Development of MKID Instrumentation for Exoplanet Direct Imaging

Rupert Dodkins

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2018

Abstract

Microwave Kinetic Inductance Detectors (MKIDs) are a highly sensitive cryogenic photodetector. These devices operate by measuring the increase in the surface impedance of a superconducting film when Cooper Pairs are broken by incident photons. This technique allows MKIDs inherent spectral resolution and instantaneous read out with essentially no false counts. In contrast to other cryogenic detectors, the intrinsic multiplexing of the pixels in frequency domain means that arrays of several tens of thousand of pixels can be readout with just a few feed-lines using room temperature radio electronics. These unique properties make MKIDs potentially transformative for exoplanet direct imaging, where residual speckle noise currently limits the minimum mass of the discernible exoplanets to several Jupiter masses. In this thesis I describe some of the developments of MKID instruments required for exoplanet imaging, and I investigate the potential of MKIDs to provide new discoveries in this field.

One of the required developments was the automation of the tuning of pixel biasing prior to taking a set of observations. Previously this required inspecting the transmission profile of each pixel by eye and then selecting the optimal power. For the large format arrays necessary for exoplanet direct imaging (on the order of tens of thousand pixels) this process can take approximately 24 human hours to complete. I show that a machine learning-based algorithm can replicate this process in a matter of minutes with equivalent accuracy to that of humans when compared to a control dataset. I then discuss an extension to this algorithm that should enhance the phase measurement performance of MKIDs, ultimately improving the limiting mass of detected exoplanets.

Simulation software was developed to make pragmatic predictions on the capabilities of MKIDs for exoplanet direct imaging. This software demonstrates the utility of Dark-Speckle Imaging discrimination technique for a realistic MKID device with appropriate limitations such as maximum count and reduced pixel yield. Possible extensions to this technique that exploit the wavelength dependence of speckles are

also explored, which are found to yield superior performance to standard differential imaging methods. I also make predictions on the performance of a next generation MKID-based instrument systems. Preliminary results show that this system can achieve the contrast performance necessary for Neptune sized exoplanets and below.

Finally, I present on-sky results of planetary nebula NGC 6751 taken with the ARCONS MKID instrument at Palomar. The inherent spectral resolution of the device meant that a low resolution spectrum of the central star could be created that provided further evidence to clarify its spectral type. The central star was also explored in the time domain at the fine resolution of MKIDs. No periodic feature was found on the time scale of the observation.

Statement of Originality

The work presented in this thesis was completed by me at the Sub-department of Astrophysics at the University of Oxford between October 2014 and September 2018 under the supervision of Dr. Kieran O'Brien and Prof. Niranjan Thatte. I certify that to the best of my knowledge, the content of this thesis is my own work, except where stated otherwise. I have not submitted any of this work for any other degree, diploma or qualification at any other institution of higher education.

Chapter 3 features work that was peer reviewed and published in Dodkins et al. [1] of which I am the lead author. The software was developed during my Long Term attachment at University of California, Santa Barbara (UCSB) using archival measurements of various MKID arrays that they designed and created. The design and implementation of these arrays is presented in: Szypryt [2] and Meeker [3]. In Chapter 4, the wavelength calibration sources were developed by myself first at UCSB and then at the University of Oxford. Chapter's 5 and 7 feature a simulator that I developed by adapting PROPER [4] and CAOS [5] and VIP [6]. I performed the data analysis of the ARCONS data in Chapter 6. I did not take the measurements and the pipeline was written by van Eyken et al. [7].

Rupert Dodkins

September 2018

Acknowledgements

First and foremost, I am very grateful to my supervisors, Dr. Kieran O'Brien and Prof. Niranjan Thatte. Their wisdom and guidance throughout my DPhil has been essential in helping me overcome all obstacles and explore new ideas. I owe a huge thank you to my collaborators at the University of California, Santa Barbara lead by Prof. Ben Mazin. Thank you for taking me into your group and allowing me to work with you on such a revolutionary project. Thank you also to my first supervisors: Prof. Patrick Leahy and Prof. Benjamin Stappers at the University of Manchester, and then Prof. Brad Whitmore at the Space Telescope Science Institute, who started me on this path of instrument research, and without whom this thesis would not be possible.

I am eternally grateful to Ashling and Garret for looking after me during my time at Oxford and magically making each of my problems go away when I needed them to.

I owe a special thank you to Dr. Sumedh Mahashabde who has been a mentor for me for several years. I always enjoyed listening to you talk, whether it was on electron-phonon interactions or the interplay of the members of the Scandinavian eurodance group, Aqua. I would like to thank all of my colleagues and friends at Oxford for their support and the fun times we had. Alex, John, Boon, David, Dani; with whom I have many cherished memories of college guest nights, balls, and office parties.

Finally, I would like to thank my mother, Georgina for her love and support. Thank you for always being there for me and listening to me as I try to explain my research problems to you.

To my mother, Georgina

Contents

1 Exoplanet Direct Detection **1**

1.1 Scientific Motivation 1

1.1.1 Exoplanetology 1

1.1.2 Methods of Detecting Exoplanets 2

1.2 High Contrast Imaging 4

1.2.1 Introduction 4

1.2.2 Important Parameters 7

1.2.3 Speckle Noise 10

1.2.4 Extreme Adaptive Optics 15

1.2.5 Coronagraphy 18

1.2.6 Post-AO Wavefront Sensing 20

1.2.7 Observing Strategies and Post Processing 22

1.3 Detectors for High Contrast Imaging 24

1.3.1 Semiconductor Detectors 25

1.3.2 Superconductor Detectors 26

1.3.3 Microwave Kinetic Inductance Detectors 27

1.4 Thesis Motivation 28

2 Microwave Kinetic Inductance Detectors **29**

2.1 Microwave Resonators 29

2.2 Kinetic Inductance 32

2.3 Photon Detection 35

2.4 Digital Readout 37

2.5 Array Parameters . . . 41
2.6 MKID Instruments . . . 43
2.6.1 Commissioned Instruments . . . 43
2.6.2 Planned Instruments . . . 44
2.7 Summary . . . 49

3 Digital Readout Tuning with Machine Learning 50
3.1 Readout Tuning . . . 50
3.2 Optimal Resonator Biasing . . . 52
3.2.1 Analytical Method . . . 53
3.2.2 Numerical Method . . . 54
3.2.3 Manual Inspection . . . 55
3.2.4 Summary . . . 55
3.2.5 Resonator Pathologies . . . 57
3.3 Resonator Biasing with Machine Learning . . . 63
3.3.1 Architecture . . . 66
3.3.2 Input data . . . 68
3.3.3 Training and Evaluation . . . 69
3.4 Results and Analysis . . . 70
3.4.1 CNN Training . . . 71
3.4.2 Accuracy Comparison . . . 73
3.4.3 Timing Comparison . . . 82
3.5 Future Plans . . . 84
3.6 Summary . . . 87

4 Calibrating MKIDs Observations 88
4.1 Introduction . . . 88
4.1.1 Data Format . . . 88
4.1.2 Data Reduction and Calibration . . . 90
4.1.3 Analysis Calibration . . . 91
4.2 Wavelength Calibration Source Development . . . 93

4.2.1 DARKNESS . . . 93
4.2.2 KIDSpec . . . 96
4.3 Summary . . . 97

5 Simulating and Processing HCI Observations 98
5.1 Motivation . . . 98
5.2 Implementation . . . 99
5.2.1 Atmosphere . . . 101
5.2.2 Telescope . . . 104
5.2.3 Detector . . . 112
5.3 Post-Processing and Analysis . . . 115
5.3.1 Contrast Curves . . . 117
5.3.2 Differential Imaging . . . 118
5.3.3 Discrimination with Speckle Statistics . . . 120
5.4 Future Plans . . . 126
5.5 Summary . . . 126

6 ARCONS Observations of NGC 6751 128
6.1 Introduction . . . 128
6.2 Analysing NGC 6751 with ARCONS . . . 130
6.2.1 Mosaic Creation . . . 132
6.2.2 Spectral and Time Analysis . . . 134
6.3 Summary . . . 137

7 Future Work and Conclusions 138
7.1 Next Generation Instrument Predictions . . . 138
7.2 Investigating MKID Parameters . . . 143
7.3 Concluding Remarks . . . 145

References 146

A MEDIS Parameters 174

List of Figures

1.1 Mass and separation distribution of the known exoplanets 5
1.2 Gallery of exoplanets 6
1.3 Contrast-separation curves of known exoplanets 9
1.4 Contrast-separation values of the predicted population of exoplanets . 9
1.5 The limiting masses of exoplanet observations 10
1.6 VLT aberrations and speckle pattern 11
1.7 The modified Rician PDF 14
1.8 The detection confidence level as a function of I_C/I_S 15
1.9 The basic implementation of adaptive optics 16
1.10 Contrast curves as a function of wavefront error 17
1.11 A schematic of the Apodized Lyot Coronagraph design for SPHERE . 19
1.12 Demonstration of long coronagraph observation 20
1.13 A schematic of one implementation of a HCI system 21
1.14 Nulling of NCPA amplitude errors with phase-shifting interferometry. 22

2.1 The architecture of a single MKID pixel 30
2.2 The transmission profile of a resonator 31
2.3 Transmission profiles for different parameters 33
2.4 Superconductor response to absorbed photons 35
2.5 The basic operation of an MKID 36
2.6 Feedline transmission spectrum 37
2.7 Block Diagram of ARCONS SDR readout 39
2.8 KIDSpec generation 1 readout board 40

2.9 Block diagram of the second generation readout unit. 40
2.10 The ARCONS array and cryostat 44
2.11 The DARKNESS array and cryostat 45
2.12 The MEC array and cryostat 45
2.13 Operating principle of KIDSpec 47
2.14 The predicted performance of KIDSpec 48
2.15 The cryostat of the KIDSpec demonstrator 48

3.1 The power handling behaviour of an ideal resonator. 53
3.2 The GUI for bias (and frequency) classifying resonators. 56
3.3 Example resonators demonstrating different types of pathologies . . . 59
3.4 Schematic of the CNN architecture for selecting the bias point of resonator powersweeps. 66
3.5 The accuracy of the CNN on the different feedlines. 72
3.6 The accuracy of the CNN after different amounts of training. 74
3.7 Confusion matrices comparing the predictions of different methods . . 76
3.8 All confusion matrices comparing each of the methods (not including NM+AM). 78
3.9 The performance of the three automated methods according to different types of powersweep behaviour. 79
3.10 The confusion matrices of one of the MI operators and the three automated methods when compared to MI_{av} and separated according to resonator behaviour. 81
3.11 A toy model of the decision process for adding powersweeps to the training data pool. 85

4.1 The layout of an ObsFile 89
4.2 Phase histogram of a single pixel 92
4.3 PCS for the KIDSpec Laser Box 95
4.4 Photo of KIDSpec Laser Box 96

5.1 An overview of MEDIS 100
5.2 MEDIS parallel processing 101
5.3 The mean profile of C_n^2 as a function of height for Paranal in 2007 . . 103
5.4 The Atmosphere CAOS worksheet for generating aberrations maps . 103
5.5 An evolving atmospheric wavefront error map 104
5.6 A schematic of the main elements of the optical train in Telescope . . 106
5.7 The power spectra used to seed the aberration maps 108
5.8 The wavefront error of differential images 109
5.9 The AO performance with different error sources 110
5.10 A schematic of the NCPA sensing method used 111
5.11 Measured phase data for a single time step when observing HD 91782 with DARKNESS 114
5.12 Overview of MKIDs module of MEDIS for 10 UMA AB 116
5.13 Demonstration of reference PSF subtraction by multichannel spectral differential imaging using MEDIS 119
5.14 Reference and angular differential imaging with MEDIS 120
5.15 A demonstration of DSI. 122
5.16 The light map created with DSI at multiple wavelengths 123
5.17 Light curve statistics at different locations in an image 125
5.18 Demonstration of SSD 126

6.1 Gemini South image of NGC 6751 129
6.2 Spectral resolution distributions of an ARCONS array and the resultant spectra of NGC 6751 131
6.3 True colour image of NGC 6751 with ARCONS 131
6.4 Mosaic images of NGC 6751 demonstrating the scalloping effect . . . 133
6.5 Mosaic Image of NGC 6751 after all processing 134
6.6 Spectra of NGC 6751 135
6.7 The light curve of the WR star at the center of NGC 6751 136
6.8 Periodogram of NGC 6751 137

7.1 The observation of a modified version of HR 8799 as captured by the next generation MKID instrument with 250 s of on-target observation time. a: Effective integration of the target. b: The difference of the mean of the reference observation and target. c: DSI was performed at one wavelength of the reference subtracted hypercube. d: The broadband DSI technique was applied to all wavelengths of the reference subtracted hypercube. The binning time of the DSI was 1 second. . . 141
7.2 Contrast-separation performance of the next generation MKID instrument on the HR 8799 system 142
7.3 SDI reference PSF subtraction for an MKID array with and without hot pixels 144
7.4 Contrast performance of a standard MKID array with various pixel yield performances 144

List of Tables

3.1 Summary of the methods used for biasing resonator powersweeps. . . 57

3.2 The occurrence statistics of the pathologies on a single feedline identified manually. 64

3.3 A summary of all of the feedlines used in training and evaluating the different methods. 68

3.4 A summary of the timing and accuracy performance of some of the methods. 82

7.1 The input parameters for the simulation of the HR 8799 system as observed by a next generation MKID instrument. 139

A.1 The turbulence profile parameters used in CAOS 175

A.2 Astrophysics configuration parameters 175

A.3 Telescope configuration parameters 176

A.4 MKIDs configuration parameters 177

Abreviations

2D	Two Dimensional.
3D	Three Dimensional.
ADI	Angular Differential Imaging.
AM	Analytical Method.
ARCONS	the Array Camera for Optical to Near-IR Spectrophotometry.
CAOS	Code for Adaptive Optics Systems.
CCD	Charge Coupled Device.
CMOS	Complementary Metal–Oxide–Semiconductor.
CNN	Convolution Neural Network.
CPA	Common Path Aberrations.
DAC/ADC	Digital to Analogue Converter and Analogue to Digital Converter board.
DARKNESS	the DARK-speckle Near-infrared Energy resolving Superconducting Spectrophotometer.
DM	Deformable Mirror.
DSI	Dark Speckle Imaging.
ELT	the Extremely Large Telescope.

FITS	Flexible Image Transport System.
FoV	Field Of View.
FPGA	Field Programmable Gate Arrays.
FPWFS	Focal Plane Wavefront Sensor.
FWHM	Full Width Half Maximum.
H2RG	HAWAII-2RG.
HCI	High Contrast Imaging.
IQ	In phase-Quadrature.
IWA	Inner Working Angle.
KIDSpec	the Kinetic Inductance Spectrograph.
LLSG	the Locally Low-rank, Sparse, and Gaussian noise components algorithm.
LTA	Long Term Attachment.
MagAO	Magellan AO.
MEC	the MKID Exoplanet Camera.
MEDIS	the MKID Exoplanet Direct Imaging Simulator.
MI	Manual Inspection.
MKID	Microwave Kinetic Inductance Detector.
MR	Modified Rician Distribution.
NCPA	Non-Common Path Aberrations.
NM	Numerical Method.

ObsFile	Observation File.
OOPP	Out Of Pupil Plane.
P1640	Project 1640.
PCS	the Planetary Camera and Spectrograph.
PDF	Probability Density Function.
PICTURE-C	the Planetary Imaging Concept Testbed Using a Recoverable Experiment – Coronagraph.
PSD	Power Spectral Density.
PSF	Point Spread Function.
PSI	the Planetary System Imager.
PtSi	Platinum Scilicide.
QE	Quantum Efficiency.
RDI	Reference Differential Imaging.
RGB	Red Green Blue.
RMS	Root Mean Square.
ROACH	Reconfigurable Open Architecture Computing Hardware.
SCExAO	the Subaru Coronagraphic Extreme Adaptive Optics system.
SCRAPS	SuperConducting Resonator Analysis and Plotting Software.
SDI	multi-channel simultaneous Spectral Differential Imaging.
SDR	Software Defined Radio.
SNR	Signal to Noise ratio.

SPHERE	the Spectro-Polarimetric High-contrast Exoplanet REsearch instrument.
SR	Strehl Ratio.
SSD	Stochastic Speckle Discrimination.
STJ	Superconducting Tunnel Junction.
TES	Transition Edge Sensor.
TiN	Titanium Nitride.
TLS	Two Level System noise.
TMT	the Thirty Meter Telescope.
UVOIR	Ultraviolet Optical and near-Infrared.
VIP	Vortex Imaging Pipeline.
VLT	Very Large Telescope.
VR	Velocity Ratio.
WFS	Wavefront Sensor.
WR	Wolf Rayet.
xAO	eXtreme Adaptive Optics.

Chapter 1

Exoplanet Direct Detection

1.1 Scientific Motivation

1.1.1 Exoplanetology

Extrasolar planets, or exoplanets, are planets that exist outside of the Solar System. In 1584 Giordano Bruno wrote that "There are countless suns and countless earths all rotating around their suns in exactly the same way as the seven planets of our system". Over 400 years later, in 1995, the first exoplanet around a main sequence star was discovered by Mayor and Queloz [8]. Since then 3,786[1] of exoplanets have been discovered – a roughly two fold increase every ∼27.3 months [9].

The observation and study of exoplanets is opening up a fascinating area of the Universe as we work towards answering some of the most profound questions such as "How does the Solar System fit in when compared to other solar systems?", "How common are Earth-like planets?" and "Is there other life in the Universe?". Recent discoveries such as the Earth-like planet in orbit around the nearest star to our own, Proxima Centauri (an M-type red dwarf) [10], and the discovery of seven planets (some in the habitable zone) around TRAPPIST-1 (also an M-dwarf) [11] demonstrate that while exo-Earths may be fairly commonplace, their worlds will be very alien to our own.

[1] http://exoplanet.eu/catalog/ as of 2 June 2018

In the last 25 years there has been an abundance of paradigm shifting discoveries: the detection of hot Jupiters with orbital periods < 10 days immediately raised questions about their formation process and evolution [12]; the ubiquity of super-Earths [13], and the concept that $\sim 5\%$ of Sun-like stars host an Earth-sized planet with orbital periods of 200-400 days [14]; the idea that there is at least one planet on average per star in the Milky Way [15]; the discovery of planetary mass companions orbiting brown dwarfs [16, 17]; the dearth of giant planets with short periods in multiple planet systems [18] as well as the existence of giant planets on ultra wide orbits [19, 20]; and the correlations of planet properties with solar mass and metallicity [21].

1.1.2 Methods of Detecting Exoplanets

Exoplanets are hard to image directly because the planet's luminosity is many orders of magnitude smaller than the parent star and the separation of the two objects is very small from our perspective. The analogy is commonly made that this observation is equivalent to observing a firefly next to a search light. Fortunately, there are several methods to bypass or overcome this challenge, each exploiting different properties of the stellar system.

The transit method exploits the planet's angular size (and albedo) compared to the star. If the planet's orbital plane is edge-on from our perspective, then as the planet passes in front of the star, the brightness of the star decreases. The transit flux depth is related to the ratio of the areas of the planet and host star. Therefore, if the stellar radius is known, then the planet radius can be determined with this technique. This is the method employed by the Kepler Space Telescope [22], which is responsible to thousands of the known exoplanet candidates, and the SuperWASP (Wide Angle Search for Planets) survey [23]. TESS (Transiting Exoplanet Survey Satellite) [24], CHEOPS (CHaracterising ExOPlanets Satellite) [25], and PLATO (PLAnetary Transits and Oscillations of stars) [26] are three space observatories dedicated to transit photometry that will come online in the coming years.

The radial velocity method exploits the effect of a planet's mass on the star. As the planet orbits the barycenter of the star system, so too does the star. If

this reflex motion has a component along the line of sight then the spectral lines of the star will be periodically Doppler shifted from our perspective. The ratio of the velocities of the two objects is the reciprocal of their mass ratio allowing the exoplanet mass to be determined with this detection method. The radial velocity method was used for the very first exoplanet detection [27], where two planets were discovered around a pulsar, as well as the exo-Jupiter discovery around 51 Pegasi [8]. This method has been responsible for $\sim$14% of the exoplanet discoveries. There are a number of upcoming instruments that are slated to find Earth analogues: ESPRESSO (Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations) [28], EXPRES (The EXtreme PREcision Spectrograph) [29], and HARPS3 (High Accuracy Radial Velocity Planet Searcher) [30].

The gravitational microlensing method also exploits a planet's mass to infer its presence. When a planet happens to pass in front of a background star, for the short duration of the alignment, the exoplanet magnifies the star's flux. Exoplanets thousands of light-years away, those with moderately wide separations, very dim planets, and planets without a parent star (rogue planets) can be discovered this way [31]. OGLE (Optical Gravitational Lensing Experiment) and MOA (Microlensing Observations in Astrophysics) are two ongoing microlensing surveys [32]. Approximately 80 of the known exoplanets were discovered with microlensing.

Direct imaging, or high contrast imaging (used here interchangeably), exploits the luminosity of the planet to infer its existence, and the parent star's luminosity serves only as a hinderance. The light from the parent star must be controlled to a very high precision and removed with a coronagraph to reveal the underlying planet. There are several dedicated instruments and campaigns that have been operational for a number of years or are about to begin: operation P1640 (Project 1640) [33], MagAO (Magellan AO) [34], SPHERE (Spectro-Polarimetric High-Contrast Exoplanet Research) [35], GPI (Gemini Planet Imager) [36], SCExAO (Subaru Coronagraphic Extreme Adaptive Optics) [37].

Currently, only a handful of discoveries have been made with direct imaging. However, non-detections have been important in placing limits on the efficiency of

planet formation at wide separations [38], and as new instruments and techniques are developed the discovery productivity of direct imaging is expected to dramatically increase [39]. JWST (James Webb Space Telescope) [40] and WFIRST (Wide Field Infrared Survey Telescope) [41] are two future space telescopes that will be fitted with coronagraphs for high contrast imaging. PCS (Planetary Camera and Spectrograph) is an instrument on the ELT (Extremely Large Telescope) dedicated for direct imaging (previously known as EPICS) [42]. ELT-PCS and WFIRST both have the potential to image the first exo-Earths.

Figure 1.1 shows the known distribution of planets colored according to the detection method used. The properties of planets that each detection method is most sensitive to is immediately apparent. Direct imaging has extended the separation axis of this plot by many orders of magnitude (and has constrained the non-existance of planets beyond that region), but so far has not been able to probe planets below the mass of Jupiter at these separations. As the time base of radial velocities surveys increase, and the performance of direct imaging instruments increase, this area should gradually fill. This will help to constrain the planet formation mechanisms at a range of temperatures and chemical environments.

1.2 High Contrast Imaging

1.2.1 Introduction

High contrast imaging (HCI) is a powerful tool in understanding the architecture and formation history of planetary systems [43]: it does not require chance orbital inclination alignment or line of sight coincidences; it is sensitive to exoplanets beyond 10 AU; it directly measures photons from the exoplanet atmosphere for composition, clouds, thermal structure and variability; it directly reveals orbital information, and direct imaging of protoplantary disks provides information on the planet formation process and substellar evolution [44].

Some interesting examples of directly imaged exoplanets are shown in Figure 1.2.

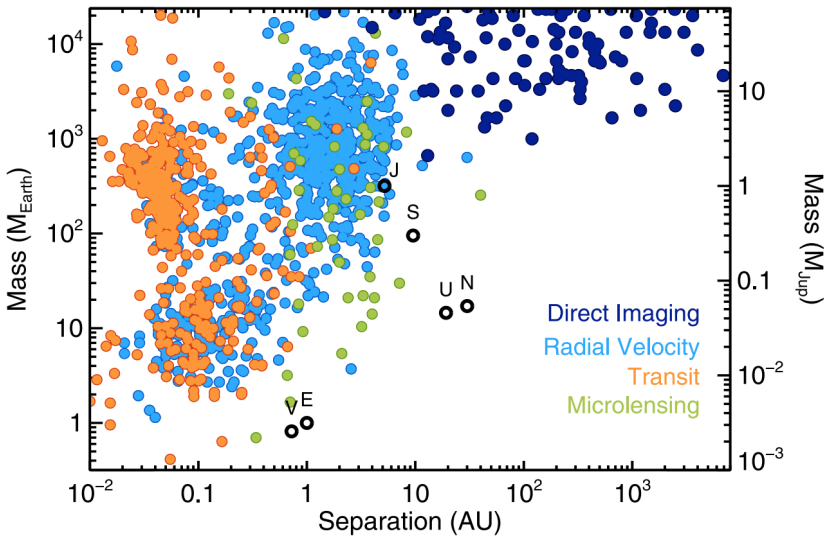

Figure 1.1: The distribution of the known exoplanets and substellar companions colored according to the discovery method. Transits and radial velocity have been high prolific. Direct imaging has had some success at wide separations. Probing the lower mass and wide separation region will help to constrain planet formation mechanisms. This image is from [38].

HR 8799 hosts a multiple planet system consisting of four known 5-10 $M_{\rm J}$ planets at orbital separations ranging from ∼20 to ∼75 AU [45, 46]. Beta Pictoris b is currently the closest exoplanet to its host star detected with direct imaging [47] and it has the shortest measured rotation period at 8.1 hours (measured using high resolution spectroscopy; [48]). Gliese 504 b is the coldest and oldest known directly imaged exoplanet at 240 °C and 100–500 Myr respectively [49]. Interestingly, HR 8799, 51 Eri [50] and HD 95086 [51, 52] all possess dust belts interior and exterior to the imaged planets [38], which is analogous to the asteroid belt and Kuiper belt that surround the gas giants in the Solar System.

The first planetary mass object detected with direct imaging was 2M1207 b, a young 5 $M_{\rm J}$ hot planet separated from the host brown dwarf by 40 AU [16]. This mass places it well below the Deuterium burning limit (13 $M_{\rm J}$) that traditionally distinguishes exoplanets from brown dwarfs [54], but the parent-planet mass ratio $q \approx 0.2$ and separation of 55 AU resemble binary star systems (the Sun-Jupiter mass ratio is $q \approx 10^{-3}$). Similarly, AB Pic b has a derived mass of approximately 13 $M_{\rm J}$

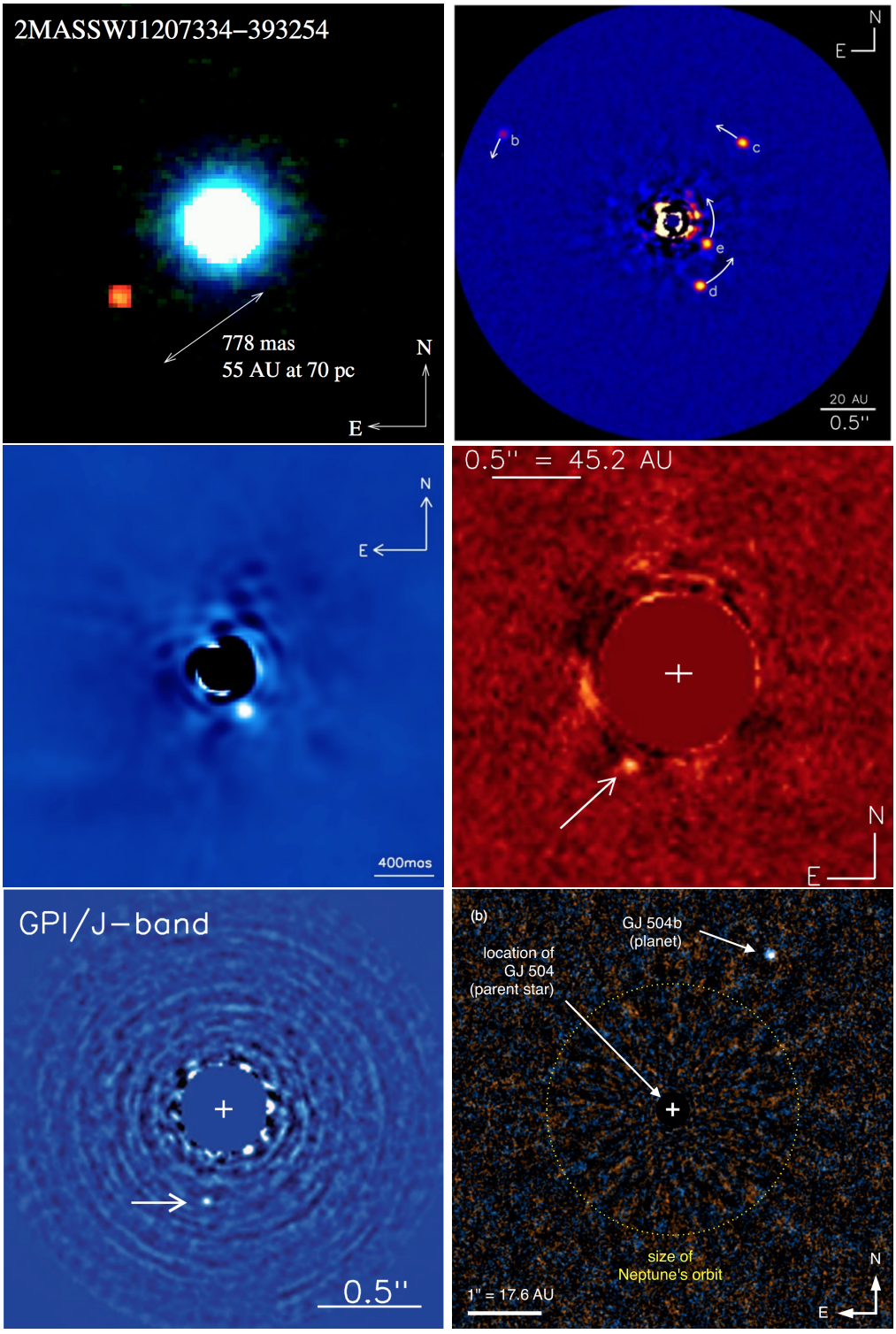

Figure 1.2: Interesting examples of directly images exoplanets. The systems are 2M1207 [16], HR 8799 [46], Beta pictoris [53], HD 95086 [51], 51 Eri [50], and GJ 504 [49]. The references contain the images shown.

placing it on the Deuterium burning limit, however the separation of $\sim$ 260 AU means that core accretion is unlikely to be the primary formation mechanism.

The results from the dozens of HCI surveys have shown that, while uncommon, massive planets can form at very wide separations (10-1000 AU) [55]. Bowler [38] predicts that the occurrence rate of 5–13 $M_{\rm J}$ exoplanets at 30–300 AU for young stars is $\sim$0.6%. These discoveries sparked a debate over the formation mechanisms of these exoplanets as core accretion simulations show that planets beyond 35 AU are unlikely to reach critical mass [56]. For example, the formation of such a rare system as HR 8799 is thought to have required a massive disk and very rapid build up of protoplanetary cores, rather than just in situ core accretion [57].

1.2.2 Important Parameters

Planets are separated from the host star by a very small angle from our perspective. The angular separation of a companion from its host star is

$$\theta = a(1+e)/d \tag{1.1}$$

where θ is the separation measured in arcseconds, a is the separation measured in AU, d is the distance to the system measured in parsecs (pc), and e is the eccentricity. This means that if observing Earth from HR 8799, which is roughly 40 pc away, at maximum Earth's separation will be 25 milliarcseconds (mas). A diffraction limited telescope observing at $\lambda = 1$ μm would have to have a primary diameter D of 8 m to have Earth be separated by $1\lambda/D$.

Contrast-ratio (or simply contrast) is the ratio of the companion (c) and star (s) brightnesses

$$C = \frac{f_c(\theta)}{f_s(\theta)} = \frac{\dot{N}_c(\theta)}{\dot{N}_s(\theta)}, \tag{1.2}$$

where f is the radiant flux measured in erg/(s cm$^2\mu$m) and $\dot{N}$ is the photon flux measured in photons/(s cm$^2\mu$m). The companion flux is the sum of the thermal and reflected fluxes. The Sun-Earth contrast at 1 μm is 10^{-9} [58]. (Contrast is also

referred to as dynamic range and is measured as the magnitude difference between the star and planet.)

The performance of a HCI observation is typically demonstrated with a 5σ contrast-separation curve, where the root mean square (RMS) luminosity in annuli at different radii are compared to the aperture photometry of the unocculted star. Chapter 5 details the process of constructing these curves (for example accounting for sample biases as small separation). Figure 1.3 shows the 5σ contrast curve for various HCI instruments and their coronagraphs. The performance of the current state-of-the art is approximately $C = 10^{-6}$ at 0.5 " separation after applying post-processing techniques.

Figure 1.4 shows the predicted exoplanet population created using MESS - a Monte Carlo simulation code that uses both statistical analysis and extrapolates the known exoplanets for its predictions [59]. The current performance of HCI corresponds to a handful of warm Jupiter detections on this diagram. To probe the underlying population of exoplanets, HCI systems will have to achieve better contrast ratios at smaller separations. A contrast ratio of 10^{-9} and a separation of 0.1 " should be sufficient for some exoplanet detections in reflected light.

Planets cool over time by radiating latent heat from their genesis. Whereas stars tend to get brighter as they evolve [61]. Therefore, it is more preferential in terms of contrast to observe younger systems. If Lithium burning, Deuterium burning, and clouds are neglected, then the luminosity of exoplanets scales as

$$L_{\text{bol}} \propto t^{-5/4} M^{5/2}, \tag{1.3}$$

where t is the object's age and M is its mass [38]. The steep mass-luminosity relationship means that a small gain in contrast in the brown dwarf regime results in a big change in the detectable mass, whereas the same amount of contrast gain in the planet regime results in a smaller change in the detectable mass.

Figure 1.5 shows the limiting masses of exoplanets as a function of contrast, age and distance [38]. When observing neighboring systems, the planets will have wider

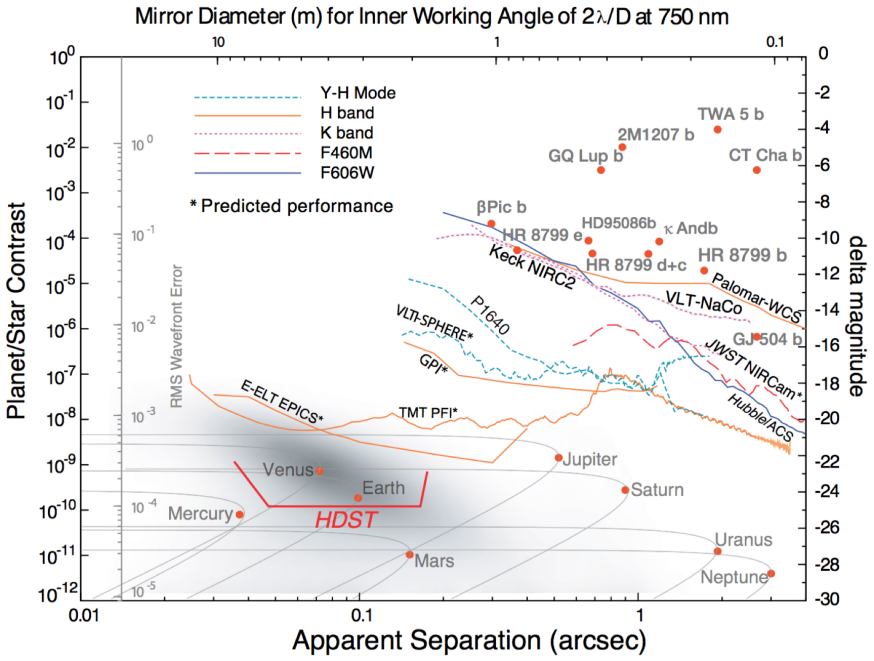

Figure 1.3: Contrast-separation curves for a range of HCI instruments and the values of some of the known exoplanets detected with high contrast imaging [60].

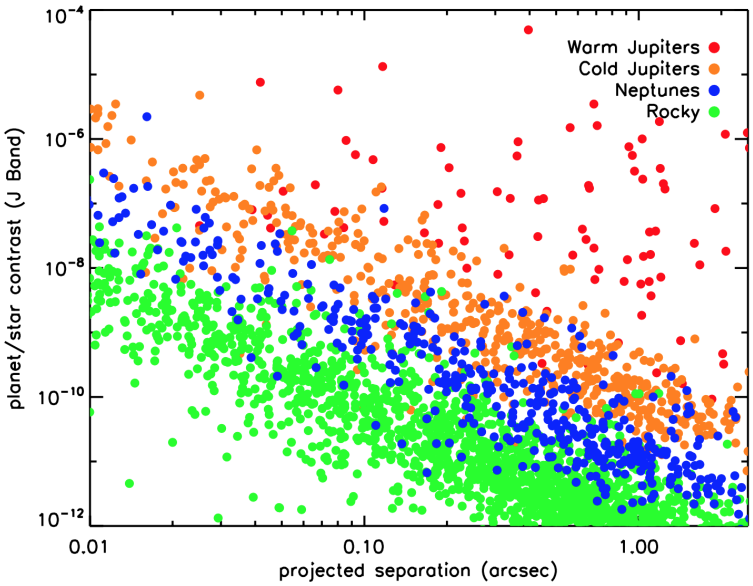

Figure 1.4: The contrast-separation values for different planet populations as simulated with MESS [59]. In order to tap into the potential pool of Earth-like planets requires sensitivity to smaller angles and lower masses.

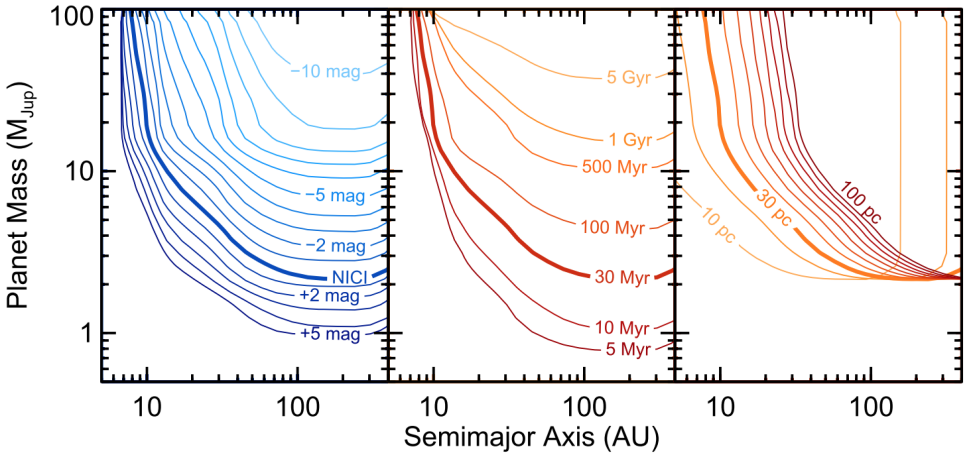

Figure 1.5: The mass sensitivity of exoplanets for different contrasts (left), ages (middle) distances (right). Direct imaging is most sensitive to young self-luminous exoplanets at close distances from the Solar System [38].

angular separations from our perspective, making the limiting mass and separation of the companion object more favourable in general.

1.2.3 Speckle Noise

If a telescope had perfect optics and there was no turbulent medium between the telescope aperture and an exoplanet system under investigation, then then parent star Point Spread Function (PSF) would be precisely known and could be subtracted from the image to reveal the location of the exoplanet. In reality, there will always be some residual aberrations from the atmosphere and in the optical train from polishing or reflectivity errors that create random wavefront errors - perturbations from the expected planar wavefront in phase and/or amplitude - that constantly vary throughout the duration of the observation. These wavefront errors create optical path differences in the beam that interfere at different locations in the focal plane, creating a speckle pattern that degrades the contrast [62]. The slowly varying errors mean the speckles are correlated (on a range of timescales depending on the origin), which means that the central limit theorem cannot be applied, and the exoplanet detection confidence level is degraded with respect to Gaussian noise [63]. The speckles also interfere with

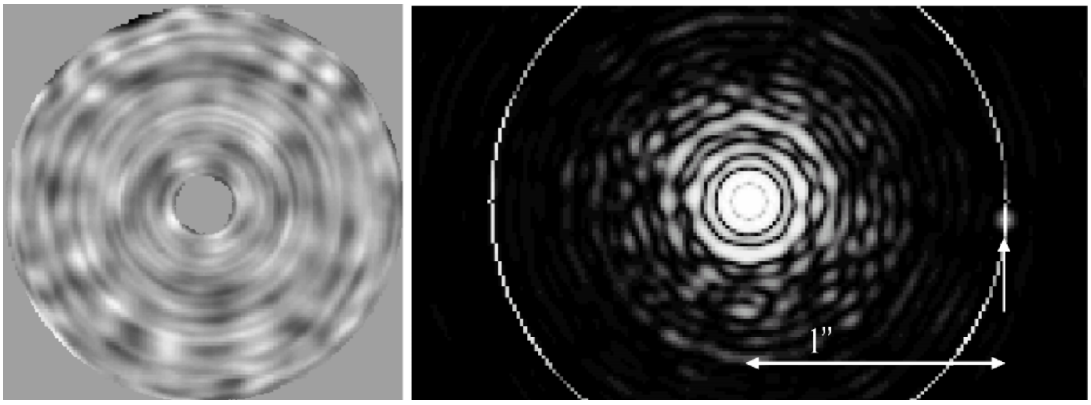

Figure 1.6: VLT mirror error map with wavefront error RMS of 43 nm (left) and the focal plane image (right) [66]. The arrow points to a binary companion, which looks indistinguishable from some of the speckles at close separations.

the deterministic diffraction pattern, which amplifies the speckles at certain parts in the image [64]. Furthermore, out-of-pupil plane optics lead to highly chromatic speckles as the wavefront evolves in the near field [65]. Figure 1.6 shows the aberrations on the primary mirror and the measured speckle pattern in a narrow band image of the focal plane.

The creation of speckles can be understood as follows. A perfect plane wavefront incident upon a circular aperture and focused on a detector will create an Airy function PSF. If instead the wavefront meeting the aperture contained a sinusoidal phase ripple, then the measured image would be a central Airy function as well as two fainter copies of the Airy function on either side of the central peak. Their distance to the center is determined by the spatial frequency of the sinusoid and the orientation of the speckle pair mirrors the orientation of the sinusoid. One can imagine a more elaborate wavefront containing lots spatial frequencies and orientations creating speckles distributed across the image. Amplitude perturbations also create speckle, but the phase relationship of each pair differs from phase-induced speckles (this difference in behavior explains why a deformable mirror cannot null all phase and amplitude aberration simultaneously).

There are various sources of wavefront error that have different characteristic timescales. These come in three categories: static, dynamic, and quasi-static. The

static aberrations are from polishing and figuring errors, as well as misalignments in the optics. Dynamic wavefront errors are the result of atmospheric turbulence and the resulting speckles have lifetimes on the order of ms. The quasi-static aberrations are the result of thermal fluctuations, mechanic deformations as the telescope tracks, and other slowly varying phenomena. These speckles can remain correlated on timescales between seconds and even hours [67].

What follows is a partial derivation of the intensity statistics in the image plane adapted from [68] and [69]. If we make the assumption that all of the aberration components can be represented in the entrance pupil, then the wavefront complex amplitude, represented in one spatial dimension x, can be expressed as

$$\begin{aligned} \Psi_1(x,t) =& [A_P + A_S(x) + a_d(x,t) + a_{qs}(x,t)]P(x) \\ =& [A(x) + a(x,t)]P(x), \end{aligned} \tag{1.4}$$

where A_P and A_S are deterministic terms correspond to the ideal plane wave and static aberrations, respectively; a_d and a_{qs} are random terms with zero mean corresponding to the dynamic and quasi-static aberrations, respectively; and P is the pupil function, where $\int P(x)dx = 1$.

The deterministic and random terms can be combined into A and a, respectively. Then, the complex amplitude in the image plane is the Fourier transform

$$\begin{aligned} \Psi_2(r,t) =& F\{\Phi_1(x,t)\} = F\{A(x)P(x)\} + F\{a(x,t)P(x)\} \\ =& C(r) + S(r,t). \end{aligned} \tag{1.5}$$

The deterministic term, $C(r)$, dictates the form of the PSF for long exposures. In the absence of static aberrations, it would be proportional to the Airy function (for a circular aperture). The random term, $S(r,t)$, has a variance that varies across the field. At all points in the image $S(r,t)$ will be zero mean circular Gaussian in the complex plane by virtue of the central limit theorem and the large number of independent a cells across the pupil [64, 68]. It follows that $\Psi_2(r,t)$ is also a circular Gaussian but centered on $C(r)$ i.e. $\langle \Psi_2(r,t) \rangle = C(r)$.

The intensity is the square modulus of the amplitude

$$|\Psi(r,t)|^2 = |C(r)|^2 + |S(r,t)|^2 + 2\mathrm{Re}\{C^*(r)S(r,t)\} \tag{1.6}$$

Since the time average of S is zero, the final term does not contribute to the mean intensity

$$I(r) = C(r)^2 + \left\langle |S(r)|^2 \right\rangle = I_C + I_S, \tag{1.7}$$

and the variance is

$$\sigma_I^2 = I_S^2 + 2I_C I_S, \tag{1.8}$$

as derived by Goodman [70] and Soummer et al. [68]. The final term describes speckle pinning, the phenomenon where speckle fluctuations are amplified by the deterministic contribution (Airy ring maxima and diffraction spikes for example) [64].

The Probability Density Function (PDF) for the intensity pattern is known as the modified Rician (MR) distribution

$$P(I) = \frac{1}{I_s} \exp\left(-\frac{I + I_c}{I_s}\right) J_1\left(2\frac{\sqrt{II_c}}{I_s}\right), \tag{1.9}$$

where J_1 is a modified Bessel function of the first kind [70] (for a derivation of equation 1.9 see [68]). The form of this function at different amounts of I_C is shown in Figure 1.7. For $I_C \ll I_S$ the distribution becomes the 'fully developed' speckle pattern with exponential statistics. This will be the regime at the Airy minima, far away from the core of the PSF, or a coronagraphic PSF. In the instance where $I_C \gg I_S$, the distribution follows a Gaussian.

The $P(I)$ along annuli also follow MR probability distributions since they all share the same ensemble mean (I_C) and same temporal fluctuations will apply to all points (if spiders and companions are masked; [63]). The less Gaussian the intensity distribution, the lower the confidence level

$$\alpha(d) = \int_{-d}^{d} P'(I) dI \tag{1.10}$$

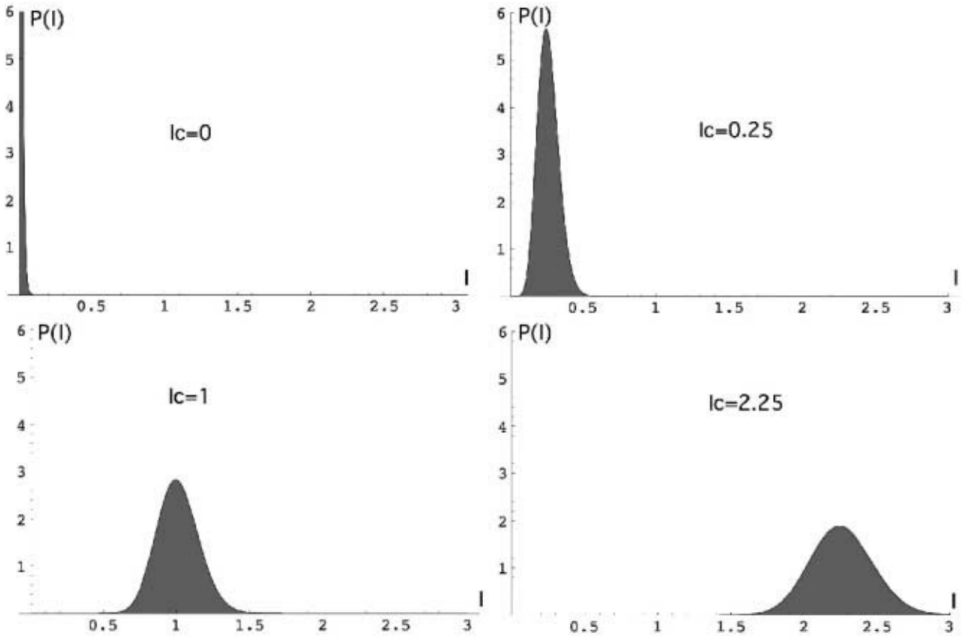

Figure 1.7: The intensity statistics for different values of I_C and $I_S = 0.1$. As I_C increase the distribution broadens in accordance with the speckle pinning phenomenon. This figure is from [64].

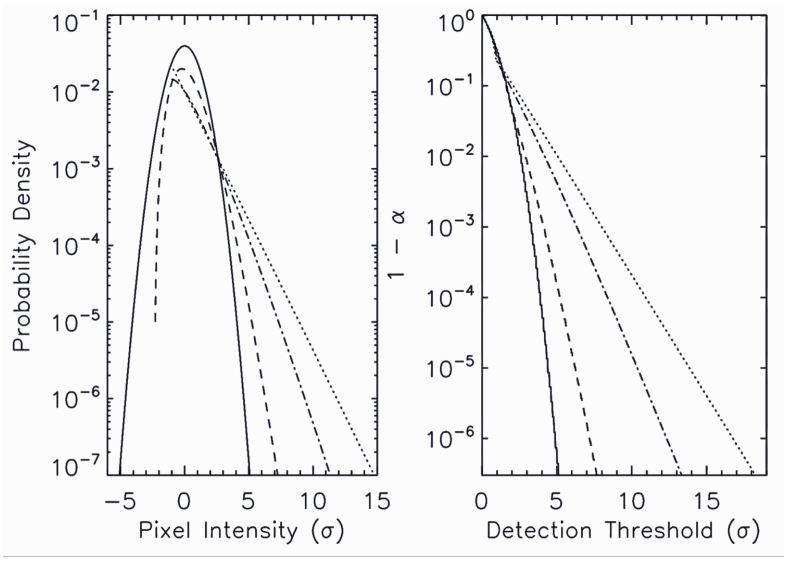

Figure 1.8: Left: the PDF for Gaussian (solid line) and MR distributions at $I_C/I_S = 10$ (dashed line), $I_C/I_S = 1$ (dot-dashed line), $I_C/I_S = 0.1$ (dotted line). Right: as I_C/I_S decreases, the confidence level at the same detection threshold also decreases. This figure is originally from [63].

where P' is the mean-subtracted PDF. Figure 1.8 shows the distributions of α for varying degrees of Gaussicity (how accurately a distribution can be described by the Gaussian function). To achieve a $1-\alpha = 10^{-6}$ requires a detection threshold increase of several orders of magnitude for MR noise compared to the Gaussian case.

1.2.4 Extreme Adaptive Optics

High performance AO is essential in minimizing the amount of wavefront error and the associated speckle noise and attaining the sufficient levels of contrast. The basic configuration of a standard AO system is shown in Figure 1.9. There are 3 key components: the wavefront sensor (WFS) that measures the wavefront errors, the real time computer (RTC) that calculates the required offsets to cancel out those wavefront errors, then the tip tilt mirror (TTM) and deformable mirror (DM) that apply the calculated offsets to the wavefront.

In the visible and near-IR there are no detectors that can measure the phase. Therefore WFS have to measure phase indirectly by changing the phase of the light

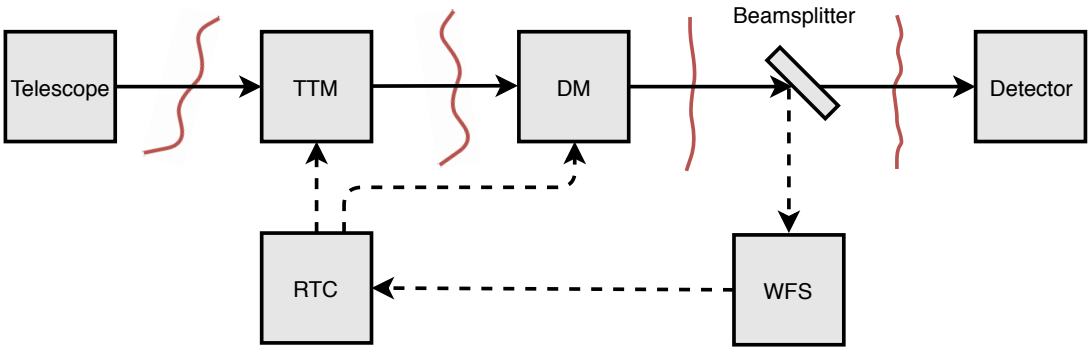

Figure 1.9: The basic implementation of adaptive optics. An initially aberrated wavefront has a correction applied to it using a tip tilt mirror and deformable mirror. Some of the light is separated, measured with a wavefront sensor, and the appropriate correction calculated using a real time computer.

and monitoring the intensity variations. These WFS have to operate at KHz frame rates and measure a large range of aberration modes without amplifying the measurement noise. In HCI this is commonly achieved with a Shack Hartmann WFS or a Pyramid WFS.

The correction is performed in two stages: a tip tilt mirror for the low order modes (for example wind vibration) and a deformable mirror that contains thousands of actuators that are fast and have high piston accuracy for the high order modes. The AO system essentially high pass filters the spatial frequencies in the pupil plane, which results in a correction in the focal plane out to a certain radius.

The Fried parameter, r_0, characterizes the atmospheric correlation length. It is defined as the characteristic length where the phase variance $\sigma_\phi^2 \approx 1$ radian. The angular resolution of the telescope (without AO) then becomes λ/r_0. The Greenwood time delay characterizes performance of an AO system: $\tau_0 = 0.314 r_0/v$ where v is the weight mean wind speed of the turbulent features; τ_0 is several milliseconds at visible wavelengths.

The Strehl Ratio, SR, quantifies the quality of the formed optical image

$$SR = \frac{I_s}{I_s^*} \tag{1.11}$$

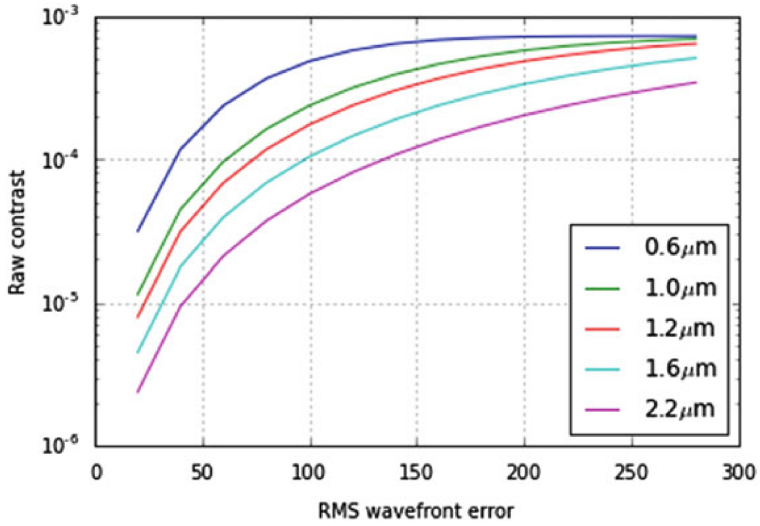

Figure 1.10: The theoretical contrast performance of an xAO HCI telescope for varying degrees of wavefront error correction. This image is from [74]

where I_s and I_s^* are the achieved central aperture intensity and theoretical maximum. The Marećhal expression provides an estimate for SR [71]

$$SR = \exp(-\sigma_\phi^2) \tag{1.12}$$

where $\sigma_\phi = \frac{2\pi\delta}{\lambda}$ and δ is the rms wavefront error. While standard AO systems typically obtain 40-60%, which corresponds to an rms error of $\sim$ 100 nm at 1 μm, whereas xAO systems frequently achieve SR $>\sim$95% ($<$ 25nm rms error; albeit over a small field of view compared to other AO types) [72]. An SR = 95% means that 5% of the starlight is distributed outside of the PSF core in the speckle pattern.

Serabyn et al. [73] provide an approximation for the contrast ratio of a high-order DM

$$C = \frac{1 - SR}{N_{act}} \tag{1.13}$$

A system observing at 1 μm with 10 nm rms wavefront errors and $N_{\rm act} = 3600$ (Palm 3000) is capable of reaching raw contrast $C = 10^{-6}$. Figure 1.10 contains a few curves demonstrating this relationship.

1.2.5 Coronagraphy

The purpose of a coronagraph is to the suppress as much of the diffracted light from an on axis source (the parent star) as possible, while retaining as much of the off-axis light (from a companion) as possible. Bernard Lyot devised the first coronagraph to observe the Stellar corona [75], and the architecture is still utilized for modern direct imaging of exoplanets. The Lyot coronagraph contains a small occulter in the focal plane that blocks the central peak. The light is then recollimated with either a lens or a mirror, which is the same as taking a Fourier transform of the complex amplitude. Since the occulter blocked the lowest frequency components, the remaining light is concentrated around the edges of the pupil plane. This ring is then blocked with an undersized pupil called a Lyot stop. An image is then recreated with another lens or mirror.

Apodized Lyot coronagraphs (ALC) are an extension of this concept. The wings of a PSF are the result of the hard edge of an entrance pupil. These wings amplify speckles and mask companions. Instead, an Apodized Pupil Lyot Coronagraph has a "softened" pupil that has graded transmission towards the edges that removes the PSF wings (at the expense of throughput). These are the type of coronagraph used in GPI for example [76]. Band-limited coronagraphs also manage the diffraction effects, but the occulter has a graded transmission instead of the pupil [77].

Another way to build on the Lyot concept is to use a phase mask instead of an amplitude mask [78]. For example a four-quadrant phase mask (4QPM) contains two sections that introduce no phase shift and 2 quadrants that introduce π phase shift [79]. The dephased beams destructively interfere with the beams that were unchanged. The diffracted starlight is then blocked with a Lyot stop. SPHERE has the option of a 4QPM or ALC [80]. The optical vortex coronagraph is an extension of the 4QPM where the mask introduces an azimuthally varying phase shift [81]. These coronagraphs offer high broadband performance, small IWA, and high throughput [82].

Coronagraphs can only remove the static intensity component from diffraction,

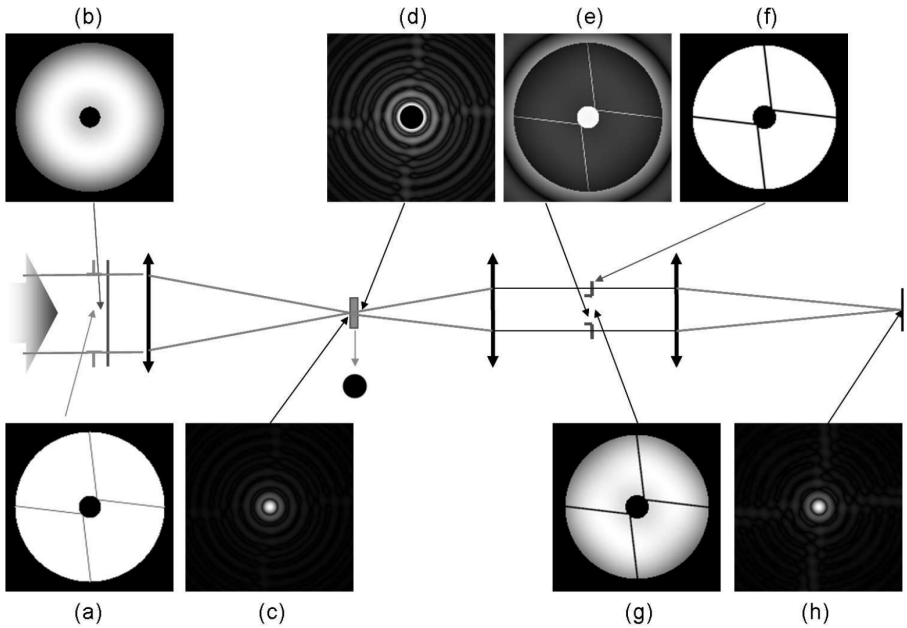

Figure 1.11: A schematic of the Apodized Lyot Coronagraph design for SPHERE [83]. The collimated light passes through an apodizer (a & b). The initially collimated wavefront is then focused and the resulting PSF has reduced intensity in its wings (c). An occulting mask removes most of the intensity of the star light (d). After recollimating the wavefront, the residual star light is concentrated at the edges of the beam (e). An undersized aperture (f) therefore blocks the majority of the remaining star light. The resulting intensity pattern in the pupil plane is shown in g. Focusing the light yields the PSF with minimal star light and maximal off-axis light (h). Each panel has a different intensity scaling.

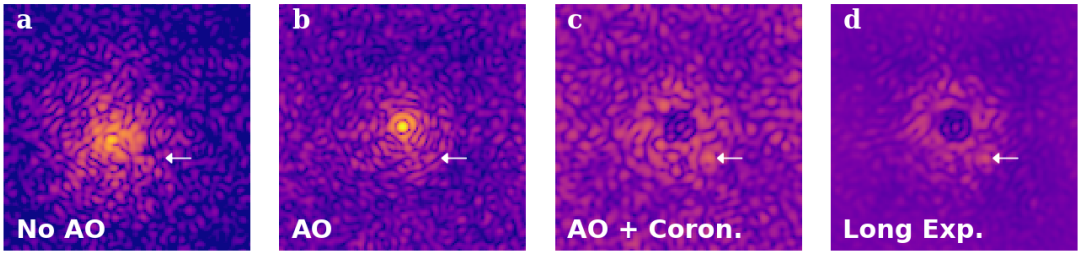

Figure 1.12: Demonstration of an observation of a system with moderate contrast requirements. The arrow points to location of a companion with brightness 1000× dimmer than the unocculted star light. Coronagraphs are only effective for the deterministic intensity contribution. Integrating for over a minute does not enhance the contrast appreciably as the quasi-static speckles coherently add. Each panel has different scaling.

and cannot remove speckles from wavefront errors. This is helpful however, because it reduces the coherent amplification of speckle variance, as well as photon noise. It also helps with dynamic range issues in detectors, which is very relevant to MKIDs. Soummer et al. [68] showed that when using a coronagraph the complex amplitude at the final focal plane $C(r)$ in equation 1.5 is simply derived by substituting C with a $\tilde{C}(r)$ term $\Psi_4(r) = \tilde{C}(r) + S(r)$ The presence of a coronagraph has the effect of removing the mean from the deterministic term C. Whereas the random term is left mostly unaffected.

Figure 1.12 demonstrates the limits of long exposure observation that contains a coronagraph. The residual speckle noise is coherent and little gain in contrast-ratio is made by integrating for a long time.

1.2.6 Post-AO Wavefront Sensing

Even after the high order correction of the xAO a quasi-static speckle pattern remains. These speckles are the result of so-called "non-common-path aberrations" (NCPAs), aberrations that exist downstream of the xAO that cannot be corrected. Additionally, some wavefront errors will remain from an imperfect correction by the xAO from: chromatic effects, insufficient frame rates, and the time lag on the servo correction.

One approach to suppress these speckles is to remove them optically through de-

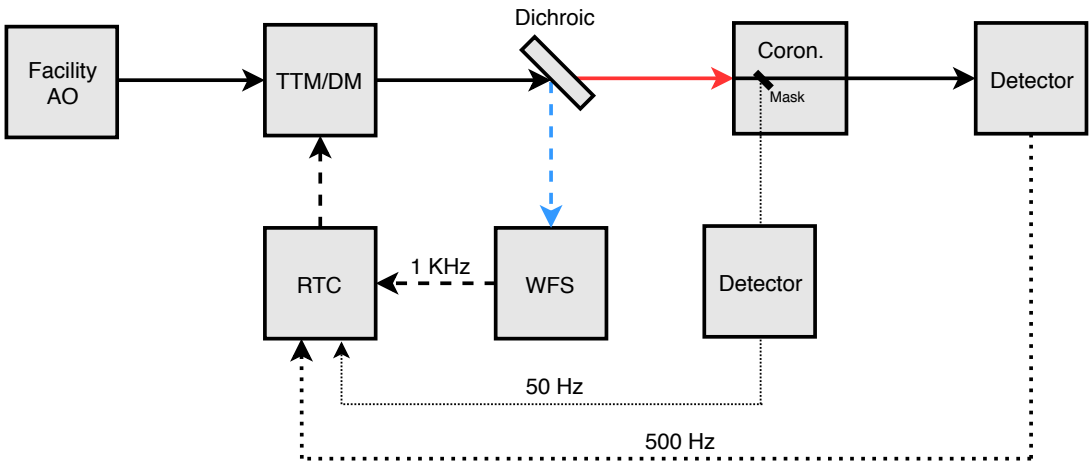

Figure 1.13: A schematic of one implementation of a xAO HCI system. After the low order correction from a facility AO, the light is corrected with an xAO that removes high order modes. The WFS using the rejected light from the coronagraph is a second feed-back loop to control low order aberations that would limit the performance of the coronagraph. The remaining NCPAs are addressed with a FPWFS.

structive interference. This also has the effect of enhancing the sensitivity by reducing the photon noise (unlike post-processing techniques as described in the following section).

Many high performance, low IWA coronagraphs are extremely sensitive to pointing errors [84]. To maintain the coronagraph performance a low order wavefront sensor measures the aberrations just prior to the coronagraph [85]. Sometimes the occulting mask itself is used as the optic to pick off the light for the WFS as shown in Figure 1.13.

A focal plane wavefront sensor, FPWFS, measures the aberrations with the science camera. In order to measure the phase (from an intensity sensor) an interference pattern is created on the science camera. This can either be produced with phase modulations from the xAO DM on the observed image [86], or interfering the observed wavefront with a separate laser source [87]. Once the phase is measured, further modulations from the DM can address the phase aberrations over the full field [88, 86, 89, 90, 91].

A HCI system will contain both phase and amplitude induced speckles. The

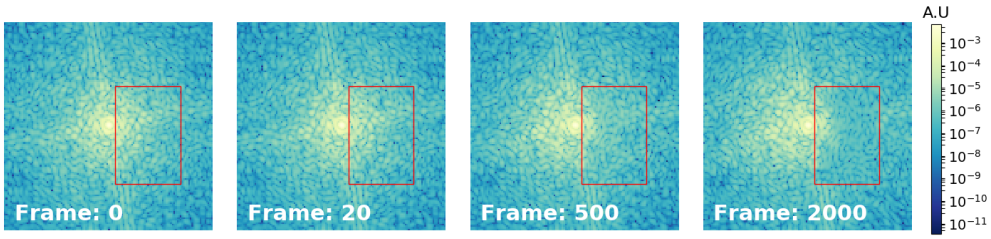

Figure 1.14: Nulling of NCPA amplitude errors with phase-shifting interferometry.

phase induced speckles are much larger in amplitude and dominate the background noise. However, once corrected, the amplitude-induced speckles remain that limit the achievable contrast. It is impossible to null all of the phase and amplitude speckles over the full field through corrections to the wavefront's phase alone. Instead, speckles in one half of the image are nulled at the expense of possibly increasing the intensity of the conjugate speckles in the other half. This is achieved by identifying the speckles based on their intensity and proximity to other speckles, and then generating a pattern on the DM that creates speckles at the same location and amplitude (along with their conjugate on the other side of the image) but with opposite phase to the original speckles. Uncertainty in the speckle phase and amplitude measurements, for example, mean that this is normally an iterative process. Figure 1.14 demonstrates FPWFSing using the active nulling technique outlined in [91].

1.2.7 Observing Strategies and Post Processing

The residual speckle pattern can be identified and subtracted with differential imaging. This has the advantage of removing the correlated component of the speckle noise transforming the MR statistics into Gaussian statistics and increasing the confidence level [92]. We will not go into too much detail about the implementation of differential imaging, as this is done in Chapter 5. In short, differential imaging exploits fixed property of the telescope to generate a reference PSF for the host star, which is then subtracted from the measured signal to reveal underlying companions. For example, reference differential imaging, RDI, uses two sets of observations. One observation is of the solar system under investigation, the second observation is of a

different star that contains no companion objects. The speckle pattern is assumed to remain constant between the two observations and the only difference in the images will be due to the presence of a celestial companion object.

Because most of the quasi-static aberrations are fixed to the pupil, HCI is often performed in 'pupil-stabilized' mode. This is done by turning off the derotator of the telescope on altitude-azimuth-mount telescopes. This means that the speckle pattern remain stationary on the detector whereas celestial objects rotate around the center. In angular differential imaging, ADI, the reference PSF is generated by derotating and combining a series of exposures [93].

Multiple channel spectral differential imaging, here called SDI, uses the wavelength dependence of the speckle pattern to discriminate it from celestial companions [94]. It requires the use of an integral field spectrograph - an instrument that measures spectra at every location in the image, yielding an observation datacube (two spatial axes and one wavelength axis). The speckle pattern linearly scales with wavelength whereas the celestial object will stay in the same location. The reference PSF can therefore be created by scaling the frames of the datacube. Furthermore, hot Jupiters contain a strong methane absorption feature around 1.6 μm. If a xAO instrument has spectroscopic capabilities, this feature can be used as evidence of the true nature of the detected point source [95]. Polarisation Differential Imaging, PDI, is where polarized reflected exoplanet light is used to discriminate against the unpolarized direct starlight [96].

The various differential imaging methods have different disadvantages. RDI suffers from the fact that the speckle pattern can evolve between the two observations and therefore the reference signal quickly decorrelates from the observation. ADI and SDI are inherently less sensitive to speckles at smaller separations, where more and smaller planets reside. PDI remains effective at small separations but is often limited by self-polarisation from the telescope and instrument optics.

There are a number of advanced algorithms for implementing differential imaging such as LOCI (locally optimized combination of images) that optimizes noise attenuation using a least squares criterion [97], KLIP (Karhunen-Loéve Image Processing)

that uses principle component analysis [98], ANDROMEDA (ANgular Differential OptiMal Exoplanet Detection Algorithm) [99] and LLSG (Local Low-rank plus Sparse Gaussian-noise decomposition) [100].

In addition to differential imaging, there are techniques that use photon statistics to discriminate speckles from celestial objects, thereby increasing the contrast performance. These techniques have the advantage of remaining effective at small separations, however fast detectors are required. Stochastic Speckle Discrimination (SSD) exploits the characteristic intensity distribution of companion planets (which have a negative skew), compared to speckles (which have a positive skew), to identify each point source [66]. Dark Speckle Imaging exploits the fact that pixels containing a modulating speckles can reach zero counts whereas those pixels that also contain a sources are prohibited from reach such levels of intensity [101]. This way sources obscured by the speckle noise can be identified.

1.3 Detectors for High Contrast Imaging

Photodetectors create an electrical signal in response to absorbed photons. Conventionally, this is accomplished with doped semiconductors and the photoelectric effect. Recently, superconductor-based detectors have been developed, which have some advantages over semiconductor detectors.

To compare the performance of different detectors there are several figures of merit. Quantum efficiency is a measure of the effectiveness of the sensor for creating an electrical response for each photon. It is a measure of what fraction of the measurable photons produce a signal. Pixel count is the total number of pixels in the device, whereas pixel yield is the fraction of those pixels that is capable of photon detection. Responsivity is the amount of current produced for a given amount of incident light power. Dark current is an electrical signal in the absence of light on the detector. It occurs because phonons excite electrons into the conduction band. Readout noise is caused by the thermal motion of electrons in the transistor amplifier. It can be mitigated with extreme cooling or reducing the measurement bandwidth.

1.3.1 Semiconductor Detectors

The first semiconductor detectors were based on charge coupled devices (CCDs) [102, 103]. Photoelectrons are generated in a silicon substrate and captured with gate electrodes that form a capacitive depletion region. After an exposure, the electric charge in each pixel is read out by sequentially transferring the contents of each capacitor to its neighbour by actuating the electrodes as a shift register. Once the charge has been amplified and converted to a voltage, it is compared to a reference signal in a technique known as correlated double sampling (CDS), which helps mitigate high frequency read out noise [104]. By taking multiple measurements of each pixel the low frequency readout noise can be reduced to achieve a readout noise of $\sim 10^{-2}$ e^-/pix, and by cooling down to 140 K the dark current can be reduced to 10^{-3} e^-/pix/day [105].

Another approach is to use a p-n junction to detect the light (a photodiode). These are the basis of CMOS (complementary metal-oxide-semiconductor) detector or active pixel sensors. At the interface of the p-n junction exists the so-called 'depletion region', an area where most of the atoms have lost their mobile charge carriers (electrons or holes) to one of the junctions. An absorbed photon with sufficient energy will make an electron-hole pair in the depletion region via the photoelectric effect. The charge carriers are swept towards the corresponding junction by the intrinsic electrical field across depletion region. In photovoltaic detectors, no external bias is applied and the voltage build up is the measured electric signal. In photoconductors, the junction is shorted so a measured current can flow. Photoconductors often reverse biased, which decreases the response time, at the expense of higher noise (dark current). Non-destructive multiple reads are also possible that further reduce the readout noise

The band gap of silicon is 1.26 eV, which means that photons of $1\mu m$ have insufficient energy to liberate electrons. To observe in the infrared regime for HCI requires alternative materials such as Mercury Cadmium Telluride (HgCdTe). In HgCdTe photodiodes, by tuning the amount of Hg, the band gap can be tailored [106]. In-

frared arrays differ from CCDs in that individual pixels are directly accessed, which makes them complex and expensive. HAWAII-2RG (H2RG) Avalanch Photodiodes (APDs) sensors achieve $\sim 10^{-2}$ e^-/pix/s dark current at 120 K, 10–20 e^-/pix readout noise per CDS read and QE $\geq 70\%$ at some wavelengths [107]. These detectors are used on SPHERE and GPI [108, 109].

1.3.2 Superconductor Detectors

Superconductors have zero DC resistance when cooled below their characteristic critical temperature. Superconductor detectors exploit the unique properties of superconducting materials to enable new modes of detection. Often at operating temperatures of below 1 K, these devices not only benefit from the reduction in background thermal noise but are more sensitive to the energy of the incident photons than in semiconductors. These devices therefore have a higher responsivity allowing for single photon detection and inherent energy resolution at ultraviolet optical and near-infrared (UVOIR) wavelengths.

As a material transitions to the superconducting state, the resistance rapidly decreases. By operating at superconducting phase transition temperature, transition-edge sensors, TES, are able to measure the small amount of heat incident on the superconductor through the sensitive changes to the resistance [110]. Superconducting quantum interference devices (SQUID) current amplifiers are used to readout the signal. TES have particularly found use in the sub-millimetre regime [111] and will play a role in cosmic microwave background satellite observations [112, 113].

Pair breaking detectors operate similar to the semiconductor detectors described above in that photons with sufficient energy liberate charge carriers that creates a measured electrical signal. In superconductors however, the band gap is many thousands of times smaller (see Chapter 2 for more detail) so the responsivity of the detector is much higher. Superconducting Tunnel Junctions, STJs, comprise of a superconductor-insulator-superconductor junction biased with a DC voltage [114]. An absorbed photon liberates charge carriers, which quantum tunnel through the junction producing a current in the external circuit.

Cryogenic detectors have been playing an increasingly significant role in astrophysics particularly in the submillimeter regime [115], however even with all the advantages listed above they have not yet replaced semiconductors in the ideal instrument design the same way CCDs replaced photographic plates in the 1980s. The primary limitation of cryogenic detectors is the difficulty to produce large pixel arrays. Since small groups of pixels require their own cryocooled readout system (with preamplifiers and wiring), the system quickly becomes very complex, technically challenging and expensive to manufacture ([116]; [117]). This has limited the efficacy of superconductor detectors in high contrast imaging.

1.3.3 Microwave Kinetic Inductance Detectors

The Microwave Kinetic Inductance Detector (MKID) is a pair breaking detector that solves the problem of producing large arrays by using frequency domain multiplexing (FDM) to read out many pixels with a single cable using room temperature electronics [118]. Instead of using an SIS junction to filter the charge carriers, which is challenging to fabricate, single layer microwave resonator circuits are used instead. The first MKID camera (ARCONS), when commissioned in 2011, was the largest non-dispersive optical/near-infrared integral field spectrograph fielded by a factor of ten ([119, 120, 121]). The current generation of UVOIR MKID devices are up to 20,000 pixels [122], and larger arrays are planned [123]. The readout scheme of MKIDs means that they do not experience dark current or readout noise like semiconductor detectors, and photons are registered instantaneously facilitating high effective frame rates and high time resolution. See Chapter 2 for a full explanation of the operating principle of MKID systems.

MKIDs therefore offer a number of unique properties that make them theoretically ideal for HCI: single photon counting, inherent energy resolution, high time resolution, no read noise or dark current, and large pixel formats. The spectral resolution removes the requirement for IFS optics and associated aberrations when performing SDI. The high frame rate means that atmospheric speckles can be captured before they decorrelate to enhance the overall contrast ratio. Additionally, the high frame

rate allows for two methods of speckle suppression that remain effective at small separations: 1) FPWFS and 2) Photon timing statistics-based discrimination techniques such as SSD and DSI.

While new instruments, such as DARKNESS and MEC [124], will soon be able to validate the transformative benefits of MKIDs on-sky, they are limited to a single configuration and cannot quantify the performance of an MKID instrument with array parameters beyond the current specificiation. MKIDs currently suffer from relatively low pixel yield (75-90%) from resonator collisions, large variations in pixel performance from fabrication uncertainties, and low maximum count rates ($\sim$2500 photons/pix/s) from the recombination time of the Cooper-pairs [125]. It is not clear which properties require the most attention to optimize the performance of MKIDs for future instruments. One goal of this thesis is to address those questions.

1.4 Thesis Motivation

This thesis aims to provide a pragmatic demonstration of the potential of MKID instruments for xAO HCI systems as well as demonstrate some of the developments to this end. This will primarily be accomplished with a simulator that I constructed for this purpose.

This thesis is structured as follows: first I will introduce the operating principle of MKIDs. Then, I will discuss my contributions to the digital readout system, and the data analysis pipeline. Then, I will outline a package I developed to simulate HCI observations with MKIDs. I will describe some on-sky observation data taken with ARCONS of planetary nebular NGC 6751 that I reduced and analyzed, before concluding with some predictions using the MKID HCI simulation package.

Chapter 2

Microwave Kinetic Inductance Detectors

In this chapter we outline the operating principle of MKID pixels, we explain how arrays of thousands of MKIDs can be readout simultaneously. We provide examples of various instruments that contain MKID arrays including those optimized for HCI. We also discuss how certain array parameters can effect a system's performance for HCI.

2.1 Microwave Resonators

An MKID device is made from a film of superconducting material patterned into an array of microwave resonators using lithography. Each microwave resonator is an inductor and a capacitor in parallel (parallel LC circuit) coupled to a transmission line for readout. A schematic of a microwave resonator is given in Figure 2.1a. These microwave resonators can either be a lumped-element resonator, which consists of an interdigitated capacitor and a meandered inductor [127], or a transmission-line resonator, which consists of a quarter length of transmission line typically shorted at one end. See [128] for a review of superconducting microwave resonators. Modern ultraviolet optical and near-infrared (UVOIR) resonators use lumped element designs.

Each resonator has a characteristic frequency $\omega_0 = 1/\sqrt{LC}$ at which a passing

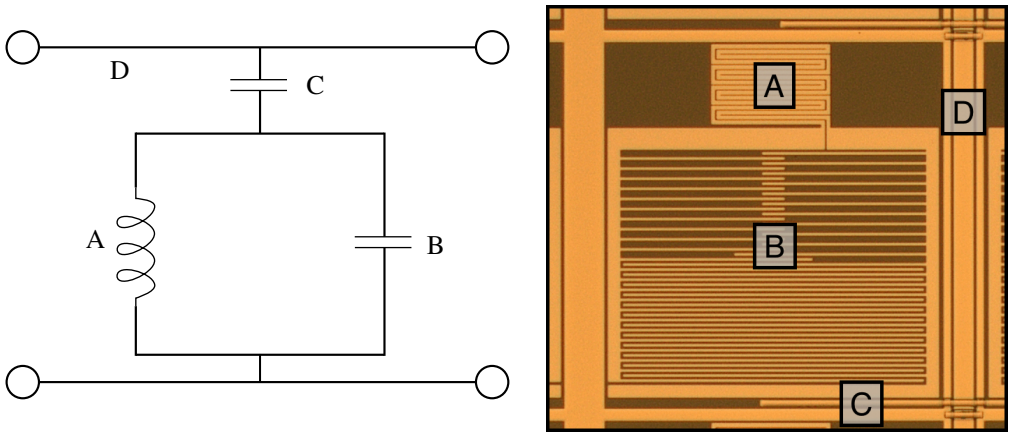

Figure 2.1: left: A schematic of an LC resonator circuit coupled to a feed-line for readout. right: Microscope image of a microwave resonator comprising an MKID pixel [126]. These resonators are lumped element, meaning they have a capacitive and an inductive region.

electric signal will form standing waves in the resonator. The complex transmission $S_{21}(\omega) = I(\omega) + iQ(\omega)$, where I and Q are the in-phase and quadrature components, is a measure of the gain in forward voltage across an electrical component. When a probe tone is swept in frequency (ω) across the resonance, the scalar transmission $|S_{21}|$ is minimized at the resonant frequency and is mostly unattenuated off-resonance. Figure 2.2 shows the transmission spectrum of a single resonator.

The figure of merit of resonators is the quality factor $Q_\mathrm{r} = \omega_\mathrm{r}/\Delta\omega$ where $\Delta\omega$ is the FWHM of the $|S_{21}|$ profile. A higher Q_r means that the resonator loses less energy with each oscillation. Q_r can be separated into two components according to

$$\frac{1}{Q_r} = \frac{1}{Q_c} + \frac{1}{Q_i} \tag{2.1}$$

where Q_c is the coupling quality factor, which is tuned with the length of the coupling bar and Q_i is the internal quality factor which accounts for the other loss channels and is less easy to control with the design of the resonators. Current UVOIR MKID resonators have Q_r in the range of $10^5 - 10^6$ when measured at $T/T_\mathrm{c} = 0.1$, where $T_\mathrm{c} \sim 1K$ is the superconductor critical temperature (described in section 2.2) [129]. For reference, the Q_r of a tuning fork is around 10^3.

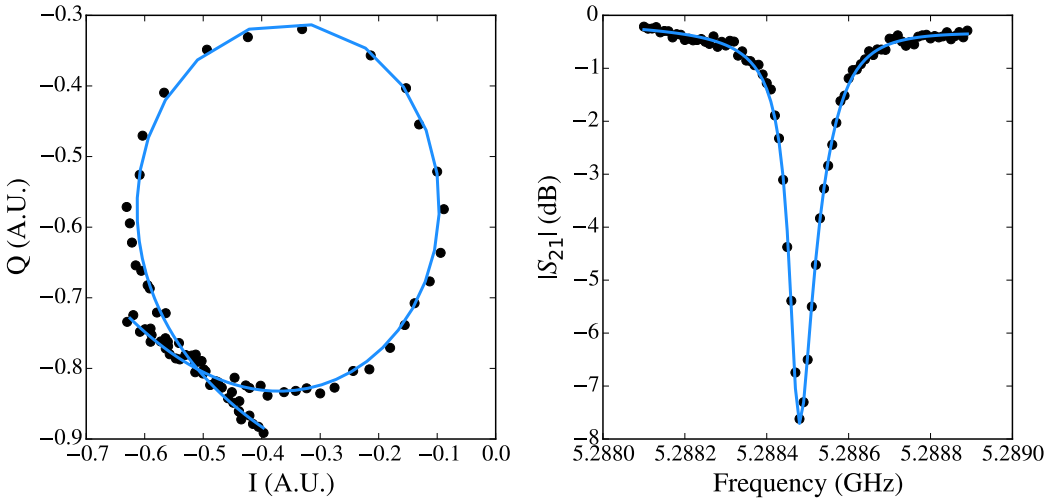

Figure 2.2: left: In the IQ plane a resonator forms a so-called resonance loop, where the resonant frequency is closest to the $I, Q = (0, 0)$ coordinate. Here I and Q are measured in arbitrary units. right: The transmission magnitude is a trough with a minimum at the resonant frequency. Both profiles are fit using equation 2.4

The complex transmission of an ideal resonator capacitively coupled to a feedline follows a Lorentzian distribution

$$S_{21}(x) = g(x)e^{i\phi(x)}\left[1 - \frac{Q_r}{Q_c}\frac{1}{1 + 2iQ_r x}\right], \tag{2.2}$$

where x is sampling frequency ω measured relative to, and scaled by, the resonant frequency at the lowest power sample ω_{r}

$$x = \frac{\omega - \omega_{\mathrm{r}}}{\omega_{\mathrm{r}}}, \tag{2.3}$$

[128]. The gain $g(x) = g_0 + g_1 x$ and phase $\phi(x) = \phi_0 + \phi_1 x$ (coefficients of Equation 2.2) account for scaling and orientation of the resonance loop due to the length of the transmission line [130].

Often there is an impedance mismatch between the resonator and feed-line that changes the S_{21} profile from the symmetrical Lorentzian. To derive this equation, Q_c is substituted with its complex form $\tilde{Q}_c$ where $\mathrm{Re}\left\{\tilde{Q}_c^{-1}\right\} = Q_c^{-1}$ and $\mathrm{Im}\left\{\tilde{Q}_c^{-1}\right\} = \delta\omega$ where $\delta\omega$ is a small frequency shift of the resonance [131]. The S_{21} profile then

becomes

$$S_{21}(x) = g(x)e^{i\phi(x)}\left[1 - \frac{\frac{Q_\mathrm{r}}{Q_\mathrm{c}}\left(1 + 2iQ_\mathrm{r}\frac{\delta\omega}{\omega}\right)}{1 + 2iQ_\mathrm{r}x}\right]. \tag{2.4}$$

The full derivation can be found in Khalil et al. [132].

Resonators exhibit a nonlinear response as a function of driving current, which can be attributed to quasiparticle heating by readout phonons [133, 134], or an intrinsic property of the kinetic inductance of superconductors at high current [135]. Given sufficient power, the resonator bifurcates into two quasi-stable states, which manifest as a discontinuity close to the resonant frequency. To account for the distortion of the S_{21} profile in the nonlinear regime, the fractional detuning becomes

$$x_p = x + \frac{a}{1 + 4Q_\mathrm{r}^2 x_p^2}, \tag{2.5}$$

where a is the nonlinearity parameter

$$a = \frac{2Q_\mathrm{r}^3 P_\mathrm{r}}{Q_\mathrm{c}\omega_\mathrm{r}E_*}, \tag{2.6}$$

where P_r is the readout power and E_* is the scaling energy from the nonlinearity [135].

The effect of varying several of the parameters of Equation 2.4 on the profile of $|S_{21}|$ can be observed in Figure 2.3.

2.2 Kinetic Inductance

When a superconducting material is cooled below its critical temperature T_c, its DC electrical resistance is zero. In this state, the electrons form bound pairs of opposite spin called Cooper pairs [136]. The driving mechanism of this attraction is due to the Coulomb attraction of electrons to the lattice vibrations in the material, known as phonons. The force of this attraction is weak enough that the thermal energy of ions at room temperature is sufficient to stop the formation of Cooper pairs. Below T_c however, the thermal energy of the lattice is below the energy gap of the Cooper

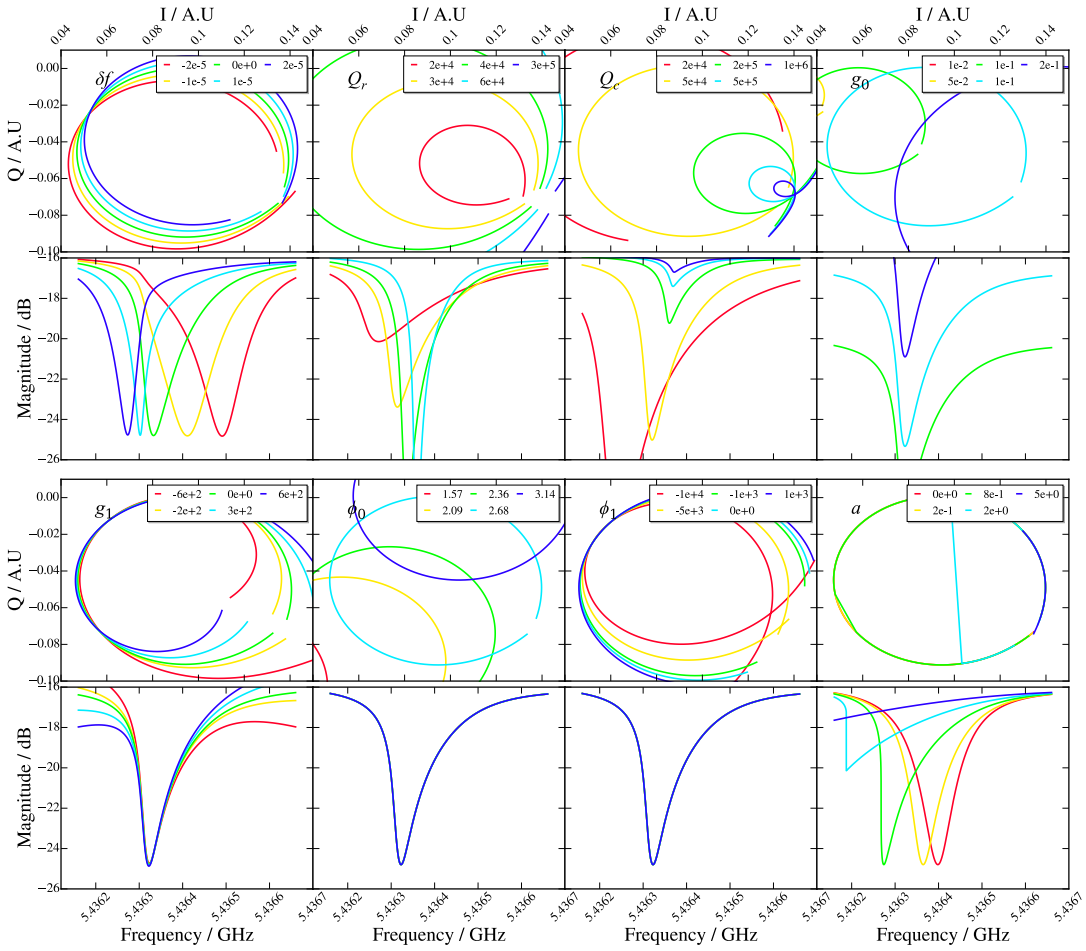

Figure 2.3: Simulated resonator profiles displaying the effect of eight of the parameters in Equation 2.4. For each resonance loop and transmission spectra pair the parameter in question is allowed to vary and the remaining parameters are held constant. The parameters are displayed in the top left corner, and the parameter ranges are displayed in the top right corner, of each resonance loop.

pairs $2\Delta \approx 3.52kT_c$ and they can flow without energy dissipation [137].

Superconductors have a non-zero impedance for alternating currents. An external electric field applied to a superconductor will accelerate Cooper pairs, providing them with kinetic energy. When the field is reversed, the Cooper pairs will experience a force in the direction opposite to their momentum. Their finite inertia will oppose the change of current and this effect is therefore equivalent to the series inductance of conventional electrical conductors. This is the kinetic inductance.

We now briefly describe the fundamental equations governing MKID responsivity drawing primarily from Mazin [138] and Szypryt [2]. According to the Drude model of electric conductivity [139], the AC conductivity of a conductor is

$$\sigma_{\mathrm{ac}}(\omega) = \frac{\sigma_{\mathrm{dc}}}{1 + i\omega\tau} \tag{2.7}$$

where ω is the frequency, τ is the scattering time. The $i\omega\tau$ term arises from the phase lag between the current and the electric field due to the inertia of the charge carriers. The scattering time is very short in metals at room temperature so the conductivity is almost completely resistive. In a superconductor both $\tau \to \infty$ and $\sigma_{\mathrm{dc}} \to \infty$, and the conductivity becomes almost completely inductive. The surface impedance is therefore

$$Z_s = R_s + i\omega L_s \tag{2.8}$$

where R_{s} is surface resistance and L_{s} is the total surface impedance. The resistive contribution arrises from the population of unbound electrons called quasiparticles. At lower temperatures, more quasiparticles are bound as Cooper pairs resulting in less dissipation.

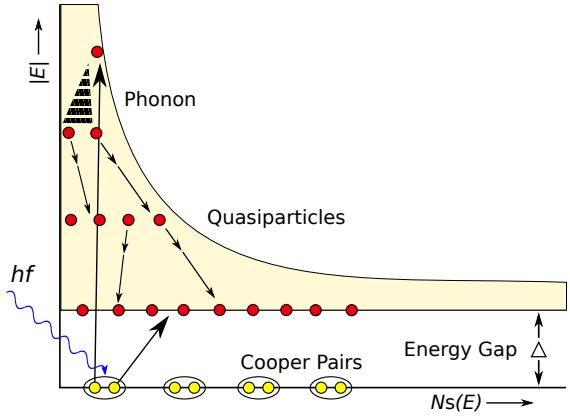

Figure 2.4: A cartoon illustration of the response of the Cooper pair and quasiparticle populations to absorbed photons. If the photon has sufficient energy $h\nu > 2\Delta$ then a number of quasiparticles $N_{qp} = \eta h\nu/\Delta$, where $\eta \approx 0.57$, will be created through a cascade.

2.3 Photon Detection

Photons with energy $h\nu > \eta 2\Delta$ will break either a Cooper pair or many through a phonon cascade. This will generate a number of quasiparticles

$$\delta N_{qp} = \eta h\nu/\Delta \tag{2.9}$$

where $\eta \approx 0.57$ is the efficiency that the photon energy is converted to quasiparticles [140]. These quasiparticles will then recombine on a timescale $\tau_{\rm qp} \approx 1\mu$s depending on the material and geometry. Figure 2.4 illustrates this effect.

The temporary increase in quasiparticle population increases the kinetic inductance (and resistance). The fractional change of the impedance roughly corresponds to the fractional change in the quasiparticle density $\delta Z_{\rm s}/Z_{\rm s} \approx \delta n_{\rm qp}/n_{\rm qp}$, leading to

$$\frac{\delta Z_s}{Z_s} \approx \frac{\delta n_{\rm qp}}{2N_0\Delta} \tag{2.10}$$

where N_0 is the single spin density of states at the Fermi level of the metal. The impedance change is very small, but sensitive measurements can be made with high quality microwave resonators.

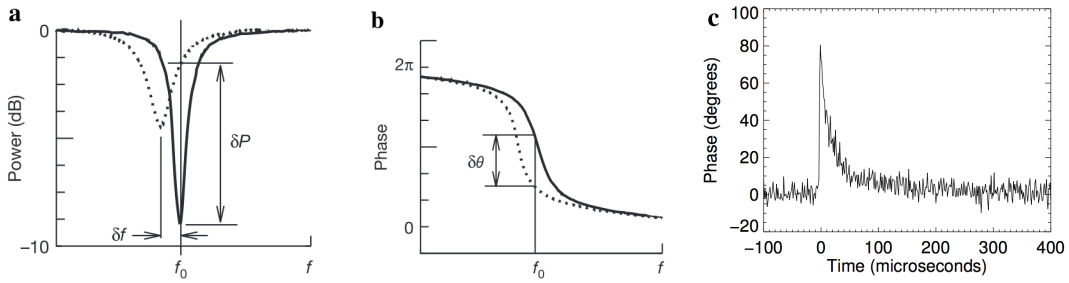

Figure 2.5: The photon detection principle of MKIDs. a: The increase in quasiparticle density in the inductor causes a decrease in the resonant frequency. b: This causes a change in the phase of the bias tone. c: The phase profile shown as a function of time has a sharp increase and then an exponential decay as the quasiparticles recombine. Figures a and b are modified from Mazin et al. [142] and Figure c is modified from Mazin et al. [143].

Figure 2.4 shows the response of an MKID pixel to an absorbed photon. The process has been explained in [141]. The increase in the quasiparticle density increases the inductance causing a decrease in the resonant frequency. The increase in the quasiparticle density also causes an increase in the resistance that causes more dissipation and decreases the resonator's quality factor.

The resonator frequency shift can be quantified by observing the phase (or amplitude) change of the tone biasing the resonator. This is measured using $\phi = \arctan\left(\frac{I-I_{\text{center}}}{Q-Q_{\text{center}}}\right)$ where I_{center} and Q_{center} are the center coordinates of the resonance loop. This measurement during a photon event is shown in Figure 2.5b. The quick decay of the phase response and the continuous read out of MKID pixels (described in the following section) allows for a time resolution on the order of μs for photon detection.

The more energy the incident photon has, the more Cooper pairs are broken, and the larger the amplitude of the phase pulse height. By measuring the height of the phase pulse at several time samples, we can obtain a (reasonably) accurate measurement for the energy of the incident photon. MKIDs therefore have (modest) intrinsic photon energy resolution. The current generation of UVOIR MKIDs achieve an average spectral resolution $R = \lambda/\Delta\lambda = 8$ at 406.6nm [129].

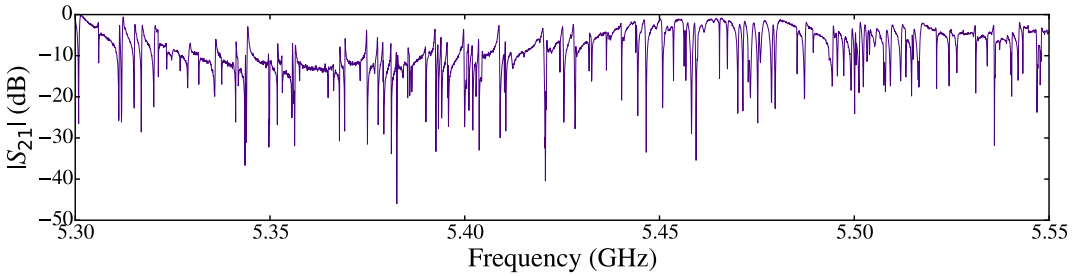

Figure 2.6: The transmission spectrum of a subsection of a typical feed-line with 2000 resonator pixels in a 4–8 GHz bandwidth sampled at a single power using a vector network analyzer. The highlighted resonators correspond to those in Figure 3.3.

From Mazin [138], the phase response for a change in the amount of quasiparticles is given by

$$\frac{d\phi}{dN_{qp}} \approx 1.63 \times 10^{-7} \frac{\alpha Q_r}{V} \tag{2.11}$$

where α is fraction of the total inductance that the kinetic inductance is responsible for and V is the volume of the resonator. In order to design responsive resonators, the selected material and geometry needs to be chosen that has a high α, Q_r and a low V. Typically, the responsivity is tuned to produce 120° phase shift for a photon at the energy of the center frequency of the band [3].

2.4 Digital Readout

UVOIR MKID arrays consists of thousands of microwave resonators, each with different resonant frequency, coupled to several feed-lines [144]. Figure 2.6 shows a subsection of the transmission spectrum of wideband frequency sweep across a single feedline where each dip is the profile of a single resonator.

Since the resonator quality factors are high, the carrier signal for a given pixel is not affected by other pixels as it propagates through the array. This means that all of the pixels can be driven simultaneously with a single signal consisting of a sum of each of the resonant frequencies – known as a frequency comb. The complexity of the readout is therefore transferred to the digital backend, which has to distinguish and separate (demultiplex) the combined signal. The transmission phase of each of these

signals can then be evaluated and monitored for photon events.

Each of these processes is carried out using custom firmware employing so-called software defined radio. Software defined radio (SDR) is where software (or firmware) components carry out the tasks typically performed with hardware. In the case of MKID digital readouts, a given set of resonant frequencies can be entered into the readout software prior to observations, the look-up table of values for the comb waveform (sum of all frequencies) is then calculated, and then that digital signal converted to a voltage using a digital to analogue converter (DAC). Conventionally this would be done with an analogue signal generator, but these are often limited to just two simultaneous frequencies. The firmware to perform the signal generation and processing can be implemented on a field programable gate array (FPGA) using the CASPER reconfigurable open architecture computing hardware (ROACH) board [145].

In order to reach the microwave frequencies of resonators, the SDR readout requires a local oscillator (LO) signal and a mixer to up-convert the signals from the DAC. After the frequency comb has passed through the array and the photons have imprinted their information on the signal, the same LO signal down-converts the signals back down to the baseband frequencies where it is digitized using an analogue to digital converter (ADC). The assembly of one ROACH, one ADC/DAC and one IF board comprises one readout cartridge.

As of writing this, two generations of readout have been developed and deployed in the field that utilize this architecture for UVOIR MKID readout. The readout architecture for the first generation is shown in Figure 2.7 [146]. The software and firmware for this system are fully open-source[1]. This system is capable of measuring 256 pixels per readout cartridge in a 512 MHz bandwidth. In order to reach higher pixel count arrays, either the quality factors of the array have to be increased so more resonators could fit in the feed-line bandwidth, or the sampling rate of the ADC has to be increased allowing for a wider bandwidth [144].

Figure 2.8 shows a photo of a single cartridge for this readout used for the KID-

[1]https://github.com/bmazin/SDR

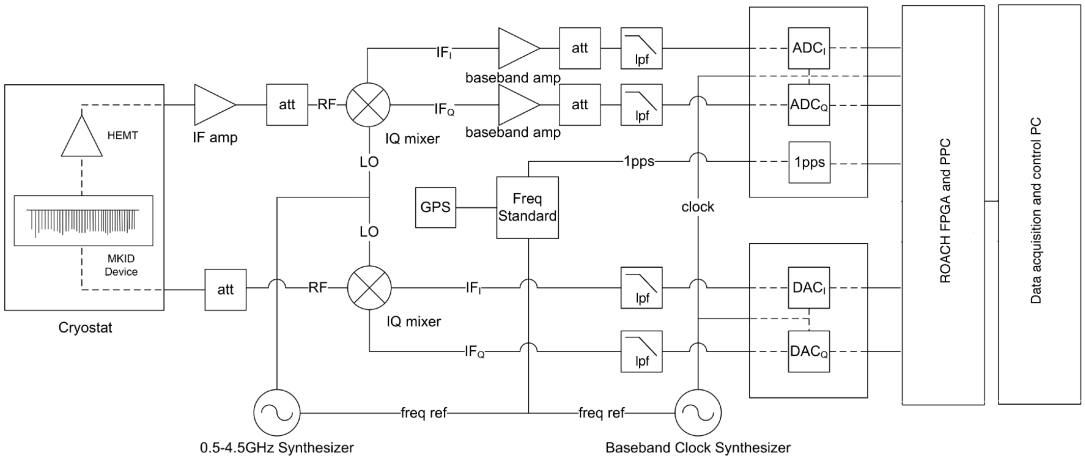

Figure 2.7: A block diagram of the SDR readout for ARCONS from [146]. To readout an MKID detector a comb signal at each of the resonator frequencies is generated and passed through the device, at which point each MKID imprints the information of any detections on the corresponding sine wave. The comb is then amplified and sent to the room temperature electronics where it is digitized and any phase modulations are recorded. Apart from the array and the low noise amplifier all analyses are done using room temperature digital electronics.

Spec demonstrator (introduced in the following section). To fit the ROACH board, ADC/DAC board and IF board in the tray, and preserve the tray walls, the IF board was installed in a non-conventional orientation.

The second generation readout system has a wider bandwidth and is therefore capable of larger amounts of resonators per feedline [144]. This system benefits from faster DACs and ADCs, with a factor of 2 and 4 fold increase in sampling rate respectively. To accommodate these higher data rates a Virtex7 FPGA was installed in the ADC/DAC board to route signals. Two readout cartridges are required per feed-line to sample the whole bandwidth of each feed-line (as opposed to the first generation that required four per feedline). Figure 3.1 shows one readout board for this system.

There are three main tasks performed by the firmware. These are: the comb generation, demultiplexing of the return signal, and phase time stream processing. For demultiplexing, the digital readout separates the tones with an initial coarse periodic binning step followed by finer fixation step that results in a narrow pass

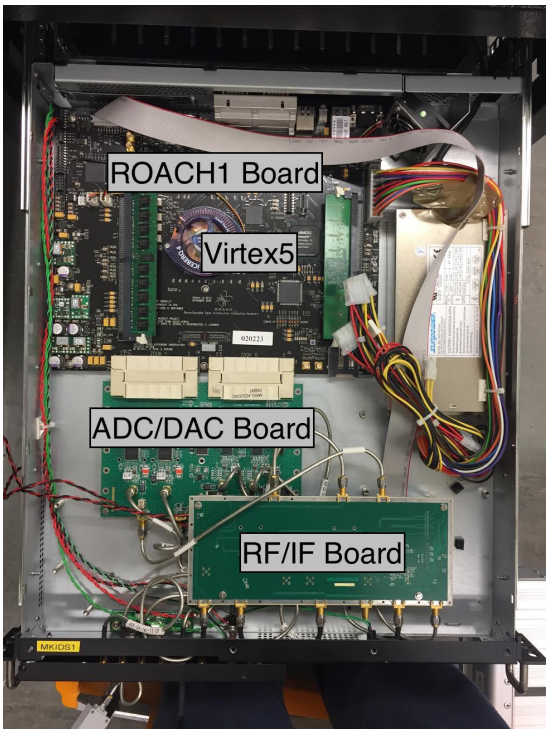

Figure 2.8: A photograph of the tray housing one generation 1 readout board for the KIDSpec demonstrator (introduced in the next section). Virtex5 is the model of the FPGA that handles the majority of the signal processing demands. The MKID array interfaces with the readout system using the SMA ports at the bottom of the image. The tray forms part of larger enclosure that houses the data acquisition PC, an ethernet switchboard and the 10 MHz/ 1 pulse per second GPS clock source.

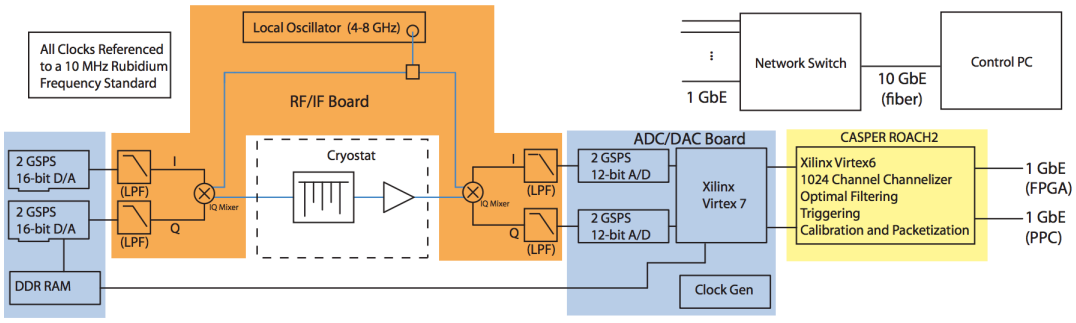

Figure 2.9: A block diagram of the second generation readout unit. Compared to the first generation readout, this design has ROACH board with more FPGA resources and an ADC/DAC board with a faster sample rate and its own FPGA, ultimately enabling more bandwidth, and therefore resonators, per readout unit.

band surrounding each tone [144]. The minimum resonator spacing, and therefore the number of pixels, is in part, dictated by the pass band around each tone after this process. If two tones lie within this pass band, this will create an oscillating 'hot pixel' that must be removed in post processing (see Chapter 4).

2.5 Array Parameters

The larger the total amount of pixels the better the sampling of speckles (if the field of view remains constant) and the more effective the speckle suppression methods can be (all other things being equal). The pixel formats of a system can be increased by simply adding more readout boards in parallel. This however linearly increases the cost and power consumption of the system. A megapixel array being processed with the second generation readout would require 50 readout crates (see Figure 2.12 in the following section for scale). Fortunately microwave electronics and SDR are of interest for many areas of industry, and therefore we can expect the multiplexing factor to increase with time. If the sampling rate capabilities of ADCs (and DACs) double every four years (as was the case for the chosen components of the first and second generation readouts) then by the year 2030 when ELT-PCS could begin development, we could expect a requirement of only ten feed-lines for the megapixel array for the same resonator spacing (equivalent amount to the current UVOIR MKID systems). Alternatively, the resonator spacing is also expected to improve as fabrication methods and materials are optimized, relaxing the requirements on the sampling rate.

The pixel yield is determined by the quality of the fabrication. Imperfections in fabrication cause random perturbations in the resonator parameters such which affect resonator spacing. Resonators will shift away from the designed frequency by differing amounts leading to frequency collisions. If a photon is incident on the higher frequency resonator pixel, that resonator can shift into the lower frequency resonator. Probe tones at both resonant frequencies will measure a phase shift and a false detection will have occurred in the lower frequency resonator pixel (despite the two pixels being well separated spatially). For example, titanium nitride, TiN, was conventionally used for

UVOIR MKIDs because the T_c of the material can be tuned by varying the nitrogen sputter rate. However, this sensitive process lead to large variations in T_c across the device, resulting in more frequency collisions and a pixel yields of $\sim$70%. Research into platinum silicide, PtSi, as an alternative superconductor material has allowed for pixel yields of 90% [2].

A higher spectral resolution, R, leads to better speckle suppression when using SDI through either the creation of more wavelength frames or less cross-talk between frames. A higher R leads to a better identification of planets based on the methane signature, and allows for better characterization of detected planets. Spectral resolution is determined by spatial variations in the Cooper pair density depending on the resonator geometry design, the quality factor of the resonators and the phase noise of the readout system [2]. The two main sources of phase noise are from two level systems, TLS, [147, 148] and the cryogenic amplifier (typically a high electron mobility transistor). Alternatively, parametric amplifiers promise quantum limited performance and are undergoing intense development [149, 150, 151]. One promising resonator geometry under investigation uses two large parallel plate capacitor regions and a small square inductance region – the resonator has the appearance of a bow tie [152]. These bow tie resonators can be driven with much more power, which helps to overcome the TLS amplifier noise. By utilizing dispersive optics, the KIDSpec demonstrator (see the next section) allows for dramatic gains in R at the expense of one dimension of spatial information [153].

The maximum count rate of an MKID array is dictated by the capabilities of the readout. If the photon arrival frequency on a pixel is too high, buffers in the readout chain overflow and the information is lost. So, a maximum count rate of 2500 cts/pixel/s is applied to all pixels. Maximum count rate can also be important for FPWFSing where more light is introduced onto the focal plane. A neutral density filter can be used to limit the number of photons received by the device so some pixels do not reach the maximum count rate and the non-linearities associated with the saturation of those pixels are avoid. However, this discards useful information for spectral suppression and it reduces the total number of photons from companion

objects, increasing the required observation time for detection. A better solution is to create larger format arrays and adjust the focal ratio accordingly such that each pixel receives less light.

Quantum efficiency, QE, is the fraction of the light incident on the detector that is registered. Exo-earths have a very low flux and therefore a good QE is important to achieve the required contrast levels for exo-Earths in an acceptable amount of time. Current UVOIR MKIDs use a microlens array to focus the incident light onto the photosensitive region and increase the fill factor [2]. Some losses occur because of imperfect alignment of the microlens array. Additionally, some light is reflected depending on the chosen material, but the absorption can be increased with an anti-reflection coating. PtSi devices currently achieve a quantum efficiency of approximately 70% between 800 and 1500 nm [2].

2.6 MKID Instruments

2.6.1 Commissioned Instruments

2.6.1.1 ARCONS

The Array Camera for Optical and Near infrared Spectrophotometer (ARCONS) was commissioned in 2011 on the Coudé focus of the Hale telescope at the Palomar Observatory. It had the capabilities to detect photons from 400 – 1100 nm simultaneously, an $R = 8$ at 400 nm and it had 2024 pixels making it the largest optical cryogenic detector at the time ([142]; [154]). The detector used lumped element resonators substoichiometric titanium nitride, TiN, on a silicon substrate [127, 155]) resulting in a pixel yield of ∼70%. ARCONS has been used to investigate optical enhancements of the Crab Pulsar [156] and period variability of an ultracompact, short period binary system [120].

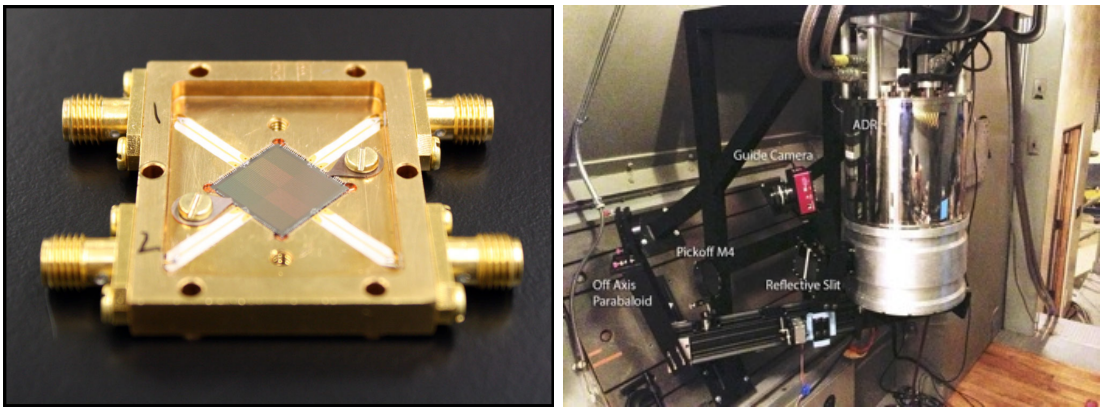

Figure 2.10: Left: A photograph of a 2024 pixel ARCONS array [154]. The array is mounted into a microwave sample box with two feedline pairs that interface to the array over wire bonds. The lid and microlens array are removed. Right: The ARCONS fridge mounted in the Coude room of Palomar observatory.

2.6.1.2 DARKNESS

The DARK speckle Near Infra Red Energy resolving Superconducting Spectrophotometer (DARKNESS) is the first MKID-based integral field unit for HCI [3]. It consists of 10000 pixels, a simultaneous wavelength coverage of 800–1400 nm and an $R \sim 8$ at 1 μm. The second generation readout architecture was developed for DARKNESS. DARKNESS integrates with the Stellar Double Coronagraph and the extreme AO Palm 3000 at the Cassegrain focus of the Palomar Observatory. The initial part of the name comes from a technique known as Dark Speckle Imaging (DSI) [101], which is described in Chapter 5. DARKNESS was first commissioned on July 2016 and has returned to Palomar on four occasions.

2.6.2 Planned Instruments

2.6.2.1 MEC and PICTURE-C

MEC, the MKID Exoplanet Camera, is currently being integrated with Subaru Coronagraphic Extreme AO [157]. SCExAO is a series of wavefront control systems and high throughput, low inner working angle coronagraphs [158]. The MEC arrays are based on the DARKNESS arrays except that they have nearly double the number

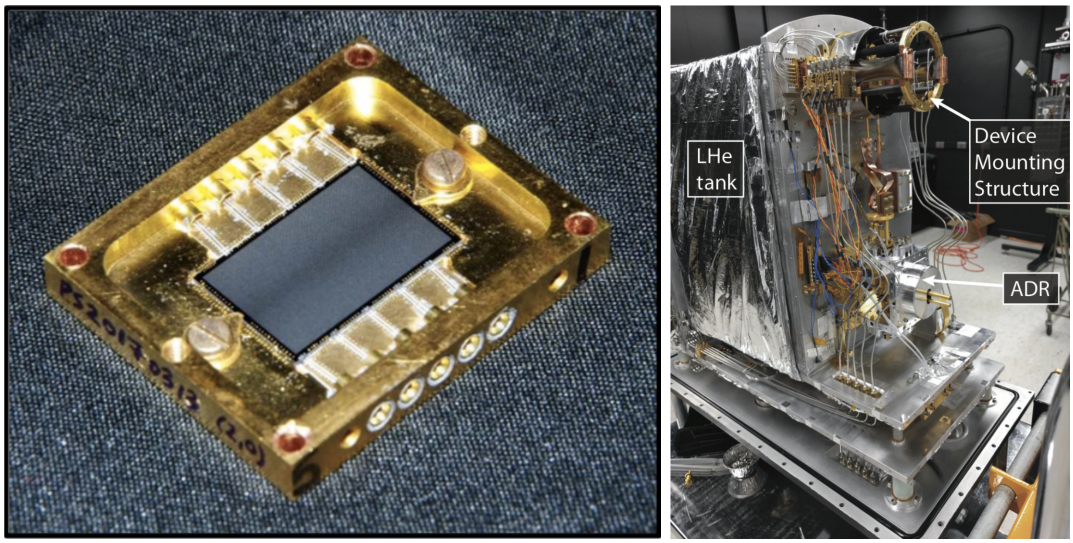

Figure 2.11: Left: A photograph of a 10000 pixel DARKNESS array [2]. The microwave connectors are G3PO for a higher density of feedlines. Right: A photograph of the different temperature stages of the cryostat that will house the DARKNESS array [3]. The cryostat uses liquid nitrogen, liquid helium and an adiabatic demagnetization refrigerator to reach 100 mK.

of detectors at 20,440. It also uses the second generation readout architecture and therefore twice the number of readout cartridges are required. With a larger primary mirror, more pixels and higher sensitivity, this system may have the potential to detect the first exoplanet using reflected star light (instead of thermal emission).

PICTURE-C, the Planetary Imaging Concept Testbed Using a Recoverable Experiment - Coronagraph, is a high altitude balloon borne mission housing an MKID instrument for HCI [159]. This mission is the successor to the PICTURE and PICTURE-B missions, which first flew in 2011, demonstrating HCI and wavefront correction on suborbital flights. PICTURE-C will fly from Columbia Scientific Ballon Facility in Fort Sumner, N.M. It will measure low resolution spectra of debris disks and exo-jupiters around a number of stars of different stellar types and ages. The MKID instrument is based on DARKNESS but adapted for the rigours of balloon flight. The high time resolution available to MKIDs will allow the PICTURE-C images to be derotated (a consequence of pointing system) without blurring or read noise.

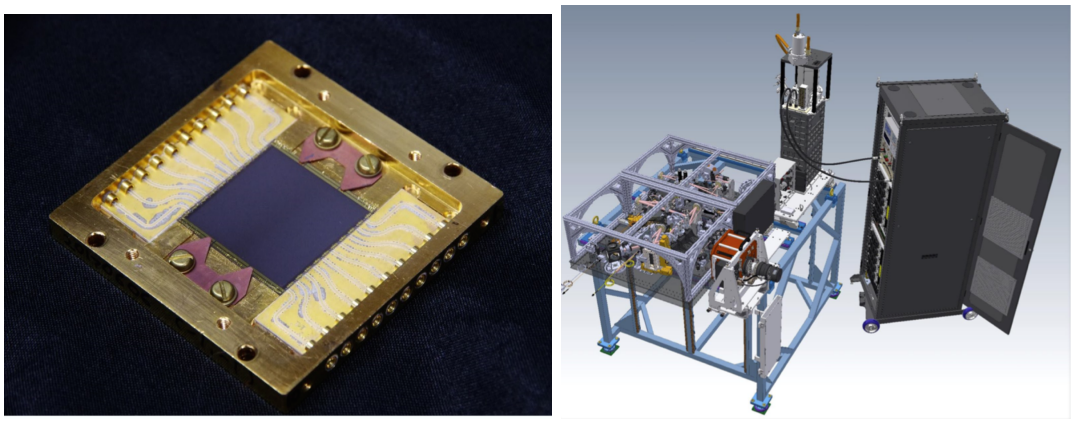

Figure 2.12: Left: A photograph of a 20,440 pixel MEC array [2]. The microwave connectors are G3PO for a higher density of feedlines. Right: A computer aided design of the MEC optical train, readout electronics crate and cryostat.

2.6.2.2 KIDSpec

An echelle-spectrograph has a dispersive element which is blazed to allow high orders followed by a low resolution cross-dispersive element to separate the individual orders, followed by a 2D CCD detector. The result is a high resolution 2D spectrum. A novel instrument has been proposed whereby the cross disperser and CCD of a typical echelle-spectrograph are replaced by a long 1D MKID array, and the intrinsic energy resolution of the array is used as the order sorter [153]. This allows for longer slits compared to cross-dispersive elements, greatly simplifies the optical design of the instrument, in addition to the photon counting capabilities inherent to MKIDs. The instrument could have energy resolution R=4000-10,000 across the entire optical and near-infrared range. Furthermore if a two-dimensional MKID array is used then the instrument would have one dimension of imaging capabilities. The concept was originally proposed by [160] using STJs instead of MKIDs as the order sorter and detector, however the device was limited by the array size and wavelength coverage attainable with STJs.

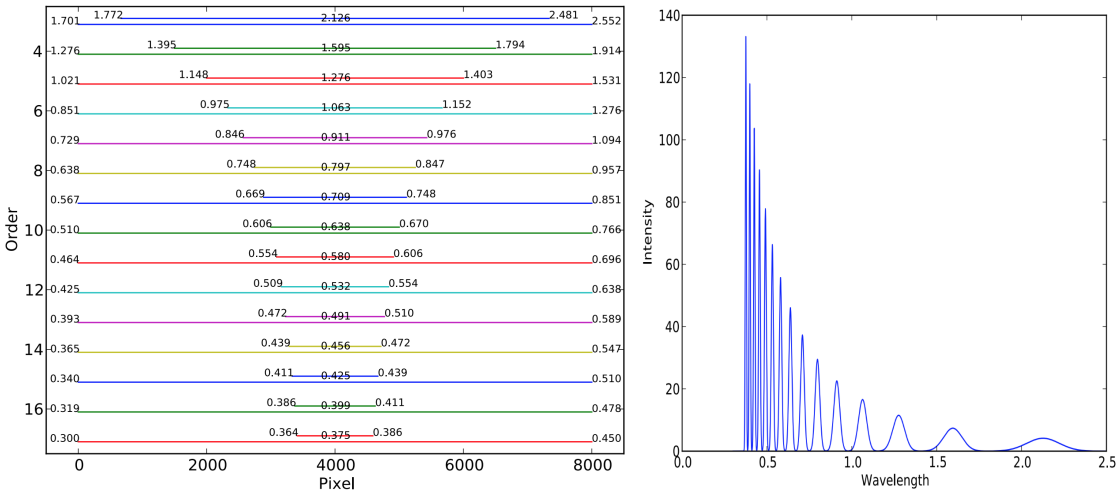

Figure 2.13: Operating principle of KIDSpec from O'Brien et al. [153]. The free spectral range of a given order are shown as well as the total range. b: shows the corresponding spectrum as measured by the center pixel if the source was uniform and an MKID array with $R = 30$. From all the measured spectra the source spectra can be reconstructed.

A simulation package was created to simulate the performance of KIDSpec for a given set of system parameters such as device format, spectral resolution. The package uses a high resolution spectrum from HIRES (HIgh-Resolution Echelle Spectrometer) [161] and then degrades the spectra accordingly. From Figure 2.14 it can be seen that a 1000 pixel linear array and R of 12 is sufficient to observe the spectral lines at 394 nm.

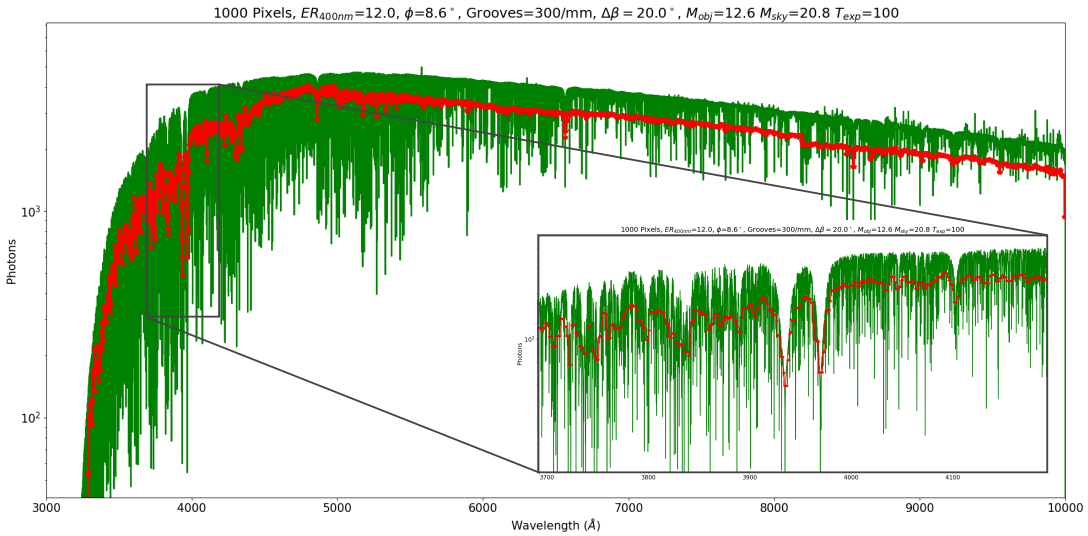

Figure 2.14: The predicted performance of KIDSpec [162]. The green curve is the original measurements from HIRES, the red is the simulated spectrum with KIDSpec. The simulation parameters are: 1000 pixels linear array, $R = 12$, a grating phase angle of 8.6° and 300 grooves/mm, an angular size of the each diffracted order sampled by the array of 20°, an exposure time of 100 s, and an absolute magnitude for the object and sky of 12.6 and 20.8 respectively.

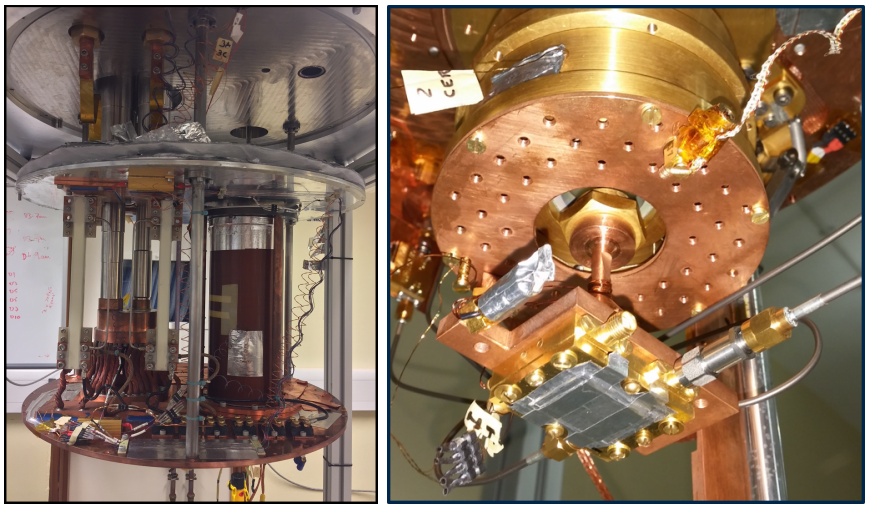

Figure 2.15: A photograph of the KIDSpec demonstrator cryostat. The cryostat uses a Pulse Tube to reach 4 K and then an adiabatic demagnetization refrigerator with a 9 T magnet to reach $T < 100$ mK. A test PtSi device is mounted on the cold finger.

2.7 Summary

In this chapter we have outlined the operating principle of MKIDs and their readout. We have discussed the relevant parameters for HCI as well as current and soon-to be instruments. With this knowledge we are now able to outline the algorithm that I developed to automate the bias point tuning for the digital readout – an important development for large format MKID arrays for HCI.

Chapter 3

Digital Readout Tuning with Machine Learning

In this chapter we introduce various required strategies for tuning the readout to the parameters array, such as resonator bias point selection. This particular process previously took several hours per one thousand MKID pixels. We detail the machine-learning based algorithm that I developed to automate this process. We conclude with plans to improve this algorithm.

This research was primarily conducted at the University of California, Santa Barbara (UCSB) during a Long Term Attachment (LTA) using archival measurements of DARKNESS arrays.

3.1 Readout Tuning

There are several necessary tuning steps that are necessary for taking observations with MKIDs. Firstly, the digital readout must be programmed with the required resonant frequencies and corresponding optimal bias power for the pixels making up the MKID array. In theory, these parameters are known a-priori from the design parameters of the array. In practice, unavoidable random errors inherent in the fabrication process cause variations in the parameters of the array. Science grade UVOIR MKID arrays consist of several independent fabrication layers, and include

spatially small ($\sim$150 μm pixel pitch) resonators closely spaced in both position and frequency space (see [2] for more detail). The imperfections inherent in the fabrication process cause resonators to shift away from the designed frequency and quality factor, causing resonator overlaps in the frequency domain which distorts the transmission profiles. These effects also bring about a variety in the power handling behavior of resonators – as will be demonstrated in section 3.2.5.

The resonant frequencies are acquired by performing a frequency sweep and measuring the transmission across the bandwidth of each readout unit using a vector network analyzer (VNA). Figure 2.6 shows a subsection of one of these measurements. The peaks are then located first by applying a matched filter tuned to the typical resonance profile, and then by human manual inspection to confirm the results and check for resonators the algorithm might have missed.

The bias selection tuning step operates in a larger parameter space and is more challenging to automate. This readout tuning process and the development of an automated algorithm will be the focus of the remaining sections of this chapter.

In order to identify phase pulses as photon events, the pulse amplitude must be beyond a defined threshold. A matched filter is created for each pixel's phase response (e.g. Figure 2.5c) to increase the SNR (Strader [144] for details). The baseline phase threshold determines the amount of false triggers from excessive phase noise. For each pixel the threshold is set to a value that discards the noise tail in the pulse height histogram assuming it is well separated from the distribution of the real triggers. This is typically four standard deviations above the noise floor.

For ARCONS and DARKNESS, the arrays were designed to have maximum physical spacing between pixels with similar resonator frequency, in order to reduce cross talk as a result of the frequency collisions. Therefore, we have to identify which resonant frequency corresponds to which pixel location. This measurement, known as a beam-map, is performed by scanning a strip of light across the array, and monitoring each pixel for a phase response. For DARKNESS, this strip of light is created by sequentially turning pixels from black to white on a LCD screen which is then imaged onto the chip with optics.

3.2 Optimal Resonator Biasing

To read out an MKID pixel, we measure the phase of the transmission at the equilibrium resonant frequency. To optimize the signal-to-noise ratio on this measurement we drive the resonators as powerfully as possible to reduce the contribution from the cryogenic amplifier (typically a High-Electron-Mobility Transistor) noise and the effect from Two Level System (TLS) noise [141]. This changes the transmission profile of the resonator.

Figure 3.1 shows the transmission through the device at frequencies ±450 kHz of the resonant frequency of a single MKID pixel at a range of readout powers. The lowest excitation used is at approximately -70 dBm. This type of measurement is hereafter termed a resonator powersweep. At low power the I and Q components trace a continuous resonance loop in the complex plane and a continuous dip in the transmission spectrum. At higher powers the resonator exhibits a nonlinear response as a function of driving current, and given sufficient power, the resonator bifurcates into two quasi-stable states, which manifest as a discontinuity close to the resonant frequency [135, 133]. In this bifurcation regime the resonator is rendered non-functional for photon detection [144]. Therefore, to bias a given resonator, the power at which the resonator first bifurcates is identified, and the power 2 dB below this value is chosen. This method is the primary criterion of this chapter and will hereafter be referred to as Rule #1.

As an aside, it has been shown by [163] that overpowering resonators can be detrimental to device sensitivity because of the increase in generation-recombination noise from readout power heating, while [135] have shown that in certain circumstances there could be benefits from operating in the non-linear regime. Ultimately, optimal operating point of a given resonator will depend on several factors such as material, geometry, quality factor and the wavelength of incident light (since UVOIR MKID detectors are not currently limited by generation-recombination noise). In this section we will assume that the sensitivity is maximized by applying Rule #1 (along with the exceptions related to pathologies described later). In section 3.5 we will discuss how

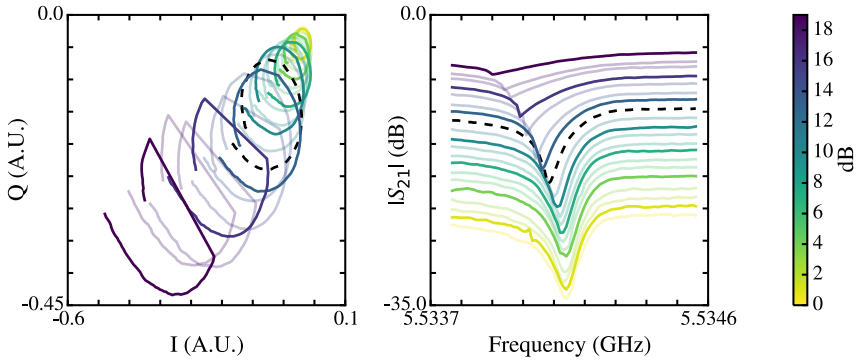

Figure 3.1: The transmission profile of a resonator, displaying ideal bifurcation behaviour at higher readout powers, plotted in the complex plane (left panel) and the magnitude spectrum (right panel). I and Q are measured in arbitrary units. The colourbar represents the amount of additional power applied to the resonator, and high saturation colours are separated by 3 dB. The lowest bias point is approximately -70 dBm. The black dashed line shows the optimal bias point chosen by an operator (human), using manual inspection, when applying Rule # 1.

this algorithm can be modified to optimize sensitivity or spectral resonation directly.

3.2.1 Analytical Method

In Chapter 2 we derived the complex transmission of resonators driven in the high power regime. Equation 2.4 depended on a the nonlinearity parameter. With this parameter, we are able to measure the degree of bifurcation of a resonator at each power sample and apply Rule #1. SCRAPS is a superconducting resonator analysis package [130] that can be used for this purpose. It has the functionality to fit measured I and Q data of a resonator powersweep with packages such as LMFIT. We substitute the default S_{21} equation with Equation 2.4 that allows for bifurcation dynamics and impedance mismatches. With a value of a at each power sample, the optimal bias is then determined by locating the lowest power when a transitions past some threshold. This analytical method will hereby be referred to as 'AM'.

3.2.2 Numerical Method

If the primary concern is the degree of bifurcation then a very simple but effective metric is to monitor the separation between the I, Q magnitudes at adjacent frequencies, here termed IQ velocity. At a single readout power, it is defined as

$$v_{IQ}(f) = \sqrt{[Q(f) - Q(f-1)]^2 + [I(f) - I(f-1)]^2}, \tag{3.1}$$

where f is the frequency index and $(f-1)$ is the previous frequency index. When a discontinuity forms during bifurcation there is a large spike in the v_{IQ} spectrum at the discontinuity frequency, and at the adjacent frequencies the v_{IQ} should remain minimal. High Q_r resonators will show large values of v_{IQ} due to the larger relative frequency sampling compared to the resonator width. However, the v_{IQ} of the adjacent frequencies will also be large. To distinguish spikes in v_{IQ} caused by a high Q_r from that of a discontinuity, the maximum v_{IQ} is compared to the mean of the surrounding values in what is here termed v_{IQ} ratio or VR

$$\mathrm{VR} = \frac{v_{IQ}(M)}{\frac{1}{N}\left(\sum_{f=M-N/2}^{M+N/2}[v_{IQ}(f)] - v_{IQ}(M)\right)}, \tag{3.2}$$

where M is the frequency index of the maximum v_{IQ} and N is the total number of the adjacent frequency samples. A tradeoff exists whereby a larger N provides more data points to overcome the noise in v_{IQ}, at the cost of increasing the amount of included baseline values around a high Q_r resonator peak and artificially increasing the VR, as well as increasing the likelihood of sampling any glitch features or multiple discontinuities (described in Section 3.2.5). Before calculating VR, it is also necessary to smooth v_{IQ} with a low-pass filter because of the measurement noise. After smoothing v_{IQ}, and applying an $N = 6$ in Equation 3.2, the majority of the false classifications at the lowest bias point are typically removed. This technique, where VR is used to identify the first bifurcation power and apply Rule #1, will hereafter be called the numerical method or NM.

3.2.3 Manual Inspection

Manual visual inspection MI consists of observing various quantities for each powersweep, by eye and then assessing the optimal power to maximize sensitivity while maintaining operability. Figure 3.2 shows the graphical user interface that displays these quantities for a single powersweep. When an operator jumps to a certain resonator, the program applies NM to the resonator to determine a first guess for the optimal bias. The operator is then presented with the resonance loops and v_{IQ} spectra for that power level and the two levels below that shown in green (or equivalently two levels higher in attenuation), as well as VR and maximum v_{IQ} at all powers. The operator can then assess if this value is correct. For example, they may look at the two green curves of the resonance loops and assess the size of the discontinuities. They may also look at the VR curve to check for a monotonic decrease past the threshold at increasing attenuation. If this is not the case, the NM may have chosen a power above the saturation power. In this case, the resonator was afflicted with a pathology.

3.2.4 Summary

Three methods have been described for classifying the optimal bias point of a resonator powersweep: AM, NM and MI. These methods are summarized in Table A.3. The metric thresholds that a resonator sampled at a given power must cross to be considered bifurcated are: $a = 4\sqrt{3}/9 \approx 0.77$ and VR = 3.5, for the AM and NM methods respectively. The a threshold is taken from the bifurcation value in [135], and the VR threshold is chosen from experience when applying the MI method, but can be adjusted depending on the amount of I and Q noise.

For both AM and NM methods, the first instance where the metric is above the threshold is used (given that the metric sometimes transitions across the threshold multiple times), and if this places bias point on or beyond the boundary ($\leq$0 dB), then 1 dB is chosen. If no metric is above the threshold then the location of highest metric is the selected bifurcation power.

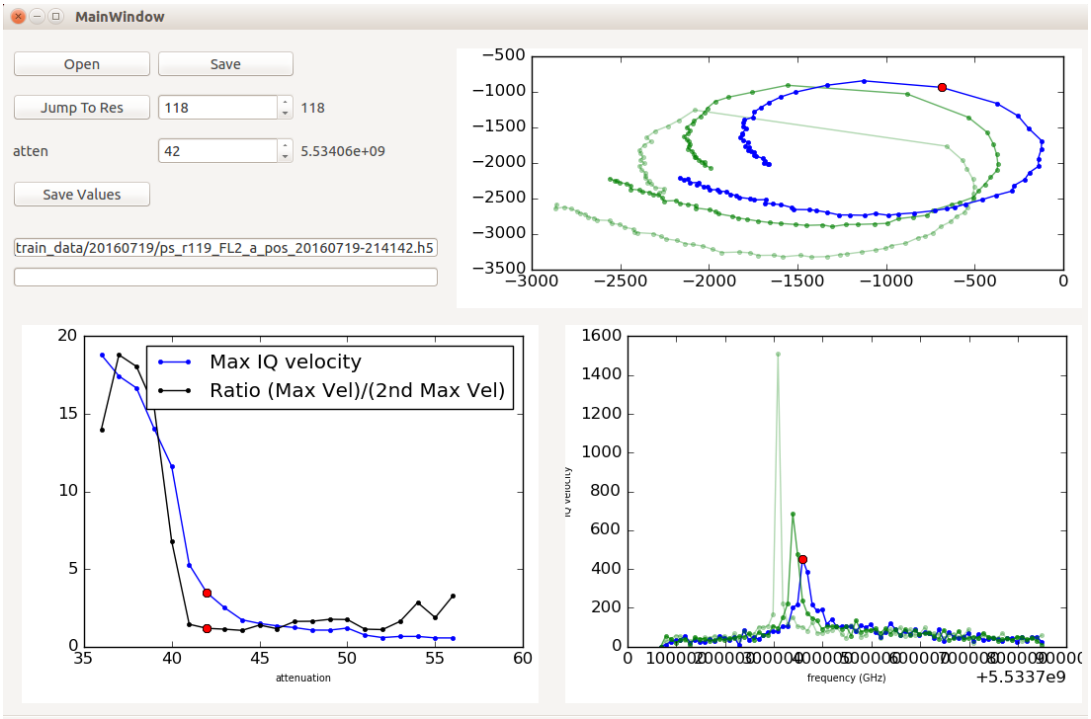

Figure 3.2: The graphical user interface for bias (and frequency) classifying resonators. The resonator shown is #118 of the feed-line (Ukko2, used in the later analysis). The top-right panel is the resonance loop at three power samples – the blue curve corresponds to 42 dB. The bottom right panel is the v_{IQ} spectra at the same three power samples. The red dot is the frequency that will be recorded. The red dot in the bottom-left panel shows the attenuation that will be recorded.

Table 3.1: Three methods for evaluating the bias point of a resonator powersweep, the associated metric parameters and the approximate time to implement the technique on a ten-kilopixel MKID array.

Method	Metric	Threshold	Time (min)
Analytical Method (AM)	a	0.77	600
Numerical Method (NM)	VR	3.5	30
Manual Inspection (MI)	I, Q, v_{IQ},VR	–	1200

Prior to applying the AM and NM methods, the I and Q amplitudes for each of the resonators at each power are filtered to suppress the noise. It was found that the optimal method involved using a convolution-based, running average method across all points excluding the maximum v_{IQ}. This technique meant the amplitude of bifurcation discontinuities remained unaffected. When performing the NM, both the convolution-based smoothing to I and Q, and the low pass filter to v_{IQ}, was applied.

3.2.5 Resonator Pathologies

There are several pathologies that can afflict a resonator, and some resonators exhibit multiple pathologies. Figure 3.3 exemplifies these pathologies. The AM, NM and MI methods have been applied to each example resonator, and the respective estimates of the bias point are shown as vertical lines in the right-most panels. The v_{IQ} parameter shown in the third column is used in the application of the MI and NM techniques. The observation of a single point above the background signifies a discontinuity and therefore bifurcation.

Each resonator in Figure 3.3 is misidentified for a specific reason (excluding (a)). One approach to this problem would be to tune a phenomenological model for all eventualities. While time consuming, it is also difficult to create exceptions that don't adversely effect the performance of other pathologies. For this reason, until now, each resonator has been inspected manually, using the NM technique as a first estimate.

In Figure 3.3a, there is only a 1 dB disagreement between both AM and NM methods and the optimum value using MI. Both AM and NM metrics show a mostly

monotonic increase with power past their respective thresholds, and the selection of the bias point is trivial.

3.2.5.1 No Bifurcation

Figure 3.3b and 3.3c display resonators that do not appear to bifurcate. The resonator in Figure 3.3b has abnormally good power handling ability, which can happen when a resonator has low quality factor Q_{r}, according to equation 2.6. When an operator is classifying this resonator through MI, they may choose the highest power, based on the degree of bifurcation and a rough extrapolation of the trend with power. However, it is also advantageous not to bias a single resonator with too much power, otherwise it will impact the dynamic range of the digital readout for lower powered resonators [144]. This second criterion is applied at the discretion of the operator and is referred to here as Rule #2.

The AM model has detected a trend towards bifurcation and matches to the bias point from MI well. The I and Q magnitudes, however, do not show any discontinuity from bifurcation, so the NM technique has no bifurcation indicator. In this instance, the bias point is actually vastly underestimated, because the NM metric, VR, has an outlier at low power. This effect can very often manifest at low powers where the signal to noise ratio is low and lead to a misclassification of otherwise ideal resonators. This issue can be mitigated by using a wider window during the initial smoothing step for each resonator, using a larger N when calculating VR, or using a larger VR threshold, however each of these processes result in misclassifications for other types of resonators.

The resonator in Figure 3.3c shows a degradation in Q_{r} with increasing readout power until the resonator has lost all appreciable magnitude without any discontinuity forming at any power. This behaviour can sometimes be the result of a second collided resonator sufficiently close, such that it is not discernible from the higher frequency resonator. In order to maintain resonator spectral resolution, one of the lower powers should be chosen during the classification, where the resonance loop shows sufficient curvature and the transmission profile shows sufficient depth, at the discretion of the

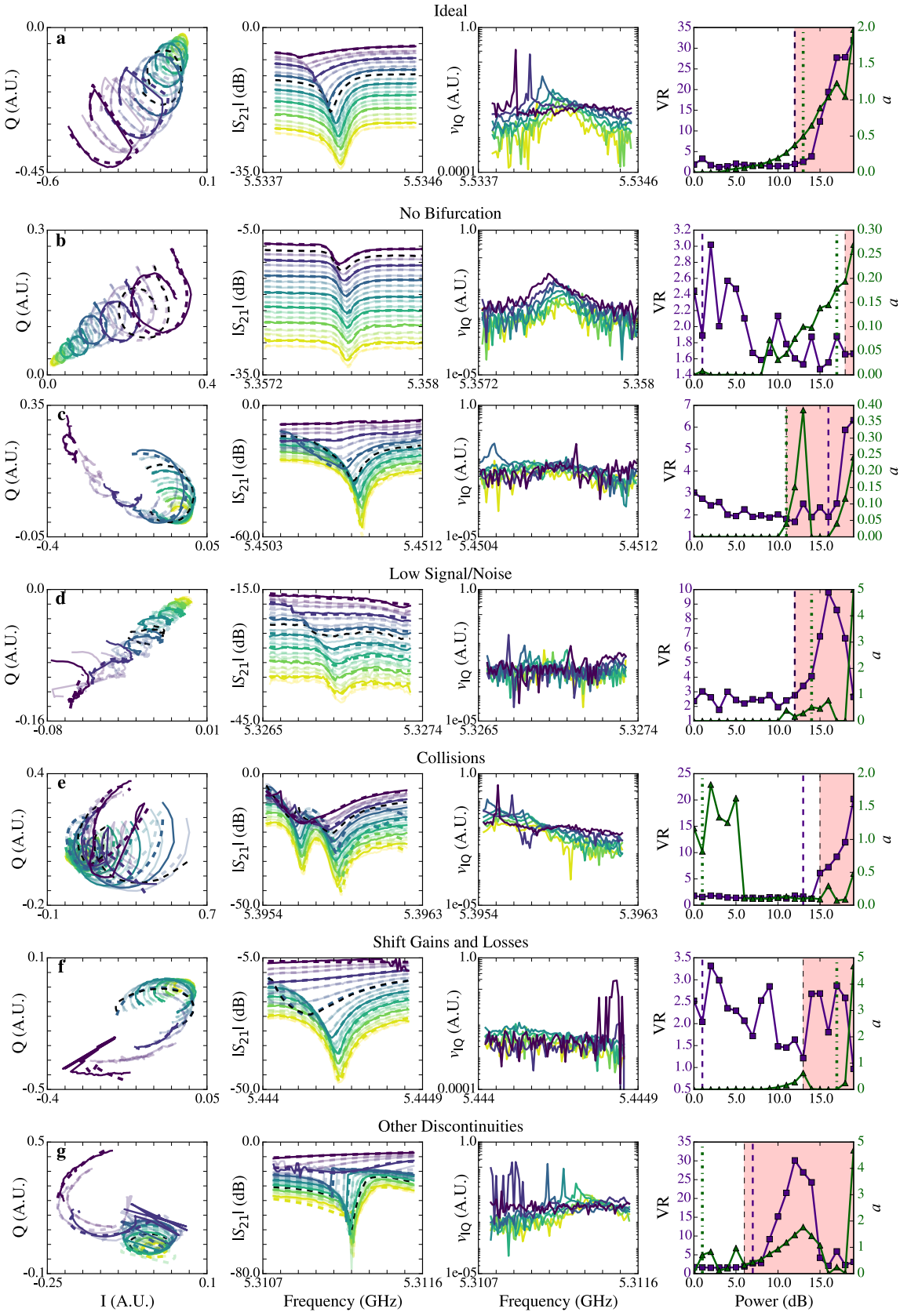

Figure 3.3: a-g: Seven measured resonator powersweeps exemplifying the pathologies described in Section 3.2.5. Full agreement between the three methods is only found on the ideal powersweep. The first two columns have the same formatting as Figure 3.1 and are superimposed with the AM fit (colored dashed lines). In the right-most column, a has triangular markers and VR has square markers. The dashed line, dot-dash line and the intersection of the bank and colored zones mark evaluated bias from the VR and a and MI respectively.

MI operator. This criterion, hereby known as Rule #3, also helps with the finite readout dynamic range.

3.2.5.2 Low Signal/Noise

The resonator in Figure 3.3d has a low signal-to-noise ratio at all powers, in part because it has a low internal quality factor Q_i compared to its coupling quality factor Q_c, and also because it is in the 5.3 – 5.45 GHz region of the feed-line with reduced transmission (see Figure 2.6d), leading to a shallow resonator transmission dip. This makes the bifurcation power harder to identify, and in severe cases, no bifurcation is visible. With the example resonator shown in Figure 3.3d, the AM and NM metrics show significant noise and nearly pass their respective thresholds at low power.

For certain measurements, such as speckle-noise suppression of high contrast imaging observations, it is often preferable to retain resonators with these characteristics in order to maximize total pixel count and retain the temporal information. In observations where high spectral resolution is of primary concern (for example measuring galactic redshifts [123]), these pixels can be discarded in post processing. Another reason for retaining these observations is that operators will want to characterize a test array.

3.2.5.3 Collisions

Resonators can shift away from the designed frequency by differing amounts [134] because of the non-uniformities in fabrication often causing collisions. Research into different superconductor materials has shown some progress in reducing this effect [129]. However, the issue persists due to cost and technology limits on readout bandwidth, pushing resonators closer together in frequency space to make larger format arrays.

Figure 3.3e shows an example of a collision. If a 1 μm photon were incident on the higher frequency resonator pixel, that resonator will shift approximately 50 kHz towards the lower frequency resonator. If resonators are too close originally, then probe tones at both resonant frequencies could measure a phase shift and a false

detection may occur in the lower frequency resonator pixel. Typically, the higher frequency resonator of a collided pair within 200 kHz of each other is used in the readout tone generation list. However, the lower frequency resonator can be used if the higher frequency resonator shows exceptionally low Q_i. This criterion, Rule #4 is applied at the discretion of the operator.

Since the neighbour modifies the profile of the chosen resonator, fitting that resonator alone may result in unreliable predictions for a. Fitting both resonators simultaneously is a larger problem requiring more parameters, meaning that the least squares algorithm is more likely to converge on a local minimum. The problem can be alleviated using maximum likelihood estimation from Markov Chain Monte Carlo sampling, at the expense of vastly longer computation times on these powersweeps.

The NM underpredicted the bias point because it does not have a means of associating the discontinuity with the lower frequency resonator, which bifurcates at a higher power. Neither the NM, nor the AM, account for both of the bifurcation metrics simultaneously, the $|S_{21}|$ differences, the quality factors, or the separation between adjacent resonators.

3.2.5.4 Sampling Window Gains and Losses

Figure 3.3f shows some ways in which the choice of sampling window, coupled with resonators that show a large frequency response with power, can make resonator biasing more challenging. The resonance is centred on the sampling window at low powers and undergoes a relatively large translation out of the sampling window at higher powers. This is an example where operator would have to apply Rule #3. Both the AM and NM metrics do not reach their respective thresholds before the resonance translates out of the sampling window and the location of the highest metric is taken. The analytical algorithm has to handle the onset of bifurcation features and then the sudden disappearance of these features. For this reason, the AM technique fits each resonator power sample independently, and no fit is performed to a as a function of power, otherwise the anomalies at high power would skew the bias point estimate. However, this approach can lead to outliers in a at low power

providing a false classification as seen in Figure 3.3e and 3.3f.

At higher powers the resonator in Figure 3.3e is afflicted by a second pathology. After the original resonator has shifted out of the sampling window, a higher frequency resonator, which is already bifurcated, shifts into the sampling window. AM is triggered by this new resonator, and a bias that is too high is chosen. In other instances when the initial resonator remains in the sampling window, if the second resonator comes within the collision threshold separation of the first resonator, then the classifier (human or an automated algorithm) could choose a lower power or apply Rule #4. If the two resonators are sufficiently separated then both resonators should be fit simultaneously and apply Rule #4.

In each of these instances the choice of sampling bandwidth and centre is important. A trade off exists whereby a wider sampling window is more likely to capture the entire resonator profile as it evolves, but this increases the risk of contamination from adjacent resonators. In practice, a bandwidth of between 0.5 and 1.5 MHz is used depending on the average Q_r of the resonators. Similarly, centering the sampling window on the initial resonance is useful for observing when higher frequency resonators pass the collision separation threshold. The quality-factor and the amount a resonator will translate at high powers (if at all, see Figure 3.3b) are not known a-priori, so a sliding sampling window has its own difficulties.

3.2.5.5 Other Discontinuities

Sometimes resonators can show behaviour that is wildly different from those on the same feed-line. The resonator in Figure 3.3g displays two seemingly independent characteristics. The first is a discontinuity forming at frequencies above the resonance, here termed backwards resonators. This type of behaviour appears to correlate with the loss of transmission between 5.3 and 5.45 GHz. This could be because of an impedance mismatch creating a severe asymmetry in the transmission profile (parameterized with $\delta\omega$ in Equation 2.4). It should be noted that, in the IQ plane, this asymmetry does not affect the bifurcation signature.

The second example of an alternative discontinuity is where a resonator shows

multiple discontinuities at higher power. This could be explained as the resonator switching between the available states after bifurcation, or sometimes glitches from digital readout induced errors.

3.2.5.6 Summary

Table 3.2 summarizes the statistics of the chosen types of resonator power-handling behaviour mentioned in this section. In this instance, the main fabrication artifact driving the variation between resonators, and causing the low amount of ideal and adequate resonators, was the film non-uniformity of stoichiometric titanium nitride (TiN) [129]. The statistics of certain types will vary between different arrays and even feed-lines of the same device. Backwards bifurcation may be more specific to this feed-line but collisions are a common occurrence in all feed-lines. There will also be other types of resonator non-ideal behaviour that are not apparent in this example feed-line.

This is why it is important to develop an algorithm that can progressively learn to account for different types of behaviour as it is exposed to them, instead of continually attempting to manually optimize a phenomenological algorithm. This type of problem, where input data has a large variety of cases, but share some embedded commonalities, is well suited to machine learning. With enough training data a deep neural network architecture should be able to develop a sufficiently sophisticated model to bias resonators to a level of human accuracy.

3.3 Resonator Biasing with Machine Learning

Neural Networks are a class of machine learning algorithms that with emergent behaviour are able to discover and model high level abstractions in multi-dimensional data [164]. As these models are exposed to increasing amounts of input data, their accuracy can progressively improve. Deep Convolution Neural Networks (CNNs) have had great success in many areas of computer vision [165], most notably image recognition [166]. These networks take a multi-dimensional vector as their input, process

Table 3.2: Occurrence statistics of the different powersweep pathologies for feed-line two of device Ukko, "Ukko2", identified manually. There are 372 resonators in total and the properties are not mutually exclusive. 'Ideal' resonators show no collisions, sampling window effects, multiple or backwards discontinuities. 'Adequate' resonators may include well separated collisions, backwards discontinuities, shift gain or multiple discontinuities after the bifurcation power. 'Shift Loss' is the where the resonance translates out of the sampling window to lower frequencies. 'Shift Gain' is when an adjacent resonance shifts into the sampling window. 'Backwards' refers to resonators with the discontinuity at frequencies higher than the curve minimum. 'Low Q_r' are those which are difficult to classify because of shallow profile of the resonator.

Type	Absolute	Percentage (%)
Adequate	185	49.9
Shift Loss	133	35.8
Multi. Disc.	108	29.1
Noise	102	27.5
Backwards	87	23.5
Shift Gain	83	22.0
Collision	71	19.1
Ideal	44	11.9
Underpowered	34	9.2
No Bifurcation	16	4.3
Low Q_r	10	2.7
Unusable	2	0.5

the input vector with a series of filters to extract the relevant features, and produce an estimate of the label that identifies the input vector [167]. In the case of image recognition tasks, the input vector is the red, green and blue channels of an image, and the label corresponds to the class of object in the image. The filters (or weight vectors) must be "learned" by training the model on many input vectors of known labels, known as a training dataset.

A CNN was utilized as a model for predicting the bias point of resonators, with an architecture similar to how a conventional image-recognition network would be structured. The power and frequency axes can be thought of as the spatial dimensions of an input image, and the I, Q, v_{IQ} channels play the same role as the red, green and blue channels. The label corresponds to the optimal bias point for the powersweep. For the training and evaluation datasets, previous manual inspection (MI) data were used.

3.3.1 Architecture

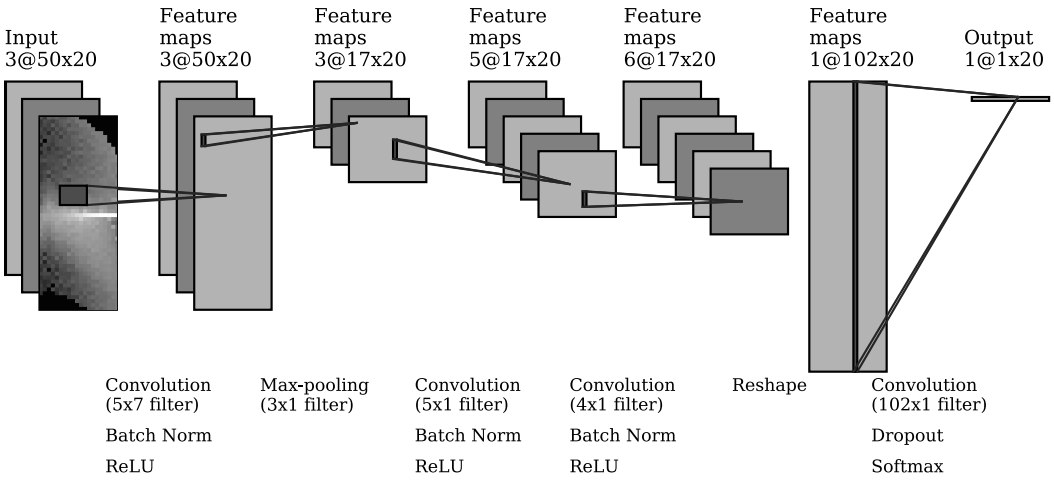

Figure 3.4: An illustration of the convolution neural network architecture used for classifying powersweeps based on the optimal bias point. The input vector is I, Q and v_{IQ} with 50 frequency samples and 20 power samples. The v_{IQ} vector is shown with example values logarithmically scaled. The output vector contains a probability for each of those power samples. The dark boxes highlighting a small area in each layer display the receptive field of each weight vector. The initial convolution layers are followed by a batch normalisation layer and a rectified linear activation layer; the final convolution layer instead is superseded by a dropout layer and a softmax activation layer. The penultimate layer is not to scale. This figure was generated by adapting the code from draw_convnet[1].

The chosen design of the CNN was primarily motivated by the risk of overfitting the model to the limited amount of available training data, which would cause the model to not generalize well to unseen data. The architecture of the CNN is shown in Figure 3.4.

First, each of the layers of the input vector is convolved with a filter that has a spatial extent known as the receptive field and a depth that controls the number of output layers. These filters extract a common feature seen throughout the spatial axis. During convolution, the input vector is zero padded along the border so the output vector retains the same size as the input vector. The first convolution step is the

[1]https://github.com/gwding/draw_convnet

only place in the network where the filters have a receptive field that extends across multiple powers. By not using a filter with a receptive field that extends across the power axis (a fully-connected layer in one dimension), the learned filters are spatially invariant, and the CNN is made insensitive to non-uniform bias distributions present in the training data.

The product of the convolution layer then undergoes a technique called batch normalisation that increases the rate of convergence during training and acts as a regularization technique to reduce overfitting [168]. After the normalisation layer, the vector is then processed by a nonlinear function that acts as a decision boundary depending on the input value, in what is called an activation layer. Rectified Linear Unit (ReLu) activation was chosen for its superior training times on CNNs [166].

The maximum pooling technique down-samples the spatial axes, prioritizing higher values, to allow larger scales to be probed by an equivalently sized filter during the next convolution. Different architectures were explored containing pooling layers at different stages. Ultimately, a single pooling layer was applied early in the network that only reduces the frequency axis. The subsequent convolution layers then detect the higher level, more abstract features, and add more parameters that increase the model sophistication. Each of these convolution layers is superseded by batch normalisation and ReLU layers.

The feature map vector with a depth of six is flattened along the power axis. The data along the frequency axis is then combined to produce a scalar value for each bias. This is achieved by applying a filter with receptive field that extends across the extent of frequency dimension. During training, in order to lessen overfitting, the regularisation technique known as dropout is applied with a 50% probability of removing a given filter unit during each pass [169]. The final layer is the softmax function that converts the scalar values for each bias into a probability. The class with the largest probability is the selected bias point.

Table 3.3: A summary of all the available feed-line datasets used in training and evaluating the full CNN. Each dataset is comprised of measured powersweeps and corresponding bias point estimates from MI that function as the labels. 'F.L.' is the assigned feed-line number and 'Repeat' is index of the powersweep measurements. Feed-line 2 of Ukko has been classified by several operators to evaluate the different biasing methods in Section 3.4.2.

Device	F.L.	Repeat	Material	Resonators	MI labels
Morpheus	2	i	TiN	198	1
Morpheus	5	i	TiN	192	1
Varuna	2	i	TiN	674	1
Faceless	3	i	PtSi	339	1
Faceless	3	ii	PtSi	390	1
Faceless	3	iii	PtSi	390	1
Faceless	2	i	PtSi	632	1
Faceless	2	ii	PtSi	632	1
Ukko	1	i	TiN	730	1
Ukko	2	i	TiN	372	4

3.3.2 Input data

The lists of bias points for feed-lines found through MI, and originally created for biasing arrays for observations, make up the training data. Table 3.3 summarizes all the input feed-line datasets used. In total, 4549 powersweeps from ten feed-line datasets were available for training (and evaluating) the CNN. These feed-line datasets came from several arrays, with data from two feed-lines, at most, for each array. Some feed-line datasets are repeat readings of powersweeps on different cool-downs. These devices are made up of different materials, three were TiN on silicon and one was platinum silicide (PtSi) on sapphire.

Each input vector was preprocessed to help the CNN converge on the relevant properties. The v_{IQ} magnitudes of each powersweep were normalized to their maxima across all powers. The I and Q magnitudes were both scaled to the maximum of $|S_{21}|$ across all powers. The power axis of all feed-line datasets were trimmed to 20 by removing the lowest power samples – and accordingly any powersweep with bias point in that range. The spectral axis was trimmed to a window (50 samples) centred on the resonance at each power. It was found that this step increased the accuracy of

the CNN by several per cent despite the loss of information about the magnitude of the resonance translation with readout power. The location of the resonant frequency was taken as the location of maximum v_{IQ} (at any power sample, fewer than 1% of resonators have a maximum v_{IQ} that is not within 50kHz of the resonance, as chosen by MI, due to excessive noise).

The label vector for each powersweep is a probability density function (PDF) with a lognormal distribution and a maximum at the MI bias point. The lognormal profile accounts for the fact that it is more detrimental to overpower a resonator and lower powers are sometimes preferable. These labels only classify a powersweep based on readout power. A CNN could be conceived that intelligently tunes both the bias point and frequency by means of a multilabel classifier architecture. This concept is saved for later work.

Data augmentation is the process of creating additional training data by performing label preserving transformations on the original training data, or synthesizing new data based on a model. Data augmentation was explored to both increase the total amount of training data and correct for the non-uniform distribution of training data for each class. These label preserving transformations included resonator phase transformations (changing the ratio of Q and I) and shifting the input data in the spectral domain by several units. No appreciable increase in accuracy was observed, and the extra training data increased the training time of the CNN. The inability of augmented data to increase the CNN accuracy is indicative of either, inconsistencies in the original training data, or the fact the CNN has learned to be insensitive to these transformations. Synthesizing powersweeps based on the AM model, and assuming some relation for each of the parameters with power, may have increased the accuracy of the CNN.

3.3.3 Training and Evaluation

The CNN was implemented using the TENSORFLOW machine learning library [170] on a computer with 16 Intel Xeon cores running at 2.2 GHz. Initially, the weights of the CNN were allocated normally distributed random values. For each training step,

batches of up to 50 input powersweeps were fed into the CNN, and the difference between the true and predicted classifications (known as loss) was measured. The weights are then adjusted in the direction that minimizes the loss, which is found using the ADAM gradient descent method [171]. As this process is repeated and CNN begins to converge on the optimal set of values, the magnitude of the correction applied to the weights was exponentially decreased to prevent overshoot.

For each investigation, separate *evaluation*, *training* and *testing* datasets were created. The evaluation dataset contains all the measured powersweeps from a single feed-line, selected for the purpose of evaluating the CNN against the other bias selection methods. The remaining nine feed-line datasets are pooled together and split (typically 95:5) to produce the training and testing data respectively. The test data helps evaluate the performance CNN on unseen data from the same feed-lines.

Ensemble techniques are where multiple models are trained separately, usually on the same training data, and combined to achieve a greater accuracy. For example, bagging involves training different models in parallel, and since the trained models are non-deterministic, the combination of the predictions should act to cancel some of the biases each model has from overfitting. This technique was used on the full CNN in Section 3.4.1.2.

3.4 Results and Analysis

A misclassication by 1 dB is often permissible because of the subjectivity that arises for some powersweeps. For example, there is some subjectivity inherent for some pathologies such as 'No Discontinuity' or 'Shift Loss' powersweeps. Another type of subjectivity is due to uncertainty in the bifurcation power because of the limited power sampling. In the event of the bifurcation power existing between two samples, a more aggressive operator may choose a higher bias point.

Therefore, a Boolean classification accuracy within 1 dB is used to assess the performance of CNN (or any of the described evaluation methods). First, the Boolean

accuracy is measured for each powersweep p according to

$$a_p(m_\mathrm{e}|m_\mathrm{t}) = \begin{cases} 100, & \text{if } |b(m_\mathrm{t}) - b(m_\mathrm{e})| \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

where b is the bias class, m_e is the evaluation method (here CNN), method m_t is assumed to produce the true values. Then, the mean a_p for all powersweeps in a feed-line dataset, yields the accuracy parameter of m_e given m_t

$$A(m_\mathrm{e}|m_\mathrm{t}) = \frac{1}{P}\sum_{p}^{p=P} a_p. \tag{3.3}$$

3.4.1 CNN Training

3.4.1.1 Input Data Investigation

To investigate the consistency of the feed-line datasets, the CNN was trained and evaluated on different combinations of them. The evaluation dataset was sequentially set to each feed-line and a CNN model was trained and tested on a combination of the powersweeps from the remaining feed-lines. Each time 3565 and 180 powersweeps were randomly selected for training and testing respectively. Each CNN was trained for 800 steps of 50 batches and then evaluated. Figure 3.5 displays each CNN's performance.

The average train data accuracy parameter is higher than the average test or evaluation data accuracy parameter , suggesting that the CNN could benefit from additional training data or more aggressive regression techniques. The evaluation accuracies for Face3ii, Face3iii and Face2i are appreciably low. The Face3ii and Face2i powersweeps had a large amount of I and Q noise because of an artifact in the digital readout – Face3iii and Face2ii are the repeat measurements after the resolution of the problem. Face3ii and iii contain higher frequency resonators which also results in more I and Q noise. Interestingly, the evaluation accuracy parameter achieved with Varu2 is among the highest despite the CNN having no experience with powersweeps

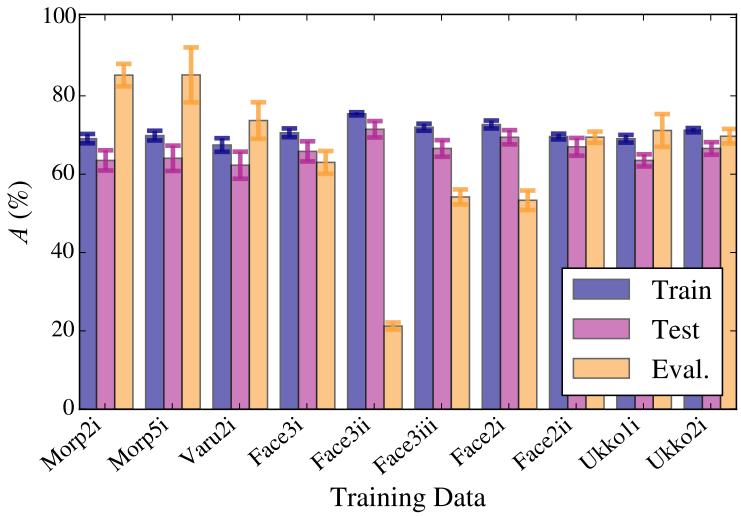

Figure 3.5: The CNN was trained on 3565 powersweeps from nine of the feed-line datasets, and using those powersweeps as m_t (Train), 180 unseen powersweeps from the same feed-lines as m_t (Test), and all of the powersweeps from the remaining feed-line dataset as m_t (Evaluation), the accuracy parameter A was evaluated from the CNN's predictions (m_e). This process was repeated for each of the ten feed-line datasets in turn. When Ukko2 was used as the evaluation dataset, the classifications from first of the four MI operators were used, $m_\mathrm{t} = \mathrm{MI_A}$. Each time the CNN was trained for 800 steps of 50 batches. These measurements were repeated ten times and the errorbars are one standard deviation from the mean of the accuracies achieved. The evaluation accuracy parameter achieved with Ukko2 indicates that that dataset is fairly representative of all feed-line datasets.

from that array, indicating the utility of the CNN for future arrays.

The evaluation accuracy parameter of Ukko2 is not far the mean achieved across all feed-lines, indicating that the difficulty of that feed-line is fairly representative of all feed-line datasets measured. Furthermore, Ukko2 is preferable for the comparison between the bias selection methods because it has a uniform distribution of bias points. For a typical feed-line, the distribution of bias points peaks around the fifth bias, because the resonators are designed to be as identical as possible. If the CNN overfits to these classes it will be most apparent in Ukko2 measurements. Therefore, Ukko2 was chosen as the evaluation dataset for all remaining investigations.

3.4.1.2 Full CNN

To investigate the training data requirements, subsets of non-Ukko2 powersweeps were randomly selected for training at increasing amounts, and a new CNN algorithm was trained each time. The testing data were also non-Ukko2 powersweeps, but the full 212 powersweeps were used for each CNN. This process continued until the CNN was trained on all 3965 powersweeps for 1200 training steps. The batch size for each step matched the amount of training data until this passed 50 powersweeps, then the number of batches remained at 50 for each step. The accuracies achieved at each training step, on each subset, are displayed in Figure 3.6.

There is a large range of final accuracies on the evaluation dataset achieved by different instances of the CNN. The ensemble technique is therefore very advantageous in guaranteeing that the optimum accuracy is achieved. It also has an appreciable impact of accuracy parameter $A = 5\%$ above the final mean.

When investigating the effect of the amount of training data, initially the training accuracy is 100%. At these amounts of training data, the number of parameters in the CNN is sufficient to essentially store the powersweeps, rather than develop a model to make predictions on them. Similarly, the test and evaluation accuracies, when using low amounts of training data, have comparatively large variety. The plateaus at large amount of training data (and training steps) indicates higher accuracies require more standardized training data or more aggressive regularization. Interestingly, using the ensemble technique on just 200 training powersweeps is sufficient to reach over accuracy parameter $A = 60\%$.

3.4.2 Accuracy Comparison

The MI feed-line datasets will suffer from human error because of telescope deadlines, lack of experience from some operators as well as the subjectivity inherent in some types of powersweeps. In order to create a more accurate dataset for comparison, the evaluation feed-line was classified four times, each time by a different operator. To account for the subjective systematic offsets between the four MI datasets, A

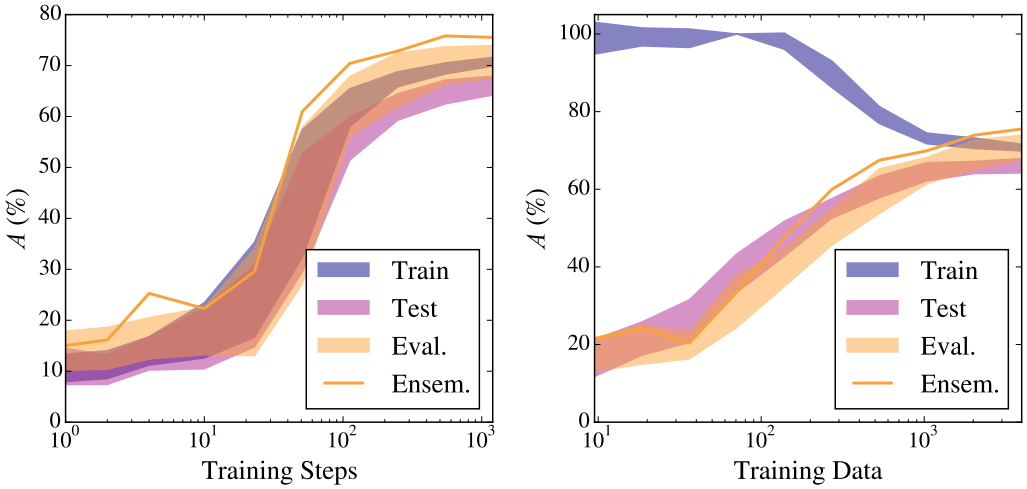

Figure 3.6: The CNN was trained on increasing amounts of randomly selected powersweeps from the non-Ukko2 datasets, and using those powersweeps as m_{t} (Train), 212 unseen powersweeps from the non-Ukko2 datasets as m_{t} (Test), and all of the powersweeps from Ukko2 as m_{t} (Evaluation), the accuracy parameter A was evaluated from the CNN's predictions (m_{e}). For the Ukko2 labels, the first of the four MI operators were used, $m_{\mathrm{t}} = \mathrm{MI}_{\mathrm{A}}$. For each training step 50 powersweeps from the training dataset were randomly selected. The Ensemble curve takes a median of all the predictions made on the evaluation dataset. left: The accuracy parameter of the CNN on the full amount of powersweeps from the training, testing and evaluation datasets, measured at different amounts of training steps. right: The achieved accuracy parameter after full amount of training steps on different amounts of training data. The extent of the filled regions are one standard deviation from the mean, of the range of accuracies achieved when repeating each set of measurements ten times, hence why A can go above 100%. The ensemble accuracies on Ukko2 demonstrate that sufficient accuracies can be achieved on minimal amounts of training data.

was measured for each of the six combinations. It was found that increasing half of labels of $\mathrm{MI_D}$ by 1 dB (increasing the average by 0.5 dB) maximized the total A of the six combinations. Three of these datasets were then combined by taking a median of the classifications for each powersweep, to produce the $\mathrm{MI_{av}}$ dataset. The remaining dataset, $\mathrm{MI_B}$, remained independent from $\mathrm{MI_{av}}$, to evaluate the MI performance against the automated methods. If any of the operators decided a powersweep was insufficient for classification, the powersweep and label were omitted from the evaluation dataset. This took the amount of powersweeps down from 377 to 340.

Confusion matrices are two-dimensional histograms used to compare the labels evaluated by two different classifier algorithms. If m_e agrees with m_t for a given powersweep, that powersweep (a true positive) will lie on the $b(m_\mathrm{e}) = b(m_\mathrm{t})$ diagonal. The false positives are all the values in the corresponding row, and the false negatives lie along the corresponding column (both excluding those located on the true positives diagonal).

Figure 3.7 shows the confusion matrices for two MI datasets, the results from the four methods when compared to the 'true' dataset $\mathrm{MI_{av}}$, and a new method that uses predictions from both NM and AM. In the top left of each panel is the accuracy parameter A. The uncertainty on A was taken to be the standard error from the a_p distribution, and the uncertainty from the quantization error in classification, added in quadrature.

$\mathrm{MI_B}$ tended to classify powersweeps more conservatively than $\mathrm{MI_A}$ by approximately 0.2 dB. This inconsistency in the training data places a limit on the attainable accuracy that the CNN can achieve.

Despite the inconsistency of the training data, the accuracy parameter achieved by the CNN was 90%, which is similar to the randomly selected operator, $\mathrm{MI_B}$. (If operator A, C or D is used as the independent operator and $\mathrm{MI_{av}}$ is recreated accordingly, the maximum difference between the accuracy parameter achieved by the CNN and MI is 2%). Only a handful of powersweeps are classified with more than 1 dB above $\mathrm{MI_{av}}$ ensuring that the vast majority of resonators are not bifurcated and will correctly operate as photon detectors. The accuracy of the CNN method is also

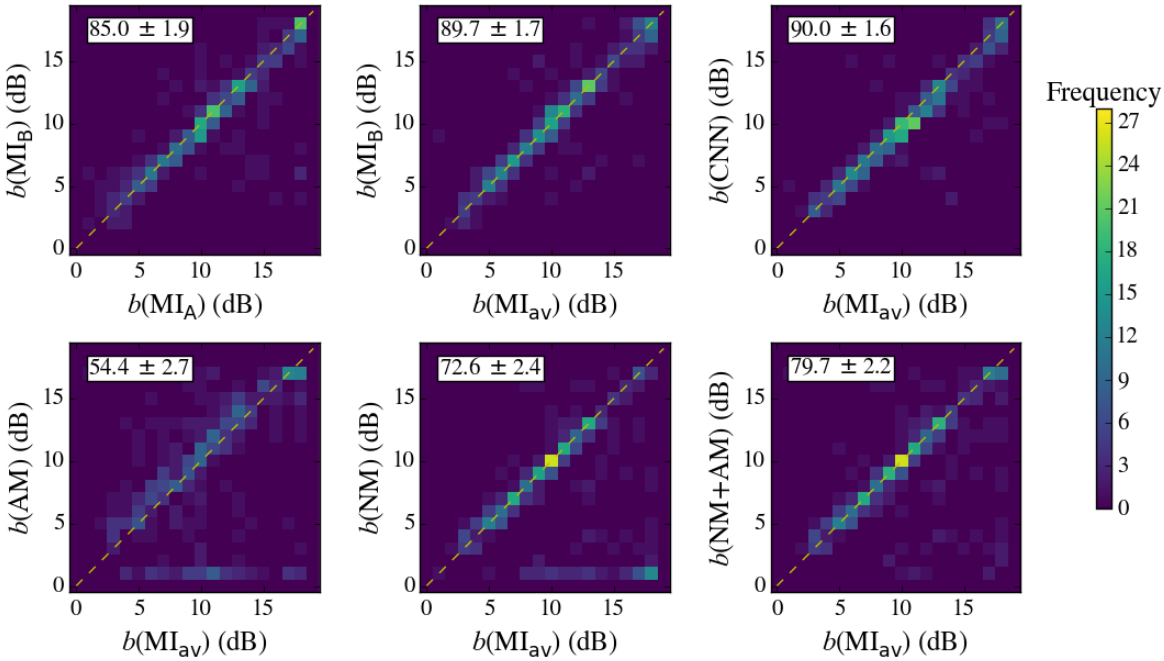

Figure 3.7: Confusion matrices comparing the evaluated bias points from different methods. $\mathrm{MI_{av}}$ is created using three of the four operators and is assumed to have the true bias values. $\mathrm{MI_A}$ and $\mathrm{MI_B}$ were the first two operators to use MI to classify the evaluation dataset. b(NM+AM) refers to a method that combines some of the predictions from both methods. The colour of the datapoints corresponds to the number of powersweeps in that bin (frequency). The dashed yellow diagonal line are true positive classifications. The accuracy parameter A, shown in the top left of each plot is the percentage of powersweeps that lie within 1dB of the true positives diagonal.

substantially higher than AM and NM, and showing no bias towards certain optimal operating points. The reason for the apparent performance improvement of the CNN between Figures 3.6 and 3.7 is because of the elimination of 37 of the questionable powersweeps from, and the improved accuracy of, the $\mathrm{MI}_{\mathrm{av}}$ labels compared to the MI_{A} labels, which are used as m_{t}. This makes all the methods appear more accurate. (If $m_{\mathrm{t}} = \mathrm{MI}_{\mathrm{A}}$, as it was in Figure 3.6, the CNN's $A = 76\%$ is still 25% and 12% above that achieved by AM and NM, respectively.)

AM classified 20% of powersweeps with >1 dB power above the $\mathrm{MI}_{\mathrm{av}}$, which would have likely rendered them unusable. Both AM and NM tend to falsely classify powersweeps in the second lowest bias. The additional noise in I and Q at these powers can trigger a false result on different pathologies (as shown in Figure 3.3). If this occurs in any of the lowest four power indices, because of the choice of implementation of Rule #1 (described in Section 3.2.4), the bias point prediction will be this value.

NM achieves an impressive 47% of powersweeps along the true positives diagonal. This could be attributed to the fact that the MI operators used the NM predictions as a first guess, and so for 'adequate' powersweeps (49.9% from Table 3.2) the operator would tend not modify those predictions. In an effort to extract the best predictions from both AM and NM, a new model was created that substituted the predictions from AM when $b(\mathrm{NM}) = 1$ dB. This method achieves an accuracy parameter $A = 80\%$, at the cost of the extra computation time of AM compared to NM.

Figure 3.8 shows all possible confusion matrices from the eight predictors. When comparing all six pairs of MI datasets the A ranged from 79% to 85% for the confusion matrix shown in Figure 3.7. The primary reason for the disagreement between operators is the subjectivity of some classifications. The CNN provides more conservative estimates than MI_{A}, MI_{B}, MI_{C}, but the A scores remain similar to those obtained from confusion matrices directly comparing MI pairs. The AM-NM matrix has false predictions along the $\mathrm{p(AM)} = 1$ column and $\mathrm{p(NM)} = 1$ row. This demonstrates that methods misclassify different types of resonators, which explains why the NM+AM technique, comparatively, is so effective.

Figure 3.9 shows how the different type of powersweep effects the accuracy pa-

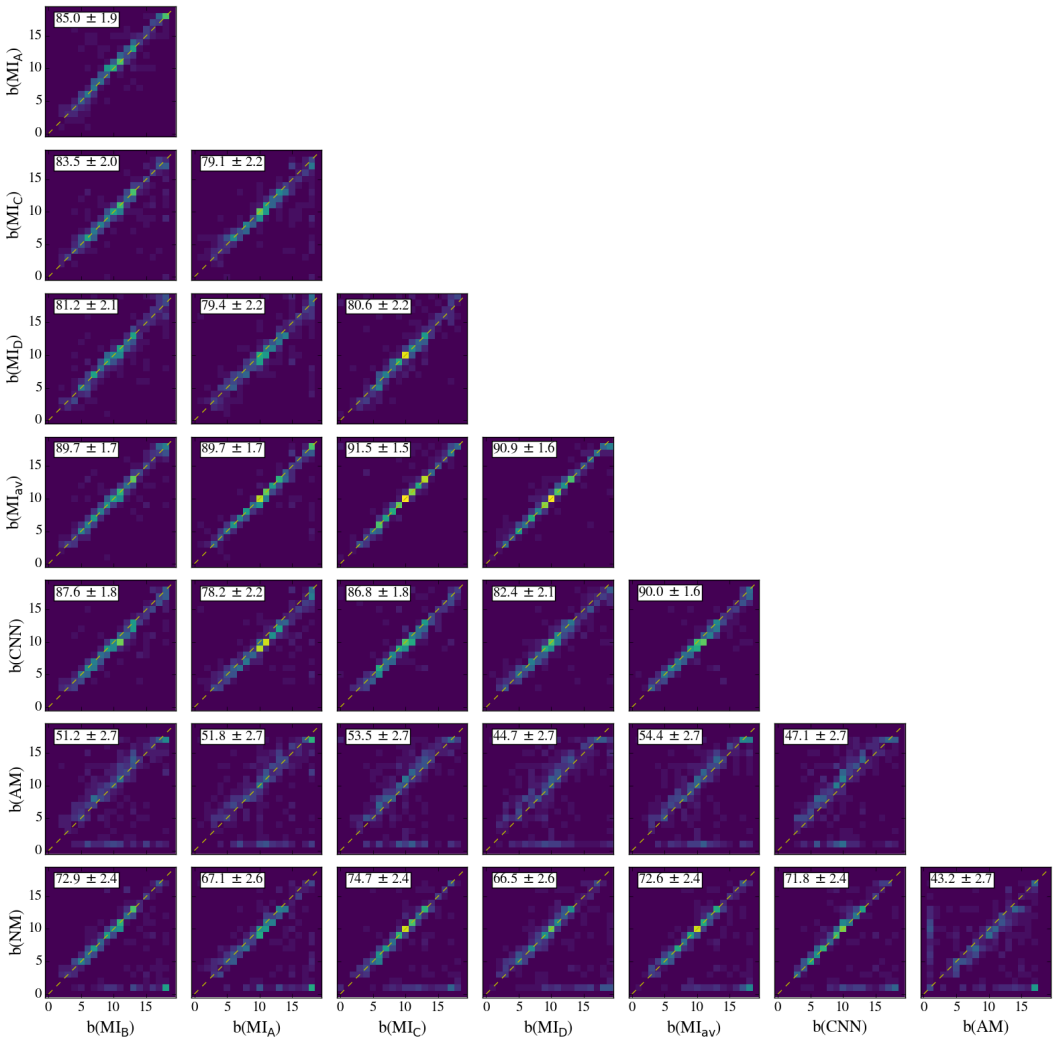

Figure 3.8: Confusion matrices comparing the consistency of the evaluated optimal power indices from different methods. The color of the datapoints corresponds to the number of resonators, or frequency, in that bin. The dashed yellow diagonal line are true positive classifications. The more resonators than lie close to the true positives diagonal, the more the two methods are in agreement. A quantative measure of this is the effective accuracy, A, which is shown in the top left of each plot. MI_A and MI_B were the first two operators to use MI to classify the evaluation dataset. MI_{av} is created using each of the four operators and is assumed to have the true optimum values. A corner plot displaying the confusion matrices for all the pairs can be found in the supplementary material.

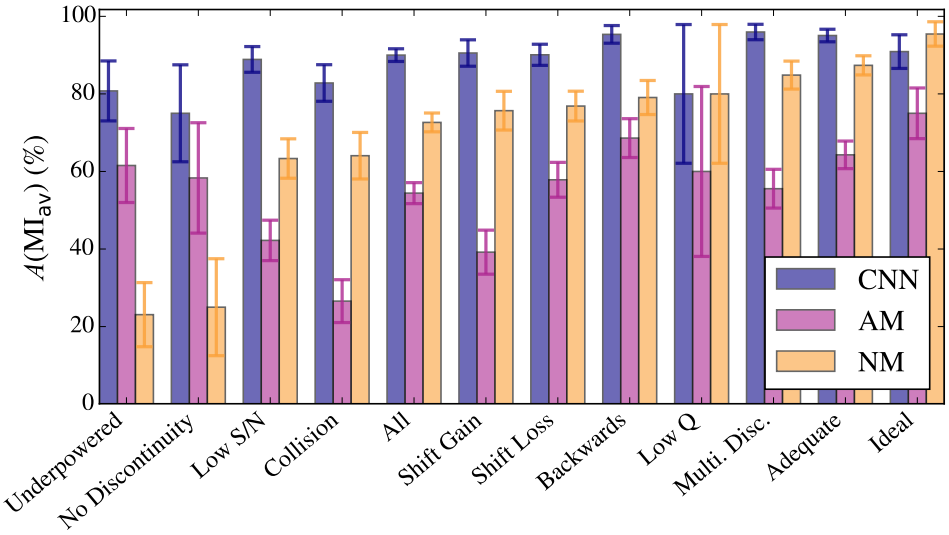

Figure 3.9: The accuracies of each of the automated methods when $m_{\mathrm{t}} = \mathrm{MI_{av}}$. The categories are the same powersweep types shown in Table 1. The associated error is the standard error from the a_p distribution and the reading error added in quadrature.

rameter of each method. The accuracies of the AM and the NM are highest for ideal powersweeps as expected. The NM is vastly superior on powersweeps showing multiple discontinuities, since these are not accounted for in the AM model. Conversely, the AM does better on underpowered powersweeps and those with no discontinuity possibly because the AM technique can identify other features for the onset of bifurcation other than the formation of a discontinuity. It is apparent that the average fit with the collision model for AM was insufficient. The accuracy of this method might therefore benefit from Markov Chain Monte Carlo sampling of the larger parameter space and maximum likelihood estimation for the parameters, at the expense of increasing the time to perform the method even further.

The CNN has superior accuracy parameter, and smaller uncertainty, across all but two powersweep types. This is partly because of the higher sophistication of the model, but also because, on powersweeps such as 'underpowered', the model has the advantage of having been exposed to these often-featureless powersweeps and learned the convention of classifying them as a high bias point.

Figure 3.10 shows the confusion matrices for $\mathrm{MI_A}$ $\mathrm{MI_{CNN}}$, $\mathrm{MI_{AM}}$, $\mathrm{MI_{NM}}$ when

compared to MI_{av} separated by the type of powersweep. As expected, the abundance of predictions at b = 1 with MI_{AM} and MI_{NM} is not present in the Ideal confusion matrices. These misclassifications do not appear to come from one type of pathology for either method. The confusion matrices for CNN and MI_A are essentially featureless across all powersweep types. (MI_A and CNN can be seen to favour certain types of powersweeps, but this simply manifests as a broadening of the distribution of predictions around the true positives diagonal rather than a more interesting feature.)

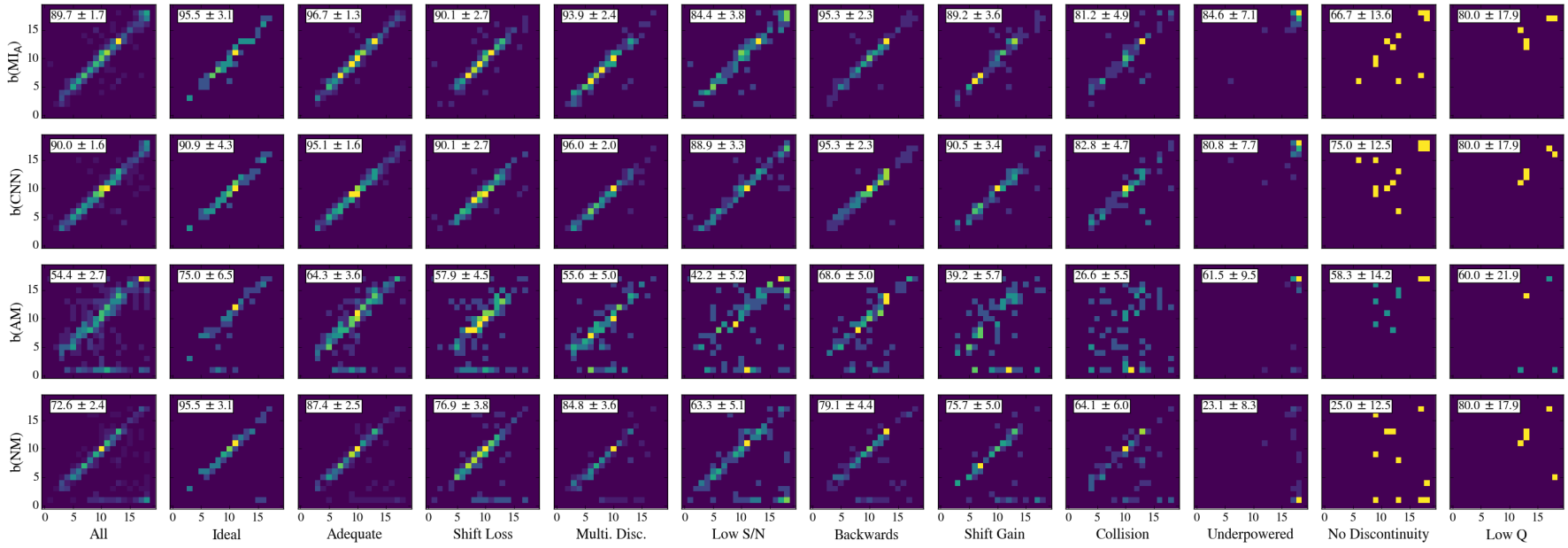

Figure 3.10: The confusion matrices of one of the MI operators and the three automated methods when compared to $\mathrm{MI_{av}}$ and separated according to resonator behaviour.

Table 3.4: The operator and computation times are for biassing 1000 pixels (half a feedline). The observation run total ('Obs. Run') assumes that the full 10,000 pixel array would need to be completely tuned two times, and that retuning is 75% faster for AM and MI. The timing was performed on a 16 core Intel Xeon processor. It was assumed that the AM method could be optimized to achieve the same accuracy parameter in 20% of the time. Rate is calculated from the time it takes to tune 1000 pixels and the accuracy parameter A. The timing each method are rounded to the nearest significant value. The time to take the power-sweep measurements are not included.

Method	Operator	Computation	Obs. Run	A (%)	Rate (m^{-1})
MI	2h	30s	25h	90	7
AM	15m	1h	16h	54	7
NM+AM	1m	11m	3h	80	66
NM	–	3m	1h	73	242
CNN	–	1m	20m	90	900

3.4.3 Timing Comparison

Table 3.4 shows predicted times to tune a ten-kilopixel array using each of the described methods by extrapolating the times values determined on Ukko2. 'Rate' is a measure of the number of correctly classified powersweeps per minute if tuning an array for the first time. For a typical observing run, some feed-lines will require multiple retunings while others require only one round of tuning. In other runs the device may need to be swapped, or there needs to be alterations to the digital read-out hardware or software settings. Depending on the severity of the changes to the resonator properties, the retuning can be made faster by using values from the first tuning round as starting points. It was found that retuning, using either AM or MI, can be performed as quick as 25% of the time of the initial tuning process. This speed enhancement has been applied to those methods for the retunings.

For the AM, some operator time is required to identify which powersweeps need to be fit with the resonator collision model (18 parameters instead of 9). A first guess can be acquired by fitting all powersweeps with the standard AM model, and then manually checking those with the goodness-of-fit parameter beyond some threshold.

The timing and the accuracy parameter displayed for the NM include the low pass filtering of v_{IQ}. The method can be made three times faster at the expense of roughly

accuracy parameter $A = 20\%$. When using MI, each powersweep is first classified using the NM for a first estimate to guide the operator. The majority of the time spent classifying powersweeps is then by the operator.

For the CNN method, the time shown is the time to classify the powersweeps. The creation of the training set and subsequent training would be done prior to an observation-run, and both processes do not have to be repeated for each new device, so those times are not included in the total observing run time. The CNN can be made ten times faster by not applying the ensemble averaging, at the expense of approximately accuracy parameter $A = 5\%$. The time to train the CNN with 1000 steps on the full training dataset took just three minutes. Furthermore, trained algorithms can be made publicly available to be used by other research laboratories[2].

For a typical telescope observing run, the total time spent tuning powersweeps with the CNN algorithm is a factor of 75 faster than using the current technique of MI, and correctly classifies more powersweeps per minute than any other of the investigated methods.

Analytical models make accurate predictions within their assumptions. Outside of this, either a more sophisticated model or human input is required. A neural network operates in the middle ground between these two techniques: it creates a very sophisticated (3×10^8 parameters) model using the human intuition (and accordingly any human error) encompassed in the training data. It finds the most optimal parameters to solve any type of powersweep as efficiently as possible, for a given number of training steps.

It has been shown that a relatively simple CNN (total training time of three minutes) can characterize a kilo-pixel MKID array in under 1 minute, with approximately equivalent 1 dB accuracy to the method of MI, and fewer over powered classifications, on a feed-line where the majority of powersweeps show non-ideal power handling behaviour. Roughly equivalent accuracies can be achieved on completely unseen devices, and competitive accuracies can be achieved on just 200 powersweeps when using ensemble training. The CNN markedly outperforms the alternative automated methods across all powersweep types and classes. This accuracy is achieved in spite of the in-

consistencies present in the training data. This further justifies the implementation of a rigorous automated algorithm to optimize the quality of MKID observational data.

The CNN algorithm has been used to set up the device and readout for each of the four previous DARKNESS observation runs at Palomar, saving potentially 100 human hours. The more and reliable training data that is accumulated, the more accurate this MKID tuning algorithm should become, which is implemented as part of the open source DARKNESS Digital Readout pipeline[2]. The ability to reliably automate the biasing of kilo-pixel arrays is vitally important as the community works towards the eventual goal of realizing mega-pixel MKID cameras [123]. Additionally, experiments can now be conducted on current devices which were unfeasible before, such as: measuring the change of the bias point of resonators across multiple feedlines with the CNN, after recycling a adiabatic demagnetization refrigerator; or using the CNN to tune a full array, and using photon data, measure the sensitivity at a range of powers around those bias points, to investigate the validity of Rule #1 on a statistically meaningful sample of resonators, free from human error.

3.5 Future Plans

Automating Training Data

The performance of the CNN is ultimately stifled by the subjectivity in MI classifications of training data. One approach to this problem is to create powersweeps using Equation 2.5 (used in fitting the AM model). Each parameter will have some dependence on power, which will have to be measured and included in the model. For example the ideal resonator in Figure 3.1 appears to have an exponential a dependence with power. Similarly, g_0 and Q_i also appear to scale with increasing power.

Once this model has been created arbitrarily large amounts of this training data can be made with ideal power handling behaviour to train the CNN to reach ∼100% accuracy on these types of resonators. Similarly, if needed, some of the pathologies

[2]https://github.com/abwalter/MkidDigitalReadout

can be included in the training data also. It would make sense to do this if the CNN consistently made very inaccurate predictions on these pathologies (sometimes overpowering resonators) and it was believed that this type of pathology will be present in future arrays.

It has been demonstrated that the CNN can achieve sufficient accuracy on unseen arrays [1]. However, it may be the case that new pathologies are introduced in future arrays. A pipeline could be conceived that uses the predictions of the 'AM+NM' method to check the CNN's predictions. If the methods do not agree then the powersweep may contain a pathology and should be checked by an operator before being added to the training data. If the methods agree then the resonator is assumed to be ideal and is added to the training data. As a quality control measure to verify that there isn't some inconsistency between AM+NM and CNN being introduced into the training data, 10% of the A >90% arm could also be sent through b(MI)

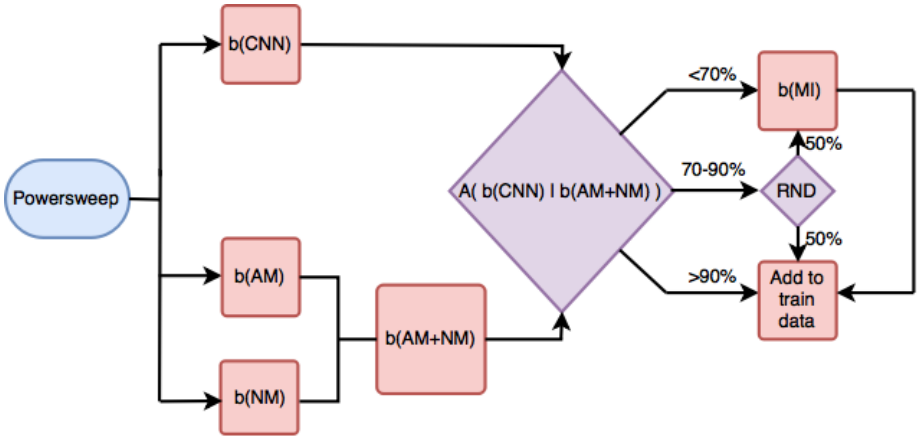

Figure 3.11: A toy model of the decision process for adding powersweeps to the training data pool. As a safe guard against the CNN introducing biases into its training data, random samples of the powersweeps that score A >90 can also be checked by MI.

Drive Power and Frequency Dual Optimization

Currently only the bias power is optimized through a machine learning algorithm and the bias frequency is determined by the location of the maximum value of v_{IQ}. As

stated in section 3.3.1, this assumption is less valid for resonators with low SNR in I/Q. It is also problematic for shift gain and loss resonators, collisions and power-sweeps showing multiple discontinuities.

As well as being useful, this algorithm should be relatively simple to implement. The feed-line datasets that made up the training (and evaluation) datasets for the original CNN model also contain optimal frequencies chosen from MI. The new model would therefore use the same training data, architecture and be developed in much the same way as the original.

The required algorithm would be a multi-label (and multi-class) classifier. The TENSORFLOW NN (neural network) library has this capability already implemented with the SIGMOID_CROSS_ENTROPY_WITH_LOGITS function. This function calculates the cross entropy (the same loss function used for the original CNN model) for independent and mutually exclusive classes

$$E = y - yy' + \log(1 + \exp(-y)) \tag{3.4}$$

where y is the label predicted by the CNN (before the Softmax activation is applied), and y' is the true label. Similar to the original CNN, the true labels would be a list of binary values, however now several elements can take on the value '1' as opposed to just a single element.

Spectral Resolution Optimization

The original CNN model was optimized using MI data that attempted to apply Rule #1 to powersweeps. This criterion was assumed to provide the maximum SNR of the phase response. An 'optimum' CNN model could be devised that is instead, optimized the labels for spectral resolution by maximizing the responsivity. A higher responsivity of the phase response would mean that longer wavelength photons could be detected above the phase noise floor. This would give the instrument sensitivity in a wavelength regime where the contrast requirements are often more favourable for high contrast imaging. Furthermore, having a larger bandwidth for the instrument

would also benefit speckle reduction algorithms.

The optimum label for a given resonator would be inferred by illuminating the pixel with with several laser lines, performing a bias powersweep, and measuring R at each power. The location of the R maximum would be the optimum power for that resonator. This process would be repeated for many resonators in order to generate the training dataset for the CNN model. In contrast to the original CNN model, this method of generating training data has the advantage of not requiring human input for each resonator, enabling the generation of much larger training datasets.

With this method, the CNN architecture would devise a model that is tailored to all parameters of a resonator rather than just the bifurcation power (as in Rule #1) and the exceptions (Rules #2-4). For example, the quality factor has a relationship to the pixel's spectral resolution. Due to the nature of machine learning algorithms, this CNN model will exploit resonator properties in a complex and non-intuitive manner, potentially finding new relationships between the parameters and providing more consistent predictions on unseen resonators.

3.6 Summary

In this chapter we introduced the digital readout system that allows thousands of MKID pixels to be probed simultaneously. We noted that in order to take observations with MKIDs, the readout tones must be tuned to the parameters of each pixel, namely the bias power. In order to automate this process a machine learning algorithm was devised based on a convolution neural network architecture. The accuracy of this algorithm was shown to be comparable to manual inspection, and the rate of resonator biasing was vastly faster. In the next chapter we will discuss how data is processed after observations.

Chapter 4

Calibrating MKIDs Observations

In this chapter we introduce the general pipeline procedures for reducing and calibrating raw photon packets produced by the MKID readout. The primary data product of the pipeline is a 4D observational *hypercube*, which has two spatial, one wavelength and one time dimension. We then describe calibration hardware that I developed for DARKNESS and KIDSpec. The DARKNESS calibration source was created during my LTA at UCSB.

4.1 Introduction

4.1.1 Data Format

When taking observations with MKIDs, each pixel is independently and continuously monitored. This means that a different approach is required to CCDs whereby the whole array is read via shift registers after each exposure or CMOS with non-destructive readout. A data reduction pipeline for MKIDs was first created for ARCONS [7] and a pipeline for DARKNESS is currently under development [3]. We will explain the differences between the pipelines in the relevant sections.

Each photon event triggers the creation of an array of values by the readout. This array contains the photon arrival time, the pixel index, the maximum of the phase response and the background phase of the incident pixel. With ARCONS,

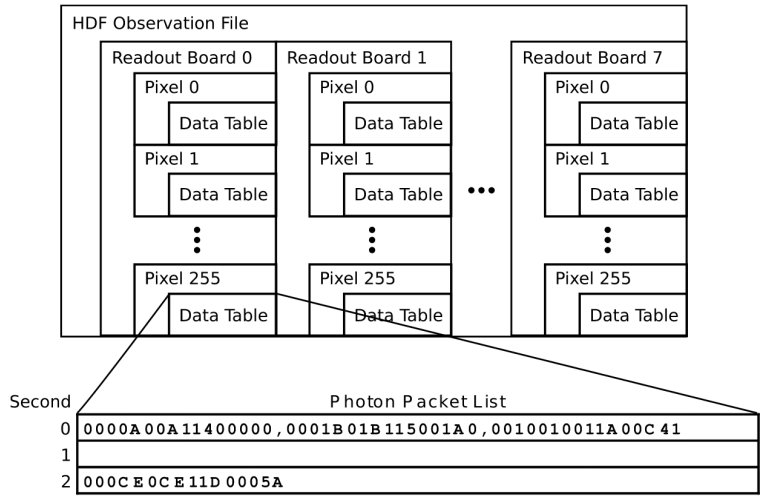

Figure 4.1: The layout of an 'observation file', originally from [7]. The data is organized first by readout board number and then by pixel number. Each pixel has a data table where rows display all detections within a certain second and each column displays a data packet of the detections that occurred in that second. The data packets are 64 bit long and consist of the pixel index, phase shift, baseline phase and a timestamp of the detection.

these photon data packets were stored in hierarchical data format (HDF5), grouped by pixel, for a duration of the observation, in what was called an *observation file*. With DARKNESS, these photon data packets are stored in binary format, with no sorting by pixel, for one second intervals to allow necessary higher data-rates for real-time communication with the xAO (this one second interval is defined in the readout software and will be reduced in later iterations). The first step in the reduction of DARKNESS data is therefore, to locate and collate the relevant binary files into an observation file. The structure of these files is shown in Figure 4.1.

As a quick aside, the MKID exoplanet direct imaging simulator (MEDIS), which is introduced in the next section, imitates observation data taken with an MKID system. It creates observation files of the same file format that can then be fed into the MKID pipeline. The plan is for MEDIS to be integrated into the MKID data pipeline to compliment observations.

4.1.2 Data Reduction and Calibration

Beammap files are recorded for each array that map each pixel's resonant frequency to its physical location on the array. A 2D array is populated with the counts at the locations determined by the beammap in order to create an image (or a datacube/hypercube).

In theory, MKIDs do not have a dark current because the lattice phonons have insufficient energy to break the Cooper pairs. In practice, due to unoptimized design and fabrication, there will be some false triggers on some pixels from excessive phase noise. Therefore, dark maps are created periodically throughout an evening of observing and these are subtracted from exposures.

Currently, MKID arrays suffer from random false measurements of very high counts rates or *hot pixels*. These were especially problematic in ARCONS because of an interaction between resonators and electrons in the silicon wafer. The transition to PtSi in the DARKNESS arrays appears to have removed the cause of this problem. However, hot pixels can still occur in DARKNESS arrays because of imperfect readout parameters. For example, if the phase threshold is set too low, the pixel will remain constantly hot. Another example is if two readout tones are too close together producing a periodic hot pixel at the beat frequency. These hot pixels can be identified and masked by comparing pixel intensities to their neighbours [7] or flagged with dark maps.

MKID arrays require flat-fielding to account for intra-pixel variations in responsivity, quantum efficiency, and transmission anisotropies in the optical train. This calibration is performed as a function of wavelength as well as the two spatial dimensions. A uniform image is created by observing the inside of the observatory dome or the twilight sky. The counts in each pixel are binned according to wavelength, and the median count rate across all pixels and each bin is calculated. Conventionally, the wavelength bin widths are equivalent to 0.1 eV (24 THz), but in principle, these bin widths can be user defined [7]. The ratio of these values to the median for each pixel produces the weights with which to scale each pixel and wavelength.

The phase response of MKIDs has a monotonic relationship with photon energy, which provides the inherent spectral resolution capabilities. To perform this calibration, the array is flood illuminated with several lasers of known wavelength and a phase histogram is constructed for each pixel. The result is several peaks with FWHM determined by the spectral resolution of the pixel at that wavelength, and a height determined by the number of registered counts during the exposure. An example spectrum is shown in Figure 4.2. The location of these peaks is measured by applying a Gaussian fit to each peak. This list of phases is then fit with a second order polynomial to determine the continuous relationship between phase and photon energy for that pixel. The width of these peaks is an estimate of the energy resolution. In the example of Figure 4.2, these widths correspond to R = 8.1, 6.3, and 5.8 at λ = 808, 980 and 1310 nm.

4.1.3 Analysis Calibration

The absolute spectral response of the array is performed by observing a spectrophotometric standard star. The measured datacube is dark subtracted, masked, flat-fielded and the median spectrum of all pixels is calculated. This spectrum is then divided into the known spectrum of the source to create the *sensitivity spectrum*[1]. Science spectra are multiplied by the sensitivity spectrum to produce absolute flux calibrated data.

To create a datacube with right ascension and declination coordinates assigned to the spatial axes, the plate scale and rotation of the array is first measured by comparing reference stars between images. Flexible Image Transport System (FITS) is a digital file format for arrays and the associated metadata which is stored in the header. The CRVAL keyword in a FITS header controls the RA and dec of a reference pixel, and the CD keyword controls the scaling allowing RA and dec values to be assigned to all pixels.

After the photon data have been reduced and calibrated, the user has an observation hypercube (two spatial, one wavelength, and one time dimension) at their

[1]http://www.eso.org/sci/observing/tools/standards/spectra/

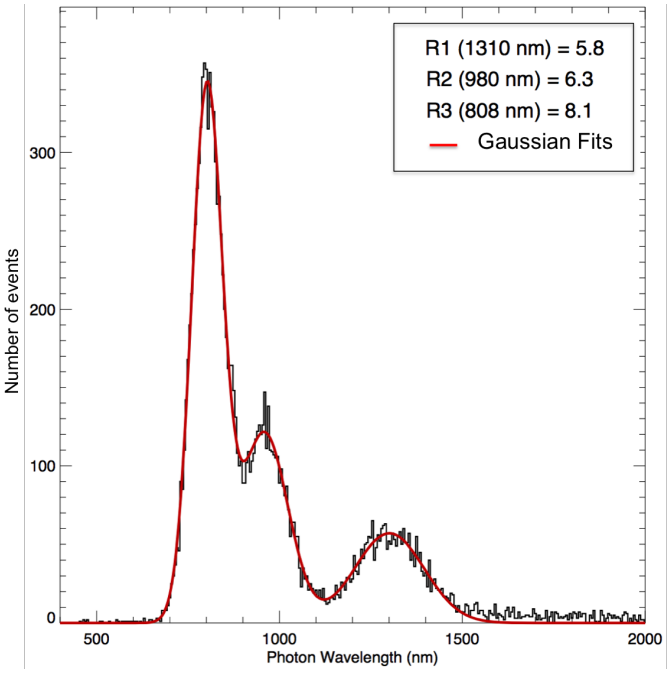

Figure 4.2: Histogram of the phase response of a single pixel when illuminated with the DARKNESS wavelength calibration source. This pixel is illuminated with the 808, 980 and 1310 nm lasers (laser lines 920 and 1120nm here not shown). From these peaks it is possible to calculate the function that relates phase response to photon wavelength and the spectral resolution spectrum for this pixel. This figure was modified from [2].

disposal. The wavelength and time bin widths are arbitrarily chosen after the observation. By selecting all photons with timestamps within a chosen time window and stacking the results according to the pixel index, an *effective* exposure is created. The optimum duration of these effective integrations can be determined ex post facto. Many effective integrations can then be combined to perform HCI post-processing techniques such as differential imaging (more detail of HCI post processing techniques are found in Chapter 5).

If during an observation, the pointing of the telescope is dithered, then several effective integrations can be stacked to form a mosaic. The resulting intensity variations are normalized by scaling each pixel by the total effective integration time of that pixel. Not only does this type of observation create a larger image, it also creates a smoother image as dead pixels in one effective integration are likely to line up with live pixels in another, removing these discontinuities.

Combining spectra from several pixels is complicated by the fact that the spectral resolution can vary considerably between pixels. One solution to this is to take the lowest spectral resolution of all the MKID pixels that contribute to a given mosaic pixel and create a Gaussian profile corresponding to that resolution. Then, convolve the spectra of the other MKID pixels with appropriate filters to degrade all spectra to the lowest resolution. In the time domain, light curves at a RA and Dec coordinate are created by concatenating the data from different pixels with an appropriate time delay. Depending on the required time resolution is is typically sufficient to just combine the time streams without performing any convolutions

4.2 Wavelength Calibration Source Development

4.2.1 DARKNESS

Since DARKNESS is sensitive to longer wavelengths than ARCONS a new calibration source was required. When creating a wavelength calibration source (also known as laser box), the following objectives have to be considered:

1. Uniform spacing between the wavelengths of the laser diodes

2. Portability

3. Remote and local control

The uniform spacing of the lasers is so that several laser diodes can illuminate the array at once with minimal overlap of the wings of the Gaussian profiles. (However, if an array shows especially poor spectral resolution, the laser diodes can be run individually, and the location of the peak identified, free of interference from adjacent laser lines.) The DARKNESS band extends from 800 to 1500 nm. In the 800 - 1100nm band several standard laser wavelengths exist that make procurement fast and inexpensive. Lasers between 1100 and 1300 nm are harder to come by, unless the purchaser is willing to pay for the wavelength to be precisely tuned. Ultimately lasers at 808, 920, 980, 1120, 1310 nm were chosen with the 1120 nm laser costing more than the others combined. The power of the lasers are roughly 10 mW placing them in the 3B British Standard and so eye protection is required when adjusting the lasers.

The calibration source is operated both in a laboratory environment when testing arrays, and on the telescope when taking observations. This necessitates portability and flexibility in the control. We created a printed circuit board (PCB) to compactly house the circuitry that controlled each of the laser diodes.

Figure 4.3 displays the circuit diagram of one laser diode channel, and the PCB configuration for all channels. A series resistor controls the magnitude of current reaching each laser diode. To complete this circuit and activate the laser diode, there are two possible routes. A mechanical switch controls one of the circuits, which is used for local control. For remote operation, a transistor-based switch controlled by an Arduino microcomputer controls the second circuit. In this mode, commands are sent via User Datagram Protocol (UDP) through an Ethernet connection. A simple control script activates each of the lasers for a defined period of time. This script was integrated into the DARKNESS readout software to perform wavelength calibrations between observations at the Palomar Observatory.

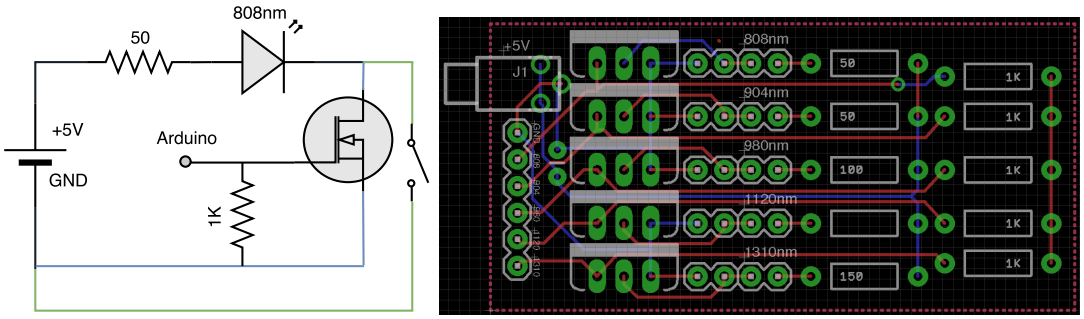

Figure 4.3: Left: a circuit diagram of the 808nm laser diode channel. The green circuit is the local control operated with a mechanical switch, and the blue circuit is the remote control operated through an Arduino microcontroller and a MOSFET transistor. The 1KΩ resistor operates as a pull-down resistor to ensure 0V reading from the Arduino remains at 0V. Both resistors are displayed in ohms. Right: each of the 5 channels and associated connectors placed placed using the PCB design software Eagle. The terminal labelled 'Arduino' in the top schematic is the '808' connector in the bottom schematic.

For the DARKNESS calibration source, the laser diodes mounted directly onto the PCB. Since the beam of the laser diodes extends across a wide solid angle, the rigidity of the laser diode legs was enough to keep them sufficiently aligned (see Figure 4.4 for the positioning of all the components in the enclosure). An integrating sphere mounted on the lid of the enclosure captures and sends the light down a fibre.

When looking at the laser lines using a spectrometer it became apparent that some of them shift between 1-10 nm sporadically a few times a minute. When ramping the laser diodes over the lazing threshold current the center frequency of the lines did shift, so one theory was that the current was close to the lazing threshold and the laser diodes were transitioning constantly. This theory was ruled out by driving one of the laser diodes at its max ratings and still the frequency jumps remained. Electrostatic damage was ruled out by replacing some of the laser diodes and taking every precaution to not damage the replacements. Not allowing the lasers sufficient time to warm up was ruled out by leaving them on for several tens of minutes. It became apparent that the lack of heat sinks for the lasers was probably causing the switching effect because the lasing wavelength is highly dependent on temperature. The spectral resolution of DARKNESS is sufficiently low that the switching effect is

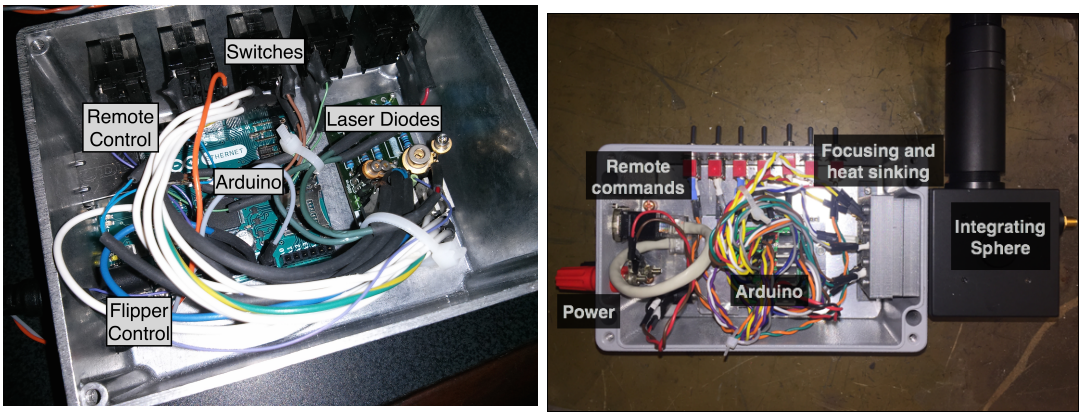

Figure 4.4: Photo of the calibration source with the lid and integrating sphere removed. The flipper mirror control plays no part in the operation of the calibration source. One of the digital output ports of the Arduino was used to remotely control a flipper mirror in the optics while observing.

negligible (see Figure 4.3). If the spectral resolution of DARKNESS arrays becomes sufficiently high then heat sinking hardware will have to be installed – such as the design used in the KIDSpec laser.

4.2.2 KIDSpec

The KIDSpec band, though currently undefined, will extend to shorter wavelengths to cover the visible regime. This means that there was scope for more laser lines to perform the calibration. This has the added benefit of implementing extra redundancy if some of the lasers show the switching effect present in the DARKNESS calibration source. Lasers at 639, 780, 980, 1310, 1550 nm were chosen for KIDSpec. Additionally, a white light emitting diode (LED) was also integrated for broadband tests of KIDSpec.

The KIDSpec wavelength calibration source also featured a 3D-printed laser mount. This part served serval purposes: it provided a heat sink for the laser diodes, it was detachable to make installation of the laser easier, it aligned the laser beams and ensured all of the light from each laser reached the integrating sphere. The last purpose was especially important as the lower wavelength lasers were more powerful and

therefore more likely to damage the user's eye. The power of the lasers ranged from 5 - 50 mW placing them in the 3B British Standard (same as the DARKNESS laser box).

4.3 Summary

In this chapter we introduced the data processing pipeline for calibrating the raw data from the MKIDs readout. We also described the development process of the wavelength calibration sources for DARKNESS and for KIDSpec. Since wavelength calibration data is required periodically throughout an evening of observing, laser boxes must have remote control capabilities. We now provide an example of data analysis of on-sky observations using the ARCONS data reduction pipeline.

Chapter 5

Simulating and Processing HCI Observations

In this chapter, we describe a package that I developed to simulate HCI observations with MKIDs (as well as an ideal detector or a H2RG) called MEDIS. We then use MEDIS to demonstrate the HCI processing techniques to take the processed photon data and create scientific products – such as 5σ contrast curves.

5.1 Motivation

To achieve the parent-planet contrast ratios and separations, necessary for small mass terrestrial planets, requires fast (ideally photon counting), large format arrays (tens of kilopixels), high QE, and low noise (sub-electron read noise and dark current) detectors. These properties both improve the level of speckle noise suppression through conventional methods with minimal speckle evolution, and enable new techniques of speckle suppression that should remain effective at small inner working angles.

DARKNESS [3], MEC [157] and PICTURE-C [159] promise to rigorously demonstrate these transformative benefits. However, it may be a number of years before these instruments are fully optimized and realized at their full potential (for example detector development is an ongoing process). To compliment these systems we have developed a general purpose end-to-end numerical simulator for high contrast

observations with MKID-based HCI systems – the so-called MKID Exoplanet Direct Imaging Simulator (MEDIS). With MEDIS we can analyse MKID-based systems in a controlled manner under a variety of conditions. We can then investigate how different parameters impact performance of the current and near future systems.

MEDIS differentiates itself from other AO simulators [172, 173, 174, 4], firstly by the fact that it is tailored to HCI and speckle reduction. For example, it has various coronagraph designs, real-time speckle control techniques, and post processing techniques. These techniques operate on wavefront errors that are introduced from a variety of sources at a range of time-scales (static, quasi-static and dynamic speckles). Secondly, MEDIS also provides a pragmatic depiction of MKIDs (and to some extent H2RGs). MKIDs are single-photon detectors, which means that a slightly different approach is required for system modeling than the conventional intensity mapping. Furthermore, MKIDs display various forms of non-ideal behavior that will ultimately limit their efficacy as detectors for HCI. MEDIS has the ability to introduce these non-ideal behaviours and quantify their effect.

5.2 Implementation

Figure 5.1 shows a schematic of the structure of MEDIS. There are four key modules, these are: Atmosphere, Telescope, Detector, and Analysis. The atmosphere module creates wavefront error maps that change rapidly between time steps according to chosen turbulence parameters. These maps seed the telescope module that creates images of the science focal plane at several wavelengths according to various optics parameters. The detector module then injects its own artifacts that would occur during the sensing of the datacube. In the case where MKIDs are the chosen detector, these deteriorations include things like variations in dead pixels, whereas if H2RGs are the chosen detector, then readout noise and dark current are introduced.

MEDIS treats light as both a wave and a particle depending on module in question. As the light propagates through the telescope, it is treated as a wavefront. If the detector is a H2RG, then the wavelength datacubes produced by Telescope at

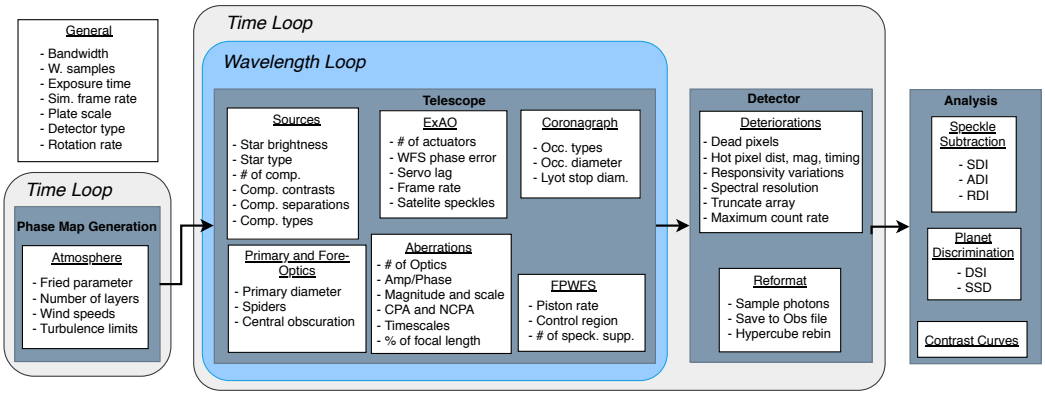

Figure 5.1: An overview of MEDIS. For each time step: the Atmosphere module creates a wavefront error map for the telescope entrance pupil; Telescope converts this and creates a focal plane intensity map at several wavelengths; this wavelength datacube seeds Detector that produces a typical HCI observation; Analysis converts the observation into science data products. The rounded squares represent loops.

each time-step are combined into a hypercube before being saved to be read in with Analysis. If the detector is an MKID array, then the wavelength datacubes are used as PDFs to source individual photons. These photons are then either added to an ObsFile or recombined into a datacube before being concatenated to other datacubes to form a hypercube in the same way as the H2RG example.

MEDIS was designed to leverage parallel processing. It does so by simulating each time step on a different computer core if they are available. Some options of Telescope may serialize the processing of different time steps (for example when an AO servo lag is introduced) but if the frames are independent then MEDIS will process time steps in parallel according to the number of cores. Figure 5.2 depicts this process for 6 time steps. An empty hypercube is created at the beginning of the simulation and then the appropriate location is populated with each datacube according to the time step index. When an ObsFile is created, the photon packets contain timestamp data, and so the photons can be ordered in post processing.

For a 40 core machine, running different time steps in parallel, a 1 hour observation can typically be simulated in approximately 5 minutes. Parallel processing is also utilized for some resource intensive algorithms in Analysis. For these algorithms each

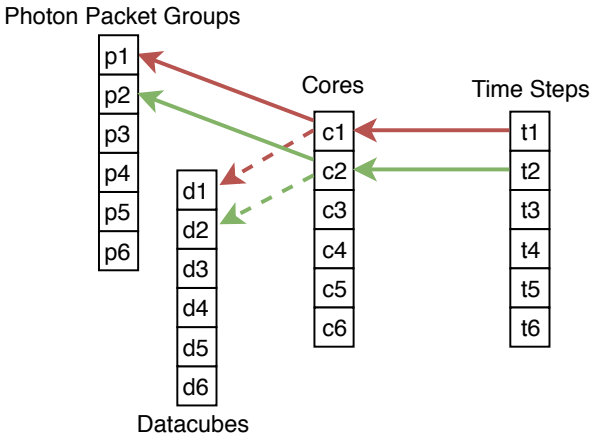

Figure 5.2: This diagram illustrates the parallel processing involved with simulating an observation. For a 100 s observation and a sampling rate of 1000 Hz, there would be 10^5 iterations, handled by up to 48 cores.

pixel is operated on in parallel.

5.2.1 Atmosphere

The AO software CAOS was utilized for the atmospheric wavefront error map generation. This software was selected because it uses multi-layer model for the atmosphere and it has the ability to model the evolution of the atmosphere with time.

In the 1940s Kolmogorov developed a theory of how energy is transported from large scales of turbulence (at $\sim$ 10-100 m) to smaller scales via the creation of progressively smaller eddies until the energy could be dissipated via the viscosity of the medium (at $\sim$ 1 mm) [175]. This model has been validated for atmospheric turbulence at La Palma [176]. The outer scale was measured for Paranal by Ali et al. [177] (the future location of the ELT, Cerro Armazones, is $\sim$ 30 km away from Paranal).

The Kolmogorov phase power spectrum is given by

$$\Phi_K(k) = 0.023 r_0^{-5/3} k^{-11/3} \tag{5.1}$$

where k is the wavenumber and r_0 is the atmospheric coherence length. From this we

can get the phase structure function according to [178]

$$D_\phi(r) = 2\int_0^\infty \Phi(k)(1-\cos(2\pi kr))dk. \tag{5.2}$$

The phase structure function is fractal, which means that it appears the same at all scales. Together the phase structure function, $D_\phi(r)$, and amplitude structure function, $D\chi(r)$ describe the total variance of the wavefront. However, the amplitude component is negligible compared to the phase component for the scales of telescope apertures used.

The issue with generating wavefront error maps from the Kolmogorov spectrum is that the large scale turbulence (those with low spatial frequency) extends beyond the width of the telescope aperture. These low frequencies are responsible for the positioning of the centroid in the focal plane and they have a large impact on the performance of the coronagraph. However, modeling both the large and small scales simultaneously for each phase screen is very resource intensive. CAOS compensates for the low-frequencies in the Kolmogorov spectrum by incorporating sub-harmonics of the aperture to more finely sample the Fourier space close to the origin [179].

Atmospheric turbulence can be assumed to be localized to a number of thin layers that move independently, and the turbulence within those layers remains frozen for the duration of the observation. The refractive-index structure constant, C_n^2, measures the amount of inhomogeneity in the local refractive index and it essentially quantifies the strength of the turbulence. It has been shown that layers located just above the telescope entrance pupil have the largest C_n^2, and this value roughly decreases with increasing altitude [177, 180] as shown in Figure 5.3. Similarly, the velocity of the higher altitude layers is large compared those just above the telescope.

These layers are generated using an ATM module. Table A.1 shows the selected values for each layer of turbulence. CAOS allows the user to control the time samples of the atmosphere evolution as well as r_0 (at 500 nm). Setting $r_0 = 0.2$ along with the parameters in Table A.1 gives an evolution time constant of the turbulence of 8.6 ms. This means that sampling the atmosphere every 1 ms is sufficient to capture the

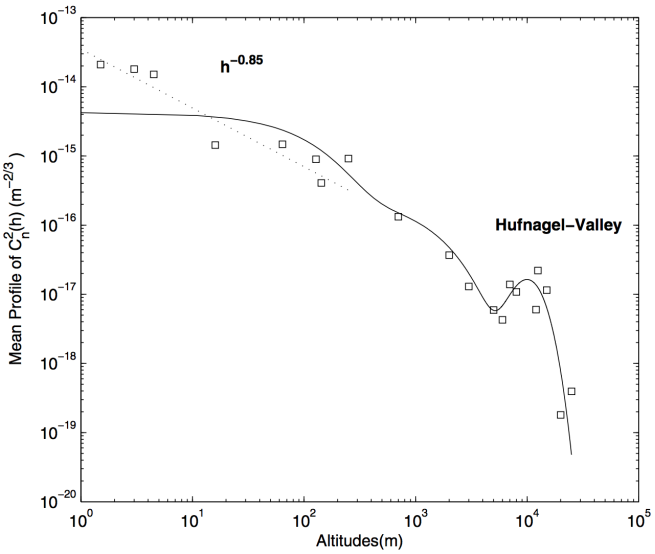

Figure 5.3: The mean profile of C_n^2 as a function of height for Paranal in 2007 [177]. The measurements are plotted with square markers and the Hufnagel-Valley fit to those data points [181]. Also included is a simple power law mode for the first meters from the surface.

full evolution of each of the layers without aliasing. In these simulations, the outer scale of the Kolmogorov spectrum was set to infinity.

These layers are then combined to produce a single phase screen as observed at the telescope aperture. To do this, a wavefront is created corresponding to a source at infinity with the SRC module. This initially unaberrated wavefront is then geometrically propagated through each of the layers and masked according to a defined telescope aperture. Figure 5.5 shows an example of these phase screens at

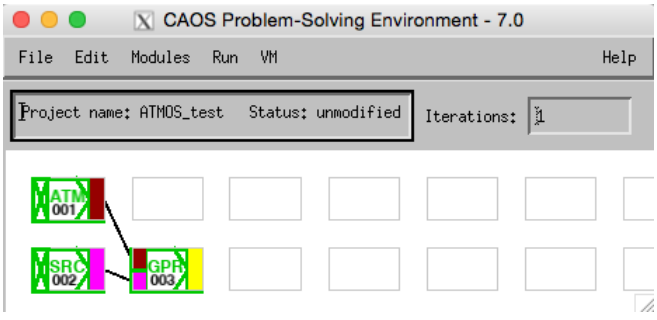

Figure 5.4: The simple block diagram constructed for the purpose of creating the atmosphere aberration maps. Each block has configurable parameters. The GPR block was modified to output the wavefront in FITS format.

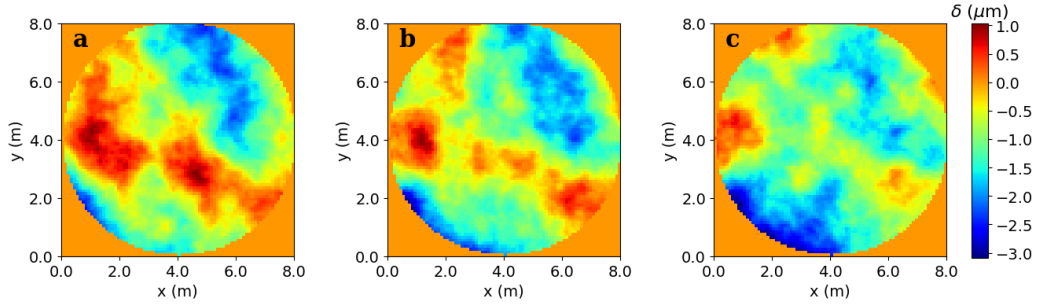

Figure 5.5: The resultant phase screens at the aperture of the telescope produced by CAOS for a telescope diameter of 10 m and $r_0(\lambda = 500\text{ nm})$ of 0.2. a: is taken at some arbitrary time. b and c are 0.1 s and 0.2 s after **a** respectively.

three instants.

CAOS is written in IDL whereas the rest of MEDIS is written in PYTHON. GPR was adapted to output the phase screens as FITS files for all time steps so that they can be read by Telescope. This way all phase screens for the duration of the simulation can be created in one go before being stored. This saves the overhead of the wavefront error map generation and the transfer between IDL and Python each time MEDIS is run.

5.2.2 Telescope

5.2.2.1 Near and Far-field Propagation

The Telescope module leverages the PROPER AO software [4]. PROPER was selected because it predicts diffraction effects of the wavefront, in both near-field and far-field conditions, as it propagates through the system. This is vitally important when attempting to numerically model the effect of different aberrations from various sources within a telescope system. Optical trains will often contain aberrations outside of pupil plane, introducing significant chromaticity into the speckles, which makes techniques SDI less effective [65].

Modeling a wavefront showing curvature on a rectangular grid can be problematic, since the phase change between adjacent pixels can be more than 2π. This causes aliasing in the phase, leading to spurious results. Instead, a reference surface is

selected (either planar or spherical), depending on whether the beam is in the near or far-field. These regimes are determined by a Gaussian pilot beam (Gaussian intensity profile) that is analytically propagated through the system.

PROPER provides a number of routines that perform a modification to the phase and/or amplitude of a wavefront. For example multiplying the wavefront by a circular obscuration or altering the curvature of the wavefront after it has passed through a lens (or equivalently a mirror). A full optical system is then simulated by combining lots of these simple routines in series.

5.2.2.2 System Overview

Figure 5.6 shows a schematic of the main elements in the Telescope module that we constructed. In the most simple configuration an unaberrated wavefront is masked by the telescope entrance pupil, focused, and the image is sent to Detector. For a full simulation: the atmosphere phase screens are read-in and applied to the wavefront created by PROPER. Then the telescope spiders and central obscuration are applied before the aperture mask of the entrance pupil. The user has the option to apply additional CPAs before the application of the xAO. Then the user can apply additional NCPAs on different timescales. After focusing the wavefront, real-time speckle control can be applied via the DM in the adaptive optics module to correct the light on the next cycle (or some user defined delay). The next group of elements make up one of several types of Lyot coronagraph architectures. The last element accounts for low order aberrations present in detector's fore optics.

In order to create a companion object a separate wavefront is generated by PROPER in addition to the on-axis one. The amplitude of this wavefront scaled to match a user-defined intensity contrast with the primary, and the direction of the wavefront is adjusted to match a user defined radial separation of the companion source in the focal plane. The direction of the wavefront is controlled by adding tip and tilt Zernike polynomial terms to the wavefront. This wavefront then passes through the same elements and is effected by the same aberrations (including those from the Atmosphere module) as the primary wavefront. Multiple wavefronts with

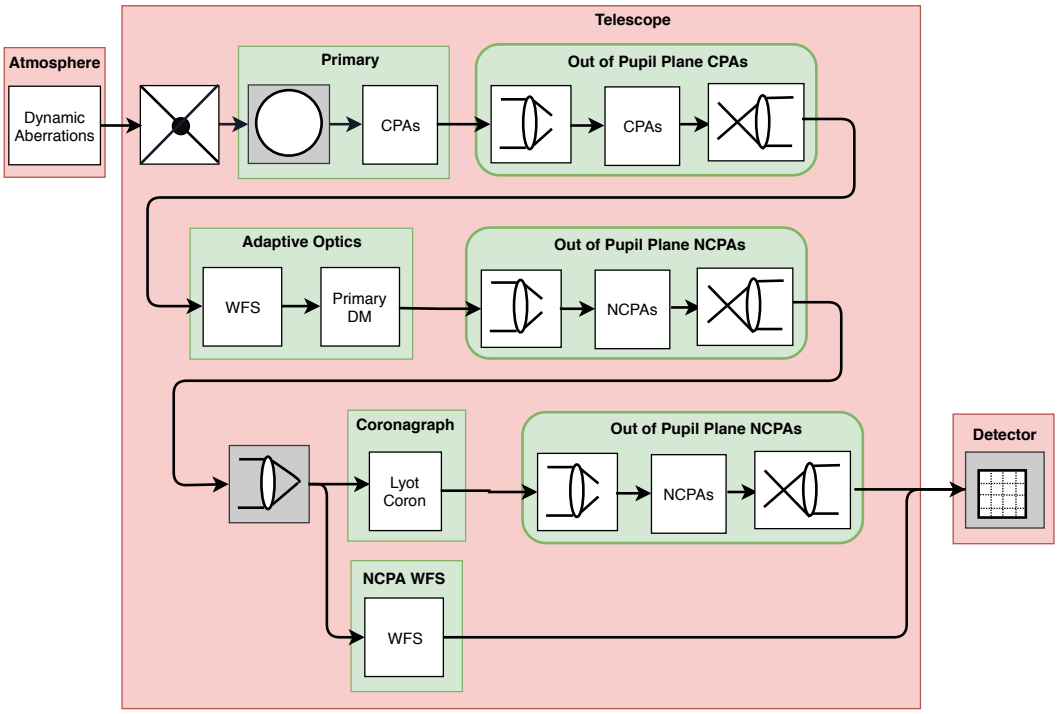

Figure 5.6: A schematic of the main elements of the optical train in Telescope. The greyed blocks are essential. Atmosphere and Detector are separate modules.

different directions are created if multiple companions are required.

If the user would like to perform ADI, then the sky observation must rotate relative to the aberrations from the optics. Both the atmosphere phase screen and companion object are made to rotate while the optics aberrations remain stationary. The atmosphere phase screen is rotated with a rotation matrix transformation, while the position of companion objects is made to move by changing the relative weighting of the tip and tilt Zernike polynomial phase contributions.

In order to take observational data at several wavelengths and perform SDI, the optics aberrations must be appropriately radially scaled. Since PROPER changes the spatial sampling (m/pix) proportionally with wavelength, the intensity map in the focal plane appears identical at all wavelengths. Instead the sampling must remain constant for all wavelength. This is achieved by scaling the beam-diameter to grid width ratio, B, inversely with wavelength. At the shortest wavelength, B, must be large enough that the focal plane is Nyquist sampled. The amplitudes of each of the

resulting monochromatic images can then be scaled according to the transmission spectrum of the system and incident spectrum of the stars.

5.2.2.3 Aberrations

Each of the types of aberrations were characterized by Equation 1.4 in Chapter 1 (shown again here)

$$\Psi_1(x,t) = [A_P + A_S(x) + a_d(x,t) + a_{qs}(x,t)]P(x) \tag{1.4}$$

A_P is determined by the size of the aperture mask, and a_d was determined by the parameters in the Atmosphere module. The remaining components, A_S and a_{qs} are created by sampling from a phase power spectrum, similar to the Atmospheric aberrations. However, these power spectra are less well defined than the Kolmogorov power spectrum, Φ_K. In MEDIS, the optic aberration maps are generated from the following profile [182]

$$\Phi_O(k) = \frac{a}{\left[1 + \left(\frac{k}{b}\right)^2\right]^{\frac{c+1}{2}}} \tag{5.3}$$

where a is the power at low frequencies, b is the correlation length (the same units as k) and defines the location of the knee of the curve, and c dictates the gradient of the curve in log space at high frequencies. Figure 5.7 shows the power spectrum used for creating aberration maps with MEDIS. The parameters values are in the top right of the figure. These values are in accordance with the model used for the VLT in Bordé and Traub [86].

As with Φ_K, the inverse power law accounts for the fractal nature of the aberrations and that the majority of errors are typically at low spatial frequency. The distribution flattens at low frequencies depending on the scale of the tools used when figuring and polishing the mirror [182].

The inset of Figure 5.7 shows a map generated with these parameters. The scale of the features match those of the measured data for the VLT mirror (Figure 1.6) reasonably well. In the real map, the features tend to follow concentric rings, as a result of

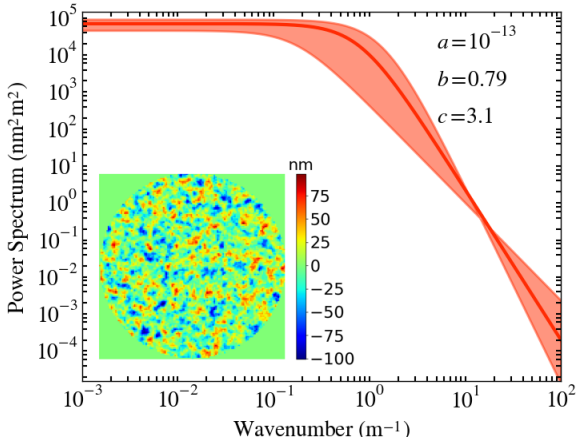

Figure 5.7: The power spectra used to seed the aberration maps. The quoted figures represent the mean input values of all optics. The shaded area is the limits achieved when drawing distribution parameters from Gaussian distributions. The parameters for this curve are provided in the figure. Inset: One example of a wavefront error map generated from the power spectrum.

figuring and polishing. In the simulated map they are randomly distributed because of the lack of phase information when seeding from a 1D power spectrum. Amplitude aberrations are less well characterized in the literature and are here assumed to follow the same profile PSD with $a = 0.05$.

As with the atmospheric aberration maps, these maps are separately and collectively created and stored in FITS format, prior to the simulation of the HCI photon data.

The user has the option to apply several aberration maps at different distances between the pupil plane and the focus. This is controlled with the OOPP and n_surfs keywords of the aber_params dictionary. OOPP is a list of values that control at what fraction of the focal length each aberration map is applied. If n_surfs is smaller than the length of OOPP then only the first n_surf distances are used.

Quasi-static aberrations are created using the same power spectrum parameters except several maps are required. According to Martinez et al. [183, 184], at each point in an image, the RMS error linearly increases with time after an initial frame. This means that reference images decorrelate with time and provide a progressively worse subtraction. To replicate this, we have the array of phases that seed the Fourier

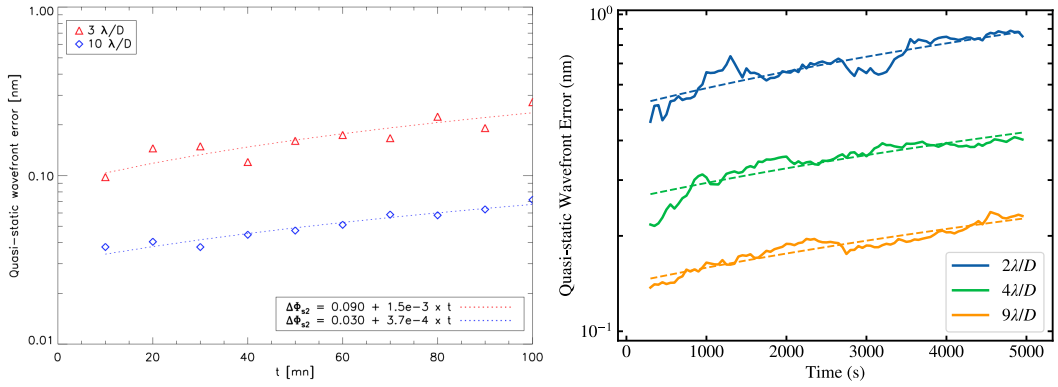

Figure 5.8: The wavefront error of a differential image increases linearly with time as the quasi-static aberrations evolve and the reference image decorrelates from the later images for SPHERE (Left; [184]) and for MEDIS (right). Furthermore the wavefront error increases for smaller radial separations.

transform perform a random walk with each iteration. The conditions used by Martinez et al. [183] were replicated in MEDIS, i.e. the system had no atmospheric aberrations, and a series of differential images were created from an initial reference image. Then, using Equation 9 of Martinez et al. [184]

$$\Delta\phi_{s2} \approx \frac{\lambda}{2\pi\sqrt{6}} \times \frac{\sigma_{\mathrm{DI}}}{\sqrt{I_c}} \tag{5.4}$$

the wavefront error for several radial separations was calculated. From Figure 5.8 it can be seen that the difference in the wavefront error from the initial frame approximately linearly increases over long periods of time.

5.2.2.4 Adaptive Optics

The phase (and amplitude) of the wavefront is already exactly known in the simulation software. Therefore, in order to simulate a measurement from the WFS, the wavefront phase is recorded and a Gaussian uncertainty is applied for each pixel. The default uncertainty is several percent.

This 2D wavefront error map is then interpolated to a grid with the dimensions equal to the number of DM actuators. The actuator height required to correct these errors is then half their value since a reflection doubles the path length.

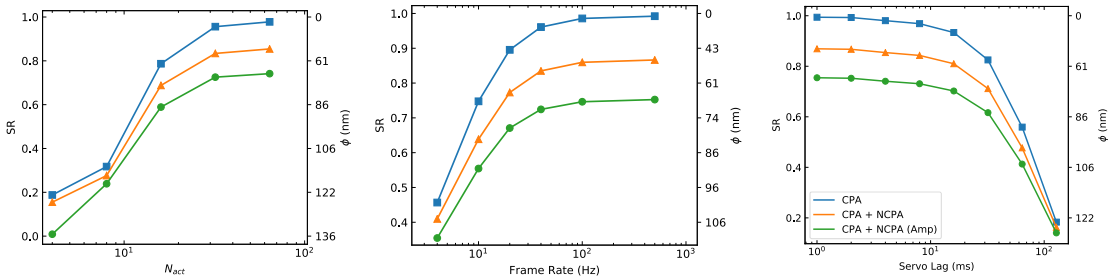

Figure 5.9: The level of correction of the AO system as different aberrations are introduced as a function of number of actuators (a), servo frame rate (b), servo time lag (c). The aberrations values are those in Table A.3. quick_ao is set to False.

Because the surface of a DM is semi-rigid, the movement of an actuator has an effect on the surface height of adjacent regions. PROPER has the functionality to reach the required heights according to the Xinetics Photonex DM influence function [185]. A Gaussian uncertainty is then applied to this array to simulate piston error.

In the default mode (quick_ao = True), the DM then applies these values and modifies the current wavefront. The WFS is placed before the DM to facilitate this technique. For a better correction, the AO module can store the DM values, measure the new residual wavefront error and update the DM values accordingly to converge on the minimum wavefront error (at the expense of serializing the simulation processes and dramatically increasing the simulation time).

Similarly, to introduce servo lag error, a time datacube of WFS measurements is created. For each time-step, the WFS measurement is placed in the datacube at the index corresponding to the time delay. For a time delay $\tau = 10$ ms and the time-step $t = 1$ ms, this would be the tenth index. The phase map at the first element is read by the DM. All the phase maps then shift one element in the time axis. For $t < \tau$, there has been insufficient time for the phase maps to propagate to first index and the DM receives an array of zeros. At $t \geq \tau$, the DM receives phase maps from time $t - \tau$. To introduce bandwidth error, the phase maps from several elements of the time datacube are mean-averaged before being sent to the DM.

Light from companions is incoherent with light from the primary for real observations. Therefore, the WFS only uses the phase measurements from the primary star.

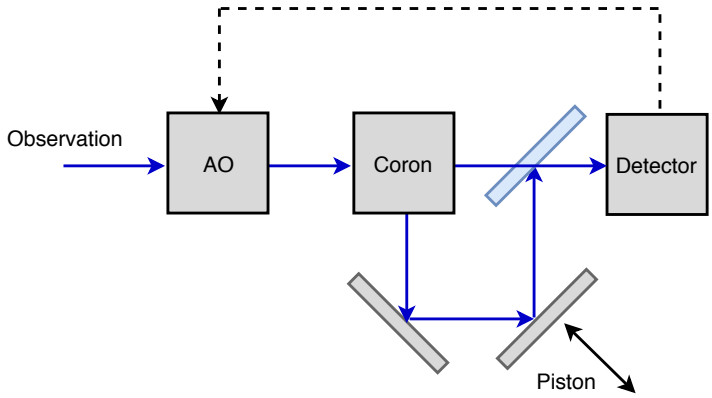

Figure 5.10: A schematic method used to detect the aberrations occurring after the AO but before the coronagraph. The planar wavefront interferes with the focused observation wavefront. The phase across the wavefront is periodically shifted with a piston. After four iterations a phase map of the observation can be determined. The type of detector used as the WFS detector can be ideal H2RG or MKID.

The DM correction is applied to both types of object.

5.2.2.5 Post-AO Wavefront Sensing

The FPWFS capabilities of MEDIS were demonstrated in Figure 1.14. To measure the phase in the focal plane, the phase shifting interferometry method from Bottom et al. [91] has been implemented. This is where some light picked off from the observation (or a laser source), is pistoned to create a phase shift, before being mixed with the original observations at a detector. The piston changes the phase of the beam in four steps of 0, $\lambda/4$, $\lambda/2$, $3\lambda/4$. This produces four intensity patters from which to calculate the phase map according to

$$\phi(x,y) = \tan^{-1}\left[\frac{I_4 - I_2}{I_1 - I_3}\right] \tag{5.5}$$

Once the phase map has been measured, the user can either feed this directly into the xAO DM and suppress the residual phase aberrations. The accuracy is determined by the precision of the piston and the precision of the DM. This correction would not null the amplitude-induced speckles, however.

It is impossible to null all of the phase and amplitude speckles through corrections

to the wavefronts phase alone. Instead, speckles on one half of the image can be completely nulled at the expense of possibly increasing the intensity of the conjugate speckles on the other half. To do this, speckles are identified in some user defined control area of the image based on brightness and proximity to one another. Then the DM pattern that is required to null each speckle is calculated and the average of all patterns pattern is applied to the DM. The efficacy of this technique is in part determined by the precision of the calibration between the sinusoidal patterns on the DM and the desired speckles.

5.2.2.6 Coronagraph

There are four types of coronagraphs that can be selected. Each of them are simple Lyot architecture coronagraphs. For the 'Solid' coronagraph, the occulter is a simple binary mask that only blocks light out to a certain radius. For 'Gaussian', the occulter has a graded transmission that follows a Gaussian profile. This means that the light is concentrated to the edges of the pupil much more effectively. The '8th_Order' coronagraph can concentrate all of the parent star light the region blocked by the Lyot stop for 100% rejection of the parent star light [186]. These three coronagraphs come with PROPER. The 'Vortex' option is a vector vortex occulter for high throughput and small inner working angle coronagraphy. This coronagraph was implemented by adapting code from the HEEPS pipeline[1]. Additionally, there exists a 'None' and a 'None (Lyot Stop)' option. 'None' bypasses the coronagraph optics entirely, wheres 'None (Lyot Stop)' includes all the optics including the Lyot stop. This is useful for creating reference PSFs for injecting fake companions or measuring the unocculted aperture photometry when creating contrast separation curves (described in the Section 5.3).

[1] https://github.com/vortex-exoplanet/HEEPS

5.2.3 Detector

After Telescope has created an ideal datacube, Detector manipulates and formats that hypercube depending on the detector. This module was developed from scratch by us. With the 'ideal' mode the datacube remains unmodified. The only process that occurs is the calibration of the intensity. The intensity will sum to unity if no light is lost in the coronagraph or if the intensity has not been scaled because of a user-defined transmission. Therefore, the intensity is calibrated to photon counts by multiplying each pixel by star_photons and scaling the different wavelength frames if necessary.

For the 'H2RG' mode, the user can set a frame rate that is lower than the rate set by Atmosphere. In this case, the intensity map is summed in the time axis (for each wavelength), downsampling the datacube to match the selected frame rate, which has the effect of blurring the image. Also for H2RG, the user can set an amount of readout noise and dark current. The readout noise is applied to each pixel of an exposure, and its value is drawn from Gaussian distribution with user-defined mean and FWHM. For dark current, the number of additional counts, drawn from Gaussian distribution with user-defined mean and FWHM, is then randomly applied to time frames depending on the frame rate of the exposures. A dark frame can then be generated, from the mean and FWHM values, and subtracted to yield only the shot noise from the dark current.

5.2.3.1 MKIDs

Before Telescope and Detector are run, the MKID array is initialized. This includes assigning spectral resolution distributions, background phase distributions, and a responsivity (from imperfect flat fielding), to each pixel. For responsivity, a value is drawn from a user-defined Gaussian distribution centered on 1 and assigned to each pixel (the responsivity is assumed to be constant for each pixel with time and wavelength). Dead pixels are assigned by setting their responsivity to zero. In practice high frequency resonators tend to be the most problematic (see Figure 5.2 of [2]), and

so placing these at the edge of the array in the design is an effective way of maintaining many operational pixels next to each other and in the center of the image. In MEDIS, there is an option to distribute the operational pixels according to a Gaussian profile centered on each feed-line. Hot pixels are assigned by adding a user-defined large value to the responsivity matrix at random points.

For background phase noise that is used to set the detection phase threshold, a scalar value is assigned to each pixel, which serves as a FWHM for a Gaussian distribution. Later, when a photon is incident on a given pixel, a value for the background is chosen from this distribution. This way, the background phase changes with each time step. The finite spectral resolution will alter the phase measurement from the true value. We assumed that the spectral resolution followed a linear trend with wavelength. Therefore, the phase distortion is drawn from a datacube (depending on the pixel location and wavelength) of FWHMs for Gaussian distributions that are sampled during each photon detection.

For example, using an observation file of HD 91782 recorded with DARKNESS, we can observe the measured phase of all pixels. Figure 5.11 shows a histogram of this data. The baseline phase is approximately symmetric about 0. Obtaining mean and standard deviation measurements to feed MEDIS would require first flat fielding and flux calibrating this observation file. The wings of the distribution imply a Lorentzian may provide a better fit to the data. For the simulations presented in this thesis the mean across the array is set sufficiently low that the chosen probability distribution has little effect on the measured performance of the device.

Figure 5.12 summarizes the steps taken by the MKIDs module. The Telescope datacube for each time-step is 3D probability map with which to sample photons from. The interp_samples flag essentially turns the datacube into a continuous distribution reducing the quantization error. Since each wavelength frame requires a new wavefront to propagate through the system, drawing photons from the datacube in this way saves a significant mount of time over generating a datacube with lots of different wavelength frames.

The datacube is first truncated to the size of the MKID array. PROPER requires

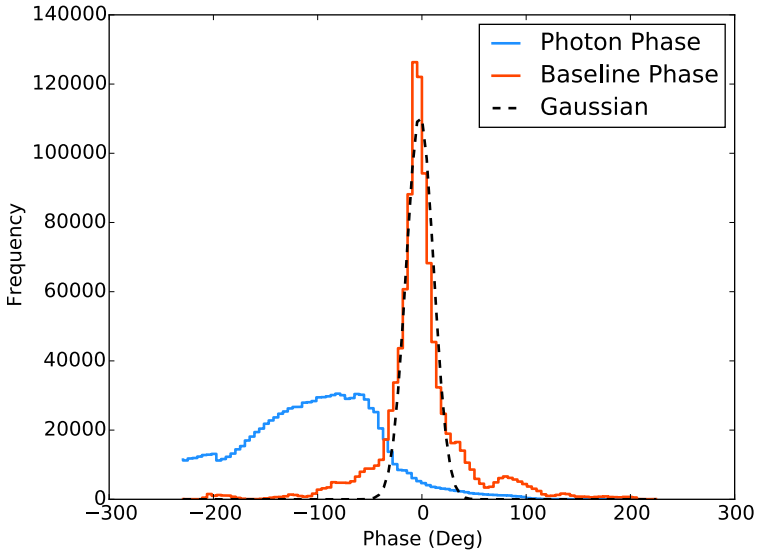

Figure 5.11: Measured phase data for a single time step when observing HD 91782 with DARKNESS. The baseline phase values are symmetric about ∼0 and they approximately follow a Gaussian distribution with standard deviation of 12.7.

that the number of pixels along one dimension of the wavefront array be a power of two. When modeling MEC (140×144 pixels) for example, a 256×256 grid size is initially required. The spatial pixels of the truncated datacube are then reassigned according to the responsivity matrix. Photons are then sampled from this datacube.

A list is created of photon packets with three elements: wavelength index, x, and y. The wavelength indices are calibrated to phase and then distorted according to each pixel's spectral resolution at that phase. If the phase is beyond the assigned threshold for that pixel, then the packet will be stored (this threshold represents some multiple of the standard deviation of the phase background). The timestamp of the datacube is applied to each stored photon packet. If the dead_time or remove_close flags are set then offending photons are rejected at this point.

The photon data table can either be saved directly to a H5 ObsFile to be processed by an MKID data reduction pipeline. Alternatively, it can be reformatted back into the hypercube format by binning the packets in wavelength and time. The parameters that characterize MKIDs is given in Table A.4.

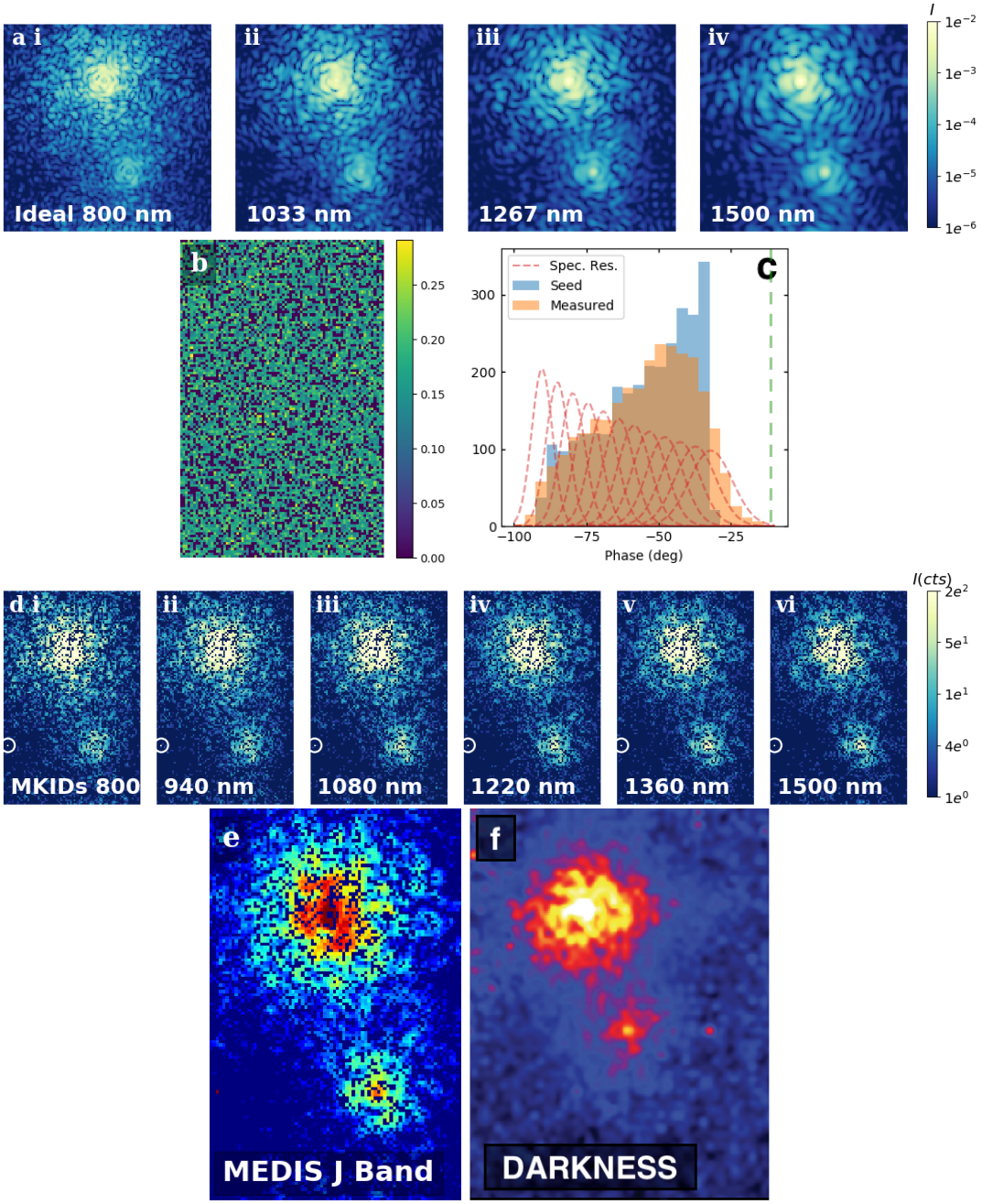

Figure 5.12: Overview of process of deteriorating creating MKIDs data with MEDIS for the 10 Uma AB binary system. a: four ideal wavelength frames spanning the bandwidth of DARKNESS created with Telescope for a 1 s exposure. b: the intensity responsivity map for each pixel. c: The phase histogram of pixels in an aperture of diameter 3 pixels centered on 10 Uma B before (Seed) and after (Measured) the phase of each pixel was altered because of the finite spectral resolution. The red dashed curves show the mean spectral resolution and the green dashed curve shows the mean phase threshold of the pixels in the aperture. d: A subset of the resulting datacube as measured by the MKID array. The white circle highlights a hot pixel. e: The wavelength frames spanning J band summed to produce the image shown. f: A mosaic image of 10 Uma AB as measured by DARKNESS November 2016 using a J band filter [124].

5.3 Post-Processing and Analysis

As of writing this, many of the exoplanet analysis tools are still being implemented in the DARKNESS pipeline. Therefore, I developed an implementation of much of the exoplanet-specific techniques described below (general observational hypercube analysis was already implemented). Some parallel efforts on the formal pipeline are being conducted and the two packages will be combined at some time in the near future. The open-source tool for high-contrast image processing: Vortex Imaging Pipeline (VIP) [6] provides tools to perform SDI, ADI and RDI and plot the contrast curves. These tools were adapted in order to implement the other speckle reduction and companion discrimination techniques described in this section. To exemplify each of these techniques we use observation data fabricated with MEDIS.

5.3.1 Contrast Curves

There are a number of ways of calculating contrast at a given radius in the processed image. Here, we will use a method that determines the sensitivity of the observation to companion objects through measurements of the noise at different radii outlined in [6]. First, resolution elements of size λ/D are organized into rings centered on the middle of the image. The RMS noise is calculated for each resolution element and the azimuthal average of each ring $\delta_{\mathrm{rms}}(r)$ is then evaluated. These data are then smoothed with a Savitzky-Golay filter.

The amount of planet signal that remains after applying a post-processing algorithm is known as throughput, T. This is quantified by injecting fake companions into each of the frames of the unprocessed hypercube at a range of radii and azimuthal angles. The fake companions are the unocculted PSF of the system appropriately scaled, which is easy to create using MEDIS. The algorithm under investigation will convert the raw hypercube into a processed image (hypercube in this context can have a length of one for the spectral or time dimensions). The brightness of each fake companion is compared to their original brightness $I_{\mathrm{out}}/I_{\mathrm{in}}$ to provide the throughput

at different points in the image. This value is azimuthally-averaged to yield $T(r)$ at a number of radii. These data points are then fit with a univariate spline to provide a continuous function for $T(r)$.

The limited sample statistics at small angles decreases the confidence level of detection (CL). Therefore to maintain an equal CL at all radii, a penalty factor is defined $k(r) = \sqrt{1 + 1/(n-1)}$ where n is the number of resolution elements in each ring [92]. The contrast-curve is then calculated with

$$C(r) = \frac{d}{I^*} \frac{k(r)\delta_{\text{rms}}(r)}{T(r)} \tag{5.6}$$

where d is the detection threshold (usually 5) and I^* is the aperture photometry measurement of the unocculted star.

5.3.2 Differential Imaging

VIP is a high contrast imaging analysis library that has tools to handle common HCI observing strategies techniques such as: SDI, ADI and RDI (multichannel spectral, angular and reference, differential imaging). In each of these techniques a PSF is generated from the range of observations that contains a minimal amount of companion light. This reference PSF is then subtracted from the raw observations to increase the relative intensity of the companion signal.

Figure 5.13 illustrates the application of SDI to the observations using an example system sampled at a range of wavelengths. Each image of a wavelength datacube is radially scaled to the longest wavelength so the stellar PSFs of each of the images line up. The most simple implementation of SDI creates the reference PSF by taking a median of each frame. In practice, principle component analysis (PCA) is used whereby, the reference PSFs are created by projecting the observations onto a low dimensional orthogonal basis (with a higher variance). The post-processing algorithm Local Low-rank, Sparse, Gaussian decomposition (LLSG) builds on this technique by implementing a three term decomposition and has shown superior subtraction at small angles [100]. These reference PSFs can then be subtracted from each frame of

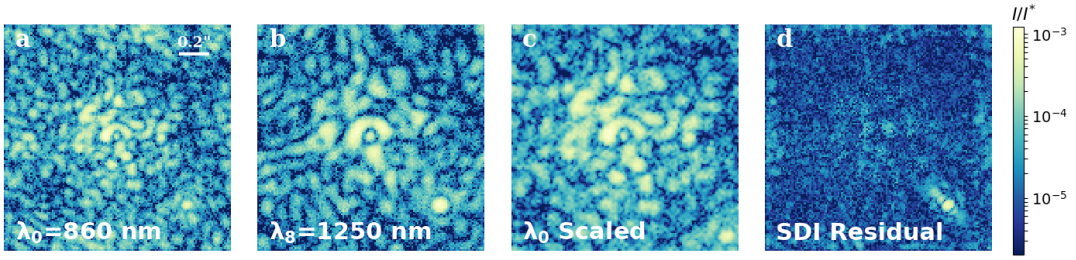

Figure 5.13: Demonstration of reference PSF subtraction by multichannel spectral differential imaging using MEDIS. The (aberrated) PSF scales linearly with wavelength (panels a and b). The frames are radially scaled to align the speckle patterns (c). Most of the initial speckle noise is removed for the SDI residual (d).

the scaled datacube, and those residuals median collapsed to form the SDI residual. In Figure 5.13d, the companion object becomes much more pronounced relative to the speckle noise.

ADI exploits the FoV rotation of a pupil stabilized observation with an altitude-azimuth telescope. First, LLSG is applied to each frame of a time datacube to create the reference PSFs. These references are subtracted from each time frame to create the residual maps. Then these residual maps are azimuthally rotated to align the companion objects before being median collapsed. Figure 5.14a shows the result of ADI on the example observation. ADI appears to provide a very good subtraction of the speckle noise at all radial separations. There is some self-subtraction of the companion even though it has a large radial separation. This is even more of a problem at smaller separations where the azimuthal rotation is smaller.

RDI involves making a second observation of a different star, which will contain a very similar quasi-static speckle pattern without the companion object. The reference observation can then simply subtracted from the science observation to reduce the speckle noise. For this demonstration, the companion object was removed for the second half of the example system observations. The evolution of the aberration pattern from the telescope slewing was produced using the method described in Section 5.2.2.3. From Figure 5.14b it is clear that there is less self-subtraction of the companion but more speckle noise at small radial separations.

Each of the techniques can be applied simultaneously to enhance contrast. For

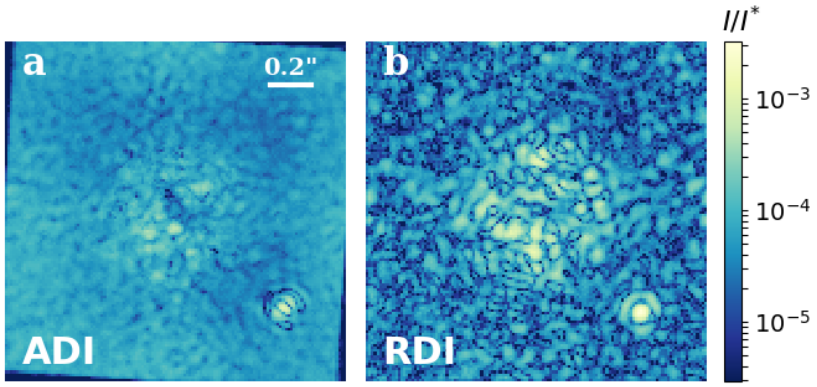

Figure 5.14: a: The residual frame after applying angular differential imaging. b: The residual frame after applying reference differential imaging. RDI is incapable of subtracting speckles that vary on time scales smaller than the slew time of the telescope.

example, through a two-step process, SDI+ADI makes use of the full MKID observational hypercube. First SDI is applied on all time samples to collapse the spectral dimension, then ADI is applied on those residuals. Similarly, ADI can be applied to both a reference observation and a science observation. The reduced reference frame is then subtracted from the reduced science frame.

5.3.3 Discrimination with Speckle Statistics

5.3.3.1 Dark Speckle Imaging

The first application of photon intensity statistics to discriminate companions from the host PSF was DSI [187, 188]. This is the technique that DARKNESS gets the initial part of its name from. In short, the intensity map is prevented from reaching the faintest levels at the location of a companion. By identifying and mapping regions that do not receive any photons during an exposure (so-called dark speckles), companion objects can be detected that have a mean brightness below the mean of speckle noise (at that annulus). MKIDs are an ideal detector for DSI since the photon counting sensitivity and the absence of read-noise or dark current enables us to reliably determine the presence of dark speckles.

DSI requires that the intensity modulates throughout the image, which is not the

case for static (or sufficiently long-lived quasi-static) speckles. This problem can be addressed by first performing a form of RDI. Two hypercubes are constructed with the same duration (and wavelength sampling), where like RDI, one the hypercubes is the target planetary system and the other is a reference. Instead of constructing a reference PSF by collapsing the reference hypercube, the two hypercubes are directly subtracted (the first effective integration of the reference is subtracted from the first effective integration of the target etc). The difference in the intensity pattern from the atmospheric evolution of the speckles between the two sets of observations can be advantageous for this technique, because it increases the rate of modulation of the speckles (particularly at small separations). The speckles can be made to modulate with a deformable mirror to serendipitously create dark-speckles or other optics based methods exist [189].

Figure 5.15 demonstrates how the data product from DSI is created – here called light maps. (The demonstration is conducted for a single wavelength – multiple wavelength DSI is investigated later.) First, the observation is split into a series of (reference subtracted) effective integrations. The duration of these effective integrations should be sufficiently long that at least a photon from the companion object can be consistently expected. If the companion brightness is unknown then several integration times can be investigated. Figure 5.15a shows the first 10 ms effective integration.

If a pixel in the effective integration has a value of 0 then 1 is added to the corresponding pixel in the dark map, otherwise the dark map pixel remains at zero (Figure 5.15b). This process is repeated for each subsequent effective integration map. Figure 5.15c shows the progress of this process after 10 effective integrations. As the speckles migrate and evolve, the dark map becomes brighter across the image except at the locations of the companion objects.

The light map is then created by taking the reciprocal of the dark map (Figure 5.15e). (Some constant small value is added to each pixel in the dark map prior to taking the reciprocal to avoid assigning infinities at the location of the zeros). For just a 2 s observation (200 of the effective integrations), the companion objects have

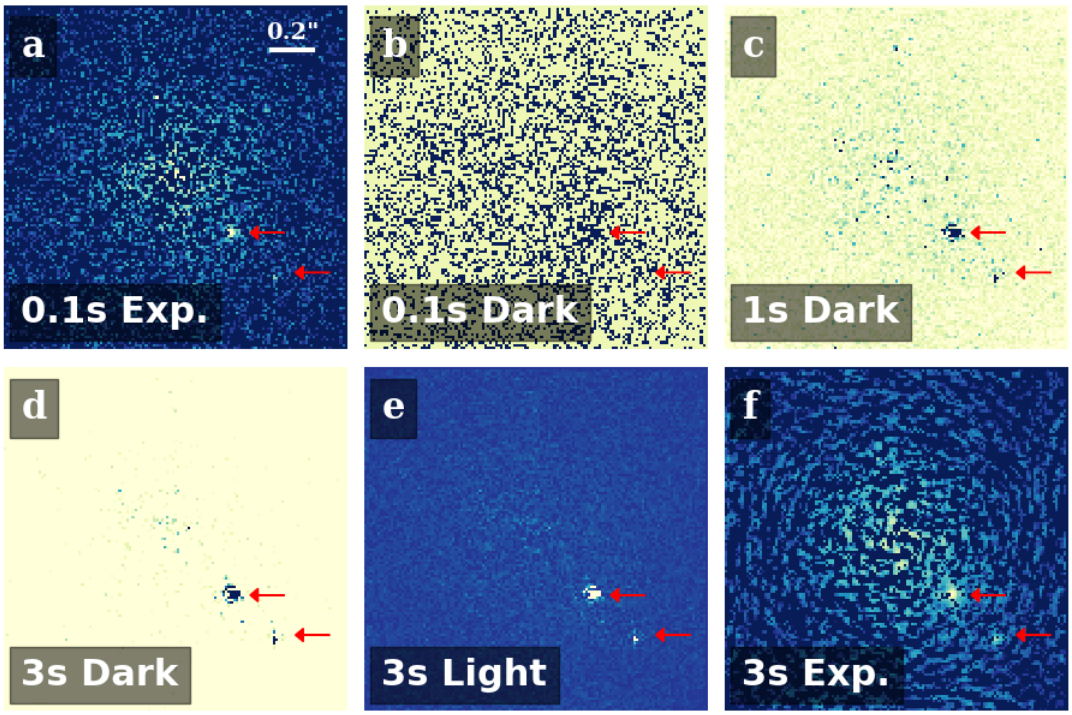

Figure 5.15: A demonstration of DSI. The observation is separated into several effective integrations (a). For each effective integration a dark map is created (b). These dark maps are stacked where companion objects should reveal themselves as dark spots in an otherwise bright image (c and d). The light map is created by taking the reciprocal of the dark map (e). The companion objects in the light map have a higher SNR compared to the integration over the same time period (f).

emerged from the speckle noise, and would have be difficult to discern from a standard integration of the same duration (Figure 5.15f). These simulations were performed with a constant Fried parameter of 0.2. The effectiveness of this technique is highly dependent on the seeing conditions. It would therefore be interesting to investigate and quantify the performance of this technique as a function of Fried parameter.

Dark speckles have a wavelength dependence that can be exploited to provide a greater suppression of the speckle noise. The dark speckles should move radially with wavelength meaning that at each moment there is a great chance of a dark speckle coinciding with the location of the companion. One method is to simply create dark maps for each wavelength and stack them – here called broadband DSI. Another technique, here called broadband-simultaneous DSI or BB sim DSI, counts

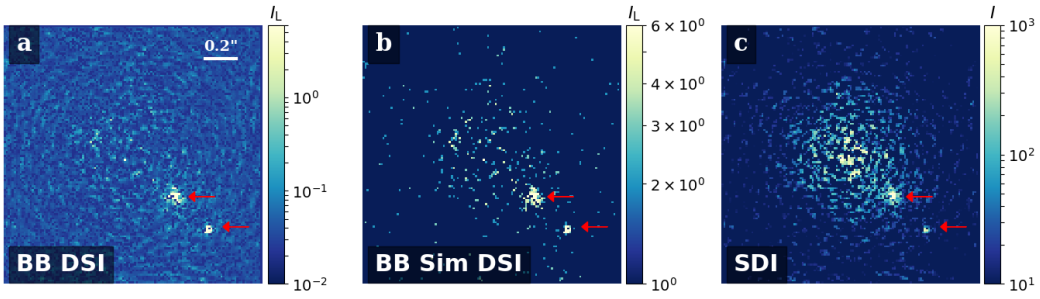

Figure 5.16: Performing DSI at multiple wavelengths after first applying RDI. a: dark maps created at several wavelengths and then stacked and inverted. b: light maps created by observing spaxels of wavelength datacubes for dark speckles. c: conventional SDI post processing.

the number of effective exposures in which a spaxel has no instances of zero counts. In Figure 5.16, these broadband DSI techniques are compared to the conventional SDI post processing technique. The BB DSI light map has lower noise especially at small separations. For future work, it would be interesting to plot contrast-separation curves for these different methods.

5.3.3.2 Stochastic Speckle Discrimination

SSD exploits the difference in the intensity probability density function (PDF) at the location of sources (on-axis) and the speckle pattern (off-axis). The PDF for off-axis speckles was introduced in Chapter 1:

$$P_{\text{off}}(I) = \frac{1}{I_s} \exp\left(-\frac{I + I_c}{I_s}\right) I_0\left(2\frac{\sqrt{I I_c}}{I_s}\right), \tag{1.9}$$

As demonstrated in Figure 1.7, at increasing I_C/I_S, the distribution tends towards a symmetric Gaussian. Conversely, as I_S dominates the distribution becomes positively skewed, meaning an excess of smaller values, ultimately converging on a negative exponential.

It has been demonstrated by several authors that the MR function cannot describe the photon statistics at the core of the image [190, 191, 192]. Instead, the intensity distribution is dictated by variations in the instantaneous Strehl ratio. In short, the phase correction applied across the pupil by the xAO can be assumed to be normally

distributed, which means that the sample phase variance is Gamma distributed over an ensemble of exposures [193]. After applying the Maréchal approximation (Equation 1.12), the PDF for the instantaneous Strehl ratio, S, is

$$P(S) = \frac{m}{2} \frac{\left[\frac{m \ln(S)}{2 \ln(\mathrm{SR})}\right]^{(m-3)/2} \exp\left[-\frac{m \ln(S)}{2 \ln(\mathrm{SR})}\right]}{\Gamma\left(\frac{m-1}{2}\right) |\ln(\mathrm{SR})| S}, \tag{5.7}$$

where m is the number of sub-apertures in the WFS, SR is the theoretical Strehl ratio from the ensemble-averaged phase variance, and the Gamma function $\Gamma = \int_0^\infty t^{x-1} e^{-t} dt$. The important feature of this equation is that it is negatively skewed at mid-high SR (where Equation 1.12 is valid).

From the definition of Strehl ratio (equation 1.11), the intensity is related to SR by the constant unabberated intensity

$$P_{\mathrm{on}}(I) = \frac{P(S)}{I_a^*}. \tag{5.8}$$

The profile of a PDF remains the same after applying a constant scaling factor, so $P_{\mathrm{on}}(I)$ is also negatively skewed for xAO.

The profile of the on-axis distribution applies to companion source as well as the central source. However, the statistics will be mixed with those of the random fluctuations. The result is a convolution between the profiles of opposite skewness that produces an approximately Gaussian distribution depending on the I_C/I_S.

Figure 5.17 demonstrates these distributions for different locations in the example HCI observation. In the core of the image, the profile is distinctly negatively skewed and the MR fit is poor resulting in what would be an under estimation of I_C/I_S. At the location of the companion, the profile is also slightly negatively skewed. At the location of the quasi-static speckle is slightly positively skewed, and the trend continues for the speckle located far from the core.

SSD therefore consists of mapping the 'beam-ratio': $r = I_C/I_S$ and observing the locations of the maxima [194]. From Equation 1.8 we can rearrange for I_S and I_C for

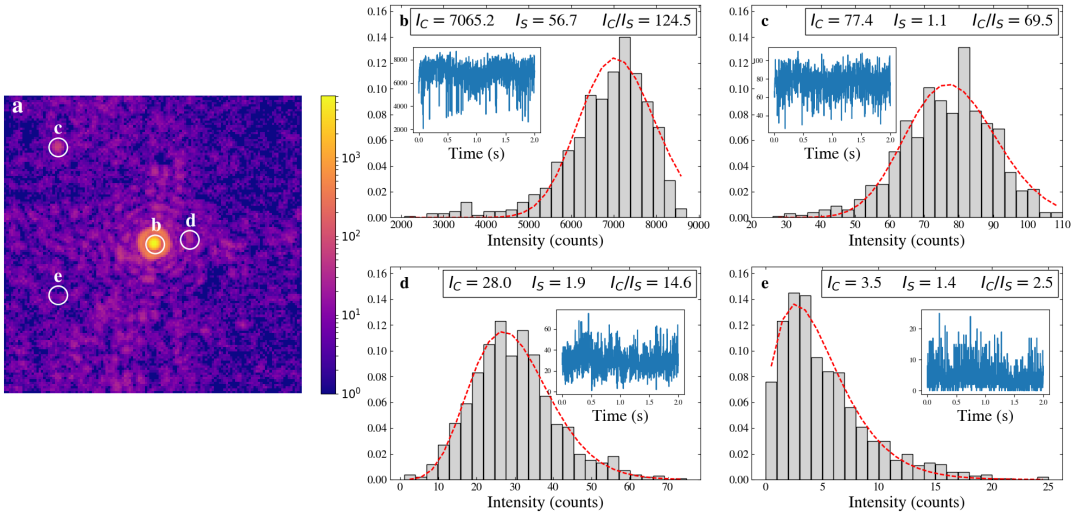

Figure 5.17: Demonstration of the speckle intensity statistics used in applying SSD on a system with xAO and a photon counting detector created using MEDIS. At each location, the intensity profile was generated from the 2 s light curve, which is shown in the inset. The fit is from an MR distribution and the parameters I_C and I_S were derived from the fit. a: The first 10 ms of the observation stacked into an effective exposure. b: The central star (unocculted). c: A companion object fainter than the host star by a factor of 100. d: A quasi-static speckle that has been 'pinned' to the intercept of one of the Airy rings and the spider diffraction spikes e: A more dynamic speckle.

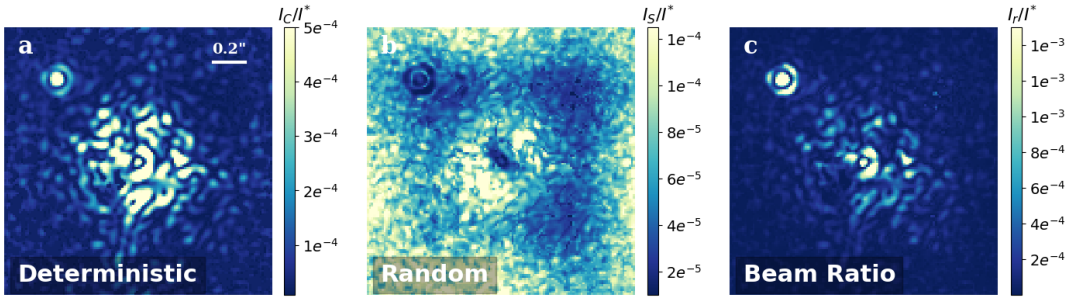

Figure 5.18: The intensity of an observation is separated into its deterministic, I_C (a) and random components I_S (b). Companion objects will have a larger beam ratio (c) compared to speckles of similar intensity.

a given light curve (photon time stream).

$$\begin{aligned} I_C &= \sqrt{\langle I \rangle^2 - \sigma^2}, \\ I_S &= \langle I \rangle - I_C, \end{aligned} \tag{5.9}$$

where $\langle I \rangle$ and σ are the mean and standard deviation of I, respectively. Alternatively, the light curve is binned to create the intensity histogram, and the curve is fit with equation 1.9 to extract I_s and I_c (this will underestimate r for negatively skewed pixels however).

Figure 5.18 demonstrates the implementation of SSD on the example observation. Companion objects will have a larger I_C, and a smaller I_S, compared to speckles of the same intensity. The ratio of these components amplifies the SNR of the companion relative to the speckle noise.

5.4 Future Plans

This chapter briefly features some data from observations of HD 91782 from a DARKNESS observing run that I attended. I selected this target for the observations because there are complimentary P1640 observations. It would be interesting to reduce these observations and compare the reduced DARKNESS data, the P1640 data and a simulated dataset created with MEDIS. This would serve as a means of verifying MEDIS and exploring the performance MKID-based instruments relative to semiconductor-

based instruments.

5.5 Summary

MEDIS has been introduced as a means of creating and analysing potential MKID HCI systems. It has been used to demonstrate the efficacy of speckle statistics-based companion discernment as well as some potential modifications to those algorithms such as the introduction of an RDI preprocessing step or the exploitation of multiple wavelength datum for DSI. The simulator provides an efficient means to develop and experiment with the new algorithms. For example, more sophisticated algorithms should be possible that exploit the intensity profile and dark speckles simultaneously, to reach even better contrast performance once developed.

Chapter 6

ARCONS Observations of NGC 6751

In this chapter we will describe the analysis of NGC 6751 observations taken with ARCONS and my contribution to the ARCONS data reduction pipeline. ARCONS was the first generation of optical MKID camera and it suffered from low pixel yield and low spectral resolution. We therefore briefly introduce the idea of time and spectral domain analysis of these data with ARCONS.

6.1 Introduction

NGC 6751, also known as the "Glowing Eye", is a multiple shell planetary nebula (PN) located at a distance of 2.7 kpc [195] in the constellation of the Eagle (Aquila). Sabbadin [196] first derived a spatiokinematical model of NGC 6751 as a prolate spheroid with 40 kms^{-1} expansion velocity. In 2009 the *Gemini South Observatory* captured an image of NGC 6751 as a result of the winning entry by high school student in the 2009 Gemini School Astronomy Contest[1], which is shown in Figure 6.1. Clark et al. [197] created a 3D morphological model of the main components using archive images from the *Hubble Space Telescope, Spitzer* and long slit spectroscopy from *Observatorio Astronòmico Nacional*. They identified a number of features, which we have labeled in Figure 6.1. Surrounding the WR star is an expanding bubble, followed by a ring and a disk. Bipolar outflows enclose this, and the halo surrounds all these

[1] `http://www.gemini.edu/node/11329`

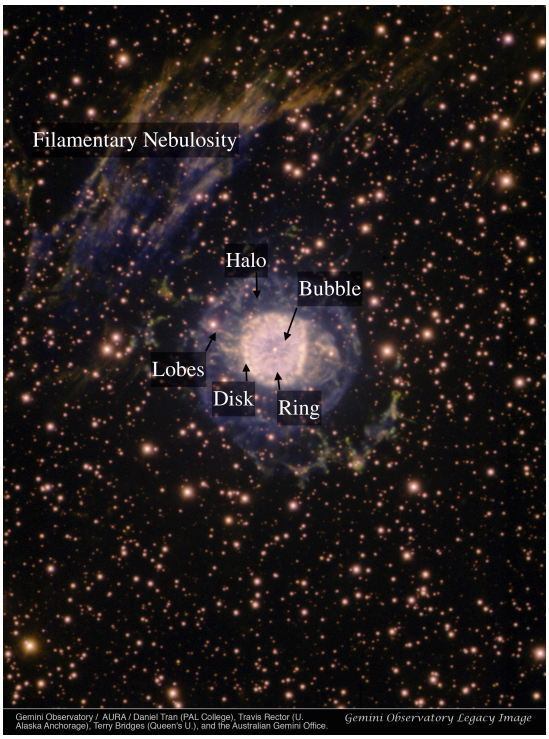

Figure 6.1: Gemini South image of NGC 6751 using narrow-band filters with the Gemini Multi-object Spectrograph (GMOS)[1]. We have added labels for the features described in the text. The Bubble is the central darker region. The Ring is the lighter blue region and the Disk is the features around the Ring coloured orange-yellow in this image. Red, blue and yellow colour coding was applied filters centered on hydrogen, ionized sulfur and doubly-ionized oxygen, respectively.

features. Beyond the halo there is filamentary extended nebulous region that has radial velocities inconsistent with NGC 6751.

The central star has been identified as a Wolf-Rayet (WR) by several authors [198, 195, 199]. These are evolved stars with strong broad emission lines of helium, nitrogen and carbon and weak or absent Hydrogen lines. Specifically, NGC 6751 is a WC 4-type star, which are characterized by a strong CIV line at 465.0 nm and a moderate OV line at 557.2–559.8 nm compared to the CIII line at 569.6 nm.

6.2 Analysing NGC 6751 with ARCONS

NGC 6751 was observed with ARCONS with the Hale 5m telescope at Palomar Observatory on the 21st of October 2014. This target was selected to demonstrate the imaging capabilities of ARCONS, which at the time, was the largest deployed superconductor-based camera. The observation consisted of a single 20 minute continuous exposure. This exposure can be separated into 38 shorter effective exposures lasting approximately 30 seconds, with the telescope pointing angle moved between exposures, forming a raster pattern across the target.

Using the ARCONS pipeline [7] each of the observation files were beam mapped, wavelength calibrated, flat fielded with a twilight exposure, flux calibrated with star SA95-42 [2], and the bad pixels time masked as explained in Chapter 4. After these calibrations about 60% of the 2024 array were usable. The majority of the pixels were lost due to magnetic flux trapping in the superconductor during cool-down and a faulty ROACH board, in addition to the resonant frequency collisions.

A spectral datacube was created for each frame by integrating the events during the effective exposure time of that frame and binning the detections in wavelength. The size of these bins (for all pixels) are 12 μm (0.1 eV) to match the wavelength bins of the flat field calibration.

The spectral resolution of MKIDs is generally low, and the spectral resolution varies between pixels. Figure 6.2a shows the distribution of R colored according to which ROACH board controlled the sub-array. These distributions are measured using a short wavelength (400 nm) laser (which for TiN produces the highest estimates of energy resolution) and yet the average energy resolution across the array is only ~ 5. The spectral resolution of each of the pixels was normalized to the lowest resolution pixel using the convolution method explained in Chapter 4.

A true colour image of the first effective exposure is shown in Figure 6.3. To create this image, the center wavelength of wavelength bin of the datacube was mapped to

[2] http://www.eso.org/sci/observing/tools/standards/spectra/

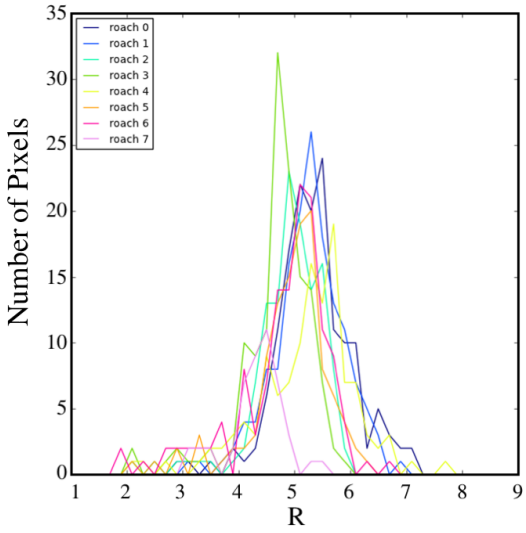

Figure 6.2: The distribution of pixel energy resolutions across the array for blue (400 nm) light.

an RGB value in 8 bit [3]. Then the mean RGB value of the spectrum for each spaxel was calculated and divided though by the average RGB values for the whole array to normalize. This effective exposure is centered on the WR star and the immediate nebulosity.

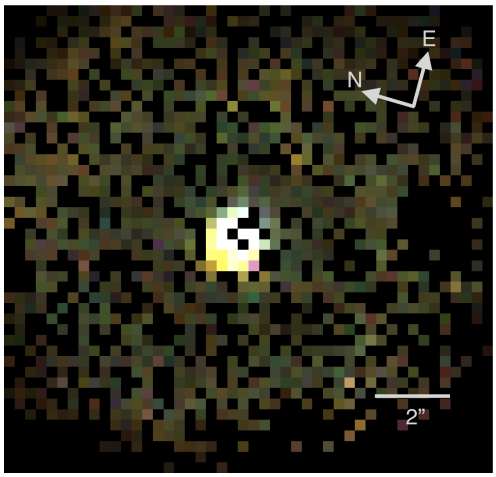

Figure 6.3: True colour image of frame 0 of the NGC 6751 mosaic created using all wavelengths of ARCONS and combining events over 8 s. This image is centered on the WR star of NGC 6751 and is located in the center of the dither pattern. 60% of the pixels was usable to construct this image after data reduction and calibration.

[3]http://www.physics.sfasu.edu/astro/color/spectra.html

6.2.1 Mosaic Creation

The displacement and angle between frames was inferred by comparing two adjacent frames (at a single wavelength) that shared a point source (no significant rotation was observed throughout the observation). The pixel location of the source was identified with a Gaussian fit in both images and the difference between the two frames recorded. The required plate scale was measured by taking the ratio of pixel offset and telescope offset from the logs. The direction of the movement between frames was calculated from the difference between arctangent of the pixel offset and arctangent of the telescope offset. To account for telescope drift, the offset of the central source between the first and last frame was measured, and for each frame, this displacement was divided by the cumulative exposure time before that frame. This assumes that the drift sums up linearly over time, which may not be true because of the offset direction changes, but provides a reasonable approximation as will be seen.

A grey scale mosaic is created by offsetting and stacking each frame (at all wavelengths), and dividing by the total effective integration time for each pixel. This first version of the mosaic suffered from some non-uniformity in each frame (Figure 6.4a). We realized that not all dead pixels had zero effective exposure time associated with them. That way, pixels in the final mosaic with contributions from these pixels would appear too dark, since they contribute no counts but were still being weighted by the exposure time as if they were. The software 'bug' was traced back to the bad wavelength calibrations masks, and effective integration time of all these flagged pixels were set to zero. After this process, a hot pixel became apparent that was missed by the hot pixel mask and remained throughout the observation, which can be seen in Figure 6.4b. This was manually removed by adding the offending pixel to the bad pixel mask.

The RGB wavelength ranges of a Hubble image were selected to create the ARCONS mosaic colour image [4]. The Hubble image was captured with Planetary Camera 1 (PC1) and the filters were Ha (f658n), Green (f555w), and OIII (f502n), for

[4] `http://www.waid-observatory.com/ngc6751-2013-02-25-TriColor-HLA.html`

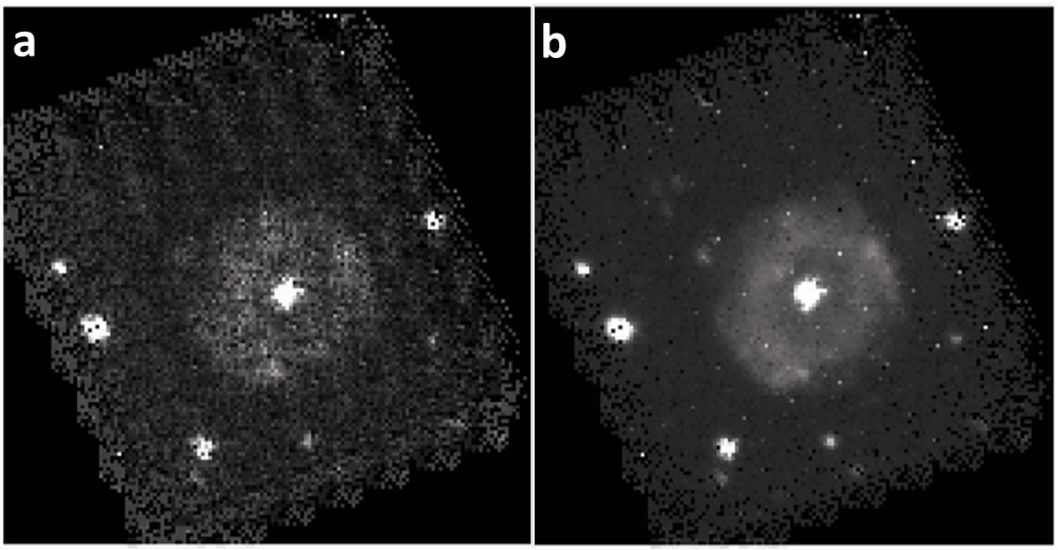

Figure 6.4: Mosaic images of planetary nebula NGC 6751 with ARCONS in the wavelength range 430–460 nm. a: The first version of the mosaic where frame offsets have been corrected for telescope drift and full calibrations (including flat fielding) have been applied. The pixel to pixel response appears to be incorrectly normalised in each frame resulting in a repeating noisy pattern. The cause of this problem was pinned down to pixels with incorrect wavelength calibrations being masked in pixel response but not effective exposure time b: The mosaic after the effective exposure time correction. A hot pixel can be observed that was was present in each effective exposure and missed by the hot pixel masking procedure.

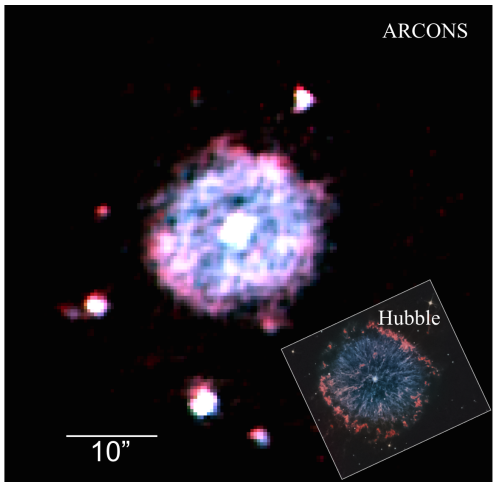

Figure 6.5: Colour image of planetary nebula NGC 6751 produced by creating red, green and blue images from the spectra and stacking those channels. The inset shows a processed Hubble image at the same orientation.

RGB respectively. The transmission profile of each these filters was multiplied by the mosaic spectra to generate the three colour channels. These three images were stacked and smoothed with a Gaussian kernel of radius 2 pixels using DS9 [200]. The result is shown in Figure 6.5. The difference in brightness between the Bubble and Ring regions is visible. Some of the filimentary structure in the Ring may also be apparent. The Disk that surrounds the Ring is clearly discernible in red. The faint Halo and Lobes are not visible in the mosaic.

6.2.2 Spectral and Time Analysis

The spectra of four components of 6.5 are given in Figure 6.6. The curves were obtained by first taking the mean spectra of apertures with radius three pixels at the five locations. Because the Bubble and Ring spectra would be contaminated by the Halo spectrum, the raw Halo spectrum was directly subtracted to produce the spectra shown. Similarly, the WR spectrum will be contaminated by both the Bubble and Halo spectra (although not Ring or Disk), so their raw measurements were subtracted to produce WR spectrum shown.

We can see that the red colour of the Disk is primarily due to the Hα line, however,

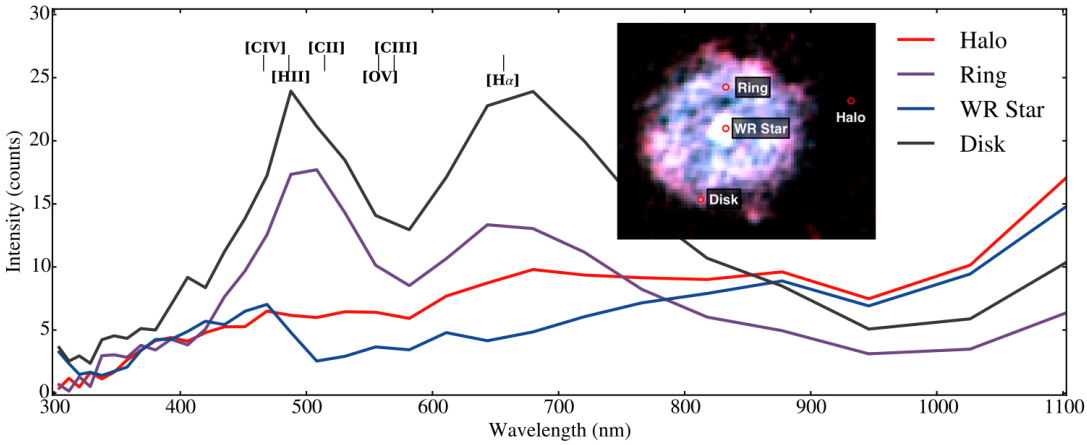

Figure 6.6: The spectra of different components of NGC 6751. The spectra are the average of apertures with radius three pixels at various locations of the mosaic. The vertical lines show the locations of relevant spectral lines.

the resolution of the spectra is insufficient to confidently discern spectral lines in any of the curves. Possible candidates for the spectral lines that act to produce the broad spectral features are included in the figure. There appears to be a deficiency of Hα in the evolved WR star compared to the Ring. Furthermore, in accordance with the WC4 classification, the WR Star appears to have a stronger CIV line, and a moderate OV line, compared to the CIII line. At the longer wavelengths there appears to be a thermal background manifesting in all spectra, however 1000-1100 nm is too short for any of the temperature stages of the instrument or the telluric contribution.

In the time domain, NGC 6751 initially appeared to exhibit a strong periodic signal at $\sim$ 1Hz from the central WR star. Publications exist ([201]; [202]) that claim WR stars could pulse at periods of several minutes or more, but no observations have ever been made of WR stars with periodicity at the capabilities of ARCONS.

Since the periodicity in the light curve is close to one second, one alternative explanation for this curve is a systematic error in saving the events in the observation files since they are formatted in rows of one second. Another explanation would be the telescope guide camera making a correction and offsetting the telescope each second.

To investigate, the light curve of one of the foreground stars in the mosaic was measured. Since the two curves correlated very strongly, the periodic nature of the

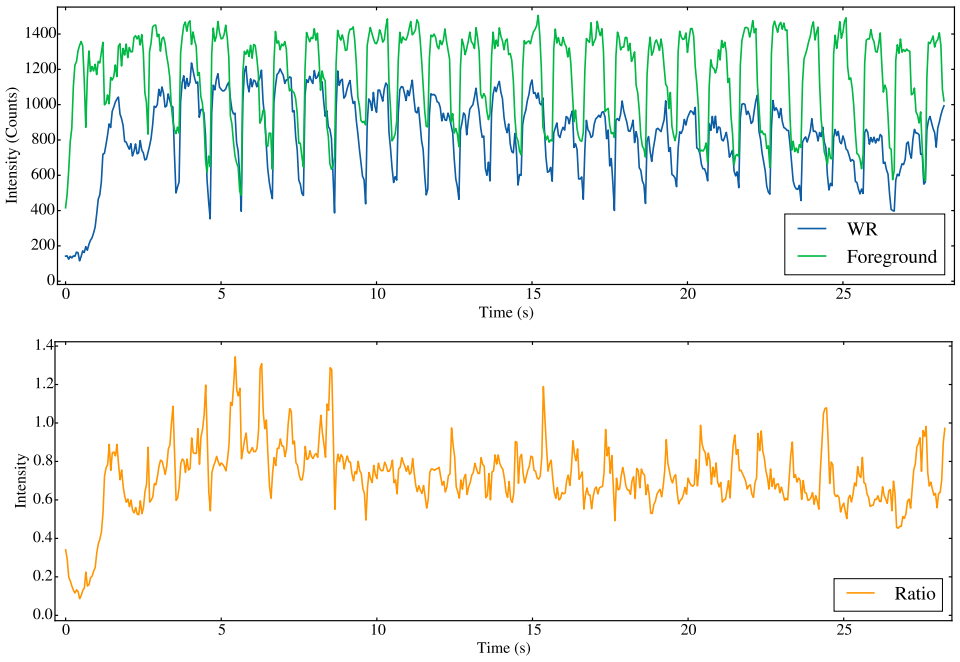

Figure 6.7: Light curve of the central WR star in NGC 6751 and a foreground star integrated over all wavelengths with apertures of radius three pixels. The initial increase in the intensity is due to the telescope settling on the target.

light curve is likely to be an artifact of the observing process and not an astrophysical phenomenon. However, the light curve of the SA95-42, which was measured approximately eight hours after the NGC 6751 observations, does not contain these artifacts indicating the artifact was a guide camera issue. The observing logs also mention an issue with the guide camera.

This artifact was addressed by dividing the WR light curve by the foreground star light curve (assuming the foreground star is not periodic on these time scales). The periodogram of each of these light curves is shown in Figure 6.8. The 1Hz feature and the resonances are apparent in the WR and foreground star curves. The ratio contains no obvious periodicity on these timescales.

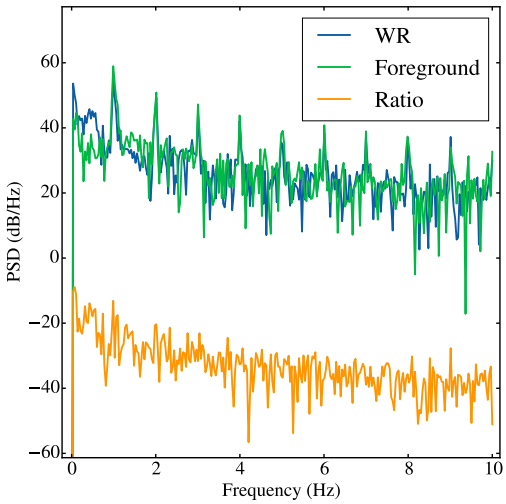

Figure 6.8: A periodogram of two structures in NGC 6751. The 1 Hz feature in the WR star and foreground star are a guide camera artifact. The WR star contains no periodicity on the timescale of the ARCONS observation.

6.3 Summary

This analysis of NGC 6751 ARCONS observations has served as an introduction to analysis of on-sky data with MKIDs. These data demonstrate the capabilities (for example inherent spectral resolution) as well as their draw-backs (for example dead pixels and low resolution).

Chapter 7

Future Work and Conclusions

In this chapter, we conclude by demonstrating the performance of a theoretical MKID HCI instrument, and make some initial predictions on how the array parameters effect the contrast performance, drawing from techniques outlined in previous chapters.

7.1 Next Generation Instrument Predictions

The Extremely Large Telescope and the Thirty Meter Telescope both have dedicated high contrast instruments planned respectively: Planetary Camera and Spectrograph [42] and Planetary Systems Imager [203]. As the field of HCI is developing rapidly, these instruments are still in the early design phase, to make use of the state-of the art technologies and methods once they become established. The primary criteria for the performance of these instruments will ultimately be their contrast-separation performance. MKIDs are being considered for such instruments and therefore detailed characterization of these devices is essential.

We theorize an MKID instrument that could exist in the year 2025 for example a possible upgrade to SPHERE. This time scale is beyond the current instruments in development but should give sufficient time to impact the design of PCS or PSI. This theoretical instrument is based on a rough extrapolation of the current trends of MKID parameters – we make no extrapolation of the performance of telescope specific parameters.

Table 7.1: The input parameters for the simulation of the HR 8799 system as observed by a next generation MKID instrument.

Parameter	Value
Number of Actuators	44×44
M1 Diameter	8 m
Throughput	10%
CPA surfaces	5
NCPA surface	5
Coronagraph	Vortex
Phase Aberrations	True
Amplitude Aberrations	False
Plate Scale	6 mas/pixel
Array size	256×256
R at 406 nm	30
Quantum Efficiency	10%
Pixel Yield	95%
Bandwidth	700–1700nm

The parameters of the simulated system are given in Table 7.1. The plate scale of 6 mas/pixel is slightly smaller than MEC's 10 mas/pixel in order to achieve better sampling of the PSF and achieve better contrast performance at small separations where less massive planets should exist. An array size of 256×256 could be achieved using several of the second generation readout electrons. This next generation array would require 3 of the MEC readout crates.

A spectral resolution of 30 would be several multiples above the current performance. However as discussed in section 2.5, there are a number of avenues being pursued to this end. A final pixel yield of 95% would require further optimization of the fabrication process, but certainly should be achievable with PtSi arrays. Extending the long wavelength end of spectral bandwidth would require increasing the SNR of the phase response of the array. The phase noise can be decreased by using parametric amplifiers instead of HEMTs, for example. Increasing the responsivity of the arrays could be achieved in the design of the resonators, or in theory using the 'optimal' bias tuning method outlined in section 3.5 whereby a laser source excites the array and the optimum bias point is chosen based on the response.

The target for the demonstration is a modified version of HR 8799, which was

introduced in Chapter 1. We use similar contrast and separations values of the HR 8799 system, except that we include extra sources at smaller separations and brightness that are currently undetectable (making no effort to test the validity of such sources existing with these parameters). Figure 1.2b shows HR 8799 as captured by Keck II.

The observation of the HR 8799 system was conducted for 530 s in total. The first 250 s tracked the target system, then a 30 s delay was implemented to account for the slewing of the telescope, then a further 250s of observation was conducted this time without companion objects. Both stars observed had the same brightness and spectral profile.

Figure 7.1 shows the observation of HR 8799 system at four stages of the data reduction process. The first panel is the effective integration of the target. The RDI panel here is created by taking the mean of the reference observation in the time domain. This reference datacube is subtracted from each time step of the target hypercube. This reference subtracted hypercube is then collapsed in the time domain by performing an effective integration. After applying DSI, all companion objects are immediately visible. The noise of the DSI images is significantly smaller than the RDI frame, especially at small separation.

Figure 7.2 shows the contrast separation curve of these plots. Conventionally, post processing algorithms decrease the noise at the radius of the companion thereby making it more prominent relative to the noise. The loss of throughput when applying RDI is negligible in this case. When applying statistic discrimination techniques such as DSI, the throughput of sources can actually increase, as well as decreasing the noise, providing huge gains in contrast performance. The noise performance of these algorithms at small separations means that the gain in contrast performance is larger at small separations where less massive planets would exist. Both companions are readily detected using DSI even if only one wavelength is used. These simulations demonstrate that Neptune sized exoplanets should be detectable with an MKID instrument.

There are a few directions where this research can be extended. Firstly the pa-

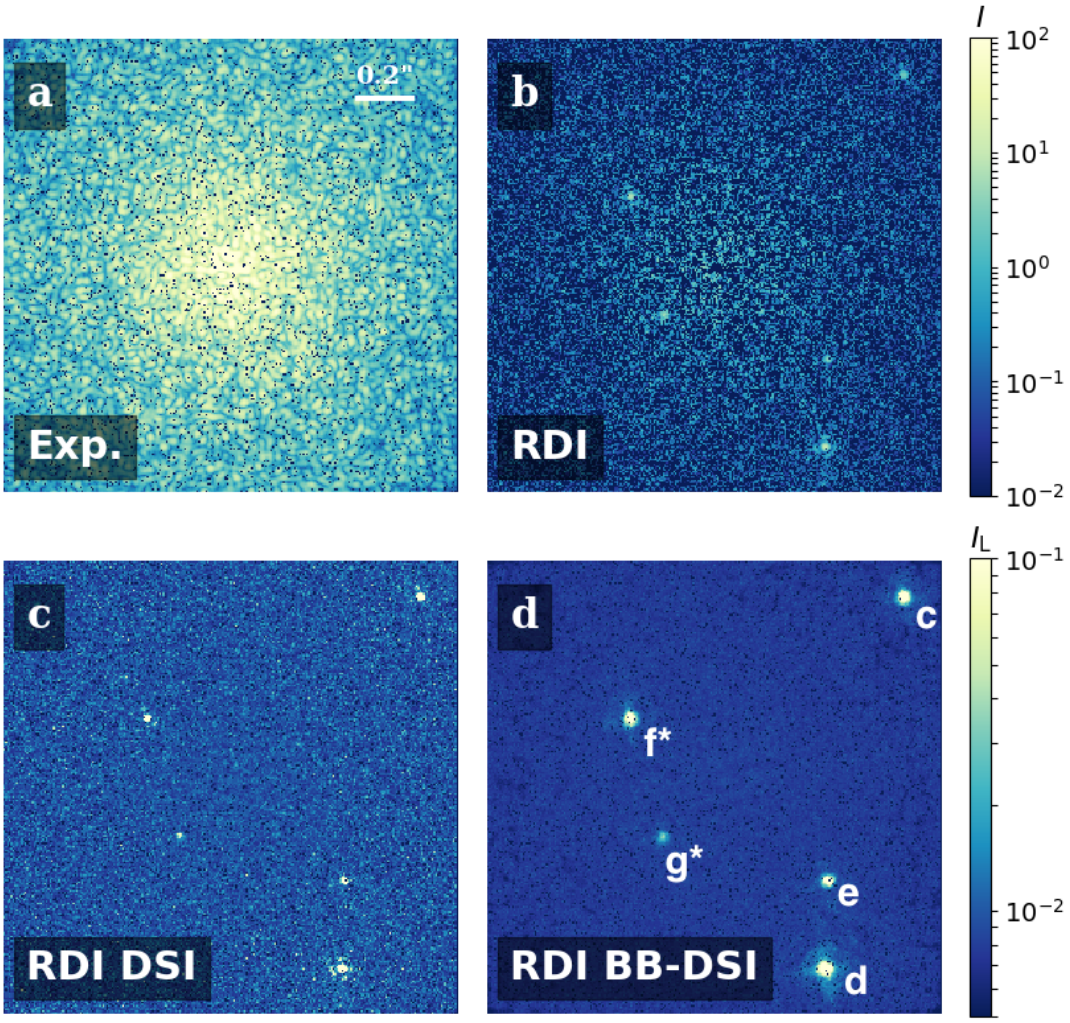

Figure 7.1: The observation of a modified version of HR 8799 as captured by the next generation MKID instrument with 250 s of on-target observation time. a: Effective integration of the target. b: The difference of the mean of the reference observation and target. c: DSI was performed at one wavelength of the reference subtracted hypercube. d: The broadband DSI technique was applied to all wavelengths of the reference subtracted hypercube. The binning time of the DSI was 1 second.

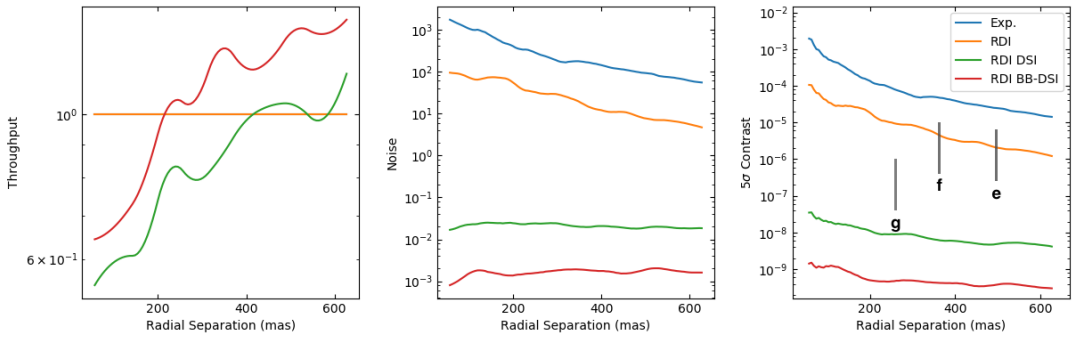

Figure 7.2: Contrast-separation performance of the next generation MKID instrument on the HR 8799 system.

rameters of the post-processing algorithms used in this simulation are not optimal. For example, the chosen length of the effective integrations has a significant effect on the performance. A larger integration means that the pixels at the location of the faint companion are more likely to receive photons and a deeper hole is created in the dark map. However, this also means that more false triggers happen, which can be observed in Figure 7.1c where isolated pixels are also 'hot'. The longer integrations times means there are fewer time samples for a given observation so these false triggers are less likely to be averaged out. In Figure 7.1d these false counts are averaged because of the extra samples from the spectral dimension. An investigation into the optimum effective integration for a given companion brightness would be useful.

Similarly, the order that the algorithms are applied in has a significant impact on the final performance. Another implementation could be to initially apply RDI (as above), then SDI on several shorter effective integrations to collapse the hypercube into a datacube with reduced length in the time dimension. Then DSI could be applied. SDI should reduce the contrast performance at small separations because of the increased self subtraction decreasing the throughput. However, there would also be an increase in the performance of the DSI algorithm as more of the static speckle intensity would be removed.

It would also be interesting to investigate combining SSD with DSI. This could be simply multiplying the result of both techniques or some simultaneous application of both. This should push the contrast performance even lower to the regime of

terrestrial planets, however a more rigorous simulation is required to verify this.

7.2 Investigating MKID Parameters

With the simulator in place it should be possible to investigate the impact of different MKID parameters on the contrast-separation performance – and consequently the exoplanet limiting mass.

Figure 7.3 illustrates an investigation into the effect of the presence of hot pixels in MKID data. Hot pixels currently manifest as a result of improper biasing parameters of the resonators [124] and are expected to be eliminated with future generations of arrays. The example system contained one companion object with 10^3 contrast at 4 λ/D. A datacube was created by integrating single photon data for 30 s and binning in wavelength. Two hot pixels with intensity equal to the maximum scalar value of the datacube were introduced into each wavelength frame of a datacube. SDI with LLSG was applied to the datacube both with and without the hot pixels. It is readily apparent that the presence of a hot pixel severely limits the performance of the SDI with LLSG algorithm.

Preliminary results have been taken for pixel yield using the parameters in Table A.4 and a simple application of the SDI algorithm. The application of SDI means that some of the dead pixels get filled. The diffraction spikes are appreciably lower magnitude as the pixel count approaches 100%.

The difference in the noise achieved is considerable resulting in nearly an order of magnitude difference in contrast at some separations. This difference may be even more prominent with the application of more post-processing algorithms.

Clearly maximizing the pixel yield is beneficial to achieving good contrast performance, but a more interesting measurement would be a surface plot of the contrast at $3\lambda/D$ vs two MKID parameters. This could be pixel yield and spectral resolution, for instance. This way the relative merits of the two parameters could be compared and quantified. It may be the case that there is a complex interdependence between the parameters, in which case three or more parameters may have to compared together

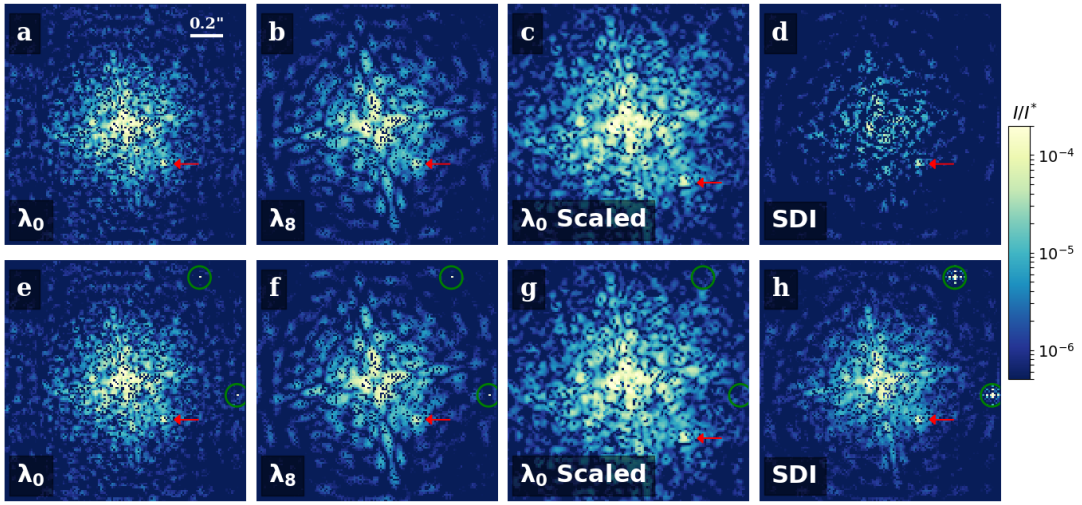

Figure 7.3: SDI reference PSF subtraction for an MKID array with and without hot pixels. The panel coloums match those in Figure 5.13. The top row does not have a hot pixel that persists the duration of the integration, the bottom row does, which are highlighted with a green circle. The companion object at 10^{-4} contrast is highlighted with a red arrow.

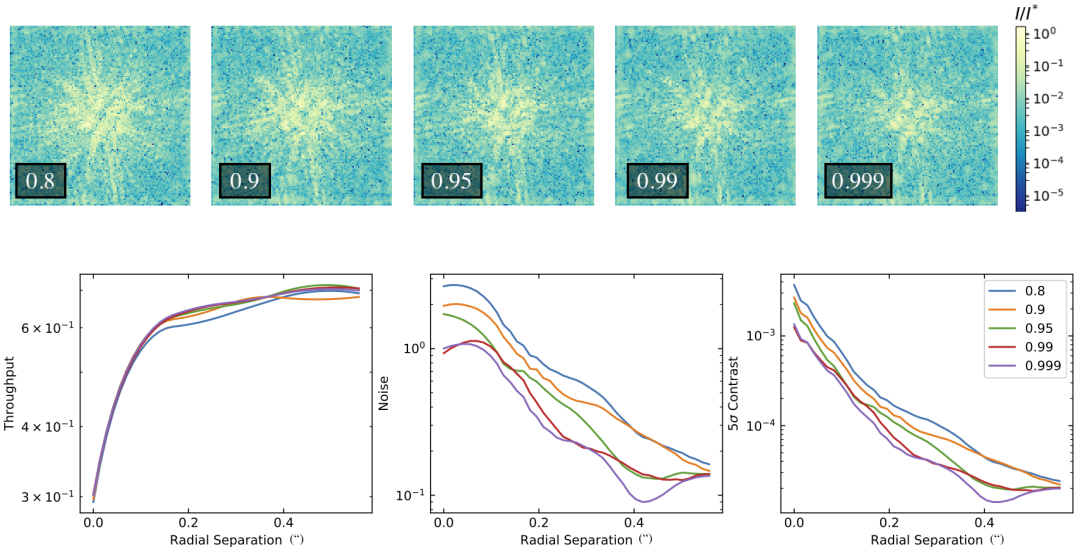

Figure 7.4: top: intensity ratio maps of the reduced image at various amounts of pixel yield. bottom: contrast separation performance for a standard MKID instrument with the parameters given in appendix A.4.

forming a multidimensional matrix of contrast values.

7.3 Concluding Remarks

MKIDs have the potential to revolutionize the field of high contrast imaging of exoplanets. I have shown that a machine learning algorithm can be created that optimizes the bias point large arrays of MKIDs. The CNN algorithm was shown to reach levels of accuracy of a human with only training data consisting of several hundred resonators. As the pixel count of MKID arrays increases with time and as this technology becomes more widely adopted, sophisticated algorithms such as the CNN will become essential for streamlining exoplanet observations. Furthermore, I outlined the next logical step in the progression of this algorithm that optimizes for pixel responsivity or spectral resolution using the laser source such as the one I created (discussed Chapter 4). This is should have an appreciable impact for exoplanet observations where long wavelength observations have more favorable contrast requirements and the spectral domain is exploited to suppress the the speckle noise through SDI or the simple extension to DSI.

I provided some characterization of MKIDs for high contrast imaging with a numerical simulator. Using this simulator I showed how some of the artifacts currently afflicting MKIDs, such as hot pixels, can impact the conventional post-processing algorithms such as SDI. If two hot pixels are missed by the hot pixel time mask, this can lead to an increase in the noise of the final image by several orders of magnitude. I verified the statistical discrimination methods that are uniquely applicable with MKIDs and showed how they achieve superior noise performance at small separations using parameters for the current generation of MKIDs as well as future extrapolations. These devices are still undergoing rapid development and hence a simulator that can make predictions based on future parameters will be a useful compliment to the current generation of MKID instruments as we work towards the coming generation of ELTs.

References

[1] R Dodkins, N. Fruitwala, N. Thatte, S. Mahashabde, K. O'Brien, S. R. Meeker, A. B. Walter, P. Szypryt, and Mazin. B. A. A Fast Machine Learning Based Algorithm for MKID Readout Power Tuning. *Proceedings of the International Symposium on Space Terahertz Technology*, 2017.

[2] Paul Szypryt. *Development of Microwave Kinetic Inductance Detectors for Applications in Optical to Near-IR Astronomy.* PhD thesis, California Institute of Technology, 2017.

[3] S. R. Meeker. *DARKNESS: The First Microwave Kinetic Inductance Detector Integral Field Spectrograph for Exoplanet Imaging.* PhD thesis, University of California, Santa Barbara, 2017.

[4] John E Krist. Proper: an optical propagation library for idl. In *Optical Modeling and Performance Predictions III*, volume 6675, page 66750P. International Society for Optics and Photonics, 2007.

[5] Marcel Carbillet, Christophe Verinaud, Mario Guarracino, Luca Fini, Olivier Lardiere, Brice Le Roux, Alfio T Puglisi, Bruno Femenia, Armando Riccardi, Barbara Anconelli, et al. Caos: a numerical simulation tool for astronomical adaptive optics(and beyond). In *Proceedings of SPIE*, volume 5490, pages 637–648, 2004.

[6] Carlos Alberto Gomez Gonzalez, Olivier Wertz, Olivier Absil, Valentin Christiaens, Denis Defrère, Dimitri Mawet, Julien Milli, Pierre-Antoine Absil, Marc Van Droogenbroeck, Faustine Cantalloube, et al. Vip: Vortex image processing

package for high-contrast direct imaging. *The Astronomical Journal*, 154(1):7, 2017.

[7] J. C. van Eyken, M. J. Strader, A. B. Walter, S. R. Meeker, P. Szypryt, C. Stoughton, K. O'Brien, D. Marsden, N. K. Rice, Y. Lin, and B. A. Mazin. The ARCONS Pipeline: Data Reduction for MKID Arrays. *The Astrophysical Journal Supplement Series*, 219(1):14, July 2015. ISSN 1538-4365. doi: 10.1088/0067-0049/219/1/14. URL `http://arxiv.org/abs/1507.05631`. arXiv: 1507.05631.

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

[9] Eric Mamajek. Cumulative number of exoplanets discoveries versus time. *URL https://figshare.com/articles/Cumulative_Number_of_Exoplanets_Discoveries_Versus_Time/4057704*, 2016.

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

[11] Michaël Gillon, Amaury HMJ Triaud, Brice-Olivier Demory, Emmanuël Jehin, Eric Agol, Katherine M Deck, Susan M Lederer, Julien De Wit, Artem Burdanov, James G Ingalls, et al. Seven temperate terrestrial planets around the nearby ultracool dwarf star trappist-1. *Nature*, 542(7642):456, 2017.

[12] Ji Wang, Debra A Fischer, Elliott P Horch, and Xu Huang. On the occurrence rate of hot jupiters in different stellar environments. *The Astrophysical Journal*, 799(2):229, 2015.

[13] Stéphane Udry, X Bonfils, X Delfosse, T Forveille, M Mayor, C Perrier, François Bouchy, C Lovis, F Pepe, D Queloz, et al. The harps search for southern extra-

solar planets-xi. super-earths (5 and 8 m $_{\oplus}$) in a 3-planet system. *Astronomy & Astrophysics*, 469(3):L43–L47, 2007.

[14] Erik A Petigura, Andrew W Howard, and Geoffrey W Marcy. Prevalence of earth-size planets orbiting sun-like stars. *Proceedings of the National Academy of Sciences*, 110(48):19273–19278, 2013.

[15] Arnaud Cassan, D Kubas, J-P Beaulieu, M Dominik, K Horne, J Greenhill, J Wambsganss, J Menzies, A Williams, Uffe Gråe Jørgensen, et al. One or more bound planets per milky way star from microlensing observations. *Nature*, 481 (7380):167, 2012.

[16] Gael Chauvin, A-M Lagrange, C Dumas, B Zuckerman, D Mouillet, I Song, J-L Beuzit, and P Lowrance. A giant planet candidate near a young brown dwarf-direct vlt/naco observations using ir wavefront sensing. *Astronomy & Astrophysics*, 425(2):L29–L32, 2004.

[17] K Todorov, KL Luhman, and KK McLeod. Discovery of a planetary-mass companion to a brown dwarf in taurus. *The Astrophysical Journal Letters*, 714 (1):L84, 2010.

[18] David W Latham, Jason F Rowe, Samuel N Quinn, Natalie M Batalha, William J Borucki, Timothy M Brown, Stephen T Bryson, Lars A Buchhave, Douglas A Caldwell, Joshua A Carter, et al. A first comparison of kepler planet candidates in single and multiple systems. *The Astrophysical Journal Letters*, 732(2):L24, 2011.

[19] KL Luhman, AJ Burgasser, and JJ Bochanski. Discovery of a candidate for the coolest known brown dwarf. *The Astrophysical Journal Letters*, 730(1):L9, 2011.

[20] Masayuki Kuzuhara, Motohide Tamura, Miki Ishii, Tomoyuki Kudo, Shogo Nishiyama, and Ryo Kandori. The widest-separation substellar companion candidate to a binary t tauri star. *The Astronomical Journal*, 141(4):119, 2011.

[21] John Asher Johnson, Kimberly M Aller, Andrew W Howard, and Justin R Crepp. Giant planet occurrence in the stellar mass-metallicity plane. *Publications of the Astronomical Society of the Pacific*, 122(894):905, 2010.

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

[23] Don L Pollacco, I Skillen, A Collier Cameron, Damian J Christian, C Hellier, J Irwin, TA Lister, RA Street, Richard G West, D Anderson, et al. The wasp project and the superwasp cameras. *Publications of the Astronomical Society of the Pacific*, 118(848):1407, 2006.

[24] George R Ricker. The transiting exoplanet survey satellite (tess): Discovering new earths and super-earths in the solar neighborhood. In *AAS/Division for Extreme Solar Systems Abstracts*, volume 3, 2015.

[25] C Broeg, A Fortier, D Ehrenreich, Yann Alibert, W Baumjohann, Willy Benz, Magali Deleuil, Michaël Gillon, A Ivanov, René Liseau, et al. Cheops: A transit photometry mission for esa's small mission programme. In *EPJ Web of Conferences*, volume 47, page 03005. EDP Sciences, 2013.

[26] Heike Rauer, C Catala, Conny Aerts, T Appourchaux, W Benz, Aexis Brandeker, J Christensen-Dalsgaard, M Deleuil, Laurent Gizon, M-J Goupil, et al. The plato 2.0 mission. *Experimental Astronomy*, 38(1-2):249–330, 2014.

[27] Aleksander Wolszczan and Dail A Frail. A planetary system around the millisecond pulsar psr1257+ 12. *Nature*, 355(6356):145, 1992.

[28] F Pepe, P Molaro, S Cristiani, R Rebolo, NC Santos, H Dekker, D Mégevand, FM Zerbi, A Cabral, P Di Marcantonio, et al. Espresso: The next european exoplanet hunter. *Astronomische Nachrichten*, 335(1):8–20, 2014.

[29] C Jurgenson, D Fischer, T McCracken, D Sawyer, A Szymkowiak, A Davis, G Muller, and F Santoro. Expres: a next generation rv spectrograph in the search for earth-like worlds. In *Ground-based and Airborne Instrumentation for Astronomy VI*, volume 9908, page 99086T. International Society for Optics and Photonics, 2016.

[30] Samantha J Thompson, Didier Queloz, Isabelle Baraffe, Martyn Brake, Andrey Dolgopolov, Martin Fisher, Michel Fleury, Joost Geelhoed, Richard Hall, Jonay I González Hernández, et al. Harps3 for a roboticized isaac newton telescope. In *Ground-based and Airborne Instrumentation for Astronomy VI*, volume 9908, page 99086F. International Society for Optics and Photonics, 2016.

[31] B Scott Gaudi. Microlensing surveys for exoplanets. *Annual Review of Astronomy and Astrophysics*, 50:411–453, 2012.

[32] Takahiro Sumi, K Kamiya, DP Bennett, IA Bond, F Abe, CS Botzler, A Fukui, K Furusawa, JB Hearnshaw, Y Itow, et al. Unbound or distant planetary mass population detected by gravitational microlensing. *Nature*, 473(7347):349, 2011.

[33] Sasha Hinkley, Ben R Oppenheimer, Neil Zimmerman, Douglas Brenner, Ian R Parry, Justin R Crepp, Gautam Vasisht, Edgar Ligon, David King, Rémi Soummer, et al. A new high contrast imaging program at palomar observatory. *Publications of the Astronomical Society of the Pacific*, 123(899):74, 2011.

[34] Laird M Close, Jared R Males, Derek A Kopon, Victor Gasho, Katherine B Follette, Phil Hinz, Katie Morzinski, Alan Uomoto, Tyson Hare, Armando Riccardi, et al. First closed-loop visible ao test results for the advanced adaptive secondary ao system for the magellan telescope: Magao's performance and status. In *Adaptive Optics Systems III*, volume 8447, page 84470X. International Society for Optics and Photonics, 2012.

[35] Kjetil Dohlen, Jean-Luc Beuzit, Markus Feldt, David Mouillet, Pascal Puget, Jacopo Antichi, Andrea Baruffolo, Pierre Baudoz, Alessandro Berton, Anthony

Boccaletti, et al. Sphere: A planet finder instrument for the vlt. In *Ground-based and Airborne Instrumentation for Astronomy*, volume 6269, page 62690Q. International Society for Optics and Photonics, 2006.

[36] Bruce Macintosh et al. The gemini planet imager. Technical report, Lawrence Livermore National Laboratory (LLNL), Livermore, CA, 2006.

[37] Thayne Currie, Olivier Guyon, Frantz Martinache, Christophe Clergeon, Michael McElwain, Christian Thalmann, Nemanja Jovanovic, Garima Singh, and Tomoyuki Kudo. Scexao: First results and on-sky performance. *Proceedings of the International Astronomical Union*, 8(S299):34–35, 2013.

[38] Brendan P Bowler. Imaging extrasolar giant planets. *Publications of the Astronomical Society of the Pacific*, 128(968):102001, 2016.

[39] Justin R Crepp and John Asher Johnson. Estimates of the planet yield from ground-based high-contrast imaging observations as a function of stellar mass. *The Astrophysical Journal*, 733(2):126, 2011.

[40] John E Krist, Kunjithapatham Balasubramanian, Charles A Beichman, Pierre M Echternach, Joseph J Green, Kurt M Liewer, Richard E Muller, Eugene Serabyn, Stuart B Shaklan, John T Trauger, et al. The jwst/nircam coronagraph: mask design and fabrication. In *Techniques and Instrumentation for Detection of Exoplanets IV*, volume 7440, page 74400W. International Society for Optics and Photonics, 2009.

[41] M Charley Noecker, Feng Zhao, Rick Demers, John Trauger, Olivier Guyon, and N Jeremy Kasdin. Coronagraph instrument for wfirst-afta. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(1):011001, 2016.

[42] Markus Kasper, Jean-Luc Beuzit, Christophe Verinaud, Raffaele G Gratton, Florian Kerber, Natalia Yaitskova, Anthony Boccaletti, Niranjan Thatte, Hans Martin Schmid, Christoph Keller, et al. Epics: direct imaging of exoplanets with the e-elt. In *Ground-based and Airborne Instrumentation for*

Astronomy III, volume 7735, page 77352E. International Society for Optics and Photonics, 2010.

[43] Ben R Oppenheimer and Sasha Hinkley. High-contrast observations in optical and infrared astronomy. *Annual Review of Astronomy and Astrophysics*, 47: 253–289, 2009.

[44] David S Spiegel and Adam Burrows. Spectral and photometric diagnostics of giant planet formation scenarios. *The Astrophysical Journal*, 745(2):174, 2012.

[45] Christian Marois, Bruce Macintosh, Travis Barman, B. Zuckerman, Inseok Song, Jennifer Patience, David Lafrenière, and René Doyon. Direct Imaging of Multiple Planets Orbiting the Star HR 8799. *Science*, 322(5906):1348–1352, November 2008. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1166585. URL `http://www.sciencemag.org/content/322/5906/1348`.

[46] Christian Marois, B Zuckerman, Quinn M Konopacky, Bruce Macintosh, and Travis Barman. Images of a fourth planet orbiting hr 8799. *Nature*, 468(7327): 1080, 2010.

[47] A-M Lagrange, D Gratadour, G Chauvin, T Fusco, D Ehrenreich, D Mouillet, G Rousset, D Rouan, F Allard, É Gendron, et al. A probable giant planet imaged in the β pictoris disk-vlt/naco deep l'-band imaging. *Astronomy & Astrophysics*, 493(2):L21–L25, 2009.

[48] Ignas AG Snellen, Bernhard R Brandl, Remco J de Kok, Matteo Brogi, Jayne Birkby, and Henriette Schwarz. Fast spin of the young extrasolar planet β pictoris b. *Nature*, 509(7498):63, 2014.

[49] M Kuzuhara, M Tamura, T Kudo, M Janson, R Kandori, TD Brandt, C Thalmann, D Spiegel, B Biller, J Carson, et al. Direct imaging of a cold jovian exoplanet in orbit around the sun-like star gj 504. *The Astrophysical Journal*, 774(1):11, 2013.

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

[51] J Rameau, G Chauvin, A-M Lagrange, A Boccaletti, Sascha P Quanz, M Bonnefoy, JH Girard, P Delorme, S Desidera, H Klahr, et al. Discovery of a probable 4-5 jupiter-mass exoplanet to hd 95086 by direct imaging. *The Astrophysical Journal Letters*, 772(2):L15, 2013.

[52] J Rameau, G Chauvin, A-M Lagrange, T Meshkat, A Boccaletti, SP Quanz, T Currie, D Mawet, JH Girard, M Bonnefoy, et al. Confirmation of the planet around hd 95086 by direct imaging. *The Astrophysical Journal Letters*, 779(2): L26, 2013.

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

[54] Iau 2006 general assembly. `https://www.iau.org/news/pressreleases/detail/iau0603/`. Accessed: 2018-06-21.

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

[56] Sarah E Dodson-Robinson, Dimitri Veras, Eric B Ford, and Charles A Beichman. The formation mechanism of gas giants on wide orbits. *The Astrophysical Journal*, 707(1):79, 2009.

[57] Thayne Currie. Hr 8799: The benchmark directly-imaged planetary system. *arXiv preprint arXiv:1607.03980*, 2016.

[58] James Kasting, W Traub, A Roberge, A Leger, A Schwartz, A Wooten, A Vosteen, A Lo, A Brack, A Tanner, et al. Exoplanet characterization and the search for life. *arXiv preprint arXiv:0911.2936*, 2009.

[59] M Bonavita, G Chauvin, S Desidera, R Gratton, M Janson, JL Beuzit, M Kasper, and C Mordasini. Mess (multi-purpose exoplanet simulation system)-a monte carlo tool for the statistical analysis and prediction of exoplanet search results. *Astronomy & Astrophysics*, 537:A67, 2012.

[60] Gilles PPL Otten, Frans Snik, Matthew A Kenworthy, Christoph U Keller, Jared R Males, Katie M Morzinski, Laird M Close, Johanan L Codona, Philip M Hinz, Kathryn J Hornburg, et al. On-sky performance analysis of the vector apodizing phase plate coronagraph on magao/clio2. *The Astrophysical Journal*, 834(2):175, 2017.

[61] Isabelle Baraffe, Derek Homeier, France Allard, and Gilles Chabrier. New evolutionary models for pre-main sequence and main sequence low-mass stars down to the hydrogen-burning limit. *Astronomy & Astrophysics*, 577:A42, 2015.

[62] René Racine, Gordon AH Walker, Daniel Nadeau, René Doyon, and Christian Marois. Speckle noise and the detection of faint companions. *Publications of the Astronomical Society of the Pacific*, 111(759):587, 1999.

[63] Christian Marois, David Lafreniere, Bruce Macintosh, and René Doyon. Confidence level and sensitivity limits in high-contrast imaging. *The Astrophysical Journal*, 673(1):647, 2008.

[64] C Aime and R Soummer. The usefulness and limits of coronagraphy in the presence of pinned speckles. *The Astrophysical Journal Letters*, 612(1):L85, 2004.

[65] Christian Marois, Don W Phillion, and Bruce Macintosh. Exoplanet detection with simultaneous spectral differential imaging: effects of out-of-pupil-plane

optical aberrations. In *Ground-based and Airborne Instrumentation for Astronomy*, volume 6269, page 62693M. International Society for Optics and Photonics, 2006.

[66] Szymon Gladysz and Julian C Christou. Detection of faint companions through stochastic speckle discrimination. *The Astrophysical Journal*, 684(2):1486, 2008.

[67] Sasha Hinkley, Ben R Oppenheimer, Rémi Soummer, Anand Sivaramakrishnan, Lewis C Roberts Jr, Jeffrey Kuhn, Russell B Makidon, Marshall D Perrin, James P Lloyd, Kaitlin Kratter, et al. Temporal evolution of coronagraphic dynamic range and constraints on companions to vega. *The Astrophysical Journal*, 654(1):633, 2007.

[68] Rémi Soummer, André Ferrari, Claude Aime, and Laurent Jolissaint. Speckle noise and dynamic range in coronagraphic images. *The Astrophysical Journal*, 669(1):642, 2007.

[69] Michael P Fitzgerald and James R Graham. Speckle statistics in adaptively corrected images. *The Astrophysical Journal*, 637(1):541, 2006.

[70] Joseph W Goodman. Statistical properties of laser speckle patterns. In *Laser speckle and related phenomena*, pages 9–75. Springer, 1975.

[71] André Maréchal. *Étude des effets combinés de la diffraction et des aberrations géométriques sur l'image d'un point lumineux...* Éditions de la Revue d'optique théorique et instrumentale, 1948.

[72] Laird M Close, Beth Biller, William F Hoffmann, Phil M Hinz, John H Bieging, Francois Wildi, Michael Lloyd-Hart, Guido Brusa, Don Fisher, Doug Miller, et al. Mid-infrared imaging of the post-asymptotic giant branch star ac herculis with the multiple mirror telescope adaptive optics system. *The Astrophysical Journal Letters*, 598(1):L35, 2003.

[73] E Serabyn, K Wallace, M Troy, B Mennesson, P Haguenauer, R Gappinger, and

R Burruss. Extreme adaptive optics imaging with a clear and well-corrected off-axis telescope subaperture. *The Astrophysical Journal*, 658(2):1386, 2007.

[74] Julien Milli, Dimitri Mawet, David Mouillet, Markus Kasper, and Julien H Girard. Adaptive optics in high-contrast imaging. In *Astronomy at High Angular Resolution*, pages 17–41. Springer, 2016.

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

[76] Bruce A Macintosh, James R Graham, David W Palmer, René Doyon, Jennifer Dunn, Donald T Gavel, James Larkin, Ben Oppenheimer, Les Saddlemyer, Anand Sivaramakrishnan, et al. The gemini planet imager: from science to design to construction. In *Adaptive Optics Systems*, volume 7015, page 701518. International Society for Optics and Photonics, 2008.

[77] Marc J Kuchner and Wesley A Traub. A coronagraph with a band-limited mask for finding terrestrial planets. *The Astrophysical Journal*, 570(2):900, 2002.

[78] Olivier Guyon, EA Pluzhnik, MJ Kuchner, B Collins, and ST Ridgway. Theoretical limits on extrasolar terrestrial planet detection with coronagraphs. *The Astrophysical Journal Supplement Series*, 167(1):81, 2006.

[79] P Riaud, A Boccaletti, D Rouan, F Lemarquis, and A Labeyrie. The four-quadrant phase-mask coronagraph. ii. simulations. *Publications of the Astronomical Society of the Pacific*, 113(787):1145, 2001.

[80] Christophe Lovis, Ignas Snellen, David Mouillet, Francesco Pepe, François Wildi, Nicola Astudillo-Defru, J-L Beuzit, Xavier Bonfils, Anthony Cheetham, Uriel Conod, et al. Atmospheric characterization of proxima b by coupling the sphere high-contrast imager to the espresso spectrograph. *Astronomy & Astrophysics*, 599:A16, 2017.

[81] D Mawet, Pierre Riaud, Olivier Absil, and Jean Surdej. Annular groove phase mask coronagraph. *The Astrophysical Journal*, 633(2):1191, 2005.

[82] D Mawet, E Serabyn, K Liewer, R Burruss, J Hickey, and D Shemo. The vector vortex coronagraph: laboratory results and first light at palomar observatory. *The Astrophysical Journal*, 709(1):53, 2009.

[83] Géraldine Guerri, Sylvie Robbe-Dubois, Jean-Baptiste Daban, Lyu Abe, Richard Douet, Philippe Bendjoya, Farrokh Vakili, Marcel Carbillet, Jean-Luc Beuzit, Pascal Puget, et al. Apodized lyot coronagraph for vlt-sphere: Laboratory tests and performances of a first prototype in the visible. In *Ground-based and Airborne Instrumentation for Astronomy II*, volume 7014, page 70143J. International Society for Optics and Photonics, 2008.

[84] Dimitri Mawet, Laurent Pueyo, Peter Lawson, Laurent Mugnier, Wesley Traub, Anthony Boccaletti, John T Trauger, Szymon Gladysz, Eugene Serabyn, Julien Milli, et al. Review of small-angle coronagraphic techniques in the wake of ground-based second-generation adaptive optics systems. In *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, volume 8442, page 844204. International Society for Optics and Photonics, 2012.

[85] Olivier Guyon, Taro Matsuo, and Roger Angel. Coronagraphic low-order wavefront sensor: principle and application to a phase-induced amplitude coronagraph. *The Astrophysical Journal*, 693(1):75, 2009.

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

[87] P Baudoz, A Boccaletti, J Baudrand, and D Rouan. The self-coherent camera: a new tool for planet detection. *Proceedings of the International Astronomical Union*, 1(C200):553–558, 2005.

[88] F Malbet, JW Yu, and M Shao. High-dynamic-range imaging using a de-

formable mirror for space coronography. *Publications of the Astronomical Society of the Pacific*, 107(710):386, 1995.

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

[90] Laurent Pueyo, Jason Kay, N Jeremy Kasdin, Tyler Groff, Michael McElwain, Amir Give'on, and Ruslan Belikov. Optimal dark hole generation via two deformable mirrors with stroke minimization. *Applied optics*, 48(32):6296–6312, 2009.

[91] Michael Bottom, J Kent Wallace, Randall D Bartos, J Chris Shelton, and Eugene Serabyn. Speckle suppression and companion detection using coherent differential imaging. *Monthly Notices of the Royal Astronomical Society*, 464 (3):2937–2951, 2016.

[92] Dimitri Mawet, Julien Milli, Zahed Wahhaj, Didier Pelat, Olivier Absil, Christian Delacroix, Anthony Boccaletti, Markus Kasper, Matthew Kenworthy, Christian Marois, et al. Fundamental limitations of high contrast imaging set by small sample statistics. *The Astrophysical Journal*, 792(2):97, 2014.

[93] Christian Marois, David Lafreniere, Rene Doyon, Bruce Macintosh, and Daniel Nadeau. Angular differential imaging: A powerful high-contrast imaging techniquebased on observations obtained at the gemini observatory, which is operated by the association of universities for research in astronomy, inc., under a cooperative agreement with the nsf on behalf of the gemini partnership: the national science foundation (united states), the particle physics and astronomy research council (united kingdom), the national research council (canada), conicyt (chile), the australian research council (australia), cnpq (brazil), and conicet (argentina). *The Astrophysical Journal*, 641(1):556, 2006.

[94] William B Sparks and Holland C Ford. Imaging spectroscopy for extrasolar planet detection. *The Astrophysical Journal*, 578(1):543, 2002.

[95] Markus Janson, Timothy D Brandt, Masayuki Kuzuhara, David S Spiegel, Christian Thalmann, Thayne Currie, Mickaël Bonnefoy, Neil Zimmerman, Satoko Sorahana, Takayuki Kotani, et al. Direct imaging detection of methane in the atmosphere of gj 504 b. *The Astrophysical Journal Letters*, 778(1):L4, 2013.

[96] Jeffrey R Kuhn, D Potter, and B Parise. Imaging polarimetric observations of a new circumstellar disk system. *The Astrophysical Journal Letters*, 553(2):L189, 2001.

[97] David Lafreniere, Christian Marois, Rene Doyon, Daniel Nadeau, and Etienne Artigau. A new algorithm for point-spread function subtraction in high-contrast imaging: a demonstration with angular differential imaging. *The Astrophysical Journal*, 660(1):770, 2007.

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

[99] F Cantalloube, D Mouillet, LM Mugnier, J Milli, Olivier Absil, CA Gomez Gonzalez, G Chauvin, J-L Beuzit, and A Cornia. Direct exoplanet detection and characterization using the andromeda method: Performance on vlt/naco data. *Astronomy & Astrophysics*, 582:A89, 2015.

[100] CA Gomez Gonzalez, Olivier Absil, P-A Absil, Marc Van Droogenbroeck, Dimitri Mawet, and Jean Surdej. Low-rank plus sparse decomposition for exoplanet detection in direct-imaging adi sequences-the llsg algorithm. *Astronomy & Astrophysics*, 589:A54, 2016.

[101] A. Labeyrie. Images of exo-planets obtainable from dark speckles in adaptive telescopes. *Astronomy and Astrophysics*, 298:544, June 1995. ISSN 0004-6361. URL `http://adsabs.harvard.edu/abs/1995A%26A...298..544L`.

[102] Willard S Boyle and George E Smith. Charge coupled semiconductor devices. *Bell Labs Technical Journal*, 49(4):587–593, 1970.

[103] James R Janesick. *Scientific charge-coupled devices*, volume 83. SPIE press, 2001.

[104] Marvin H White, Donald R Lampe, Franklyn C Blaha, and Ingham A Mack. Characterization of surface channel ccd image arrays at low light levels. *IEEE Journal of Solid-State Circuits*, 9(1):1–12, 1974.

[105] Javier Tiffenberg, Miguel Sofo-Haro, Alex Drlica-Wagner, Rouven Essig, Yann Guardincerri, Steve Holland, Tomer Volansky, and Tien-Tien Yu. Single-electron and single-photon sensitivity with a silicon skipper ccd. *Physical review letters*, 119(13):131802, 2017.

[106] Antony Rogalski. Hgcdte infrared detector material: history, status and outlook. *Reports on Progress in Physics*, 68(10):2267, 2005.

[107] *HAWAII-2RG Visible & Infrared Focal Plane Array.* Teledyne Imaging Sensors, 2018. URL `http://www.oir.caltech.edu/twiki_oir/pub/Keck/NGAO/NIRTTS/H2RG_Brochure_-_Approved_for_Public_Release.pdf`.

[108] Jean-Luc Beuzit, Markus Feldt, Kjetil Dohlen, David Mouillet, Pascal Puget, Francois Wildi, Lyu Abe, Jacopo Antichi, Andrea Baruffolo, Pierre Baudoz, et al. Sphere: a planet finder instrument for the vlt. In *Ground-based and airborne instrumentation for astronomy II*, volume 7014, page 701418. International Society for Optics and Photonics, 2008.

[109] James E Larkin, Jeffrey K Chilcote, Theodore Aliado, Brian J Bauman, George Brims, John M Canfield, Andrew Cardwell, Daren Dillon, René Doyon, Jennifer Dunn, et al. The integral field spectrograph for the gemini planet imager. In *Ground-based and Airborne Instrumentation for Astronomy V*, volume 9147, page 91471K. International Society for Optics and Photonics, 2014.

[110] Kent D Irwin and Gene C Hilton. Transition-edge sensors. In *Cryogenic particle detection*, pages 63–150. Springer, 2005.

[111] WS Holland, D Bintley, EL Chapin, Antonio Chrysostomou, GR Davis, JT Dempsey, WD Duncan, M Fich, P Friberg, M Halpern, et al. Scuba-2: the 10 000 pixel bolometer camera on the james clerk maxwell telescope. *Monthly Notices of the Royal Astronomical Society*, 430(4):2513–2533, 2013.

[112] B Westbrook, A Cukierman, A Lee, A Suzuki, C Raum, and W Holzapfel. Development of the next generation of multi-chroic antenna-coupled transition edge sensor detectors for cmb polarimetry. *Journal of Low Temperature Physics*, 184(1-2):74–81, 2016.

[113] Gianfranco De Zotti. Prospects for next generation cosmic microwave background experiments. *arXiv preprint arXiv:1802.04753*, 2018.

[114] A Peacock, P Verhoeve, N Rando, A Van Dordrecht, BG Taylor, C Erd, MAC Perryman, R Venn, J Howlett, DJ Goldie, et al. Single optical photon detection with a superconducting tunnel junction. *Nature*, 381(6578):135, 1996.

[115] J. Zmuidzinas and P.L. Richards. Superconducting detectors and mixers for millimeter and submillimeter astrophysics. *Proceedings of the IEEE*, 92(10): 1597–1616, October 2004. ISSN 0018-9219. doi: 10.1109/JPROC.2004.833670.

[116] W. S. Holland, D. Bintley, E. L. Chapin, A. Chrysostomou, G. R. Davis, J. T. Dempsey, W. D. Duncan, M. Fich, P. Friberg, M. Halpern, K. D. Irwin, T. Jenness, B. D. Kelly, M. J. MacIntosh, E. I. Robson, D. Scott, P. A. R. Ade, E. Atad-Ettedgui, D. S. Berry, S. C. Craig, X. Gao, A. G. Gibb, G. C. Hilton, M. I. Hollister, J. B. Kycia, D. W. Lunney, H. McGregor, D. Montgomery, W. Parkes, R. P. J. Tilanus, J. N. Ullom, C. A. Walther, A. J. Walton, A. L. Woodcraft, M. Amiri, D. Atkinson, B. Burger, T. Chuter, I. M. Coulson, W. B. Doriese, C. Dunare, F. Economou, M. D. Niemack, H. A. L. Parsons, C. D. Reintsema, B. Sibthorpe, I. Smail, R. Sudiwala, and H. S. Thomas.

SCUBA-2: The 10000 pixel bolometer camera on the James Clerk Maxwell Telescope. *Monthly Notices of the Royal Astronomical Society*, 430(4):2513–2533, April 2013. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/sts612. URL http://arxiv.org/abs/1301.3650. arXiv: 1301.3650.

[117] Go Fujii, Masahiro Ukibe, Shigetomo Shiki, and Masataka Ohkubo. Development of Array Detectors with Three-Dimensional Structure toward 1000 Pixels of Superconducting Tunnel Junctions. *IEICE Transactions on Electronics*, E98.C(3):192–195, 2015. doi: 10.1587/transele.E98.C.192.

[118] Peter K. Day, Henry G. LeDuc, Benjamin A. Mazin, Anastasios Vayonakis, and Jonas Zmuidzinas. A broadband superconducting detector suitable for use in large arrays. *Nature*, 425(6960):817–821, October 2003. ISSN 0028-0836, 1476-4679. doi: 10.1038/nature02037. URL http://www.nature.com/doifinder/10.1038/nature02037.

[119] BA Mazin, Seth R Meeker, MJ Strader, P Szypryt, D Marsden, JC van Eyken, GE Duggan, AB Walter, G Ulbricht, M Johnson, et al. Arcons: A 2024 pixel optical through near-ir cryogenic imaging spectrophotometer. *Publications of the Astronomical Society of the Pacific*, 125(933):1348, 2013.

[120] P. Szypryt, G. E. Duggan, B. A. Mazin, S. R. Meeker, M. J. Strader, J. C. van Eyken, D. Marsden, K. O'Brien, A. B. Walter, G. Ulbricht, T. A. Prince, C. Stoughton, and B. Bumble. Direct detection of SDSS J0926+3624 orbital expansion with ARCONS. *Monthly Notices of the Royal Astronomical Society*, 439(3):2765–2770, November 2014. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stu137. URL http://mnras.oxfordjournals.org/content/439/3/2765.

[121] MJ Strader, AM Archibald, SR Meeker, P Szypryt, AB Walter, JC van Eyken, G Ulbricht, C Stoughton, B Bumble, DL Kaplan, et al. Search for optical pulsations in psr j0337+ 1715. *Monthly Notices of the Royal Astronomical Society*, 459(1):427–430, 2016.

[122] Seth Meeker, Benjamin Mazin, Rebecca Jensen-Clem, Alex Walter, Paul Szypryt, Matthew Strader, and Clint Bockstiegel. Design and Development Status of MKID Integral Field Spectrographs for High Contrast Imaging. *Adaptive Optics for Extremely Large Telescopes 4 - Conference Proceedings*, 1(1), January 2015. doi: 10.20353/K3T4CP1131701. URL `http://escholarship.org/uc/item/217686nz`.

[123] Danica W. Marsden, Benjamin A. Mazin, Kieran O'Brien, and Chris Hirata. Giga-z: A 100,000 OBJECT SUPERCONDUCTING SPECTROPHOTOMETER FOR LSST FOLLOW-UP. *The Astrophysical Journal Supplement Series*, 208(1):8, September 2013. ISSN 0067-0049, 1538-4365. doi: 10.1088/0067-0049/208/1/8. URL `http://stacks.iop.org/0067-0049/208/i=1/a=8?key=crossref.67caf0d5ab9df225b42df73326f3802a`.

[124] Seth Meeker, Ben Mazin, Alex Walter, Mathew Strader, Neelay Fruitwala, Clint Bockstiegel, Paul Szypryt, Gerhard Ulbricht, Gregoire Coiffard, Bruce Bumble, Guila Collura, and **Rupert Dodkins** et al. DARKNESS: A Microwave Kinetic Inductance Detector Integral Field Spectrograph for High-Contrast Astronomy. *Publications of the Astronomical Society of the Pacific*, submitted.

[125] P Szypryt, SR Meeker, G Coiffard, N Fruitwala, B Bumble, G Ulbricht, AB Walter, M Daal, C Bockstiegel, G Collura, et al. Large-format platinum silicide microwave kinetic inductance detectors for optical to near-ir astronomy. *arXiv preprint arXiv:1710.07318*, 2017.

[126] Ulbricht Gerhard. personal communication.

[127] S. Doyle, P. Mauskopf, J. Naylon, A. Porch, and C. Duncombe. Lumped Element Kinetic Inductance Detectors. *Journal of Low Temperature Physics*, 151(1-2):530–536, April 2008. ISSN 0022-2291, 1573-7357. doi: 10.1007/s10909-007-9685-2. URL `http://link.springer.com/10.1007/s10909-007-9685-2`.

[128] Jonas Zmuidzinas. Superconducting Microresonators: Physics and Applications. *Annual Review of Condensed Matter Physics*, 3(1):169–214, 2012. doi: 10.1146/annurev-conmatphys-020911-125022.

[129] P. Szypryt, B. A. Mazin, G. Ulbricht, B. Bumble, S. R. Meeker, C. Bockstiegel, and A. B. Walter. High quality factor platinum silicide microwave kinetic inductance detectors. *Applied Physics Letters*, 109(15):151102, 2016. URL `http://aip.scitation.org/doi/abs/10.1063/1.4964665`.

[130] F. W. Carter, T. S. Khaire, V. Novosad, and C. L. Chang. scraps: An Open-Source Python-Based Analysis Package for Analyzing and Plotting Superconducting Resonator Data. *IEEE Transactions on Applied Superconductivity*, 27 (4):1–5, June 2017. ISSN 1051-8223. doi: 10.1109/TASC.2016.2625767.

[131] Kurtis Lee Geerlings. *Improving coherence of superconducting qubits and resonators*. Yale University, 2013.

[132] MS Khalil, MJA Stoutimore, FC Wellstood, and KD Osborn. An analysis method for asymmetric resonator transmission applied to superconducting devices. *Journal of Applied Physics*, 111(5):054510, 2012.

[133] C. N. Thomas, S. Withington, and D. J. Goldie. Electrothermal model of kinetic inductance detectors. *Superconductor Science and Technology*, 28(4): 045012, 2015. ISSN 0953-2048. doi: 10.1088/0953-2048/28/4/045012. URL `http://stacks.iop.org/0953-2048/28/i=4/a=045012`.

[134] AV Semenov, IA Devyatov, PJ de Visser, and TM Klapwijk. Coherent excited states in superconductors due to a microwave field. *Physical review letters*, 117 (4):047002, 2016.

[135] L. J. Swenson, P. K. Day, B. H. Eom, H. G. Leduc, N. Llombart, C. M. McKenney, O. Noroozian, and J. Zmuidzinas. Operation of a titanium nitride superconducting microresonator detector in the nonlinear regime. *Journal of Applied Physics*, 113(10):104501, 2013. ISSN 0021-8979. doi: 10.1063/1.4794808.

[136] DC Mattis and John Bardeen. Theory of the anomalous skin effect in normal and superconducting metals. *Physical Review*, 111(2):412, 1958.

[137] John Bardeen, Leon N Cooper, and John Robert Schrieffer. Theory of superconductivity. *Physical Review*, 108(5):1175, 1957.

[138] Benjamin A. Mazin. *Microwave kinetic inductance detectors*. PhD thesis, California Institute of Technology, 2005. URL `http://thesis.library.caltech.edu/3910/`.

[139] Paul Drude. Zur elektronentheorie der metalle. *Annalen der Physik*, 306(3): 566–613, 1900.

[140] AG Kozorezov, AF Volkov, JK Wigmore, A Peacock, A Poelaert, and R Den Hartog. Quasiparticle-phonon downconversion in nonequilibrium superconductors. *Physical Review B*, 61(17):11807, 2000.

[141] Jiansong Gao. *The physics of superconducting microwave resonators*. PhD thesis, California Institute of Technology, 2008.

[142] Benjamin A. Mazin, Bruce Bumble, Seth R. Meeker, Kieran O'Brien, Sean McHugh, and Eric Langman. A superconducting focal plane array for ultraviolet, optical, and near-infrared astrophysics. *Optics Express*, 20(2):1503, January 2012. ISSN 1094-4087. doi: 10.1364/OE.20.001503. URL `http://arxiv.org/abs/1112.0004`. arXiv: 1112.0004.

[143] Benjamin A Mazin, Bruce Bumble, Seth R Meeker, Kieran O'Brien, Sean McHugh, and Eric Langman. A superconducting focal plane array for ultraviolet, optical, and near-infrared astrophysics. *Optics express*, 20(2):1503–1511, 2012.

[144] M. J. Strader. *Digital Readout for Microwave Kinetic Inductance Detectors and Applications in High Time Resolution Astronomy*. PhD thesis, California Institute of Technology, 2016.

[145] Aaron Parsons, Donald Backer, Chen Chang, Daniel Chapman, Henry Chen, Patrick Crescini, Christina De Jesus, Chris Dick, Pierre Droz, David MacMahon, et al. Petaop/second fpga signal processing for seti and radio astronomy. In *Signals, Systems and Computers, 2006. ACSSC'06. Fortieth Asilomar Conference on*, pages 2031–2035. IEEE, 2006.

[146] Sean McHugh, Benjamin A. Mazin, Bruno Serfass, Seth Meeker, Kieran O'Brien, Ran Duan, Rick Raffanti, and Dan Werthimer. A readout for large arrays of Microwave Kinetic Inductance Detectors. *Review of Scientific Instruments*, 83(4):044702, 2012. ISSN 00346748. doi: 10.1063/1.3700812. URL `http://arxiv.org/abs/1203.5861`. arXiv: 1203.5861.

[147] Jiansong Gao, Miguel Daal, John M Martinis, Anastasios Vayonakis, Jonas Zmuidzinas, Bernard Sadoulet, Benjamin A Mazin, Peter K Day, and Henry G Leduc. A semiempirical model for two-level system noise in superconducting microresonators. *Applied Physics Letters*, 92(21):212504, 2008.

[148] Omid Noroozian, Jiansong Gao, Jonas Zmuidzinas, Henry G LeDuc, and Benjamin A Mazin. Two-level system noise reduction for microwave kinetic inductance detectors. In *AIP Conference Proceedings*, number 1, pages 148–151. AIP, 2009.

[149] Byeong Ho Eom, Peter K Day, Henry G LeDuc, and Jonas Zmuidzinas. A wideband, low-noise superconducting amplifier with high dynamic range. *Nature Physics*, 8(8):623, 2012.

[150] Saptarshi Chaudhuri, Dale Li, KD Irwin, Clint Bockstiegel, Johannes Hubmayr, JN Ullom, MR Vissers, and Jiansong Gao. Broadband parametric amplifiers based on nonlinear kinetic inductance artificial transmission lines. *Applied Physics Letters*, 110(15):152601, 2017.

[151] Boon-Kok Tan and Ghassan Yassin. Design of a uniplanar resonance phase-matched josephson traveling-wave parametric amplifier. In *Millimetre Waves*

and Terahertz Technologies (UCMMT), 2017 10th UK-Europe-China Workshop on, pages 1–4. IEEE, 2017.

[152] Gregoire Coiffard, Benjamin Mazin, Miguel Daal, Paul Szypryt, Gerhard Ulbricht, and Nicholas Zobrist. Parallel plate microwave kinetic inductance detectors. *Applied Physics Letters*, 2017 submitted.

[153] Kieran O'Brien, Niranjan Thatte, and Benjamin Mazin. KIDSpec: an MKID based medium resolution integral field spectrograph. volume 9147, pages 91470G–91470G–8, 2014. doi: 10.1117/12.2056297.

[154] B. A. Mazin, S. R. Meeker, M. J. Strader, P. Szypryt, D. Marsden, J. C. van Eyken, G. E. Duggan, A. B. Walter, G. Ulbricht, M. Johnson, B. Bumble, K. O'Brien, and C. Stoughton. ARCONS: A 2024 Pixel Optical through Near-IR Cryogenic Imaging Spectrophotometer. *Publications of the Astronomical Society of the Pacific*, 125(933):1348–1361, November 2013. ISSN 00046280, 15383873. doi: 10.1086/674013. URL `http://www.jstor.org/stable/info/10.1086/674013`.

[155] Henry G Leduc, Bruce Bumble, Peter K Day, Byeong Ho Eom, Jiansong Gao, Sunil Golwala, Benjamin A Mazin, Sean McHugh, Andrew Merrill, David C Moore, et al. Titanium nitride films for ultrasensitive microresonator detectors. *Applied Physics Letters*, 97(10):102509, 2010.

[156] M. J. Strader, M. D. Johnson, B. A. Mazin, G. V. Spiro Jaeger, C. R. Gwinn, S. R. Meeker, P. Szypryt, J. C. van Eyken, D. Marsden, K. O'Brien, A. B. Walter, G. Ulbricht, C. Stoughton, and B. Bumble. EXCESS OPTICAL ENHANCEMENT OBSERVED WITH ARCONS FOR EARLY CRAB GIANT PULSES. *The Astrophysical Journal*, 779(1):L12, December 2013. ISSN 2041-8205, 2041-8213. doi: 10.1088/2041-8205/779/1/L12.

[157] A Walter, BA Mazin, C Bockstiegel, et al. Mec: the mkid exoplanet camera for high contrast astronomy at subaru. *Adaptive Optics System VI], Proc. SPIE*, pages 10702–31, 2018.

[158] N Jovanovic, Frantz Martinache, Olivier Guyon, Christophe Clergeon, Garima Singh, Tomoyuki Kudo, Vincent Garrel, Kevin Newman, D Doughty, Julien Lozi, et al. The subaru coronagraphic extreme adaptive optics system: enabling high-contrast imaging on solar-system scales. *Publications of the Astronomical Society of the Pacific*, 127(955):890, 2015.

[159] Christopher Mendillo, Kuravi Hewawasam, Glenn A Howe, Jason Martel, Susanna Finn, Timothy Cook, and Supriya Chakrabarti. Picture-c: A nasa balloon mission to directly image exozodiacal dust around nearby stars. In *American Astronomical Society Meeting Abstracts*, volume 231, 2018.

[160] Mark Cropper, M. Barlow, M. A. C. Perryman, Keith Horne, R. Bingham, M. Page, P. Guttridge, A. Smith, A. Peacock, D. Walker, and P. Charles. A concept for a superconducting tunnelling junction based spectrograph. *Monthly Notices of the Royal Astronomical Society*, 344:33–44, September 2003. ISSN 0035-8711. doi: 10.1046/j.1365-8711.2003.06772.x. URL `http://adsabs.harvard.edu/abs/2003MNRAS.344...33C`.

[161] Steven S Vogt, Steven L Allen, Bruce C Bigelow, L Bresee, William E Brown, T Cantrall, Albert Conrad, M Couture, C Delaney, Harland W Epps, et al. Hires: the high-resolution echelle spectrometer on the keck 10-m telescope. In *Instrumentation in Astronomy VIII*, volume 2198, pages 362–376. International Society for Optics and Photonics, 1994.

[162] Sumedh Mahashabde, O'Brien Kieran, Dodkins Rupert, Hepburn Iain, Bartlett Jo, Hardy Graham, Mazin Ben, and Yassin Ghassan. Poster: Development of the kinetic inductance detector spectrograph (kidspec) prototype. 2016.

[163] PJ de Visser, S Withington, and DJ Goldie. Readout-power heating and hysteretic switching between thermal quasiparticle states in kinetic inductance detectors. *Journal of Applied Physics*, 108(11):114504, 2010.

[164] Yoshua Bengio, Yann LeCun, et al. Scaling learning algorithms towards ai. *Large-scale kernel machines*, 34(5):1–41, 2007.

[165] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521 (7553):436–444, 2015.

[166] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[167] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10): 1995, 1995.

[168] Sergey Ioffe and Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv:1502.03167 [cs]*, February 2015. URL `http://arxiv.org/abs/1502.03167`. 01099 arXiv: 1502.03167.

[169] Nitish Srivastava, Geoffrey E Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *Journal of machine learning research*, 15(1):1929–1958, 2014.

[170] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *preprint arXiv:1603.04467*, 2016.

[171] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *preprint arXiv:1412.6980*, 2014.

[172] Laurent Jolissaint, Jean-Pierre Véran, and Rodolphe Conan. Analytical modeling of adaptive optics: foundations of the phase spatial power spectrum approach. *JOSA A*, 23(2):382–394, 2006.

[173] François Rigaut1a and Marcos Van Dam. Simulating astronomical adaptive optics systems using yao. 2013.

[174] M Carbillet, C Verinaud, B Femenia, A Riccardi, and L Fini. Caos: Code for adaptive optics systems. *Astrophysics Source Code Library*, 2011.

[175] Andrey Nikolaevich Kolmogorov. Dissipation of energy in locally isotropic turbulence. In *Dokl. Akad. Nauk SSSR*, volume 32, pages 16–18, 1941.

[176] NS Nightingale and DF Buscher. Interferometric seeing measurements at the la palma observatory. *Monthly Notices of the Royal Astronomical Society*, 251 (1):155–166, 1991.

[177] W Dali Ali, A Ziad, A Berdja, J Maire, J Borgnino, M Sarazin, G Lombardi, J Navarrete, H Vazquez Ramio, M Reyes, et al. Multi-instrument measurement campaign at paranal in 2007-characterization of the outer scale and the seeing of the surface layer. *Astronomy & Astrophysics*, 524:A73, 2010.

[178] Changhui Rao, Wenhan Jiang, and Ning Ling. Spatial and temporal characterization of phase fluctuations in non-kolmogorov atmospheric turbulence. *Journal of Modern Optics*, 47(6):1111–1126, 2000.

[179] RG Lane, A Glindemann, JC Dainty, et al. Simulation of a kolmogorov phase screen. *Waves in random media*, 2(3):209–224, 1992.

[180] Andrei Tokovinin. Where is the surface-layer turbulence? In *Ground-based and Airborne Telescopes III*, volume 7733, page 77331N. International Society for Optics and Photonics, 2010.

[181] RE Hufnagel. Optical propagation through turbulence. *Digest of Technical Papers*, 69, 1974.

[182] Eugene L Church and Peter Z Takacs. Optimal estimation of finish parameters. In *Optical Scatter: Applications, Measurement, and Theory*, volume 1530, pages 71–86. International Society for Optics and Photonics, 1991.

[183] P Martinez, C Loose, E Aller Carpentier, and M Kasper. Speckle temporal stability in xao coronagraphic images. *Astronomy & Astrophysics*, 541:A136, 2012.

[184] P Martinez, M Kasper, A Costille, JF Sauvage, K Dohlen, P Puget, and JL Beuzit. Speckle temporal stability in xao coronagraphic images-ii. refine model for quasi-static speckle temporal evolution for vlt/sphere. *Astronomy & Astrophysics*, 554:A41, 2013.

[185] Mark A Ealey and John T Trauger. High-density deformable mirrors to enable coronographic planet detection. In *UV/Optical/IR Space Telescopes: Innovative Technologies and Concepts*, volume 5166, pages 172–180. International Society for Optics and Photonics, 2004.

[186] Marc J Kuchner, Justin Crepp, and Jian Ge. Eighth-order image masks for terrestrial planet finding. *The Astrophysical Journal*, 628(1):466, 2005.

[187] A Labeyrie. Images of exo-planets obtainable from dark speckles in adaptive telescopes. *Astronomy and Astrophysics*, 298:544, 1995.

[188] A Boccaletti, A Labeyrie, and R Ragazzoni. Preliminary results of dark-speckle stellar coronography. *arXiv preprint astro-ph/9806144*, 1998.

[189] Erez N Ribak and Szymon Gladysz. Fainter and closer: finding planets by symmetry breaking. *Optics Express*, 16(20):15553–15562, 2008.

[190] Szymon Gladysz, Julian C Christou, and Michael Redfern. Characterization of the lick adaptive optics point spread function. In *Advances in Adaptive Optics II*, volume 6272, page 62720J. International Society for Optics and Photonics, 2006.

[191] J Christou, S Gladysz, M Redfern, L Bradford, and L Roberts Jr. Characterization of the variability of the strehl ratio of adaptive optics point spread functionbs. In *The Advanced Maui Optical and Space Surveillance Technologies Conference*, page E57, 2006.

[192] Rémi Soummer and André Ferrari. The strehl ratio in adaptive optics images: statistics and estimation. *The Astrophysical Journal Letters*, 663(1):L49, 2007.

[193] Szymon Gladysz, Julian C Christou, L William Bradford, and Lewis C Roberts. Temporal variability and statistics of the strehl ratio in adaptive-optics images. *Publications of the Astronomical Society of the Pacific*, 120(872):1132, 2008.

[194] Szymon Gladysz, Natalia Yaitskova, and Julian C Christou. Statistics of intensity in adaptive-optics images and their usefulness for detection and photometry of exoplanets. *JOSA A*, 27(11):A64–A75, 2010.

[195] You-Hua Chu, Arturo Manchado, George H Jacoby, and Karen B Kwitter. The multiple-shell structure of the planetary nebula ngc 6751. *The Astrophysical Journal*, 376:150–160, 1991.

[196] Franco Sabbadin. Spatiokinematical models of five planetary nebulae. *Monthly Notices of the Royal Astronomical Society*, 210(2):341–358, 1984.

[197] DM Clark, Ma T García-Díaz, JA López, WG Steffen, and MG Richer. Shaping the glowing eye planetary nebula, ngc 6751. *The Astrophysical Journal*, 722(2): 1260, 2010.

[198] Lawrence Hugh Aller. Central stars of planetary nebulae. *Memoires of the Societe Royale des Sciences de Liege*, 9:271–299, 1975.

[199] L Koesterke and W-R Hamann. Spectral analyses of central stars of planetary nebulae of early wc-type ngc 6751 and sanduleak 3. *Astronomy and Astrophysics*, 320:91–100, 1997.

[200] WA Joye. New features of saoimage ds9. In *Astronomical Data Analysis Software and Systems XV*, volume 351, page 574, 2006.

[201] A. Blecha, G. Schaller, and A. Maeder. Fast pulsations in a Wolf–Rayet star. *Nature*, 360(6402):320–321, November 1992. doi: 10.1038/360320a0. URL `http://www.nature.com/nature/journal/v360/n6402/abs/360320a0.html`.

[202] André-Nicolas Chené and Anthony F. J. Moffat. Pulsations in Wolf-Rayet stars: observations with MOST. *Proceedings of the International Astronomical Union*, 6(S272):445–450, July 2010. ISSN 1743-9213, 1743-9221. doi: 10.1017/S1743921311011082. URL http://www.journals.cambridge.org/abstract_S1743921311011082.

[203] Andrew J Skemer, Deno Stelter, Dimitri Mawet, Michael Fitzgerald, Benjamin Mazin, Olivier Guyon, Christian Marois, Zackery Briesemeister, Timothy Brandt, Jeffrey Chilcote, et al. The planetary systems imager: 2-5 micron channel. In *Ground-based and Airborne Instrumentation for Astronomy VII*, volume 10702, page 10702A5. International Society for Optics and Photonics, 2018.

[204] Benjamin A Mazin, George D Becker, Gustavo Cancelo, Kevin France, Wesley C Fraser, Tucker Jones, Seth R Meeker, Kieran O'Brien, J Xavier Prochaska, Shriharsh Tendulkar, et al. Krakens: a superconducting mkid integral field spectrograph concept for the keck i telescope. In *Ground-based and Airborne Instrumentation for Astronomy VII*, volume 10702, page 107020H. International Society for Optics and Photonics, 2018.

Appendix A

MEDIS Parameters

Table A.1: The turbulence profile parameters used in CAOS. The turbulence is assumed to be localized to several layers and the turbulence within those layers is frozen in time. h is the height above the telescope aperture, C_n^2 ratio sets the relative strength of each of the layers. The wind directions are restricted to be along the x and y axes.

Layer	h (m)	C_n^2 ratio	v (m/s)	Direction (°)
0	0	0.20	5	0
1	10	0.20	5	90
2	100	0.20	5	180
3	500	0.10	5	0
4	700	0.05	5	270
5	1000	0.05	10	90
6	5000	0.05	10	0
7	7000	0.05	10	180
8	10000	0.05	15	0
9	15000	0.05	15	270

Table A.2: Astrophysics configuration parameters. a: the location of the companion object in units of λ/D. b: the total number of simulated steps. c: contrast is currently assumed to decrease linearly towards longer wavelengths. This parameter dictates the gradient of this transition across the band. d: the calibration applied to each intensity map. e: dictates which phase maps are applied for the first simulation (useful for reference differential imaging)

Parameter	Options	Example
companion	Bool	True
contrast	[Float, ...]	[5×10^{-4}]
lods[a]	[Float, Float], ...]	[[-1.0, 1.0]]
numframes[b]	Float	1000
C_spec[c]	[Float, ...]	1.5
star_photons[d]	Float	1×10^{6}
startframe[e]	Float	0

Table A.3: Telescope configuration parameters. a: The characteristic lifetime of the quasi-static aberrations. b: the number of actuators across one dimension. c: the beam ratio at the lowest wavelength. d: the focal ratio. e: the number of pixels across one axis. f: secondary DM. g: rate of azimuthal rotation. h: the sampling of the atmosphere maps generated in Atmosphere. i: servo lag and minimum sampling time.

Parameter	Options	Example
CPA_type	None/'Static'/'Quasi'	'Static'
NCPA_type	None/'Static'/'Quasi'	'Quasi'
aber_time[a]	[Float, ...]	[100s]
active_null	Bool	False
ao_act[b]	Int	64
band	[Float, Float]	[800 nm, 1400 nm]
beam_ratio[c]	Float	0.3
detector	'Ideal'/'H2RG'/'MKIDs'	'MKIDs'
diam	Float	8 m
exposure_time	Float	0.001
f_lens[d]	Float	322
grid_size[e]	Int	128
n_surfs	Float	2
'OOPP'	[Float, ...]	[8,4]
null_ao_act[f]	int	32
nwsamp	Float	8
occult_loc	[Int,Int]	[0, 0]
occulter_type	None/'Vortex'/'Solid'/'Gauss'/'8th Order'	'8th Order'
piston_error	Bool	True
platescale	Float	13.61 mas/px
rot_rate[g]	Float	0
samp[h]	Float	0.125 m/px
satelite_speck	Bool	True
servo_error[i]	[Float, Float]	[10 ms, 10 ms]
speck_locs	[[Float, Float], ...]	[[50, 60]]
speck_peakIs	[Float, ...]	[0.05]
speck_phases	[Float, ...]	$[\pi/2]$
use_ao	Bool	True
use_apod	Bool	False
use_atmos	Bool	True
use_hex	Bool	False
use_spiders	Bool	True
use_zern_ab	Bool	False
wfs_measurement_error	Bool	False

Table A.4: MKIDs configuration parameters. a: the center of the Gaussian distribution with which to draw pixel spectral resolutions from. b: array size to project data from Telescope onto. c: option to mask 1- pix_yield pixels. d: the center of the Gaussian with which to draw pixel background phases from. e: apply a dead time after each count − similar to max_count. f: center of Gaussian for pixel responsivities. g: sample photons from a continuous datacube. h: number of elements that make up the pixel sample distributions. i: Threshold SNR for detection. j: linear relation for phase assumed – here 800 nm corresponds to -90° and 1500 nm corresponds to -30°

Parameter	Options	Example
R_mean[a]	Float	10
R_sig	Float	2
array_size[b]	[Int, Int]	[80, 125]
bad_pix[c]	Bool	False
bg_mean[d]	Float	-10°
bg_sig	Float	15°
dead_time[e]	Float	4×10^{-4}
r_mean[f]	Float	0.95
r_sig	Float	0.025
hot_pix	Float	0.01
interp_sample[g]	Bool	True
max_count	Float	2500 cts/s
phase_background	Bool	False
phase_uncertainty	Bool	False
pix_yield	Float	0.9
remove_close	Bool	False
res_elements[h]	Int	128
threshold_phase[i]	Float	3
wavecal_coeffs[j]	[Float, Float]	[1/12, -157]