USING MODERN MATHEMATICAL AND COMPUTATIONAL TOOLS FOR IMAGE PROCESSING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jacob Daniel Shapiro

December 2020

USING MODERN MATHEMATICAL AND COMPUTATIONAL TOOLS FOR IMAGE PROCESSING

Jacob Daniel Shapiro, Ph.D.

Cornell University 2020

In the search for exoplanets, many methods have proven fruitful. All involve careful observation of the host star and complex post-processing algorithms to identify if any planets are part of the system. One methodology that has shown promise, yet currently yields relatively few results, is direct imaging.

The first part of this dissertation showcases the development of a novel technique to identify planets in post-processing of direct imaging data. It leverages the common spatial pattern filtering algorithm in combination with a forward modeled matched filter. I compare the algorithm to other leading techniques.

The second part develops the tools and software for generalizing this approach to many different datasets. This allows for systematic, large-scale statistical analyses of the CSP method applied to a variety of stars and injected data. I present results for multiple sets of observations and show how the new technique can be expected to perform.

Finally other avenues of image processing are explored both for use in a new type of filter and for the development of advanced self-assembling telescopes in space.

BIOGRAPHICAL SKETCH

Jacob Shapiro received a Bachelor of Science in Aerospace Engineering from the University of Florida in May, 2015. He complemented this degree with minors in Business Administration and Astronomy. He has worked as an intern at Johnson and Johnson: Vision Care (2014), as well as at SpaceX (2015), where he developed parts for the M1D engine. He began his PhD program at Cornell in Fall of 2015, where he joined the Space Imaging and Optical Systems Lab.

To the Farm House and everyone under its roof.

ACKNOWLEDGEMENTS

While I may be the author of this dissertation, this document and the work it represents couldn't exist without so many people and their help along the way.

Dmitry Savransky has been an unyielding force of support during my time at Cornell. Whether I needed technical expertise, emotional support, or even a bit of push to keep me going, I could always count on him to be my ally. Dmitry is one of the best professors I have worked with, and it has been an honor to be his pupil.

Much of this work is based on published works and group efforts. I am grateful for the help of my coauthors J.B. Ruffio, and Bruce Macintosh, and to the the GPIES team. I am also happy to have been a part of the Space Imaging and Optical Systems Lab here at Cornell. My labmates, Daniel Garrett, Joyce Fang, Gabriel Soto, Dean Kiethly, Corey Spohn, Duan Li, Katie Summey, and Zvonimir Stojanovski have been great resources to me, and have made the work environment a place filled with friends.

My friends at home and in Ithaca have kept me as close to sane as possible the past five years. In particular, thank you to Gabriel Soto, Kyle Wellmerling, Matthew Walsh, Julian Morelli, and Joseph Corbett-Davies. I hope our companionship will remain for many years to come.

Of everyone, my parents and sister have been by my side the longest, and deserve so many thanks for putting me on the path that led me here. They surrounded me with encouragement, instilled in me a passion for learning, and gave me so much unconditional love.

To everyone else who has helped along the way, I couldn't have done it without you. Thank you.

This work is based on observations obtained at the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy under a cooperative agreement with NSF on behalf of the Gemini partnership, whose membership includes: NSF (United States); the National Research Council (Canada); the Comisión Nacional de Investigación Científica y Tecnológica (Chile); the Australian Research Council (Australia); the Ministério da Ciência, Tecnologia e Inovação (Brazil); and Ministerio de Ciencia, Tecnología e Innovación Productiva(Argentina). GPI data are archived at the Gemini Observatory Archive: https://archive.gemini.edu/searchform.

Additionally, this work is supported by NASA Space Technology Grant Numbers NNX16AI13G and 80NSSC20K0068, as well as NASA Innovative Advanced Concepts Grant 80NSSC18K0869.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Dissertation Overview

While this work exists to support several disparate projects, all of it is united by a common theme: using modern mathematical and computational tools to assist image processing for space exploration. The overarching problem that links each project is the search for the faintest, most elusive signals in imaging. To take pictures of planets orbiting other stars, we need the best technology at every level of the endeavor: designing instrumentation, collecting data, and processing the results. At each stage, we need to leverage cutting edge methodology. This is precisely what has been done for the work covered in this dissertation.

The remainder of **Chapter 1** discusses some of the necessary background information for the work in the succeeding chapters. It establishes important techniques, definitions, and mathematical concepts.

Chapter 2 develops a new algorithm for post-processing of direct imaging data, with the express goal of extracting planet signals. I present the application of Common Spatial Pattern filtering to the detection of point sources in high-contrast astronomical images. The data is pre-processed in two different ways: one copy of the data set keeps the point source in the same spatial location over many images through time, while a second copy of the dataset moves the point source azimuthally through the image as in Angular Differential Imaging. The differences between these two datasets are exploited via Common Spatial Pat-

tern filtering, highlighting the point source. I develop a forward model for this new process, which is then used to predict the result of applying this method to data with a point source at a particular spatial location. I validate the forward model numerically and use the prediction as template in a matched filter with the actual CSP modes. I present results for multiple sets of observational data and show that the new CSP Forward Model Matched Filter performs well in terms of Signal-to-Noise ratio and planet signal recovery.

Given that common spatial pattern forward model matched filters have been shown to be an effective method for detection of point sources in specific examples of high-contrast astronomical images, **Chapter 3** develops tools and software for generalizing this approach to many different datasets. Furthermore, the software is compatible with the same input/output architecture as pyKLIP, a software package for analyzing exoplanet data in python with Principal Component Analysis. This allows for systematic, large-scale statistical analyses of the CSP method applied to a variety of stars and injected data. I present results for multiple sets of observational data and show how CSP FMMF performs in terms of SNR.

Chapter 4 examines the optical design of a large space observatory. Space observatories have many advantages over ground-based telescopes. However, constructing and launching large space telescopes remains a significant challenge. A solution to this problem lies in autonomous, in-space assembly. To benefit from efficiencies of scale and mass production, a modular telescope assembled in space can be constructed from identical mirror segments. These identical segments must then be deformed to an appropriate shape in space. This work examines the optical feasibility of such a project, using a 31-meter

Ritchey-Chrétien telescope composed of about 1,000 1-m mirrors as a case study. In particular, this chapter examines the shape of the telescope optics through Zernike decomposition and computes the physical optics propagation of such a system to analyze the resultant PSF with simulation in Zemax OpticStudio.

Chapter 5 examines a novel approach to using image processing as part of the measurement for a non-Gaussian filter. I present a portion of this larger project to use imaging as an update for state-variables. I develop the necessary models to express the camera state and show the effects of deformation, as well as the propagation of a SIFT algorithm as part of the measurement function. I also show that the resultant data from this model is non-Gaussian and requires a unique filter.

In **Chapter 6**, I summarize the entire dissertation and acknowledge the future work necessary to each of these projects.

The **Appendix** contains Table 1, which summarizes the common notation and symbols in **Chapter 1**, **Chapter 2**, and **Chapter 3**.

1.2 Direct Imaging

There are a variety of techniques used to find exoplanets. To date, the planets discovered are shown in Figure 1.1.

As pictured, direct imaging occupies a unique portion of the phase space: large, bright planets far from their host star. Capturing more of this data by discovering more planets would give humanity a better picture of our universal environment, and perhaps more insight into the formation of our home solar

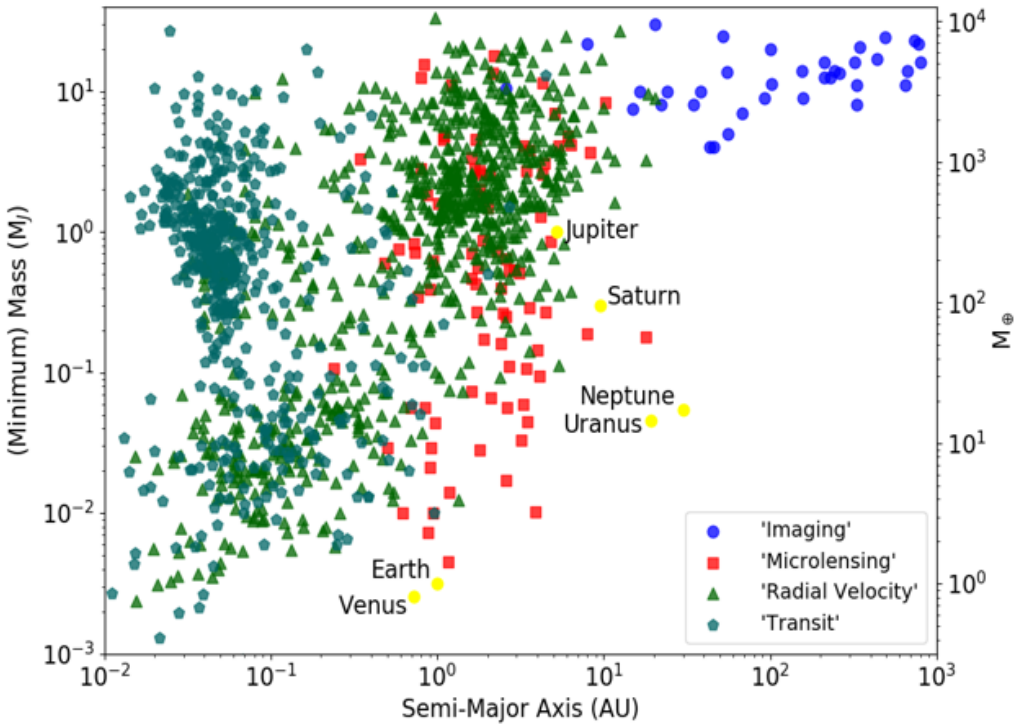

Figure 1.1: Plot of all currently known exoplanets[1]. Direct Imaging is shown in blue circles, and stands unique from other members of the population.

system and planets. It places our knowledge into context.

Direct imaging of exoplanets requires a combination of high-contrast instrumentation, adaptive optics (AO), and advanced data processing techniques. Despite continued advances in all of these areas, extracting the planet signal from a set of images remains incredibly challenging. A multitude of noise sources can be found within high contrast astronomical images, including bright, spatially correlated noise called speckles [22, 18]. Speckles are a result of internal refraction errors within the telescope, and (for ground-based imaging) residuals from AO corrections of the atmosphere. They are nearly unavoidable and are particularly troublesome because they can strongly resemble the planet signals of interest. Even after post-processing, target signal fluxes are often of the same order of magnitude as the residual noise, making unbiased planet signal

extraction very challenging.

Fortunately, speckles and planet signals behave differently across the spectral and spatial domains. This key difference has given rise to the method of speckle subtraction. Speckle subtraction from one observation primarily relies on two methods of observing: Angular Differential Imaging (ADI) [14] and Spectral Differential Imaging (SDI) [22]. ADI is accomplished by allowing the telescope to rotate relative to the sky by switching off the derotator on ground-based telescopes or by rolling space observatories about the line of sight. This ensures that the desired target signal moves azimuthally from image to image over time. Meanwhile, sources of noise from within the telescope optics, namely speckle noise, remain fixed in the telescope frame. This only allows for temporal evolution of speckles rather than changes caused by the rotation of the telescope.

SDI exploits the fact that speckles are correlated across a spectral band such that the speckle pattern scales linearly with wavelength. This is in direct contrast to the planet signal, which exhibits no spatial scaling with wavelength, although its flux may vary across the band. By employing an Integral Field Spectrograph (IFS), we can simultaneously obtain images at a variety of wavelengths and use this information to our advantage.

Both of these observing strategies (ADI and SDI) that induce spatial diversity between signal and noise allow for development of a speckle noise model that ideally does not include the planet signal. Upon subtracting this model of the noise from the science images, what remains is the planet signal. A common method for building the model of the speckle noise is based on Principal Component Analysis (PCA) [23].

1.3 Principal Component Analysis and Common Spatial Pattern Filtering

Principal Component Analysis is the foundation of techniques like Karhunen-Loève Image Processing (KLIP) [32]. The Karhunen-Loève decomposition is the most information-compact representation of a dataset; it reproduces the data in a new, orthogonal basis in which the modes that contain the most information possible are organized in descending order.

Common Spatial Pattern Filtering (CSP) [3] and Principal Component Analysis are both examples from a class of algorithms known as Blind Source Separation (BSS). Ultimately, both PCA and CSP are both linear, least-squares algorithms.

Principal Component Analysis is a technique generally used for dimensionality reduction. It finds the dimensions of highest variance in a data set and uses those as a basis set to reproduce the data. In the context of direct imaging, these highest-variance dimensions correspond to the most informational-compact modes. The highest information mode is the most similar mode to the data set. When computed from a set of images that does not include the target image, the weighted combination of every PCA mode is the least-squares approximation of the target image based on a linear combination of the images in the reference set. The principal components $\mathbf{w}$ of a dataset $\mathbf{X}$ are defined by solving

$$\arg\max_{\mathbf{w}} \left\|\mathbf{X}^T\mathbf{w}\right\|^2 \quad \text{s.t.} \quad \|\mathbf{w}\| = 1\,. \tag{1.1}$$

The maximizing values of $\mathbf{w}$ are equal to the eigenvectors of the covariance of

X. Because the eigenvalues are associated with the variance of the dataset along their associated eigenvector **w**, the eigenvectors associated with largest eigenvalues are the msot prominent modes from the dataset. These more prominent modes contain more information about the dataset.

Common Spatial Pattern filtering is defined by a maximization that is very similar to PCA's Equation (1.1), except that it is formulated as a comparison between two different sets of data

$$\arg\max_{\mathbf{w}} \frac{\left\|\mathbf{X}_1^T\mathbf{w}\right\|^2}{\left\|\mathbf{X}_2^T\mathbf{w}\right\|^2} \quad \text{s.t.} \quad \|\mathbf{w}\| = 1 \,. \tag{1.2}$$

This is remarkably similar to PCA; if $\mathbf{X}_2$ is the identity matrix (or a multiple thereof), then the expression reduces to PCA exactly. However, the two methods are conceptually different. In PCA, the **w** vectors are the directions of maximum variance in one set of data. Conversely, in CSP, the **w** vectors maximize the difference between two datasets $\mathbf{X}_1$ and $\mathbf{X}_2$ [7]. To find the **w** values that satisfy Equation (1.2), we must rewrite it in a form that is more easily evaluated. We first do this by finding the covariance matrices for the two datasets as in Equation (2.1). We sum the two covariance matrices to form a composite covariance matrix, which we then decompose into eigenvalues and eigenvectors

$$\mathbf{C}_+ = \mathbf{C}_1 + \mathbf{C}_2 = \mathbf{U}\mathbf{\Gamma}\mathbf{U}^T \,. \tag{1.3}$$

Here, **U** is a unitary matrix of the eigenvectors of $\mathbf{C}_+$, and $\mathbf{\Gamma}$ is the diagonal matrix of eigenvalues. The decomposition can be used to construct a whitening matrix, **P**

$$\mathbf{P} = \mathbf{\Gamma}^{-1/2}\mathbf{U}^T \,. \tag{1.4}$$

Throughout this paper, the square root operator as applied to matrices will only be used for purely diagonal matrices. This implies that the root of the matrix

is also a diagonal matrix of the square roots of the diagonal elements of the original matrix. This operator is not defined (in this work) for non-diagonal matrices. If the whitening matrix is applied to each covariance matrix,

$$\bar{\mathbf{C}}_n = \mathbf{P}\mathbf{C}_n\mathbf{P}^T , \tag{1.5}$$

where n represents either dataset 1 or 2. Then, the sum of the whitened, normalized covariance matrices must be the identity matrix

$$\bar{\mathbf{C}}_1 + \bar{\mathbf{C}}_2 = \mathbf{I} . \tag{1.6}$$

The whitened data can be used to reformulate Equation (1.2). Given that the sum whitened covariance matrices must be the identity matrix, as $\left\|\bar{\mathbf{X}}_1^T\mathbf{w}\right\|^2$ increases, $\left\|\bar{\mathbf{X}}_2^T\mathbf{w}\right\|^2$ must simultaneously decrease. Thus, Equation (1.2) is equivalent to

$$\arg\max_{\mathbf{w}} \left\|\bar{\mathbf{X}}_1^T\mathbf{w}\right\|^2 \quad \text{s.t.} \quad \mathbf{w}^T(\bar{\mathbf{C}}_1 + \bar{\mathbf{C}}_2)\mathbf{w} = \mathbf{I} \tag{1.7}$$

Using Lagrange multipliers to solve Equation (1.7), and noting that $\bar{\mathbf{C}}_1^T = \bar{\mathbf{C}}_1$, the optimization reduces to an eigenvalue problem

$$L(\mathbf{w}, \phi) = \mathbf{w}^T\bar{\mathbf{X}}_1\bar{\mathbf{X}}_1^T\mathbf{w} - \phi[\mathbf{w}^T\left(\bar{\mathbf{C}}_1 + \bar{\mathbf{C}}_2\right)\mathbf{w} - \mathbf{I}] \tag{1.8}$$

$$\frac{\delta L}{\delta \mathbf{w}} = \bar{\mathbf{C}}_1\mathbf{w} + \bar{\mathbf{C}}_1^T\mathbf{w} - 2\phi\mathbf{w} = 0 \tag{1.9}$$

$$\bar{\mathbf{C}}_1\mathbf{w} = \phi\mathbf{w} . \tag{1.10}$$

Thus, the eigenvectors of the whitened covariance matrix from the first dataset are the directions that maximize the difference between $\mathbf{X}_1$ and $\mathbf{X}_2$.

CHAPTER 2

THE ORIGINAL COMMON SPATIAL PATTERN FILTERING FORWARD MODEL MATCHED FILTER ALGORITHM FOR DIRECT IMAGING DETECTION

Common spatial pattern processing represents a new class of direct imaging algorithms. It offers a unique approach to planet signal extraction: rather than subtracting away the speckle noise from images, the planetary signal itself can be directly modelled.

In this chapter, I present one particular methodology to use CSP as a direct imaging reduction tool and specific examples of applications of this technique. The real data and injected signal used as test cases is described in §2.1, as well as the details concerning images as data structures. Our specific algorithm that uses CSP for direct imaging signal extraction is elaborated on in §2.2. The effects of the CSP process on a point source is examined in §2.3, which develops the forward model for CSP and discusses its accuracy and implications for tuning parameters. I discuss using the Forward Model Matched Filter (FMMF) in §2.4. Finally, I compare the CSP FMMF results with that obtained from a current KLIP reduction in §2.5. Throughout this chapter, bold symbols refer to matrices and capital boldface will distinguish two-dimensional matrices from one-dimensional vectors, which will be shown in lower case font.

This chapter is primarily based on work published in [30].

2.1 Data Sources and Structure

2.1.1 Images as Data

An observing sequence is assumed to be a set of N images corresponding to individual observations of the same target star over a short span of time. Each image is a spatio-spectral data cube like those generated by the Gemini Planet Imager data reduction pipeline [19]. Each data cube consists of a set of 2D images in different wavelengths, all captured at the same time. Moving from slice to slice is equivalent to looking at the same spatial data in a different wavelength. Each image (datacube slice) from the full set of datacubes encodes a unique combination of time and wavelength and therefore a different realization of background noise and different spatial location of the target planet signal, if one is present.

We can write each of the N images, each containing P pixels, as a column vector. The specific vectorization does not matter as long as the scheme to vectorize the images is applied consistently to each image. Ensuring the vectorization technique is consistent from image to image guarantees that data from the same spatial locations of the image are represented in the same location in each corresponding vector. For this analysis, each column of the image is appended to the bottom of the previous column, from left to right within an image. A simple example is shown in Figure 2.1.

Any given vector-mean subtracted image is denoted $\mathbf{i}$, of dimension $P \times 1$. A collection of images can be represented as an $N \times P$ dataset $\mathbf{X}$, where $\mathbf{X} \triangleq$

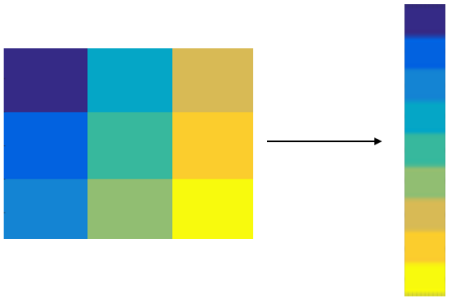

Figure 2.1: An example of vectorizing a 2D image into a column vector. In this example, columns are appended vertically from left-to-right.

$[\mathbf{i}_1 \quad \mathbf{i}_2 \quad \mathbf{i}_3 \quad ... \quad \mathbf{i}_N]^T$. Thus, the image-to-image sample covariance of the dataset is given by

$$\mathbf{C} = \frac{\mathbf{X}\mathbf{X}^T}{P-1}. \tag{2.1}$$

2.1.2 Data Source and Synthetic Planet Signal Injection

To gather and test real data, observing sequences were selected from the GPI Exoplanet Survey (GPIES)[11]. Specifically, I selected the November 2015 observation of β Pic b [35] and the August 2017 non-detection observation of HD 14706. This allows for one set of real science data with a known exoplanet signal and another with no known exoplanet as contrasting examples. Both data sets were truncated to the central 75×75 pixels to expedite computations. Additionally, the known signal falls within this image region.

In addition to the original data, we can also inject a known planet signal into the images to be used for analysis and comparison. To recreate a valid planetary Point Spread Function (PSF), we must examine both wavelength dependence and spatial structure. In our injections, planet spectra are selected from atmospheric models developed by [27]. In particular, we have used an L-Type

model.

To develop an appropriate spatial model of the planet PSF, we turn to information already contained in the science images. The GPIES data has four scaled representations of the unblocked stellar PSF in that particular image, called satellite spots [31, 15]. These four satellite spots are median combined to act as a model for the injected planetary PSF in each image. This spatial model is scaled by the spectra at each wavelength. It can be seen in Figure 2.2 that the spatial model deviates significantly from an idealized 2D Gaussian.

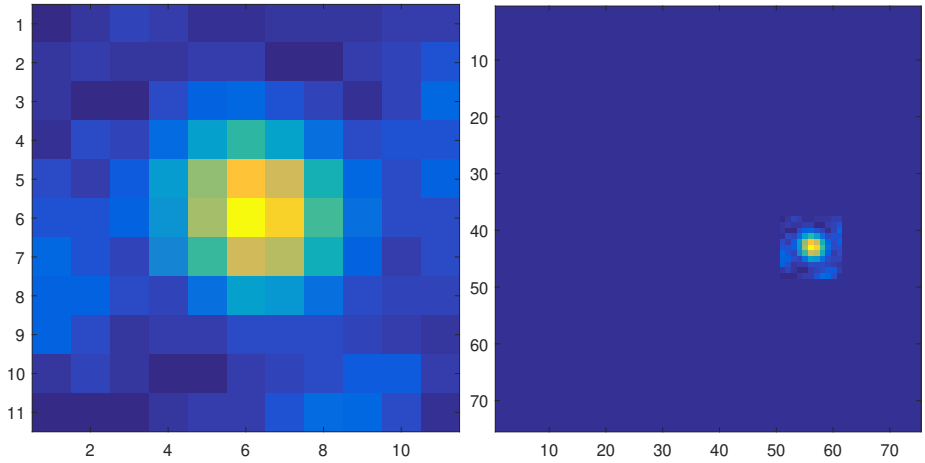

Figure 2.2: Detail and full view of injected planet data at one slice. The planet signal used is based on the satellite spots of the original data, and is non-Gaussian.

The injected planets cannot be seen in the raw data of either β Pic data or the HD 14706 data across multiple different image scaling laws, as shown in Figure 2.3.

The planet injected into the β Pic dataset has a relative flux of 1.7×10^{-4} relative to the total background flux of the entire 75×75 stamp. This value is a cumulative sum from all slices of every datacube. The actual planet, β Pic b, can be very faintly seen on the left of the squared scaling law. The fact that this can

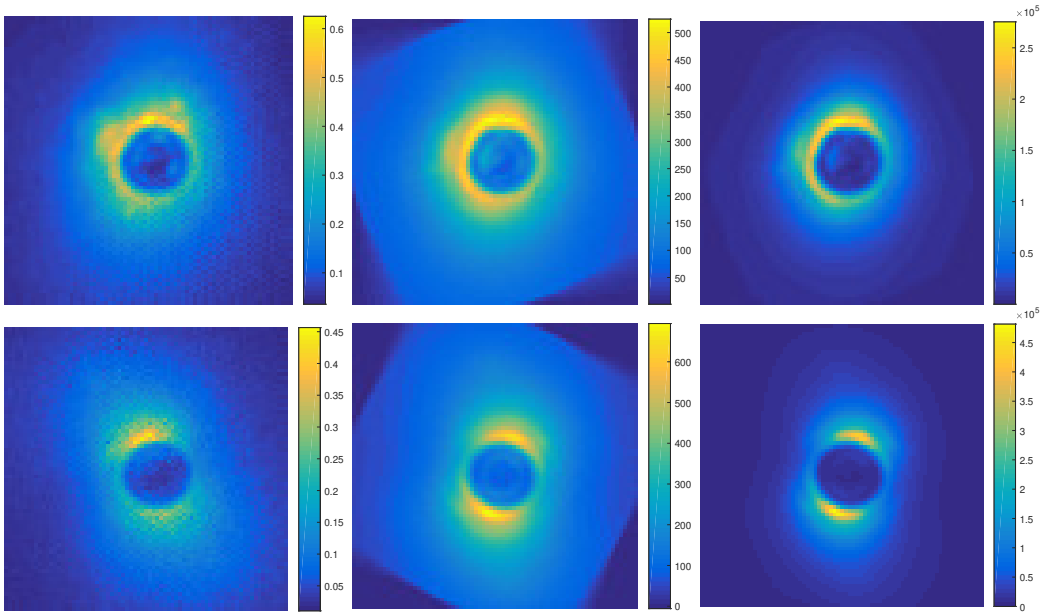

Figure 2.3: Top row: β Pic. Bottom row: HD14706. Left column: single image example. Middle column: linear scaling law. Right column: squared scaling law. The planet signal has been injected into both datasets, and is not visible to the naked eye regardless of scaling law.

be seen, but the injected planet cannot shows that the injected planet is injected at a lower raw SNR than β Pic b. Similarly, the relative flux of the planet in the HD 14706 data is 1.8×10^{-4}.

2.2 CSP Methodology

In this section, I discuss our particular methodology for using CSP to analyze the direct imaging data. As KLIP is a particular implementation of PCA on imaging data, this method too is simply one possible use of CSP to detect exoplanets. I believe that other methodologies and other selections of $\mathbf{X}_1$ and $\mathbf{X}_2$ could be useful approaches to identifying exoplanets.

In the case of direct imaging, we begin by subdividing the data by wave-

length. At each individual wavelength we have a sequence of images at differing parallactic angles. We conduct CSP independently on the set of images corresponding to each wavelength. While the CSP process can operate using all the images in a dataset (number of wavelength $\times$ number of timesteps), the error in the forward model used has quadratic growth with the number of images, as shown in §2.3.3. Since a forward model matched filter will yield the best results, as it does with KLIP [24], we must modify our dataset used for CSP to reflect the Forward Model results. A natural way to limit the number of images used in the algorithm is to subdivide by wavelength, evaluate the process for each subdivided set, and sum each of the separate results together. Given the simplified problem (only accounting for changes in wavelength), we can create the two CSP datasets, $\mathbf{X}_1$ and $\mathbf{X}_2$ as follows.

The $\mathbf{X}_1$ Dataset: In this dataset, the planet signal, if one exists, is distributed spatially from image to image. This relies primarily on ADI. Due to Angular Differential Imaging, at each time step, the planet signal will be slightly displaced azimuthally about the line of sight. For one given wavelength, we select all the images corresponding to different times in the dataset.

The $\mathbf{X}_2$ Dataset: This dataset uses the same images as those in $\mathbf{X}_1$. However, here we wish to create a substantial difference from $\mathbf{X}_1$ that somehow relates to the planet signal. This difference ensures that the information of the planet signal is a component within the projection matrix formed by the vectors $\mathbf{w}$. To accomplish this, we design the preprocessing of the dataset opposite that of $\mathbf{X}_1$: the planet signal is in the same spatial location in every image. We must derotate each image by the parallactic angle of the time to the north-up position.

If we are to conduct CSP with all of the collected images at once, the two

datasets can be modified to include multiple wavelengths as well. For every wavelength image at a single timestep, the planet signal is in the same spatial location. For $\mathbf{X}_1$, we would use the wavelength information to distribute the signal radially along the image. If we choose a reference image wavelength, λ_0, typically in the middle of the wavelength range, we can scale the size of all of the images to this value by λ_0/λ_i. At wavelengths larger than the reference, the planet signal moves closer to the parent star, and the at smaller wavelengths, it will be moved out closer to the edge of the image. Adding wavelength images to the $\mathbf{X}_2$ datatset requires no modifications, as the planet signal is already spatially co-located.

Once the two datasets are created, we step through the CSP process as outlined in §1.3. After finding the $\mathbf{w}$ vectors that solve the eigenvalue problem posed in Equation (1.10), we can use these to build a new projection matrix

$$\mathbf{Z} = \mathbf{\Phi}^{-1/2}\mathbf{W}^T\mathbf{P}\mathbf{X}_1 \,. \tag{2.2}$$

Here, $\mathbf{\Phi}$ represents the diagonal matrix of eigenvalues of $\bar{\mathbf{C}}_1$ and $\mathbf{W}$ is the matrix of its eigenvectors. $\mathbf{P}$ is the whitening matrix as defined in Equation (1.4). This makes $\mathbf{Z}$ an $N \times P$ matrix where the rows represent the modes of $\bar{\mathbf{X}}_1$ in decreasing order of information content.

To find the signal of an injected planet, we compare the actual modes of $\mathbf{Z}$ with modes predicted by a forward model that includes a planet signal. If the predicted and actual modes do not align (the matched filter doesn't yield a strong result), then there is not likely to be an astrophysical signal at that particular location. Importantly, not all modes of $\mathbf{Z}$ are used; we limit the modes used for the matched filter to an optimal set of K modes, associated with both high- and low-valued eigenvalues. It is in these modes that the planet signal

itself is prominent, as shown in §2.3.4. bf This methodology is different from KLIP. In KLIP, the first K modes are used for subtraction of the noise from the image. In this CSP technique, we used primarily the end modes directly as well as other modes distributed throughout. These modes most highly contain the planet signal itself.

2.3 Forward Modeling

2.3.1 Overview

Recall that PCA is the maximally information-compact dimensionality reduction. If an observing dataset is treated with this technique and a target image is projected onto these modes, the first eigenvectors contains most of the structure in the image. In direct imaging, these modes can include a wind pattern from ground-based telescopes and many of the static components of speckles. Unfortunately, the speckles are only quasi-static throughout an observation. However, the planet signal has been designed to move from image to image. Speckle information is therefore more likely to be contained in the higher-information modes than the planet signal despite some slight time variation. By subtracting away these modes of information from different target images, what remains is the planet signal, if it exists. Nonetheless, the planet signal can be significantly altered by this process. It has been shown that the planet signal undergoes over-subtraction and self-subtraction [21]. Over-subtraction is the result of fitting the astrophysical signal with the speckle data, causing subtraction and distortion. Self-subtraction is the phenomenon where the signal subtracts itself as it moves

due to ADI and SDI.

These effects on the target planet signal due to KLIP warranted further study. These effects were examined rigorously by [21] by propagating a point source model through the complete KLIP process. The forward model provides an accurate prediction of the result of KLIP on a planet source, should it exist. In turn, this prediction can be used as a template in a matched filter to scan through completed KLIP reductions and find matches to the predicted planetary signals. The forward model therefore further helps distinguish the planet signal from speckle residuals [24]. Much of the CSP forward modeling in §2.3.2 is based on the work done on the KLIP Forward Model in [21].

2.3.2 Developing the Model

In developing the forward model, we assume that it can be conducted across both ADI and SDI simultaneously. In practice, however, we limit the analysis to ADI. This reflects the decision to limit the number of images per forward model calculation made in §2.2. Nonetheless, the forward model development presented here is representative of a general case.

Following the formalism developed in [21], we let each image value at a time t, wavelength λ, and spatial location x be represented as

$$\mathbf{i}_{\lambda,t}(x) = \mathbf{s}_{\lambda,t}\left(\frac{x}{\lambda}\right) + \epsilon R_{\theta_t}\left[\mathbf{a}_{\lambda}(x)\right] . \tag{2.3}$$

where $\mathbf{s}$ is the speckle noise and $\mathbf{a}$ is the planet signal. To align these two components at the same spatial location, the speckle noise is scaled by wavelength while the planet signal must be rotated by a rotation matrix R, which is a function of the parallactic angle θ at time t. The ϵ multiplier represents the contrast

ratio of the planet relative to the hypothetical off-axis star. As the planet is typically quite faint compared to the background in any given image, ϵ is treated as a small quantity throughout. These definitions are consistent with changes to an image induced both by ADI and SDI.

The sets of images can then be transformed into two datasets for CSP, as described in §2.2. After applying initial CSP transformations to the images, the equations representing the two datasets are

$$\mathbf{i}_{1_{\lambda,t}}(x) = \mathbf{i}_{\lambda,t}\left(x\frac{\lambda}{\lambda_0}\right) = \mathbf{s}_{\lambda,t}\left(\frac{x}{\lambda_0}\right) + \epsilon R_{\theta_t}\left[\mathbf{a}_\lambda\left(x\frac{\lambda}{\lambda_0}\right)\right] \tag{2.4}$$

$$\mathbf{i}_{2_{\lambda,t}}(x) = R_{-\theta_t}\left[\mathbf{i}_{\lambda,t}(x)\right] = R_{-\theta_t}\left[\mathbf{s}_{\lambda,t}\left(\frac{x}{\lambda}\right)\right] + \epsilon\mathbf{a}_\lambda(x)\,. \tag{2.5}$$

For notation purposes, we simplify many of these transformations to be represented with subscripts: we let $\mathbf{s}_n$ be speckle data from an image transformed to match a particular dataset, and $\mathbf{a}_n$ correspond to the transformed planet signal data. The images belonging to each dataset can then be written as

$$\begin{aligned}\mathbf{i}_{1_{\lambda,t}} &= \mathbf{s}_1 + \epsilon\mathbf{a}_1\\ \mathbf{i}_{2_{\lambda,t}} &= \mathbf{s}_2 + \epsilon\mathbf{a}_2\,.\end{aligned} \tag{2.6}$$

Combining each of the individual vectors into matrices, and taking their transposes gives the expression for both transformed datasets

$$\begin{aligned}\mathbf{X}_1(x) &= \mathbf{S}_1 + \epsilon\mathbf{A}_1\\ \mathbf{X}_2(x) &= \mathbf{S}_2 + \epsilon\mathbf{A}_2\,,\end{aligned} \tag{2.7}$$

where the capital letters correspond to matrices of which the rows are vectorized images (the lowercase counterparts). This is consistent with the previous definition of a dataset $\mathbf{X}$ as an $N \times P$ matrix consisting of rows of images. The covariances of these matrices are thus

$$\mathbf{C}_1 = \frac{1}{P-1}\left(\mathbf{S}_1\mathbf{S}_1^T + \epsilon(\mathbf{A}_1\mathbf{S}_1^T) + \epsilon(\mathbf{S}_1\mathbf{A}_1^T) + \epsilon^2(\mathbf{A}_1\mathbf{A}_1^T)\right) = \mathbf{C}_{S_1S_1} + \epsilon\mathbf{C}_{AS_1} + O(\epsilon^2) \tag{2.8}$$

$$\mathbf{C}_2 = \frac{1}{P-1}\left(\mathbf{S}_2\mathbf{S}_2^T + \epsilon(\mathbf{A}_2\mathbf{S}_2^T) + \epsilon(\mathbf{S}_2\mathbf{A}_2^T) + \epsilon^2(\mathbf{A}_2\mathbf{A}_2^T)\right) = \mathbf{C}_{S_2S_2} + \epsilon\mathbf{C}_{AS_2} + O(\epsilon^2)\,. \quad (2.9)$$

At this point, assuming that the flux of the target signal is very small relative to the rest of the flux in the image, we can disregard terms of order ϵ^2. If ϵ were to be large enough that this assumption becomes unreasonable, then significant post-processing would no longer be necessary—the signal would be visible in the raw data. By defining

$$\begin{aligned} \mathbf{C}_{S_nS_n} &\triangleq \frac{1}{P-1}\left(\mathbf{S}_n\mathbf{S}_n^T\right) \\ \mathbf{C}_{AS_n} &\triangleq \frac{1}{P-1}\left(\mathbf{A}_n\mathbf{S}_n^T + \mathbf{S}_n\mathbf{A}_n^T\right), \end{aligned} \quad (2.10)$$

we can now rewrite the composite covariance matrix and its eigenvalue decomposition as

$$\begin{aligned} \mathbf{C}_+ &= \mathbf{C}_1 + \mathbf{C}_2 \\ &= \mathbf{C}_{S_1S_1} + \mathbf{C}_{S_2S_2} + \epsilon(\mathbf{C}_{AS_1} + \mathbf{C}_{AS_2}) \\ &= \mathbf{U}\mathbf{\Gamma}\mathbf{U}^T\,. \end{aligned} \quad (2.11)$$

This decomposition is important because we wish to find an expression for the whitening matrix $\mathbf{P}$ in order to find the whitened composite covariance matrix as in Equation (1.5). To determine expressions for $\mathbf{U}$ and $\mathbf{\Gamma}$, we assume that there is a component corresponding to the background noise, and a perturbation to those values caused by the signal. We assume the following definitions

$$\Gamma_k \triangleq \Lambda_k + \epsilon\delta\Lambda_k \quad (2.12)$$

$$\mathbf{u}_k \triangleq \mathbf{v}_k + \epsilon\delta\mathbf{v}_k \quad (2.13)$$

$$\Gamma_k\mathbf{u}_k \triangleq (\mathbf{C}_{S_1S_1} + \mathbf{C}_{S_2S_2} + \epsilon\mathbf{C}_{AS})\mathbf{u}_k \quad (2.14)$$

$$\Lambda_k\mathbf{v}_k \triangleq (\mathbf{C}_{S_1S_1} + \mathbf{C}_{S_2S_2})\mathbf{v}_k\,, \quad (2.15)$$

where $\mathbf{C}_{AS}$ is equal to $\mathbf{C}_{AS_1} + \mathbf{C}_{AS_2}$. Here, $\mathbf{u}_k$ and Γ_k are respectively the individual eigenvectors and eigenvalues of the whole system, whereas $\mathbf{v}_k$ and Λ_k are

the eigenvectors and eigenvalues of the unperturbed system. $\mathbf{U}$, $\mathbf{\Gamma}$, $\mathbf{V}$, and $\mathbf{\Lambda}$ are their matrix counterparts, respectively. The values associated with the unperturbed system, as well as the models for the noise and the planet PSFs are all quantities known a priori.

We also note that the perturbations to the eigenvectors can be expressed in terms of a sum of the eigenvectors

$$\delta\mathbf{v}_k = \sum_{p=1}^{N} c_{k,p}\mathbf{v}_p \,. \tag{2.16}$$

Here, $c_{k,p}$ is simply a coefficient for the linear combination of all eignevctors p to create the perturbation to eigenvector k. This can be done because the covariance matrix $\mathbf{C}_+$ is positive definite, thus each of the eigenvectors is linearly independent and form a complete basis set. The perturbations can thus be expressed in this basis set.

With this information, we can rewrite Equation (2.14) as

$$(\mathbf{C}_{S_1S_1} + \mathbf{C}_{S_2S_2} + \epsilon\mathbf{C}_{AS})(\mathbf{v}_k + \epsilon\delta\mathbf{v}_k) = (\Lambda_k + \epsilon\delta\Lambda_k)(\mathbf{v}_k + \epsilon\delta\mathbf{v}_k) \tag{2.17}$$

Isolating first order terms of epsilon yields

$$(\mathbf{C}_{S_1S_1} + \mathbf{C}_{S_2S_2})\,\delta\mathbf{v}_k + \mathbf{C}_{AS}\mathbf{v}_k = \delta\Lambda_k\mathbf{v}_k + \Lambda_k\delta\mathbf{v}_k \,. \tag{2.18}$$

Using the identities from Equation (2.15) and Equation (2.16), we can simplify to

$$\mathbf{C}_{AS}\mathbf{v}_k + \sum_{p=1}^{N} c_{k,p}\Lambda_p\mathbf{v}_p = \delta\Lambda_k\mathbf{v}_k + \Lambda_k\sum_{p=1}^{N} c_{k,p}\mathbf{v}_p \,. \tag{2.19}$$

Equation (2.19) can be solved to find expressions for $\delta\mathbf{v}_k$ and $c_{k,p}$. We know that the eigenvectors $\mathbf{V}$ form an orthonormal basis, so that $\mathbf{v}_k^T\mathbf{v}_p = 1$ for $k = p$, and zero otherwise. Using this information, we left multiply Equation (2.19) by $\mathbf{v}_k^T$,

causing multiple terms to evaluate to 0 or 1, and leaving

$$\delta\Lambda_k = \mathbf{v}_k^T \mathbf{C}_{AS} \mathbf{v}_k \,. \tag{2.20}$$

Alternatively, we can left-multiply the same equation by $\mathbf{v}_p^T$, specifically where $p \neq k$ to generate the equation for $c_{k,p}$,

$$c_{k,p} = \frac{\mathbf{v}_p^T \mathbf{C}_{AS} \mathbf{v}_k}{\Lambda_k - \Lambda_p} \,. \tag{2.21}$$

Now, we have complete expressions for Γ_k and $\mathbf{u}_k$. We can then find $\bar{\mathbf{C}}_1$ as

$$\begin{aligned} \bar{\mathbf{C}}_1 &= \mathbf{\Gamma}^{-1/2} \mathbf{U}^T \mathbf{C}_1 \mathbf{U} \mathbf{\Gamma}^{-1/2} \\ &= (\mathbf{\Lambda} + \epsilon\delta\mathbf{\Lambda})^{-1/2} (\mathbf{V} + \epsilon\delta\mathbf{V})^T (\mathbf{C}_{S_1S_1} + \epsilon\mathbf{C}_{AS_1}) (\mathbf{V} + \epsilon\delta\mathbf{V}) (\mathbf{\Lambda} + \epsilon\delta\mathbf{\Lambda})^{-1/2} \,. \end{aligned} \tag{2.22}$$

The next step is to group terms by powers of ϵ, and to then drop higher order terms. This is a straightforward process, excluding the $(\mathbf{\Lambda} + \epsilon\delta\mathbf{\Lambda})^{\mathbf{-1/2}}$ term. Given that ϵ is very small, we can substitute the binomial series expansion of this expression. Furthermore, since each of the matrices is a real diagonal matrix, this expansion is well defined as:

$$(\mathbf{\Lambda} + \epsilon\delta\mathbf{\Lambda})^{-1/2} \approx \mathbf{\Lambda}^{-1/2} - \frac{1}{2}\epsilon\mathbf{\Lambda}^{-3/2}\delta\mathbf{\Lambda} \,, \tag{2.23}$$

where we have once again dropped all higher order terms in ϵ.

If we take $\bar{\mathbf{C}}_1$ to be the sum of components associated with the speckle noise, and others associated with perturbations due to the planet signal,

$$\bar{\mathbf{C}}_1 = \bar{\mathbf{C}}_S + \epsilon\bar{\mathbf{C}}_A \tag{2.24}$$

then

$$\bar{\mathbf{C}}_S = \mathbf{\Lambda}^{-1/2} \mathbf{V}^T \mathbf{C}_{S_1S_1} \mathbf{V} \mathbf{\Lambda}^{-1/2} \tag{2.25}$$

and

$$\begin{aligned} \bar{\mathbf{C}}_A = \mathbf{\Gamma}^{-1/2} &(\mathbf{V}^T \mathbf{C}_{S_1S_1} \delta\mathbf{V} + \delta\mathbf{V}^T \mathbf{C}_{S_1S_1} \mathbf{V} + \mathbf{U}^T \mathbf{C}_{AS} \mathbf{U}) \mathbf{\Gamma}^{-1/2} \\ &- \frac{1}{2}\mathbf{\Lambda}^{-1/2} \mathbf{V}^T \mathbf{C}_{S_1S_1} \mathbf{V} \mathbf{\Lambda}^{-3/2} \delta\mathbf{\Lambda} - \frac{1}{2}\mathbf{\Lambda}^{-3/2} \delta\mathbf{\Lambda} \mathbf{V}^T \mathbf{C}_{S_1S_1} \mathbf{V} \mathbf{\Lambda}^{-1/2} \,. \end{aligned} \tag{2.26}$$

We now have an expression for $\bar{\mathbf{C}}_1$ in terms of the unperturbed data, the csp reduction of the unperturbed data, and our model of the planet signal. These are all terms we can pre-calculate. We must now evaluate the eigenvalue problem posed in Equation (1.10) using these terms. Again, we assume that the quantity we wish to decompose, $\bar{\mathbf{C}}_1$, has eigenvalues and eigenvectors composed of unperturbed components and a perturbation due to the planet signal

$$\Phi_k \triangleq \Omega_k + \epsilon\delta\Omega_k \tag{2.27}$$

$$\mathbf{w}_k \triangleq \mathbf{y}_k + \epsilon\delta\mathbf{y}_k \tag{2.28}$$

$$\bar{\mathbf{C}}_1\mathbf{w}_k = (\bar{\mathbf{C}}_S + \epsilon\bar{\mathbf{C}}_A)\mathbf{w}_k \triangleq \Phi_k\mathbf{w}_k \tag{2.29}$$

$$\bar{\mathbf{C}}_S\mathbf{y}_k \triangleq \Omega_k\mathbf{y}_k\,. \tag{2.30}$$

Once again, Φ_k and $\mathbf{w}_k$ are the individual eigenvalues and eigenvectors of the whole system, whereas Ω_k and $\mathbf{y}_k$ are the eigenvalues and eigenvectors of the unperturbed system. $\mathbf{\Phi}$, $\mathbf{W}$, $\mathbf{\Omega}$, and $\mathbf{Y}$ are their matrix counterparts, respectively.

We can use the same methods here to find $\delta\Omega_k$ and $\delta\mathbf{y}_k$ that were used to generate Equation (2.20) and Equation (2.21). This results in

$$\delta\Omega_k = \mathbf{y}_k^T\bar{\mathbf{C}}_A\mathbf{y}_k \tag{2.31}$$

$$\delta\mathbf{y}_k = \sum_{p=1}^{N}\frac{\mathbf{y}_p^T\bar{\mathbf{C}}_A\mathbf{y}_k}{\Omega_k - \Omega_p}\mathbf{y}_p\,. \tag{2.32}$$

These perturbations can be used to find the complete equations for the eigenvalue decomposition of $\bar{\mathbf{C}}_1$. We can now use these terms to find the new CSP projection matrix

$$\begin{aligned}\mathbf{Z} &= \mathbf{\Phi}^{-1/2}\mathbf{W}^T\mathbf{P}\mathbf{X}_1 \\ &= \mathbf{\Phi}^{-1/2}\mathbf{W}^T\mathbf{\Gamma}^{-1/2}\mathbf{U}^T\mathbf{X}_1\,.\end{aligned} \tag{2.33}$$

We replace $\mathbf{W}$, $\mathbf{\Phi}$, $\mathbf{\Gamma}$, and $\mathbf{U}$ with the known noise CSP reduction plus perturbations. After this, we separate like terms such that $\mathbf{Z}$ is also a sum of the known reduction and its perturbations

$$\mathbf{Z} = \mathbf{Z}_s + \epsilon\delta\mathbf{Z}\,. \tag{2.34}$$

After separating like terms, and again using the binomial expansion for certain terms, the two components of $\mathbf{Z}$ can be written as

$$\mathbf{Z}_s = \mathbf{\Omega}^{-1/2}\mathbf{W}^T\mathbf{\Lambda}^{-1/2}\mathbf{V}^T\mathbf{S}_1 \tag{2.35}$$

$$\begin{aligned}\delta\mathbf{Z} = &\underbrace{\mathbf{\Omega}^{-1/2}\mathbf{Y}^T\mathbf{\Lambda}^{-1/2}\mathbf{V}^T\mathbf{A}_1}_{\text{Model Interaction with Speckle Components}} + \underbrace{\mathbf{\Omega}^{-1/2}\mathbf{Y}^T\mathbf{\Lambda}^{-1/2}\delta\mathbf{V}^T\mathbf{S}_1 - \frac{1}{2}\mathbf{\Omega}^{-1/2}\mathbf{Y}^T\mathbf{\Lambda}^{-3/2}\delta\mathbf{\Lambda}\mathbf{V}^T\mathbf{S}_1}_{\text{Effect on Composite Covariance Eigen Decomposition}} \\ &+ \underbrace{\mathbf{\Omega}^{-1/2}\delta\mathbf{Y}^T\mathbf{\Lambda}^{-1/2}\mathbf{V}^T\mathbf{S}_1 - \frac{1}{2}\mathbf{\Omega}^{-3/2}\delta\mathbf{\Omega}\mathbf{Y}^T\mathbf{\Lambda}^{-1/2}\mathbf{V}^T\mathbf{S}_1}_{\text{Effect on Whitened } \mathbf{X_1} \text{ Covariance Eigen Decomposition}}\,.\end{aligned} \tag{2.36}$$

This is the first-order accurate model of the impact the planet signal has on the CSP modes, $\delta\mathbf{Z}$.

2.3.3 Model Results

Model Accuracy

The methods described in §2.2 yield results for the β pic data set shown in Figure 2.4. This is the sum of the final two modes of $\mathbf{Z}$, calculated using one wavelength. An injected planet can be faintly seen at pixel location [55,51] and the known exoplanet is shown at [23,32]. Additionally, many residual speckles are

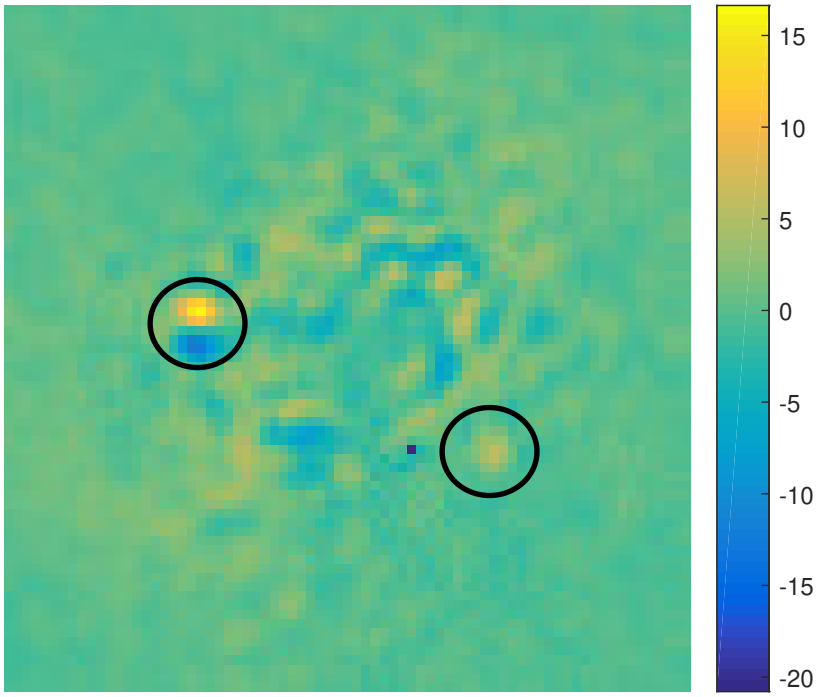

Figure 2.4: This is the sum of the final two modes of the CSP reduction of the β Pic dataset at a wavelength of 1,587.1 nm. The known planet signal is circled on the left, and the injected planet is seen more faintly on the right. Similar speckle noise throughout the image indicates a need for summation of multiple wavelengths and a matched filter approach to increase SNR.

present that can be construed as false positives. Upon addition of similar results at each wavelength, the signals would appear at a higher signal-to-noise ratio. Nonetheless, background speckles would still be present. This necessitates the use of a matched filter between the forward model results of $\delta\mathbf{Z}$, and the calculated $\mathbf{Z}$ from the CSP reduction. This approach only works, however, if the forward model itself is accurate. A quantitative analysis of the errors associated with the forward model is discussed in this section.

We use two metrics to determine the accuracy of the forward model. The first is a comparison between the estimated value of $\mathbf{Z}$ from the forward model, and the values from evaluating the CSP process. To do so, we calculate the percent

error of the estimated matrix compared to the value of the known matrix

$$\% \text{ error} = \frac{\mathbf{Z}_{(i,j),CSP} - \mathbf{Z}_{(i,j),FM}}{\mathbf{Z}_{(i,j),CSP}} 100\% \tag{2.37}$$

for every entry (i, j) in the matrix. The FM subscript indicates a forward model result, and the CSP subscript similarly identifies the CSP reduction result. An evaluation of all the errors in $\mathbf{Z}$ for the HD 14706 dataset is shown in Figure 2.5. The errors appear to be roughly normally distributed in $\log(\%error)$. In the

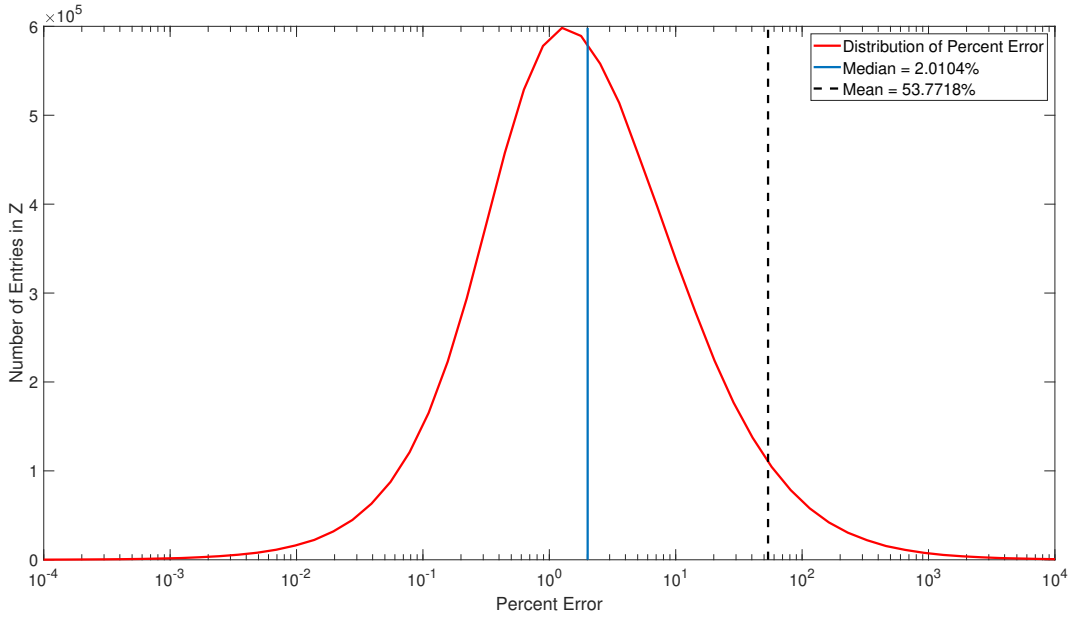

Figure 2.5: Analysis of error in the forward model of the projection matrix, $\mathbf{Z}$, compared to its true value. Many errors are small, and those that are very high likely correspond to entries in the matrix that have very low absolute values.

$\mathbf{Z}$ matrix, many entries are very small values. Thus a slight absolute error can yield very large relative errors, which is shown in Figure 2.5. However, the important components of our analysis are the elements of $\mathbf{Z}$ that are not inherently small. While some entries of $\mathbf{Z}$ may have high individual errors, their very small absolute values will likely not propagate towards the actual signal. Thus, it is a much more relevant analysis to determine the difference between the FM signal and the CSP evaluation in an area of interest - surrounding the planet signal. A

comparison of the final mode of both CSP and the forward model is shown in Figure 2.6. The values for entries in the modes centered around the target signal are directly compared. In these images, it can clearly be seen that there is signif-

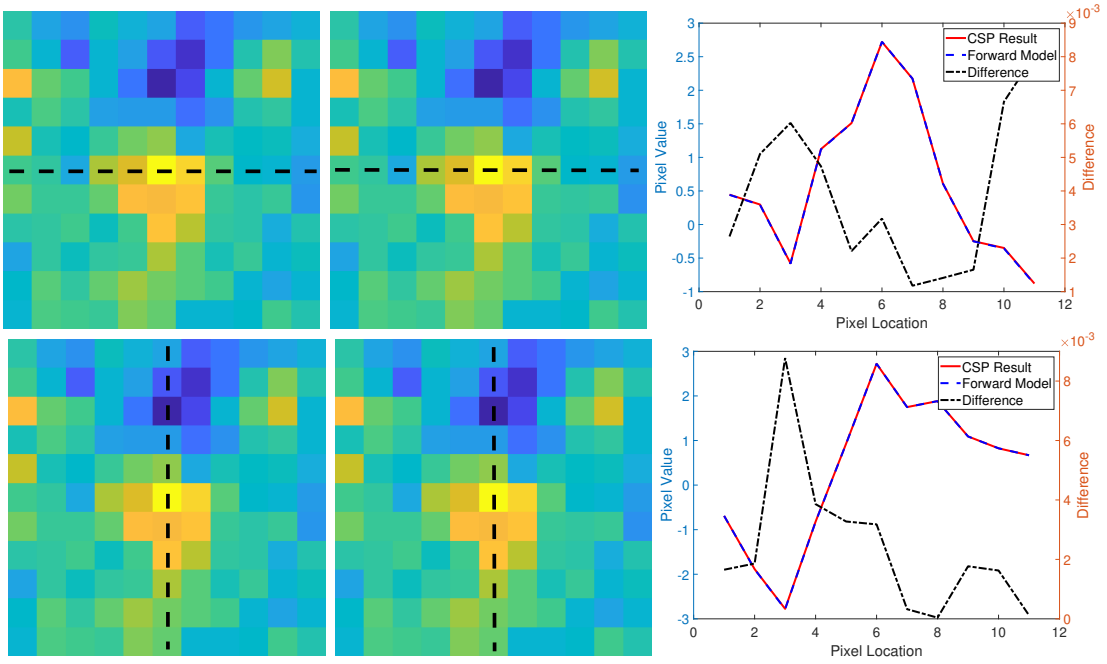

Figure 2.6: The forward model results compared to the CSP results. Top row: horizontal cut. Bottom row: vertical cut. Left column: CSP result. Middle column: forward model result. Right column: pixel comparison. The forward model very accurately resembles the actual CSP results. The difference between the two in terms of pixel value is shown with the dotted black line on the right y-axis.

icant overlap between the true CSP result and the predictive forward model. By both metrics, the median of $\mathbf{Z}$ error and the comparison in signals, the forward model appears to perform well. Furthermore, we can describe the normalized maximum likelihood estimation computed on the above window between the forward model and the actual planet signal as

$$F = \frac{\sum_{i=1}^{121} y_i x_i}{\sum_{i=1}^{121} x_i^2}, \tag{2.38}$$

where x is the CSP results and y is the forward model prediction. Thus, the

photometric error introduced by the forward model can be described as

$$|F - 1| \tag{2.39}$$

which evaluates to 0.0020 for the data shown in Figure 2.6 across the whole image. This is a strong indicator that the forward model predicts the behavior very well. Nonetheless, a true test of its usefulness will be when it is used as a matched filter on a known signal, as demonstrated in §2.4.

Image Number Error

One key result is that the forward model scales very poorly with the number of images. The photometric error in **Z** rises drastically, shown in Figure 2.7.

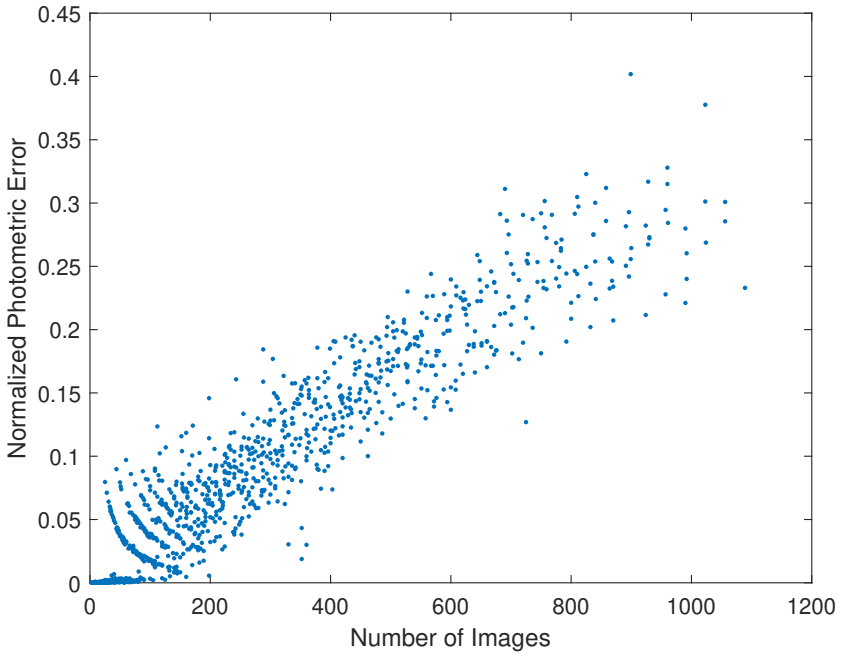

Figure 2.7: Forward Model error as a function of number of images. As the number of images used to calculate the forward model increases, the photometric error of the forward model as described in Equation (2.39) increases linearly.

The data suggests that the error grows with the number of images. This in-

crease in error is due to the accumulation of small errors in an increasing number of small eigenvalues. These eigenvalues correspond to the low-information modes. Any given image from an observing sequence bears resemblance to and shares some information with the other images from the full dataset. Thus, adding an image to a dataset does not add a full image's worth of information to the dataset. As the number of images increase, the total number of modes needed to express the information does not increase linearly. This results in an increasing fraction of very small eigenvalues as more images from the same observation are used in a dataset. In turn, these small numbers are subtracted and used in denominators as in Equation (2.21) and Equation (2.36). Furthermore, these small errors are a direct result of truncating to a first order model.

We can solve the issue of scaling error by simply using fewer images in the CSP process. We still span the entire image space by conducting the CSP process multiple times and summing the result, without the drawback of accumulated error.

2.3.4 Informing Mode Selections

Another useful result of the forward model is its ability to help guide our decisions in tuning the parameters of the CSP reduction. Namely, the model helps us select which modes to use in the matched filter of $\mathbf{Z}$. To do so, we can directly examine the modes from $\delta\mathbf{Z}$ and $\mathbf{Z}$. For the smallest wavelength in the β Pic data (1,494.6 nm), every mode from $\delta\mathbf{Z}$ is shown in Figure 2.8.

The planet signal can be seen in the bottom right quadrant of nearly every image. It is often displaced to various locations that span the azimuthal rotation

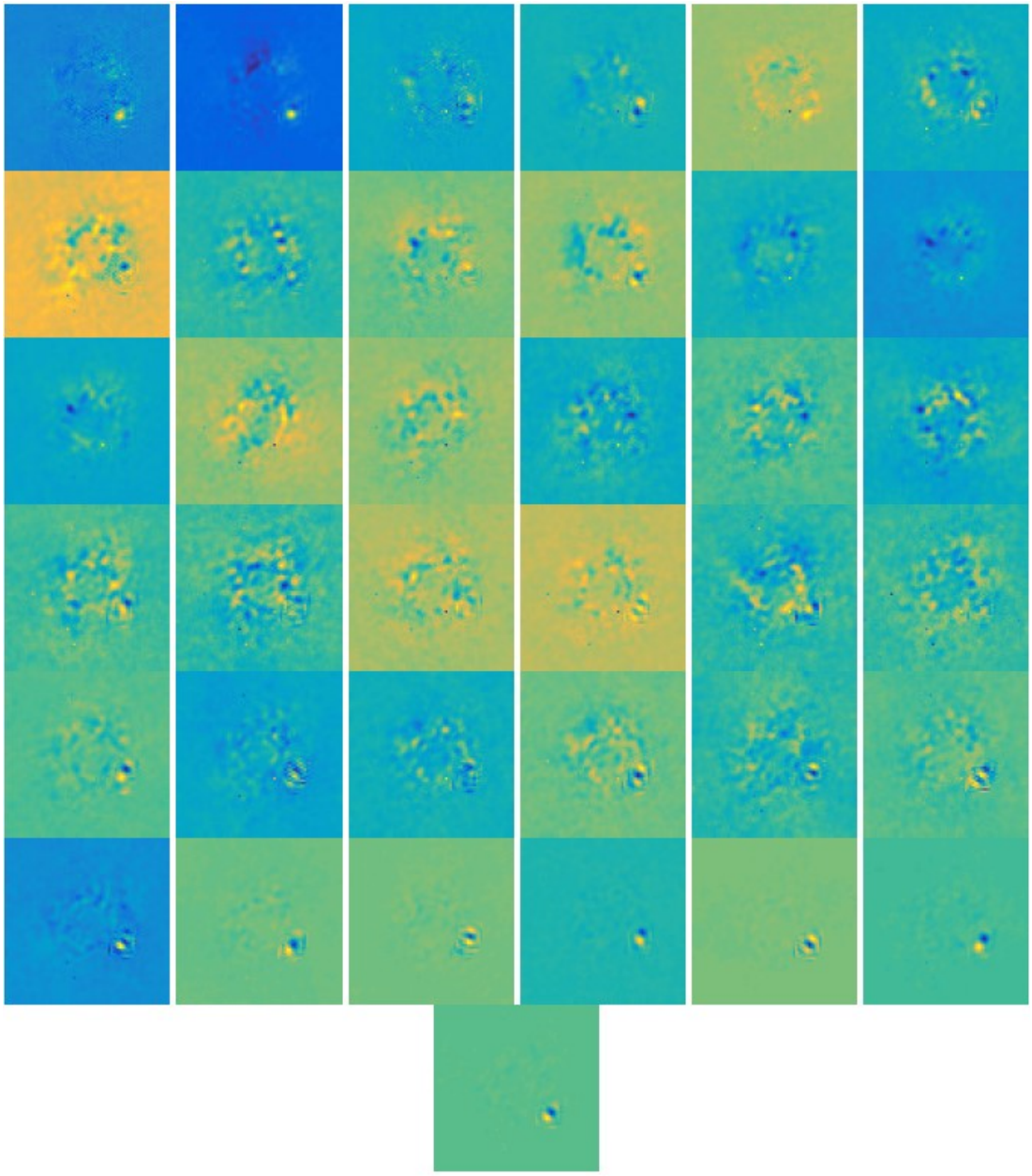

Figure 2.8: Each of the modes of $\delta\mathbf{Z}$. While the signal can be seen in most modes, it is isolated from other speckle effects in the final modes.

from ADI. The movement of the signal from the injected location is particularly indicative of the need for a forward model matched filter for CSP. Using it, we can achieve a significantly more accurate astrometric result for the location of the planet.

Furthermore, the planet signal is mostly accompanied by other sources of noise in the data. It is only in the final modes (32-37), that the planet signal is mostly isolated from other sources of noise. However, the planet signal is also very strong in the first few modes, and prevalent in others as well. This suggests that there is an optimal set of modes to combine, but it is not limited to simply the last set of modes (or the first set, for that matter). The actual modes used will vary from system to system.

2.4 Matched Filter

Having derived a model for the planet from the satellite spots, we can compute a forward model for the planet signal at any location in the image. We accomplish this by representing a given image $\mathbf{i}$ as $\mathbf{i} = \mathbf{i} - \epsilon\mathbf{a} + \epsilon\mathbf{a}$, where $\mathbf{i} - \epsilon\mathbf{a}$ is the model for speckles ($\mathbf{s}$ in Equation (2.3)). We do this at every spatial location in a set of images. We then compute the forward model at every location, providing estimates in terms of $\mathbf{Z}_s$ and $\delta\mathbf{Z}$ for every possible target signal location. Whichever of the modes of $\delta\mathbf{Z}$ matches best (as evaluated via a matched filter) with the corresponding CSP reduction modes of $\mathbf{Z}$ is the most likely location of the planet.

The template used for the matched filter at spatial position x is a 21×21 stamp of pixels centered at x within mode $\delta\mathbf{z}_k$ based on the forward model from a signal injection centered at x. This is matched with the corresponding 21×21 stamp from the mode $\mathbf{z}_k$.

We calculate the matched filter result at x by computing a cross-correlation. We multiply each pixel p in the forward model stamp with the corresponding

pixel in actual reduction. We sum the products for each pixel location to get the matched filter result at a single pixel. We repeat this for every spatial location x. This process generates a matched filter map for each of the modes for a given CSP reduction. By adding each of these maps, we retrieve the signal for one CSP processing. Recall that we conduct CSP for each wavelength independently. Thus, we further stack each result from independent wavelengths, improving the signal-to-noise ratio (SNR) of the result. The sum $\hat{\mathbf{s}}$ for a dataset with l different wavelengths and K different modes used

$$\hat{\mathbf{s}}(x) = \sum_{\lambda=\lambda_1}^{\lambda_l} \sum_{k=1}^{K} \sum_{p=1}^{441} p_{FM_{\lambda,k}} p_{CSP_{\lambda,k}} \tag{2.40}$$

represents the final image with which to evaluate for exoplanets.

The final step is to determine which modes are used to evaluate with a matched filter. The goal is to use an optimal set of modes for SNR recovery, knowing that the optimal set of modes is not always exclusively the low-eigenvalue modes. However, given that the final modes are uniformly good for this process, we begin by automatically including the last mode. For our purposes, the signal to noise ratio is defined as the matched filter result of a single pixel location divided by the variance of the the matched filter result for every spatial location in an annulus around the star that intersects the target signal. The known target signals are masked out of this calculation. Mathematically, we define

$$SNR = \frac{\hat{\mathbf{s}}(x) - \mu(x)}{\sigma(x)}, \tag{2.41}$$

where $\mu(x)$ is the mean of the annulus for the pixel at x and the $\sigma(x)$ is the standard deviation of that annulus. The signal estimate and noise values are all taken after convolution with the matched filter.

If the dataset contains n timesteps, it will also contain n modes. This final nth mode begins as our analysis image. We proceed by adding the $n-1$ mode as well. If the SNR at a given pixel location improves, we set the new analysis image to be the sum of the modes n and $n-1$. Otherwise, we keep the analysis image as just mode n. This process repeats with subsequently higher and higher eigenvalue modes, adding them to the analysis image if and only if the SNR of the selected spatial location increases. In each calculation of SNR, the values for the mean and standard deviation of the noise as well as the signal estimate is recomputed iteratively, based on the included modes.

When this process is completed, we have the optimal SNR combination for that location. This process, starting with just mode n, must be replicated for every pixel. This results in a different optimal image for every single spatial location. Essentially, we create an image significance map, full with the peak SNR of every single pixel. This is shown in Figure 2.9 for both images.

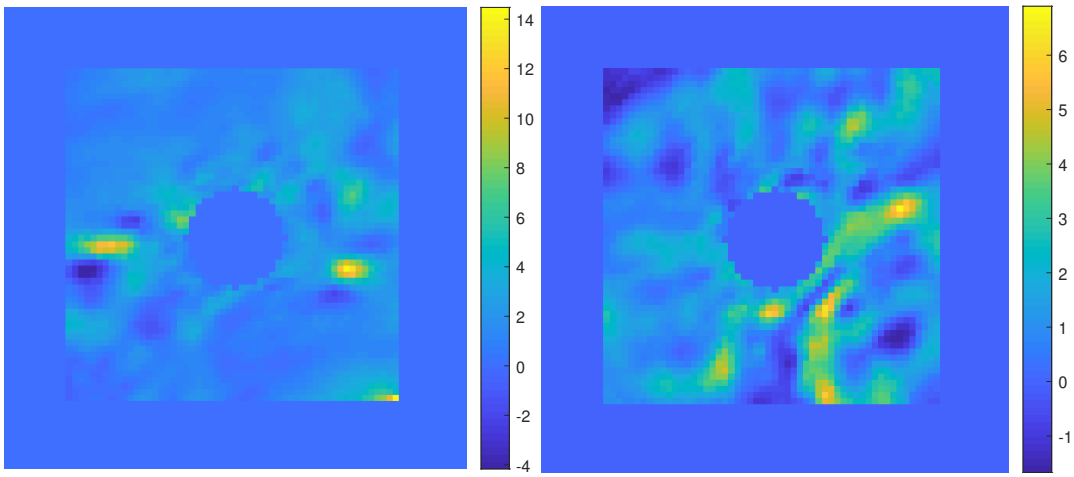

Figure 2.9: This shows the optimal SNR map for every pixel in both data sets. The left image is the β Pic data, and the right is the HD 14706 data. In both datasets, the planet locations have the highest SNR. However, the method of maximizing SNR can also increase the rate of false positives, as the HD 14706 data shows an increase in SNR for the correlated, speckle noise as well.

For computational efficiency, we do not in practice compute the forward model for every possible signal location. Instead, we compute a base forward model for one single location. We can then rotate and scale this model to each new location. Rotation is computed based on azimuthal location, and the template is scaled linearly based on the distance from star's center to the injected signals' spatial location in the image. This method demonstrates its efficacy by correctly identifying the β Pic b known exoplanet, despite its differences in angular separation and azimuthal location.

2.5 Results

To examine the performance of the CSP FMMF algorithm, we must evaluate its results in the context of the rest of the field. For direct imaging post-processing, a recent advance in the state of the field is KLIP-FMMF. Thus, I have analyzed the two datasets described in §2.1.2 with both CSP-FMMF and KLIP-FMMF. The latter was completed using the FMMF module from pyKLIP, a collection of codes used as an implementation of the KLIP algorithm [36].

Both results, for both datasets, are shown together in Figure 2.10. In the β Pic data, the injected planet can be seen at on the right, and the known exoplanet is visible on the left of the image. The signal injected into HD 14706 data can be found on the right as well. In these reductions, every wavelength was treated with $K = 2$ modes. In these CSP-FMMF results, a border with a width equal to half of the stamp size (in this case, 10 pixels) is not evaluated as the template would go beyond the edge of the data otherwise.

Comparing the results, the background speckles of the CSP-FMMF are much

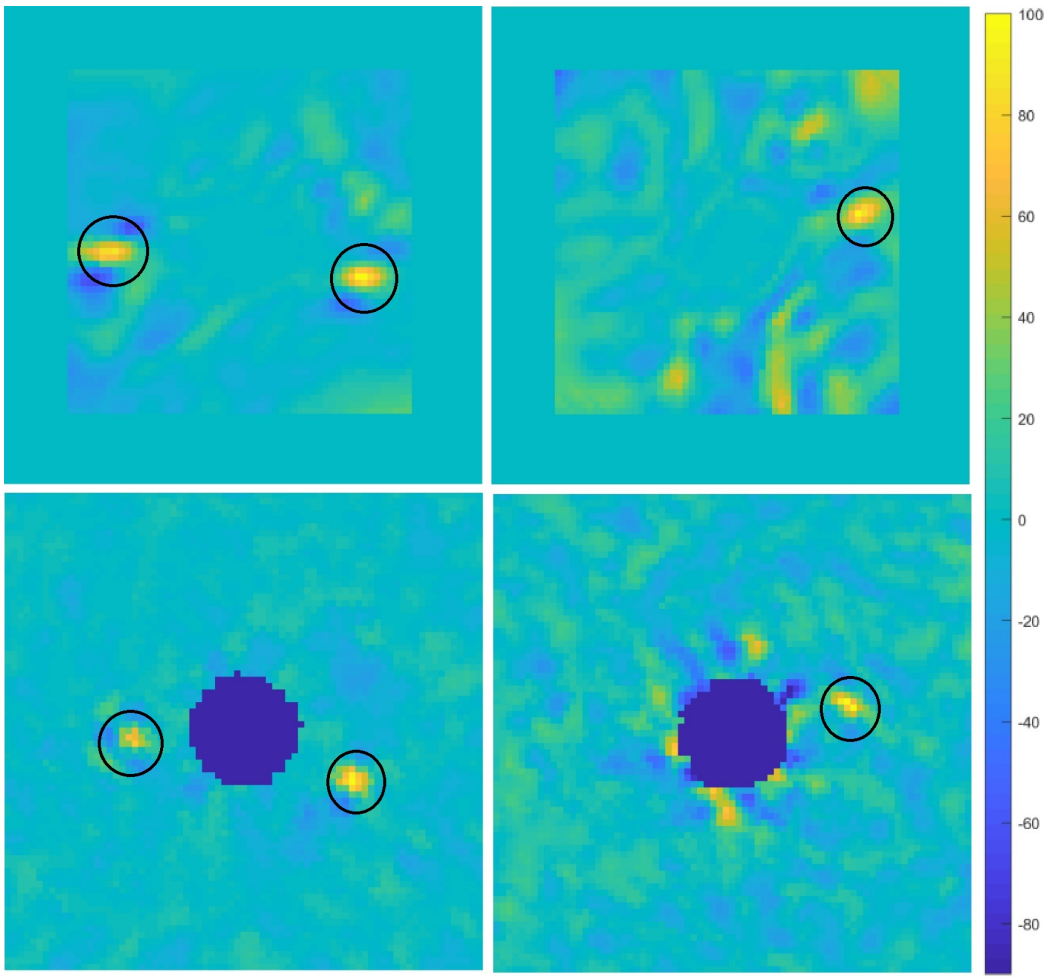

Figure 2.10: Top: CSP FMMF Reduction. Bottom: PCA FMMF Reduction. Left: β Pic data. Right: HD 14706 data. Each image is shown in a linear scaling law. All images clearly show the injected planet signal, and the left images show β Pic b. CSP-FMMF recovers the signal at slightly higher SNR, and much worse astrometric error. Background noise (leading to more false positives) are also more present in the CSP-FMMF data.

less uniform than from pyKLIP, but the signals appear much stronger; the quantification of this difference is the SNR.

In quantitatively comparing to previously existing methods, two essential metrics are important: signal recovery (while minimizing false positives) and astrometric biasing. Signal recovery can be defined by the SNR, whereas astrometric biasing is the ability to accurately determine the location of the planet

and is characterized by the error from the true location. The SNR results for the three known planets are summarized in Table 2.1. The CSP-FMMF result shows

	SNR	
	CSP	KLIP
HD 14706 Injected	6.90	10.65
β Pic Injected	14.47	13.05
β Pic b	11.79	9.64

Table 2.1: Comparison between KLIP and CSP Processing. The CSP results shown are from the optimal analysis for the injected planet in both images.

a higher SNR for both planets in the β Pic data, but not in the HD 14706 data.

A note of caution here is that the KLIP-FMMF algorithm was designed and tuned for detections of planets dimmer than β Pic b. This is a primary reason for the low SNR of that detection in this case. Further study is warranted with a wider range of planet signals. Furthermore, none of the distributions of noise in the annuli around the planet signals in any image, for CSP or KLIP, are Gaussian. This tells us the variance in the noise, by itself, is not sufficient to describe the noise distribution. This indicates that the SNR metric is not perfectly appropriate for either distribution. However, all six annuli have noise distributions that belong the the Pearson Type IV distribution [17]. Each of the Pearson distributions uses 4 moments to describe a distribution - skewness and kurtosis in addition to the mean and standard deviation from a Gaussian description. In each of the six annuli, the sample skewness (β_1) and kurtosis (β_2) both fall in the Type IV region of differing Pearson distribution types. This is purely based on the classification system for Pearson distributions on the $\beta_1 - \beta_2$ phase space [2]. These Type IV distributions typically have fatter tails than standard Gaussian distributions, which indicates that SNR as a metric is too lenient in terms of a simple probability estimate that the observation is a part of the noise. For analysis, however, I retain the convenient, common practice of using exclusively

the standard deviations to compute the SNR. I nonetheless caution the reader against a straightforward interpretation of these SNR values.

I also note that the CSP-FMMF results also have much stronger false positives than the KLIP-FMMF results. This is particularly evident in the HD 14706 data. Typically, this false alarm rate would be analyzed with a Receiver Operating Characteristic [4]; however, in these 4 images the planet signals are all at significantly higher SNR than any potential false positives, rendering this particular analysis unhelpful.

In regards to the astrometric biasing, the centroids of the KLIP-FMMF resultant spots are within sub-pixel accuracy to the known location of the two injected planets. Meanwhile, the CSP-FMMF results deviate from the known location by more than a pixel. In the $\delta\mathbf{Z}$ modes of CSP-FMMF, the planet signal appears in a variety of locations along the azimuthal motion caused by ADI. The location of the signal per mode, however, is not directly related to the parallactic angle. For example, two signals that are initially injected at differing locations may show the planet signal in the final mode in the same spatial location, assuming that that location is part of the motion of both signals through time. Thus, the recovery of the location of a signal with CSP-FMMF is much more difficult. Other methods, like KLIP-BKA are a superior technique in this regard [35].

2.6 Conclusion

In this chapter, I have shown a new method to be used for planetary signal extraction from direct imaging data of exoplanets based on the algorithm for

Common Spatial Pattern filtering to detect differences between to datasets. Furthermore, I have developed forward model that can predict the effects of CSP on a point source in data. This forward model has been used as a template for a matched filter with a blind CSP reduction to create a significance map. This map can be used to determine whether a planet exists at a specific location.

Nonetheless, this work demonstrates that Common Spatial Pattern filtering is an effective planetary signal extraction technique and, rather than subtracting away speckles from science images, it represents a way to directly model planet signals in the data.

CHAPTER 3

DATA ANALYSIS OF THE CSPFMMF ALGORITHM

In **Chapter 2**, common spatial pattern filtering with a forward model matched filter was shown to be effective at identifying planets. This technique directly models the planet signal rather than modellng the noise and subtracting it away, as in other popular techniques. This chapter introduces a more robust and refined version of that algorithm, particularly in relation to computing across an entire image and for analyzing varied datasets. This algorithm is released in its entirety in a code base to be used in tandem with pyKLIP [36]. I use this software to investigate the proper parameters and ways to use this algorithm. With these tools, we have analyzed a large portion of GPIES [11] data, with injected planet signals. Finally, I show the resultant statistical analyses that can quantify the performance of this new algorithm.

To illustrate some results initially, I have provided some end products for two common astrophysical signals, shown in Figure 3.1s. The planet signals are all readily visible.

3.1 Common Spatial Pattern Filtering Forward Model Matched Filter

3.1.1 Algorithmic Updates

The algorithm itself has undergone many improvements from the processes described in **Chapter 2**. These updates accomplish the goals of:

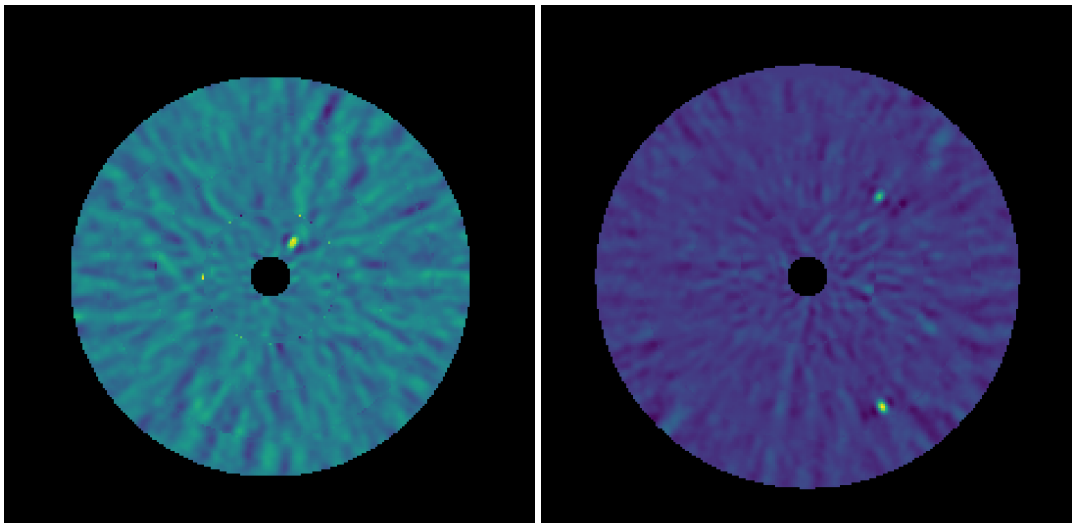

Figure 3.1: Examples of well-known directly imaged exoplanets. Beta Pictorus b is on the left, and two of the companions to HR 8799 are shown on the right.

- scaling for full-sized datasets instead of smaller images
- easy integration with pyKLIP codebase
- robustness and precision improvements
- computational efficiency.

These changes are both to the code structure and the mathematical foundation of the algorithm. The primary structural improvements are:

- being rewritten in Python
- parallelization
- reliance on pyKLIP
- removing the 'best SNR Matching' phase for computational efficiency.

Additionally, the direct algorithmic changes are:

- segmenting each image into annuli

- applying a threshold for the matched filter template
- treating each end result with an SNR map.

Subsections

In many cases, including when using standard GPIES data, an entire image cannot be treated at once. The matrices involved in the CSP process all scale geometrically with the number of pixels used per image. As this number increases, computational requirements and time to completion skyrocket. A simple solution is to subdivide the images.

Noise regimes are assumed to be constant in annuli of equal separation in a given observation. Furthermore, the expected motion of the planet signal is in an arc of equal separation. Several annuli of equal width make for natural divisions of an image. The pyKLIP software subdivides each image similarly, with additional segmentations of the ring into into arcs. This is done to further reduce the size of each section. Unfortunately, this is not a viable solution with CSPFMMF, due to fundamental differences in the algorithms. The CSPFMMF process relies on the signal being spatially diverse, unlike KLIP, which derotates the signal to a single spatial location. Thus, our signals are dispersed over an arc, exemplified in Figure 3.2.

Without knowing where a signal may appear a priori, we cannot set limits on the edges of such a segment. Thus, they must span the entire annulus.

The software as written allows for any number of annuli, but for the purposes of this chapter, 4 segments were used for all analyses.

Figure 3.2: The signal is dispersed in an arc that follows the spatial motion of the planet during observation. This was found as a prediction with the forward model.

Thresholding

While the signal itself is readily visible in most modes, like that shown in Figure 3.2, a lot of noise remains in the rest of the section. If we were to blindly match this to the corresponding CSP mode, the chance of this noise correlating with some from that mode could dilute the recovered FMMF result. A sample equivalent CSP-reduced mode is shown in Figure 3.3.

Figure 3.3: CSP mode from second annular ring. No planets are visible, but the annulus still contains correlated noise

The challenge, then, is to isolate the planet signal in the FMMF mode. This is done by thresholding the template developed from the forward model. All pixel locations are set to 0 if they don't reach the threshold. The value is set as n standard deviations from the median value for all pixels in the data set. The median was chosen such that the planet signals do not bias the noise statistics. The example from Figure 3.2 is shown along with its thresholded version in Figure 3.4.

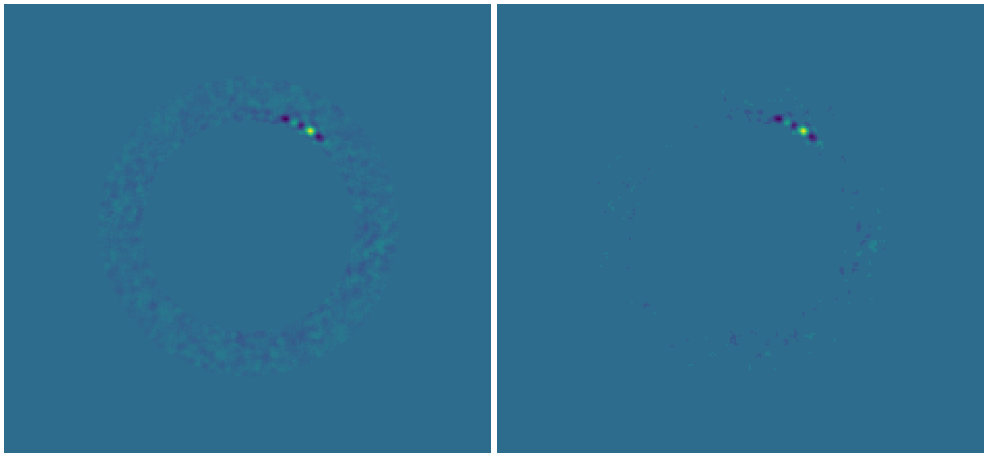

Figure 3.4: The full forward model predicted template, and version with a 5-sigma threshold.

As pictured, the threshold primarily works to set the majority of the noise to zero, while preserving the signal. The threshold can be set to any multiples of standard deviations from the median. As the threshold value increases, the noise in the segment decreases. However, this could potentially come at a sacrifice: more of the signal itself is lost.

Significance Mapping

After setting segments and applying a threshold to the template from the forward model, the matched filter can be computed. This is done across all wave-

lengths and a set number of modes. Afterwards, segmnents are combined and each map is summed together. An example of a summed result can be seen in Figure 3.5. This is injected at contrast of 5×10^{-5}, which is why it is so readily visible.

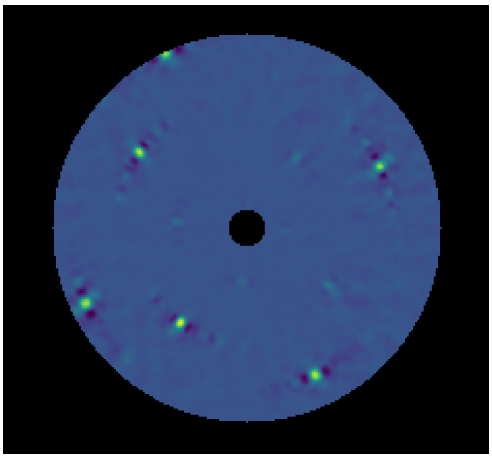

Figure 3.5: Wavelength-summed FMMF map for HD 3888. Each subsection is independent of others. Inner rings are much harder to see.

Due to the segmented processing, each annulus is entirely independent of the others. As can easily be seen, the inner rings are significantly darker than their outer counterparts. This is partially due to the increased impact of coronal noise on rings closer to the inner working angle. However, their proximity is not the only contributing factor. Should that effect be entirely nullified, the inner rings would remain darker. This is thought to be due to the differing areas used for matched filters. Even after thresholding, signals at larger separations sweep out a larger area, leading to an overall stronger signal. This area cannot be known a priori, and thus the difference cannot be compensated for.

This doesn't mean that the rings cannot be compared, nor that inner data is less valuable; each ring simply must be treated independently. Within the environment of one segment, the ratio between signal and noise is appropriately scaled, as seen in Figure 3.6, which is the second segment from Figure 3.5.

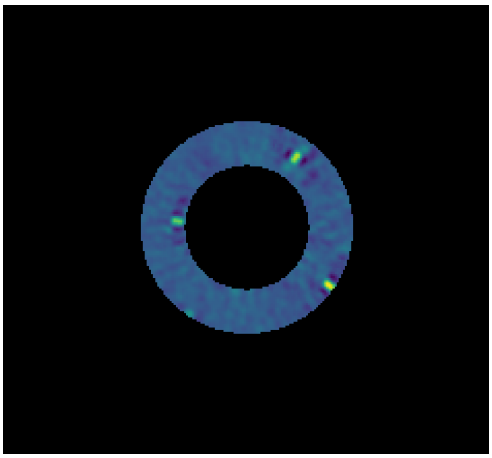

Figure 3.6: The second segment of a CSP result, with 3 identifiable signals.

The planet signals in this ring, when viewed alone, are easily identified. To view an entire dataset on the same scale, one comparison that can be made is a signal-noise map, where each pixel is compared to others at the same separation. The equation for SNR is

$$SNR = \frac{s - \mu}{\sigma}, \tag{3.1}$$

where s is the signal strength, μ is the mean noise in the annulus, and *sigma* the standard deviation of that noise. The particular noise region is set to be an annulus with width twice the Full-Width Half-Max (FWHM) of the instrument PSF. This noise annulus is also bounded by segment borders. Other injection sites are masked out as well. This scheme is shown in Figure 3.7.

The result of applying this strategy to the same dataset can be seen in Figure 3.8. The signals in the inner rings are much more prominently shown, and the resultant map can be compared across separation.

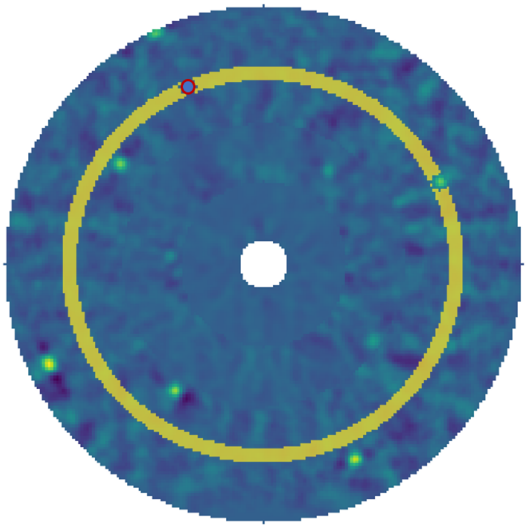

Figure 3.7: SNR Masking. In this example, the SNR is calculated for the point at the center of the red circle, and noise is calculated using every pixel in the yellow region. Note that the planet injection site on the right has been masked out as well.

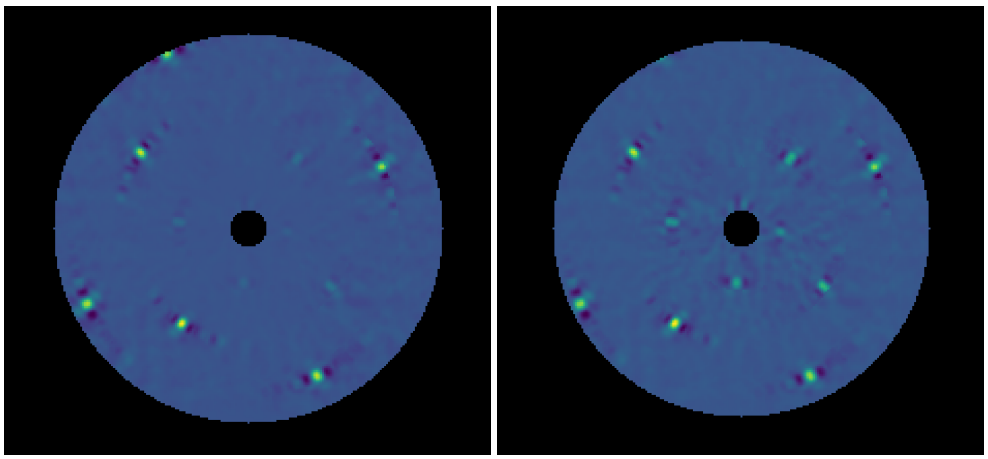

Figure 3.8: Left: Raw CSPFMMF result. Right: SNR Map. When using an SNR map, the entire region is visible. Inner rings are still fainter because they are closer to noise generated by the star.

3.2 Observations and Data Reduction

3.2.1 Observations

A large number of statistics from real, scientific data is necessary to truly develop a good picture of the performance of this algorithm, This is accomplished by injecting planet signals into many, many observations. All original datasets come from the GPIES campaign. However, this observing campaign targeted many different stars and objects in the sky. In order to compare like to like, I set limits on the types of observations to include in analysis. Each observation must:

- be a GPIES target
- have been observed with an H-band filter
- have between 30 and 45 images
- have an average integration time per image between 59 and 60 seconds
- not have observing errors (marked bad)
- not contain any debris or dust disks to bias noise measurement

In all, this led to a group of 337 different observations.

3.2.2 Injected Planets

In addition to the science data from each observation, I have injected known, fake planet signals into each image to be used for analysis and comparison.

To use a valid planetary Point Spread Function (PSF), we recreate behavior from both wavelength dependence and spatial structure. The particular L-Type model planet spectra used were selected from atmospheric models developed by Didier Saumon [27].

The science images contain the information to develop an accurate spatial model for each planet signal. Each GPIES image has four scaled representations of the unblocked stellar PSF, called satellite spots [31, 15]. The model of the planetary signal is created via a median combination of each of the four spots. The wavelength of each image scales the spatial model.

To demonstrate the injected pattern, one dataset was injected with planets far too bright for ordinary analysis. This dataset was just used for illustrative purposes, as in Figure 3.9.

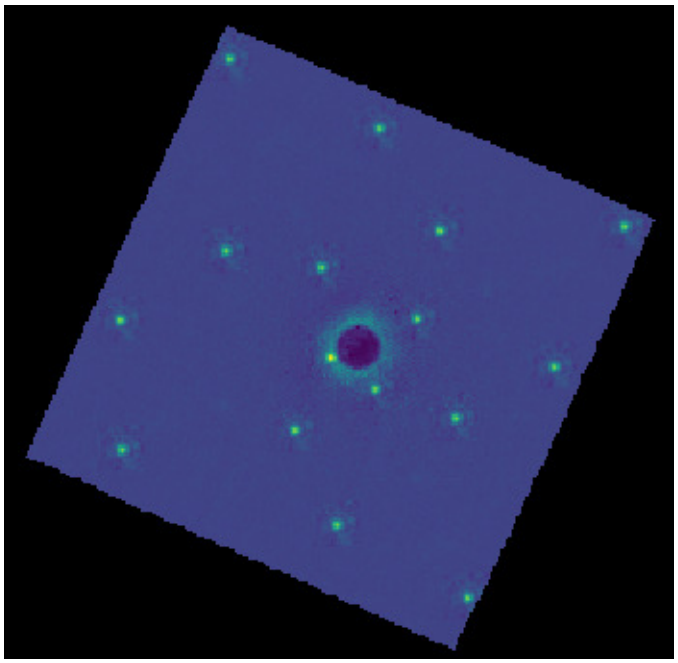

Figure 3.9: Highly exaggerated version for illustration purposes. This is the pattern of injected planets. The square limits on the actual injected planet are also visible.

Normally, the planet signals are significantly dimmer. For analysis, each observation was reduced three times. Each reduction was done with planets that were injected at a different contrast: 5×10^{-5}, 5×10^{-6}, and 5×10^{-7}. This allows for comparison across different strength signals. The contrasts were selected such that the brightest were always visible, the middle strength sometimes visible, and the weakest were very difficult to detect.

3.3 Global Parameter Selection

CSPFFMF uses a combination of parameters to process the data. Two of these variables influence the result significantly: the number of modes to keep and the threshold for the matched filter template. In this section, I examine how to optimally select the value of these parameters from a variety of combinations. One particularly useful example would come from an observation with excellent seeing conditions, such as the first observation of C Eri [13]. From this observation, I have examined the resultant SNR of injected planets at two separations: one midway through the viewing window, and another closer to the outer working angle.

This singular analysis represents an optimum for 2 injections in a dataset. The larger trends, however, must be quantified at more separations. The peaks from 3 sample datasets across all separations are shown in Figure 3.11. The same samples of modes kept and thresholds set as in Figure 3.10.

As can be seen, the most common is a threshold of 5 standard deviations from the mean and 10 modes kept. However, this alone is not overwhelming evidence that these selections are superior. One option to combat this partic-

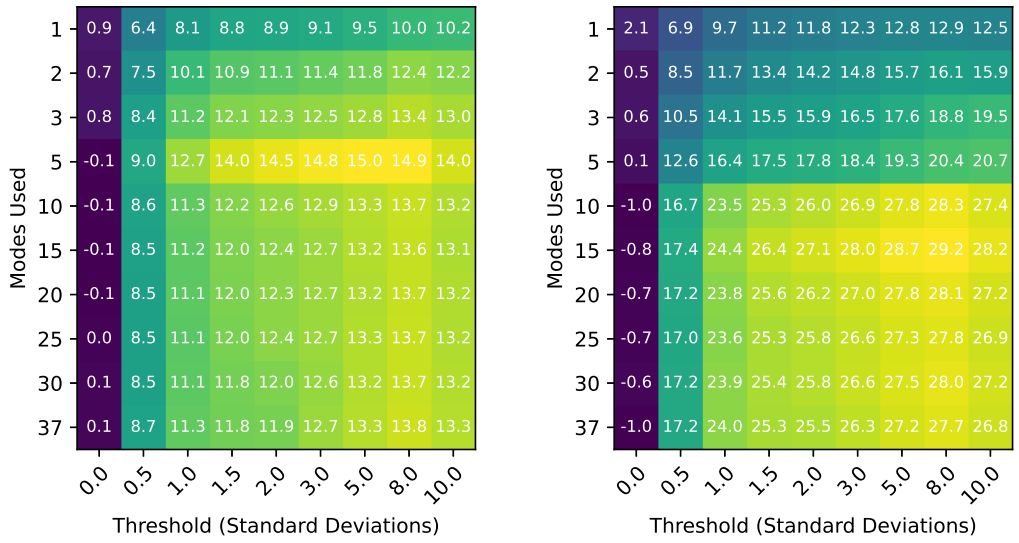

Figure 3.10: Resultant SNRs for two injections with a variety of initial parameters. The highest SNR for the first injection is 14.9 with a 3 threshold and 5 modes used. The highest SNR for the second injection is 29.2 with an 8 threshold and 15 modes used.

ular problem is to simply calculate the result at each possible combination of parameters, and then to take the peak SNR at each pixel location for the given SNR. This strategy was compared to just the 10 modes and 5 threshold strategy for a subset of 35 stars from the dataset. The Receiver Operating Characteristic (ROC) curve is shown for both strategies in Figure 3.12. The exact process for developing these ROC curves is discussed in §3.5.1.

In the long run, the Maximal SNR from all slices captures more of the true planets. However, this is not particularly useful. In real-world use of this algorithm, we only wish to operate in the region of reasonably high SNRs - 5 and above, for example. In this region, shown in more detail by the zoomed-in graph, the single slice metric performs much better. It identifies more planets with fewer false positives. This is likely because the maximum SNR increases the strength of the noise as well as the signal, increasing the number of false

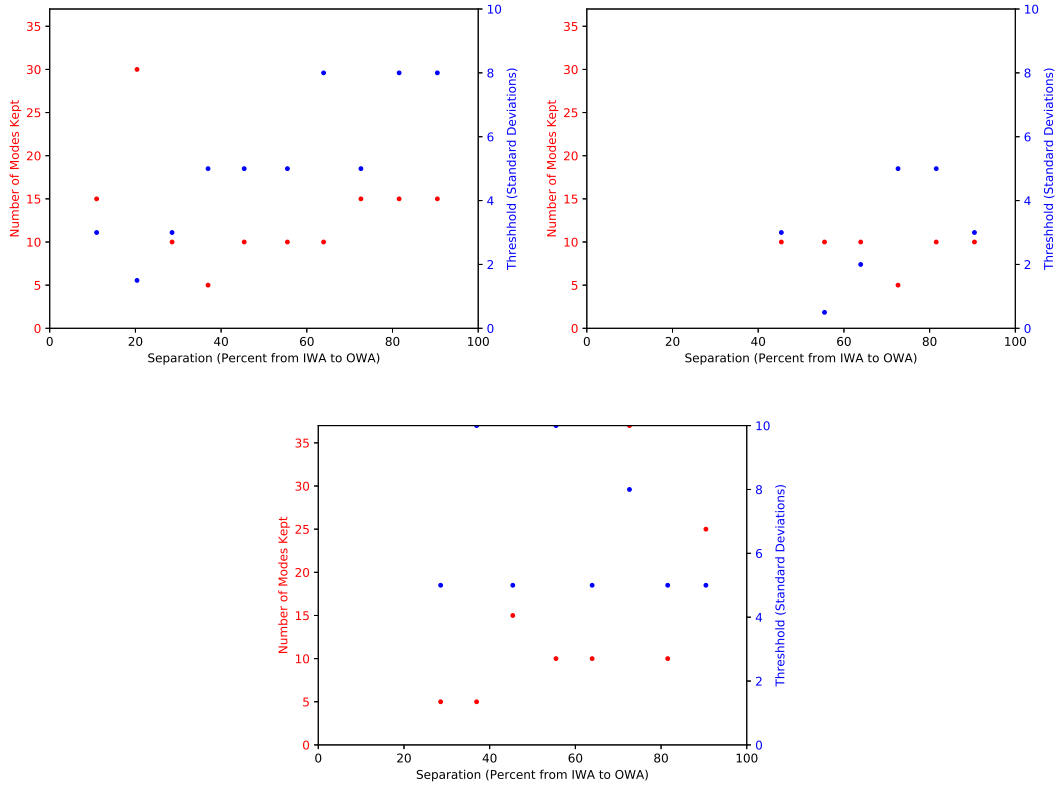

Figure 3.11: Plots of the best combination of parameters for three example observations (C Eri, HR 4372, and HR 7012).

positives. Thus, it is more advantageous to use just one set of parameters. This tactic is also less computationally expensive as well. This was done for the remaining datasets beyond the initial 35 used for testing.

3.4 Circumstellar Disk Use

One other potential use of this algorithm is to examine circumstellar disks. As a brief example to assess performance, I put an observation HD 4796 through the CSPFMMF process. The resultant SNR map is shown in Figure 3.13.

Clearly, the algorithm identifies and pulls out the disk. Any individual point

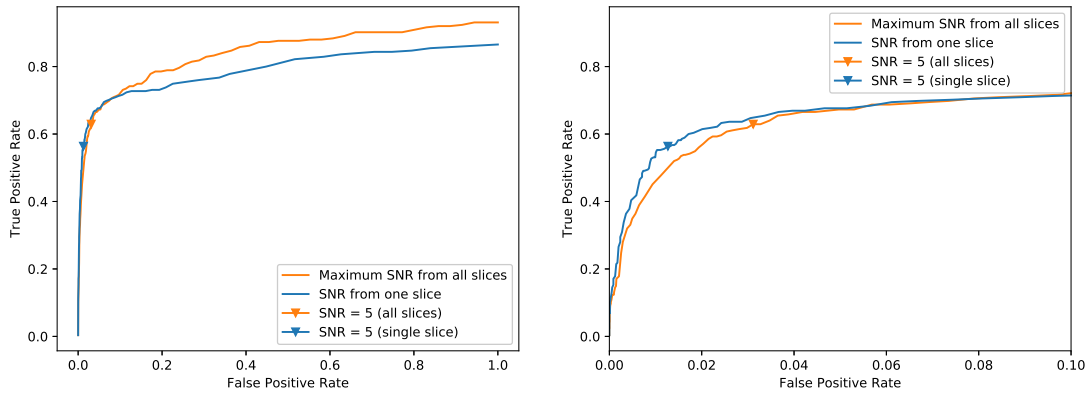

Figure 3.12: ROC curves comparing using a select set of parameters (10 modes kept and a threshold of 5) against picking the highest SNR point from every combination of parameters. The single selection metric captures fewer true positives in the long run, but performs better in the region we are interested in: above 5 SNR.

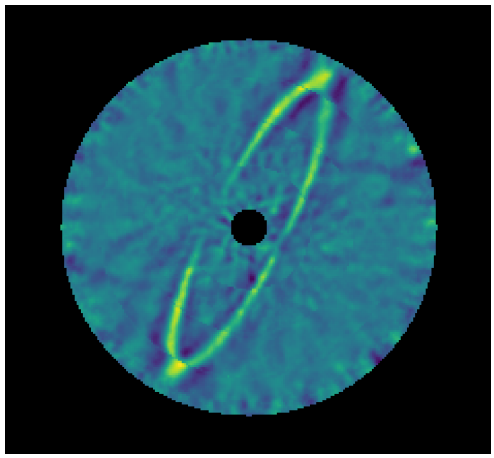

Figure 3.13: The HR 4796a disk is readily visible.

on the disk has the same spatial disparity from image to image in the observation sequence as a point source at that location. The entire disk can be thought of as a collection of point sources. However, the description used for a point source in the forward model assumes the region surrounding the signal drops to the noise floor rather than continuing as an extended source. This likely results in a sub-optimal representation of extended sources, and is not the intended purpose of the CSPFMMF algorithm.

3.5 Results

3.5.1 Receiver Operating Characteristic

In general, an ROC curve is a measure comparing true positive rate to the false positive rate, parameterized by some threshold [4]. In this case, that threshold is an SNR such that every location above that SNR is considered a detection. As the threshold lowers, more and more points are viewed as detections. Some of these are actually planet locations, where many others are simply false positives. The rate for true positives is determined by the number of true detections divided by the number of known planets injected into the data.

Whenever any signal is detected, the surrounding resolution element is masked out from the image and the threshold is lowered, whether or not the signal was a true positive. Thus, the total number of false positives is approximately the number of resolution elements in an image multiplied by the number of images used. The false positive rate is the number of false positives detected divided by the total number of false positives. In general, the closer a curve gets to 100 percent true positives and 0 percent false positives, the better the performance of the tested algorithm. The ROC curves for each of the three injection strengths are shown in Figure 3.14.

As expected, the 5×10^{-5} performs well, as it is an extremely strong signal. Likewise, 5×10^{-7} performs incredibly poorly. The true measure of the efficacy of the algorithm comes when examining the 5×10^{-6}. This result shows a very small false positive rate before an SNR of 5, and a relatively high true positive rate. This shows that CSPFMMF is a very robust algorithm.

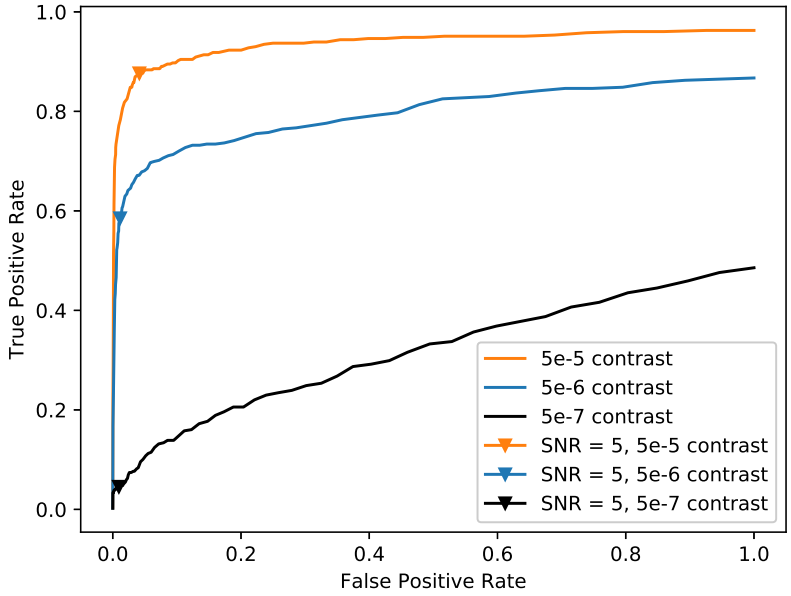

Figure 3.14: Receiver Operating Characteristic curve for all processed data sets.

3.5.2 Separation-Contrast Maps

While the ROC curve is adept at determining the efficacy of the algorithm, a different analysis helps predict what it is that the algorithm might detect. In this section, I develop separation-contrast maps with SNR contours to make those predictions.

Looking at all observations, we can compare the calculated SNR and injected contrast for a target at a given separation. Two of these plots are shown in Figure 3.15.

As expected, the SNR increases when the planet is brighter. From this, we can infer trend lines at that particular separation. This tells us the SNR we can reasonably expect for a given speration and contrast. This trendline is computed using all of the datapoints with a linear least-squares regression.

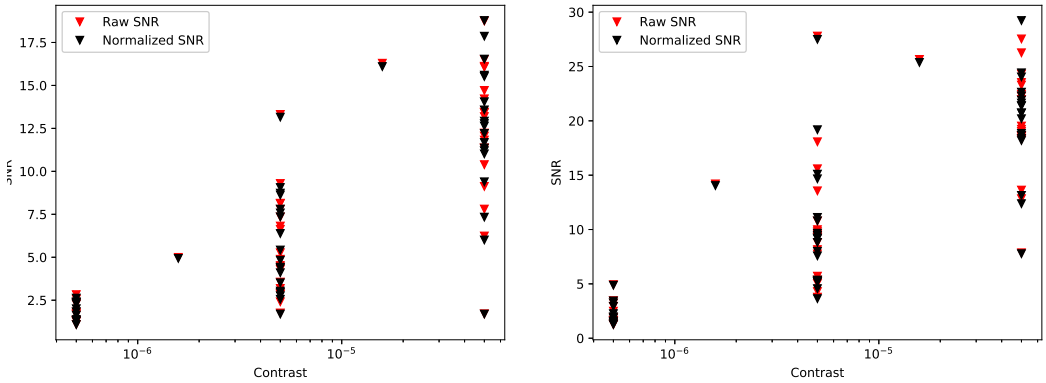

Figure 3.15: A comparison of the SNR and injected contrast for all observations for two given separations.

What is also important to note is that each observation of a target may have a differing number of image; for our purposes, this was between 30 and 45. This corresponds directly with integration time, as each image was taken over the course of about a minute. Given that SNR is generally proportional to the square root of integration time, the datasets were normalized by that amount. Specifically, the SNR for a given datapoint was divided by the square root of the number of images from its original dataset and then multiplied by the square root of the average observation time of all targets in the set.

To determine if this was a valid tactic, the R^2 value of each regression line was computed both for the raw and normalized data. Figure 3.16 shows the two values as they change with separation.

Data so close to the inner working angle is predictably unreliable in both cases. However, the linear trend becomes more and more reliable the farther in separation a target planet is. The variation from a true linear relationship can be explained at least in part by differences in seeing, telescope conditions, and target star properties. Furthermore, the normalized data performs about

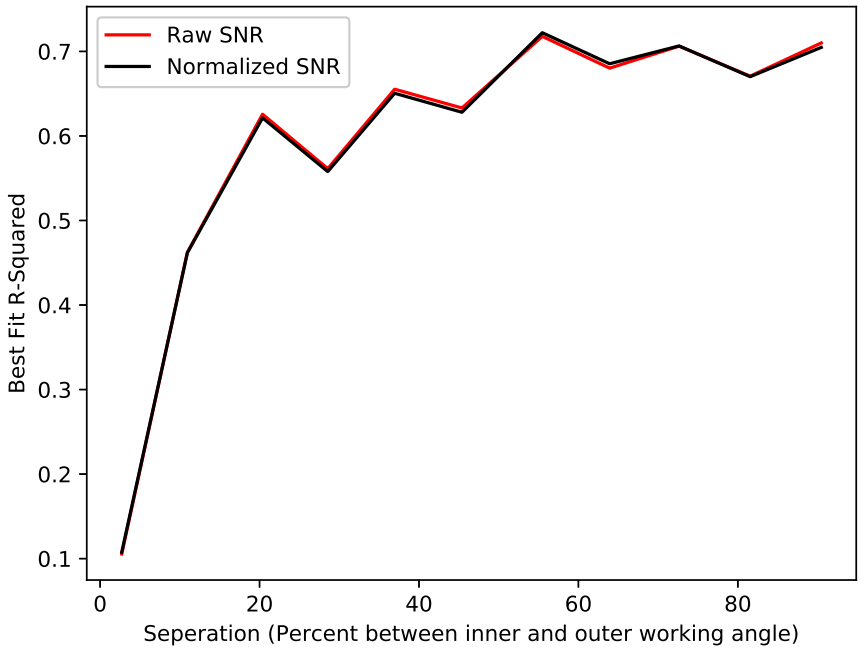

Figure 3.16: Correlation coefficient vs separation.

the same as the raw data within a small margin of error across the entire field of view.

Knowing that this is a valid tactic, we can then use the data to predict what SNR a target will appear at for a given contrast and separation, assuming an average integration time. This is shown in Figure 3.17.

As expected, larger contrasts are required for greater resultant SNRs, and planets are more visible at larger separations.

3.6 Conclusion

In this chapter, I have improved the CSPFMMF algorithm to detect planets from direct imaging data. This was done by parallelizing the code in Python to work with pyKLIP, alongside algorithmic updates like segmented images, thresholds

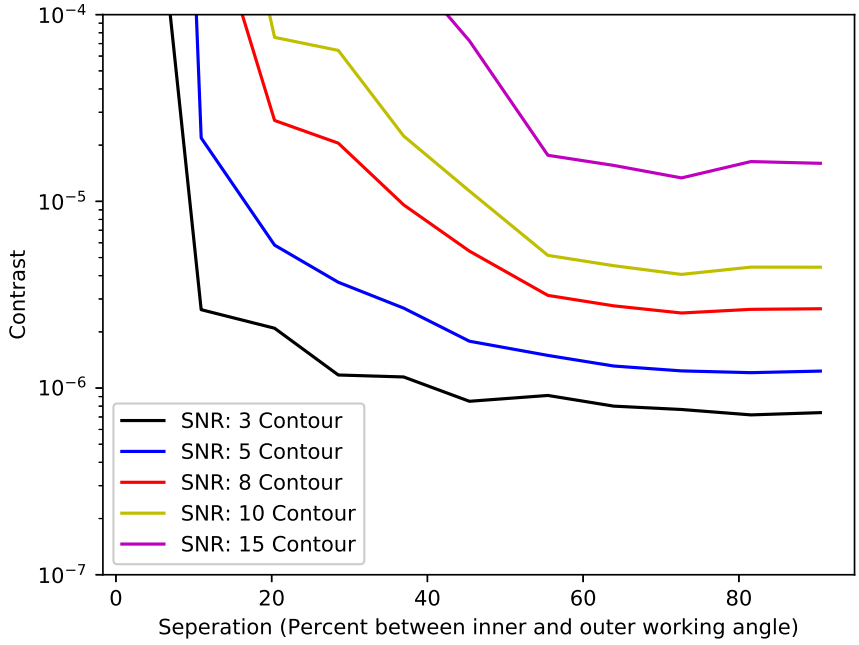

Figure 3.17: Separation vs Contrast map with SNR contours.

on matched-filter templates, and SNR maps. Furthermore, I have shown that selecting one set of optimal parameters will outperform blindly picking the best from a muiltitude of parameter sets. Finally, I have analyzed the performance of the algorithm with ROC curves and separation-contrast maps. These have shown CSPFMMF to be a robust algorithm adept at identifying planets with predictable results.

This code will be made publicly available as part of the pyKLIP distribution [36].

CHAPTER 4

OPTICAL DESIGN OF A LARGE, SEGMENTED SPACE TELESCOPE

As the field of space science is maturing, it requires bigger and bigger telescopes, free from the constraints of Earth and its atmosphere. This allows for observing at a higher resolution, with no atmospheric disturbances, and with longer observing blocks. An orbiting telescope is not limited to night observing, but instead by its specific orbital parameters, which are often more forgiving. By collecting more light, fainter objects and those farther away are more readily observable. These are just a few of the myriad reasons to use large space telescopes. In particular as it relates to finding faint signals like those of exoplanets.

However, large space observatories have their own set of drawbacks. They're expensive, they can't be built easily on site by people, they likely must be delivered in many launch vehicles, and servicing the observatories is very difficult.

In an attempt to solve some of these issues while maintaining the advantages of the large space observatory, we have created a first order approximation of a design reference mission for a telescope of this sort: segmented with hundreds of mirrors and assembled autonomously in space [28]. The size of our telescope in comparison to other space telescopes is shown in Figure 4.1

This chapter will primarily address the optical components of this design. It covers the overview of the mission, the requirements driving the shape of the mirrors, segmentation, and the best method of approximating the desired shape of the mirrors. Finally, we model the proposed system in Zemax OpticStudio,

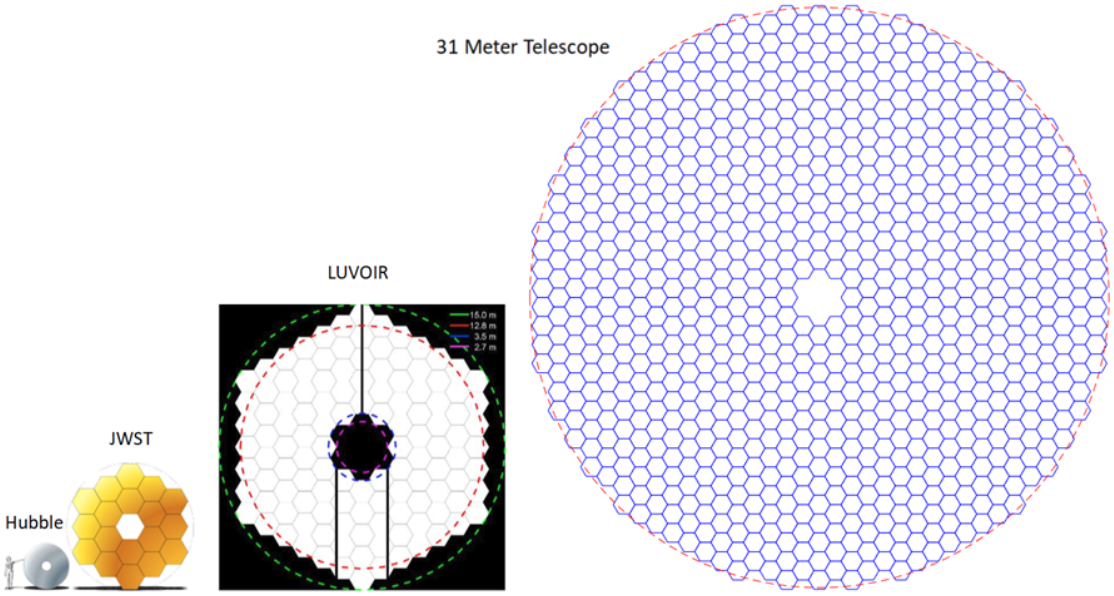

Figure 4.1: To-scale size comparison between the proposed telescope, the Hubble Space Telescope, the James Webb Space Telescope and the Large Ultraviolet/Optical/Infrared surveyor, which is one of the largest aperture mission concept currently in development [6, 16]

.

and use that model for physical optics propagation to measure and examine optical properties like the wavefront and the resultant point spread function. This chapter is primarily based on my previous NIAC work from[28] and [29].

4.1 Mission Concept and Architecture

Everything in this design flows from a core tenet: identical, mass-produced mirrors launch independently and self-assemble in space. Each mirror is a part of an identical spacecraft module. Each module is a self-contained spacecraft, and is propelled by a solar sail for navigation from earth to the L2 Lagrange point. There, the spacecraft will be placed on a Lissajous parking orbit. The quasi-periodic nature of the parking orbit allows for the the modules to get into close proximity to each other fairly regularly, which will lead to docking maneuvers.

These are described in Figure 4.2.

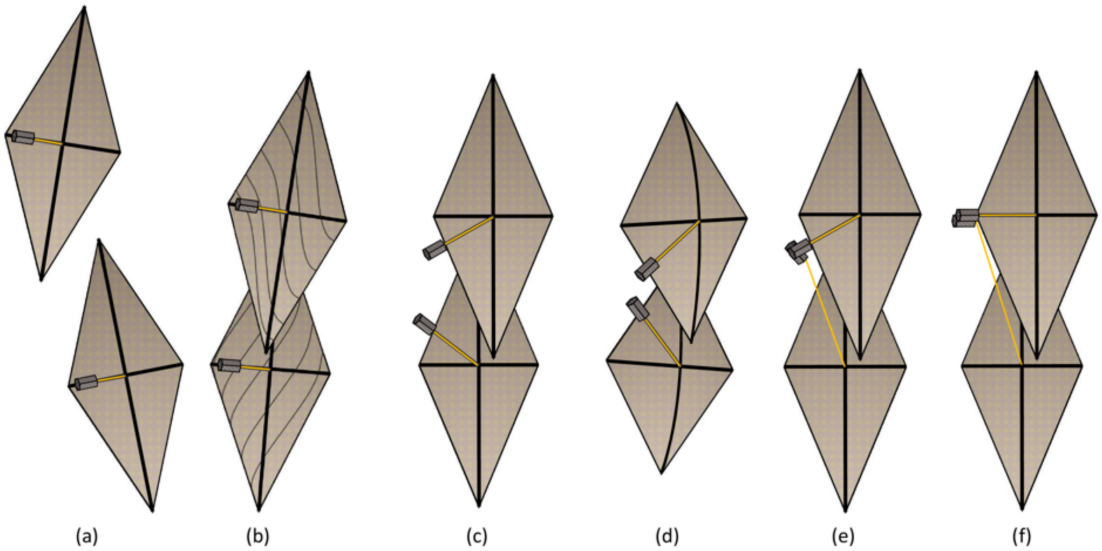

Figure 4.2: Displays docking of independent modules. (a) Spacecraft divert to an intercept trajectory with another module or sub assembly. (b) The sails attach via hook-and-loop material on the sails. (c-d) Reaction wheels trigger a moment of inertia and bring the spacecraft closer together, while the connection on the sails provides support. (e) Controllable electromagnets align the spacecraft properly. (f) The spacecraft interlock with mechanical apparati, and the sails of one spacecraft are released such that they are no longer hard-linked.

The sails will form a conglomerate with the other sails. Together, they act as a sun shade during telescope operation. Each of the modules will continue to join others until a vast primary mirror is formed. Afterwards, the primary mirror assembly will be joined by independent spacecraft: the secondary mirror and the instrument package. The craft is shown in Figure 4.3.

This methodology ensures that we can counteract some of the drawbacks of large space telescopes. The mirrors can be mass produced, increasing efficiency and lowering costs. Furthermore, the small size of a single module means they can be launched either as a set of primary payloads, or as payloads of opportunity, lessening the launch costs. Furthermore, if each mirror is identical, a single point of failure on one module does not spell failure for the entire mission. In-

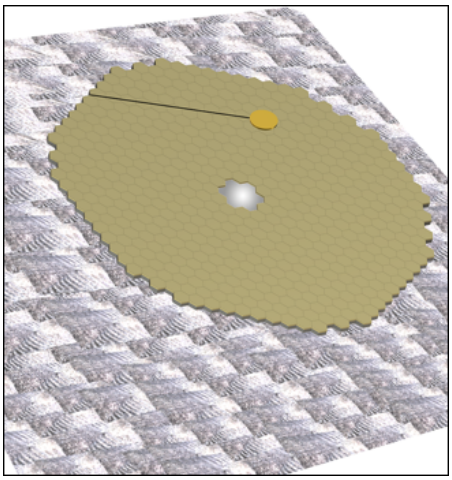

Figure 4.3: The combined telescope, with an assembled primary, a separate secondary, and the sun-shield of solar sails.

dividual modules can be relaunched and replace any faulty ones.

The optical design of the telescope is driven by the new challenges that arise from this concept of operations: autonomous self assembly in space of identical mirror segments on independent spacecraft. We address each of these challenges in turn, and show a complete analysis of the point spread function of the resultant design. The general workflow of this portion of the project occurred in 5 steps:

1. Select telescope parameters
2. Design ideal primary shape
3. Decompose each segment shape into Zernike modes
4. Model the whole system
5. Simulate the system point spread function (PSF)

The last two steps are completed in Zemax OpticStudio, a commercial optical propagation software package.

4.2 Requirements

There are several driving factors that influence the design of the telescope and mirror shape.

- Identical mirrors. Because each mirror may be in any location in the array, they cannot be made to a prescribed shape. This also greatly reduces manufacturing costs by relying on mass production.
- Feasible actuation scheme. Once in place, each mirror must then conform to a desired shape to create the primary mirror. This will require a control scheme with set actuators. However, the actuator design must be reliable and employ known technology. Specifically, we will be re-implementing the actuator design for the James Webb Space Telescope (JWST) [33].
- Static wavefront error < 9.5 nm root-mean-square (RMS) error. This number is adopted from the LUVOIR interim report [6], which states their static wavefront error is ¡38 nm RMS. Given that our mission targets a larger LUVOIR architecture, we adopt this number with a factor of safety of 4.
- Focused point spread function (PSF). This will be validated quantitatively with the Strehl ratio.

4.3 Telescope Design

After a careful review of the designs of both current space telescopes, including Hubble [8] and JWST [20], and plans for giant, ground-based telescopes, such as

the Thirty Meter Telescope [26], a Ritchey-Chrètien Cassegrain design was selected. This particular design, with two hyperbolic mirrors, removes coma and spherical aberrations better than a standard parabolic Cassegrain. These comparisons to similar telescopes and our design goals have lead to the selection of the following parameters:

- Primary Aperture $d = 31m$
- Primary focal length F_P of $f/2$
- Distance from primary to secondary $D = 0.9F_P$
- Total effective focal length F of $f/5.6$

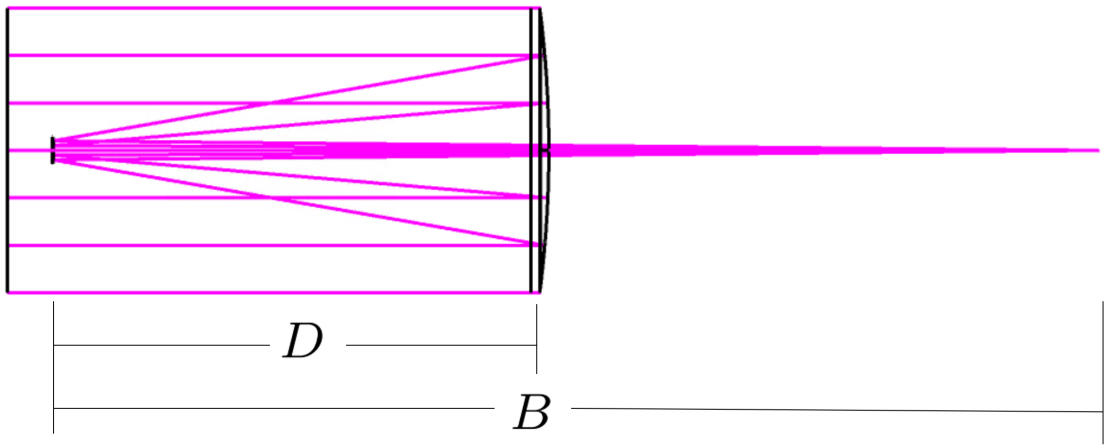

Figure 4.4: Side view of extended focus, with dimension labels.

While the size of the primary mirror is set by the mission concept, we also wish to keep the secondary mirror to within approximately 3 m in diameter, such that it could be manufactured as a monolithic structure. Following the examples of the aforementioned telescopes, which place the secondary mirror at approximately 90% of the primary focal distance, places strict limits on the focal length of our system. These two requirements of secondary size and position translate to a maximum $f/2$ system. In comparison, the TMT telescope is $f/1$.

At exactly 90% of the distance to an $f/2$ focal point, the secondary mirror is 3.06 m in diameter, meeting the sizing goal. The image plane is placed at $2d$ behind the primary mirror. The complete optical path is shown in Figure 4.4.

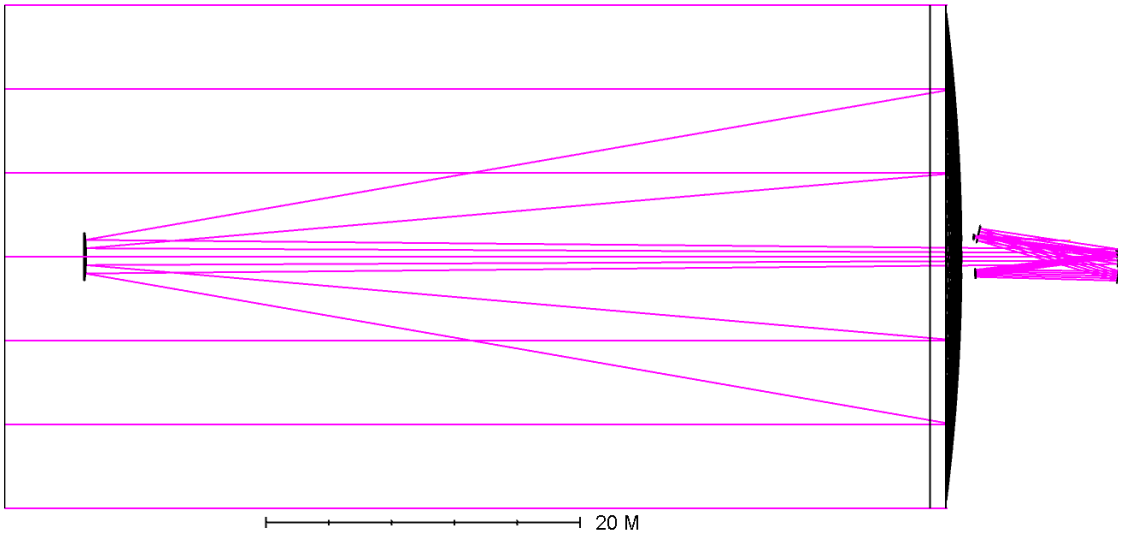

Figure 4.5: A sample ray trace of the full telescope design, including a folded focus behind the primary mirror.

The intent is that the beam path will be folded within the instrument package. However, as the instrument package was not the focus of this phase of research, and the specific folding of this path has no impact on determining the feasibility of constructing the primary mirror in space, no detailed analysis was carried out on the instrument package itself. A sample folded focus, that has not been analyzed in detail nor optimized, is shown in Figure 4.5

Using the above parameters, we can evaluate the equations that determine the shape of the mirrors:

- Primary Radius of Curvature

$$R_1 = \frac{DF}{F - B} \tag{4.1}$$

- Primary Conic Constant

$$K_1 = -1 - \frac{2B}{M^3 D} \tag{4.2}$$

- Secondary Radius of Curvature

$$R_2 = \frac{2DB}{F - B - D} \tag{4.3}$$

- Secondary Conic Constant

$$K_2 = -1 - \frac{2}{(M-1)^3}\left[M\left(2M - 1) + \frac{B}{D}\right)\right] \tag{4.4}$$

- Magnification

$$M = \frac{F - B}{D} \tag{4.5}$$

These equations describe the radius of curvature at the center of each mirror, and the conic constant for each mirror in a perfect system. Evaluating these equations yields the results in Table 4.1.

Table 4.1: Table of values defining the shape of the primary and secondary mirrors

	Radius of Curvature (m)	Conic Constant
Primary	124	-1.000615
Secondary	13.0888888889	-1.241808

The final design produces a primary mirror that is relatively flat, with a large radius of curvature. It also has a nearly parabolic shape, with a conic constant just less than -1. These parameters ensure that segmentation, discussed in Section 4.4, will be a reasonable endeavor. In contrast, the secondary mirror has a much more pronounced curve and is highly hyperbolic, strengthening the case for manufacturing the secondary as a monolithic mirror.

Using these results, we can describe the shape of a given mirror with

$$z = \frac{R + \sqrt{R^2 - (K+1)r^2}}{K+1}. \tag{4.6}$$

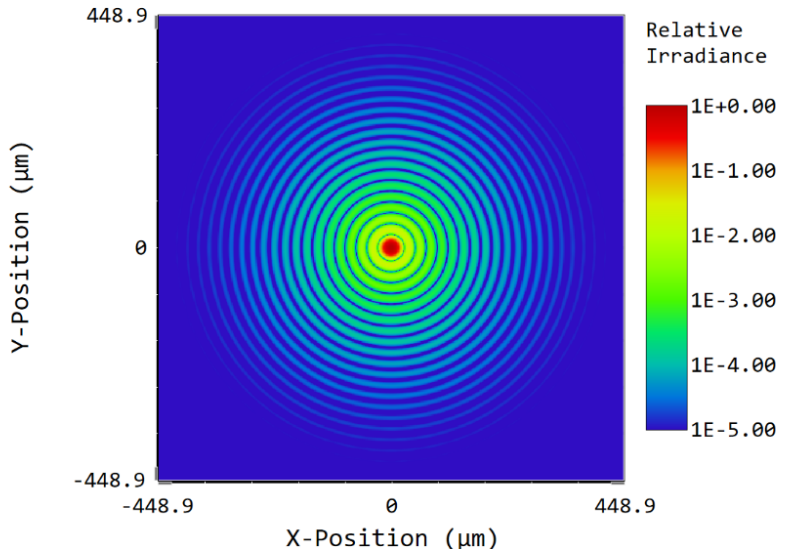

Figure 4.6: Analysis from Zemax of the point spread function of the ideal mirror in log scale. A central, circular region is shown with a uniform value of one, surrounded by concentric rings of diminishing strength, as expected.

where r is the radial distance from the center of each mirror in the xy plane. This design, with all parameters, was modeled and analyzed in Zemax OpticStudio, assuming perfect, monolithic mirrors in both cases. The resultant point spread function from this system, shown in Figure 4.6, verifies that the system performs ideally.

4.4 Segmentation

Next, the ideal primary shape was decomposed into hexagonal segments sized 1-m flat-to-flat. This sizing lies within the bounds of the design phase space, which was predetermined by other portions of the mission. This was selected as the point design for optical analysis.

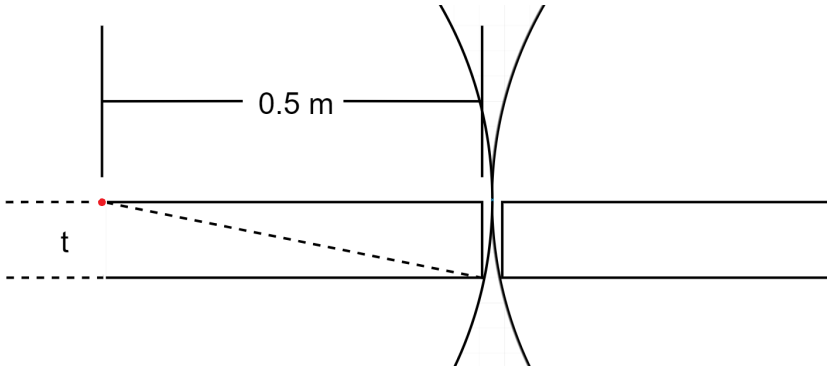

Figure 4.7: A side view of two adjacent mirrors. They are spaced sufficiently far such that the arcs drawn by the farthest point in worst-case scenario tilt do not intercept each other. This determines the gap spacing.

Gap Sizes

The first step is to determine the spacing of the mirrors, placing them as close as possible to each other while guaranteeing that the segments do not impact one another when tip/tilt corrections are applied to generate the final mirror shape.

The gap between the mirror segments was calculated purely from the mirror geometry, assuming that each mirror had the worst case tip/tilt applied relative to its neighbor. As seen in the Figure 4.7, the footprint for each mirror's possible position is carved out by the arc with a radius from the center of the surface of a mirror (the point of rotation) to the bottom corner of the mirror. Thus, the worst case scenario for gap size necessary would be described by

$$0.5^2 + t^2 = \left(0.5 + \frac{g}{2}\right)^2 \tag{4.7}$$

where g represents the full gap between the two mirrors. The thickness of the mirror, t was calculated by scaling the thickness of the James Webb mirrors based on the mirror diameter of the two designs. This process results in a required gap size of 6 mm.

Center Gap and Edges

Ideally, the primary mirror would have a perfectly circular aperture. However, because of the segmentation and the telescope design, it would have two significant deviations from the ideal circular shape: near the center and at the edge of the aperture.

Because of the Cassegrain design, light must be allowed to pass through the primary to reach the image plane. If all mirrors are the same size, the only solution is to eliminate rings of mirrors near the center. We have calculated the most extreme point where the beam would cross a complete primary, and exclude the segments it intercepts and all mirrors interior.

The edge of the aperture as well will no longer resemble a perfect circle. Along the hexagonal spokes of the primary, targeting a diameter of 31 m is accomplished by placing 15 segments on all sides of a central hex (had the central hex not been removed). However, to more closely approximate a circular aperture, more segments are included beyond the hexagonal pattern in between these spokes. We continue to add more layers beyond 15, eliminating mirrors as needed. The criteria we used was to continue adding segments as long as at least approximately 2/3 of the area of each hexagon is inside the 31-m diameter, leading to the final design shown in Figure 4.8.

4.4.1 Mirror Modal Decomposition

In most segmented telescopes, each mirror is polished to a specified shape corresponding to its given location in the array. However, our design requires that

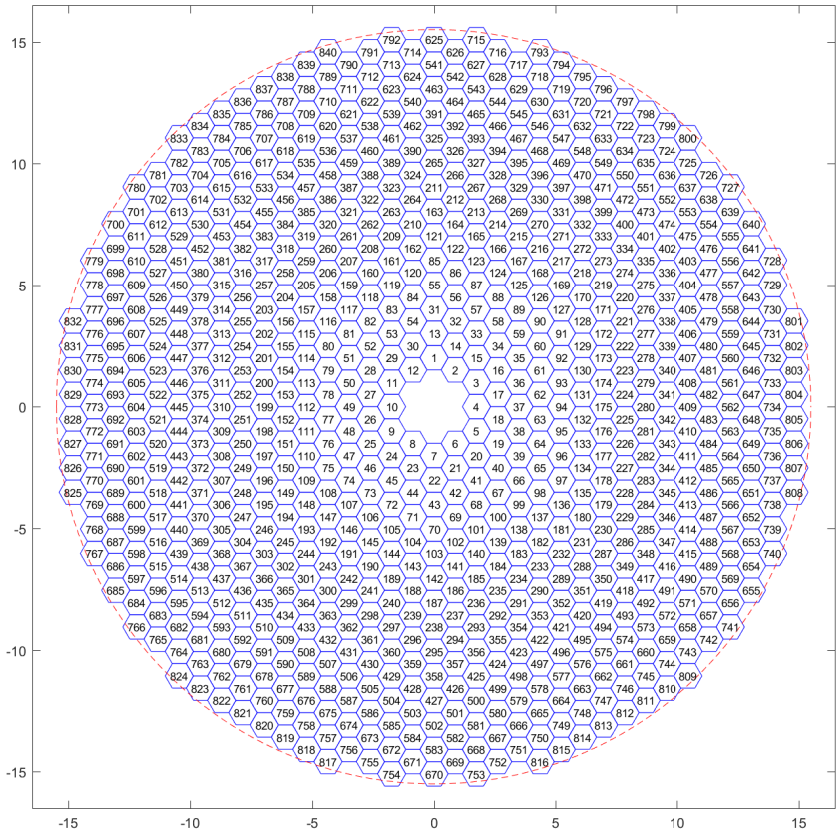

Figure 4.8: A scheme for segmentation and placement of individual mirrors on the primary assembly. Seven central mirror segments are missing (one exactly center and the surrounding ring of 6). A dotted red line show a 31 meter diameter circle. Each mirror is numbered in a clockwise spiral pattern, ending at mirror 840. The size scale is in meters

the segments are randomly placed into the array via docking on the Lissajous orbit. We cannot know beforehand what location each mirror will ultimately be placed in. Thus, each mirror must initially be identical prior to self-assembly. To transform each to the desired shape, we must first identify the differences between that ideal shape, and the stock, flat shape.

We accomplish this by decomposing each mirror into Zernike polynomials. These polynomials are a standard set of two-dimensional functions that together form an orthogonal basis on the unit circle. Generally, the Zernike polynomials are defined by

$$
\begin{aligned}
Z_n^m(\rho,\theta) &= R_n^m(\rho)cos(m\theta) \\
Z_n^{-m}(\rho,\theta) &= R_n^m(\rho)sin(m\theta) \\
R_n^m &= \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k(n-k)!}{k!\left(\frac{n+m}{2}-k\right)!\left(\frac{n-m}{2}-k\right)!}\rho^{n-2k}
\end{aligned}
\tag{4.8}
$$

where ρ, θ are the polar coordinates defined on the unit circle, and n, m are non-negative integers. Each polynomial is normalized such that

$$
\int_0^{2\pi}\int_0^1 Z_n^{m2}\rho d\rho d\theta = \pi. \tag{4.9}
$$

We modified this slightly, to conform to the hexagonal mirrors: each point within the domain but outside the hexagonal surface is set to zero. The polynomials themselves are shown in their original, circular domains in Figure 4.9.

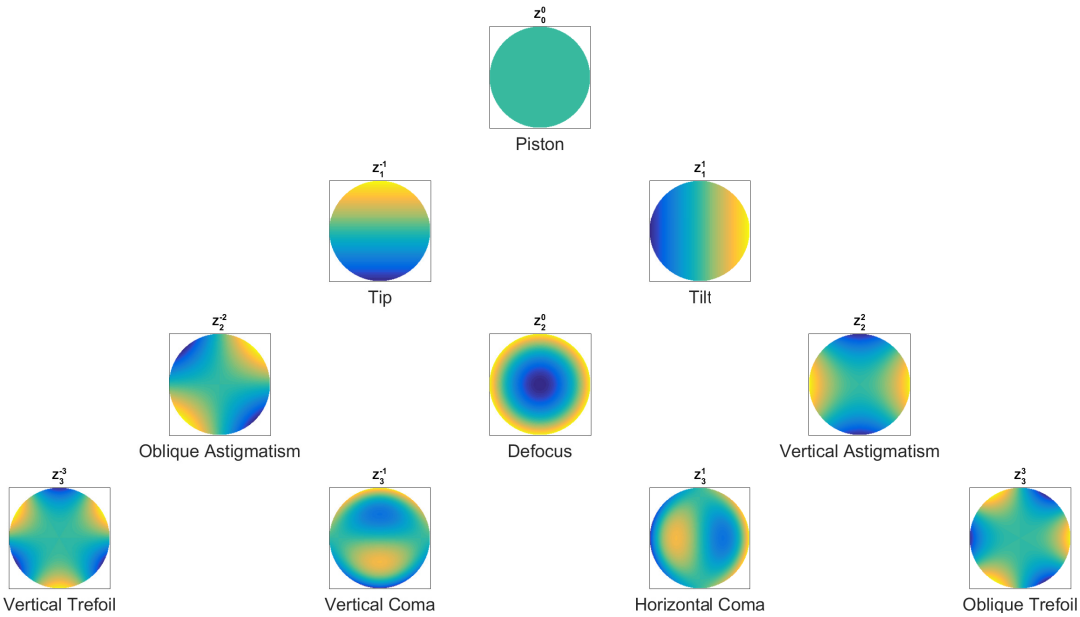

Figure 4.9: The first 10 Zernike modes. Each is a two dimensional polynomial in the unit circle, and orthogonal to the others.

The ideal shape of each mirror segment was decomposed into Zernike modes by least-squares fitting to the polynomial functions. The weights of the

least-squares fit corresponds to the intensity of each mode in the desired shape. The weights of the first five modes for every mirror are shown in Figure 4.10.

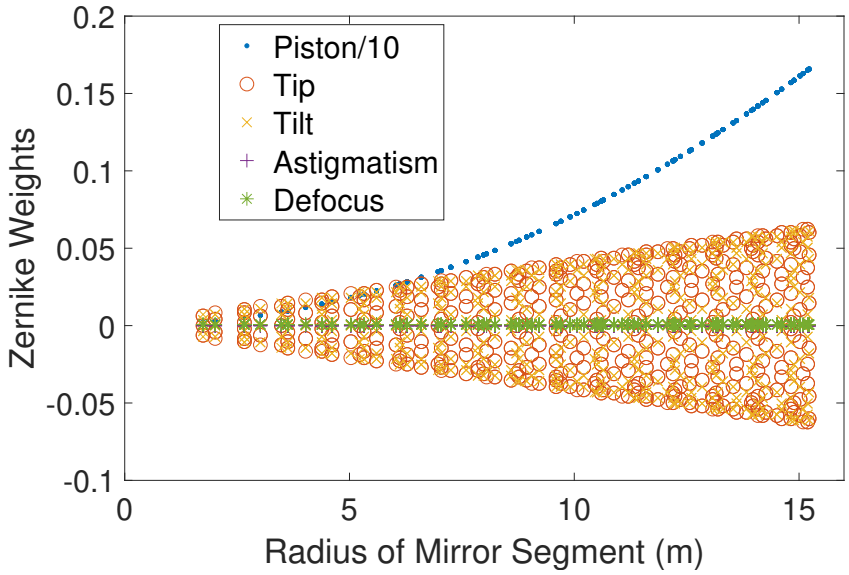

Figure 4.10: The weights of the Zernike modes for each mirror segment. As the distance of the mirror from the center of the assembly increases, the piston mode, maximum angle (tip and tilt), and defocus increase. These modes have the highest impact on the mirror shape.

Most prominent is the piston mode (the primary driver of the hyperbolic shape of the mirror). The next most prominent modes are the tip and tilt of each mirror. Each of these first three modes is easy to correct for by using the established technology of a Stewart Platform as part of each spacecraft module, providing six degrees of freedom. The next highest mode is the defocus mode, corresponding to a perfectly symmetric spherical curvature error. This is similarly easy to correct for, using the same technology as tested on the JWST: a strut to induce curvature on a mirror segment [33]. Beyond these major modes, other Zernikes are less readily corrected for. Likely, this would require a separate array of many different actuators on the back-plate of the mirror, and would be inefficient for the differing needs of the mirrors across the primary.

Thus, based on the major modes of the decomposition, and the corresponding fact that these modes have already-tested and proven technology, we limit our reconstruction to the first four Zernike modes: piston, tip, tilt, and defocus. This model will be referred to as the "fully reconstructed model" henceforth. Furthermore, we can examine the defocus mode more carefully - the associated radius of curvature change is marginal across all mirrors, as shown in Figure 4.11. This similarity allows for another method to ease the shape requirements.

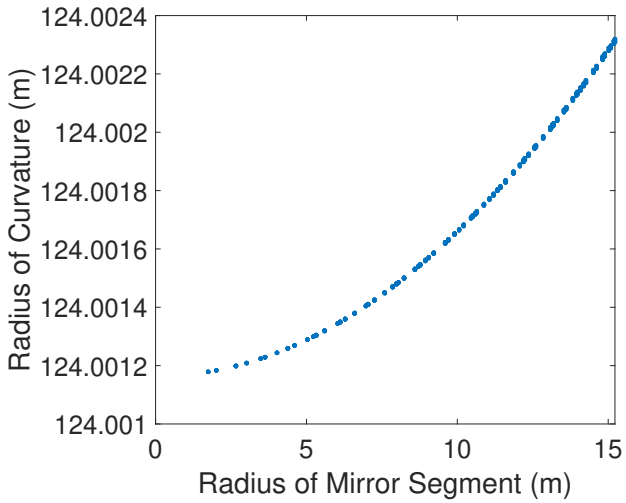

Figure 4.11: The radius of curvature for each mirror segment generated from its defocus mode. All are slightly larger than 124 m, and the difference between the largest and smallest curvatures is approximately 0.0011 m.

All of the mirrors can be manufactured to the mean curvature of each of the mirrors, as opposed to a flat surface. Using all mirrors set to a mean curvature does not dramatically reduce the quality of the optical system, as will be shown through the analysis in §4.5. The model that specified only a mean defocus to every mirror, in addition to piston, tip, and tilt, will be referred to as the "mean defocus model."

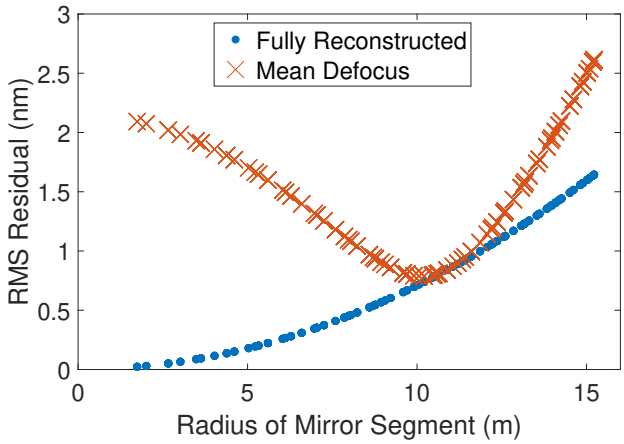

Figure 4.12: RMS errors for both the fully reconstructed model and the mean defocus model. Both are within 3 nm RMS, meeting requirements.

4.5 Analysis

4.5.1 Mirror RMS Error

The ideal model of every mirror from the hyperbolic equations of a Ritchey-Chrétian telescope is known, as is the reconstructed shape of each mirror using exclusively piston, tip, tilt, and a specified mean defocus. In comparing the two, we can determine the error associated with both the fully reconstructed model and the mean defocus model. The RMS of each mirror for each model is shown in Figure 4.12. Figure 4.13 also shows specific examples of the residuals for two mirrors, taken from the fully reconstructed model. In particular, these demonstrate that the remaining error corresponds to the next-order modes, astigmatism, which we expect. We have explicitly not corrected for terms beyond astigmatism. The more central mirror has a much lower residual than the outer mirror, also meeting expectations.

At a maximum RMS error of about 2.5 nm, both of the models meet the

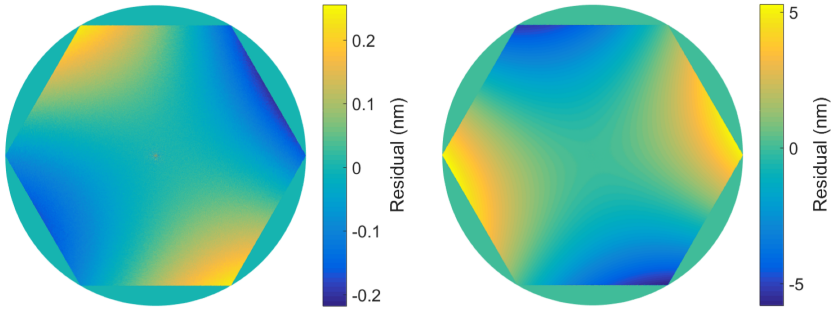

Figure 4.13: Left: Inner mirror, with low residual errors. Right: Outer mirror, with more significant errors. Both mirrors have errors that resemble astigmatism modes.

requirement set in §4.2. Treating the entire primary as one mirror yields an RMS of 0.971 nm. This is less than 10% of the error budget for the whole assembly.

4.5.2 Zemax OpticStudio Model

The subsequent analysis of the optical design was completed in Zemax Opticstudio. An option in this software allows specifications of mirror shapes through Zernike weights, fitting our needs perfectly. The model shown in Figure 4.5 was generated in this software. Importantly, it can calculate diffraction effects, which is particularly relevant considering the gaps in the primary mirror (illustrated in Figure 4.14).

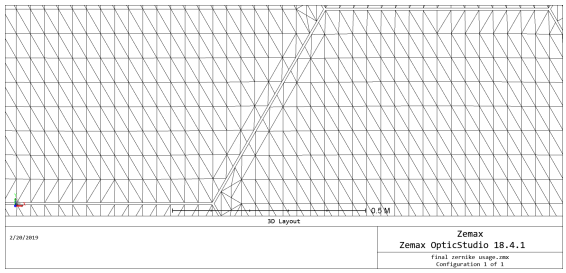

Figure 4.14: A zoomed-in view of the Zemax OpticStudio model of the primary mirror assembly showing the 6 mm gaps between each mirror segment.

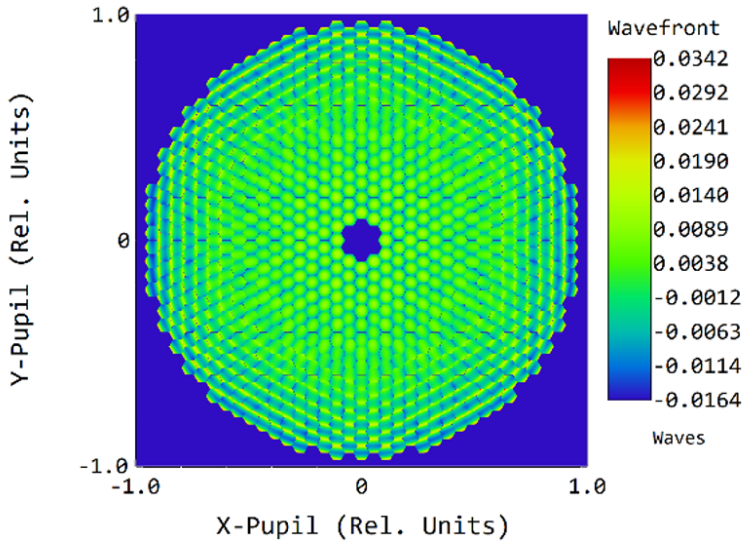

Figure 4.15: The wavefront error as determined by the Zemax OpticStudio Model of the mean defocus model. The wavelength used is 550 nm.

Additionally, by examining the wavefront directly in the Zemax model, we can confirm our calculations for the RMS of the telescope. The wavefront error on the image plane, shown in Figure 4.15, verifies our previous results.

An additional analysis is necessary to determine and quantify the effects of the gaps and segmentation compared to effects from reconstruction of the mirror shape. To do so, we created two models: one with the ideal primary mirror shape, with the hexagonal geometry and gaps overlaid, and another model with each segment reconstructed based entirely on the first four Zernike weights for that mirror.

4.5.3 PSF and Spot Diagram

The Huygens Point Spread Function, which accounts for diffraction in non-sequential components, was calculated in OpticStudio for both the ideal and

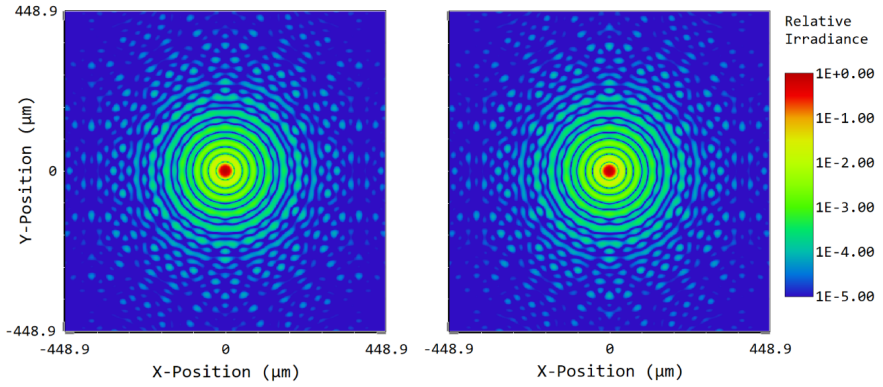

Figure 4.16: Left: PSF from the model with perfect mirror shape and gaps. Right: PSF from model with both reconstructed mirror shapes and gaps. There is strong agreement between the two models.

reconstructed models, as shown in Figure 4.16, demonstrating excellent agreement. The ideal mirror, with gaps, represents the best achievable PSF given this setup. As the surface figure improves on each segment, the fully reconstructed model will approach greater similarity with the ideal model. In both models, the core PSF is very clearly visible, as is the expected hexagonal structure as a result of segmentation.

In addition to a qualitative, visual analysis, we have computed the Strehl ratio of the reconstructed PSF. This ratio is calculated via comparing the peak intensity of the reconstructed PSF compared to the peak intensity of the completely non-diffracted, ideal PSF (previously shown in Figure 4.6). The computed Strehl ratio was 0.9986, which is exceedingly high. We expect this number to decrease with both manufacturing and alignment errors, which have not been analyzed in the scope of this paper.

We additionally sought to characterize off-axis point sources, and so multiple fields were added about 10 resolution elements away from the center of the image. The resulting spot diagram of all field points is shown in Figure 4.17.

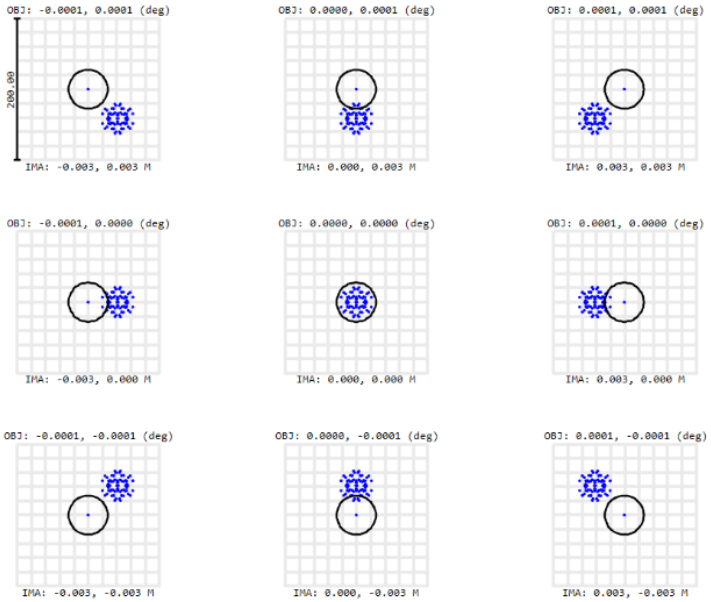

Figure 4.17: Spot diagram of the true, segmented primary model. This includes 1 axial field, and 8 other fields offset by approximately 10 $\frac{\lambda}{d}$. The black circle in each field is the size of one resolution element. In each case, the diffraction of the spot is approximately the size of the circle. Off-axis fields have a slight displacement towards the center of the image plane compared to their true location.

These results indicate that simply using a Stewart platform for piston, tip, and tilt for each mirror, coupled with a mirror-ground curvature and a curvature strut, will sufficiently meet the needs of the optical design for this telescope in ideal situations. These ideal situations include neither manufacturing nor alignment error. Actuators and active control will still be necessary for distortions in the mirrors and to correct for mid-spatial frequency wavefront errors.

4.6 Conclusion

In order to conduct more and more accurate space science, the need for large space observatories is paramount. However, accomplishing such a task brings a new set of challenges. We propose an architecture for a self-assembling, modular space telescope. In this design, each of the mirrors must be manufactured identically, creating a unique optical design problem.

In addressing this issue, and designing the telescope, we initially sought to compare to similar, modern designs: TMT, Hubble, LUVOIR, etc. In doing so, we selected to analyze a 31-m $f/5.6$ Ritchey-Chrétian Cassegrain telescope. To achieve the desired shape for every mirror in the primary array, each mirror was decomposed into Zernike modes. The first four modes correspond directly to the actuator design for the mirrors of the JWST. Thus we limited the reconstruction to piston, tip, tilt, and defocus.

Modeling this system in Zemax Opticstudio allowed us to conduct an analysis on the performance of the telescope. This showed strong agreement between the PSF of this model and the ideal system, slight off-axis abberations, very low RMS errors on each mirror and the telescope as a whole, and a high Strehl ratio. This indicates that within the constraints of this model (perfect alignment, perfect manufacturing, no specified instrument package, and no need for dynamic wavefront control), the optical design of such a telescope is feasible. Thus, we believe this design and approach of such an observatory warrants further exploration.

CHAPTER 5

DETECTING STATE VARIABLES WITH AN IMAGE MEASUREMENT FUNCTION

Detecting exoplanets not only requires great telescopes and advanced data reduction algorithms, but also sophisticated methodology for actually collecting data. Instead of using image processing as an end of itself (**Chapter 2** and **Chapter 3**) or as a tool for design (**Chapter 4**), this chapter examines the filtering component of signal detection, by way of a portion of a currently unpublished project. Here, image processing is used as an essential intermediate step for a novel filter.

5.1 Overview

5.1.1 Project Goals

The objective of the larger project is to create a new, autonomous on-orbit calibration scheme for a constellation of imaging satellites. Any given primary satellite will use its own imagery for cross-calibration with other satellites, rather than dedicated measurements. This will enable continuous calibration and measurement simultaneously.

This is accomplished by implementing a filter whose state is a full descriptor of satellite dynamics and the camera state. The satellites will all share feature sets from their image and state estimates. The given state is

$$\mathbf{x} = [\mathbf{q}\ \mathbf{r}\ \mathbf{c}]^{\mathbf{T}}, \tag{5.1}$$

where **q** is the set of quaternions relating the orientation from the camera frame to the Earth Centered Inertial Frame, **r** is the position vector of the camera relative to the Earth center of mass, and **c** is the camera model. The camera model, which we are interested in for this chapter, can be expanded to

$$\mathbf{c} = [f\ c_1\ c_2\ c_3]^T. \tag{5.2}$$

This is the vector of focus and camera distortion parameters. These, along with the rest of the state, will be used to develop the imaging measurement function. The outputs of this function are SIFT keypoint information, discussed in detail in §5.2.

5.1.2 Camera Parameters

For testing the measurement function, several parameters for a camera model were necessary to define. These are summarized in Table 5.1. These values either came from other portions of the project, or were modelled after the Dove spacecraft [25].

Note that the detector size in pixels and ρ are both significantly smaller than expected, to help with computational purposes for this early demonstration.

	Symbol	Value
Distance from camera aperture to observation target	u	200 km
Camera focal length	f	50 cm (nominal)
Distance from mirror plane to image plane	s	50 cm
Camera aperture	A	0.55 m
Detector size	-	200×200
Pixel-physical units conversion	ρ	1000 pixel/m

Table 5.1: Camera parameters used for examples.

5.1.3 Data Sources

The ultimate goal is to take a ground-truth state of the world and distort it through the camera parameters defined by the state. The difficulty lies in determining a ground-truth model for each image. In the final use case, the proposal is that the constellation of satellites constantly updates and modifies a consensus version of what each image 'originally' looks like.

For early testing purposes, we are using data from LANDSAT 8 Analysis Ready Data [5] as a ground-truth model for input into the image measurement function. Each image conveniently comes with GPS information for each corner of the image. A sample image from this dataset, to be used and manipulated through testing, is shown in Figure 5.1.

5.2 SIFT

The Scale-Invariant Feature Transform (SIFT) algorithm [10] is used to identify specific features (keypoints) in digital images. As the name suggests, the identification of features ideally remains constant. The prepackaged software [9] provides a list of each keypoint with location information and a descriptor. The

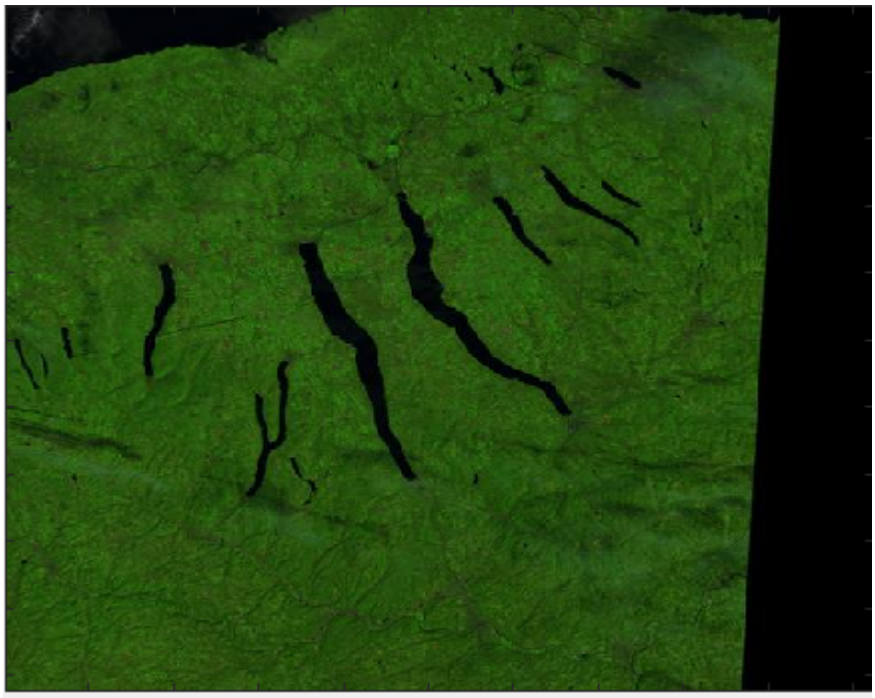

Figure 5.1: Sample image from LANDSAT 8 Analysis Ready Data, situated above the New York Finger Lakes.

location information consists of 4 parts: row, column, scale, and orientation. The descriptor is a 128-long unit vector. The measure of similarity in descriptors is the tool used to match keypoints from differing images despite different location data. A sample set of images with matched keypoints is shown in Figure 5.2.

5.3 Camera Model

5.3.1 Defocus

The change in an image due to defocus can reasonably be modelled with a Gaussian blur kernel [12]. Furthermore, the standard deviation of the Gaussian can

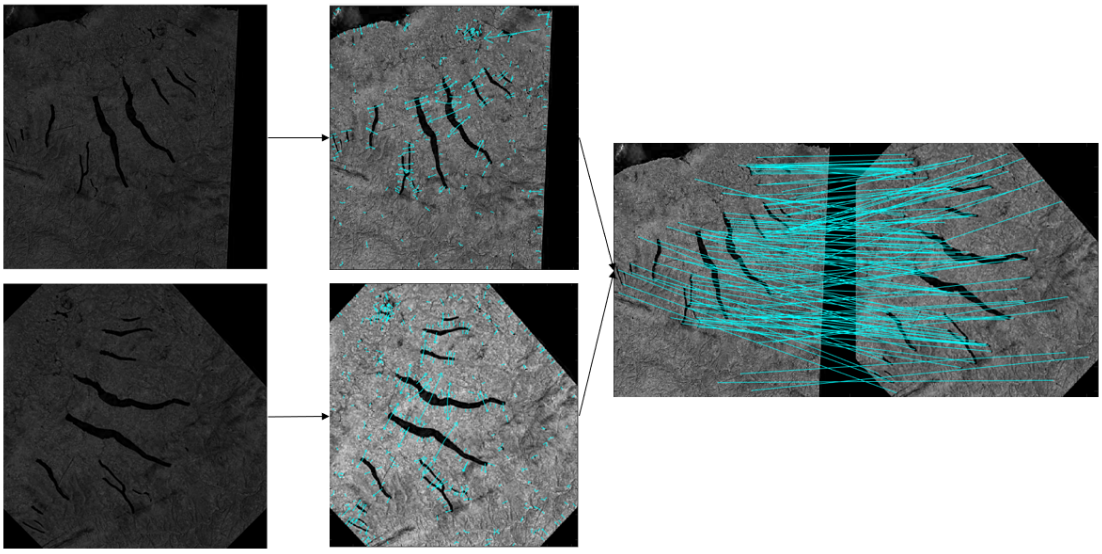

Figure 5.2: Example of detecting and matching SIFT points from two images of the same scene. The bottom row is a rotated and scaled version of the top row.

be described as

$$\sigma_G = \frac{\sigma}{\sqrt{2}}, \tag{5.3}$$

where

$$\sigma = \rho \frac{fs}{2N} \left(\frac{1}{f} - \frac{1}{u} - \frac{1}{s} \right). \tag{5.4}$$

Here, N is the f-number, described by $\frac{f}{A}$. An example of this defocus model using a portion of the a the sample LANDSAT image is shown in Figure 5.3. In this example, the focal length is offset from the image plane by 2.5mm.

5.3.2 Distortion Model

For initial feasibility, we are using a symmetric radial distortion model for the camera [34]:

$$r_2 = r_1 \left(1 + r_1 c_1 + r_1 + c_2 {r_1}^2 + \dots \right) \tag{5.5}$$

We are using three coefficients in our model. As an example of this model,

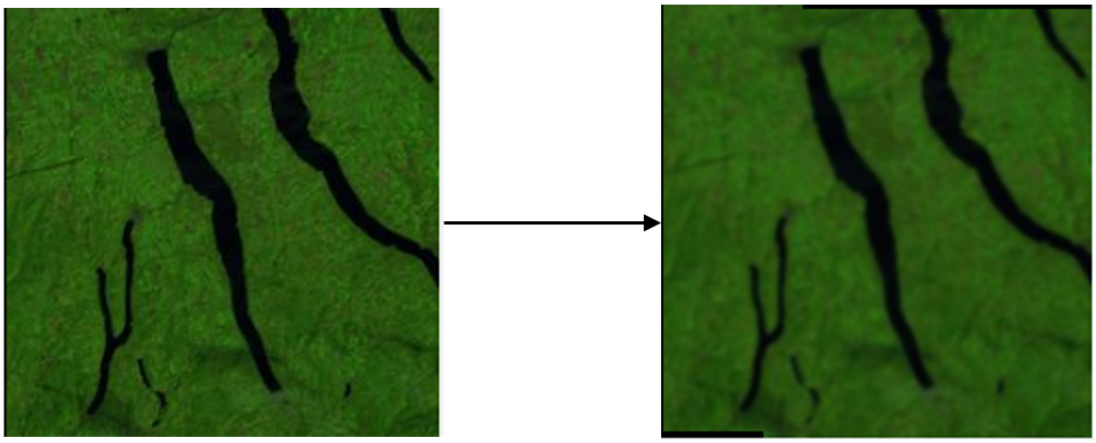

Figure 5.3: Left: Sample Image. Right: Gaussian blurred sample image based on an offset focal length of 2.5 mm.

an undistorted image is shown in Figure 5.4, and different types of possible distortions are shown in Figure 5.5.

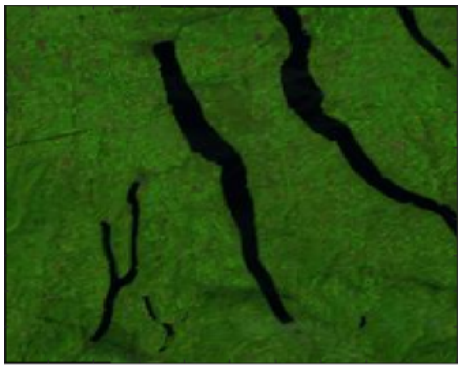

Figure 5.4: Undistorted sample image.

.

Depending on the signs of the coefficients (all positive, negative, or mixed), the resultant distortion can take different forms. The pincushion distortion present in the leftmost image is difficult to see, but most evident in the corners,

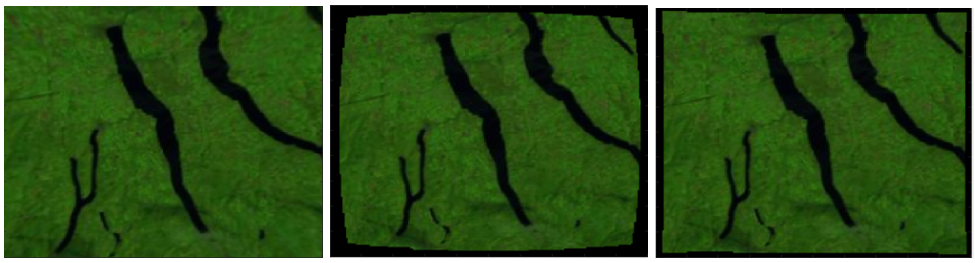

Figure 5.5: Demonstration of the variable types of symmetric radial distortions possible with this model. Left: Pincushion, Middle: Barrel, Right: Mustache

where rivers have been displaced from their original locations.

5.4 Measurement Function

The measurement function uses an input of a ground-truth model of the world (LANDSAT images in testing) and the state variables to predict what the detected imaged will look like. The steps are:

1. Input original image
2. Find boundaries of camera FOV
3. Homogeneous coordinate transform to square detector pixels
4. Apply gaussian blurring
5. Appply camera distortion model
6. Identify SIFT points

The portion not discussed thus far is the homogeneous transform of coordinates to the new image (going from step 2 to step 3). This is completed by beginning with four points $p_1 - p_4$ as the corners of the camera field of view in terms of image coordinates from the original image. The goal is to transform

these locations to a new grid that is the dimension of the detector ($npix \times npix$) such that p_1 maps to $(1,1)$, p_2 maps to $(npix, 1)$, p_3 maps to $(npix, npix)$, and p_4 maps to $(1, npix)$. To do so, let

$$A^* = \begin{bmatrix} p_{1x} & p_{2x} & p_{3x} \\ p_{1y} & p_{2y} & p_{3y} \\ 1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} p_{4x} \\ p_{4y} \\ 1 \end{bmatrix} \tag{5.6}$$

$$A = \begin{bmatrix} A_1^* p_{1x} & A_2^* p_{2x} & A_3^* p_{3x} \\ A_1^* p_{1y} & A_2^* p_{2y} & A_3^* p_{3y} \\ A_1^* & A_2^* & A_3^* \end{bmatrix}. \tag{5.7}$$

Similarly,

$$B^* = \begin{bmatrix} 1 & npix & npix \\ 1 & 1 & npix \\ 1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ npix \\ 1 \end{bmatrix} \tag{5.8}$$

$$B = \begin{bmatrix} B_1^* & B_2^* npix & B_3^* npix \\ B_1^* & B_2^* & B_3^* npix \\ B_1^* & B_2^* & B_3^* \end{bmatrix} \tag{5.9}$$

Let $C = BA^{-1}$. Now, any coordinates $(x, y, 1)$ can be mapped to the new position in the destination image (x', y', z') via

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = C \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{5.10}$$

The final destination in the new image (x'', y'') is normalized by

$$x'' = \frac{x'}{z'} \tag{5.11}$$

$$y'' = \frac{y'}{z'}. \tag{5.12}$$

The new data on the destination image are fit to a grid with a bicubic interpolant. An example of this step, alongside each of the other steps in the entire measurement function process, is shown in Figure 5.6.

5.5 Non-Gaussianity of Measurements

The larger process necessitates the ability to use sigma points in the filter to propagate forward that state information and its uncertainty. These sigma points are determined by the noise distribution of the resultant measurement functions. Thus, it is essential to determine if the distributions of noise on the image measurement function is Gaussian.

This was completed by calculating the entire measurement function results (locations and keypoint descriptors) for a single image 500 times with Gaussian noise on the input state vector. The results for one keypoint, circled in red in Figure 5.7, were measured and used for test of Gaussianity. The resultant distributions are shown in Figure 5.8 and Figure 5.9.

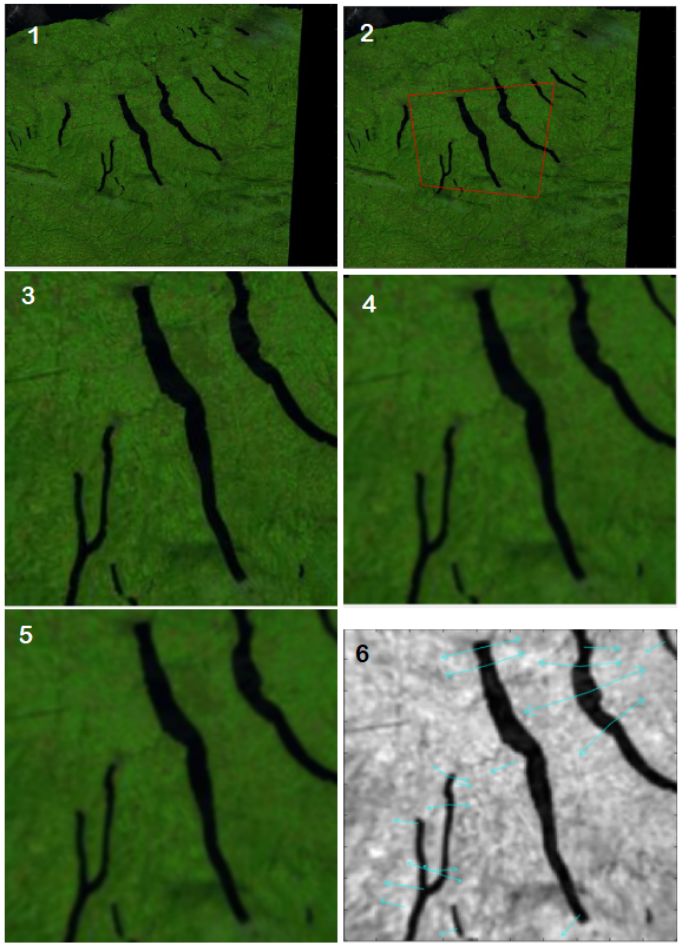

Figure 5.6: Each step in the process of computing the measurement function.

In each graph, the distributions are decidedly non-Gaussian. Of particular interest is the distribution of the similarity between keypoint descriptors. In each result, the descriptor was compared via dot product to the undisturbed result. Because each descriptor is a unit vector in length, this distribution has a maximum value of 1, and thus is inherently skewed to the left. The values describing the first four moments of each distribution are shown in Table 5.2

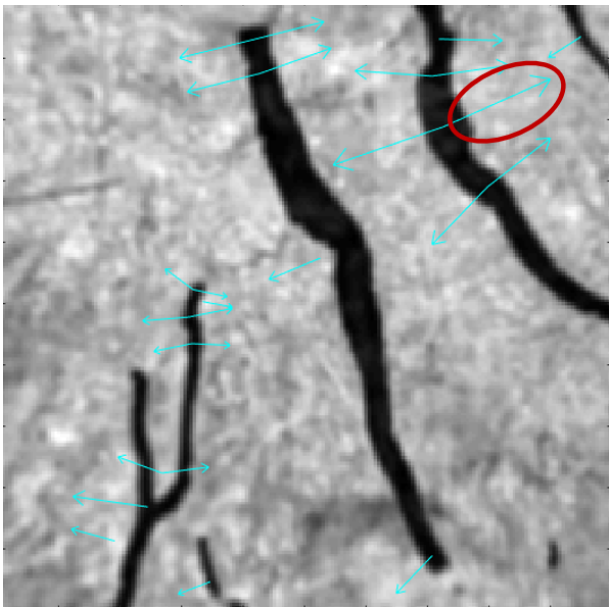

Figure 5.7: Highlighting the particular sift point used for comparison.

	Mean	Standard Deviation	Skewness	Kurtosis
Row Location	42.2	0.73	-0.007	2.416
Column Location	145.1	0.69	-0.048	2.50
Orientation	0.418	0.005	-.506	2.80
Scale	6.39	0.05	-0.041	2.80
Descriptor Similarity	0.994	3.5×10^{-4}	-.094	3.67

Table 5.2: Values of the first four moments of the distributions for a single keypoint.

5.6 Conclusions

In this chapter, we have shown how image processing using modern methods can be vital as an intermediate step towards a novel type of filter. This was done with modern satellite data, distorted with a specific model for the camera, and processed using a recent computer-vision algorithm to identify keypoints in the images. Furthermore, we have shown that the resultant measurement function data is non-Gaussian. This indicates that for a filter to be more successful, it must employ a 4-moment model of uncertainty and noise.

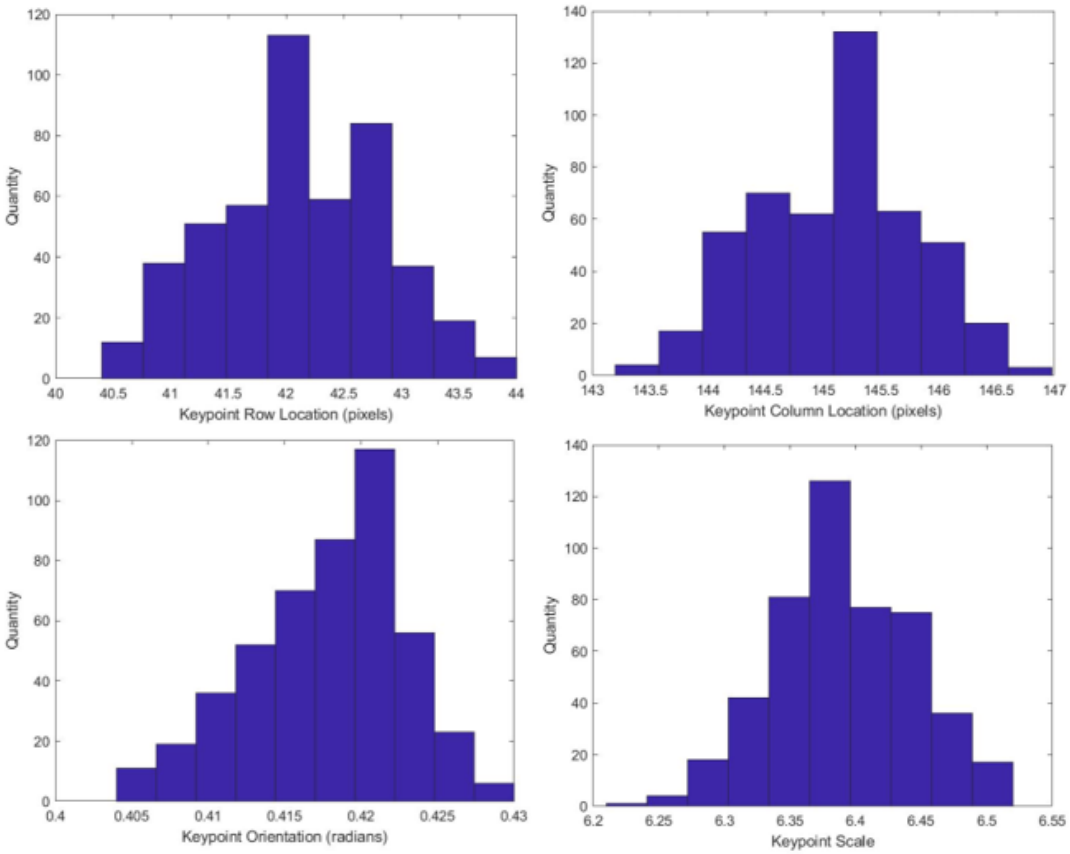

Figure 5.8: The distributions of the 4 location parameters. Each is non-Gaussian.

Certainly, there is much to do in this portion of the project before larger results are attainable, but this has laid the groundwork necessary for future work on the program.

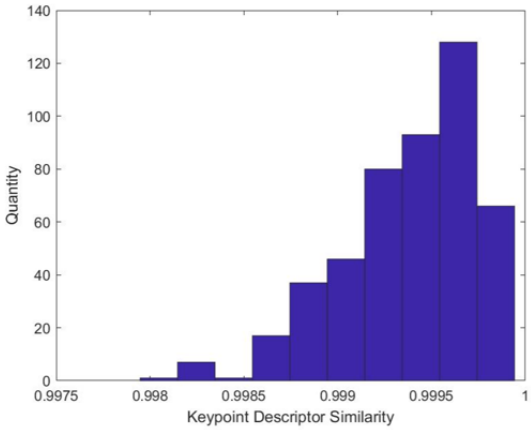

Figure 5.9: The distribution of SIFT Keypoint Similarity.

CHAPTER 6

CONCLUSION

Directly imaging exoplanets requires excellent technology: the very best telescopes employing sophisticated techniques during observation leading to complex, multi-faceted data analysis algorithms. This endeavor is clearly challenging, yet could yield a new understanding of planets in our universe as more and more are discovered. This dissertation is directly in service of that goal. I have leveraged modern mathematical and computational techniques for image processing in the service of a variety of projects, beginning with post-processing exoplanet direct imaging data.

In **Chapter 2** and **Chapter 3**, I have developed a novel technique for extracting planet signals from high-contrast direct imaging data. This was done by developing a Common Spatial Pattern technique and applying a matched filter to the results with a forward model developed from mathematical first principals as a template. I have shown the accuracy of this forward model, and demonstrated its effectiveness. The algorithm identifies planets with efficacy comparable to other state-of-the-art methods. CSPFMMF represents an entirely new tool available for anyone in the direct imaging scientific community to use.

After an update to the model in Python, I was able to test the method over a large variety of real stellar targets. I have shown that the the ideal way to use this new tool is to select a specific set of modes to keep and to set a specific threshold for the matched filter template. By using this set of parameters, I have shown with an ROC curve that the useful region of this method (approximately a 10^{-6} contrast) has a particularly high detection rate before identifying many false positives. Finally, I have correlated the detection SNR with a given planet's

flux contrast and angular separation from the host star. These results indicate that CSPFMMF is a robust, predictable algorithm that will be a useful analysis for direct imaging scientists.

Before post-processing can even take place, we must make stellar observations, and those observations should be made using the most sophisticated technology available. An imaginative concept to create a self-assembling 31-meter space telescope serves that end. Using a massive aperture in the least atmospherically noisy environment enhances our ability to detect the faintest signals, like exoplanets. The work in **Chapter 4** shows this approach. Here, modern software for optical propagation and resultant PSF analysis were used to inform the design of mirrors for a such a telescope. I leveraged Zernike decomposition and optical theory to design a mirror shape that leads to high fidelity images while adhering to physical limitations of hardware. This has shown that such an enormous project as the proposed telescope may be optically feasible.

The last examined piece of the space imaging puzzle comes from the actual collection of data. Currently, calibration images, state estimation, and science images are mostly all treated independently. The work in **Chapter 5** looked to change that. I employed image processing as a vital step to a novel filtering algorithm. I examined how models of cameras and distortion can lead to accurate predictions of non-Gaussian image features in a complex model. This lays the groundwork and establishes the necessity of a 4-moment model of uncertainty propagation in a filter that uses SIFT points as part of the measurement function that can determine state elements of the telescope from live science images.

Each of these projects, though disparate in nature, fall under the umbrella of space image processing. They serve the goal of developing our technology and

techniques to best detect small, faint signals throughout our universe. Image processing has been used at the start of an engineering project to influence the design of a telescope, as an intermediate step in a new type of filter, and in post-processing of complex direct imaging datasets. This dissertation represents several small steps to a future where we can more easily take pictures of Earth-analogs throughout our galaxy.

Appendix: Summary of Mathematical Notation

We present a table here summarizing each of the various notations and symbols used throughout §1 and §2. As a general note, bold values indicate matrices and vectors. With the set of bold symbols, upper case differentiates matrices from vectors.

Table 1: Summary of Symbols and Notation.

Symbol	Dimension	Description
	Descriptions of Data and Transformations	
N	1×1	number of images in a dataset
P	1×1	number of pixels within an image
λ	1×1	wavelength of an image
λ_0	1×1	reference wavelength
ϵ	1×1	scaling of target signal compared to noise
x	1×1	spatial position within an image
t	1×1	time image was taken
$\mathbf{i}$	$P \times 1$	vector-mean subtracted image
$\mathbf{t}$	$P \times 1$	specific, target image
$\mathbf{X}$	$N \times P$	dataset containing images as row vectors
$\mathbf{C}$	$N \times N$	image-to-image covariance matrix of a given dataset
$\mathbf{s}_{\lambda,t}$	$P \times 1$	speckle noise in a specific image
$\mathbf{S}_1$	$N \times P$	matrix of transposed $\mathbf{s}$ vectors in dataset 1
$\mathbf{S}_2$	$N \times P$	matrix of transposed $\mathbf{s}$ vectors in dataset 2
$\mathbf{a}_\lambda$	$P \times 1$	planet signal model in a specific wavelength
$\mathbf{s}_n$	$P \times 1$	single image speckle data in dataset n
$\mathbf{a}_n$	$P \times 1$	single image astrophysical signal model in dataset n

Continuation of Table 1

Symbol	Dimension	Description
$\mathbf{A}_1$	$N \times P$	matrix of transposed **a** vectors in dataset 1
$\mathbf{A}_2$	$N \times P$	matrix of transposed **a** vectors in dataset 2
$\mathbf{i}_{1_\lambda,t}$	$P \times 1$	image in dataset 1 after transformations
$\mathbf{i}_{2_\lambda,t}$	$P \times 1$	image in dataset 2 after transformations
$\mathbf{X}_1$	$N \times P$	dataset with spatially distributed target signals
$\mathbf{X}_2$	$N \times P$	dataset with spatially overlapping target signals
θ_t	1×1	parallactic angle of an image associated with time t
R_{θ_t}	-	operator indicating rotation of the argument by θ
		CSP and Forward Model Terms
$\mathbf{I}$	-	identity matrix: pure diagonal matrix of ones
$\mathbf{w}$	$N \times 1$	vector multiplier in PCA and CSP
$\mathbf{C}_1$	$N \times N$	sample image covariance matrix of $\mathbf{X}_1$
$\mathbf{C}_2$	$N \times N$	sample image covariance matrix of $\mathbf{X}_2$
$\mathbf{C}_+$	$N \times N$	composite covariance Matrix, sum of $\mathbf{C}_1$ and $\mathbf{C}_2$
$\bar{\mathbf{X}}_n$	$N \times P$	whitened dataset n
$\bar{\mathbf{C}}_n$	$N \times N$	whitened covariance matrix from dataset n
$\mathbf{u}_k$	$N \times 1$	eigenvector k of $\mathbf{C}_+$
$\mathbf{U}$	$N \times N$	eigenvector matrix of $\mathbf{C}_+$
Γ_k	1×1	eigenvalue k of $\mathbf{C}_+$
$\mathbf{\Gamma}$	$N \times N$	pure diagonal matrix of eigenvalues of $\mathbf{C}_+$
$\mathbf{P}$	$N \times N$	whitening matrix of $\mathbf{C}_+$
ϕ	1×1	lagrange multiplier
Φ_k	1×1	eigenvalue k of $\bar{\mathbf{C}}_1$
$\mathbf{\Phi}$	$N \times N$	pure diagonal matrix of eigenvalues of $\bar{\mathbf{C}}_1$

Continuation of Table 1

Symbol	Dimension	Description
$\mathbf{w}_k$	$N \times 1$	eigenvector k of $\bar{\mathbf{C}}_1$
$\mathbf{W}$	$N \times N$	eigenvector matrix of $\bar{\mathbf{C}}_1$
$\mathbf{Z}$	$N \times P$	projection matrix used to find planet signal
K	1×1	number of modes of $\mathbf{Z}$ used
$\mathbf{z}_k$	$K \times P$	row K of $\mathbf{Z}$, a single mode
$\delta\mathbf{z}_k$	$K \times P$	row K of $\delta\mathbf{Z}$, a single perturbation mode
$\hat{\mathbf{s}}$	$P \times 1$	planet signal estimate
$\mathbf{C}_{S_nS_n}$	$N \times N$	simplification of $\frac{1}{P-1}\left(\mathbf{S}_n\mathbf{S}_n^T\right)$
$\mathbf{C}_{AS_n}$	$N \times N$	simplification of $\frac{1}{P-1}\left(\mathbf{A}_n\mathbf{S}_n^T + \mathbf{S}_n\mathbf{A}_n^T\right)$
Λ_k	1×1	eigenvalue k of $\mathbf{C}_{S_1S_1} + \mathbf{C}_{S_2S_2}$
$\mathbf{\Lambda}$	$N \times N$	pure diagonal matrix of eigenvalues of $\mathbf{C}_{S_1S_1} + \mathbf{C}_{S_2S_2}$
$\mathbf{v}_k$	$N \times 1$	eigenvector k of $\mathbf{C}_{S_1S_1} + \mathbf{C}_{S_2S_2}$
$\mathbf{V}$	$N \times N$	eigenvector matrix of $\mathbf{C}_{S_1S_1} + \mathbf{C}_{S_2S_2}$
$\bar{\mathbf{C}}_S$	$N \times N$	the component of $\bar{\mathbf{C}}_1$ due to speckle noise
$\bar{\mathbf{C}}_A$	$N \times N$	the component of $\bar{\mathbf{C}}_1$ due to perturbations
Ω_k	1×1	eigenvalue k of $\bar{\mathbf{C}}_S$
$\mathbf{\Omega}$	$N \times N$	pure diagonal matrix of eigenvalues of $\bar{\mathbf{C}}_S$
$\mathbf{y}_k$	$N \times 1$	eigenvector k of $\bar{\mathbf{C}}_S$
$\mathbf{Y}$	$N \times N$	eigenvector matrix of $\bar{\mathbf{C}}_S$
$\mathbf{Z}_s$	$K \times P$	the component of $\mathbf{Z}_K$ due to speckle noise
$\delta\mathbf{Z}$	$K \times P$	the component of $\mathbf{Z}_K$ due to perturbations
$\mathbf{Z}_{(i,j),CSP}$	1×1	actual element of $\mathbf{Z}$ generated by CSP
$\mathbf{Z}_{(i,j),FM}$	1×1	predicted element of $\mathbf{Z}$ generated by the Forward model
l	1×1	number of wavelengths in a data cube

Continuation of Table 1

Symbol	Dimension	Description
$p_{CSP_{\lambda,k}}$	1×1	mode $\mathbf{Z}_k$ pixel value from CSP
$p_{FM_{\lambda,k}}$	1×1	mode $\delta\mathbf{Z}_k$pixel value from FM

BIBLIOGRAPHY

[1] R. L. Akeson, X. Chen, D. Ciardi, and M. et al Crane. The NASA exoplanet archive: Data and tools for exoplanet research. *Publications of the Astronomical Society of the Pacific*, 125:989, 2013.

[2] Andriy Andreev, Antti Kanto, and Pekka Malo. Computational examples of a new method for distribution selection in the pearson system. *Journal of applied Statistics*, 34(4):487–506, 2007.

[3] Kai Keng Ang, Zheng Yang Chin, Haihong Zhang, and Cuntai Guan. Filter bank common spatial pattern (fbcsp) in brain-computer interface. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pages 2390–2397, June 2008.

[4] Elizabeth R. DeLong, David M. DeLong, and Daniel L. Clarke-Pearson. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*, 44(3):837–845, 1988.

[5] John L Dwyer, David P Roy, Brian Sauer, Calli B Jenkerson, Hankui K Zhang, and Leo Lymburner. Analysis ready data: enabling analysis of the landsat archive. *Remote Sensing*, 10(9):1363, 2018.

[6] D. A. Fischer, B. Peterson, LUVOIR Team, et al. The luvoir mission concept study interim report. *arXiv preprint arXiv:1809.09668*, 2018.

[7] Zoltan J. Koles, Michael S. Lazar, and Steven Z. Zhou. Spatial patterns underlying population differences in the background eeg. *Brain Topography*, 2(4):275–284, Jun 1990.

[8] Matthew D Lallo. Experience with the hubble space telescope: 20 years of an archetype. *Optical Engineering*, 51(1), 2012.

[9] David Lowe. Siftdemo, Jul 2005.

[10] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[11] Bruce Macintosh, James R. Graham, Patrick Ingraham, Quinn Konopacky, Christian Marois, Marshall Perrin, Lisa Poyneer, Brian Bauman, Travis

Barman, Adam S. Burrows, Andrew Cardwell, Jeffrey Chilcote, Robert J. De Rosa, Daren Dillon, Rene Doyon, Jennifer Dunn, Darren Erikson, Michael P. Fitzgerald, Donald Gavel, Stephen Goodsell, Markus Hartung, Pascale Hibon, Paul Kalas, James Larkin, Jerome Maire, Franck Marchis, Mark S. Marley, James McBride, Max Millar-Blanchaer, Katie Morzinski, Andrew Norton, B. R. Oppenheimer, David Palmer, Jennifer Patience, Laurent Pueyo, Fredrik Rantakyro, Naru Sadakuni, Leslie Saddlemyer, Dmitry Savransky, Andrew Serio, Remi Soummer, Anand Sivaramakrishnan, Inseok Song, Sandrine Thomas, J. Kent Wallace, Sloane Wiktorowicz, and Schuyler Wolff. First light of the gemini planet imager. *Proceedings of the National Academy of Sciences*, 111(35):12661–12666, 2014.

[12] Fahim Mannan and Michael S Langer. What is a good model for depth from defocus? In *2016 13th Conference on Computer and Robot Vision (CRV)*, pages 273–280. IEEE, 2016.

[13] Franck Marchis, Eric L Nielsen, Sarah Blunt, Julien Rameau, Travis Barman, Robert J De Rosa, Quinn Konopacky, Mark S Marley, Christian Marois, Laurent Pueyo, et al. Discovery and follow-up of 51 eri b, a directly-imaged jupiter-like exoplanet and status of the gpies campaign. In *DPS*, pages 501–01, 2015.

[14] C. Marois, D. Lafreniere, R. Doyon, B. Macintosh, and D. Nadeau. Angular Differential Imaging: A Powerful High-Contrast Imaging Technique1. *The Astrophysical Journal*, 641:556–564, 2006.

[15] Christian Marois, David Lafrenière, Bruce Macintosh, and René Doyon. Accurate astrometry and photometry of saturated and coronagraphic point spread functions. *The Astrophysical Journal*, 647(1):612, 2006.

[16] NASA. Observatory Mirrors, 2019. [Online; accessed 1-Aug-2019].

[17] Karl Pearson. Mathematical contributions to the theory of evolution. xix. second supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 216:429–457, 1916.

[18] M. D. Perrin, A. Sivaramakrishnan, R. B. Makidon, B. R. Oppenheimer, and J. R. Graham. The structure of high strehl ratio point-spread functions. *The Astrophysical Journal*, 596:702, 2003.

[19] Marshall Perrin, Jerome Maire, Patrick J. Ingraham, Dmitry Savransky, Max Millar-Blanchaer, Schuyler G. Wolff, Jean-Baptiste Ruffio, Jason J.

Wang, Zachary H. Draper, Naru Sadakuni, Christian Marois, Abhijith Rajan, Michael P. Fitzgerald, Bruce Macintosh, James R. Graham, Rene Doyon, James E. Larkin, Jeffrey K. Chilcote, Stephen J. Goodsell, David W. Palmer, Kathleen Labrie, Mathilde Beaulieau, Robert J. De Rosa, Alexandra Z. Greenbaum, Markus Hartung, Pascale Hibon, Quinn M. Konopacky, David Lafreniere, Jean-Francois Lavigne, Franck Marchis, Jenny Patience, Laurent A. Pueyo, Fredrik Rantakyro, Remi Soummer, Anand Sivaramakrishnan, Sandrine J. Thomas, Kim Ward-Duong, and Sloane Wiktorowicz. Gemini Planet Imager Observational Calibrations I: overview of the GPI data reduction pipeline. In *Proc. SPIE*, volume 9147, 2014.

[20] Marshall D Perrin, Anand Sivaramakrishnan, Charles-Philippe Lajoie, Erin Elliott, Laurent Pueyo, Swara Ravindranath, and Loïc Albert. Updated point spread function simulations for jwst with webbpsf. In *Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*, volume 9143, page 91433X. International Society for Optics and Photonics, 2014.

[21] Laurent Pueyo. Detection and characterization of exoplanets using projections on karhunen–loeve eigenimages: Forward modeling. *The Astrophysical Journal*, 824(2):117, 2016.

[22] René Racine, Gordon A. H. 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.

[23] Kamisetty Ramam Rao and Patrick C Yip. *The transform and data compression handbook*. CRC press, 2000.

[24] J.-B. Ruffio, B. Macintosh, J. J. Wang, L. Pueyo, E. L. Nielsen, R. J. De Rosa, I. Czekala, M. S. Marley, P. Arriaga, V. P. Bailey, T. Barman, J. Bulger, J. Chilcote, T. Cotten, R. Doyon, G. Duchêne, M. P. Fitzgerald, K. B. Follette, B. L. Gerard, S. J. Goodsell, J. R. Graham, A. Z. Greenbaum, P. Hibon, L.-W. Hung, P. Ingraham, P. Kalas, Q. Konopacky, J. E. Larkin, J. Maire, F. Marchis, C. Marois, S. Metchev, M. A. Millar-Blanchaer, K. M. Morzinski, R. Oppenheimer, D. Palmer, J. Patience, M. Perrin, L. Poyneer, A. Rajan, J. Rameau, F. T. Rantakyrö, D. Savransky, A. C. Schneider, A. Sivaramakrishnan, I. Song, R. Soummer, S. Thomas, J. K. Wallace, K. Ward-Duong, S. Wiktorowicz, and S. Wolff. Improving and Assessing Planet Sensitivity of the GPI Exoplanet Survey with a Forward Model Matched Filter. *Astrophysical Journal*, 842:14, June 2017.

[25] Mike Safyan. Planet's dove satellite constellation. *Handbook of Small Satel-*

lites: Technology, Design, Manufacture, Applications, Economics and Regulation, pages 1–17, 2020.

[26] Gary H Sanders. The thirty meter telescope (tmt): An international observatory. *Journal of Astrophysics and Astronomy*, 34(2):81–86, 2013.

[27] D. Saumon, M. S. Marley, M. Abel, L. Frommhold, and R. S. Freedman. New H_2 Collision-induced Absorption and NH_3 Opacity and the Spectra of the Coolest Brown Dwarfs. *The Astrophysical Journal*, 750:74, May 2012.

[28] Dmitry Savransky, Dean Keithly, Jacob Shapiro, Gabriel Soto, Erik Gustafson, Kevin Liu, and Christopher Della Sanatina. Modular active self-assembling space telescope swarms. niac phase i, final report. Technical report, 2019.

[29] Jacob Shapiro, Dean Keithly, Gabriel Soto, Dmitry Savransky, and Erik Gustafson. Optical design of a modular segmented space telescope. In *Astronomical Optics: Design, Manufacture, and Test of Space and Ground Systems II*, volume 11116, page 111160D. International Society for Optics and Photonics, 2019.

[30] Jacob Shapiro, Dmitry Savransky, Jean-Baptiste Ruffio, Nikhil Ranganathan, and Bruce Macintosh. Detecting planets from direct-imaging observations using common spatial pattern filtering. *The Astronomical Journal*, 158(3):125, 2019.

[31] Anand Sivaramakrishnan and Ben R. Oppenheimer. Astrometry and photometry with coronagraphs. *The Astrophysical Journal*, 647(1):620, 2006.

[32] R. Soummer, L. Pueyo, and J. Larkin. Detection and characterization of exoplanets and disks using projections on Karhunen-Loeve eigenimages. *The Astrophysical Journal Letters*, 755(2):L28, 2012.

[33] H Philip Stahl. Jwst mirror technology development results. In *Proc. SPIE*, volume 6671, pages 667102–1, 2007.

[34] Zhongwei Tang, Rafael Grompone von Gioi, Pascal Monasse, and Jean-Michel Morel. A precision analysis of camera distortion models. *IEEE Transactions on Image Processing*, 26(6):2694–2704, 2017.

[35] J. J. Wang, J. R. Graham, L. Pueyo, P. Kalas, M. A. Millar-Blanchaer, J.-B. Ruffio, R. J. De Rosa, S. M. Ammons, P. Arriaga, V. P. Bailey, T. S. Barman,

J. Bulger, A. S. Burrows, A. Cardwell, C. H. Chen, J. K. Chilcote, T. Cotten, M. P. Fitzgerald, K. B. Follette, R. Doyon, G. Duchêne, A. Z. Greenbaum, P. Hibon, L.-W. Hung, P. Ingraham, Q. M. Konopacky, J. E. Larkin, B. Macintosh, J. Maire, F. Marchis, M. S. Marley, C. Marois, S. Metchev, E. L. Nielsen, R. Oppenheimer, D. W. Palmer, R. Patel, J. Patience, M. D. Perrin, L. A. Poyneer, A. Rajan, J. Rameau, F. T. Rantakyrö, D. Savransky, A. Sivaramakrishnan, I. Song, R. Soummer, S. Thomas, G. Vasisht, D. Vega, J. K. Wallace, K. Ward-Duong, S. J. Wiktorowicz, and S. G. Wolff. The Orbit and Transit Prospects for β Pictoris b Constrained with One Milliarcsecond Astrometry. *The Astronomical Journal*, 152:97, October 2016.

[36] J. J. Wang, J.-B. Ruffio, R. J. De Rosa, J. Aguilar, S. G. Wolff, and L. Pueyo. pyklip: Psf subtraction for exoplanets and disks. Astrophysics Source Code Library, June 2015.