C., Henning, T., Hirata, C., Ho, S., Huff, E., Hutter, A., Jansen, R., Jha, S., Johnson, S., Jones, D., Kasdin, J., Kelly, P., Kirshner, R., Koekemoer, A., Kruk, J., Lewis, N., Macintosh, B., Madau, P., Malhotra, S., Mandel, K., Massara, E., Masters, D., McEnery, J., McQuinn, K., Melchior, P., Melton, M., Mennesson, B., Peeples, M., Penny, M., Perlmutter, S., Pisani, A., Plazas, A., Poleski, R., Postman, M., Ranc, C., Rauscher, B., Rest, A., Roberge, A., Robertson, B., Rodney, S., Rhoads, J., Rhodes, J., Ryan, Jr., R., Sahu, K., Sand, D., Scolnic, D., Seth, A., Shvartzvald, Y., Siellez, K., Smith, A., Spergel, D., Stassun, K., Street, R., Strolger, L.-G., Szalay, A., Trauger, J., Troxel, M. A., Turnbull, M., van der Marel, R., von der Linden, A., Wang, Y., Weinberg, D., Williams, B., Windhorst, R., Wollack, E., Wu, H.-Y., Yee, J., & Zimmerman, N. (2019). *The Wide Field Infrared Survey Telescope: 100 Hubbles for the 2020s*. Technical report. Publication Title: arXiv e-prints ADS Bibcode: 2019arXiv190205569A Type: article.

- [8] Bailey, V. P., Bottom, M., Cady, E., Cantalloube, F., Boer, J. d., Groff, T., Krist, J., Millar-Blanchaer, M. A., Vigan, A., Chilcote, J., Choquet, E., Rosa, R. J. D., Girard, J. H., Guyon, O., Kern, B., Lagrange, A.-M., Macintosh, B., Males, J. R., Marois, C., Meshkat, T., Milli, J., N'Diaye, M., Ngo, H., Nielsen, E. L., Rhodes, J., Ruane, G., Holstein, R. G. v., Wang, J. J., & Xuan, W. (2018). Lessons for WFIRST CGI from ground-based high-contrast systems. In *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*, volume 10698 (pp. 1913–1925): SPIE.
- [9] Baxter, W., Potier, A., Ruane, G., & Prada, C. M. (2021). Design and commissioning of an in-air coronagraph testbed in the HCIT facility at NASA's Jet Propulsion Laboratory. In *Techniques and Instrumentation for Detection of Exoplanets X*, volume 11823 (pp. 574–585): SPIE.
- [10] Bertin, E. (2011). Automated Morphometry with SExtractor and PSFEx. 442, 435. Conference Name: Astronomical Data Analysis Software and Systems XX ADS Bibcode: 2011ASPC..442..435B.
- [11] Bertin, E. & Arnouts, S. (1996). SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series*, 117(2), 393–404. Number: 2 Publisher: EDP Sciences.
- [12] Bolcar, M. R., Crooke, J., Hylan, J. E., Bronke, G., Collins, C., Corsetti, J., Genetie, J., Gong, Q., Groff, T., Hayden, W., Jones, A., Matonak, B., Park, S., Sacks, L., West, G., Yang, K., & Zimmerman, N. (2018). The large UV/optical/infrared surveyor (LUVOIR): decadal mission study update. In *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*, volume 10698 (pp. 232–244): SPIE.
- [13] Bonse, M. J., Garvin, E. O., Gebhard, T. D., Dannert, F. A., Cantalloube, F., Cugno, G., Absil, O., Hayoz, J., Milli, J., Kasper, M., & Quanz, S. P. (2023). Comparing Apples with Apples: Robust Detection Limits for Exoplanet High-contrast Imaging in the Presence of Non-Gaussian Noise. *The Astronomical Journal*, 166(2), 71. Publisher: The American Astronomical Society.
- [14] Bordé, P. J. & Traub, W. A. (2006). High-Contrast Imaging from Space: Speckle Nulling in a Low-Aberration Regime. *The Astrophysical Journal*, 638(1), 488. Publisher: IOP Publishing.
- [15] Bottom, M., Wallace, J. K., Bartos, R. D., Shelton, J. C., & Serabyn, E. (2017). Speckle suppression and companion detection using coherent differential imaging. *Monthly Notices of the Royal Astronomical Society*, 464(3), 2937–2951.

- [16] Bowens-Rubin, R., Dillon, D., Hinz, P. M., & Kuiper, S. (2021). Performance of Large-Format Deformable Mirrors Constructed with Hybrid Variable Reluctance Actuators II: Initial Lab Results from FLASH. In *Techniques and Instrumentation for Detection of Exoplanets X* (pp.69). arXiv:2110.01693 [astro-ph].
- [17] Cady, E. J., Prada, C. A. M., An, X., Balasubramanian, K., Diaz, R. T., Kasdin, N. J., Kern, B. D., Kuhnert, A. C., Nemati, B., Poberezhskiy, I., Riggs, A. J. E., Zimmer, R. P., & Zimmerman, N. T. (2015). Demonstration of high contrast with an obscured aperture with the WFIRST-AFTA shaped pupil coronagraph. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(1), 011004. Publisher: SPIE.
- [18] Cantalloube, F., Mouillet, D., Mugnier, L. M., Milli, J., Absil, O., Gonzalez, C. A. G., Chauvin, G., Beuzit, J.-L., & Cornia, A. (2015). Direct exoplanet detection and characterization using the ANDROMEDA method: Performance on VLT/NaCo data. *Astronomy & Astrophysics*, 582, A89. Publisher: EDP Sciences.
- [19] Carter, A. L., Hinkley, S., Kammerer, J., Skemer, A., Biller, B. A., Leisenring, J. M., Millar-Blanchaer, M. A., Petrus, S., Stone, J. M., Ward-Duong, K., Wang, J. J., Girard, J. H., Hines, D. C., Perrin, M. D., Pueyo, L., Balmer, W. O., Bonavita, M., Bonnefoy, M., Chauvin, G., Choquet, E., Christiaens, V., Danielski, C., Kennedy, G. M., Matthews, E. C., Miles, B. E., Patapis, P., Ray, S., Rickman, E., Sallum, S., Stapelfeldt, K. R., Whiteford, N., Zhou, Y., Absil, O., Boccaletti, A., Booth, M., Bowler, B. P., Chen, C. H., Currie, T., Fortney, J. J., Grady, C. A., Greebaum, A. Z., Henning, T., Hoch, K. K. W., Janson, M., Kalas, P., Kenworthy, M. A., Kervella, P., Kraus, A. L., Lagage, P.-O., Liu, M. C., Macintosh, B., Marino, S., Marley, M. S., Marois, C., Matthews, B. C., Mawet, D., McElwain, M. W., Metchev, S., Meyer, M. R., Molliere, P., Moran, S. E., Morley, C. V., Mukherjee, S., Pantin, E., Quirrenbach, A., Rebollido, I., Ren, B. B., Schneider, G., Vasist, M., Worthen, K., Wyatt, M. C., Briesemeister, Z. W., Bryan, M. L., Calissendorff, P., Cantalloube, F., Cugno, G., Furio, M. D., Dupuy, T. J., Factor, S. M., Faherty, J. K., Fitzgerald, M. P., Franson, K., Gonzales, E. C., Hood, C. E., Howe, A. R., Kuzuhara, M., Lagrange, A.-M., Lawson, K., Lazzoni, C., Lew, B. W. P., Liu, P., Llop-Sayson, J., Lloyd, J. P., Martinez, R. A., Mazoyer, J., Palma-Bifani, P., Quanz, S. P., Redai, J. A., Samland, M., Schlieder, J. E., Tamura, M., Tan, X., Uyama, T., Vigan, A., Vos, J. M., Wagner, K., Wolff, S. G., Ygouf, M., Zhang, X., Zhang, K., & Zhang, Z. (2023). The JWST Early Release Science Program for Direct Observations of Exoplanetary Systems I: High-contrast Imaging of the Exoplanet HIP 65426 b from 2 to 16 μm . *The Astrophysical Journal Letters*, 951(1), L20. Publisher: The American Astronomical Society.

- [20] Chonis, T. S., Gallagher, B. B., Knight, J. S., Acton, D. S., Smith, K. Z., Wolf, E., Coppock, E., Tersigni, J., & Comeau, T. (2018). Characterization and calibration of the James Webb space telescope mirror actuators fine stage motion. In *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*, volume 10698 (pp. 106983S): International Society for Optics and Photonics.
- [21] Choquet, , Perrin, M. D., Chen, C. H., Soummer, R., Pueyo, L., Hagan, J. B., Gofas-Salas, E., Rajan, A., Golimowski, D. A., Hines, D. C., Schneider, G., Mazoyer, J., Augereau, J.-C., Debes, J., Stark, C. C., Wolff, S., N'Diaye, M., & Hsiao, K. (2016). FIRST IMAGES OF DEBRIS DISKS AROUND TWA 7, TWA 25, HD 35650, AND HD 377. *The Astrophysical Journal Letters*, 817(1), L2. Publisher: The American Astronomical Society.
- [22] Coyle, L. E., Knight, J. S., Pueyo, L., Arenberg, J., Bluth, M., East, M., Patton, K., & Bolcar, M. R. (2019). Large ultra-stable telescope system study. In *UV/Optical/IR Space Telescopes and Instruments: Innovative Technologies and Concepts IX*, volume 11115 (pp. 111150R): International Society for Optics and Photonics.
- [23] Danielski, C., Baudino, J.-L., Lagage, P.-O., Boccaletti, A., Gastaud, R., Coulais, A., & Bézard, B. (2018). Atmospheric Characterization of Directly Imaged Exoplanets with JWST/MIRI. *The Astronomical Journal*, 156(6), 276. Publisher: American Astronomical Society.
- [24] Debes, J. H., Ren, B., & Schneider, G. (2019). Pushing the limits of the coronagraphic occulters on Hubble Space Telescope/Space Telescope Imaging Spectrograph. *Journal of Astronomical Telescopes, Instruments, and Systems*, 5, 035003.
- [25] Doble, N., Miller, D. T., Yoon, G., & Williams, D. R. (2007). Requirements for discrete actuator and segmented wavefront correctors for aberration compensation in two large populations of human eyes. *Applied Optics*, 46(20), 4501–4514. Publisher: Optica Publishing Group.
- [26] Eggers, J. (2016). NASA Completes Balloon Technology Test Flight.
- [27] ESA (2018). Hubble's Instruments: WFPC1 — Wide Field and Planetary Camera 1.
- [28] Fowler, J., Noss, J., Laginja, I., Soummer, R., & Perrin, M. (2020). The Generalized Lab Architecture for Restructured optical Experiments (GLARE). 235, 373.09. Conference Name: American Astronomical Society Meeting Abstracts #235.

- [29] Fraquelli, D. A., Schultz, A. B., Bushouse, H., Hart, H. M., & Vener, P. (2003). NICMOS Coronagraphy. *Publications of the Astronomical Society of the Pacific*, 116(815), 55. Publisher: The University of Chicago Press.
- [30] Galicher, R., Baudoz, P., Rousset, G., Totems, J., & Mas, M. (2010). Self-coherent camera as a focal plane wavefront sensor: simulations. *Astronomy and Astrophysics*, 509, A31.
- [31] Galloway, M. N., Benton, S. J., Clark, P., Damaren, C. J., Eifler, T., Hartley, J. W., Jones, W. C., Li, L., Luu, T. V. T., Massey, R. J., Barth Netterfield, C., Padilla, I. L., Redmond, S., Rhodes, J. D. Romualdez, L. J., & Schmoll, J. (2017). *SuperBIT: Diffraction Limited Visible/Near UV Imaging from the Stratosphere*. Canadian Astronomical Society Annual Meeting 2017-30-05 - 2017-01-06.
- [32] Garner, R. (2023). SuperBIT – Super Pressure Balloon.
- [33] Gaudi, B. S., Seager, S., Kiessling, A., Mennesson, B., & Warfield, K. (2019). The Habitable Exoplanet Observatory (HabEx). *Astro2020 APC White Paper*, Volume 51, Number 7.
- [34] Give'on, A., Kern, B., Shaklan, S., Moody, D. C., & Pueyo, L. (2007). Broadband wavefront correction algorithm for high-contrast imaging systems. In *Astronomical Adaptive Optics Systems and Applications III*, volume 6691 (pp. 66910A).: International Society for Optics and Photonics.
- [35] Goodman, J. W. (2005). *Introduction to Fourier optics*. Roberts and Company Publishers.
- [36] Groff, T. D., Kasdin, N. J., Carlotti, A., & Riggs, A. J. E. (2012). Broadband focal plane wavefront control of amplitude and phase aberrations. In *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, volume 8442 (pp. 84420C).: International Society for Optics and Photonics.
- [37] Groff, T. D., Riggs, A. J. E., Kern, B., & Kasdin, N. J. (2015). Methods and limitations of focal plane sensing, estimation, and control in high-contrast imaging. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(1), 011009. Publisher: International Society for Optics and Photonics.
- [38] Hecht, E. (2017). *Optics*. Pearson, 5 edition.
- [39] Hertz, P. & Clampin, M. (2023). NASA Town Hall 241st AAS Meeting.

- [40] Heymans, C., Van Waerbeke, L., Bacon, D., Berge, J., Bernstein, G., Bertin, E., Bridle, S., Brown, M. L., Clowe, D., Dahle, H., Erben, T., Gray, M., Hettterscheidt, M., Hoekstra, H., Hudelot, P., Jarvis, M., Kuijken, K., Margoniner, V., Massey, R., Mellier, Y., Nakajima, R., Refregier, A., Rhodes, J., Schrabback, T., & Wittman, D. (2006). The Shear Testing Programme - I. Weak lensing analysis of simulated ground-based observations. *Monthly Notices of the Royal Astronomical Society*, 368, 1323–1339. ADS Bibcode: 2006MNRAS.368.1323H.
- [41] Jensen-Clem, R., Mawet, D., Gonzalez, C. A. G., Absil, O., Belikov, R., Currie, T., Kenworthy, M. A., Marois, C., Mazoyer, J., Ruane, G., Tanner, A., & Cantalloube, F. (2017). A New Standard for Assessing the Performance of High Contrast Imaging Systems. 155(1), 19. Publisher: American Astronomical Society.
- [42] Juanola-Parramon, R., Zimmerman, N. T., Pueyo, L., Bolcar, M., Ruane, G., Krist, J., & Groff, T. (2019). The LUVOIR Extreme Coronagraph for Living Planetary Systems (ECLIPS) II. Performance evaluation, aberration sensitivity analysis and exoplanet detection simulations. In *Techniques and Instrumentation for Detection of Exoplanets IX*, volume 11117 (pp. 21–36): SPIE.
- [43] Kaiser, N., Squires, G., & Broadhurst, T. (1995). A Method for Weak Lensing Observations. *The Astrophysical Journal*, 449, 460. ADS Bibcode: 1995ApJ...449..460K.
- [44] Kasdin, N. J., Bailey, V. P., Mennesson, B., Zellem, R. T., Ygouf, M., Rhodes, J., Luchik, T., Zhao, F., Riggs, A. J. E., Seo, B.-J., Krist, J., Kern, B., Tang, H., Nemati, B., Groff, T. D., Zimmerman, N., Macintosh, B., Turnbull, M., Debes, J., Douglas, E. S., & Lupu, R. E. (2020). The Nancy Grace Roman Space Telescope Coronagraph Instrument (CGI) technology demonstration. In *Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave*, volume 11443 (pp. 300–313): SPIE.
- [45] Kasdin, N. J., Vanderbei, R. J., Spergel, D. N., & Littman, M. G. (2003). Extrasolar Planet Finding via Optimal Apodized-Pupil and Shaped-Pupil Coronagraphs. *The Astrophysical Journal*, 582(2), 1147. Publisher: IOP Publishing.
- [46] Kasting, J., Traub, W., Roberge, A., Leger, A., Schwartz, A., Wootten, A., Vosteen, A., Lo, A., Brack, A., Tanner, A., Coustenis, A., Lane, B., Oppenheimer, B., Mennesson, B., Lopez, B., Grillmair, C., Beichman, C., Cockell, C., Hanot, C., McCarthy, C., Stark, C., Marois, C., Aime, C., Angerhausen, D., Montes, D., Wilner, D., Defrere, D., Mourard, D., Lin, D., Kite, E., Chassefiere, E., Malbet, F., Tian, F.,

- Westall, F., Illingworth, G., Vasisht, G., Serabyn, G., Marcy, G., Bryden, G., White, G., Laughlin, G., Torres, G., Hammel, H., Ferguson, H., Shibai, H., Rottgering, H., Surdej, J., Wiseman, J., Ge, J., Bally, J., Krist, J., Monnier, J., Trauger, J., Horner, J., Catanzarite, J., Harrington, J., Nishikawa, J., Stapelfeldt, K., von Braun, K., Biazzo, K., Carpenter, K., Balasubramanian, K., Kaltenecker, L., Postman, M., Spaans, M., Turnbull, M., Levine, M., Burchell, M., Ealey, M., Kuchner, M., Marley, M., Dominik, M., Mountain, M., Kenworthy, M., Muterspaugh, M., Shao, M., Zhao, M., Tamura, M., Kasdin, N., Haghhighipour, N., Kiang, N., Elias, N., Wolf, N., Mason, N., Absil, O., Guyon, O., Lay, O., Borde, P., Fouque, P., Kalas, P., Lowrance, P., Plavchan, P., Hinz, P., Kervella, P., Chen, P., Akeson, R., Soummer, R., Waters, R., Barry, R., Kendrick, R., Brown, R., Vanderbei, R., Woodruff, R., Danner, R., Allen, R., Polidan, R., Seager, S., MacPhee, S., Hosseini, S., Metchev, S., Kafka, S., Ridgway, S., Rinehart, S., Unwin, S., Shaklan, S., ten Brummelaar, T., Mazeh, T., Meadows, V., Weiss, W., Danchi, W., Ip, W., & Rabbia, Y. (2009). Exoplanet Characterization and the Search for Life. 2010, 151. Conference Name: astro2010: The Astronomy and Astrophysics Decadal Survey Place: eprint: arXiv:0911.2936.
- [47] Kiefer, S., Bohn, A. J., Quanz, S. P., Kenworthy, M., & Stolker, T. (2021). Spectral and angular differential imaging with SPHERE/IFS - Assessing the performance of various PCA-based approaches to PSF subtraction. *Astronomy & Astrophysics*, 652, A33. Publisher: EDP Sciences.
- [48] Kitching, T. D., Balan, S. T., Bridle, S., Cantale, N., Courbin, F., Eifler, T., Gentile, M., Gill, M. S. S., Harmeling, S., Heymans, C., Hirsch, M., Honscheid, K., Kacprzak, T., Kirkby, D., Margala, D., Massey, R. J., Melchior, P., Nurbaeva, G., Patton, K., Rhodes, J., Rowe, B. T. P., Taylor, A. N., Tewes, M., Viola, M., Witherick, D., Voigt, L., Young, J., & Zuntz, J. (2012). Image analysis for cosmology: results from the GREAT10 Galaxy Challenge. *Monthly Notices of the Royal Astronomical Society*, 423, 3163–3208. ADS Bibcode: 2012MNRAS.423.3163K.
- [49] Krist, J. (2020). WFIRST CGI OS9 Time Series Simulations.
- [50] Krist, J. (2022). Observing Scenario (OS) 11 time series simulations for the Hybrid Lyot Coronagraph Band 1.
- [51] Krist, J., Nemati, B., & Mennesson, B. (2016). Numerical modeling of the proposed WFIRST-AFTA coronagraphs and their predicted performances. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2, 011003.

- [52] Lafrenière, D., Marois, C., Doyon, R., & Barman, T. (2009). HST/NICMOS DETECTION OF HR 8799 b IN 1998. *The Astrophysical Journal*, 694(2), L148. Publisher: The American Astronomical Society.
- [53] Laginja, I. (2020). PASTIS: v2.0.0 Semi-analytical PASTIS for LUVOIR.
- [54] Laginja, I., Sauvage, J.-F., Mugnier, L. M., Pueyo, L., Perrin, M. D., Noss, J., Will, S. D., Brooks, K. J., Por, E. H., Petrone, P., & Soummer, R. (2021a). Wavefront tolerances of space-based segmented telescopes at very high contrast: Experimental validation. *Astronomy & Astrophysics*. Publisher: EDP Sciences.
- [55] Laginja, I., Soummer, R., Mugnier, L. M., Pueyo, L. A., Sauvage, J.-F., Leboulleux, L., Coyle, L. E., & Knight, J. S. (2021b). Analytical tolerancing of segmented telescope co-phasing for exo-Earth high-contrast imaging. *Journal of Astronomical Telescopes, Instruments, and Systems*, 7(1), 015004. Publisher: SPIE.
- [56] Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. (2010). ASTROMETRY.NET: BLIND ASTROMETRIC CALIBRATION OF ARBITRARY ASTRONOMICAL IMAGES. *The Astronomical Journal*, 139(5), 1782. Publisher: The American Astronomical Society.
- [57] Leboulleux, L., Sauvage, J.-F., Pueyo, L. A., Fusco, T., Soummer, R., Mazoyer, J., Sivaramakrishnan, A., N'Diaye, M., & Fauvarque, O. (2018). Pair-based Analytical model for Segmented Telescopes Imaging from Space for sensitivity analysis. *Journal of Astronomical Telescopes, Instruments, and Systems*, 4(03), 1.
- [Li] Li, N. Gravitational Lensing.
- [59] Lin, C.-A. (2016). *Cosmology with weak-lensing peak counts*. PhD thesis, École doctorale Astronomie et Astrophysique d'Île-de-France. arXiv:1612.04041 [astro-ph].
- [60] Loh, E. D. & Spillar, E. J. (1986). Photometric Redshifts of Galaxies. *The Astrophysical Journal*, 303, 154.
- [61] Lyot, B. (1939). The study of the solar corona and prominences without eclipses (George Darwin Lecture, 1939). *Monthly Notices of the Royal Astronomical Society*, 99, 580. ADS Bibcode: 1939MNRAS..99..580L.
- [62] Malbet, F., Yu, J. W., & Shao, M. (1995). High-Dynamic-Range Imaging Using a Deformable Mirror for Space Coronagraphy. *Publications of the Astronomical Society of the Pacific*, 107, 386. ADS Bibcode: 1995PASP..107..386M.

- [63] Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. (2006). Angular Differential Imaging: A Powerful High-Contrast Imaging Technique*. *The Astrophysical Journal*, 641(1), 556. Publisher: IOP Publishing.
- [64] Marois, C., Lafrenière, D., Macintosh, B., & Doyon, R. (2008). Confidence Level and Sensitivity Limits in High-Contrast Imaging*. *The Astrophysical Journal*, 673(1), 647. Publisher: IOP Publishing.
- [65] Massey, R., Heymans, C., Bergé, J., Bernstein, G., Bridle, S., Clowe, D., Dahle, H., Ellis, R., Erben, T., Hettterscheidt, M., High, F. W., Hirata, C., Hoekstra, H., Hudelot, P., Jarvis, M., Johnston, D., Kuijken, K., Margoniner, V., Mandelbaum, R., Mellier, Y., Nakajima, R., Paulin-Henriksson, S., Peeples, M., Roat, C., Refregier, A., Rhodes, J., Schrabback, T., Schirmer, M., Seljak, U., Semboloni, E., & Van Waerbeke, L. (2007). The Shear Testing Programme 2: Factors affecting high-precision weak-lensing analyses. *Monthly Notices of the Royal Astronomical Society*, 376(1), 13–38.
- [66] Mawet, D., Milli, J., Wahhaj, Z., Pelat, D., Absil, O., Delacroix, C., Boccaletti, A., Kasper, M., Kenworthy, M., Marois, C., Mennesson, B., & Pueyo, L. (2014). FUNDAMENTAL LIMITATIONS OF HIGH CONTRAST IMAGING SET BY SMALL SAMPLE STATISTICS. *The Astrophysical Journal*, 792(2), 97. Publisher: The American Astronomical Society.
- [67] McCleary, J. E., Everett, S. W., Shaaban, M. M., Gill, A. S., Vassilakis, G. N., Huff, E. M., Massey, R. J., Benton, S. J., Brown, A. M., Clark, P., Holder, B., Fraisse, A. A., Jauzac, M., Jones, W. C., Lagattuta, D., Leung, J. S.-Y., Li, L., Luu, T. V. T., Nagy, J. M., Netterfield, C. B., Paracha, E., Redmond, S. F., Rhodes, J. D., Schmoll, J., Sirks, E., & Tam, S. I. (2023). Lensing in the Blue II: Estimating the Sensitivity of Stratospheric Balloons to Weak Gravitational Lensing. arXiv:2307.03295 [astro-ph].
- [68] Mendillo, C. B., Hewawasam, K., Martel, J., Potter, T., Cook, T. A., & Chakrabarti, S. (2022). The PICTURE-C exoplanetary imaging balloon mission: laboratory coronagraph demonstrations of high-contrast imaging and low-order wavefront control. In *Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave*, volume 12180 (pp. 776–785).: SPIE.
- [69] Mennesson, B., Gaudi, B. S., Seager, S., Kiessling, A., & Warfield, K. (2020a). The Habitable Exoplanet Observatory mission concept. In M. Lystrup, N. Batalha, E. C. Tong, N. Siegler, & M. D. Perrin (Eds.), *Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave* (pp. 200). Online Only, United States: SPIE.

- [70] Mennesson, B., Juanola-Parramon, R., Nemati, B., Ruane, G., Bailey, V. P., Bolcar, M., Martin, S., Zimmerman, N., Stark, C., Pueyo, L., Benford, D., Cady, E., Crill, B., Douglas, E., Gaudi, B. S., Kasdin, J., Kern, B., Krist, J., Kruk, J., Luchik, T., Macintosh, B., Mandell, A., Mawet, D., McEnery, J., Meshkat, T., Poberezhskiy, I., Rhodes, J., Riggs, A. J., Turnbull, M., Roberge, A., Shi, F., Siegler, N., Stapelfeldt, K., Ygouf, M., Zellem, R., & Zhao, F. (2020b). Paving the Way to Future Missions: the Roman Space Telescope Coronagraph Technology Demonstration. *arXiv:2008.05624 [astro-ph]*. arXiv: 2008.05624.
- [71] Miyazaki, S., Hamana, T., Ellis, R. S., Kashikawa, N., Massey, R. J., Taylor, J., & Refregier, A. (2007). A Subaru Weak-Lensing Survey. I. Cluster Candidates and Spectroscopic Verification. *The Astrophysical Journal*, 669(2), 714. Publisher: IOP Publishing.
- [72] Moriarty, C., Brooks, K., Soummer, R., Perrin, M., Comeau, T., Brady, G., Gontrum, R., & Petrone, P. (2018). High-contrast Imager for Complex Aperture Telescopes (HiCAT): 6. Software Control Infrastructure and Calibration. *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*, (pp. 176). arXiv: 1903.03192.
- [73] {National Academies of Sciences Engineering} (2021). Interactive Overview: Pathways to Discovery in Astronomy and Astrophysics for the 2020s.
- [74] N'Diaye, M., Choquet, E., Pueyo, L., Elliot, E., Perrin, M. D., Wallace, J. K., Groff, T., Carlotti, A., Mawet, D., Sheckells, M., Shaklan, S., Macintosh, B., Kasdin, N. J., & Soummer, R. (2013). High-contrast imager for complex aperture telescopes (HiCAT): 1. testbed design. In *Techniques and Instrumentation for Detection of Exoplanets VI*, volume 8864 (pp. 558–567): SPIE.
- [75] Nemati, B. (2020). Photon counting and precision photometry for the Roman Space Telescope Coronagraph. In *Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave*, volume 11443 (pp. 884–895): SPIE.
- [76] Nemati, Bijan, N. & Miller, S. (2022). EMCCD Detect.
- [77] Nemati, Bijan, N., Miller, S., & Ludwick, K. (2022). Photon Count.
- [78] Nemiroff, R. & Bonnell, J. (2006). Astronomy Picture of the Day.
- [79] Netterfield, C. B. (2012). A Balloon-borne Near-UV/Visible Light Telescope and Test-bed. Published: CSA Class Grant and Contribution Program to Support Research, Awareness and Learning in Space Science and Technology.

- [80] Netterfield, C. B. (2022). *GigaBIT: A high resolution, wide field telescope operating on the edge of space*. CFI Proposal, University of Toronto.
- [81] Noss, J., Fowler, J., Moriarty, C., Brooks, K. J., Laginja, I., Perrin, M. D., Soummer, R., Comeau, T., & Olszewski, H. (2021). spacetelescope/catkit: vo.36.1.
- [82] Okolski, G. (2008). A Brief History of the Hubble Space Telescope.
- [83] Peacock, J. (1999). *Cosmological Physics*. Cambridge University Press.
- [84] Perrin, M. D., Pueyo, L., Gorkom, K. V., Brooks, K., Rajan, A., Girard, J., & Lajoie, C.-P. (2018). Updated optical modeling of JWST coronagraph performance contrast, stability, and strategies. In *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*, volume 10698 (pp. 1069809): International Society for Optics and Photonics.
- [85] Perrin, M. D., Sivaramakrishnan, A., Makidon, R. B., Oppenheimer, B. R., & Graham, J. R. (2003). The Structure of High Strehl Ratio Point-Spread Functions. *The Astrophysical Journal*, 596(1), 702. Publisher: IOP Publishing.
- [86] Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Bacigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J.-F., Carron, J., Challinor, A., Chiang, H. C., Chluba, J., Colombo, L. P. L., Combet, C., Contreras, D., Crill, B. P., Cuttaia, F., de Bernardis, P., de Zotti, G., Delabrouille, J., Delouis, J.-M., Di Valentino, E., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Dusini, S., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Fantaye, Y., Farhang, M., Fergusson, J., Fernandez-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W., Herranz, D., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Karacici, A., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J.-M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Lemos, P., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Lilley, M., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y.-Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marcos-Caballero, A., Maris, M., Martin, P. G., Martinelli, M., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D.,

- Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M.-A., Molinari, D., Montier, L., Morgante, G., Moss, A., Natoli, P., Nørgaard-Nielsen, H. U., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Peiris, H. V., Perrotta, F., Pettorino, V., Piacentini, F., Polastri, L., Polenta, G., Puget, J.-L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rosset, C., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Shellard, E. P. S., Sirignano, C., Sirri, G., Spencer, L. D., Sunyaev, R., Suur-Uski, A.-S., Tauber, J. A., Tavagnacco, D., Tenti, M., Toffolatti, L., Tomasi, M., Trombetti, T., Valenziano, L., Valiviita, J., Van Tent, B., Vibert, L., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Zacchei, A., & Zonca, A. (2018). Planck 2018 results. VI. Cosmological parameters. *ArXiv e-prints*. _eprint: 1807.06209.
- [87] Poberezhskiy, I., Luchik, T., Zhao, F., Frerking, M., Basinger, S., Cady, E., Colavita, M. M., Creager, B., Fathpour, N., Goullioud, R., Groff, T., Morrissey, P., Kempenaar, J., Kern, B., Koch, T., Krist, J., Mok, F., Muliere, D., Nemati, B., Riggs, A. J., Seo, B.-J., Shi, F., Shreckengost, B., Steeves, J., & Tang, H. (2021). Roman space telescope coronagraph: engineering design and operating concept. In *Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave*, volume 11443 (pp. 114431V): International Society for Optics and Photonics.
- [88] Pogorelyuk, L. & Kasdin, N. J. (2019). Dark Hole Maintenance and A Posteriori Intensity Estimation in the Presence of Speckle Drift in a High-contrast Space Coronagraph. *The Astrophysical Journal*, 873(1), 95. Publisher: American Astronomical Society.
- [89] Pogorelyuk, L., Kasdin, N. J., & Rowley, C. W. (2019). Reduced Order Estimation of the Speckle Electric Field History for Space-based Coronagraphs. *The Astrophysical Journal*, 881(2), 126. Publisher: American Astronomical Society.
- [90] Pogorelyuk, L., Pueyo, L., & Kasdin, N. J. (2020). On the effects of pointing jitter, actuator drift, telescope rolls, and broadband detectors in dark hole maintenance and electric field order reduction. *Journal of Astronomical Telescopes, Instruments, and Systems*, 6(3), 039001. Publisher: International Society for Optics and Photonics.
- [91] Pueyo, L. (2016). DETECTION AND CHARACTERIZATION OF EXOPLANETS USING PROJECTIONS ON KARHUNEN-LOEVE EIGENIMAGES: FORWARD MODELING. *The Astrophysical Journal*, 824(2), 117. Publisher: The American Astronomical Society.

- [92] Pueyo, L., Kay, J., Kasdin, N. J., Groff, T., McElwain, M., Give'on, A., & Belikov, R. (2009). Optimal dark hole generation via two deformable mirrors with stroke minimization. *Applied Optics*, 48(32), 6296–6312. Publisher: Optical Society of America.
- [93] Pueyo, L., Stark, C., Juanola-Parramon, R., Zimmerman, N., Bolcar, M., Roberge, A., Arney, G., Ruane, G., Riggs, A. J., Belikov, R., Sirbu, D., Redding, D., Soummer, R., Laginja, I., & Will, S. (2019). The LUVOIR Extreme Coronagraph for Living Planetary Systems (ECLIPS) I: searching and characterizing exoplanetary gems. In *Techniques and Instrumentation for Detection of Exoplanets IX*, volume 11117 (pp. 1111703).: International Society for Optics and Photonics.
- [94] Racine, R., Walker, G. A. H., Nadeau, D., Doyon, R., & Marois, C. (1999). Speckle Noise and the Detection of Faint Companions. *Publications of the Astronomical Society of the Pacific*, 111(759), 587. Publisher: IOP Publishing.
- [95] Redmond, S. (2018). *Thermal Design and Control for Stratospheric Balloon-borne Telescopes*. Thesis. Accepted: 2018-11-15T21:02:35Z.
- [96] Redmond, S. F., Pogorelyuk, L., Pueyo, L., Por, E., Noss, J., Will, S. D., Laginja, I., Brooks, K., Maclay, M., Fowler, J., Kasdin, N. J., Perrin, M. D., & Soummer, R. (2022a). Implementation of a dark zone maintenance algorithm for speckle drift correction in a high contrast space coronagraph. *Journal of Astronomical Telescopes, Instruments, and Systems*, 8(3), 035001. Publisher: SPIE.
- [97] Redmond, S. F., Pueyo, L., Pogorelyuk, L., Noss, J., Will, S. D., Laginja, I., Kasdin, N. J., Perrin, M. D., & Soummer, R. (2021a). Implementation of a broadband focal plane estimator for high-contrast dark zones. In *Techniques and Instrumentation for Detection of Exoplanets X*, volume 11823 (pp. 543–553).: SPIE.
- [98] Redmond, S. F., Pueyo, L., Pogorelyuk, L., Por, E., Noss, J., Laginja, I., Brooks, K., Perrin, M., Soummer, R., & Kasdin, J. (2021b). Dark zone maintenance results for segmented aperture wavefront error drift in a high contrast space coronagraph. In *Techniques and Instrumentation for Detection of Exoplanets X*, volume 11823 (pp. 491–504).: SPIE.
- [99] Redmond, S. F., Pueyo, L., Pogorelyuk, L., Por, E. H., Noss, J., Brooks, K., Laginja, I., Perrin, M. D., Soummer, R., & Kasdin, N. J. (2022b). Dark zone maintenance for future coronagraphic space missions. In *Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave*, volume 12180 (pp. 869–877).: SPIE.

- [100] Riggs, A. J. E. (2016). *Integrated Wavefront Correction and Bias Estimation for the High-Contrast Imaging of Exoplanets*. Ph.D., Princeton University, United States – New Jersey. ISBN: 9781339815220 Publication Title: ProQuest Dissertations and Theses.
- [101] Riggs, A. J. E., Bailey, V., Moody, D. C., Sidick, E., Balasubramanian, K., Moore, D. M., Wilson, D. W., Ruane, G., Sirbu, D., Gersh-Range, J., Trauger, J., Mennesson, B., Siegler, N., Bendek, E., Groff, T. D., Zimmerman, N. T., Debes, J., Basinger, S. A., & Kasdin, N. J. (2021). Flight mask designs of the Roman Space Telescope coronagraph instrument. In *Techniques and Instrumentation for Detection of Exoplanets X*, volume 11823 (pp. 611–633): SPIE.
- [102] Riggs, A. J. E., Kasdin, N. J., & Groff, T. D. (2016). Recursive Starlight and Bias Estimation for High-Contrast Imaging with an Extended Kalman Filter. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(1), 011017. arXiv: 1602.02044.
- [103] Robbins, M. & Hadwen, B. (2003). The noise performance of electron multiplying charge-coupled devices. *IEEE Transactions on Electron Devices*, 50(5), 1227–1232. Conference Name: IEEE Transactions on Electron Devices.
- [104] Roberge, A., Chen, C. H., Millan-Gabet, R., Weinberger, A. J., Hinz, P. M., Stapelfeldt, K. R., Absil, O., Kuchner, M. J., & Bryden, G. (2012). The Exozodiacal Dust Problem for Direct Observations of Exo-Earths. *Publications of the Astronomical Society of the Pacific*, 124(918), 799. Publisher: IOP Publishing.
- [105] Romualdez, L., Benton, S., Brown, A., Clark, P., Damaren, C., Eifler, T., Fraisse, A., Galloway, M., Hartley, J., Jauzac, M., Jones, W., Li, S., Luu, V., Massey, R., McCleary, J., Netterfield, C., Redmond, S., Rhodes, J., Schmoll, J., & Tam, S.-I. (2018). *Overview, design, and flight results from SuperBIT: a high-resolution, wide-field, visible-to-near-UV balloon-borne astronomical telescope*.
- [106] Romualdez, L. J. (2017). *Design, Implementation, and Operational Methodologies for Sub-Arcsecond Attitude Determination, Control, and Stabilization of the Superpressure Balloon-borne Imaging Telescope (SuperBIT)*. PhD Thesis, University of Toronto.
- [107] Romualdez, L. J., Benton, S. J., Brown, A. M., Clark, P., Damaren, C. J., Eifler, T., Fraisse, A. A., Galloway, M. N., Gill, A., Hartley, J. W., Holder, B., Huff, E. M., Jauzac, M., Jones, W. C., Lagattuta, D., Leung, J. S.-Y., Li, L., Luu, T. V. T., Massey, R. J., McCleary, J., Mullaney, J., Nagy, J. M., Netterfield, C. B., Redmond, S., Rhodes, J. D., Schmoll, J., Shaaban, M. M., Sirks, E., & Tam, S.-I. (2020). Robust

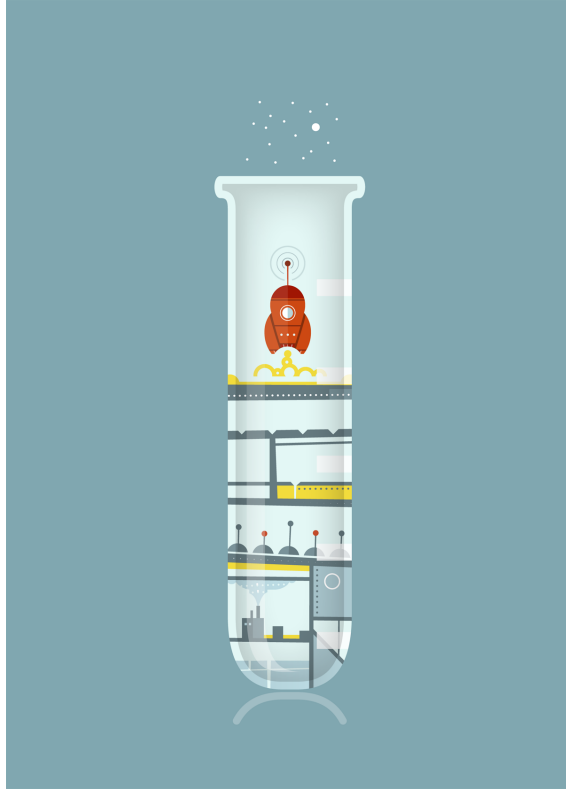
- diffraction-limited near-infrared-to-near-ultraviolet wide-field imaging from stratospheric balloon-borne platforms—Super-pressure Balloon-borne Imaging Telescope performance. *Review of Scientific Instruments*, 91(3), 034501. Publisher: American Institute of Physics.
- [108] Rosenblatt, F. (1971). A two-color photometric method for detection of extra-solar planetary systems. *Icarus*, 14(1), 71–93.
- [109] Ruane, G., Riggs, A. J. E., Serabyn, E., Baxter, W., Prada, C. M., Mawet, D., Noyes, M., Poon, P. K., & Tabiryan, N. (2022). Broadband Vector Vortex Coronagraph Testing at NASA’s High Contrast Imaging Testbed Facility. arXiv:2207.13742 [astro-ph, physics:physics].
- [110] Sahoo, A., Laginja, I., Pueyo, L., Soummer, R., Coyle, L. E., Knight, J. S., & East, M. (2021). LUVOIR A Segment level thermo-mechanical stability requirements. In *Techniques and Instrumentation for Detection of Exoplanets X*, volume 11823 (pp. 118231N): SPIE.
- [111] Schneider, G. & Silverstone, M. D. (2003). Coronagraphy with HST/NICMOS: detectability is a sensitive issue. In *High-Contrast Imaging for Exo-Planet Detection*, volume 4860 (pp. 1–9): SPIE.
- [112] Schwieterman, E. W., Kiang, N. Y., Parenteau, M. N., Harman, C. E., DasSarma, S., Fisher, T. M., Arney, G. N., Hartnett, H. E., Reinhard, C. T., Olson, S. L., Meadows, V. S., Cockell, C. S., Walker, S. I., Grenfell, J. L., Hegde, S., Rugheimer, S., Hu, R., & Lyons, T. W. (2018). Exoplanet Biosignatures: A Review of Remotely Detectable Signs of Life. *Astrobiology*, 18(6), 663–708. Publisher: Mary Ann Liebert, Inc., publishers.
- [113] Scientific Volume Imaging B.V. (2018). Point Spread Function.
- [114] Seo, B.-J., Patterson, K., Balasubramanian, K., Crill, B., Chui, T., Echeverri, D., Kern, B., Marx, D., Moody, D., Prada, C. M., Ruane, G., Shi, F., Shaw, J., Siegler, N., Tang, H., Trauger, J., Wilson, D., & Zimmer, R. (2019). Testbed demonstration of high-contrast coronagraph imaging in search for Earth-like exoplanets. In *Techniques and Instrumentation for Detection of Exoplanets IX*, volume 11117 (pp. 111171V): International Society for Optics and Photonics.
- [115] Shaaban, M. M., Gill, A. S., McCleary, J., Massey, R. J., Benton, S. J., Brown, A. M., Damaren, C. J., Eifler, T., Fraise, A. A., Everett, S., Galloway, M. N., Henderson, M., Holder, B., Huff, E. M., Jauzac, M., Jones, W. C., Lagattuta, D., Leung, J. S.-Y.,

- Li, L., Luu, T. V. T., Nagy, J. M., Netterfield, C. B., Redmond, S. F., Rhodes, J. D., Robertson, A., Schmoll, J., Sirks, E., & Sivanandam, S. (2022). Weak Lensing in the Blue: A Counter-intuitive Strategy for Stratospheric Observations. *The Astronomical Journal*, 164(6), 245. Publisher: The American Astronomical Society.
- [116] Shaklan, S. B. & Green, J. J. (2006). Reflectivity and optical surface height requirements in a broadband coronagraph. 1. Contrast floor due to controllable spatial frequencies. *Applied Optics*, 45(21), 5143–5153. Publisher: Optica Publishing Group.
- [117] Shapiro, J., Savransky, D., Ruffio, J.-B., Ranganathan, N., & Macintosh, B. (2019). Detecting Planets from Direct-imaging Observations Using Common Spatial Pattern Filtering. *The Astronomical Journal*, 158(3), 125. Publisher: The American Astronomical Society.
- [118] Shi, F., An, X., Balasubramanian, K., Cady, E., Kern, B., Lam, R., Marx, D., Prada, C., Moody, D., Patterson, K., Poberezhskiy, I., Seo, B.-J., Shields, J., Sidick, E., Tang, H., Trauger, J., Truong, T., White, V., Wilson, D., & Zhou, H. (2017). *Dynamic testbed demonstration of WFIRST coronagraph low order wavefront sensing and control (LOWFS/C)*.
- [119] Shi, F., Balasubramanian, K., Bartos, R., Hein, R., Kern, B., Krist, J., Lam, R., Moore, D., Moore, J., Patterson, K., Poberezhskiy, I., Shields, J., Sidick, E., Tang, H., Truong, T., Wallace, K., Wang, X., & Wilson, D. (2015). Low order wavefront sensing and control for WFIRST-AFTA coronagraph. In *Techniques and Instrumentation for Detection of Exoplanets VII*, volume 9605 (pp. 960509): International Society for Optics and Photonics.
- [120] Simon, D. (2006). *Optimal state estimation : Kalman, H [infinity] and nonlinear approaches*. Hoboken, N.J.: Wiley-Interscience.
- [121] Sivaramakrishnan, A., Soummer, R., Pueyo, L., Wallace, J. K., & Shao, M. (2008). Sensing Phase Aberrations behind Lyot Coronagraphs. *The Astrophysical Journal*, 688, 701–708. ADS Bibcode: 2008ApJ...688..701S.
- [122] Smith, W. H. (1987). SPECTRAL DIFFERENTIAL IMAGING DETECTION OF PLANETS ABOUT NEARBY STARS. *Publications of the Astronomical Society of the Pacific*, 99(622), 1344. Publisher: IOP Publishing.
- [123] Soummer, R., Brady, G. R., Brooks, K., Comeau, T., Choquet, E., Dillon, T., Egron, S., Gontrum, R., Hagopian, J., Laginja, I., Lebouilleux, L., Perrin, M. D.,

- Petrone, P., Pueyo, L., Mazoyer, J., N'Diaye, M., Riggs, A. J. E., Shiri, R., Sivaramakrishnan, A., Laurent, K. S., Valenzuela, A.-M., & Zimmerman, N. T. (2018). High-contrast imager for complex aperture telescopes (HiCAT): 5. first results with segmented-aperture coronagraph and wavefront control. In *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*, volume 10698 (pp. 106981O).: International Society for Optics and Photonics.
- [124] Soummer, R., Ferrari, A., Aime, C., & Jolissaint, L. (2007). Speckle Noise and Dynamic Range in Coronagraphic Images. *The Astrophysical Journal*, 669, 642–656.
- [125] Soummer, R., Hagan, J. B., Pueyo, L., Thormann, A., Rajan, A., & Marois, C. (2011). ORBITAL MOTION OF HR 8799 b, c, d USING HUBBLE SPACE TELESCOPE DATA FROM 1998: CONSTRAINTS ON INCLINATION, ECCENTRICITY, AND STABILITY. *The Astrophysical Journal*, 741(1), 55. Publisher: The American Astronomical Society.
- [126] Soummer, R., Laginja, I., Will, S., Juanola-Parramon, R., Iii, P. P., Brady, G., Noss, J., Perrin, M. D., Fowler, J., Kurtz, H., Laurent, K. S., Fogarty, K., McChesney, E., Scott, N., Brooks, K., Comeau, T., Ferrari, M., Gontrum, R., Hagopian, J., Hugot, E., Leboulleux, L., Mazoyer, J., Mugnier, L., N'Diaye, M., Pueyo, L., Sauvage, J.-F., Shiri, R., Sivaramakrishnan, A., Valenzuela, A.-M., & Zimmerman, N. T. (2019). High-contrast imager for complex aperture telescopes (HiCAT): 6. Two deformable mirror wavefront control (Conference Presentation). In *Techniques and Instrumentation for Detection of Exoplanets IX*, volume 11117 (pp. 111171Y).: International Society for Optics and Photonics.
- [127] Soummer, R., Por, E. H., Pourcelot, R., Redmond, S., Laginja, I., Will, S. D., Perrin, M. D., Pueyo, L., Sahoo, A., Petrone, P., Brooks, K. J., Fox, R., Klein, A., Nickson, B., Comeau, T., Ferrari, M., Gontrum, R., Hagopian, J., Leboulleux, L., Leongomez, D., Lugten, J., Mugnier, L. M., N'Diaye, M., Nguyen, M., Noss, J., Sauvage, J.-F., Scott, N., Sivaramakrishnan, A., Subedi, H. B., & Weinstock, S. (2022). High-contrast imager for complex aperture telescopes (HiCAT): 8. Dark zone demonstration with simultaneous closed-loop low-order wavefront sensing and control. In *Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave*, volume 12180 (pp. 816–832).: SPIE.
- [128] Soummer, R., Pueyo, L., Ferrari, A., Aime, C., Sivaramakrishnan, A., & Yaitskova, N. (2009). Apodized Pupil Lyot Coronagraphs for Arbitrary Apertures. II. Theoretical Properties and Application to Extremely Large Telescopes. *The Astrophysical Journal*, 695, 695.

- [129] Soummer, R., Pueyo, L., & Larkin, J. (2012). DETECTION AND CHARACTERIZATION OF EXOPLANETS AND DISKS USING PROJECTIONS ON KARHUNEN-LOÈVE EIGENIMAGES. *The Astrophysical Journal*, 755(2), L28. Publisher: American Astronomical Society.
- [130] Spalding, E. (2018). It's a bird, it's a planet, it's a ... speckle?
- [131] Spergel, D., Gehrels, N., Baltay, C., Bennett, D., Breckinridge, J., Donahue, M., Dressler, A., Gaudi, B. S., Greene, T., Guyon, O., Hirata, C., Kalirai, J., Kasdin, N. J., Macintosh, B., Moos, W., Perlmutter, S., Postman, M., Rauscher, B., Rhodes, J., Wang, Y., Weinberg, D., Benford, D., Hudson, M., Jeong, W. S., Mellier, Y., Traub, W., Yamada, T., Capak, P., Colbert, J., Masters, D., Penny, M., Savransky, D., Stern, D., Zimmerman, N., Barry, R., Bartusek, L., Carpenter, K., Cheng, E., Content, D., Dekens, F., Demers, R., Grady, K., Jackson, C., Kuan, G., Kruk, J., Melton, M., Nemati, B., Parvin, B., Poberezhskiy, I., Peddie, C., Ruffa, J., Wallace, J. K., Whipple, A., Wollack, E., & Zhao, F. (2015). Wide-Field Infrared Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report. Publication Title: arXiv e-prints ADS Bibcode: 2015arXiv150303757S.
- [132] Subedi, H. B., Juanola-Parramon, R., & Groff, T. D. (2019). : San Diego, CA. NTRS Author Affiliations: NASA Goddard Space Flight Center, Maryland Univ. NTRS Report/Patent Number: GSFC-E-DAA-TN72786 NTRS Document ID: 20190030904 NTRS Research Center: Goddard Space Flight Center (GSFC).
- [133] Sun, H. (2019). *Efficient Wavefront Sensing and Control for Space-based High-contrast Imaging*. Ph.D., Princeton University, United States – New Jersey. ISBN: 9781392678770.
- [134] Sun, H., Kasdin, N. J., & Vanderbei, R. (2018). Identification and adaptive control of a high-contrast focal plane wavefront correction system. *Journal of Astronomical Telescopes, Instruments, and Systems*, 4(4), 049006. Publisher: International Society for Optics and Photonics.
- [135] Sun, H., Sun, H., Goun, A., Redmond, S., Galvin, M., Groff, T., Rizzo, M., & Kasdin, N. J. (2020). High-contrast integral field spectrograph (HCIFS): multi-spectral wavefront control and reduced-dimensional system identification. *Optics Express*, 28(15), 22412–22423. Publisher: Optical Society of America.
- [136] Taboga, M. (2017). Properties of the expected value. In *Lectures on probability theory and mathematical statistics, Third edition*. Kindle Direct Publishing. Online appendix.

- [137] Talbot, W. H. F. (1836). Facts relating to optical science, No. IV. *Philos. Mag.* 9, (pp. 401–407).
- [138] Team, T. L. (2019). The LUVOIR Mission Concept Study Final Report. *arXiv:1912.06219 [astro-ph]*. arXiv: 1912.06219.
- [139] Trauger, J., Moody, D., Gordon, B., Krist, J., & Mawet, D. (2011). A hybrid Lyot coronagraph for the direct imaging and spectroscopy of exoplanet systems: recent results and prospects. In *Techniques and Instrumentation for Detection of Exoplanets V*, volume 8151 (pp. 81510G): International Society for Optics and Photonics.
- [140] Uyama, T., Xie, C., Aoyama, Y., Beichman, C. A., Hashimoto, J., Dong, R., Hasegawa, Y., Ikoma, M., Mawet, D., McElwain, M. W., Ruffio, J.-B., Wagner, K. R., Wang, J. J., & Zhou, Y. (2021). Keck/OSIRIS Paß High-contrast Imaging and Updated Constraints on PDS 70b. *The Astronomical Journal*, 162(5), 214. Publisher: The American Astronomical Society.
- [141] Wang, J., Delorme, J.-R., Ruffio, J.-B., Morris, E., Jovanovic, N., Echeverri, D., Schofield, T., Pezzato, J., Skemer, A., & Mawet, D. (2021). *High resolution spectroscopy of directly imaged exoplanets with KPIC*.
- [142] Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., Hirata, C., Riess, A. G., & Rozo, E. (2013). Observational probes of cosmic acceleration. *Physics Reports*, 530, 87–255. _eprint: 1201.2434.
- [143] Wright, J. & Gaudi, B. (2013). : (pp. 489–540).
- [144] Ygouf, M., Beichman, C., Hodapp, K., & Roellig, T. (2017). *High-contrast imaging with the JWST-NIRSpec Integral Field Unit*. Conference Name: SF2A-2017: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics Pages: Di ADS Bibcode: 2017sf2a.conf..325Y.
- [145] Ygouf, M., Zimmerman, N., Bailey, V., Krist, J., Zellem, R., & Debes, J. (2021). *Roman Coronagraph Instrument Post Processing Report - OS9 HLC Distribution*. Technical report, California Institute of Technology.
- [146] Zackay, B. & Ofek, E. O. (2017). How to COAAD Images. II. A Coaddition Image that is Optimal for Any Purpose in the Background-dominated Noise Limit. *The Astrophysical Journal*, 836(2), 188. Publisher: American Astronomical Society.
- [147] Zimmerman, N., Rizzo, M., Stark, C., & Bogat, E. (2021). Exoscene: v1.2 simulate direct images of exoplanetary systems.



THIS THESIS WAS TYPESET using L^AT_EX, originally developed by Leslie Lamport and based on Donald Knuth's T_EX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. The above illustration, "Science Experiment 02", was created by Ben Schlitter and released under [CC BY-NC-ND 3.0](#). A template that can be used to format a PhD thesis with this look and feel has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/Dissertate or from its author, Jordan Suchow, at suchow@post.harvard.edu.