Title
Footage of Other Worlds: Unveiling the Dynamical Architecture of Young Exoplanetary Systems

Author
Wang, Jason

Publication Date
2018

Footage of Other Worlds: Unveiling the Dynamical Architecture of Young Exoplanetary Systems

By

Jason Jinfei Wang

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Astrophysics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor James R. Graham, Chair
Professor Eugene Chiang
Professor Paul G. Kalas
Professor Philip Stark

Summer 2018

Footage of Other Worlds: Unveiling the Dynamical Architecture of Young Exoplanetary Systems

Abstract

Footage of Other Worlds: Unveiling the Dynamical Architecture of Young Exoplanetary Systems

by

Jason Jinfei Wang

Doctor of Philosophy in Astrophysics

University of California, Berkeley

Professor James R. Graham, Chair

This thesis focuses around using the Gemini Planet Imager (GPI), an instrument designed to directly-image exoplanets, as part of the GPI Exoplanet Survey (GPIES), a multi-year survey to image and characterize Jovian exoplanets orbiting at Solar System Scales (5-100 au). The first half of this thesis is on building the data analysis infrastructure for GPIES. I helped commission GPI by characterizing the ability to measure the position and flux of a star when it is placed behind an occulting mask specifically designed to suppress starlight. I also led the development of the automated data processing infrastructure for GPIES that handles all of the data storage, indexing, and processing. This data infrastructure has optimized survey execution, ensured uniform data products, and characterized instrument performance. In the second half of this thesis, I characterize two notable exoplanetary systems through precise astrometry and orbital analysis. I present improvements to astrometric data analysis that allow us to measure the position of β Pic b to one milliarcsecond, the most precise astrometry of an exoplanet to date. I tighten constraints on its orbit, including timing a window in which circumplanetary material of this young exoplanet could be transiting the star. Lastly, I apply the same astrometry techniques to the HR 8799 system which harbors four Jovian exoplanets. I add dynamical stability priors on my orbit fits using an N-body integrator to find stable orbits in the system. Through this, I explore possible orbital resonances and place dynamical constraints on the masses of the planets.

I dedicate this dissertation to my family.

Contents

List of Figures

List of Tables

Acknowledgments

This thesis would not be possible without the collaboration and support from so many people. I will try my best to credit them here.

I would first like to thank my advisors James Graham and Paul Kalas. James has taught me everything I know about optics, statistics, and orbits, which comprises the majority of my thesis and the backbone of all my research. He also generously provided me the opportunities and resources to thrive, which is not limited to buying me a computer with 32 cores, 256 GB of RAM, and 40 TB of RAID 5 storage which was fundamental for conducting three quarters of this thesis. Paul has taught me the ins and outs of proposal writing (especially HST proposals) and of working with a large research collaboration. He also has been constantly available through my graduate school career either keeping his door open so I can bounce ideas off him or checking up on me just to see how I was doing.

I also want to thank the instrument team that built GPI and the GPIES collaboration. Without GPI, this entire thesis would not exist, and many years of work went into designing and building the instrument before I arrived and benefited from using it. It has been a tremendous pleasure to work with the GPIES collaboration. I was welcomed with open-arms from day one, and, ever since, has been a great source of mentoring, knowledge, and collaboration. I thank the GPIES steering committee for making the rules and setting a great example to enable this positive atmosphere. I also worked with many great scientists. Rob De Rosa has helped me through all of my technical data reduction challenges throughout the years, including pointing out that I almost missed 51 Eri b, GPI's first exoplanet discovery. Laurent Pueyo and I have formed a great duo where he derives the math for data analysis and I implement it in software. He also shares many short and long term goals that have helped shaped my research direction. I have also been on collaborative projects with many GPIES graduate students: Schuyler Wolff, Jean-Baptise Ruffio, Alex Greenbaum, Pauline Arriaga, Abhi Rajan, and Max Millar-Blanchaer. Lastly, I benefited greatly from the work of Bruce Mactintosh, Marshall Perrin, Dmitry Savransky, Quinn Konopacky, and Jérôme Maire. They developed the foundation from which my work is built off of.

I would like to also thank the rest of my thesis committee outside of James and Paul. Eugene Chiang has provided insightful advice on all topics throughout my graduate career. Philip Stark has forced me to think about all the assumptions astronomers often take for granted when doing statistics, especially priors.

The staff in the Berkeley Astronomy Department has been super helpful all the time. I

am not sure how many questions Dexter Stewart has answered from me or how many hours she has put in to make sure I get paid every year, but she did all of that with a smile (enjoy your retirement!). I would also like to thank Bill Boyd in particular for helping me set up all my computing needs (and there were many), including running all of the department machines which I often hog with hundreds of Python processes.

Fellow BADGrads present and past, thanks for joining me in this marathon called graduate school. Especially to Carina Cheng, Zaki Ali, Tom Zick, Eve Lee, Dominic Ryan, and Lea Hirsch, thanks for making it fun while we did this together. I also benefited from the wisdom of past graduate students. Isaac Shivvers and Erik Petigura gave me valuable advice on succeeding in graduate school and becoming a real researcher. Thanks also to everyone that has contributed knowledge to the BADGrads Wiki and to Peter Williams for developing the ucastrothesis LaTeXthesis template.

I also want to acknowledge my past mentors. Terry Herter gave me my first serious research project when I was an undergraduate. He got me hooked on astronomy with my first observing experience (which just happened to be on the SOFIA airborne observatory) and the data pipeline work that followed. I am thankfully for working with him and Peter Gierasch on my first science project studying the atmosphere of Jupiter in the mid-infrared. I also want to thank my high school Latin teacher Bob Durrett, who not only taught me epic poetry but also instilled in me a sense of adventure to try interesting things.

I would also like to thank everyone outside of astronomy that have kept me sane during my PhD. Florine Pascal has been a constant companion through this endeavor. Michael Nazario, Dylan Beal, Patrick Druggan, Marston Bender, Matt Luscher, and Ellie Taylor have been enthusiastic in joining me for adventures both in the Bay Area and to wherever my various travels to telescopes and conferences take me.

To my parents, I would not have gotten here without you. You have always supported my education, no matter the cost. I am grateful for the opportunity to pursue my studies while also eating your tasty food. Thank you for everything.

Chapter 1

Introduction

1.1 Imaging Extrasolar Planets

Unlike stars that shine bright from nuclear fusion, planets are faint. Even in our own Solar System, the discovery of the furthest-out planets required distinguishing faint sources of light that slowly orbit the Sun from near stationary stars in the background and highly eccentric comets. As detecting planets in our Solar System is based on observing sunlight reflecting off of these worlds, planets further from the Sun become harder to find and require better telescopes. Due to this, the efforts to image all of the planets in our own Solar System has spanned many centuries (Herschel & Watson 1781; Galle 1846; Leonard 1930). Even now, the largest telescopes in the world have been searching the skies for a possible ninth planet in our Solar System that has eluded centuries of telescopes (Batygin & Brown 2016).

Imaging exoplanets comes with the added difficulty that the light from the faint planet resides next to a bright host star in the sky. Unlike Solar System planets which can be on the other side of the sky from the Sun, exoplanets can never get far from their stars from our point of view on Earth. A planet at 5 AU around a nearby star at 10 pc would only have an angular separation in the sky of $0.5''$. At these separations, detecting reflected starlight off the atmospheres or surfaces of these planets is impractical with our current instrumentation. Jupiter is a billion times fainter than the Sun, Earth is another ten times fainter than Jupiter, and both are inaccessible until future space missions (Crill & Siegler 2017).

Thankfully, we are not limited to only detecting mature exoplanets through reflected starlight. As a planet forms, it is heated by the material it accretes. Over time, it radiates away its heat from formation as it cools. Even though the exact physics of planet formation and accretion of materials is uncertain, planets that are only ten million years old are hot ($\sim 500 - 1500$ K) and are only 10^{-4}-10^{-6} times fainter in the near-infrared than stars similar to the Sun (Baraffe et al. 2003; Marley et al. 2007). This is several orders of magnitude brighter than they would be if they only reflected starlight. Thus, nearly all attempts to image exoplanets to date have targeted young systems to look for hot, young exoplanets (e.g., McBride et al. 2011).

With a perfect circular telescope with no optical aberrations and with no atmospheric

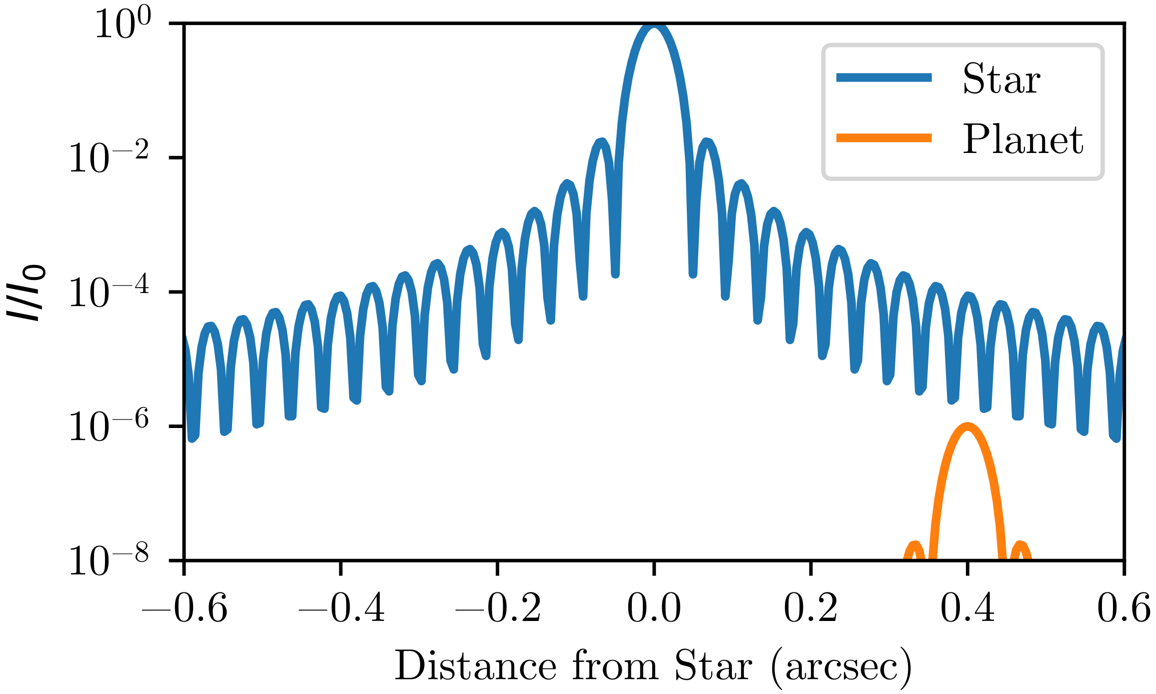


Figure 1.1: One-dimensional plot of the signals of a star and planet. Fluxes are normalized to the peak flux of the star. The planet is 10^{-6} times fainter than the star and separated by 0.4″. This assumes a perfect 8-m circular telescope in the absence of both optical aberrations and an atmosphere.

turbulence, a young Jovian planet that is within an arcsecond of a star is still extremely hard to image. Figure 1.1 shows the relative signal of a planet and a star for a case representative of the faintest planets that can be detected by current imaging instruments. The ringing signal of the star is due to the diffraction of starlight through a perfect circular aperture (i.e., due to the fundamental wavelike property of light). To see these planets, a coronagraph is needed to suppress diffraction and keep starlight away from the regions in the image where we are looking for planets.

Very few coronagraphs have been used in space for exoplanet imaging. With the coronagraphs on the *Hubble Space Telescope* ending operations in 2008, we will need to wait until the launch of the *James Webb Space Telescope* for another space coronagraph (Beichman et al. 2010). Otherwise, we must be able to handle atmospheric turbulence in order to use coronagraphs with ground-based telescopes, which have done the bulk of exoplanet imaging. Turbulence in the Earth's atmosphere will smear out the light from each source by ≈ 1″. This may be tolerable for an equal-brightness stellar binary, but will render a faint planet invisible. The planet both will have its light spread over a large portion of the image and will have the bright blur of the star on top of that. To mitigate this issue, ground-based telescopes use adaptive optics (AO) systems to correct for distortions in the wavefront of light due to Earth's atmosphere in order to spatially separate the light of the planet from the light of the star. Note that for eventual imaging of exo-Earths, even space telescopes will need AO systems to correct for optical aberrations due to thermal drifts in the instrument (Crill & Siegler 2017).

Not all exoplanets are unveiled simply by taking images behind a coronagraph and AO system. The majority of imaged exoplanet discoveries also required further removal of stellar glare by post-processing the data. Post-processing relies on observing strategies that decorrelate the signals of a planet from its star in the images. This thesis will focus on two strategies. The first is angular differential imaging (ADI), where the sky rotates due to the Earth's rotation so planets around a star rotate over the course of a night while the diffraction pattern of the star stays constant over time since it depends only on the orientation of the instrument (Marois et al. 2006a). The second is spectral differential imaging (SDI), where the stellar diffraction pattern moves out radially from the location of the star with increasing wavelength, while the signal of the planet stays at a constant location (Marois et al. 2000). In both ADI and SDI, the way and amount the star and planet decorrelate from each other is known: in ADI, we know how much a planet rotates based on its location in the sky; in SDI, we know how much the stellar diffraction moves based on the wavelengths the data were taken at. This information can then be put into algorithms that seek to remove the stellar light.

Many post-processing algorithms have been used. For this thesis, we use Karhunen-Loève Image Projection (KLIP; Soummer et al. 2012; Pueyo et al. 2015). KLIP uses principal component analysis (PCA) to model and subtract the signal of the star in the data. One advantage of KLIP is that the number of PCA modes to use is tunable, which limits overfitting the data and removing the signal of planets. This is an issue with ADI and SDI since these observing strategies do not remove the planet in the images, but rather mostly decorrelate the signal from the stellar signal. Thus, planet signal will significantly leak into the model of the star, but only after a certain number of PCA modes. Thus, KLIP provides a tunable parameter, the number of PCA modes, which allows us to trade of between maximizing the signal of the planet and suppressing stellar diffraction.

Thus, to image other worlds, we require the combination of AO systems to correct for atmospheric turbulence, coronagraphs to suppress the glare of the star, and post-processing techniques that leverage specific observation strategies to remove the residual signal of the star. With the current technology available to us, we are able to image young, Jovian planets and study how they formed.

1.2 Studying Imaged Worlds

Because of the technical challenges in directly imaging planets, we have found < 10 planets with orbital separations < 100 au (Bowler 2016). Despite the paucity of imaged planets, we can characterize in detail each planet we image by studying its orbit and spectrum. First, we must correctly extract out the position, flux, or spectrum of the planet in the data. This is particularly difficult, as post-processing distorts the signal of the planet due to overfitting the stellar diffraction pattern it was trying to subtract. The standard in the field has been using simulated planets to calibrate this effect or to inject a negative planet profile with opposite flux at the location of planet so that after post-processing there is only noise at the location (Marois et al. 2010a; Lagrange et al. 2010). Most recently, analytical methods have been developed that can calculate the distortions induced by algorithms such as KLIP (Pueyo 2016). While a technical detail at times, properly measuring the planet in the data is crucial to understanding their nature (Nielsen et al. 2017).

Direct imaging and transit spectroscopy are the only ways to study the atmosphere of exoplanets, but the two probe completely different planets: direct-imaging is sensitive to planets at large separations ($\gtrsim 5$ au) whereas transit spectroscopy only studies planets that are close in to the star with short orbital periods ($\lesssim 100$ days). By directly obtaining the spectra of giant planets and measuring molecular absorption features, we can infer their abundances and thus the composition of the planet. Since these giant planets lie between different snow lines (radial distances from the star beyond which the equilibrium temperature is cold enough that certain molecules solidify), the composition of the planets can inform us on their formation mechanisms and accretion histories (Öberg et al. 2011). Due to uncertainties in the atmospheric models, these analyses are difficult, but the work to date has favored core accretion (Pollack et al. 1996) for forming these directly imaged planets (Konopacky et al. 2013; Lavie et al. 2017).

By obtaining images of planets over time, we can observe the motion of these exoplanets and gain insights into the dynamics of the system. However, directly-imaged exoplanets orbit at at separations > 5 au from their stars, with the shortest known period to be $\sim$20 years (Wang et al. 2016). No exoplanet has been directly observed for a full orbit yet. Inferring the orbits from partial orbital arcs requires precise astrometry and robust statistical methods. Precise astrometry requires high angular resolution, good coronagraphs, and well calibrated instruments. The statistical precision, σ_x, of an astrometric measurement is approximated as

$$\sigma_x = \sigma_{PSF}/SNR. \tag{1.1}$$

Here, we assumed that the point spread function (PSF) of the planet can be approximated by a symmetric 2-D Gaussian function with standard deviation σ_{PSF} (Birney et al. 2006). The planet is detected with a certain signal-to-noise ratio (SNR) that depends on the brightness of the planet and the amplitude of the noise, which has contributions both from the ability to suppress the stellar glare and any detector or photon noise. If we assume we are using a really good AO system on a 8-m telescope observing at the diffraction limit at 1.6 μm ($\sigma_{PSF} \approx 20$ mas), we need a coronagraphic instrument and post-processing algorithms good enough to achieve a SNR of at least 20 to obtain 1 mas astrometry of the planet. Astrometry is often obtained by multiple instruments, especially for astrometric baselines that span many years, so calibrating multiple instruments together is another challenge. Finally, when we measure the planet's position, it needs to be in terms of relative offset from its host star

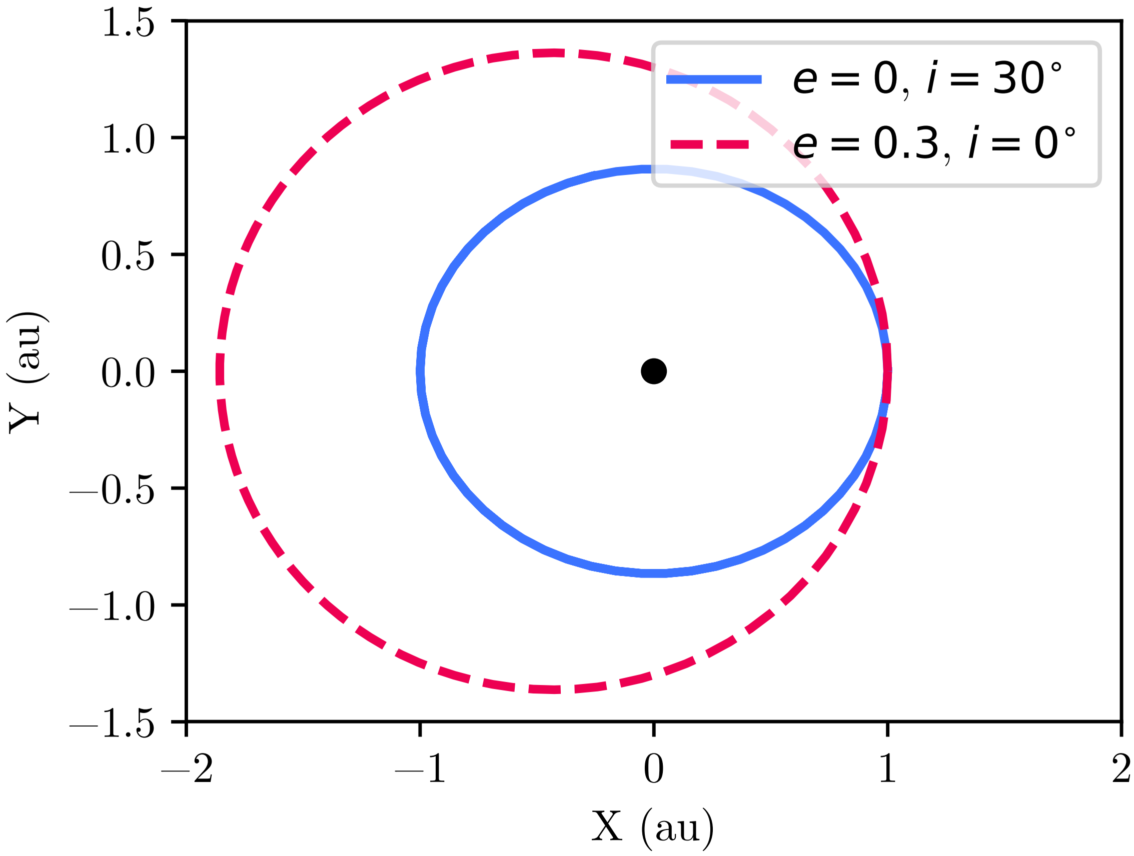


Figure 1.2: Two orbits plotted in 2-D projection to show a degeneracy in the orbital parameters. The orbit in the solid blue line is circular, but inclined 30° along the y-axis. The dotted red orbit is face-on but has an eccentricity of 0.3 with periastron along y=0. The projections of the orbits onto the 2-D sky plane are very similar on the right side.

to have physical meaning (e.g., for Keplerian orbits). As we typically place stars behind a coronagraphic mask, locating the star is equally important as locating the planet. In the design of recent exoplanet imaging instruments, fiducial spots are created by diffracting starlight before it is occulted by the coronagraph, a technique we will discuss further in this thesis (Marois et al. 2006b; Sivaramakrishnan & Oppenheimer 2006).

With partial orbit arcs, and without much, if any, knowledge of the position or velocity of the planet towards/away from Earth, many orbital solutions can satisfy the same data. At least six orbital parameters are required to describe a planet orbiting a star (Green 1985). Typically, there are many local minima in this six-dimensional space. For example, Figure 1.2 shows how, with a small orbital arc, degeneracies exist between an inclined or eccentric orbit. Many studies in the literature have used least squares fits, which currently in implementation only probe finite families of orbits (e.g., Marois et al. 2008; Zurlo et al. 2016). Recently, Bayesian techniques that use Markov chain Monte Carlo (MCMC) methods have been developed and can consider all the possible orbital degeneracies (Kalas et al. 2013; Macintosh et al. 2014; Pueyo et al. 2015; Millar-Blanchaer et al. 2015).

After we overcome the technical challenges in measuring the orbit of a directly imaged exoplanet, we are rewarded with insights into the dynamics of the system. The exoplanet β Pic b is one of the best examples of this (Lagrange et al. 2010). The star is known to harbor a debris disk with a strong warp feature (Burrows et al. 1995; Mouillet et al. 1997; Heap et al. 2000). By measuring the orientation of its orbit as well as excluding other possible planets through stability arguments, we know that β Pic b is responsible for driven the warp feature (Dawson et al. 2011; Lagrange et al. 2012). It is also launching comets into the star (Thébault & Beust 2001; Millar-Blanchaer et al. 2015), but it alone may not be responsible for cleaning out the cavity of the debris disk (Millar-Blanchaer et al. 2015). Lastly, the planet's orbit is so nearly edge-on that, while the planet does not pass in front of the star, its Hill sphere does and offers one of the best opportunities to probe the circumplanetary environment of a young exoplanet (Stuik et al. 2017; Mékarnia et al. 2017; de Mooij et al. 2017).

While the dynamics in other systems might be less clear than in the β Pic system, orbital analysis still can provide many insights. The HR 8799 system harbors four super-Jupiters orbiting between 15 and 70 au (Marois et al. 2008, 2010b). Tightly packing planets between $\sim$ 5-7M_{Jup} likely requires orbital resonances and offers an unique opportunity to constrain the masses of these planets dynamically (Fabrycky & Murray-Clay 2010; Goździewski & Migaszewski 2014). Peculiarly, the orbit of Fomalhaut b looks to cross the debris disk in the system, but the debris disk shows no signs of disturbance, hinting that the planet may be low-mass and bright due to light reflected off of a shroud of dust around the planet (Kalas et al. 2013). By studying both the orbit of HD 95086 b and the debris disk, we know that the gap in the debris disk is too wide for the one known planet to carve by itself and there should be other unseen planets closer in to the star (Rameau et al. 2016).

Finally, establishing long-baselines to constrain the orbit of these planets will provide synergies with the Gaia mission (Perryman et al. 2014). Gaia will measure the astrometric reflex motion of the host star due to planets orbiting it, leading to model-independent measurements of the masses of these planets. However, its five year mission is short, and will only measure a partial orbital arc for these known systems. By combining Gaia measurements with longer-term astrometry from direct imaging, we can have better constraints on the masses of these planets (Sozzetti et al. 2016). These dynamical masses can then be used to constrain models of planet formation, as the mass and luminosity of the planet encodes information on how the planet formed (Baraffe et al. 2003; Marley et al. 2007).

1.3 This Dissertation

This dissertation revolves around using the Gemini Planet Imager (GPI; Macintosh et al. 2014). GPI combines a high-order AO system (Poyneer et al. 2016), an apodized-pupil Lyot coronagraph (Soummer et al. 2011b), and an integral field spectrograph (IFS) that takes simultaneous images and spectra between 1 and 2.4 μm with spectral resolution between 30 and 80 (Chilcote et al. 2012; Larkin et al. 2014). The IFS can alternatively be used for imaging polarimetry instead of spectroscopy to look for polarized light from debris disks and the atmospheres of exoplanets (Perrin et al. 2015). GPI was finished, delivered to the Gemini South telescope in Chile, and began its commissioning in 2013. After a year of commissioning, the GPI Exoplanet Survey (GPIES) began a multi-year search around 600 young and nearby stars for Jupiter-like exoplanets orbiting beyond 5 au (McBride et al. 2011). GPIES aims to discover new planets with its improved performance over previous instruments, characterize imaged exoplanets through measuring their orbits and spectra, constrain the frequency of giant planets at Solar System scales, and resolve the circumstellar dust that is left over from planet formation.

The first half of this dissertation can generally be described as developing data analysis instrumentation: building and understanding the infrastructure for GPI data reduction. Chapter 2 is instrument commissioning work on understanding the astrometric and spectrophotometric calibration of the star, from which relative measurements of the planets properties are made, and is published in Wang et al. (2014). Chapter 3 describes the automated data processing system that was developed for GPIES that handles all data processing, storage, and indexing and is published in Wang et al. (2018). The second half of the thesis uses GPI to make the most precise astrometric measurements of directly-imaged exoplanets and characterizes the dynamics of exoplanetary systems through orbital analysis. Chapter 4 introduces data analysis improvements to obtain one milliarcsecond astrometry of exoplanets and applies it to the well-studied exoplanet β Pic b. This work was published in Wang et al. (2016). Finally, Chapter 5 explores the orbits of the four planets orbiting HR 8799, and in particular how the application of dynamical stability constraints can strongly constrain both the orbits and masses of the planets in the system. As of writing, this work has been reviewed by the GPIES collaboration and submitted to a refereed journal. This thesis also contains two appendices. Appendix A is a published appendix to Chapter 4. Appendix B details a frequentist approach to measuring the astrometry of directly imaged exoplanets, from which we can properly combine uncertainties on the planet's location with frequentist astrometric calibration uncertainties.

During my thesis, I have also written an open-source Python implementation of KLIP called pyKLIP (Wang et al. 2015). pyKLIP supports data from most major high-contrast imaging instruments and has a `GenericData` interface for everything else. In addition to removing the glare of the star with KLIP, it provides the tools to do scientific analysis such as searching for planets, quantifying sensitivity, extracting the position and spectrum of a planet, and characterizing the morphology of debris disks. We also have a full site of documentation on how to use and develop for pyKLIP. Unfortunately, I did not have the time to write a section on it as part of my thesis, but felt it was important to point out this contribution to the field. As of writing this thesis, pyKLIP has been used in over 25 publications.

Chapter 2

Gemini Planet Imager Observational Calibrations VIII: Characterization and Role of Satellite Spots

A version of this chapter was first published in *Proceedings of the SPIE*: Wang et al. (2014).

The Gemini Planet Imager (GPI) combines extreme adaptive optics, an integral field spectrograph, and a high performance coronagraph to directly image extrasolar planets in the near-infrared. Because the coronagraph blocks most of the light from the star, it prevents the properties of the host star from being measured directly. Satellite spots created by diffraction from a square grid in the pupil plane can be used to locate the star and extract its spectrum. We describe the techniques implemented into the GPI Data Reduction Pipeline to measure the properties of the satellite spots and discuss the precision of the reconstructed astrometry and spectrophotometry of the occulted star. We find the astrometric precision of the satellite spots in an H-band datacube to be 0.05 pixels and is best when individual satellite spots have a signal to noise ratio (SNR) of > 20. In regards to satellite spot spectrophotometry, we find that the total flux from the satellite spots is stable to $\sim 7\%$ and scales linearly with central star brightness and that the shape of the satellite spot spectrum varies on the 2% level.

2.1 Introduction

The Gemini Planet Image (GPI) is a new facility-class adaptive optics (AO) instrument for the Gemini Observatory that is optimized for coronagraphic observations of bright ($I <$ 9 mag.) natural guide stars. The principal science goal of GPI is imaging the environment of young (~ 125 Myr) stars in the solar neighborhood (~ 50 pc) to discover and characterize self-luminous planets and planetary debris disks. GPI uses low resolving power ($R \sim 45$) spectra to study exoplanet atmospheres and astrometry to quantify their kinematics (McBride et al.

2011).

Astrometry of exoplanets is necessary to discriminate between background objects and physically associated bodies and to measure their orbital elements. The semimajor axis, orbital inclination (relative to other components of the system), and eccentricity all carry important clues to their origin and evolution. Due to the large semimajor axis separation of GPI-detected planets, orbital periods are long (10—1000 yr) and orbital coverage is necessarily incomplete. While such fragmentary data can be used to reconstruct the orbit using modern statistical techniques (Kalas et al. 2013; Macintosh et al. 2014), the lack of a complete orbit places milli-acrsecond requirements on the accuracy and precision (Graham 2009).

Near-IR ground-based spectrophotometry has to be corrected for strong wavelength dependent absorption by the terrestrial atmosphere (notably by H_2O and CO_2). Traditionally, this correction is constructed from contemporaneous observations of calibration stars of known spectral type (typically A or G dwarf stars) at similar airmass. However, this requires additional planning and overhead for each star that is to be observed. It is more efficient when this calibration is done simultaneously with the science observations.

Coronagraphic observations pose a unique problem for differential spectrophotometry and astrometry of an exoplanet relative to its primary, which is occulted. To overcome this obstacle, GPI includes a square grid in a pupil plane which acts as a two-dimensional amplitude grating. The grid is superimposed on the the pre-occulter pupil apodizer (left image of Figure 2.1). Diffraction of starlight from this grating injects first-order diffraction spots into the field of view for a given wavelength. In each wavelength channel in spectral mode, this creates reference spots that we term “satellite spots” (center image of Figure 2.1). In broadband polarimetry mode, the satellite spots become streaks extending radially outward from the location of the star (right image of Figure 2.1). These satellite spots preserve the information needed to reconstruct the spectrum and location of the occulted star: the diffraction pattern is centered on the true stellar position and is imprinted with an attenuated version of the stellar spectrum (Sivaramakrishnan & Oppenheimer 2006; Marois et al. 2006b). In this report we discuss how we measure the satellite spots and investigate the spectrophotometric and astrometric properties of these diffraction spots.

2.2 Satellite Spot Extraction

The raw data from the integral field unit is reduced using the GPI Data Reduction Pipeline (DRP; Perrin et al. 2014). For both spectral and polarimetry mode, the data are dark-subtracted, corrected for flexure, and extracted into three-dimensional datacubes, where the x and y dimensions are angular coordinates on the sky. In spectral mode, the z-dimension of the datacube consists of the flux in 37 wavelength channels; in polarimetry mode, the z-dimension consists of the broadband flux in two orthogonal polarization modes.

There are a couple of steps that are important to properly measuring the satellite spots. Extracting the flux from each microlens (particularly in spectral mode) is challenging, espe-

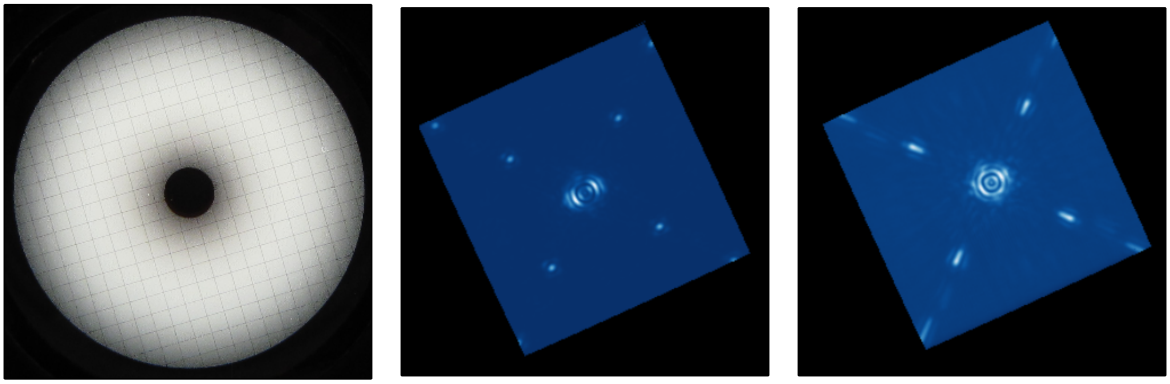


Figure 2.1: *Left*: An image of the H-band apodized pupil mask fabricated by Aktiwave LLC, Rochester, NY. *Center*: A wavelength slice from an H-band spectral datacube taken in the lab showing both the first order diffraction spots used for calibration and the second order spots that peek out from the edge of the detector. *Right*: A total intensity H-band image from a polarimetry mode datacube taken in the lab showing both the first-order and second-order diffraction spots.

cially in the presence of instrument flexure between the microlens array and the dectector (Wolff et al. 2014). We characterize flexure shifts by taking argon arc lamp spectra after the telescope slews to a new target. Additionally, we apply a geometric distortion correction to all the images to rectify the position of the satellite spots (Konopacky et al. 2014).

2.2.1 Spectral Mode

Locating & Measuring the Satellite Spots

Locating the satellite spots in the reduced spectral data cubes is typically achieved in two steps. First, an approximate set of spot locations is determined from one wavelength slice of the cube (this step can be bypassed if guesses for the locations of the satellite spots are manually supplied). These approximate locations are then scaled linearly with wavelength to find initial positions for all cube slices, which are then used to fit the precise location of each spot in each slice independently.

The first step of the procedure is described in detail in Savransky et al. (2013). Briefly, we assume that the four satellite spots form a perfect square in the image, and that they are relatively bright as compared with their neighboring pixels (although they need not be the brightest features in the image, nor even brighter than the median image value). These assumptions allow us to carry out an efficient search of the image. We iteratively construct a list of candidate spot locations, identifying the current brightest point in the image, and then removing it, along with a radius of x pixels about it, from the search region (the value of x is determined by the average size of a satellite spot in all cube wavelengths). These bright images are stored in the expanding set $\{\mathbf{r}_i\}_{i=1}^N$, where each vector ($\mathbf{r}_i$) is the two-dimentional pixel coordinate of the candidate location. As the list is being built, we conduct a breadth-first search for pixels sets forming perfect squares, where each node is a subtree whose branches represent all of the combinations of the root node with three other elements of $\{\mathbf{r}_i\}$. The value of all terminal nodes is binary, given by operating on the set of six distances between any four points ($\{\mathbf{r}_i\}_{i=1}^4$):

$$\{d_k\}_{k=1}^6 = \{\|\mathbf{r}_i - \mathbf{r}_j\| : i, j \in \{_4C_2\}\} \ , \tag{2.1}$$

where $\{_4C_2\}$ is the set of two element sets representing all combinations of four elements taken two at a time, without repetition. Assuming the set of distances is ordered by increasing magnitude, the four points define a square if and only if the first four distances ($\{d_k\}_{k=1}^4$) are equal, the last two distances ($\{d_k\}_{k=5}^6$) are equal, and the ratio of the magnitudes of the two subsets of distances is $\sqrt{2}$.

This search is made much more efficient by introducing a pruning heuristic, namely that all perfect squares contain right triangles. For any subset of three of the four candidate vertices, the corresponding subset of distances:

$$\{d'_k\}_{k=1}^3 = \{\|\mathbf{r}_i - \mathbf{r}_j\| : i, j \in \{_3C_2\}\} \ , \tag{2.2}$$

must contain two elements that are equal, and one element that is $\sqrt{2}$ times larger. This is a necessary, but not sufficient, condition for defining a square. This condition allows us to prune whole branches of the tree, and greatly improve the search efficiency. This is also the smallest subtree that can be tested, as any two arbitrary points can form an edge of a square.

Once identified, the approximate satellite spot locations can be refined by fitting a template point spread function (PSF) to the area around the approximate location. The usual approach is to assume a fixed-width Gaussian PSF as the template, and to use this in a matched filter (convolving with the pixels around the approximate location to find the best-fit offset).

When extracting the peak flux of a satellite spot, we first fit the full width half maximum (FWHM) to the actual data to calculate the standard devation of the Gaussian template. We then use this Gaussian template of fixed width in the convolution and fit for the height of the Gaussian. This fitting step also allows us to relax the assumption of a symmetric PSF, as we can fit the FWHM in two orthogonal directions and thereby account for any PSF ellipticity.

In order to ensure that our brightness assumption holds, we typically apply a high-pass filter to the image prior to beginning the satellite spot search. The high-pass filter is also applied when estimating the precise spot locations (so as not to bias the astrometry with a background gradient) and can be optionally applied when fitting the spot peaks to exclude the residual AO halo from contrast calculations. We can also include specific constraints on the admissible size of elements in the distance sets in Eqs. (2.1) and (2.2), based on prior knowledge of the approximate relative separation of satellite spots in each wavelength band, which further allows us to prune the search tree and increase the efficiency of the original search.

Locating the Central Star

In the current GPI pipeline, the central star is located by taking a simple mean of all 4×37 satellite spot positions. We rely on the fact that after distortion correction, opposite pairs of satellite spots are displaced equally and in the opposite direction in both x and y from the central star. For lab data, this technique works well as any single poorly measured satellite spot will have little effect on the measured central star position.

However, most on-sky data taken during GPI commissioning runs were without the atmospheric dispersion corrector (ADC) and suffers from atmospheric differential refraction (ADR)(Hibon et al. 2014). Due to this, the central star is in a different position in each wavelength slice and a simple mean of all satellite spots can be biased in the case where a companion and its primary have different spectral shapes. The simplest approach to mitigate ADR effects is to measure the star position in each slice individually. This technique works well as long as the satellite spots are bright in each individual frame. We use this technique for all of the astrometric stability analysis in this paper to disentangle ADR effects from our satellite spot measurements. An illustrative description on how we locate the occulted star position is described in section 2.3 and Figure 2.3.

2.2.2 Polarimetry Mode

In broadband polarimetric images, the satellite spots are smeared radially outwards from the central star. The true location of the star would have all four streaks pointing towards it, excluding effects from ADR which will be discussed in section 2.3.2.

Locating the Central Star

To take advantage of the streaks of satellite spots, we take a simplified Radon transform of the image, implementing a technique similar to (Pueyo et al. 2015) with P1640 images. We will summarize the procedure here. First, we combine the two orthogonal polarization channels to form a total intensity image. We also run a high pass filter on the image by subtracting a median filtered image created with a 9×9 pixel box to calculate the median. If there are any visible objects (e.g. binaries) in the image, we manually mask them out at this point as they will interfere will the Radon transform. Then, we begin the Radon transform technique. We pick a guess for the center of the image and draw straight lines that go through this point. For each line, we compute the line integral of the image along that line (this corresponds to one point in Radon space). We then sum the square of the line integral along each of the lines (summing the square of each of the corresponding points in Radon space). Thus, we have gotten a measure of how much light is “pointed" at this pixel. We do this for each pixel in a small search area around the initial guessed center of the image

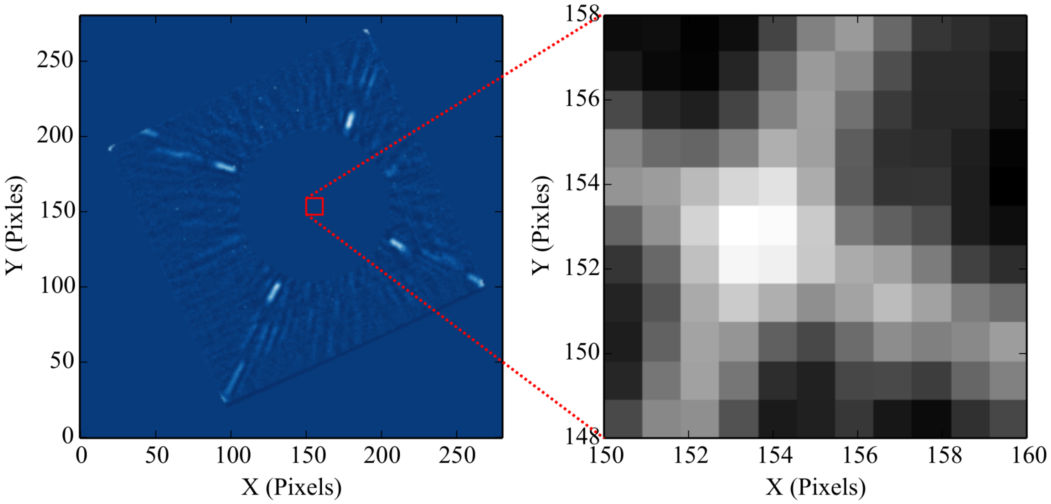


Figure 2.2: *Left*: A total intensity image of HIP 118666 in polarimetry mode with the central star masked. The image has been high-pass filtered using a median filter. The red box indicates the search area for the star center. Note that atmospheric differential refraction effects are apparent as the satellite spots do not all point back to the same location. *Right*: Map of the search box where each pixel represents the total intensity of all straight lines that pass through that pixel which we calculate using a Radon transform. This is before we interpolate this map to a finer resolution.

(see the left image of Figure 2.2). At this point, we have created a map of the intensity of the image in Radon space for each pixel (see right image of Figure 2.2). The pixel with the most power harbors the location of the central star. To achieve subpixel accuracy, we then interpolate this map to a finer resolution and pick out the subpixel that has the greatest value.

Several of the steps we have outlined are introduced because of computational efficiency. Ideally, given infinite computational resources, we would interpolate the image before performing the Radon transform to achieve subpixel resolution. However, if we did do this, the computational time increases by a factor of order N^4 where N is the over sampling factor (i.e. $N = 2$ means doubling the image sampling). This is because the Radon transform method we use scales as N^2 and we need to do it for ever pixel in a search box which also grows as N^2. We also limit our search box to a small area near the true center for the same reason. Using the default 280×280 pixel image and an 11×11 pixel search box, this routine runs in a few seconds on a 3 GHz core.

Measuring Satellite Spot Fluxes

Because the satellite spots are smeared into streaks, a precise flux measurement requires the development of new techniques to address this issue. We defer this work to a future paper and will not present any characterization of satellite spot photometry in polarimetry mode.

2.3 Astrometric Stability

In the following section, we present on the ability of the satellite spots to locate the occulted star. Because we can extract more information out of the satellite spots in spectral mode since they are not smeared out, we will focus most of our analysis on the spectral mode of GPI. We will then compare the spectral mode stability to the polarimetry mode stability.

First, we note that it is difficult to establish the absolute accuracy and precision of the satellite spots. Images taken without the coronagraph that reveal both the star and the satellite spots would establish the reliability of reconstructing the star position. However, we have found it is impossible to achieve good signal to noise on the satellite spots without saturating the central star. Another technique of selecting long sequences of coronagraphic images and measuring the occulted star position is strongly affected by uncorrected tip/tip errors that cause the star to move and also cannot be used for this task.

Using known astrometric binaries was a promising method we investigated. By measuring the relative separation of the binary companion and the occulted star, we could characterize the satellite spot astrometric accuracy. However, this method presents similar challenges. Known companions that fit into GPI's field of view are bright and saturate before we can obtain good signal to noise on the satellite spots. Thus, we have not been able to perform an absolute calibration of our ability to locate the occulted star. After an extensive search of all binaries, we found that the binary star system HIP 70931 has a faint enough companion

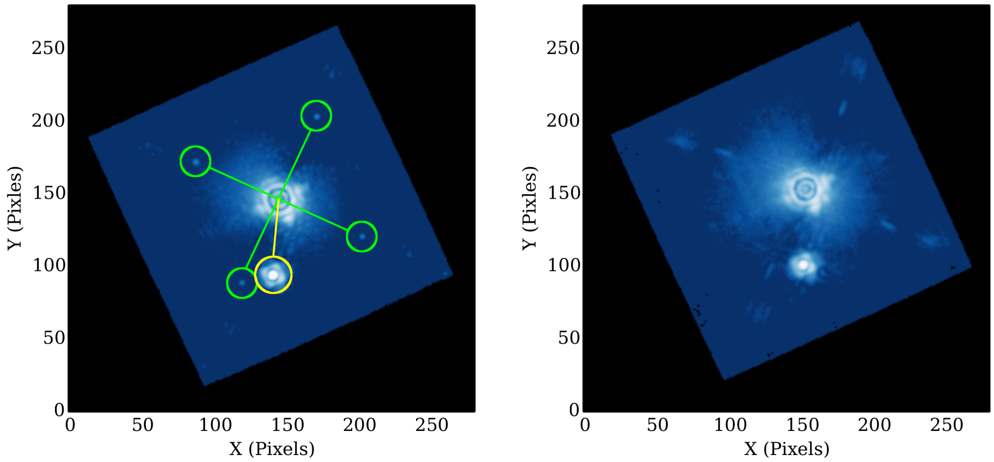


Figure 2.3: Sample GPI images of the HIP 70931 binary system. These images have gone through the basic steps in the pipeline to be extracted into datacubes, but no further processing of the data has been done. *Left*: The 1.663 μm wavelength slice of the first image taken in spectral mode on 2014 March 24 08:48 of HIP 70931. Overlaid on top is a schematic of how we measure the binary separation. The occulted star position is located using the satellite spots in green. From the inferred star position, the binary separation can then be measured (yellow). *Right*: A total intensity H-band image from the first exposure of the 2014 March 24 polarimetry mode dataset of HIP 70931.

and is therefore a suitable target to perform a binary separation analysis. Figure 2.3 shows some GPI data for this binary and a schematic of how we perform the astrometry analysis. The orbital parameters of this system have not yet been well characterized and cannot be used for an absolute calibration. We will instead present our relative stability analysis.

2.3.1 Spectral Mode

In spectral mode, we measure the separation of the binary system in each wavelength slice of each datacube. Then, we average the separations calculated for each datacube to obtain the measured separation of that particular datacube. We do this to obtain similar centroiding precision as the regular pipeline processing step of finding the star center using all the wavelength slices together while allowing us to disentangle effects due to ADR. For HIP 70931, the companion and satellite spots are sufficiently bright to work with individual wavelength slices. The companion is also much brighter than the satellite spots ($\sim$ 10 times brighter than the combined flux from all four satellite spots), so the uncertainty in the measured separation should probe errors in the satellite spot centroids. We use the same Gaussian matched filter approach for locating the companion as the satellite spots to reduce systematic biases.

HIP 70931 was observed on four separate nights in spectral mode (see Table 2.1 for a listing of nights and bands). We note that on 2014 May 10 and 2014 May 11, cloudy weather interfered with the observations and forced us to increase exposure times to compensate for extinction. Data on 2014 May 13 was taken during AO testing, causing the data quality to vary as AO parameters were adjusted. To mitigate this, we discarded any frames where the FWHM of the binary companion was greater than four pixels. We also discarded any frames in any of the datasets where the binary companion overlaps with a satellite spot.

In Figure 2.4, we plot the measured separation in each wavelength slice of each datacube as a function of average satellite spot signal to noise ratio (SNR). SNR was calculated by aperture photometry using an aperture radius of 5 pixels to measure the flux, a sky annulus between 8 and 12 pixels to measure the sky background, and the assumption that the scatter is Poisson noise. We see that the astrometric precision of the satellite spots decreases significantly when the average SNR of the four satellite spots in a single wavelength slice falls below ~ 20. In the following analysis of the astrometric precision of the satellite spots, we do not consider any individual wavelength slice where the average satellite spot SNR < 20. This SNR corresponds to a photon noise accuracy limit of < 0.2 pixels as given by FWHM/SNR for an individual satellite spot in H-band (typical FWHM of < 4.0 pixels). If we combine the four satellite spots, this would then limit the uncertainty on the position of the occulted star to < 0.1 pixels. This leads us to conclude that future observations taken with GPI should ensure SNR of individual satellite spots to be above 20 to ensure the central star can be located precisely.

Within a single exposure, the errors seem to be dominated by photon noise. In Table 2.1, we log the average scatter in the separation measured in each slice of a datacube and the theoretical limit on centroiding accuracy due to purely photon noise for a single slice. We calculate this limit using the expression (FWHM/fSNR)/2. fSNR is the average signal to noise of the four satellite spots after running a high-pass filter (median filter subtraction with a 9×9 pixel box) on the image, since this is done before locating the spots in our algorithm. We divide the centroid precision of a single satellite spot by $\sqrt{4} = 2$ to get the precision of locating the occulted star since we take the mean of the four satellite spots to locate the star. We find that the scatter in the measured separation approaches the photon noise limit of the central star position. Thus, within a single datacube we conclude that the majority of the error in centroiding on the central star is dominated by photon noise in the satellite spots.

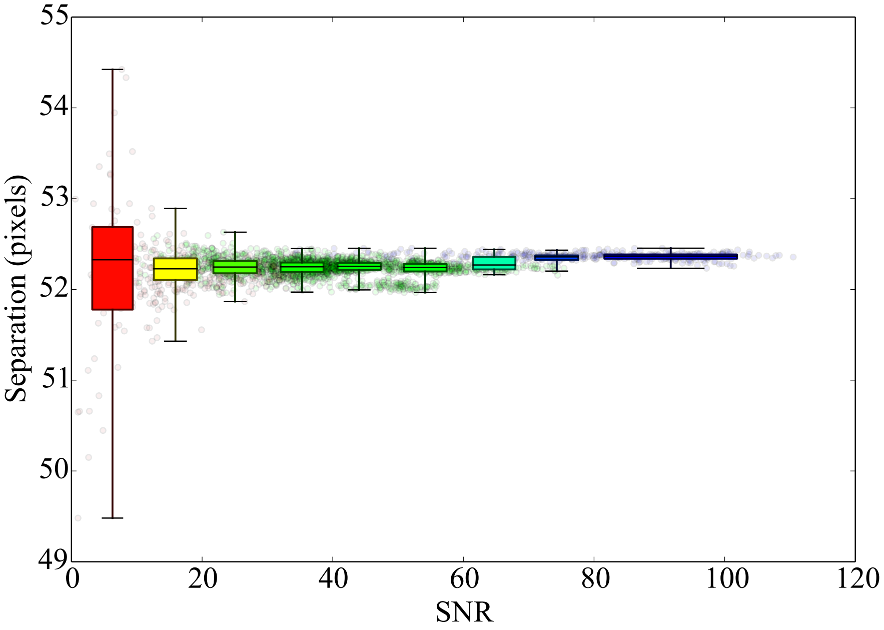


Figure 2.4: Separation of the binary system HIP 70931 as a function of the average signal to noise ratio (SNR) of the four satellite spots in a wavelength slice of a datacube. Each faded circle in the background represents a separation calculated for an individual wavelength slice of a datacube, with color representing wavelength (red is J-band, green is H-band, and blue is *K1*, the first half of K-band). Overlaid on top are boxplots with the center of the box being the median, the edges of the box being the first and third quartiles of the data, and the edges of the whiskers being the most extreme points. Each boxplot represents a SNR range of 10, except for the last box, which represents SNR $>$ 80 and is wider. The color of the box indicates the fraction of points in each bin from each wavelength band (i.e. the red bin has almost exclusively points from *K1* and the yellow bin is a mix of H and *K1* data).

Table 2.1: Measurement of the HIP 70931 binary separation for various observation sequences in spectral mode. We list the average internal scatter of measurements from wavelength slices within a single datacube ("Internal Scatter"), the photon noise limit for centroiding precision of occulted star for an individual slice ("(FWHM/fSNR)/2"), and the mean separation and sample standard deviation as calculated from combining datacubes in an observing sequence ("Measured Separation").

Observation Start (UT)	Band	Frames	Internal Scatter[1] (pixels)	(FWHM/fSNR)/2 (pixels)	Measured Separation[2] (pixels)
2014 March 24 08:48	*H*	7	0.03	0.03	52.26 ± 0.01
2014 March 24 09:07	*H*	4	0.03	0.03	52.04 ± 0.02
2014 May 10 03:52	*H*	10	0.04	0.03	52.26 ± 0.05
2014 May 10 05:46	*J*	9	0.03	0.02	52.36 ± 0.03
2014 May 11 04:22	*H*	6	0.08	0.06	52.26 ± 0.04
2014 May 11 05:08	*K1*	7	0.10	0.06	52.17 ± 0.03
2014 May 13 03:28	*H*	31	0.04	0.04	52.25 ± 0.05
All *H* Combined	*H*	58	0.05	0.04	52.24 ± 0.07
Clipped *H* Combined[3]	*H*	54	0.05	0.04	52.26 ± 0.05

We can no longer reach down to the photon noise limit when averaging datacubes together. When looking at the measured separation between datacubes in a single epoch, we would expect the photon noise floor for the sample standard deviation to be around $(\mathrm{FWHM}/\mathrm{fSNR})/\sqrt{37 \times 4} \approx 0.005$ pixels in H-band since we are now averaging over all satellite spots in all slices of a datacube. The last column of values in Table 2.1 shows that we do not get down to that level of precision, hitting a noise floor between 0.01 and 0.05 pixels. In fact, some observing sequences saw an increase in scatter when combining values from different datacubes. Looking at Figure 2.5, we see trends in the measured separation as a function of time. It is especially pronounced in the last four exposures taken on 2014 March 24 and in the dip in the measured separation on 2014 May 13.

We speculate this could be due to one or more of the following effects. Qualitatively from looking at the images as they evolve in time, we noticed that the PSF changes due to uncorrected wavefront aberrations. This could cause biases in our centroiding routine. In particular, the satellite spots, which are a diffraction phenomenon, do not have the same PSF as a star imaged on the detector. The satellite spots are slightly elongated along one axis due to diffraction since the wavelength channels span a finite wavelength range. Additionally, speckle noise near the satellite spots and diffraction spikes from the occulted star, which are aligned with the satellite spots, vary in time and can bias the centroiding if positioned correctly. However, we have not determined if any of these effects is the cause of the variations we see. We do note that we do not see any trends in the measured separation as a function of wavelength or as a function of the satellite spot distance from the occulted star (a proxy for plate scale).

Altogether, in the four different epochs of observation in H-band plotted in Figure 2.5, we find that despite the uncertainties, the data agree very well. The mean separations listed in Table 2.1 are almost identical for these four epochs. When we combine these four nights of observations, we find the uncertainty in the occulted star position is 0.07 pixels. If we eliminate the four outlier points from 09:00 on 2014 March 24, the uncertainty decreases down to 0.05 pixels (or ~ 0.7 mas), which we take as the uncertainty in the central star position for an H-band datacube along one axis. When moving to other bands (J and $K1$), we see a systematic disagreement with H-band on the order of 0.1 pixels. Since almost all of the astrometric calibration and analysis has been done at H-band, it is not surprising that we have not accounted for some chromatic effects.

Our uncertainties in satellite spot astrometry should not greatly impact GPI's science performance though. This error term is below the 1.8 mas per epoch requirement for GPI astrometry. Astrometric uncertainties in plate scale and distortion are of comparable magnitude (Konopacky et al. 2014). Additionally, when it comes to GPI's main science goal, exoplanets, we will likely be limited by the low SNR of the exoplanet itself - we would need a SNR of greater than 30 for centroid errors due to photon noise from the exoplanet to fall

---

[1] rms scatter of separations measured between slices of each individual datacube

[2] error quoted is sample standard deviation in the measured separation among datacubes

[3] removed data from the four exposures taken at 2014 March 24 09:07 that appear to be outliers (see Figure 2.5)

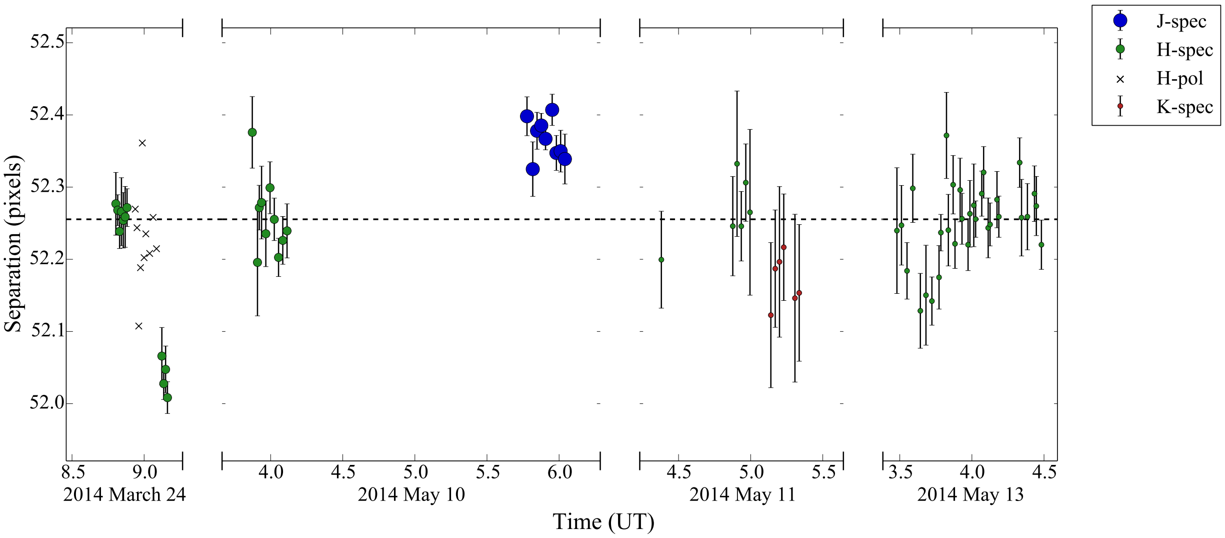


Figure 2.5: Measured separation of the binary system HIP 70931 as a function of time. Each circle represents the mean separation measured in a single datacube, with error bars representing the rms scatter within the datacube and color representing wavelength of the data as denoted in the legend. The three different sizes of markers represent 3 different signal to noise bins: the smallest symbols indicate datacubes with a median satellite spot SNR of less than 40, medium sized ones are for SNR between 40 and 60, and the largest markers for SNR greater than 60. Black crosses represent separations calculated in polarimetry mode in H-band. The horizontal dashed line represents the mean separation calculated with all of the H-band spectral data, excluding the four images with abnormally low separation taken 2014 March 24.

under 0.1 pixels in H-band.

2.3.2 Polarimetry Mode

Limited data are available to characterize satellite spot astrometry in polarimetry mode. We will focus on ten observations of the binary HIP 70931 taken on 24 March 2014 in between the two spectral mode observations discussed in the previous section. Using the Radon transform technique described in section 2.2.2 to find the occulted star, we measure a binary separation of 52.23 ± 0.06 pixels. The measured separation and error is comparable to the spectral data (see Figure 2.5).

We note the lack of systematic bias between the polarimetry and spectral data. All of the HIP 70931 data taken on 24 March 2014 are without the atmospheric dispersion corrector (ADC) at an airmass of $\sim$ 1.15. In spectral mode, we were able to mitigate chromatic effects by analyzing each wavelength slice independently, something we are not able to do in polarimetry mode. Despite that, finding comparable separations indicate that atmospheric differential refraction (ADR) does not significantly affect the satellite spots astrometric precision in polarimetry mode, at least not for binary star systems taken in H-band at airmasses less than about 1.15. It remains to be seen how great of an effect ADR is for planets, which differ more in color with respect to their host stars, and at low elevations when ADR effects are the largest. This however should be mitigated with the commissioning of the ADC (Hibon et al. 2014).

Table 2.2: Measurements of the grid ratio (ratio of brightness of the occulted star to the total brightness of all four satellite spots) and corresponding magnitude difference

Target	Date	Filter	Grid Ratio	Δm (mags)
β Pic	2013-11-18	H	5433 ± 1790	9.3 ± 0.4
β Pic	2013-12-10	H	5395 ± 953	9.3 ± 0.2
β Pic	2013-12-11	H	4569 ± 1936	9.2 ± 0.5
HD 118335	2014-03-25	H	6147 ± 964	9.5 ± 0.2

2.4 Spectrophotometric Stability

We used data collected on targets observed in the coronagraphic mode, during the first two commissioning runs in November and December of 2013, in order to measure the spot flux as a function of target magnitude (Figure 2.6). The observations were reduced using the typical reduction primitives within the GPI DRP. The background within each image was subtracted by applying a high-pass filter, with the satellite spots being masked to prevent self-subtraction. Aperture photometry was then performed on the four satellite spots within each wavelength slice of the reduced data cube. These measurements, obtained with the H and $K1$ filters, show a linear trend of decreasing satellite spot flux as a function of stellar flux.

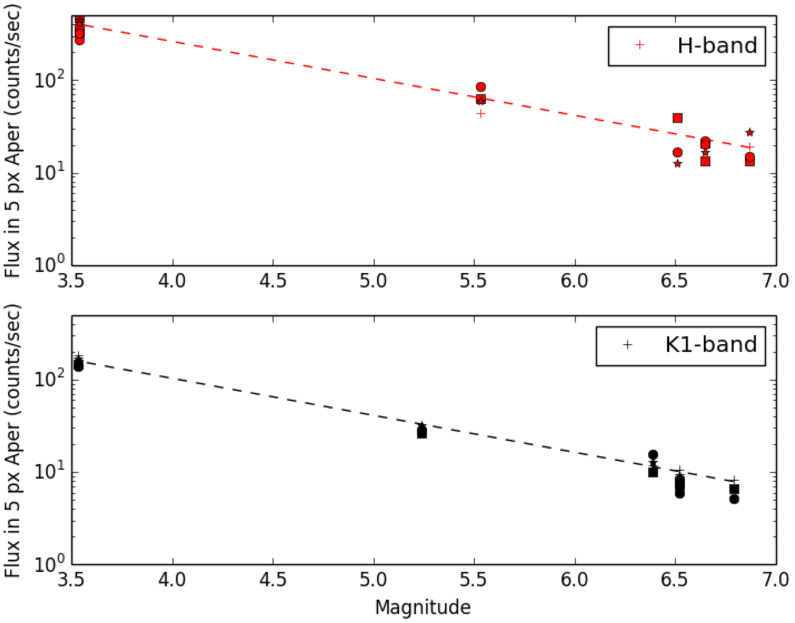


Figure 2.6: The average flux measured within the 11th to 14th wavelength slice of the reduced data cubes for the four satellite spots (differentiated by symbol) in the H-band (red dashed line and symbols), and $K1$-band (black dashed line and symbols). The fits to the measured fluxes as a function of magnitude are consistent with the expected relation of $m = -2.5\log(f) + c$.

The ratio of the flux of satellite spots, where we average the flux of all four spots, to the flux of the star behind the occulting mask was also estimated using two different techniques and is presented in Table 2.2. A detailed discussion of all the photometry results from the GPI photometric calibration program presented at this conference (Maire et al. 2014).

2.4.1 Long duration monitoring

To measure the spectrophotometric stability of GPI, data taken on multiple stellar targets with time-series durations spanning several minutes to multiple nights were analyzed. The data presented in this paper were collected primarily in the first and second calibration runs over Nov, Dec 2013 and Mar 2014.

Figure 2.7 and 2.8 show preliminary results on the long-term stability of the GPI spot data for β Pic, used as a proxy for the stability of the integral field spectrograph (IFS). In Figure 2.7 we have analyzed a single slice near the middle of the wavelength cube (1.662 μm), to estimate the stability of the spot photometry. The aperture photometry was performed using a 5-pixel radius aperture on unsharp-masked (9 px box median filtered) images. Plotted in the figure are the trends for each of the individual satellite spots, the sum of the diagonal spots and the sum of all four spots in the field. The sum of all four spots has the lowest scatter amongst the different trends plotted on the figure. The sum of spots is used for the measurement of spectrophotometric monitoring over shorter time durations in the next section. Figure 2.8 shows the same dataset as the one used for Figure 2.7, however it plots the spectral response of the sum of all four satellite spots for each of the spectral cubes observed over the three nights. The spectra have been all been normalized over the same spectral range to enable comparison of the scatter in the dataset, and the residual scatter over the three nights is ~2% indicating that the spot spectra are fairly stable and consistent across the three nights.

2.4.2 Short duration monitoring

For the J-band, two separate continuous data sets were identified, each of which span time periods of ~15-25 min. These data were reduced using standard methods as described in Section 2, and then we measured aperture photometry of the spots using `DAOPHOT`. Since the data is collected over a short duration and taken at low airmass, we chose to not correct the photometry for airmass or seeing variations. For the J-band data we chose two separate targets, β Pic and HIP 47115, which is of similar brightness as HR 8799. Figure 2.9 shows the sum of the spots (which was found to be the most stable choice) over the time duration. For the stack of 60s exposures, the β Pic J-band data shows ~6% scatter over the duration of the data and for HIP 47115 it is ~10%.

For the H-band, we utilized data taken on β Pic and HD 8049 both of which have companions visible within the GPI field-of-view. Figure 2.10 shows the variations in the data. The normalized spot flux has variations of $\sim$ 7% in the longer duration β Pic A spot

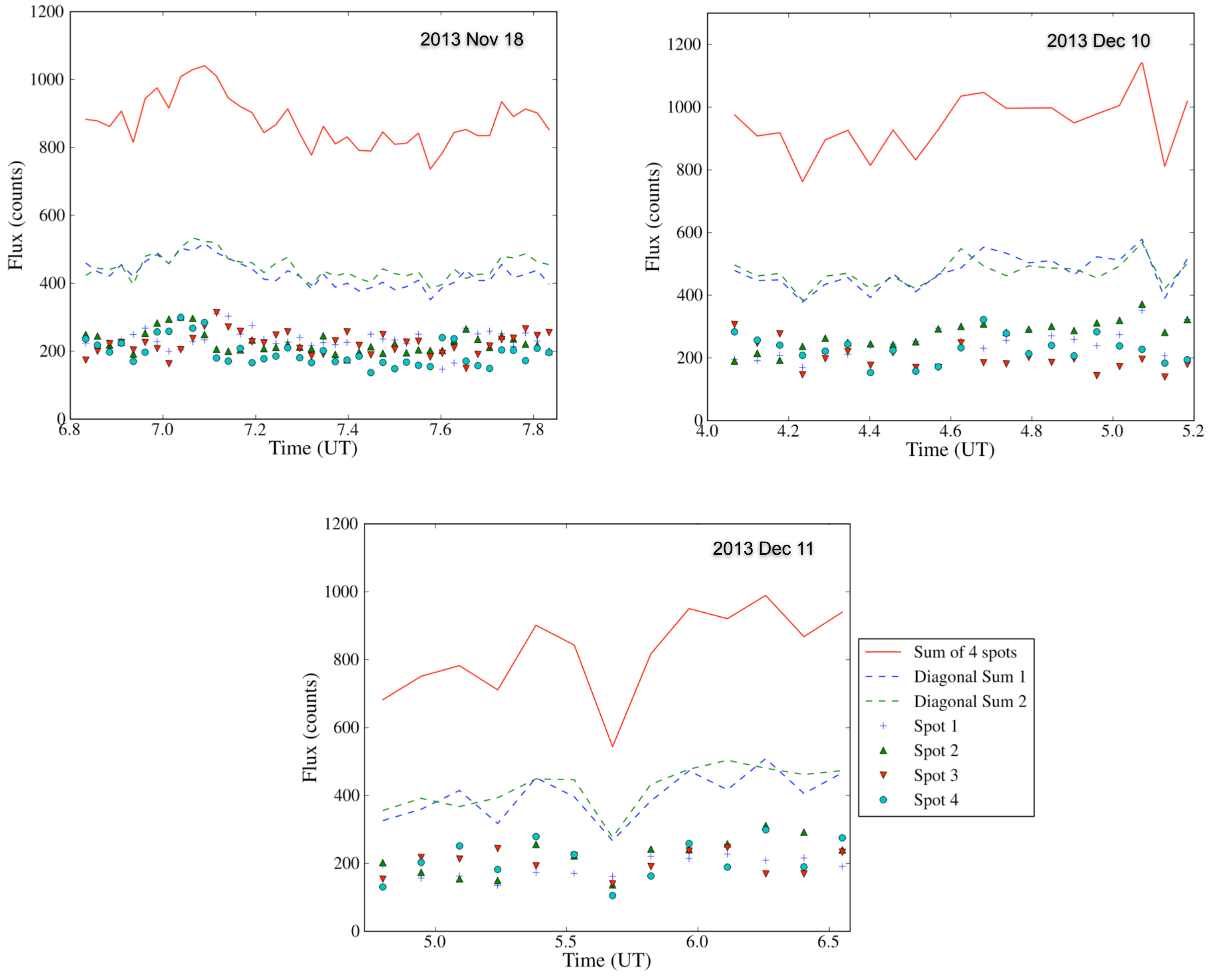


Figure 2.7: GPI photometric monitoring of β Pic. Plotted in the figure are the photometric data of individual spots in a 5 pixel aperture for a single wavelength slice. Also plotted are the sum of the two diagonal spots and all four spots represented by the dashed and solid lines. The sum of all four spots has the lowest relative scatter amongst the various combinations.

light curves and ~2% and ~1% variations in the spots and companion in the HD 8049 system respectively.

2.5 Future Work

Future work on improving satellite spot astrometric precision will proceed in three main areas. First, we will investigate the time-varying errors that are likely hindering the satellite spot astrometric precision. We will look at logs from the AO system to try to correlate

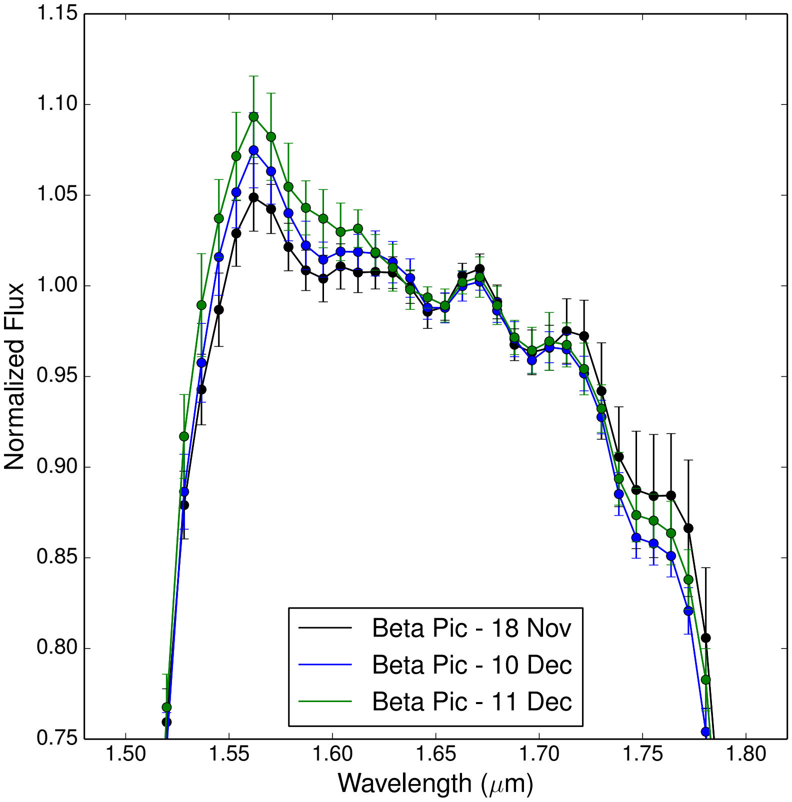


Figure 2.8: GPI spectrophotometric monitoring of β Pic. Plotted in the figure is the spot spectrum taken over three epochs during the GPI commissioning run. The data are all normalized to the same region of the spectrum and the error bars are indicative of the scatter and the stability of the instrument over these short exposure times.

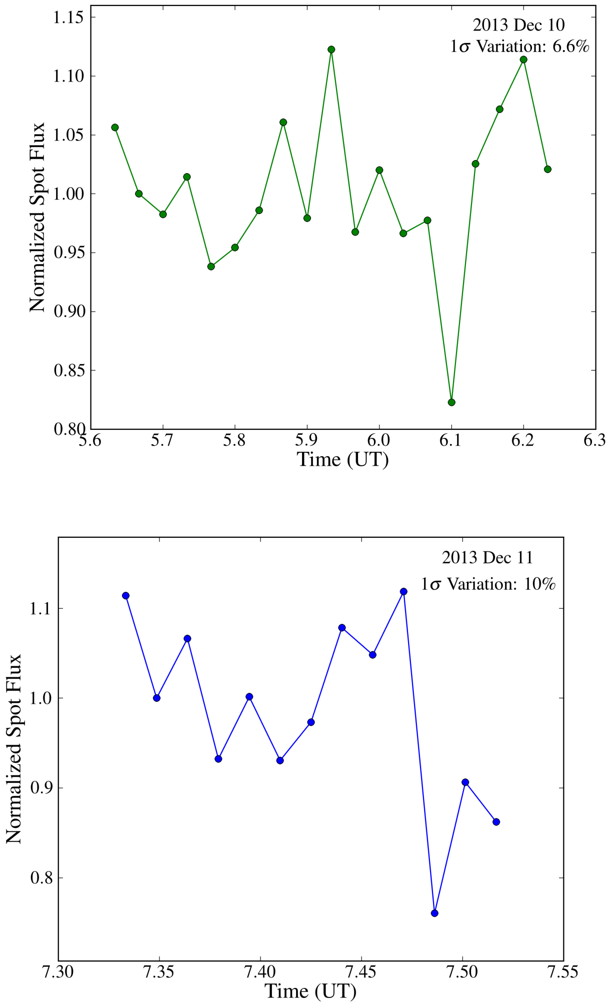


Figure 2.9: GPI photometric stability in the J-band. Measurements involve two different stars (β Pic and HIP 47115) with short time-series taken on the 10 and 12 Dec 2013. Aperture photometry with a 5 px aperture was used to estimate the flux for all the data.

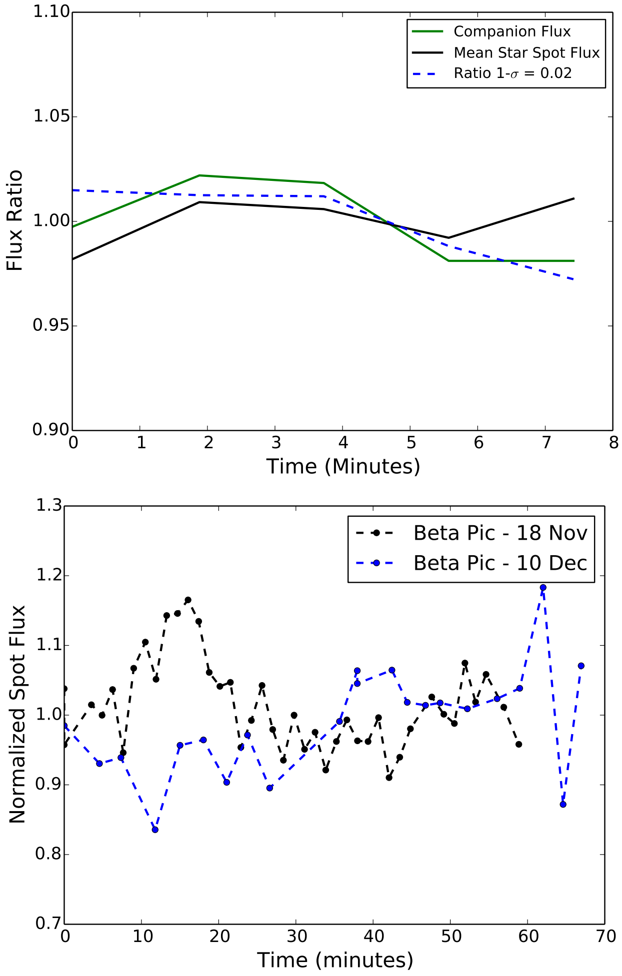


Figure 2.10: GPI photometric stability in the *H*-band. Measurements involve using two different stars (β Pic and HD 8049) with short and long time baseline observations. Aperture photometry with a 5 px aperture was used to estimate the flux for the entire data set.

the variations we see in the PSFs and the measured binary separation to anything physical. Next, we can improve the theoretical photon noise limit on relative centroiding within a datacube by combining satellite spots measured in all wavelength channels in a datacube to locate the occulted star. We will develop a least-squares fit to all the satellite spots to calculate a center, compensating for ADR or residual ADR if data is taken with the ADC. Having a model that connects the measurements from all 37 wavelength channels can improve robustness in locating the central star (e.g. insensitive to one bad satellite spot location) and help with images with low SNR satellite spots where the four satellite spots in a single wavelength slice are not precise enough. Lastly, we will continue searching for an astrometric calibrator suitable for GPI that will allow us to make an absolute calibration on satellite spot astrometry.

Future work on measuring the spectrophotometric stability of the GPI spots will require longer duration datasets, preferably over multiple hours in all the wavelengths. The length of the datasets we have now, less than ~ 40 minutes in all wavelengths aside from the H-band, currently limits the measurements. Furthermore we need to measure the stability not just as a function of time but also across the different wavelengths. Some multi-epoch and multi-wavelength data exists for this work and further data will be collected over the final commissioning run to completely characterize the spot spectrophotometric stability. Additionally, we will develop techniques to measure and characterize satellite spot fluxes in polarimetry mode.

Acknowledgements

The Gemini Observatory is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), Ministério da Ciência, Tecnologia e Inovação (Brazil), and Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina). J.R.G., J.J.W., and P.K. thank support from NASA NNX11AD21G, NSF AST-0909188, and the University of California LFRP-118057. This work is also supported in part by NASA grant APRA08-0117 and the STScI Director's Discretionary Research Fund. A.Z.G. is supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1232825

Chapter 3

Automated data processing architecture for the Gemini Planet Imager Exoplanet Survey

A version of this chapter was first published in *Journal of Astronomical Telescopes, Instruments, and Systems*: Wang et al. (2018).

The Gemini Planet Imager Exoplanet Survey (GPIES) is a multi-year direct imaging survey of 600 stars to discover and characterize young Jovian exoplanets and their environments. We have developed an automated data architecture to process and index all data related to the survey uniformly. An automated and flexible data processing framework, which we term the Data Cruncher, combines multiple data reduction pipelines together to process all spectroscopic, polarimetric, and calibration data taken with GPIES. With no human intervention, fully reduced and calibrated data products are available less than an hour after the data are taken to expedite follow-up on potential objects of interest. The Data Cruncher can run on a supercomputer to reprocess all GPIES data in a single day as improvements are made to our data reduction pipelines. A backend MySQL database indexes all files, which are synced to the cloud, and a front-end web server allows for easy browsing of all files associated with GPIES. To help observers, quicklook displays show reduced data as they are processed in real-time, and chatbots on Slack post observing information as well as reduced data products. Together, the GPIES automated data processing architecture reduces our workload, provides real-time data reduction, optimizes our observing strategy, and maintains a homogeneously reduced dataset to study planet occurrence and instrument performance.

3.1 INTRODUCTION

The Gemini Planet Imager (GPI) is a high contrast imaging instrument on the Gemini South telescope designed to directly image young, recently formed exoplanets and their planet-forming environments (Graham et al. 2007; Macintosh et al. 2014). To suppress the glare of the bright host star to see faint planets and circumstellar material, GPI is equipped with a high-order adaptive optics (AO) system to correct for atmospheric turbulence (Poyneer et al. 2016) and an apodized-pupil Lyot coronagraph to suppress diffraction from the star (Soummer et al. 2011b). To image and characterize planets and disks, an infrared integral field spectrograph (IFS) sits behind the AO system and coronagraph. The IFS uses a microlens array to disperse light from a 2.8 × 2.8 arcsecond field of view onto a HAWAII 2-RG detector in one of five filters, Y, J, H, $K1$, and $K2$, where K-band has been split into two filters. The IFS supports both a spectral mode, with a spectral resolving power R between 30 and 80 depending on the wavelength, and a polarimetry mode that allows for broadband imaging polarimetry using a Wollaston prism and a rotating waveplate(Chilcote et al. 2012; Larkin et al. 2014; Perrin et al. 2015).

The GPI Exoplanet Survey (GPIES) is a multi-year survey of 600 young, nearby stars with GPI to discover new planets, characterize the orbits and atmospheres of directly-imaged planets, constrain giant planet occurrence rates, and resolve planet-forming environments. GPIES is searching for planets in spectral mode in H-band around 600 stars, following-up and characterizing planet candidates in multiple wavelength bands at multiple epochs, and looking for resolved circumstellar material in polarized, scattered light around a subset of stars.

From both a technical and scientific perspective, GPIES requires an automated infrastructure to handle the data associated with the survey. Technically, both the size and complexity of data processing requirements make manual bookkeeping impractical. Over several years, GPIES will accumulate more than 30,000 raw frames of science data and a roughly equal number of raw frames of calibration data. While this is not close to the scale of the largest surveys in astronomy, it is large enough that we need automated methods to track and organize the data. Additionally, data reduction for GPIES is non-trivial, requiring a pipeline to reduce IFS data and complex algorithms to model and subtract out the diffracted light from the host stars. For each raw IFS frame, we must calibrate and extract ∼35,000 microspectra from the detector to make a single spectral datacube. If we consider each wavelength slice of each spectral datacube as an individual frame for data processing, which effectively is how it gets treated, then we will need to subtract the diffracted light of the star from over one million frames using multiple algorithms to mitigate algorithmic biases. To maintain data consistency, we need an automated system to reprocess all of our data as upgrades and bug fixes are applied to our data reduction pipelines.

Scientifically, we need the infrastructure to make sense of all the data from GPIES. Consistently processed data is important to maintain a homogeneous dataset to use for statistical planet occurrence studies. Being able to easily access specific data from across the entire survey is important to understand trends in instrument performance, inform observing

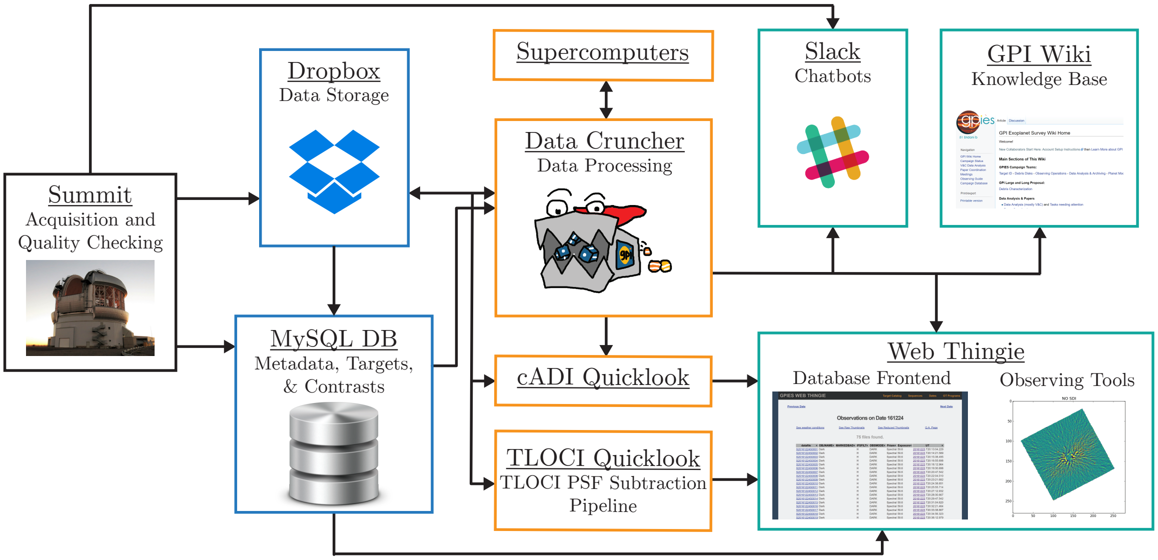


Figure 3.1: Schematic of the GPIES automated data processing infrastructure. Boxes represent the different components of the infrastructure that are described in Section 3.3. The boxes are colored so that black represents the telescope, blue represents data storage, orange represents data processing modules, and teal represents user-facing services. Arrows indicate the flow of data or information from one component to another.

strategies, and prioritize future instrument upgrades. Finally, having real-time processing during the observing night is extremely advantageous for optimizing observing strategy. Fully-reduced data that is processed and displayed in real time helps observers assess the data quality being achieved. It also allows observers to identify candidate companions within an hour of the observations so that the most promising candidates can be followed up the same night or the following night.

Thus, we have built an automated data processing infrastructure capable of storing all of the data, processing all science and calibration data in real-time, reprocessing the entire campaign as necessary, and providing powerful but easy-to-use tools to search across all data from the survey. The individual components and how they fit together are displayed in Figure 3.1.

Section 3.2 details all of the data reduction steps for GPIES, which we have automated. The different components that comprise the data infrastructure, including the Data Cruncher, a key component that automatically reduces all GPIES data, are described in Section 3.3. In Section 3.4, we describe the benefits of this automated architecture for observing, data processing efficiency, instrument performance, and survey statistics. Finally, in Section 3.5, we conclude with plans for future steps to take for the automated data processing architecture.

3.2 Data Reduction for GPIES

3.2.1 Calibration Data

All processing of calibration files uses the Data Parser module of the open-source and publicly available GPI DRP (Perrin et al. 2014, 2016). The Data Parser can be used through its GUI interface or programmatically in IDL using the `parse_fileset_to_recipes()` function that is associated with the `parsercore` object. Either way, the Data Parser is automatically able to identify and create data reduction steps for all of our calibration data.

In addition to the standard observatory calibration sequence of dark frames that are taken at a variety of exposure times, GPIES takes a sequence of 80 60 s dark frames at the end of each night, as most of the campaign uses 60 s exposure times. Each sequence of darks of the same exposure time is combined to produce one dark frame(Ingraham et al. 2014).

Wavelength calibration frames, or “wavecals," are taken at two times: during the day and right before each science sequence. During the day, deep sequences are taken with both the argon and xenon arc lamps at each band as part of the standard observatory calibration sequence. These are processed to be “master wavecals" since they are of high SNR and allow for the wavelength solution of each microspectra to be computed individually. Before each science sequence, GPIES takes a 30 s argon arc lamp observation at the band the science data is to be observed in (except in K-band, where an H-band arc lamp frame is taken). This frame is processed using the “Quick Wavelength Solution" recipe template, which measures shifts in the microspectra due to instrument flexure by using a master wavecal and computing the global offset in both X and Y necessary to shift the master wavecal to align with the microspectra of the argon arc lamp. This frame is only used to correct for flexure, and relies on the master wavecal for the rest of the wavelength solution. The accuracy of the GPI wavelength solution is < 2 nm in each band (Wolff et al. 2014).

For polarimetry data, we need calibration files that specify the locations and the shapes (fit as 2-D Gaussians) of the polarization spot pairs for each lenslet. Each spot in the pair corresponds to the intensity in a given polarization channel of a spatial pixel in a polarimetry datacube. The calibration files provide the required mapping between each spot on the detector and its location within the final datacube (where the 3rd dimension represents the polarization state). Calibration files are generated in each band for the polarimetry spot calibration using the flat field lamp of the Gemini calibration unit, GCAL, and processed using the GPI DRP Data Parser.

3.2.2 Spectral Mode

The majority of data from GPIES comes in the form of IFS data, where the raw data consists of ~35,000 microspectra spread across the detector.

Constructing Spectral Datacubes

To extract the microspectra into a more useful data product, we use the GPI DRP to construct spectral datacubes. The datacubes are constructed and calibrated using the steps listed in Table 3.1, which are slightly different than those of the default recipes offered as part of the GPI DRP. Each step corresponds to a primitive, one specific reduction task, that can be run in the GPI DRP.

While most of the primitives are straightforward, there are a few points to clarify. The wavelength calibration that is loaded in is typically the argon arc snapshot taken right before each science sequence described in Section 3.2.1. In K-band, the master wavecal is loaded in since H-band arc snapshots are taken instead. To then correct for instrument flexure, we use the “BandShift" method of the “Update Spot Shifts for Flexure" primitive. For sequences not in K-band, this only corrects for shifts of the microspectra due to flexure caused by changes in the elevation of the telescope between the snapshot arc and the current science frame. For a K-band sequence, this feature also corrects for the shift between the current data and the master K-band wavecal that is loaded in by measuring the offset between the H-band master wavecal and the H-band arc snapshot that was taken before the sequence. This requires the H- and K-band master wavecals to be taken together without the telescope having been moved in between so that they experience the same flexure.

Due to a few large clusters of bad pixels on the detector, we are not able to interpolate over all bad pixels in the 2-D frame. Some bad pixels are propagated to the datacube, where we can use neighboring spatial pixels in the datacube that are far apart in the 2-D frame to fix the remaining bad pixels.

The satellite spots, four fiducial spots created by diffraction of starlight off of a grid and centered on the location of the star (Marois et al. 2006b; Sivaramakrishnan & Oppenheimer 2006), are used for locating the occulted star and calibrating the data photometrically. The GPI DRP measures the position and flux of each satellite spot, and stores them to the header of the datacube. The mean of the positions of the four spots in each wavelength channel is also written into the header as the location of the star at each wavelength. It is up to the various post-processing pipelines whether to use these numbers or to recalculate them. In particular, it is important to measure the flux of the satellite spots and astrophysical sources in the same way to mitigate biases, so often the satellite spot fluxes are recomputed. However, the values stored here are used to calibrate the single-frame contrast curves, so they are important to log.

Stellar PSF Subtraction

Taking all of the datacubes obtained in a sequence on a given target, we use angular differential imaging (ADI; Marois et al. 2006a) and spectral differential imaging (SDI; Marois et al. 2000) to distinguish the light of the star from the light of other astrophysical sources, so that we can model and subtract out the stellar point spread function (“stellar PSF subtraction"). To avoid dependence on a single stellar PSF subtraction algorithm, we use both the classical ADI algorithm (cADI)(Marois et al. 2006a) and `pyKLIP`(Wang et al. 2015), an

Table 3.1: GPI DRP processing steps to make a spectral datacube.

Primitive Name	**Purpose**
Load Wavelength Calibration	Reads in the appropriate wavelength calibration file
Subtract Dark Background	Finds an appropriate dark frame and subtracts it from the raw 2-D image
Update Spot Shifts for Flexure	Corrects for shifts in the microspectra due to instrument flexure(Wolff et al. 2014)
Interpolate Bad Pixels in 2D Frame	Identifies and fixes bad pixels(Ingraham et al. 2014)
Assemble Spectral Datacube	Extracts the data into a 3-D cube using a 3-pixel wide moving box
Interpolate Wavelength Axis	Interpolates the wavelength dimension to be 37 equally-spaced channels
Interpolate Bad Pixels in Cube	Fixes any remaining bad pixels using spatially nearby pixels
Correct Distortion	Corrects for optical distortion. The datacubes are saved after this step, and following steps only modify the header information.
Measure Satellite Spot Locations	Automatically finds satellite spots using a computer vision algorithm and fits their locations with a Gaussian(Savransky et al. 2013; Wang et al. 2014)
Filter Datacube Spatially	Subtracts off a 15-pixel median filter from the image to high-pass filter the image for the following steps. The filtered image is not saved to disk.
Measure Satellite Spot Peak Fluxes	Uses a Gaussian matched filter and the location of the satellite spots in the header to measure the flux of each satellite spot
Measure Contrast	Measures the single-frame contrast. Then it saves the 1-D contrast curve to a FITS file and contrasts at three fiducial separations to the header of the datacube

open-source Python implementation of Karhunen-Loève Image Projection (KLIP)(Soummer et al. 2012; Pueyo et al. 2015), to provide two separate subtractions of the stellar PSF. The cADI pipeline is quick so it is used for real-time reductions, while the `pyKLIP` pipeline is used for our final sensitivity analysis.

Using cADI, we perform two subtractions. The first is a basic ADI reduction where the images are collapsed into broadband images using the mean, high-pass filtered using a Fourier filter with a smooth cutoff frequency of four spatial cycles, subtracted with the mean of the broadband image sequence, rotated North-up, and collapsed in time to form one 2-D stellar-PSF-subtracted image. The second reduction is similar to the first except before collapsing into broadband images, it includes an SDI pre-processing step where in each spectral cube, each frame is subtracted by the median of all frames in the same spectral cube. The latter subtraction is better suited for planets with sharp molecular absorption features in the spectrum, while the former works better for all other cases.

Using `pyKLIP`, we perform three reductions to search for astrophysical signals: one general reduction, one optimized for T-dwarfs with strong molecular absorption features, and one optimized for circumstellar disks. In the first general-purpose reduction, we first high-pass filter the images using an 11-pixel full-width half-maximum (FWHM) Gaussian filter in Fourier space to remove broad features caused by the seeing and wind-shake while minimizing attenuation of high-frequency point source signals of planets. Gaussians with smaller FWHMs were found to subtract out too much point source signal. Then, the images are duplicated once for each spectral channel to be aligned to a common center and magnified using bicubic spline interpolation so that the speckles are aligned at that wavelength (i.e., so that the data at that wavelength do not need to be magnified through interpolation). Then, we break the images into nine annuli that increase logarithmically with separation and break each annulus into four azimuthal sectors. For each frame in a given sector, we build our Karhunen-Loève (KL) modes using the 300 most correlated reference frames where a hypothetical planet at the center of the sector would have moved at least one pixel due to a combination of ADI field rotation and SDI rescaling of the speckles. One pixel, which is about a third of the FWHM of the planet PSF in H-band, was chosen empirically to maximize the SNR of potential planets. Since the number of KL-modes to use varies depending on the dataset, we save images that use 1, 10, 20, 30, and 50 KL-modes to reconstruct and subtract off the stellar PSF. The images are all rotated to be North-up. We then save six separate data products. Five are spectral datacubes, one for each KL-mode cutoff, where we have collapsed all the data in time. The last one is a KL-mode cube where we have collapsed the images in both time and wavelength, leaving the third dimension to be the number of KL-modes used to model the stellar PSF. This cube allow for quick visual inspection of the dataset to determine what the optimal number of KL-modes to use is and to search for planets.

The second reduction optimized for T-dwarfs is similar to the first except for two changes. First, the exclusion criterion of one pixel is weighted so that reference images where the planet is faint due to molecular absorption is preferred. This means that even if the hypothetical planet did not move one pixel due to ADI and SDI, as long as it is faint enough at that wavelength compared to at the wavelength we are trying to perform stellar PSF subtraction, the image will be included in the reference library. Appendix A.1.3 of Ruffio et al. (2017) explains this in quantitative detail and illustrates this selection process for a typical GPI dataset. Second, the KL-mode cube is collapsed in wavelength using a weighted mean where the weights are the fluxes for each spectral channel of a model T-dwarf atmosphere.

The third reduction is optimized for disks and again is similar to the first reduction except for three changes: the images are not high-pass filtered to avoid removing disk structure, the images only use ADI so only images at the same wavelength are used to build the reference data, and more conservative KL-mode cutoffs of 1, 3, 10, 20, 50 are saved.

Contrast Curves

After stellar PSF subtraction, we compute the sensitivity to planets in the dataset. In this paper, we will use the term “contrast" to refer to the flux ratio between the faintest planet we can detect and its star, and the term “contrast curve" to refer to our achieved contrasts in a dataset as a function of projected separation from the star.

For GPIES, we typically are sensitive to planets with L-type and T-type spectra. Due to the strong molecular absorption of methane in T-type spectra, we have improved sensitivity to T-type planets since we can use all the spectral channels where the planet is dim due to methane absorption to be more aggressive in modelling and subtracting out the stellar PSF without significant self-subtraction of the planet itself (see Section 3.2.2). Thus, due to the differing sensitivities, we produce contrast curves for both L- and T-type planets.

For L-type and T-type planets, we use the general reduction and the T-type optimized reduction respectively to estimate the level of the noise in the image. We then cross correlate each reduction with a Gaussian function that has a FWHM of 3.5 pixels to smooth out high-frequency noise and maximize the signal of potential planets, which have PSFs that are roughly Gaussian (Ruffio et al. 2017). This cross correlation map is then used to estimate the noise level as a function of separation by computing the standard deviation of concentric annuli centered on the star. As we have a limited number of independent noise realizations, we correct for small sample statistics assuming the noise distribution is Gaussian (Mawet et al. 2014) and compute the planet brightness that corresponds to the 5σ-equivalent false positive probability of 2.9×10^{-7}, and we take this to be our sensitivity limit (i.e., achieved contrast). Thus, we produce a contrast curve for each planet spectral type that is uncalibrated for flux biases introduced by the data reduction steps described in this section and the previous section.

For each type of planetary spectrum, we then perform two more stellar PSF subtractions using the same parameters as those used in Section 3.2.2, but after injecting simulated planet signals. Simulated planets of that particular spectrum are injected as 2-D Gaussian signals and are used to quantify flux biases due to stellar PSF subtraction. For these simulated planet reductions, we inject planets at nine separations corresponding to the center of each of the nine annuli and at four different position angles so that the simulated planets spiral outward to avoid significantly influencing each other when using SDI. The second simulated planet reduction has the spiral arms offset by 45° with respect to the first. We use the same Gaussian cross correlation routine that was used to estimate the noise level to measure the flux of each simulated planet after stellar PSF subtraction and calibrate flux measurement biases induced by our data reduction process. Then for each point in the contrast curve, we correct for these flux measurement biases. For the contrast within 40 pixels (566 mas), the calibration term is calculated as the linear interpolation between the two closest measured flux calibration factors. Outside of 40 pixels, the correction factor applied is the average correction factor for all simulated planets with separations greater than 40 pixels, as we determined empirically that the calibration factor is constant in this regime to measurement uncertainty ($\sim 20\%$). This results in one T-type contrast curve and one L-type contrast curve for each dataset.

3.2.3 Polarimetry Mode

Constructing Polarimetry Datacubes

Instead of a 37-channel spectral datacube, we use the GPI DRP to generate datacubes where the third axis contains the two orthogonal polarizations of light. These polarimetry datacubes are an intermediate product that we then use for stellar PSF subtraction. The steps performed by the GPI DRP to generate polarimetry datacubes are listed in Table 3.2.

Stellar PSF Subtraction

For polarimetry data, we subtract the diffracted light of the host star using two methods: polarimetric differential imaging (PDI) to look for polarized scattered light from circumstellar dust around an unpolarized star and ADI to look for all scattered light from circumstellar material.

To perform PDI, we use the GPI DRP to execute the steps listed in Table 3.3. This outputs 3-D Stokes cubes where the third dimension contains the four components of the Stokes vectors. We also generate radial Stokes cubes by transforming the Stokes vectors into a radial Stokes basis (Schmid et al. 2006). Since light from most debris disks is only scattered once, and thus has a tangential polarization, the radial Stokes basis typically has all of the polarized astrophysical signal in one Stokes parameter. For PDI, to account for instrumental polarization, we measure and subtract out the apparent stellar polarization by measuring the polarized signal behind the focal plane mask. One advantage of this technique is that it is robust against apparent stellar polarization due to interstellar polarization, unresolved disk structure, or other potential astrophysical polarization sources. Polarization from these sources would also affect the stellar speckles and needs to be subtracted out since we are only interested in the resolved polarized component in GPI's field of view.

To look for an unpolarized astrophysical signal next to the star in our polarimetry data, we sum the two orthogonal polarizations in each polarimetry cube and treat each cube as a single broadband image. Then, `pyKLIP` uses ADI to model and subtract out the stellar PSF using two different sets of reduction parameters. The images are not high-pass filtered before stellar

Table 3.2: GPI DRP processing steps for polarimetry datacubes.

Primitive Name	**Purpose**
Load Polarimetry Spot Calibration	Loads in the appropriate calibration file with the positions of the polarization spot pairs
Subtract Dark Background	Finds an appropriate dark frame and subtracts it from the raw 2-D image
Flexure 2D X Correlation with Polcal	Corrects for instrument flexure by cross correlating the spot calibration with the data and finding the optimal global offset in X and Y
Destripe Science Image	Models and removes vibration-induced microphonics (Ingraham et al. 2014)
Interpolate Bad Pixels in 2D Frame	Identifies and fixes bad pixels (Ingraham et al. 2014)
Assemble Polarization Cube	Constructs a 3-D datacube where the third dimension is the two orthogonal polarizations of light
Interpolate Bad Pixels in Cube	Fixes any remaining bad pixels using spatially nearby pixels
Correct Distortion	Corrects for optical distortion (Konopacky et al. 2014)
Measure Star Position for Polarimetry	Locates the position of the star using a Radon-transform based algorithm (Wang et al. 2014)
Measure Satellite Spot Flux in Polarimetry	Measures the flux of each satellite spot using aperture photometry. The aperture is elongated radially to match the shape of the satellite spots (Hung et al. 2016).
Measure Contrast in Pol Mode	Measures the contrast achieved in total intensity in a single frame of observation (Millar-Blanchaer et al. 2016)

Table 3.3: GPI DRP processing steps for PDI.

Primitive Name	**Purpose**
Accumulate Images	Gathers together all the images in memory and demarcates that the following primitives will be run on a series of images, rather than on each image individually
Clean Polarization Pairs via Double Difference	Removes polarized and unpolarized speckles in the images using data taken with different half-wave plate orientations(Perrin et al. 2015)
Smooth a 3D Cube	Convolves each polarization image with a Gaussian with a FWHM of 1 pixel to suppress pixel-to-pixel noise and improve the noise properties in the final Stokes Cube without sacrificing significant spatial information.
Subtract Mean Stellar Polarization from podc	Uses an annulus inside the occulting mask between 7 and 13 pixels from the star to measure the stellar polarization and subtract it off (Millar-Blanchaer et al. 2016). The results are saved and can be referenced at a later time.
Rotate North Up	Rotates each image so that north is up and east is left
Combine Polarization Sequence	Takes the entire sequence and makes a single 3-D datacube where the third dimension is the four Stokes parameters. This Stokes cube is then saved to disk.
Convert Stokes Cube to Radial	Transforms the Stokes cube into a radial Stokes basis, which is also saved to disk
Get Statistics on Polarimetry Vectors	Generates a histogram of polarization directions across the entire field of view and saves the polarimetry quicklook display

PSF subtraction to preserve extended disk emission. The first reduction, the conservative one, divides the image into seven annuli, uses images where potential astrophysical sources have moved by at least 3 pixels due to ADI to be used to model the stellar PSF, and reconstructs the stellar PSF using 1, 3, 5, 10, 20, and 50 KL modes. The second reduction, the aggressive one, divides the image into nine annuli, uses images where potential astrophysical sources have moved by at least 1 pixel, and reconstructs the stellar PSF using 1, 2, 3, 4, 5, 10, 20, and 50 KL modes. We found that these two sets of parameters provide a conservative and aggressive reduction for studying circumstellar materials. The conservative reduction tries to mitigate stellar PSF subtraction biases on circumstellar material to preserve the diffuse emission and disk morphology. The aggressive reduction allows one to search for faint structure close in to the star.

3.3 The Components of the Automated GPIES Data Infrastructure

To automatically download, index, process, and display all the data from GPIES, multiple components have been integrated together to form the automated data infrastructure. Figure 3.1 provides a schematic of the various parts of the infrastructure, which we will discuss in detail here. Roughly, the infrastructure is divided into data acquisition and storage, data processing, and front-end interfaces. In Section 3.3.1, we will describe the data acquisition and storage: data tools for quicklook quality checking on the summit (Section 3.3.1), the database that stores all of the metadata, target information, and planet sensitivity curves (Section 3.3.1), and Dropbox, which stores the raw and reduced images as well as AO telemetry (Section 3.3.1). Next in Section 3.3.2, we will describe the functionality and software implementation of the Data Cruncher, the automated data processing framework that automatically processes all science and calibration data from GPIES. Quicklook software, including one from the separate TLOCI pipeline (Marois et al. 2014), that creates real time displays of the current observing sequence are discussed in Section 3.3.3. The rest of this section is dedicated to our front-end interfaces. Our main user-facing front-end is the Web Thingie, which hosts views into our database as well as observing tools (Section 3.3.4). We also describe how collaborative tools such as Slack and our internal wiki are integrated into our automated data infrastructure (Section 3.3.5).

3.3.1 Data Acquisition and Storage

A substantial amount of data is being generated by GPIES. About three years into the survey, we have accumulated ∼23,000 raw science files and ∼26,000 raw calibration files. Including reprocessed data, we have generated ∼27,000 quicklook datacubes, ∼5,600 reduced calibration files, ∼80,000 science-grade datacubes, ∼68,000 stellar-PSF-subtracted images, and ∼61 million contrast curve data points. In addition to the science data, we also have ∼73,000 raw telemetry files produced by the AO system to monitor its performance, the observing status of all 600 targets in the survey, and also information on the target stars themselves. To handle all of the data while also making it available to the entire collaboration, we use a combination of Dropbox and a MySQL database to store the data.

Summit Quicklooks and Data Download

After new IFS data are taken on the summit, an instance of the GPI DRP running on the summit uses the GPI DRP Autoreducer module to automatically perform quicklook reductions of the data. These quicklook reductions allow observers to assess data quality and, for science data, provide a contrast curve to demonstrate the sensitivity achieved in a single frame of data for understanding observing conditions. If the data are rendered unusable due to issues like AO control loops opening or wind-shake moving the star off of the coronagraph mask, the observer can log that particular file as bad through the GPItv interface in the GPI DRP. The summit quicklook reductions provide observers the basic tools to assess data quality so that observing can continue in the unlikely case the observatory network connection fails and renders the rest of the data infrastructure ineffective. During an observing sequence, the observer also periodically takes AO telemetry data every 5-10 minutes to allow for further analysis of AO system performance.

To move the data off the summit, automated scripts download the raw and quicklook science data as well as the AO telemetry data to a server located at Cornell University that hosts the MySQL database. While AO telemetry data is downloaded during daytime hours to avoid saturating the network bandwidth during the night, IFS frames are written at a rate of fewer than one per minute, so it is downloaded in real time along with the bad files log without using significant network resources. The server at Cornell University then uploads the data and metadata to Dropbox and the database respectively.

The Database

After receiving new data, the Cornell server adds entries for all of the data into a MySQL database. For all of the science and calibration data taken by the IFS, header information and metadata gets uploaded into the raw and reduced data tables. The raw table contains one row for each raw file, with one column for each of the fields in the file headers (e.g., observing mode, wavelength band, time of observation), along with a column for the unique identifier (ID) for each file. The reduced table contains one row for each reduced file, produced either by the quicklook GPI DRP on the summit or by the Data Cruncher afterwards. The reduced data table contains information after some data processing has happened such as whether the data product is a quicklook or science-grade reduction, the sensitivity achieved at some fiducial separations, flux calibration conversions, as well as a unique ID for the reduced file. To link the reduced data to their original raw data products, a third “Raw2Reduced" table is a two column table where each row associates one raw file ID with one reduced file ID. Multiple rows can specify that multiple raw files went into producing one reduced file. Similarly, multiple rows could contain multiple reduced files that all used a single raw file, but not necessarily exclusively that raw file. Contrast curves associated with a given reduction (either quicklook contrast curves from the summit or Data Cruncher contrast curves) are stored in a contrast curve table where each row is the contrast for a given reduced file at a given separation. Lastly, the bad files log is appended into the observing notes table, which contains notes on each data frame. A row consists of a text comment, whether the bad file flag was enabled, the user that submitted it, and time it was submitted, and the associated raw data file it pertains to. Multiple rows can correspond to a single file to give a history of bad-file marking.

Similarly, the AO data has a raw data table and a reduced data table in a similar fashion as the IFS data, but with AO-relevant metadata stored in each of the tables (e.g., the wavefront sensor frame rate, spatial filter size, and seeing). To make the most out of the AO data, there exist tables that link the raw AO data to reduced AO data and the raw AO telemetry to raw IFS data. The latter table allows AO performance metrics to be linked to final sensitivity to planets.

In the database there also exist tables for GPIES targets. To keep track of the large GPIES target list, we use several target tables. The basic target table contains information on the target such as the celestial coordinates, an estimate for the age of the system, and the target ID number. Then, there is a table with all of the photometry in the literature on each target, a table of name aliases for each target, a table of binary companions associated with each target, a table with spectral energy distribution (SED) fits for each target, a table with the Gemini Observing Tool observing sequence numbers for each target, and a table of observing statuses (e.g., observed, incomplete, candidate companion).

Dropbox

After the data are logged into the database, the raw and quicklook reductions are uploaded onto Dropbox. The raw and summit-reduced quicklook data are put into a read-only directory, as they are not meant to be modified again. After the raw data are synced to Dropbox, the Data Cruncher can then read and process the data automatically. The automatically reduced data from the Data Cruncher also gets synced back onto Dropbox, and then their metadata also stored into the database. All of the AO telemetry is also stored and synced on Dropbox in its own folder.

Sometimes, it is necessary to reduce data by hand, especially when there is an astrophysical source that requires further analysis. In these cases, individuals also upload their own reductions and analysis outputs onto Dropbox so that they can be shared with others. The only restriction is that these data products lie outside of the automatically generated directories, to avoid being erased or moved. This is because after the Data Cruncher reprocesses the entire campaign, all of the automatically generated folders gets moved into an archive section of Dropbox, and new reductions are moved into its place. The archive contains subdirectories, each with an older version of all of the data reduced in the survey by the Data Cruncher.

3.3.2 The Data Cruncher

The Data Cruncher is a Python framework that automates all of the data processing steps detailed in Section 3.2. The Data Cruncher is comprised of two parts: the “Processing Backend" runs the various data reduction pipelines and controls the flow of data through the pipelines, while a series of “Instructors" sends commands to the Processing Backend about what data to process. Figure 3.2 illustrates how the Instructors and Processing Backend fit together to generate data products for the survey. By separating the code between a unit that focuses on the data processing and a unit that focuses on what data to process, we are able to develop a modular and flexible framework. This allows us to scale the Data Cruncher arbitrarily: multiple Instructors can use the same Processing Backend to process data for multiple purposes, or a single Instructor can talk to many Processing Backends when a large amount of data needs to be processed in a parallel fashion.

First, we will focus on the functionality of the Data Cruncher for GPIES. Without delving into the technical details, one can think of the Data Cruncher as being able to produce all of the desired data products given some raw GPI data. The technical software implementation of the Data Cruncher, including the details on the Instructors and Processing Backends, follows in Section 3.3.2 after we have provided some context for the software design.

Real-Time Reductions

When a raw data file is uploaded to Dropbox, the “Realtime Scanner" Instructor is alerted and decides how to instruct the Processing Backend to process the file depending on the context in which the file is taken. For example, if a new raw science frame is downloaded, the Realtime Scanner will just instruct the generation of a datacube from the raw 2-D file and the updating of the quicklook stellar PSF subtractions (described in Section 3.3.3) if the object name in its file header is the same as the previous file's. However, it will also send instructions for the full stellar PSF subtraction of the current data sequence if the object name is different, since this indicates that the observers have moved on to a new target. The Realtime Scanner is able to handle all standard observing procedures for all spectral, polarimetry, and calibration data taken as part of GPIES. The only exception is that the Realtime Scanner is not programmed to handle processing thermal background frames and subtracting them from the entire sequence. However, we find that thermal background subtraction is only necessary in K-band for extended sources such as circumstellar disks, for which we need to distinguish between the large-scale astrophysical emission of the disk from the smooth thermal background. K-band imaging of disks is not a science-goal of GPIES, so this only happens when the Data Cruncher processes queue programs led by GPIES members. In these cases, manual reductions need to be performed.

On a 32-core machine with AMD Opteron 6378 processors clocked at 2.3 GHz, the Data Cruncher generates quicklook reductions within 1 minute of receiving the data, generates the first `pyKLIP` reduction 10 minutes after a sequence finishes, and generates fully-calibrated contrast curves for both L- and T-type planets 1 hour after a sequence is complete.

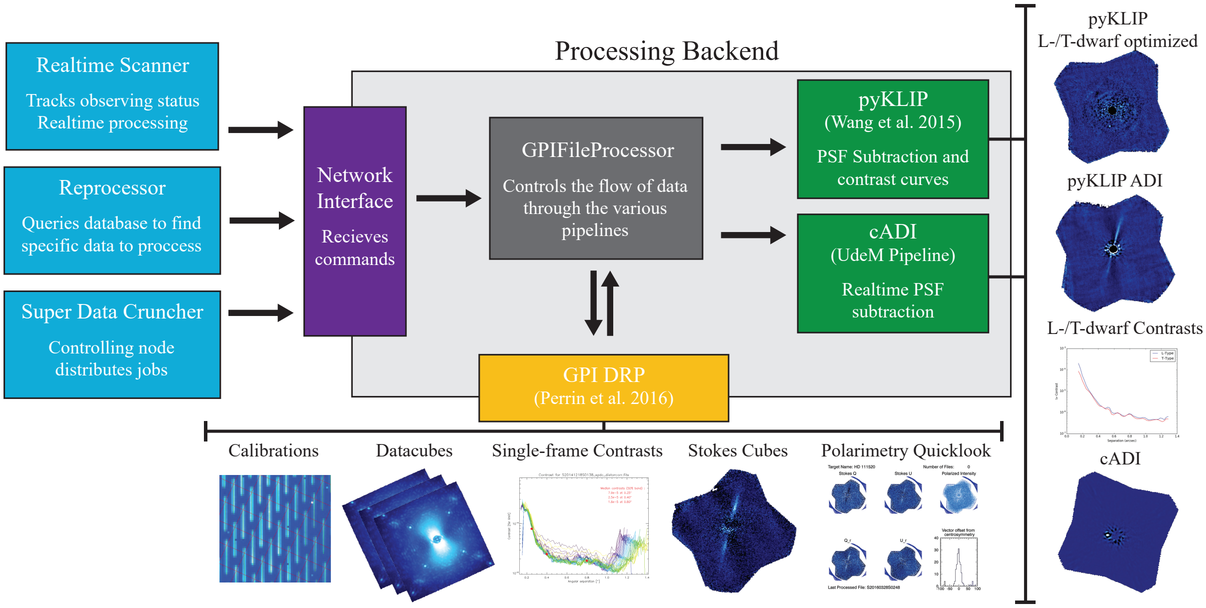


Figure 3.2: Schematic of the Data Cruncher architecture. Boxes represent the different modules of the Data Cruncher, which are discussed in Appendix 3.3.2. The light blue boxes represent the various Instructors. The large light gray box represents the Processing Backend, which consists of several smaller modules. Arrows indicate the flow of data or information from one module to another. Samples of the different data products are shown and grouped by the pipeline that produced them (either the GPI DRP or the stellar PSF subtraction pipelines).

Reprocessing Individual Datasets

Sometimes, a single dataset needs to be reprocessed in order to exclude bad frames or to fix bugs in the reduction pipelines. The Reprocessor module allows us to query the database for certain datasets, find the corresponding files, and process them. For a specified target, the user can request for data taken at a given wavelength band, on a given date, in a given observing mode, and in a given observing program (whether GPIES or an affiliated proposal by GPIES team members) to be processed. The user can also specify whether to process the data completely from raw data, or to use the reduced datacubes and only run specified post-processing algorithms. For example, a user can ask to reprocess all data on a target taken in polarimetry mode when upgrades to the GPI DRP's polarimetry reduction functions happen. A user could also ask to regenerate the contrast curves for a given target at a given band to exclude files where the data quality was poor.

Reprocessing the Entire Campaign

When the Data Cruncher runs on a supercomputer, we call the framework the Super Data Cruncher. Many nodes, each a single computer, are requested, and each node runs one instance of the Processing Backend to parallelize the embarrassingly parallel problem of reprocessing the entire campaign. A single node is designated the controlling node and sends commands to all of the Processing Backends to split up the work. The Super Data Cruncher has successfully run on the Edison and Cori machines operated by the National Energy Research Scientific Computing Center (NERSC) as well as the Comet machine operated by the San Diego Supercomputer Center (SDSC) as part of the Extreme Science and Engineering Discovery Environment (XSEDE; Towns et al. 2014). On Comet, using 30 nodes, with 24 cores per node for a total of 720 cores, we used 16,560 CPU hours, corresponding to 22.9 hours of wall-clock time, to reprocess all survey data up to the end of 2016 (17,008 frames of raw IFS data).

Data Cruncher Software Implementation

Here, we will discuss how we implemented the Data Cruncher framework in software. In general, thread synchronization, network communication, and file system interrupts are key techniques for this implementation.

Processing Backend The Processing Backend manages the data flow through the three data reduction pipelines discussed in Section 3.2 (GPI DRP, `pyKLIP`, and cADI). The Processing Backend is broken up into Python threads that each manage one task and are depicted as the boxes inside the Processing Backend box in Figure 3.2. The threads communicate with each other by passing messages through queues and use monitors (implemented using the `Lock()` and `Condition()` objects available in Python's `threading` library) as the synchronization construct to block threads that are waiting for new messages to be passed into the queue. For example, after one pipeline finishes processing some intermediate data products, it can send a message about these data products to another thread via their communication queue, alerting this second thread, which was previously sleeping since it had nothing to do, of the new data to process. The Processing Backend has four different kinds of threads: a network interface thread, the GPI DRP thread, stellar PSF subtraction threads, and the `GPIFileProcessor`, the backbone that manages the data flow.

The network interface thread receives commands from Instructor processes. Instructors pass messages that contain the file or files to be processed, the output directory, whether the input data are raw 2-D images or 3-D datacubes, what kind of data it is (i.e., spectral, polarimetry, or calibration data), and optional parameters such as which stellar PSF subtraction algorithms to use. There are three network interface threads available: web socket, Message Passing Interface (MPI), or a regular queue. These threads are mutually exclusive so only one thread ever runs. The web socket thread is used in most contexts, except when run on a supercomputer. The network interface's main purpose is to parse the received messages and send the messages to the `GPIFileProcessor`.

The `GPIFileProcessor` thread manages all of the various pipeline threads. It receives messages from the network interface and parses the instructions into work units that need to be passed into the various pipelines. It then checks against the database to ensure that the files are not marked as bad, discarding instructions for bad data as it goes along. All good, raw, 2-D data are sent to the GPI DRP for data processing.

Because the GPI DRP was not designed to be fully automated to this degree and does not fully expose all of its features programmatically, a few workarounds were required. The GPI DRP uses a recipe directory, where each job is an XML file detailing the data reduction steps that need to be executed on some data. These recipe files exist in three states: waiting to be executed, in the processes of being executed, and finished being executed (either successful or failed). The `GPIFileProcessor` writes XML recipe files into the queue directory, which queues a job for the GPI DRP. The recipe for science data are custom recipes specified in Section 3.2. For calibration files, an instance of the GPI DRP Data Parser is created, and the Data Parser generates the appropriate recipes to be written into the queue.

Once a recipe file is written into the queue, we need to identify when the GPI DRP finished that recipe so we can then run the appropriate post-processing algorithms. The GPI DRP thread's sole purpose is to track this. When the GPI DRP finishes processing a recipe, it updates the recipe file to indicate whether the recipe succeeded or failed. Thus, the GPI DRP thread uses the Python `watchdog` package to receive file system interrupts. When a recipe file has been updated, the GPI DRP thread is awakened to run a function that notifies the `GPIFileProcessor` that data has been processed.

Except in the case of quicklook reductions, the `GPIFileProcessor` waits on the condition that all raw data are finished being processed by the GPI DRP, regardless of whether all files were successfully processed, before running stellar PSF subtraction. When it is notified by the GPI DRP thread that all queued raw files have been processed, then it passes jobs to the stellar PSF subtraction threads. Generally, there are two stellar PSF subtraction threads that run in parallel: one for `pyKLIP` and one for cADI. Both threads run their respective pipelines as subprocesses rather than subthreads to avoid the Python global interpreter lock and to improve memory efficiency by relying on the operating system to free memory rather than the Python garbage collector. As cADI does not require many computational resources, it is practical to run the two threads in parallel. The cADI thread runs two reductions as discussed in Section 3.2.2. The pyKLIP thread receives eight jobs for each spectral mode dataset: three different stellar PSF subtractions optimized for different science objectives as discussed in Section 3.2.2, four reductions to inject and recover fake planets to calibrate out flux biases induced by the data processing, and one calculation of the contrast curve as detailed in Section 3.2.2. These work units are broken up and prioritized so that if multiple datasets have stellar PSF subtractions queued up, a single KLIP reduction for each dataset is prioritized first, allowing for a fast initial look in real time during an observing night. For polarimetry mode data, two `pyKLIP` reductions (as discussed in Section 3.2.3) are queued up to look for total intensity disk signal. PDI reductions are also queued into the GPI DRP queue to subtract the unpolarized stellar signal to look for polarized astrophysical emission.

For quicklook reductions, only cADI is used for spectral mode data and the GPI DRP still performs the quicklook PDI reductions. Quicklook reductions are specified with a flag in the instructions received by the network interface. Upon seeing this flag, the `GPIFileProcessor` does not wait for all raw files to be finished by the GPI DRP. Rather, the goal is to have the quicklook reduction done as fast as possible in real time with however many files have already been processed. Because of this, when multiple files are downloaded by the Data Cruncher at once due to latency in the file syncing, duplicate quicklook stellar PSF subtraction jobs are created. To avoid having the same instruction multiple times in the stellar PSF subtraction queue to improve efficiency, the queue through which jobs are passed to the stellar PSF subtraction threads disallows duplicates.

Instructors Currently, we use three different Instructor interfaces that send commands to the Processing Backends. Multiple Instructors can talk to the same Processing Backend, and to practically handle this, we do not leave web sockets open, closing them immediately after sending instructions so that another Instructor is able to connect without timing out. Often times, the Instructor and the Processing Backend live on the same machine as the Instructors do not consume much computing resources.

The Realtime Scanner module handles all of the real-time processing. It uses the Python `watchdog` module to receive alerts from the operating system when new data are synced to the computer and written to disk. Upon being alerted of a new file, the Realtime Scanner decides how to process the file depending on the context in which the file is taken. To keep track of the state of observing, the logic inside of the Realtime Scanner is implemented as a finite state machine. The finite state machine logic handles almost all standard observing procedures for spectral, polarimetry, and calibration data taken as part of GPIES. The only exception are thermal background frames taken in K-band.

The Reprocessor Instructor is a series of Python functions that can be called on demand to perform individual tasks such as reprocessing a single dataset on a target or processing a list of raw files. Each function uses the `mysql` Python library to query the database for the desired files requested by the user and sends commands to the Processing Backend to process the files appropriately. The Reprocessor can also query for all data from the campaign, generate instructions to reprocess all of the raw data from scratch, and save the commands to text files to be uploaded onto a supercomputing cluster that doesn't have direct database access.

The third Instructor is the controlling node for the Super Data Cruncher, which is what the Data Cruncher is called when it runs on a supercomputer. Each node on the supercomputer runs one instance of the Processing Backend. The previous two instructors use web sockets to communicate to the Processing Backend, but in supercomputer clusters, MPI is the network interface of choice. The exception is the controlling node itself, which uses a simple queue interface to pass instructions to its own Processing Backend, avoiding the overhead of MPI. The controlling node reads instructions that have been pre-generated in a text file to avoid setting up a connection to the database from the supercomputing center. It then distributes the instructions across all of the nodes with a granularity level of a full sequence on a single target.

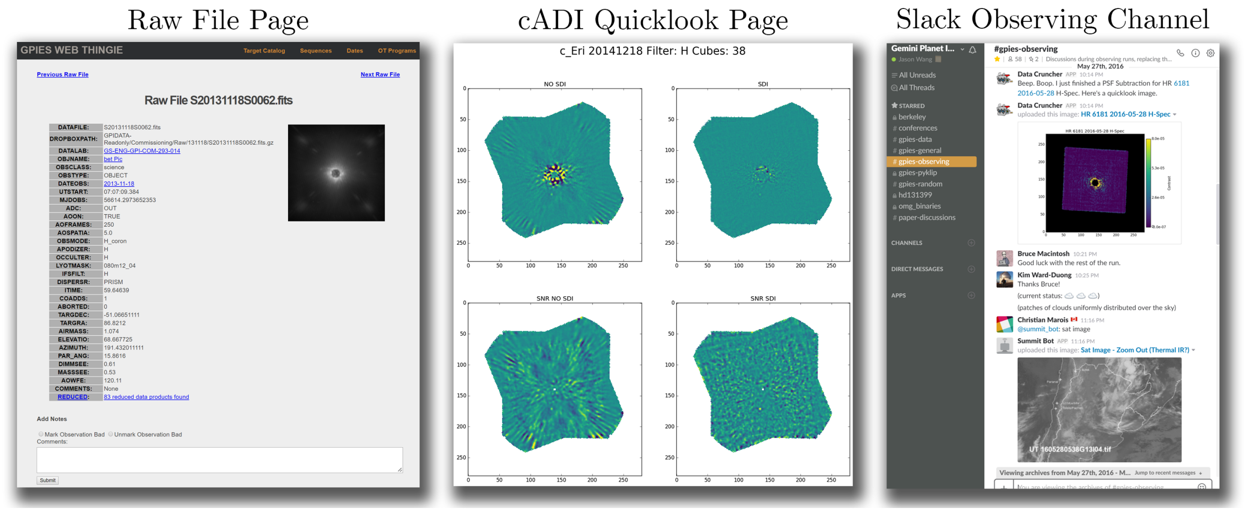
Practically, the reprocessing of the entire campaign is done in two phases: once for the raw 2-D data to make datacubes, and once to run stellar PSF subtraction on the datacubes. As processing the raw data requires running the GPI DRP, and thus requiring IDL licenses for each node, we typically only use ~20 nodes to do this step. Afterwards, a script is run to quality check the reductions, ensuring that calibration issues like flexure offsets are handled properly. Then, the Super Data Cruncher runs stellar PSF subtraction and contrast curve generation for all datasets. Due to IDL licensing issues, only a small subset of nodes are designated to run all of the cADI reductions, but since the cADI reductions are fast to compute, this does not slow down the reprocessing. `pyKLIP` stellar PSF subtraction and contrast curve generation is evenly distributed across all of the nodes.

3.3.3 Quicklooks

During the observing night as data are taken and reduced in real time, quicklook tools translate the data into easy-to-view images that are constantly refreshed as new data comes in. Currently, we run three quicklook tools, which update every minute as new data are taken. Each quicklook tool syncs images to a certain location on Dropbox so that it can be picked up and displayed on web pages on the Web Thingie.

The first is the cADI quicklook tool. The Data Cruncher produces the two cADI reductions of spectral mode data described in Section 3.2.2 every time new frames are taken and uploaded to Dropbox. The cADI quicklook tool takes these reductions and creates a four-panel plot to display stellar-PSF-subtracted images (Figure 3.3). The top two images of the plot are just the data without any further processing. The bottom two images are the signal-to-noise ratio (SNR) maps of the data after convolving the data with a Gaussian function and weighing each pixel by the inverse of the noise in pixels of similar separations from the star.

The second is the TLOCI quicklook display, which uses stellar-PSF-subtracted spectral mode data from the TLOCI pipeline (Marois et al. 2014), independent of the Data Cruncher.


Raw File Page

cADI Quicklook Page

Slack Observing Channel

Raw File S20131118S0062.fits

DATAFILE:	S20131118S0062.fits
DROPBOXPATH:	GPIDATA-Readonly/Commissioning/Raw/131118/S20131118S0062.fits.gz
DATALAB:	GS-ENG-GPI-COM-293-014
OBJNAME:	bet Pic
OBSCLASS:	science
OBSTYPE:	OBJECT
DATEOBS:	2013-11-18
UTSTART:	07:07:09.384
MJDOBS:	56614.2973652353
ADC:	OUT
AOON:	TRUE
AOFRAMES:	250
AOSPATIA:	5.0
OBSMODE:	H_coron
APODIZER:	H
OCCULTER:	H
LYOTMASK:	080m12_04
IFSFILT:	H
DISPERSR:	PRISM
ITIME:	59.64639
COADDS:	1
ABORTED:	0
TARGDEC:	-51.06651111
TARGRA:	86.8212
AIRMASS:	1.074
ELEVATIO:	68.667725
AZIMUTH:	191.432011111
PAR_ANG:	15.8616
DIMMSEE:	0.61
MASSSEE:	0.53
AOWFE:	120.11
COMMENTS:	None
REDUCED:	83 reduced data products found

Add Notes

○ Mark Observation Bad ○ Unmark Observation Bad

Comments:

Submit

#gpies-observing

58 | 2 | Discussions during observing runs, replacing th...

May 27th, 2016

Data Cruncher APP 10:14 PM
Beep. Boop. I just finished a PSF Subtraction for HR 6181 2016-05-28 H-Spec. Here's a quicklook image.

Data Cruncher APP 10:14 PM
uploaded this image: HR 6181 2016-05-28 H-Spec

Bruce Macintosh 10:21 PM
Good luck with the rest of the run.

Kim Ward-Duong 10:25 PM
Thanks Bruce!
(current status: ☁ ☁ ☁)
(patches of clouds uniformly distributed over the sky)

Christian Marois 11:16 PM
@summit_bot: sat image

Summit Bot APP 11:16 PM
uploaded this image: Sat Image - Zoom Out (Thermal IR?)

Viewing archives from May 27th, 2016 - M...

You are viewing the archives of #gpies-observing

Figure 3.3: Example screenshots of the user-facing web front ends that are integrated with the data processing infrastructure. *Left*: A dynamically generated page from the Web Thingie for one single raw data file. This is one of multiple types of pages the Web Thingie can generate. *Center*: The real-time cADI quicklook display that is automatically generated and updated in real time on the Web Thingie. *Right*: An example of the observing support provided by the Slack chatbots. Even though the weather was poor, the Data Cruncher continued to produce science-grade stellar-PSF-subtracted data and the Summit bot provided updates on the weather conditions as requested.

The TLOCI pipeline typically uses the quicklook spectral datacubes made on the summit so that it does not rely on the Data Cruncher for datacubes, but can be modified to point at the spectral datacubes made by the Data Cruncher. The TLOCI quicklook display offers four plots, two for L-dwarf reductions and two for T-dwarf reductions. For each spectral type, there is a reduction done by a quick simple code, and a slower, but more sophisticated code. TLOCI quicklook display also offers a simple planet detection code and quicklook astrometry and spectra of candidates flagged by the planet detection code.

The last is the polarimetry quicklook display generated by the Data Cruncher through the GPI DRP. This updates after every four images (corresponding to a full cycle of waveplate positions) and displays images from both the linear and radial Stokes cubes produced by the Data Cruncher that are described in Section 3.2.3.

3.3.4 The Web Thingie: Web Front End for Database and Observing Tools

Between all of the raw, reduced, calibration, and AO telemetry data generated by the GPIES campaign, it quickly becomes hard to track down the specific files taken on specific dates, and the associated reductions. Instead of requiring users to make SQL queries to the database themselves when they need to look for data, we have created a user-friendly web front end, named the Web Thingie, to allow users to find data, mark files as bad, pick targets, and look at real-time stellar PSF subtractions.

The Web Thingie is run on an Apache-based web server gateway interface using Flask, a Python, Werkzeug-based web development framework. It uses Jinja2 to dynamically generate web pages from database queries. For example, instead of static web pages, the pages that list all raw files taken on each date are generated from one template page with the date populated by the user requested date and rows of a raw files table populated from a database query. The Web Thingie can a page for any individual raw or reduced file stored on Dropbox, for raw files taken on a specific date, for raw files corresponding to a specific object, for raw files of a given calibration file type, for a calendar indicating dates under each month in which GPI data was taken, among others.

Typically users are interested in data taken on a specific date or of a specific object, so going to those pages will list all data from that date or object including the instrument configuration, wavelength, time, and the data quality of the file. Once they do click on an individual raw file, a web page following the same template as the one shown in Figure 3.3 is generated. There is a thumbnail image of the raw data frame, a list of important header keywords, a link to a page that lists all the reduced data products that used this raw data frame, existing notes on this data frame, and the ability to add more notes and mark whether this frame is bad or not. If the data frame is bad, a large text banner in red font on the top of the page indicates that. When the user clicks through to the list of all the reduced data, they can scroll through and click on the desired reduced file to go to its page for more information on the reduction.

Each target also gets its own page to help observers during a run. On the top of the page, there are links to the SIMBAD page for the target, the internal wiki page for the target, and the finder chart for crowded field stars. Then, the status of the target is listed regarding whether spectral or polarimetry data needs to be taken. Basic information on the star as well as binary companions and their separation and position angle are listed next, to help observers pick the right star at the telescope. The page ends with a series of data quality plots, which are generated dynamically. The first set of plots shows single-frame contrast, AO wavefront error, and satellite spot flux as a function of image number for all observing sequences on this target. These metrics indicate how conditions change over an observing sequence and allow for comparison to previous observing sequences on the same target. The second set of plots displays the histograms of all single-frame contrasts for all targets in the database, all targets within 0.25 magnitudes, and all previous data from the same star. It also displays the 16th, 50th, and 84th percentile single frame contrasts for all targets within 0.25 magnitudes. This allows observers to quickly compare the quality of the data from the summit quicklook reductions that they are currently taking and see how they compare to the rest of the campaign, allowing for a straightforward assessment of data quality.

To maintain target selection flexibility, dynamically generated target selection tools are also available on the Web Thingie. At any given time of the night, a dynamic target list page automatically orders targets by a parameterized score (McBride et al. 2011; Savransky et al.

2013) that is a combination of the inherent proprieties of the system (e.g., age, distance) as well as how easily observable the target is (e.g., targets that transit directly overhead need to be observed just before transit for maximum field rotation for ADI). The list can also be trimmed to only look at bright stars, or to avoid targets that require the telescope to point in the current direction of the wind. The list also flags whether polarimetry observations are needed, alerts observers to binary stars with colorfully flashing star icons to ensure that the correct star is observed, and compares how much field rotation is achieved observing it immediately versus the maximum achievable field rotation in a one-hour observing sequence. Additionally, any target can be plotted on the “Nightly Planning Tool" page, which plots the elevation, cumulative field rotation, and instantaneous field rotation over time to determine the feasibility of observing the target at any given time. An offline target list also exists in case the Web Thingie is ever inaccessible.

Lastly, the Web Thingie hosts three quicklook stellar PSF subtraction monitors to allow observers to search for candidates and assess stellar-PSF-subtracted data quality in real time as discussed in Section 3.3.3.

3.3.5 Collaborative Tools

To streamline collaboration across an international team, a series of tools are used. To improve their utility, we have integrated some of these tools with the automated infrastructure.

Wiki

An internal wiki run on the Dokuwiki platform is used for documentation. The wiki is used as a knowledge-base for information like troubleshooting the instrument. For data analysis, a target page is automatically generated for each target, with an observing summary of that target automatically generated by the Data Cruncher. Users can then post additional data analysis numbers, such as astrometry and photometry of candidates, as well as summaries of the current state of data analysis.

Slack

The messaging application Slack is used to communicate efficiently between collaborators located all across the world. The “`#gpies-observing`" channel allows for focused discussion during the current observing run. Slack also has a programmatic API, allowing for the development of software integrations. As shown in Figure 3.3, The Data Cruncher has a chatbot interface (source code available at Github) which posts `pyKLIP` reductions as it finishes them to `#gpies-observing`. When messaged, it also can retrieve images of previous `pyKLIP` reductions on GPIES targets, tell the time in any time zone, provide sunrise and sunset times for Gemini South, and display the current moon phase using an emoji. For telescope and observing information, a “Summit Bot," also shown in Figure 3.3, was forked from the Data Cruncher chatbot to run on the Gemini South summit computers and thus can access the state of the telescope, weather monitors, and cloud cameras, which have proved extremely useful for remote observing when no one from the GPIES team is on the summit. Together, the two chatbots assist in taking observations and analyzing data, reducing the manpower needs for executing the survey. They also allow GPIES members outside of the observing team to easily stay updated on observing status: within half an hour after an observing sequence on a target is complete, GPIES members can look at the fully reduced data the Data Cruncher chatbot posted onto Slack, without anyone in GPIES having to do any data processing.

3.4 Utility of the Data Infrastructure

A fully automated data architecture from start to finish brings several benefits to GPIES regarding observing efficiency, data processing, data archiving, instrumental performance, and survey statistics.

3.4.1 Tools for Observers

A whole suite of tools has been developed to inform observers in real time, reducing the complexity of decisions that need to be made during the night. Before and during the night, the Summit Bot on Slack can be messaged to display environmental monitors and webcams from the summit. This allows the observers to look at the current time series of data from seeing monitors as well as look at in which directions in the sky there might be clouds. Since all of this data appear on the `#gpies-observing` Slack channel, members of GPIES outside from the core observing team can chime in, providing advice and improving the cohesion of the team. Once conditions are good for observing, the automated target selection tool on the Web Thingie allows for flexible scheduling of targets, giving observers the highest priority target at any given time, after accounting for pointing restrictions. Once a target is picked, and data are being acquired, the GPI DRP on the summit generates quicklook datacubes to check for star alignment and image quality. Obviously bad data, such as when the AO control loops open, can be marked as bad in the GPI DRP and the bad file flags will be propagated into the database. Quicklook single-frame contrasts for the new data can be compared to the histogram of all single-frame contrast of similar targets. This allows observers to understand in what percentile the current observing conditions are, since single-frame contrasts are the best predictors of final stellar-PSF-subtracted contrast for a whole dataset (Bailey et al. 2016).

As we accumulate frames on a target, the quicklook stellar PSF subtraction pages on the Web Thingie update in real-time. This is particularly useful when following up known systems or candidate planets. Because of different observing conditions affecting AO performance and stellar PSF stability, the exposure times to achieve a desired SNR for confirmation or characterization of astrophysical sources will vary. The quicklook reductions, even though they do not achieve the best stellar PSF subtraction, give us an excellent measure. If a candidate companion does not show up in the quicklook tools after a standard one-hour follow-up sequence, observers can decide to add an extra hour to the follow-up observations. Contrarily, if conditions are excellent and the quicklooks already detect the source at high SNR, then observers can choose to cut a long sequence short, and move on to other targets.

After a sequence is complete, observers can use the Web Thingie to change the target status flag, flagging it as complete, incomplete, or requiring follow-up of candidates. Typically, ten minutes after a sequence is complete and with no human intervention, the Data Cruncher Slack bot posts the final `pyKLIP` stellar-PSF-subtracted data onto `#gpies-observing`, allowing for convenient viewing of the previous dataset and discussion of possible candidates in the data.

3.4.2 Data Processing Speed and Consistency

To make inferences from a large survey, it is important that the data are processed uniformly with the latest calibrations and bug-fixes. With such a large amount of data over several years, an automated architecture to handle all of the various data types and data processing needs can mitigate errors and reduce the workload, particularly for junior members of the GPI collaboration who often are the ones reducing the data.

Using the Data Parser functionality of the GPI DRP, the Data Cruncher is able to identify and process all calibration files taken before or during an observing run. Without the need to trigger human intervention to process these data, we always have the latest calibration files processed before each run and available on Dropbox. This also means that there is no need to reprocess the data again immediately after the run. The real-time reductions generated by the Data Cruncher are of science-grade and are used in papers.

Because of this, we have an extremely quick turnaround when we do detect interesting objects. For example, when GPIES discovered HR 2562 B (Konopacky et al. 2016b), we could watch in real time using the quicklooks on the Web Thingie as the companion appeared in our data over the course of an hour. Preliminary analysis of the companion indicated a peaky H-band spectrum indicative of a bound, L-type companion rather a background star, which likely has a monotonically smooth spectrum. We had confidence in our preliminary analysis since the wavelength calibration and spectral datacubes produced by the Data Cruncher were of science grade. We knew we wanted to follow it up immediately to characterize the spectrum at other wavelength bands, and we did so at the next possible opportunity during the same observing run. Similar situations have happened with background objects, for which spectra and astrometry were extracted within hours of observation and candidates were identified as background objects within the same night by comparing with archival data. The fast turnaround time on our data processing means we can quickly identify and prioritize the most interesting objects found by GPIES. This becomes important when planets are rare and weather conditions are bad.

When upgrades or bug-fixes to the various pipelines occur, the Super Data Cruncher is able to reprocess the entire campaign on a supercomputer within a single day. This way, we can maintain a homogeneous reduction process for all data, and observations collected at different times are not subject to varying systematic errors associated with changes in the reduction process. Eliminating these systematic bugs in a uniform manner is necessary when performing statistical analysis on the campaign.

3.4.3 Large-Scale Data Analysis

A large dataset that is uniformly processed enables a wide range of analyses that allow us to better understand planets, our algorithms, and GPI itself. A critical final goal of the survey is to place limits on the occurrence of giant planets at Solar System scales by comparing the number of planets imaged to the sensitivity to planets based off contrast curves generated by the Data Cruncher. Because the contrast curves are also stored in the database, both single-frame and final stellar-PSF-subtracted contrasts can be correlated with target and atmospheric keywords stored in the data headers. This allows us to understand instrument performance by linking metrics generated by the AO system to metrics expressing planet sensitivity generated by the Data Cruncher (Poyneer et al. 2016; Bailey et al. 2016). For example in Bailey et al. (2016), we were able to construct the histogram of GPI performance for stars of different brightness as well as show GPI's AO correction is dependent on the speed of atmospheric turbulence rather than just seeing. As another example, the measurements of the apparent stellar polarization obtained throughout GPIES can be used to improve the characterization of the instrumental polarization over previous methods (Millar-Blanchaer et al. 2017).

The Data Cruncher cannot satisfy every data analysis need, since there are always additional analyses to be done. While rarely are datacubes regenerated, manual stellar-PSF-subtractions are done for scientifically interesting objects to optimize the reduction for that object. Additionally, as new planet detection algorithms are developed, such as Forward Model Matched Filter (FMMF; Ruffio et al. 2017), they are run using the uniformly processed datacubes made by the Data Cruncher to characterize algorithm performance. While the Data Cruncher is not the solution to all data reduction needs, it offers a solid infrastructure for others to run analysis on.

Also, because all data from GPIES are processed uniformly and available on Dropbox in an organized fashion, publicly releasing GPIES data after the campaign will be straightforward as sharing the link to the Dropbox folder containing the most up-to-date reductions from the Data Cruncher.

Most importantly, all of this data analysis is possible without a large effort by many members of GPIES reducing the data by hand. A standard 1 hour GPIES observing sequence takes 1 hour to manually generate all of the reduced data products. If follow-up observations are included, ∼430 datasets have been observed, so assuming an eight-hour workday, it would have required ∼54 days worth of work to manually reduce all of the data taken so far once. Reprocessing the campaign would be incredibly time-consuming since each reprocessing would take just as many work hours. Instead, the processing and reprocessing of GPIES data is accomplished with minimal time investment by humans. This has allowed GPIES members to focus on higher-level analysis and writing papers rather than processing data.

3.5 Future Steps

The successful design and implementation of the automated data architecture has been one of the major accomplishments of the GPIES collaboration. A few additional features remain desirable. Our top priority is to transition from planet detections by-eye to the FMMF algorithm, which can automatically flag candidate companions with low false positive rates (Ruffio et al. 2017). Automated application of FMMF has not been implemented, partially due to the substantial computation cost of the current version of the algorithm. Simpler planet detection algorithms have been implemented and tested in the Data Cruncher and TLOCI pipeline, but these create more work than they save because they flag too many false positives. After an automated planet detection algorithm is implemented into the Data Cruncher, automated astrometry and spectrophotometry of candidates will follow. However, given the rate of planet candidate detections (a few per observing run), it has not been an efficient use of time to implement all of these features.

A natural application of the GPIES architecture would be to other direct imaging surveys, past, present, and future. Given the similar data products and data processing needs, it would be relatively straightforward to adapt many of our tools for these surveys. Using such an infrastructure to uniformly process all the data would ensure archival data is processed with the latest stellar PSF subtraction algorithms and contrast curves are removed of biases between data reduction pipelines when combining results from multiple surveys.

Although GPIES has very specific data processing and technical needs, some of the infrastructure from GPIES could be broadly applicable to other surveys of similar scales, where the data volume is large enough that doing everything by hand is impractical, but small enough that it lacks a dedicated team of professionals to manage all of the data. Pieces of the infrastructure or general principles could be applicable to other surveys. Much of the architecture described here can credit its success to the integration and design of the entire system. Features such as linking all reduced data products to raw data products to instrumental and weather condition keywords in a database can useful for other surveys, but perhaps in different forms (e.g., there might be several intermediate data products that need to be tracked). Another area to emphasize are the real-time processing and observing tools that reduce the complexity of observing and streamline optimization of observing strategy. Aggregating observing statistics and providing real-time observing conditions in convenient ways such as a website or Slack can benefit other surveys too.

In the mean time, the GPIES team will enjoy the automation of the data processing for the rest of the survey. The lack of updates to the Data Cruncher recently (∼2 commits a month on average to our Git repository in 2017) is evidence that the infrastructure is running smoothly, and that the team is reaping the benefits of the work put into the automation.

Acknowledgements

The Gemini Observatory is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), Ministério da Ciéncia, Tecnologia e Inovaçāo (Brazil), and Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina). This work was supported in part by NASA's NExSS program, grant number NNX15AD95G. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. Support for MMB's work was provided by NASA through Hubble Fellowship grant #51378.01-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. Portions of this work were performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

We thank the two anonymous referees for their thorough review and suggestions. We thank the NERSC and SDSC staff for their helpful support in providing computational resources and technical help. We also thank Barry Mieny for granting us permission to use the database icon used in Figure 3.1. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013a).

Chapter 4

The Orbit and Transit Prospects for β Pictoris b constrained with One Milliarcsecond Astrometry

A version of this chapter was first published in *The Astronomical Journal*: Wang et al. (2016).

A principal scientific goal of the Gemini Planet Imager (GPI) is obtaining milliarcsecond astrometry to constrain exoplanet orbits. However, astrometry of directly imaged exoplanets is subject to biases, systematic errors, and speckle noise. Here we describe an analytical procedure to forward model the signal of an exoplanet that accounts for both the observing strategy (angular and spectral differential imaging) and the data reduction method (Karhunen-Loève Image Projection algorithm). We use this forward model to measure the position of an exoplanet in a Bayesian framework employing Gaussian processes and Markov-chain Monte Carlo (MCMC) to account for correlated noise. In the case of GPI data on β Pic b, this technique, which we call Bayesian KLIP-FM Astrometry (BKA), outperforms previous techniques and yields 1σ errors at or below the one milliarcsecond level. We validate BKA by fitting a Keplerian orbit to 12 GPI observations along with previous astrometry from other instruments. The statistical properties of the residuals confirm that BKA is accurate and correctly estimates astrometric errors. Our constraints on the orbit of β Pic b firmly rule out the possibility of a transit of the planet at 10σ significance. However, we confirm that the Hill sphere of β Pic b will transit, giving us a rare chance to probe the circumplanetary environment of a young, evolving exoplanet. We provide an ephemeris for photometric monitoring of the Hill sphere transit event, which will begin at the start of April in 2017 and finish at the end of January in 2018.

4.1 Introduction

Astrometry is an essential tool for characterizing directly imaged exoplanets and their physical relationship to other elements of the planetary system in which they reside. The Gemini Planet Imager (GPI; Macintosh et al. 2014) was designed with a goal of achieving ≤1.8 mas astrometric accuracy (Graham 2009), which is necessary for characterizing the eccentricity distribution of exoplanet orbits from the GPI Exoplanet Survey (Konopacky et al. 2014). To do so, the astrometric calibration of GPI has continually been benchmarked to well calibrated astrometric fields (Konopacky et al. 2014). This had led to some of the most precise astrometry on directly imaged exoplanet systems to date (Millar-Blanchaer et al. 2015; De Rosa et al. 2015; Rameau et al. 2016), allowing us to constrain or fit the first ever orbit of some of these directly imaged exoplanets. However, limited by either the signal-to-noise ratio (SNR) of these exoplanets or by biases in the various data analysis algorithms, so far no astrometric study with GPI has achieved the design goal of 1.8 mas precision.

The importance of understanding planetary orbits is highlighted by the β Pictoris system, a young ($\sim$ 23 Myr; Mamajek & Bell 2014; Binks & Jeffries 2016) and nearby (19.3 pc; van Leeuwen 2007) system that has been extensively studied. β Pic harbors a near edge-on debris disk that was first imaged by Smith & Terrile (1984) and which was subsequently observed to have a warp thought to be induced by a planet whose orbit is inclined relative to the debris disk (Burrows et al. 1995; Mouillet et al. 1997; Heap et al. 2000). Additional indirect signatures of a planet were derived from variable spectral features modeled as infalling comets (Beust & Morbidelli 2000) and a peculiar light curve anomaly detected in 1981 (Lecavelier Des Etangs et al. 1997).

Lagrange et al. (2009, 2010) then discovered β Pic b, a planet at an appropriate mass ($\sim 10\ M_{\mathrm{Jup}}$) and semi-major axis ($8 - 13$ AU) to be responsible for the previously observed indirect signatures of planets. A key focus of subsequent observations was determining the alignment of β Pic b relative to the main outer disk and the warp to determine if β Pic b is causing the warp. By observing the disk and planet simultaneously, Lagrange et al. (2012) concluded that the planet is misaligned from the main outer disk and consistent with being responsible for the warp. Additionally, Dawson et al. (2011) ruled out the possibility of having another giant planet in the system massive enough to cause the warp instead. Thus, β Pic b is responsible for the warp in the debris disk. This was confirmed in astrometric monitoring campaigns (Chauvin et al. 2012; Nielsen et al. 2014; Millar-Blanchaer et al. 2015) which used homogeneous datasets to limit systematics and constrain the orbit of β Pic b.

Refining the orbital elements of β Pic b is not only crucial for investigating the dynamical link between the planet and the disk warp, but also because β Pic b may transit its host star once every ~20 years. To date, there are no other known systems where the physical properties of an exoplanet can be characterized by using both the direct imaging and transit techniques. Currently, the tightest constraint on the probability of transit is ~0.06%, obtained with a dedicated astrometric monitoring campaign with GPI (Millar-Blanchaer et al. 2015). However, Lecavelier des Etangs & Vidal-Madjar (2016) point out that GPI measurements from Millar-Blanchaer et al. (2015) have a higher position angle (PA) than astrometry from previous measurements, which could arise from a possible systematic calibration offset between GPI and previous instruments instead of actually due to β Pic b's orbit being slightly inclined away from edge on. We note there is currently no evidence of a PA offset with the GPI astrometric calibration, and the GPI astrometry of HD 95086 b are consistent with previous astrometry from other instruments (Rameau et al. 2016). However, it is important to more accurately compute the orbit of β Pic b because in late 2017 to early 2018 (Millar-Blanchaer et al. 2015), it will be at its closest projected separation from the star. The transit of the planet and/or any circumplanetary material orbiting around it could therefore be detectable.

One of the obstacles in characterizing directly imaged exoplanets is that even with the newest instrumentation, the glare of the host star covers the signal of the planet. In order to subtract the point-spread function (PSF) of the star and maximize the SNR of the planet, observing techniques such as Angular Differential Imaging (ADI; Marois et al. 2006a) and Spectral Differential Imaging (SDI; Marois et al. 2000) and data reduction algorithms like Karhunen-Loève Image Projection (KLIP; Soummer et al. 2012; Pueyo et al. 2015) are used in combination to disentangle the PSF of the star from potential astrophysical sources. However, these techniques distort the planet signal and create data reduction artifacts, which are usually nuisance parameters that need to be calibrated out to obtain unbiased astrometry.

Forward modeling effects of observing techniques and data reduction algorithms on the PSF of the planet was first done in the context of ADI and LOCI, showing significant improvements in astrometry and photometry for simulated planets (Marois et al. 2010a; Galicher & Marois 2011). In similar contexts with ADI and LOCI, Brandt et al. (2013) and Esposito et al. (2014) used forward modeling to correct for the flux loss of exoplanets and the flux and morphology of disks, respectively. For classical ADI (cADI), Cantalloube et al. (2015) showed that forward modeling techniques can improve the sensitivity of cADI and mitigate biases. The use of Markov-chain Monte Carlo (MCMC) in conjunction with forward modeling was presented in Bottom et al. (2015) for reference differential imaging.

Recently, Pueyo (2016) introduced a method called KLIP-FM to analytically forward model the degradation of a faint astrophysical signal that occurs when using least squares-based PSF subtraction algorithms such as KLIP that is also generally applicable to any observing strategy. Additionally, the computation of the forward model with KLIP-FM is much quicker than negative fake planet injection methods (Marois et al. 2010a; Lagrange et al. 2010), as the stellar PSF subtraction algorithm, KLIP, needs to be run only once. In this paper, we demonstrate the advantages of KLIP-FM for precise astrometry and constraining planetary orbits by applying it to GPI observations of β Pic b reduced using KLIP and ADI+SDI. In Section 4.2, we describe our new astrometry technique, Bayesian KLIP-FM Astrometry (BKA), in which we forward model the PSF of the planet with KLIP-FM and then use the forward model in a Bayesian framework to measure the position of the planet while also modeling the correlated nature of the noise. In Section 4.3, we validate our technique by fitting an orbit to the data and showing that our astrometry and uncertainties are consistent with Keplerian motion with no obvious systematics. Finally, in Section 4.4, we apply our new astrometry to constrain the orbit of β Pic b and place the tightest constraints yet on the transit of the planet and its Hill sphere.

4.2 Observations and Data Reduction

4.2.1 Observations

To obtain a large temporal baseline of GPI astrometric points, we compiled GPI data of β Pic from commissioning (Gemini program GS-ENG-GPI-COM), an astrometric monitoring campaign of β Pic b (Gemini programs GS-2015A-Q-21 and GS-2015B-Q-9; PI: Graham), the GPI Exoplanet Survey (GS-2014B-Q-500; PI: Macintosh), and a Gemini Large and Long Program to characterize debris disks (GS-2015B-LP-6: PI: Chen). All the data used in the following analysis are listed in Table 4.1. Most of these data were published and analyzed in a previous study characterizing the orbit of β Pic b by Millar-Blanchaer et al. (2015). We have reprocessed those data with the new astrometry methods presented in this paper and combined them with five additional new epochs.

There were three datasets in Millar-Blanchaer et al. (2015) that we did not use. We did not use the polarimetry dataset as there was no instrumental PSF obtained with the data to forward model. The dataset on 2014 March 23 was taken during tests of the adaptive optics (AO) system, causing the instrumental PSF to be highly varying and making it unsuitable for forward modeling. The 2015 January 24 dataset contained artifacts in the construction of the forward model that we could not remove. We chose not to include the measured astrometry from Millar-Blanchaer et al. (2015) for the datasets we omitted in order to maintain homogeneity in astrometric datasets and reduce potential systematic errors. We also note that in the last dataset taken on 2016 January 21, there was saturation on the edge of the occulting mask due to bad seeing. This affected the forward modeling of β Pic b, which was also near the occulting mask so a large portion of the frames taken could not be used in this analysis.

Table 4.1: Observations and Data Reduction Parameters for GPI Data on β Pic

UT date	Program	Filter	Exposure time (s)	Field rotation (°)	Average seeing[a] (″)	KL modes	Exclusion criteria (pixels)	High-pass filtered	Fitting box size (pixels)
2013 Nov 16	GS-ENG-GPI-COM	*K1*	1789	26	1.09	7	4	No	13
2013 Nov 16	GS-ENG-GPI-COM	*K2*	1253	18	0.93	7	4	No	13
2013 Nov 18	GS-ENG-GPI-COM	*H*	2446	32	0.68	7	4	No	13
2013 Dec 10	GS-ENG-GPI-COM	*H*	1312	39	0.77	7	4	No	13
2013 Dec 10	GS-ENG-GPI-COM	*J*	1597	19	0.70	7	4	No	13
2013 Dec 11	GS-ENG-GPI-COM	*H*	410	65	0.46	7	4	No	11
2014 Nov 8	GS-ENG-GPI-COM	*H*	2147	25	0.77	7	4	No	11
2015 Apr 2	GS-2015A-Q-21	*H*	1312	10	0.51	2	2	No	11
2015 Nov 6	GS-2014B-Q-500	*H*	2207	28	-[b]	7	4	No	11
2015 Dec 5	GS-2015B-LP-6	*J*	4948	66	0.92	7	4	Yes	11
2015 Dec 22	GS-2015B-Q-9	*H*	2088	19	0.76	7	4	No	11
2016 Jan 21	GS-2015B-Q-9	*H*	954	17	1.18	10	2	Yes	7

[a] Measured by the Gemini DIMM

[b] Seeing monitor data were not available for this observation

4.2.2 Reducing Raw GPI Data

The raw integral field spectrograph (IFS) data were reduced to form three-dimensional (3-D) spectral data cubes using the GPI Data Reduction Pipeline (DRP) version 1.2.1 or 1.3 (Perrin et al. 2014). There were no significant changes between the two versions of the GPI DRP that impacted the astrometry of β Pic b. We used the same data reduction process as in Millar-Blanchaer et al. (2015) and will summarize them here. First, dark subtraction and bad pixel correction were applied to each two-dimensional (2-D) image. For earlier datasets in which cryocooler vibration induced correlated noise on the detector, the frames were "destriped" to remove this noise (Ingraham et al. 2014). Then, we corrected for instrument flexure using an argon arc lamp taken before each sequence to align each individual spectrum for extraction (Wolff et al. 2014). Then each 2-D frame was turned into a spectral data cube, corrected for any remaining bad pixels, and corrected for distortion (Konopacky et al. 2014). For the K-band data, thermal background frames were taken along the sequence. We constructed thermal background cubes in the same fashion and subtracted them from the K-band data.

To spatially register and photometrically calibrate our data, we used the GPI DRP to measure the flux and location of the "satellite" spots, which are centered on the occulted star and imprinted with its attenuated spectrum (Marois et al. 2006b; Sivaramakrishnan & Oppenheimer 2006; Wang et al. 2014). The satellite spot fluxes were used to derive a flux calibration in each spectral channel, which we used when constructing the forward models in Section 4.2.4.

The location of the occulted star at each wavelength in each datacube was found using a least squares fit to all of the satellite spots' positions and the magnitude of the atmospheric differential refraction. The occulted star center is used to align all the images together before PSF subtraction, and is crucial for determining the astrometry of β Pic b relative to its host star. The precision on the star center is 0.05 pixels (0.7 mas) for satellite spots with signal-to-noise ratios (SNR) > 20 (Wang et al. 2014), which is certainly the case for all of our data on bright stars like β Pic.

4.2.3 PSF Subtraction

To subtract the stellar PSF from each individual datacube, we used `pyKLIP` (Wang et al. 2015), a Python implementation of the KLIP algorithm. We used both ADI and SDI to decorrelate the stellar PSF from the PSF of β Pic b. As we were only concerned with β Pic b, we applied our PSF subtraction only on the annulus that included the planet, rather than the full image.

We adjusted three main parameters for the PSF subtraction, depending on the dataset. The first was the number of modes used from the Karhunen-Loève (KL) transform to model the stellar PSF. The second was an exclusion criteria for picking reference PSFs. The exclusion criteria is similar to the quantity N_δ in Lafrenière et al. (2007) and is defined by the number of pixels that β Pic b would move azimuthally and radially in an observing sequence due to ADI and SDI. Thirdly, we toggled an ∼11 pixel wide spatial high-pass filter that was applied to some datasets before PSF subtraction. The high-pass filter was implemented using a Gaussian filter in Fourier space and the ∼11 pixel size was chosen to remove low-frequency background without significantly distorting any point sources in the image. In general, increasing the number of KL modes, decreasing the exclusion criteria, and applying a high-pass filter improves the subtraction of the stellar PSF. However, taken to the extreme, all three options attenuate signal from the planet. Additionally, for a planet as bright as β Pic b, the forward modeling described in Section 4.2.4 may not be valid when the PSF subtraction becomes too aggressive (Pueyo 2016). Thus, to optimize the signal of the planet while maintaining the validity of the forward modeling, we varied these parameters for each dataset. As β Pic b is bright, in most of the datasets we used an exclusion criteria of four pixels, which is slightly greater than 1 λ/D, where λ is the wavelength and D is the diameter of the telescope. When there was little field rotation for ADI, we decreased the exclusion criteria to not overly restrict our reference PSFs, but also decreased the number of KL modes used to avoid being too aggressive. When the speckle noise was too bright due to observing at a shorter wavelength or when the planet moved closer to the star, we applied a high-pass filter to remove some of the diffracted starlight. We list the chosen parameters for each dataset in Table 4.1.

After PSF subtraction, all images were rotated so that north is up and east is left. Then, all the frames were mean combined in both the spectral and temporal directions, resulting in a single PSF-subtracted frame for each dataset.

4.2.4 Constructing the Forward Model

After stellar PSF subtraction, the PSF of the planet is distorted by over-subtraction, caused by the presence of the planet in the data we are subtracting, and self-subtraction, caused by the presence of the planet in the reference images. Over- and self-subtraction perturbs the astrometric and photometric properties of the planet's PSF and prevents a straightforward measurement of its position. Typically, the biases and uncertainties in the astrometry caused by stellar PSF subtraction are estimated by injecting fake planets into the data at other PAs and comparing the retrieved position to the injected position. However, the over-subtraction and self-subtraction that distorts the planet PSF is deterministically caused by the existence of a planet in the data and its apparent motion in the reference images induced by ADI and SDI. In turn, these features, if modeled, can inform us about the location of the planet and improve our astrometric precision and accuracy.

Recently, Pueyo (2016) introduced KLIP-FM, an analytic framework to compute the effect of a planet on the KL modes and use these perturbations to reconstruct the over-subtraction and self-subtraction features. Using this technique, we are able to generate the PSF of the planet after PSF subtraction (see Appendix A.1 for a detailed description of the procedure). Briefly, we use the PSF of the planet before PSF subtraction, the apparent motion of the planet due to ADI and SDI, and a model of the planet spectrum to compute the distorted PSF of the planet after PSF subtraction. The result is a 2-D broadband planet

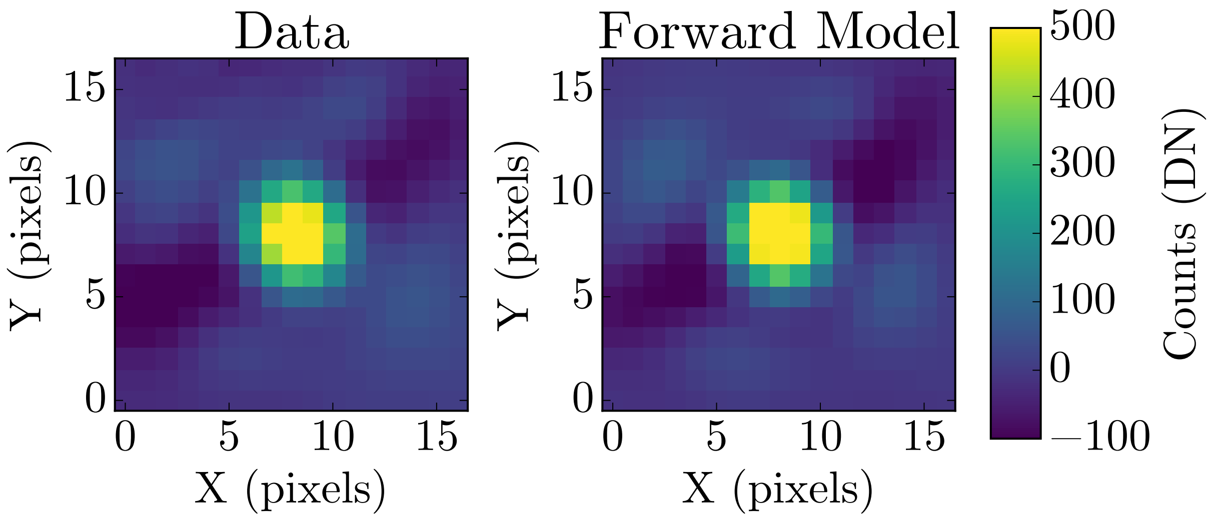


Figure 4.1: *(Left)* An image of β Pic b from the 2013 November 18 GPI H-band data after stellar PSF subtraction. *(Right)* Unoptimized forward model of β Pic b for the same dataset. The forward model has not been optimized to fit the data yet using the MCMC procedure discussed in Section 4.2.5, but should already be accurate to within a pixel. The best fit forward models and residuals for each GPI dataset are in Appendix A.2. The star is to the upper left of the planet and far outside the region shown here.

PSF, which we call F, centered at (x_0, y_0), our initial estimate for the location of the planet.

For our GPI data, the forward models were generated using the implementation of KLIP-FM in `pyKLIP`. For each dataset, we used the same parameters as our PSF subtraction to construct the forward model. To construct a model of the instrumental PSF at each wavelength, we used the average PSF of the satellite spots across all images for that wavelength. We used a basic center of light centroiding routine to measuring an approximate x and y positions of the planet (x_0 and y_0) in the PSF-subtracted image. This initial estimate for the position is good to within a pixel for our GPI data. For the input planet spectrum of the forward model, we used the normalized best-fit model spectrum from Chilcote et al. (2015), which has an effective temperature of 1650 K and a $\log(g) = 4.0$ (cgs units). We scaled the spectrum by eye to approximately match the contrast of β Pic b in our data and used the satellite spot fluxes to convert from contrast to digital numbers. We found that varying the spectrum and photometry had negligible impacts on our measured astrometry. Even a spectrum differing by $\sim$25% in shape changed the astrometry by < 0.2 mas, significantly smaller than the uncertainties we find in the following analysis.

With this input information, we used `pyKLIP` to generate one forward-modeled PSF, F, for each dataset. In Figure 4.1, we show an example forward-modeled PSF (not optimized to fit the data) and comparison to data for our 2014 November 18 dataset on β Pic b.

Qualitatively, the forward model matches all the features seen in data, including the prominent negative self-subtraction lobes on either side of the planet. As `pyKLIP` parallelizes the computation, the generation of the forward models is quick. On a 32 core machine with AMD Opteron 6378 processors clocked at 2.3 GHz, forward models for all 37 channels of a representative 37 cube GPI dataset were generated in 15 minutes: 4 minutes of overheads for preprocessing and generating the instrumental PSFs and 11 minutes to execute KLIP-FM and create the forward model. We note that we chose to run the forward model on a large annulus to examine noise properties in the data and that the computation time for KLIP-FM decreases ∼20-30% if a small sector around the planet was used instead.

4.2.5 Locating the Planet with Bayesian Parameter Estimation

To use the forward-modeled PSF, F, to perform astrometry, we developed a Bayesian framework to fit F to the data, account for correlated noise, and estimate our fitting uncertainties. First, ignoring the correlated noise, we can use three parameters to fit the forward model to the data: the location on the planet in x (x_p), the location of the planet in y (y_p), and a scale factor (α) to match the flux of F to the data. We can then write the posterior probability for x_p, y_p, and α given the data D using Bayes' Theorem as

$$P(x_p, y_p, \alpha | D) = P(D | x_p, y_p, \alpha) P(x_p, y_p, \alpha), \tag{4.1}$$

ignoring normalization constants. The first term on the right-hand side is the likelihood and the second is the prior.

To construct the likelihood, we must first use our input parameters and F to generate a model to compare to the data. We scale F by α and recenter it from its guessed location (x_0, y_0) to (x_p, y_p). The residual, R, between the model and the data, D, in fitting region $\mathcal{F}$ is calculated by

$$R \equiv R(x_p, y_p, \alpha) = (D - \alpha F(x_p, y_p))_{\mathcal{F}}\,, \tag{4.2}$$

where fitting region $\mathcal{F}$ is a fixed rectangular box centered at the approximate location of the planet in the data. We pick the size of the fitting region to be a few λ/D to encompass the PSF and the self-subtraction wings. We varied the fitting box size for each of our datasets in order to keep the fit focused on the area where signal from the planet can be seen. We list the size of the fitting box for each dataset in Table 4.1. Note that we do not fit any background term to the image, as one of the first steps of KLIP is to subtract off the mean of the image, removing any spatially constant background.

For data with uncorrelated Gaussian errors, the log of the likelihood function of the parameters for a particular set of data is

$$\begin{aligned} P(D | x_p, y_p, \alpha) &= \ln \mathcal{L}(x_p, y_p, \alpha) \\ &= -\frac{1}{2} \sum_{i}^{N_{\text{pix}}} \left[\frac{R_i^2(x_p, y_p, \alpha)}{\sigma_i^2} + \ln(2\pi\sigma_i) \right], \end{aligned} \tag{4.3}$$

where σ_i is the uncertainty in pixel i and $N_{\rm pix}$ is the number of pixels in the fitting region. However, the assumption of uncorrelated errors does not hold for images limited by speckle noise. Except in the cases of very aggressive PSF subtraction or at much greater separations from the star, the residual noise after PSF subtraction in GPI data of a bright star like β Pic is still dominated by correlated speckle noise, which has a correlation scale that depends on the aggressiveness of the PSF subtraction. Due to the bright nature of β Pic b, we could not use a very aggressive reduction to ensure that the planet remained a perturbation on our KL modes, thereby preserving the validity of the analytical forward modeling technique. The conservative PSF subtraction combined with the close separation of β Pic b from its bright host star required us to capture the correlated nature of the noise. We thus write our likelihood function instead as

$$\ln \mathcal{L} = -\frac{1}{2}(R^T C^{-1} R + \ln(\det C) + N_{\rm pix}\ln(2\pi)), \tag{4.4}$$

where C is the covariance matrix of size $N_{\rm pix} \times N_{\rm pix}$ and R is a $N_{\rm pix} \times 1$ matrix.

We applied a Gaussian process framework to characterize the covariance in the noise (see Czekala et al. (2015) for an in-depth explanation of the application of Gaussian processes to astronomical data). We only aim to model the correlations within each individual speckle, which spans λ/D in spatial extent. While in a single unprocessed frame speckle noise also has additional correlations on much larger scales, there are no significant correlations between speckles at larger spatial scales in our small fitting region due to stellar PSF subtraction with KLIP and averaging uncorrelated speckles together when collapsing the frames in our ADI+SDI sequence. Thus, within our fitting region, the dominant correlation in our noise is from pixels within a single speckle, and thus is the one correlation we modeled.

Following the procedure in Czekala et al. (2015) for fitting one-dimensional (1-D) correlations in stellar spectra, we used the Matérn covariance function parametrized with $\nu = 3/2$ to model the correlated speckle noise. We chose the Matérn function with $\nu = 3/2$ as it better fits the correlations at larger separations compared to a simple squared exponential relation. However, we note that switching between covariance functions that have similar shapes does not significantly alter the error bars. We also chose to assume symmetric noise as we did not find any difference in the correlation scale of our noise in the radial and azimuthal directions. This is likely be due to the fact that we used both ADI and SDI to model the stellar PSF and thus are better at subtracting speckle noise, which, before PSF subtraction, is more correlated radially than azimuthally due to the finite spectral bandwidth of the data. In instruments that are not able to utilize SDI, it might be necessary to separate the noise into radial and azimuthal components, each with its own characteristic correlation length. Thus, for our purposes, we chose the symmetric Matérn covariance function with $\nu = 3/2$, which defines the covariance between two pixels i and j separated by a distance of $r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ as

$$C_{ij} = \sigma_i \sigma_j \left(1 + \frac{\sqrt{3} r_{ij}}{l}\right) \exp\left(-\frac{\sqrt{3} r_{ij}}{l}\right), \tag{4.5}$$

where σ_i and σ_j are the noise for each pixel and l is a characteristic correlation length scale that increases when noise is correlated at larger spatial scales. We calculated the uncertainty for each pixel by computing the standard deviation of pixels in an annulus with a width of $\sim 2\lambda/D$ (6 pixels), centered on the star, and with a mean radius equal to the distance between that pixel and the star. Any pixels containing signal from the planet were masked and not used in estimating the noise in the annulus.

The correlation length l is not fixed, but rather kept as a nuisance parameter in our Bayesian framework that we will marginalize over in the end. Thus, the final posterior we are trying to calculate is really

$$P(x_p, y_p, \alpha, l|D) = P(D|x_p, y_p, \alpha, l)P(x_p, y_p, \alpha, l), \tag{4.6}$$

where α and l are both nuisance parameter as we are only interested in the astrometry.

Compared to our likelihood function, our prior is relatively straightforward. We use uniform priors in x and y within 2 pixels from our initial guess location. Similarly, we use an uniform prior between 0 and 5 for α to determine how much to scale F, which was already scaled to an approximate contrast of β Pic b. The correlation length l has a uniform prior between 0 and 10 pixels, which provides an ample range to explore the correlations within individual speckles of size $\lambda/D \approx 3$ pixels.

We used the Goodman & Weare (2010) Affine Invariant MCMC sampler implemented in the `emcee` Python package (Foreman-Mackey et al. 2013) to sample the posterior distribution and custom `cython` code to quickly generate the covariance matrix as we vary l. The MCMC sampler was run for 800 steps using 100 walkers, with a "burn-in" of 200 steps beforehand. In Figure 4.2, we show the posterior distributions from the MCMC fit for the 2014 November 18 H-band dataset as a representative posterior distribution. The value of l is close to $\lambda/D \approx 3$ pixels, indicating that we are accurately fitting the correlated speckle noise. We use the resulting posterior distributions to calculate the most likely values and uncertainties for the location of β Pic b in our image at each epoch.

We note that this method's estimated uncertainties are Bayesian, and that it is incommensurable to include frequentist uncertainities used in the astrometric calibration or compare with frequentist methods from previous analyses. Thus, the statistical analysis in the following paragraphs of this section is improper. In Appendix B, we offer an alternative frequentist method for estimating the location of the planet in the data that still includes the forward modeling of the planet PSF and the handling of correlated noise. We show in Appendix B that we can reach a similar conclusion when properly combining uncertainties.

To convert our results to more useful physical values, we need to convert our measured location of β Pic b to its positional offset from its host star in right ascension (RA) and declination (Dec). As our images are already rotated so that $-x$ is positive RA and $+y$ is positive Dec, it is straightforward to convert from image fates to sky coordinates. We use the satellite spots to measure the location of the star behind the coronagraph, which has a precision of 0.7 mas (Wang et al. 2014). This allows us to derive the separation of the planet from the star in pixels. To convert from pixel separations to physical separations in RA and Dec, we use the most recent astrometric calibration numbers from De Rosa et al. (2015): a plate scale of 14.166 ± 0.007 mas lenslet^{-1} and a residual north angle offset of $0\overset{\circ}{.}10 \pm 0\overset{\circ}{.}13$ from the north angle value used in the GPI DRP (versions 1.2.1 to 1.3, the current version). These astrometric calibration numbers show no significant change over time, so we apply them to all our epochs of data. Then, we assume all of these error terms are uncorrelated and add them in quadrature with our measurement errors from our MCMC fit to determine our full astrometric precision.

The combination of the forward-modeled PSF and the Bayesian framework makes up the Bayesian KLIP-FM Astrometry technique we introduce in this work. We apply BKA for all 12 GPI datasets and report the measured astrometry and error budgets in Table 4.2. The best-fit forward models and residuals to the fit are shown in Appendix A.2. On most of our datasets, we are not limited just by the uncertainty in the location of the planet, which was as low as 0.3 mas. The uncertainties in the location of the star and North angle also make significant contributions to the error budget. Typically, we achieved ∼1 mas precision on the relative astrometry between β Pic b and its host star. This is a factor of ∼2–4 improvement over previous techniques (Millar-Blanchaer et al. 2015) using the same data, indicating that this technique can be useful in reanalysis of archival data to obtain better astrometry in cases where the limiting factor is the uncertainty on the planet position. In two of the later datasets where the planet is observed closer in, we were limited by the SNR of the planet and unable to achieve 1 mas precision. In the 2015 December 5 dataset, the noise was higher due to the planet being fainter relative to the star in J-band. In the 2016 January 21 dataset, a combination of poor seeing and a small amount of usable data limited our astrometric precision.

Overall, though, this GPI β Pic b data is an excellent demonstration for Bayesian KLIP-FM Astrometry as the planet is bright enough that the extended PSF features, such as the negative self-subtraction lobes, are clearly seen and provide significant information to constrain the position of the planet. For fainter planets, the extended features are harder to distinguish from the noise. As one of the main advantages of BKA over techniques that do not forward model the PSF is being able to forward model the extended self-subtraction lobes, the astrometric improvement would not be as large for lower signal-to-noise ratio planets. There still should be some improvement, though due to accurately modeling the over-subtraction on the core of the PSF and small contributions from the extended features even if they are hard to distinguish from noise. Regardless, in addition to the improved precision, BKA should also more accurately estimate the uncertainties as it fits for the correlation scale of the noise at the location of the planet.

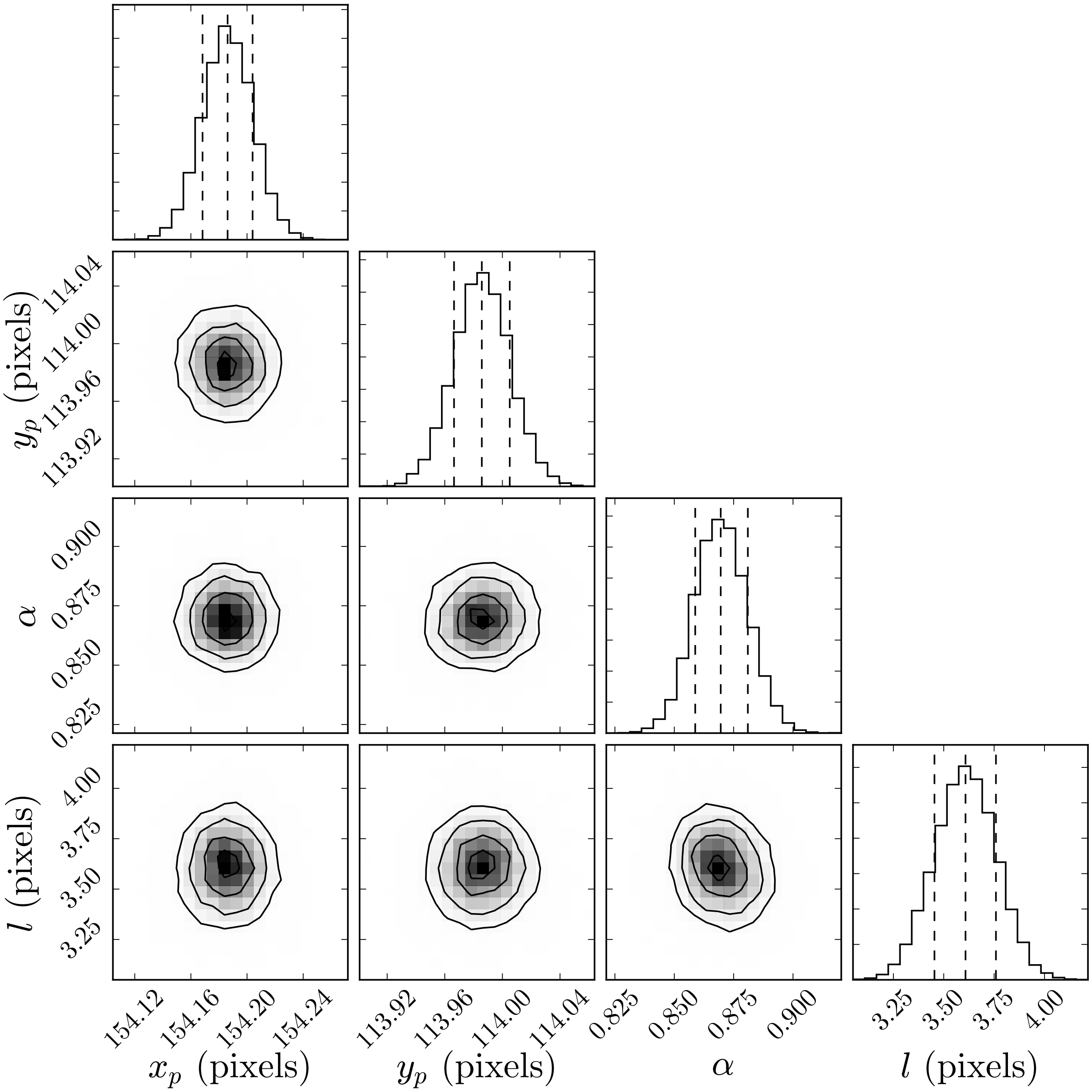


Figure 4.2: Posterior distribution of the four parameters in the MCMC fit for the astrometry for the 2014 November 18 epoch. The vertical dashed lines in the marginalized posterior distribution plots indicate the 16th, 50th, and 84th percentile values.

Table 4.2: Astrometric Error Budget and Measured Astrometry of β Pic b

	Contributions to Error Budget				Measured Astrometry			
Dataset	Planet x/y (mas)	Star x/y (mas)	Plate Scale (mas)	PA (°)	ΔRA (mas)	ΔDec (mas)	Radial Separation (mas)	Position Angle (°)
2013 Nov 16 $K1$	0.6/0.7	0.7/0.7	0.3	0.13	-228.5 ± 1.3	-366.2 ± 1.1	431.6 ± 1.0	212.0 ± 0.2
2013 Nov 16 $K2$	0.5/0.4	0.7/0.7	0.3	0.13	-229.2 ± 1.2	-364.5 ± 1.0	430.6 ± 0.9	212.2 ± 0.2
2013 Nov 18 H	0.3/0.3	0.7/0.7	0.3	0.13	-229.1 ± 1.1	-364.7 ± 1.0	430.6 ± 0.8	212.1 ± 0.2
2013 Dec 10 H	0.4/0.4	0.7/0.7	0.3	0.13	-227.9 ± 1.2	-359.9 ± 1.0	426.0 ± 0.9	212.3 ± 0.2
2013 Dec 10 J	0.6/0.7	0.7/0.7	0.3	0.13	-227.2 ± 1.3	-361.1 ± 1.2	426.6 ± 1.1	212.2 ± 0.2
2013 Dec 11 H	0.5/0.4	0.7/0.7	0.3	0.13	-227.8 ± 1.2	-359.2 ± 1.0	425.4 ± 0.9	212.4 ± 0.2
2014 Nov 8 H	0.5/0.5	0.7/0.7	0.2	0.13	-194.0 ± 1.1	-299.1 ± 1.0	356.5 ± 0.9	213.0 ± 0.2
2015 Apr 2 H	0.4/0.5	0.7/0.7	0.2	0.13	-172.1 ± 1.0	-266.5 ± 1.0	317.2 ± 0.9	212.9 ± 0.2
2015 Nov 6 H	0.7/0.7	0.7/0.7	0.2	0.13	-137.8 ± 1.1	-207.2 ± 1.0	248.8 ± 1.0	213.6 ± 0.3
2015 Dec 5 J	1.2/1.3	0.7/0.7	0.2	0.13	-133.9 ± 1.5	-200.5 ± 1.5	241.1 ± 1.4	213.7 ± 0.4
2015 Dec 22 H	0.5/0.5	0.7/0.7	0.2	0.13	-130.0 ± 1.0	-194.7 ± 0.9	234.1 ± 0.9	213.7 ± 0.2
2016 Jan 21 H	1.8/1.6	0.7/0.7	0.2	0.13	-126.8 ± 2.0	-185.8 ± 1.8	225.0 ± 1.8	214.3 ± 0.5

Table 4.3: Orbital Parameters of β Pic b

				Posterior Percentiles		
Parameter	Unit	Prior Range	Prior Distribution	16	50	84
Semi-major axis (a)	au	$4 - 40$	Uniform in $\log a$	9.02	9.66	10.78
Epoch of Periastron (τ)	-	$-1.0 - 1.0$	Uniform in τ	0.32	0.73	0.87
Argument of Periastron (ω)	°	$-360 - 360$	Uniform in ω	192.8	205.8	258.4
Longitude of the Ascending Node (Ω)	°	$25 - 85$	Uniform in Ω	31.67	31.76	31.84
Inclination (i)	°	$81 - 99$	Uniform in $\cos i$	88.70	88.81	88.93
Eccentricity (e)	-	$0.00001 - 0.99$	Uniform in e	0.027	0.080	0.171
Total Mass (M_T)	$M_\odot$	$0 - 3$	Uniform in $M_\odot$	1.76	1.80	1.83
Derived Parameters						
Period (P)	years	-	-	20.21	22.47	26.24
Hill Sphere Ingress	MJD	-	-	57,840	57,846	57,854
1/2 Hill Sphere Ingress	MJD	-	-	57,916	57,924	57,934
Closest Approach Date	MJD	-	-	57,986	57,996	58,008
1/2 Hill Sphere Egress	MJD	-	-	58,056	58,069	58,082
Hill Sphere Egress	MJD	-	-	58,132	58,147	58,162

4.3 Validation through Orbit Fitting

To explore the validity of our new astrometry technique, we fit a Keplerian orbit to our 12 epochs of astrometry. Since each epoch of astrometry is fit independently, and since the Keplerian orbit is agnostic toward the exact data-reduction methods, having all 12 epochs of astrometry fit the Keperlain orbit would only be possible if all of the astrometry is accurate and precise. If there are errors in estimating the magnitude of the uncertainties or any remaining biases in our measurements, we expect them to become evident in the residuals of the Keplerian fit either as systematic trends or fit outliers.

We followed the same analysis as in Millar-Blanchaer et al. (2015) to obtain the orbital elements of β Pic b using a MCMC fit with the parallel-tempered sampler implemented in `emcee`. We combined the 12 GPI astrometric points presented here along with the datasets presented in Chauvin et al. (2012) and Nielsen et al. (2014). Unlike Millar-Blanchaer et al. (2015), we did not explicitly include the radial velocity measurement of β Pic b from Snellen et al. (2014) in order to limit potential systematics between instruments, but we do use it to constrain the prior on the PA of the ascending node and thus the direction of the orbit (i.e., we know that that β Pic b has been moving toward us since ∼2007). Our model of the orbit fits the same seven parameters as Millar-Blanchaer et al. (2015). For convenience, the parameters are listed in Table 4.3.

The MCMC sampler was run for 30,000 steps using 1024 walkers at each of the 20 temperatures after 30,000 steps of "burn-in" to allow the walkers to converge. We thinned the chains to remove any remaining correlations, keeping every 75th step to result in effectively 400 samples per walker for a total of 409,600 samples to construct our posterior.

The posteriors on the seven parameters in our model are shown in Figure 4.3 and the 16%, 50%, and 84% percentiles of the marginalized posterior distribution for each parameter is listed in Table 4.3. For comparison, we plot 500 randomly chosen possible orbits along with the measured astrometry in Figure 4.4. The fit residuals of our measured GPI points are shown in Figure 4.5. The residuals in RA and Dec offset (which are the parameters we use in our MCMC) are consistent with zero and do not show any systematics. The residuals in radial separation and PA, neither of which are used in our fit, are slightly further away from zero, but do not indicate any obvious errors in either our astrometry or our error estimates.

To quantitatively assess the validity of our measurements, we use the reduced chi-squared (χ^2_{red}) statistic, which measures how consistent our astrometry is with a Keplerian orbit. Ideally, χ^2_{red} would be unity if all measurements and uncertainties were accurate. However, biases and improperly estimated errors would cause it to deviate from unity. Due to systematic astrometric calibration offsets between instruments that have not been characterized, we expect χ^2_{red} to be slightly above unity. For example, Millar-Blanchaer et al. (2015) reported a χ^2_{red} of $1.55^{+0.09}_{-0.05}$ when combining GPI astrometry, measured using standard techniques in the field, with previous astrometric measurements

Even with ∼2–4 times smaller error bars on the GPI points, the χ^2_{red} of our accepted orbits is $1.53^{+0.08}_{-0.06}$, unchanged from the $1.55^{+0.09}_{-0.05}$ reported in Millar-Blanchaer et al. (2015). With BKA contributing 12 out of the 30 astrometry measurements, if the BKA technique contained biases larger than one milliarcsecond, they would have caused a significant increase in χ^2_{red}. Likewise, if we had been too optimistic with our error estimates, χ^2_{red} should have also increased as the reported uncertainties would not have matched the scatter in the measurements. The lack of increase in χ^2_{red} indicates that the more precise astrometry from BKA is not biased and has accurate uncertainties.

In addition to the Keplerian orbit fit, we can examine the accuracy of our estimated one milliarcsecond uncertainties by looking at the astrometry measured on the same or consecutive days. As we do not expect the planet's position to change significantly in the span of a single day, the three measurements in 2013 November and the three measurements in 2013 December ought to be consistent with each other. Indeed, our measurements in Table 4.2 indicate that in both sets of astrometry, the measurements agree at the milliarcsecond level, confirming that our estimated uncertainties are accurate.

Thus, the well behaved residuals of our GPI measurements and the lack of change in χ^2_{red} from Millar-Blanchaer et al. (2015) even with significantly smaller error bars lead us to conclude the measured astrometry using BKA are accurate and free from biases. The fact the residuals are consistent with zero, the lack of change in χ^2_{red}, and the consistency of repeated measurements taken around the same time all indicate that the one milliarcsecond uncertainties estimated from BKA are also accurate. Together, these multiple assessments of the validity of BKA all indicate that this technique is both accurate and precise.

4.4 Discussion

4.4.1 The Orbit of β Pic b

Having demonstrated the accuracy and precision of this new technique, we now analyze the new constraints on the orbit of β Pic b. Not surprisingly, the estimates for a and e have not changed significantly from Millar-Blanchaer et al. (2015) since the GPI points reanalyzed in this paper do not have a sufficiently long time baseline to constrain these parameters. As seen in Figure 4.4, all but one astrometric measurement is on one-half of the orbit curve. The other half of the orbit is not as well constrained, leaving a degeneracy in a and e. This degeneracy can be broken with more measurements obtained when the planet appears on the other side of the star. Better constraints on a and e will provide better insight on how β Pic b interacts with the debris disk and potential unseen planets in the system. The new total mass of the system, $M_T = 1.80^{+0.03}_{-0.04}\ M_\odot$, is significantly higher than the $M_T = 1.61 \pm 0.05\ M_\odot$ from Millar-Blanchaer et al. (2015). This new total system mass, which effectively measures the mass of the star at this precision, is in better agreement with the stellar mass of 1.75 $M_\odot$ derived from stellar photometry (Crifo et al. 1997). Our measurement of the position angle of the ascending node, Ω, slightly improves upon the value obtained by Millar-Blanchaer et al. (2015), but the overall value remains consistent. Thus, β Pic b is still consistent with being the planet responsible for the known warp in the debris disk (e.g. Dawson et al. 2011). Compared to Millar-Blanchaer et al. (2015), the argument of the periastron, ω in our fit has increased from $156^{\circ}{}^{+33}_{-76}$ to $206^{\circ}{}^{+52}_{-13}$. This new value is consistent with $\omega = 200^{\circ} \pm 20^{\circ}$ that is required for the falling evaporating bodies (FEBs) scenario proposed by Thébault & Beust (2001) to explain redshifted absorption features in β Pic's spectrum. Note that under previous definitions of the orbital parameters, this has been expressed as $\omega = -70^{\circ} \pm 20^{\circ}$ from the line of sight. This scenario also requires a slightly eccentric orbit, which is consistent with our derived orbital parameters. For a more in-depth discussion of β Pic b's relationship to the debris disk and the FEB scenario, we direct the reader to Millar-Blanchaer et al. (2015).

The biggest improvement in our understanding of the orbit of β Pic b is the improved constraint on the inclination of the orbit. We find the inclination to be $i = 88\overset{\circ}{.}81^{+0.12}_{-0.11}$ which allows us to place the tightest constraints on the probability that β Pic b will transit its host star. Assuming an angular diameter of the star of 0.736 mas (Defrère et al. 2012) and considering the range of a from our orbit fit, we find that in most cases, we need $|i - 90^{\circ}| < 0.05^{\circ}$ in order for the planet to transit. With our current constraints on the inclination, we have ruled out the possibility that β Pic b will transit at 10σ significance. This tight constraint on the inclination and transit probability is due to the slightly longer time baseline and the improved precision in the measured position angle of the GPI astrometry compared to Millar-Blanchaer et al. (2015). For an edge-on orbit that transits the star, we should see no significant change in PA over time. However, Figure 4.4 shows that the GPI points alone reveal a significant increase in PA over time. Thus, regardless of systematic astrometric calibration errors between instruments, we conclude that β Pic b will not transit its star.

4.4.2 Hill Sphere Transit

Unlike the planet, β Pic b's Hill sphere, the region around the planet that could contain gravitationally bound circumplanetary material, will transit the star. We define the radius of the Hill sphere as $r_H \approx a(1-e)\sqrt[3]{\left(\frac{m}{3M}\right)}$ using the approximate form proposed in Hamilton & Burns (1992) where m is the mass of the planet and M is the mass of the star. Assuming β Pic b follows a"hot-start" evolutionary track with a mass of $12.7 \pm 0.3\ M_{\mathrm{Jup}}$ (Morzinski et al. 2015) and using the range of semi-major axes, eccentricities, and stellar masses from our MCMC orbit fit, $r_H = 1.165^{+0.013}_{-0.016}$ au ($59.9^{+0.7}_{-0.9}$ mas). Given that our prediction for the closest approach of β Pic b will be $9.9^{+0.9}_{-0.8}$ mas (0.19 ± 0.02 projected au) from the star, the Hill sphere of β Pic b will transit the star, as shown in Figure 4.6.

From our MCMC orbit fit, we can compute when the transit of the Hill sphere will occur. We pick five notable events to focus on: two are the ingress and egress of the Hill sphere which are the extrema in time between which any circumplanetary material could transit; another is the date of closest approach, which gives the opportunity to probe material closest in to the planet; and the last two are the transit of the sphere that is $1/2\ r_H$ in radial extent. Almost all stable prograde circumplanetary orbits reside within $1/2\ r_H$ (Shen & Tremaine 2008), so it is more likely to find material within half a Hill sphere. In Table 4.3, we list our constraints on the date of these events, and in Figure 4.3, we plot the posterior distribution for these events. The duration of the Hill sphere transit will be long: ingress is ~2017 April 3 and egress is ~2018 January 29. The 1/2 Hill sphere begins transit ~2017 June 20 and ends transit on ~2017 November 12, with closest approach on ~2017 August 31.

Given the RA of β Pic ($05^h47^m17^s$), the star will not be visible from most ground-based observatories during almost the entire time between ingress of the Hill sphere and closest approach. Ground-based telescopes in Antarctica, airborne observatories capable of travelling to Antarctica, and space-based observatories would provide the only opportunities to observe β Pic during this time period. The second half of the Hill sphere transit will be visible from most southern hemisphere ground-based observatories, making β Pic a well suited candidate for photometric monitoring in late 2017.

As it is a rare opportunity to probe circumplanetary material, it is not certain what will be seen when the Hill sphere of β Pic b transits the star. One possibility is that satellites could reside in the Hill sphere. To approximate the photometric transit depths, the 0.736 mas angular diameter of the star corresponds to 1.53 $R_\odot$ or 1.065×10^6 km. Moons as large as Ganymede (r = 2630 km) or Io (r = 1820 km) would give transit depths of 6×10^{-6} and 3×10^{-6}, respectively. Detecting these photometric signatures will require a high cadence, as any single satellite orbiting β Pic b will have a transit duration of ~2 days. Additionally, β Pic is a variable star with pulsation timescales of ~0.5 hours and variability amplitudes in B-band of < 5 mmags (Koen et al. 2003), so careful modeling of stellar activity is necessary to be sensitive to these transit depths.

Another possibility is that, as β Pic b is still young and evolving, it may harbor a circumplanetary disk or ring system comprised of leftover material from planet formation. Such a hypothesis is not unprecedented as Kenworthy & Mamajek (2015) found evidence for a large circumplanetary disk around an unseen planet in the 1SWASP J140747.93-394542.6 (hereafter J1407) system, which has a similar age (~16 Myr; Mamajek et al. 2012) and thus likely at a similar stage in its evolution. Kenworthy & Mamajek (2015) interpreted the series of complex and deep eclipses in the J1407 light curve as from a circumplanetary disk 0.6 au in radial extent in the process of forming rings due to newly formed satellites. It is plausible that β Pic b can harbor a similar disk as a 0.6 au disk would be 1/2 r_H in extent and consistent with where we would expect stable orbits to reside around β Pic b. Additionally, since β Pic b is young, it is plausible that there is a large amount of circumplanetary material which has yet to be cleared out dynamically. Such a large disk would transit the star and be suitable for detection through photometric monitoring of the host star if the disk is inclined by $\gtrsim 18°$ with respect to the orbital plane of β Pic b. To maintain such an inclination, the disk needs to be massive enough to prevent stellar tidal forces from aligning the disk to the orbital plane (Zanazzi & Lai 2017).

So far, there have not been many observational constraints of circumplanetary material around β Pic b. Lecavelier Des Etangs et al. (1997) reported a photometric event in 1981 and hypothesized it could be due to the transit of a planet that cleared out a hole in the debris disk around its Hill sphere. However, Millar-Blanchaer et al. (2015) show that the planet is not embedded in the debris disk and we definitely show the planet itself will not transit.

Still, our orbit models give a 8% and 4% chance the photometric event coincided with the transit of the Hill sphere and 1/2 Hill sphere respectively, during which time material around the planet could have passed in front of the star. Additionally, a circumplanetary disk or ring may also be detectable through the planet's spectral energy distribution (SED). The planet's near-infrared spectrum would experience extinction if the dust resides between us and the planet, with the magnitude of extinction depending on the amount of dust. In the near infrared, the extinction would be greater at shorter wavelengths due to the increased scattering and absorption by dust and would produce a spectral slope in the planet's near-infrared SED. Dust around the planet would also scatter starlight, causing the planet's SED to appear brighter in the optical, as has been postulated for Fomalhaut b (Kalas et al. 2008). Furthermore, the dust will produce millimeter emission. However, detecting circumplanetary material in the planet's SED will require being able to distinguish it from the circumstellar disk with precise spectral data.

4.5 Conclusions

In the first part of this work, we have presented a new technique for more precise and accurate astrometry of directly imaged exoplanets using a new analytical forward modeling approach in a robust statistical framework.

- Using the KLIP-FM framework presented in Pueyo (2016), we are able to analytically forward model the PSF of the planet through the data reduction process, giving us better information on the location of the planet. We apply KLIP-FM to GPI data on β Pic and forward model the PSF of β Pic b using the open source `pyKLIP` package.

- For a close-in planet orbiting a bright star like in the case of β Pic b, we are limited by correlated speckle noise in our data. We developed a Bayesian framework utilizing Gaussian processes and MCMC to account for the correlated noise and to find the position of the planet simultaneously.

- With this technique, we have achieved the most precise astrometry on β Pic b to date. On most of our GPI datasets, we achieve ∼1 mas precision on the relative separation between β Pic b and its host star, a ∼2-4 fold improvement over previous techniques using the same data (Millar-Blanchaer et al. 2015).

- In datasets where the astrometry is limited by noise and not by astrometric calibration uncertainty, Bayesian KLIP-FM Astrometry approach should improve astrometric precision.

In the second part of this work, we apply our Bayesian KLIP-FM Astrometry technique to the orbit of β Pic b.

- To validate this new astrometric technique, we used it to measure the position of β Pic b in 12 epochs of GPI data. We combined these 12 astrometric points with

2 previous astrometric monitoring campaigns and fit a Keplerian orbit using MCMC methods. We find the residuals to the fit are consistent with zero and show no apparent systematic trends, indicating that our fit is accurate and the uncertainties we estimate are reliable.

- Due to the improved PA measurements from our technique, we have the tightest constraints on the inclination of the orbit and can exclude a possible transit of β Pic b at 10σ significance.
- While the planet will not transit, we are confident the Hill sphere around β Pic b will transit. The Hill sphere will begin transit at the start of April in 2017 and finish transiting at the end of January in 2018 with closest approach in the end of August in 2017. The transit of β Pic b's Hill sphere should be our best chance in the near future to investigate young circumplanetary material.

In the future, this MCMC forward modeling technique can be applied to photometry and spectral extraction alongside of astrometry of directly imaged exoplanets, allowing for improved characterization of their atmospheres. For β Pic b, continued monitoring of its orbit will yield more insight into the dynamics of the star system, although the planet will soon be too close to its star to be seen with current direct imaging instrumentation. However, once the planet appears on the other side, continued astrometric monitoring should be able to constrain the semi-major axis and eccentricity of the orbit much better, which will improve our understanding of how β Pic b perturbs the disk and if there are other planets perturbing β Pic b.

Acknowledgements

The GPI project has been supported by Gemini Observatory, which is operated by AURA, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the NSF (USA), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), MCTI (Brazil) and MINCYT (Argentina). We thank the National Science Foundation (NSF AST-1518332) and NASA (NNX15AC89G and NExSS program NNX15AD95G) for contributing to support of this research. Portions of this work were performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France. The posterior distribution plots were made with `corner` (Foreman-Mackey et al. 2016).

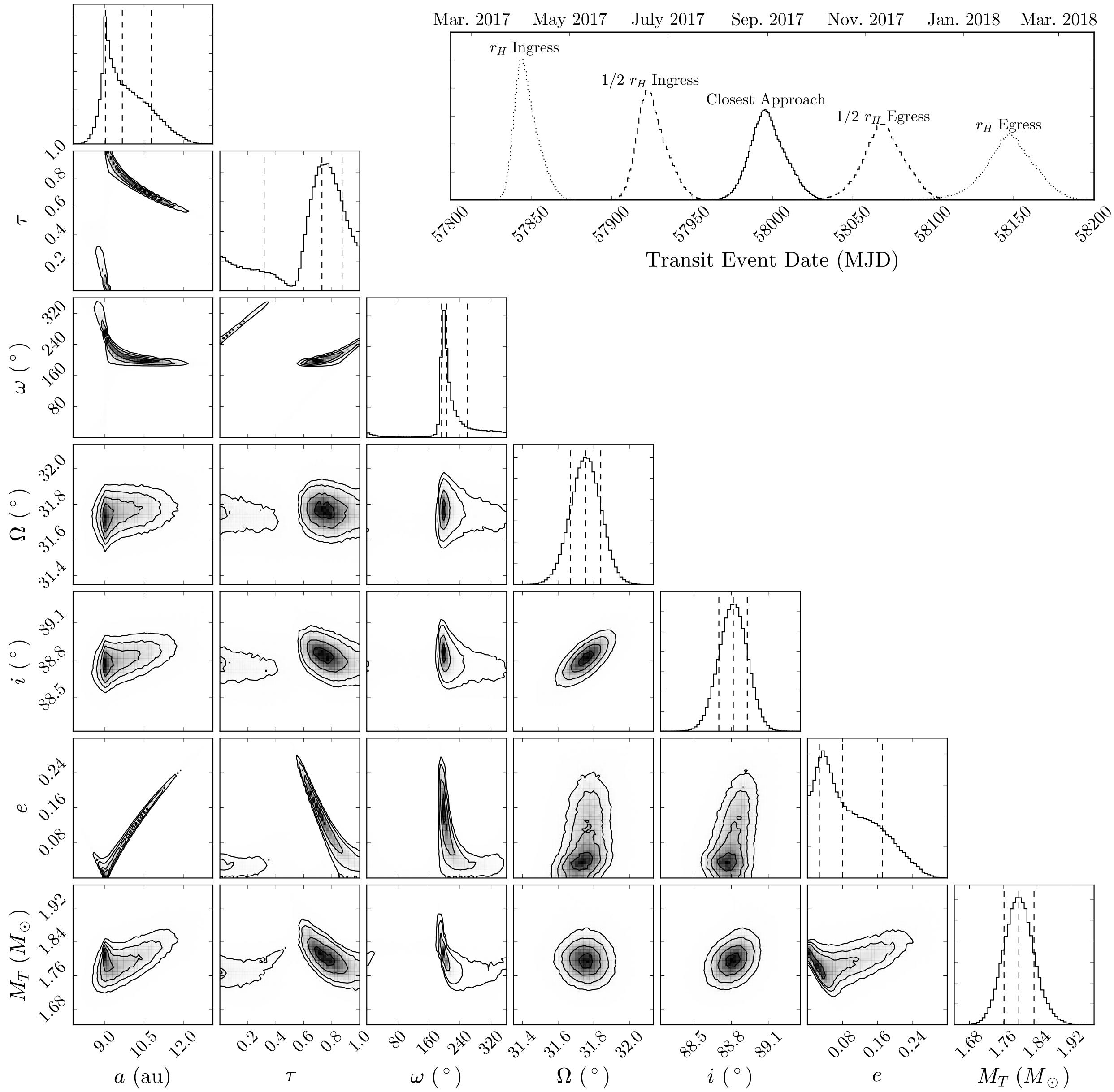


Figure 4.3: Posterior distributions for the seven orbital elements in our Keplerian orbit model along with inferred distributions of possible dates for transit events in the top right corner. The vertical dashed lines in the marginalized posterior distribution plots indicate the 16th, 50th, and 84th percentile values. For the transit events, the dotted line corresponds to the ingress and egress of the full Hill sphere, the dashed line corresponds to the ingress and egress of the half Hill sphere, and the solid line corresponds to the date of closest approach.

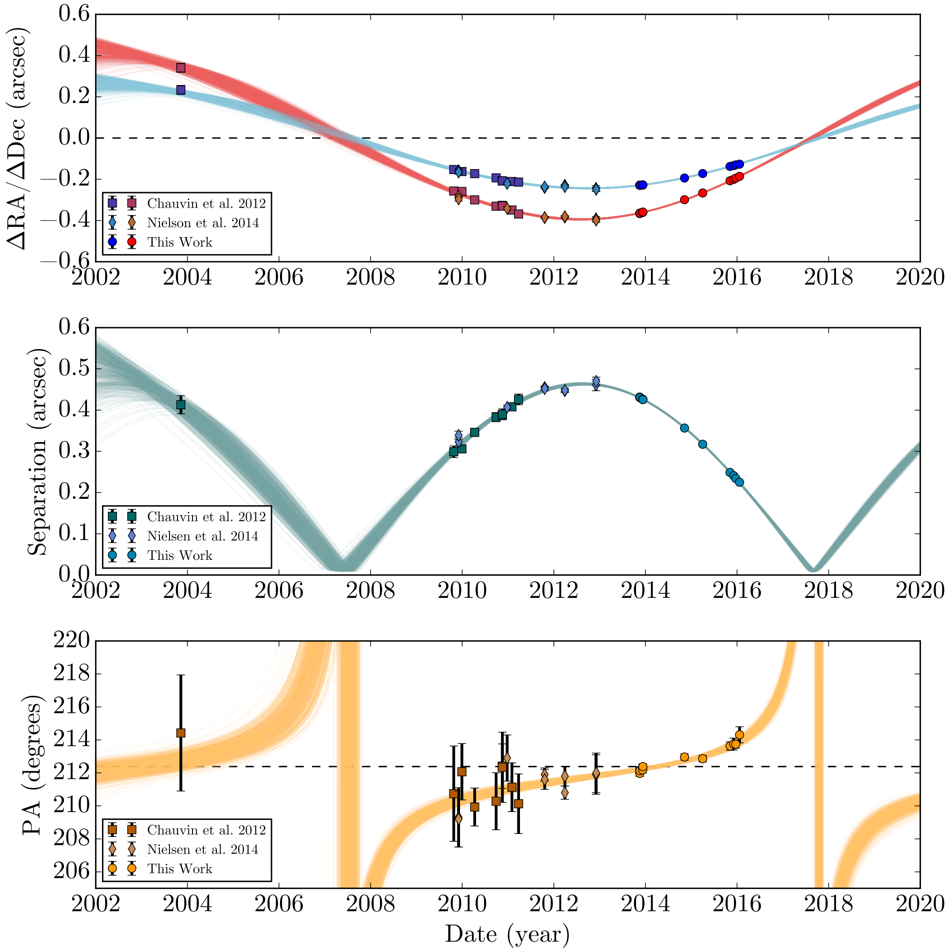


Figure 4.4: *(Top)* Offset of β Pic b in right ascension (blue) and declination (red) with respect to β Pic as a function of time. We have plotted the measured data and 500 randomly chosen accepted orbits from the MCMC sampler. *(Middle)* Radial separation of β Pic b from the star as a function of time. The same 500 randomly-chosen orbits have are also plotted *(Bottom)* PA as a function of time for the data and the 500 randomly chosen orbits. To keep the data compact, we have wrapped PA by 180° to only consider position angles between 180° and 360°. This allows easy comparison to the 2003 point, which is nominally at a PA of 34.°4 but here displayed at a PA of 214.°4. The dashed black line indicates a constant PA of 212.°4, the weighted mean of all points. If the planet were to transit, we would not be able to see a significant deviation from constant PA in time. For all the plots, error bars are also plotted but many are too small to be seen on this scale.

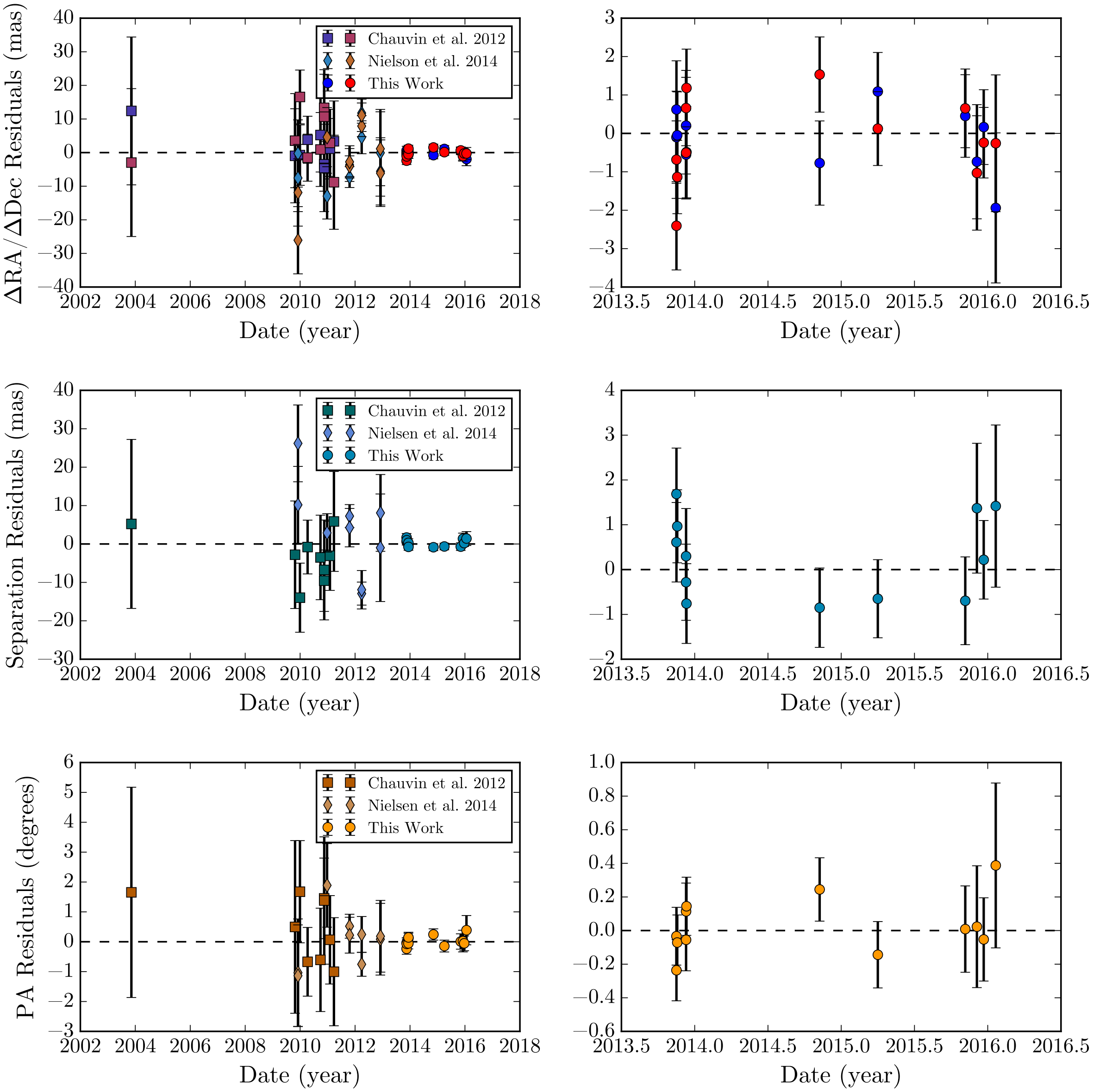


Figure 4.5: Residuals to the orbit fit for the average of 500 randomly chosen accepted orbits. The top row shows the residuals in ΔRA (blue) and ΔDec (red) offset, which are the coordinates used in the MCMC analysis to fit the orbit. We also plot radial separation (middle row) and position angle (bottom row). The separation and position angle residuals were not optimized in the MCMC fit. The right column is a zoom-in of the left column showing only the residuals of the GPI astrometry.

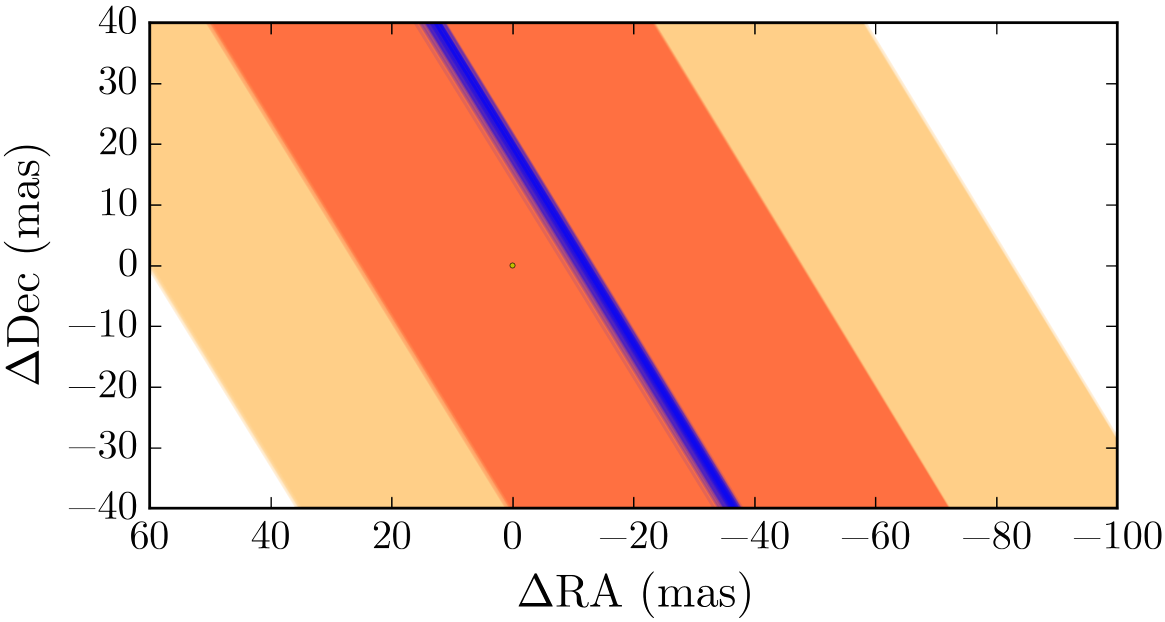


Figure 4.6: Region in the sky that the Hill sphere of β Pic b will sweep across during the closest approach of the planet in 2017. One-hundred randomly chosen accepted orbits (blue) are plotted along with the angular extents of their Hill spheres (light orange) and 1/2 Hill spheres (dark orange). The star, β Pic, is shown in its true angular size (small yellow dot). Both the Hill sphere and 1/2 Hill sphere will pass in front of the star. Note that for clarity we are not plotting the orbital path for β Pic b when it comes back in its orbit and passes behind the star.

Chapter 5

Dynamical Constraints on the HR 8799 Planets with GPI

A version of this chapter has been submitted for publication in a refereed journal.

The HR 8799 system uniquely harbors four young super-Jupiters whose orbits can provide insights into the system's dynamical history and constrain the masses of the planets themselves. Using the Gemini Planet Imager (GPI), we obtained one milliarcsecond precision on the astrometry of these planets. We assessed four-planet orbit models with different levels of constraints and found that assuming the planets are near 1:2:4:8 period commensurabilities, or are coplanar, does not worsen the fit. We added the prior that the planets must have been stable for the age of the system (40 Myr) by running orbit configurations from our posteriors through N-body simulations and varying the masses of the planets. We found that only assuming the planets are both coplanar and near 1:2:4:8 period commensurabilities produces dynamically stable orbits in large quantities. Our posterior of stable coplanar orbits tightly constrains the planets' orbits, and we discuss implications for the outermost planet b shaping the debris disk. A four-planet resonance lock is not necessary for stability up to now. However, planet pairs d and e, and c and d, are each likely locked in two-body resonances for stability if their component masses are above 6 M_{Jup} and 7 M_{Jup}, respectively. We combined the dynamical and luminosity constraints on the masses using hot-start evolutionary models and found the mass of planet b to be 5.8 ± 0.5 M_{Jup}, and the masses of planets c, d, and e to be $7.2^{+0.6}_{-0.7}$ M_{Jup} each.

5.1 Introduction

High-contrast imaging spatially separates the faint light of planets from the bright glare of their host star. By monitoring exoplanetary systems with high-contrast imaging, we are able to obtain footage of these exoplanets in motion and trace out their orbits. Orbit analysis has been a powerful tool in characterizing the dynamics of directly-imaged systems. Through orbital monitoring of β Pic b, we now know that the planet is responsible for inducing the observed warp in the circumstellar debris disk (Dawson et al. 2011; Lagrange et al. 2012), although it may not be alone in clearing out the cavity of the disk (Millar-Blanchaer et al. 2015). Precise orbital determination also has timed the Hill sphere transit of the planet to between April of 2017 to January of 2018 (Wang et al. 2016), which offered a unique opportunity to probe the circumplanetary environment of a young exoplanet (Stuik et al. 2017; Mékarnia et al. 2017; de Mooij et al. 2017). For HD 95086 b, by combining orbit fits with constraints on the debris disk geometry, Rameau et al. (2016) showed that the planet alone cannot be clearing out the gap in the system, and that additional planets reside closer in to the star. The orbit of Fomalhaut b was shown to cross the debris disk in the system, revealing that the planet cannot be a massive Jupiter-like planet, but rather a dwarf planet shrouded by dust (Kalas et al. 2013). Finally, future orbital monitoring of 51 Eri b could shed light on the interactions between the planet and the wide-separation binary GJ 3305 (De Rosa et al. 2015).

Long-term orbital monitoring can also lead to dynamical mass measurements of the planets themselves, which will assess evolutionary models of young giant planets that all current mass estimates of directly-imaged exoplanets are based on (Baraffe et al. 2003; Marley et al. 2007). In the coming years, Gaia will measure the astrometric reflex motion of stars hosting planets (Perryman et al. 2014). Gaia astrometry combined with long-term orbital monitoring from direct imaging will provide the tightest model-independent constraints on the masses of the planets (Sozzetti et al. 2016). Alternatively, multi-planet systems where planets mutually perturb their orbits provide another way to constrain the masses of the planets in the system. In resonant systems where the dynamical timescales are close to the orbital timescales, such mutual perturbations have been measured in short period planets as variations in the host star's radial velocity signature (e.g., Marcy et al. 2001; Rivera et al. 2010) and as transit timing variations (e.g., Agol et al. 2005; Holman & Murray 2005), leading to direct measurements of the masses. Due to the long orbital periods of known directly-imaged systems, such a direct measurement of the mutual perturbations on the orbits has been impossible with the current observational baselines, none of which span a full orbital period. Still, upper limits on the masses of the planets based on dynamical stability can be obtained. Stability mass constraints have been used to characterize exoplanets discovered in compact systems, such as TRAPPIST-1 (e.g., Quarles et al. 2017; Tamayo et al. 2017), Kepler-36 (Deck et al. 2012), and the HR 8799 system discussed in this paper.

HR 8799 is unique among directly-imaged systems as it is the only one known to harbor four planets (Marois et al. 2008, 2010b). The planets orbit ~15-70 au from the star between two rings of rocky bodies, similar to the configuration of the giant planets in our own Solar System (Su et al. 2009). The outer belt has been resolved with far-infrared and millimeter observations, although the exact orientation and inner edge of the disk are not entirely agreed upon (Hughes et al. 2011; Matthews et al. 2014; Booth et al. 2016; Wilner et al. 2018). Assuming "hot-start" evolutionary models and an age of 30 Myr, Marois et al. (2008, 2010b) translated the planet luminosities into masses: planet b is ~5 $M_{\rm Jup}$ while the inner three planets are ~7 $M_{\rm Jup}$ (Marois et al. 2008, 2010b). However, as the evolutionary models are uncertain at these early ages, so are the exact masses of the planets. Fortunately, dynamics can provide an additional constraint on the masses of the planets, even if their long orbital periods mean we cannot detect planet-planet interactions and fully constrain the masses this way.

Since the discovery of the HR 8799 planets, their orbits have been closely monitored. Keplerian orbits have been fit to the astrometry obtained from many instruments using least-squares techniques that look for families of orbits or Bayesian parameter estimation with Markov-chain Monte Carlo (MCMC) methods that explore the full posterior of orbital parameters (Soummer et al. 2011a; Currie et al. 2012; Esposito et al. 2013; Maire et al. 2015; Pueyo et al. 2015; Zurlo et al. 2016; Konopacky et al. 2016a; Wertz et al. 2017). Fitting the planets independently, some studies have reported planet d to be misaligned in its orbit relative to the other planets (Currie et al. 2012; Esposito et al. 2013; Pueyo et al. 2015) or one of the inner planets having eccentricities above 0.2 (Maire et al. 2015; Wertz et al. 2017). However, several of the authors have noted that unaccounted astrometric calibration offsets between instruments may be inducing inclination and eccentricity biases (Pueyo et al. 2015; Maire et al. 2015; Konopacky et al. 2016a). Recently, Konopacky et al. (2016a) presented self-consistent astrometry using only measurements from Keck and found that coplanar and low-eccentricity solutions were consistent with the data. Despite the uniform analysis, the 7 years of Keck data still only cover a short arc of these orbits that have periods between ∼40-400 years, leaving many possible orbital configurations.

The measured astrometry is not the only constraint on the orbit of these planets. HR 8799 is part of the Columba moving group (Zuckerman et al. 2011), a group of stars that formed together 42^{+6}_{-4} Myr ago (Bell et al. 2015). Thus the four planets need to be stable dynamically for almost the same amount of time. Studies using N-body simulations have explored the dynamical constraints on the orbital parameters and masses. These studies have found stable orbits using the nominal luminosity-derived masses from Marois et al. (2010b) without invoking orbital resonances (Sudol & Haghighipour 2012; Götberg et al. 2016) or to even higher masses assuming long-term resonance lock of the planets (Fabrycky & Murray-Clay 2010; Marois et al. 2010b; Goździewski & Migaszewski 2014; Gozdziewski & Migaszewski 2018). However, many of these studies initialize or fit the simulated orbits to one astrometric measurement, leaving a gap between orbit fits from the data and dynamical constraints from simulations (Fabrycky & Murray-Clay 2010; Marois et al. 2010b; Sudol & Haghighipour 2012; Götberg et al. 2016). To connect simulations to the data more rigorously, Goździewski & Migaszewski (2014) developed a novel technique to lock the planets into resonance and then search for times and orientations that matched all of the available data. Their orbit and mass constraints though only apply to the family of orbits that slowly migrated into a four-planet resonance lock.

A few attempts have been made to include stability in the orbit fitting of this system. Analytical prescriptions have been used to remove the orbits that are most obviously not dynamically stable (Pueyo et al. 2015; Konopacky et al. 2016a). Esposito et al. (2013) ran N-body simulations on their orbital fits from a least-squares algorithm and only found stable orbits up to 5 M_{Jup}. In general, finding stable orbits in the orbit fits has been impractical

Table 5.1: GPI Observations of HR 8799

UT Date	Filter	Exposure Time (s)	Field Rotation (°)	Planets Imaged
2013 Nov 17	*K1*	2130	17	cde
2014 Sep 12	*H*	3107	19	bcd
2016 Sep 19	*H*	3579	21	cde

with short orbital arcs. Having only the 2-D sky projection of an arc of an orbit, even with milliarcsecond-level precision, cannot break many degeneracies in the orbital parameters resulting in too wide a variety of orbital solutions which are nearly all unstable.

In this paper, we present an analysis that better bridges the gap between orbit fits and dynamical constraints by incorporating N-body simulations as a rejection sampling step of our Bayesian orbit fit to enforce stability. In Section 5.2, we show we have obtained one milliarcsecond astrometry of all four planets using the Gemini Planet Imager (GPI; Macintosh et al. 2014) and the open-source `pyKLIP` data reduction package (Wang et al. 2015). In Section 5.3, we combine the precise GPI measurements with the uniformly-reduced Keck astrometry measured by Konopacky et al. (2016a) and fit multiple orbital models with different assumptions about coplanarity and resonance using MCMC techniques that sample the full posterior of possible orbital configurations. In Section 5.4, we take the posteriors of orbits from our Bayesian analysis and simulate them for 40 Myr using the `REBOUND` N-body integrator (Rein & Liu 2012) to find the posterior of stable orbits after applying a dynamical stability prior. We discuss the consequences of our results, such as planets shaping the cold debris disk, the necessity of orbital resonances for stability, dynamical limits on the masses of the planets, and the future stability of the system.

5.2 Observations and Data Reduction

To obtain astrometry of the planets, we used three epochs of observations of HR 8799 taken with the integral field spectroscopy (IFS) mode of GPI. Two epochs were from instrument commissioning (Gemini program GS-ENG-GPI-COM) and one epoch from the GPI Exoplanet Survey (Gemini program GS-2015B-Q-500; PI: Macintosh). Details of the three observations are listed in Table 5.1. While HR 8799 b is normally located outside of the field of view of GPI, we steered the field of view on the detector during the 2014 September 12 observations to see planet b, although the conditions in this dataset were too poor to see planet e.

Raw IFS data from each epoch were processed to create 3-D spectral datacubes using the automated data reduction system for the GPI Exoplanet Survey (Wang et al. 2018). Briefly, the data were dark subtracted, individual micro-spectra on the detector were extracted to form spectral datacubes, bad pixels were corrected, distortion in the image was corrected,

Table 5.2: Astrometric Measurements of the HR 8799 planets

UT Date	Planet	KL Modes	Exclusion Criterion (pixels)	Radial Separation (mas)	Position Angle (°)
2013 Nov 17	c	10	3	949.5 ± 0.9	325.18 ± 0.14
	d	10	3	654.6 ± 0.9	214.15 ± 0.15
	e	20	1.5	382.6 ± 2.1	265.13 ± 0.24
2014 Sep 12	b	10	1.5	1721.2 ± 1.4	65.46 ± 0.14
	c	10	1.5	949.0 ± 1.1	326.53 ± 0.14
	d	10	1.5	662.5 ± 1.3	216.57 ± 0.17
2016 Sep 19	c	10	2	944.2 ± 1.0	330.01 ± 0.14
	d	10	2	674.5 ± 1.0	221.81 ± 0.15
	e	10	1	384.8 ± 1.7	281.68 ± 0.25

and satellite spots, fiducial diffraction spots centered about the location of the star, were located. See Appendix A of Wang et al. (2018) for details.

We used the Karhunen-Loève Image Projection algorithm (KLIP; Soummer et al. 2012; Pueyo et al. 2015) to subtract off the stellar glare and the Bayesian KLIP-FM Astrometry (BKA) technique (Wang et al. 2016) to measure the astrometry of each planet. BKA forward models the distortions to the planet point spread function (PSF) induced by KLIP in subtracting the stellar PSF and fits for the planet position while also accounting for the correlated noise in the image as a Gaussian process. In Wang et al. (2016), we used this technique to obtain one milliarcsecond astrometry on β Pic b. We used the KLIP and BKA implementations available in the `pyKLIP` package (Wang et al. 2015) from commit `4f56e34`. For all the reductions, we first ran a high-pass filter to suppress the low spatial frequency background, constructed the instrumental PSF from the satellite spots, selected an annulus containing each planet to run KLIP on, and averaged the data in time and wavelength. To optimize the detection of each planet, we varied the number of Karhunen-Loève (KL) modes to model the stellar PSF, and the minimum number of pixels the planet needed to move in the reference images due to angular differential imaging (Marois et al. 2006a) and spectral differential imaging (Marois et al. 2000). We listed these parameters in Table 5.2. To measure the planets' astrometry, we used the `emcee` package (Foreman-Mackey et al. 2013) to sample the posterior distribution for the location of the planet while also fitting the noise as a Gaussian process with spatial correlation described by the same Matérn covariance function as used in Wang et al. (2016). For each planet, our Markov-chain Monte Carlo sampler used 100 walkers, and each walker was run for 800 steps, with a "burn-in" of 300 steps beforehand that corresponded to at least three autocorrelation times for any chain. We then added additional terms in our astrometric error budget in quadrature: a 0.05 pixel uncertainty in locating the central star (Wang et al. 2014); a plate scale of 14.166 ± 0.007 mas lenslet^{-1};

and a residual North offset of $0\overset{\circ}{.}10 \pm 0\overset{\circ}{.}13$ (De Rosa et al. 2015). Our final astrometric results are listed in Table 5.2.

We note that it is improper to combine our Bayesian uncertainties on the planet position with the frequentist uncertainties of our astrometric calibration. In Appendix B, we present a frequentist alternative that performs the same analysis, but with a commensurable interpretation.

5.3 Orbit Fitting

To investigate the possible orbital solutions for the HR 8799 planets, we combined our GPI measurements with those from Keck that were reported in Konopacky et al. (2016a). We chose to consider only these two datasets to minimize unknown systematic errors in the astrometric calibration across instruments. Specifically, the astrometric calibration of GPI is derived from the same instrument at Keck that is used for their observations of HR 8799, so systematic offsets between the two datasets are minimized (Konopacky et al. 2014; De Rosa et al. 2015). While *Hubble Space Telescope* data from 1998 provides an additional 6 years of baseline, the 20-30 mas 1σ uncertainties are not particularly constraining so we did not use them (Soummer et al. 2011a).

In this section, we fit the four planet orbits to four orbital configurations with increasing constraints: first, four Keplerian orbits that share the same parallax and stellar mass (Section 5.3.1); second, forcing coplanarity of the four planets (Section 5.3.2); third, forcing the four planets to be near 1:2:4:8 period commensurabilities but with no coplanarity constraints (Section 5.3.3); lastly, forcing both coplanarity and the periods to be near a 1:2:4:8 ratio (Section 5.3.4). The constraints are intended to tighten the parameter space around stable orbits, but we are not directly considering stable orbits in these orbit fits. Dynamical stability constraints will be added in Section 5.4.

5.3.1 Unconstrained Orbits

First, we fit four independent Keplerian orbits to the data. We employed the same Bayesian framework as Wang et al. (2016) that used Markov-chain Monte Carlo (MCMC) to sample the posterior distribution of orbital elements. For each planet, we fit for the conventional Keplerian orbital elements: semi-major axis (a), epoch of periastron after MJD 50,000 in units of fractional orbital period (τ), argument of periastron (ω), longitude of the ascending node (Ω), inclination (i), and eccentricity (e). Our conventions follow those defined in Alzner & Argyle (2012) for binary stars. In this approach each planet's orbital properties are independent, except we require that the four planets' orbits use the same parallax and total system mass, which we take to be the stellar mass. To account for the uncertainties in the parallax and stellar mass, we assumed a Gaussian prior for the system parallax of 24.76 ± 0.64 mas (Gaia Collaboration et al. 2016) and a Gaussian prior for the stellar mass of $1.52 \pm 0.15\ M_{\odot}$, which is mass reported by Baines et al. (2012) but with an

Table 5.3: Orbital Parameters of HR 8799 bcde from Different Models

Body	Orbital Element	Unconstrained	Coplanar	Near 1:2:4:8	Near 1:2:4:8 Coplanar Low-e
b	a_b (au)	$69.5^{+9.3}_{-7.0}$	$66.4^{+4.1}_{-3.6}$	$69.4^{+3.1}_{-4.0}$	$69.5^{+2.6}_{-2.8}$
	τ_b	$0.54^{+0.14}_{-0.16}$	$0.46^{+0.05}_{-0.06}$	$0.40^{+0.11}_{-0.15}$	$0.38^{+0.11}_{-0.12}$
	ω_b (°)	92^{+30}_{-34}	92 ± 15	95^{+49}_{-41}	102^{+39}_{-46}
	Ω_b (°)	127^{+32}_{-93}	126^{+12}_{-14}	82^{+36}_{-16}	78^{+13}_{-10}
	i_b (°)	29^{+7}_{-8}	23 ± 5	23^{+4}_{-6}	24 ± 3
	e_b	0.15 ± 0.05	0.15 ± 0.06	$0.07^{+0.06}_{-0.05}$	$0.05^{+0.04}_{-0.03}$
c	a_c (au)	$37.6^{+2.2}_{-1.7}$	$40.5^{+2.7}_{-1.7}$	$41.2^{+2.3}_{-1.6}$	$43.3^{+1.9}_{-1.7}$
	τ_c	$0.50^{+0.10}_{-0.18}$	$0.14^{+0.18}_{-0.11}$	$0.13^{+0.15}_{-0.09}$	$0.09^{+0.02}_{-0.07}$
	ω_c (°)	65^{+59}_{-29}	52^{+83}_{-35}	48^{+62}_{-29}	63^{+60}_{-29}
	Ω_c (°)	110^{+38}_{-47}	126^{+12}_{-14}	112^{+17}_{-26}	78^{+13}_{-10}
	i_c (°)	20^{+4}_{-5}	23 ± 5	21^{+3}_{-4}	24 ± 3
	e_c	0.09 ± 0.04	$0.05^{+0.05}_{-0.03}$	$0.04^{+0.05}_{-0.03}$	$0.03^{+0.04}_{-0.02}$
d	a_d (au)	$27.7^{+2.2}_{-1.7}$	$25.3^{+1.3}_{-1.1}$	$25.6^{+1.2}_{-1.3}$	$25.6^{+1.0}_{-0.9}$
	τ_d	$0.79^{+0.07}_{-0.18}$	$0.872^{+0.019}_{-0.016}$	0.85 ± 0.03	0.839 ± 0.20
	ω_d (°)	144^{+13}_{-23}	133^{+15}_{-11}	148^{+22}_{-137}	165^{+11}_{-157}
	Ω_d (°)	92^{+27}_{-15}	126^{+12}_{-14}	86^{+26}_{-16}	78^{+13}_{-10}
	i_d (°)	33 ± 4	23 ± 5	23^{+5}_{-6}	24 ± 3
	e_d	0.15 ± 0.11	0.28 ± 0.04	0.20 ± 0.05	$0.18^{+0.02}_{-0.03}$
e	a_e (au)	$15.3^{+1.4}_{-1.1}$	$14.0^{+0.7}_{-0.6}$	$15.7^{+0.06}_{-0.07}$	15.4 ± 0.06
	τ_e	$0.71^{+0.17}_{-0.33}$	$0.91^{+0.05}_{-0.06}$	$0.10^{+0.15}_{-0.06}$	$0.07^{+0.05}_{-0.04}$
	ω_e (°)	100^{+27}_{-49}	128^{+25}_{-18}	86^{+44}_{-36}	76^{+25}_{-21}
	Ω_e (°)	117 ± 17	126^{+12}_{-14}	90^{+9}_{-19}	78^{+13}_{-10}
	i_e (°)	31 ± 5	23 ± 5	28^{+3}_{-4}	24 ± 3
	e_e	$0.13^{+0.06}_{-0.05}$	0.16 ± 0.05	$0.05^{+0.07}_{-0.04}$	$0.08^{+0.03}_{-0.04}$
A	Parallax (mas)	24.70 ± 0.16	24.60 ± 0.56	24.38 ± 0.62	$24.38^{+0.54}_{-0.53}$
	$M_\star$ ($M_\odot$)	$1.48^{+0.05}_{-0.04}$	$1.46^{+0.12}_{-0.11}$	$1.42^{+0.12}_{-0.11}$	1.40 ± 0.10
	χ^2_ν	$1.01^{+0.10}_{-0.07}$	0.88 ± 0.07	$0.95^{+0.07}_{-0.06}$	$0.94^{+0.12}_{-0.05}$
	ΔBIC	0^{+7}_{-6}	-34 ± 0.06	-4^{+5}_{-6}	-29^{+6}_{-4}
Stable Orbits (first 10^6 draws)		0	0	1	441

Note. — The quoted values for ω and Ω are wrapped to be between 0° and 180° so posterior percentiles describe one of the two symmetric peaks. For each parameter, the median value is reported with the superscript and subscript corresponding to the 84th and 16th percentiles of the distribution respectively. For a normal distribution, these values correspond to the mean and 1σ range.

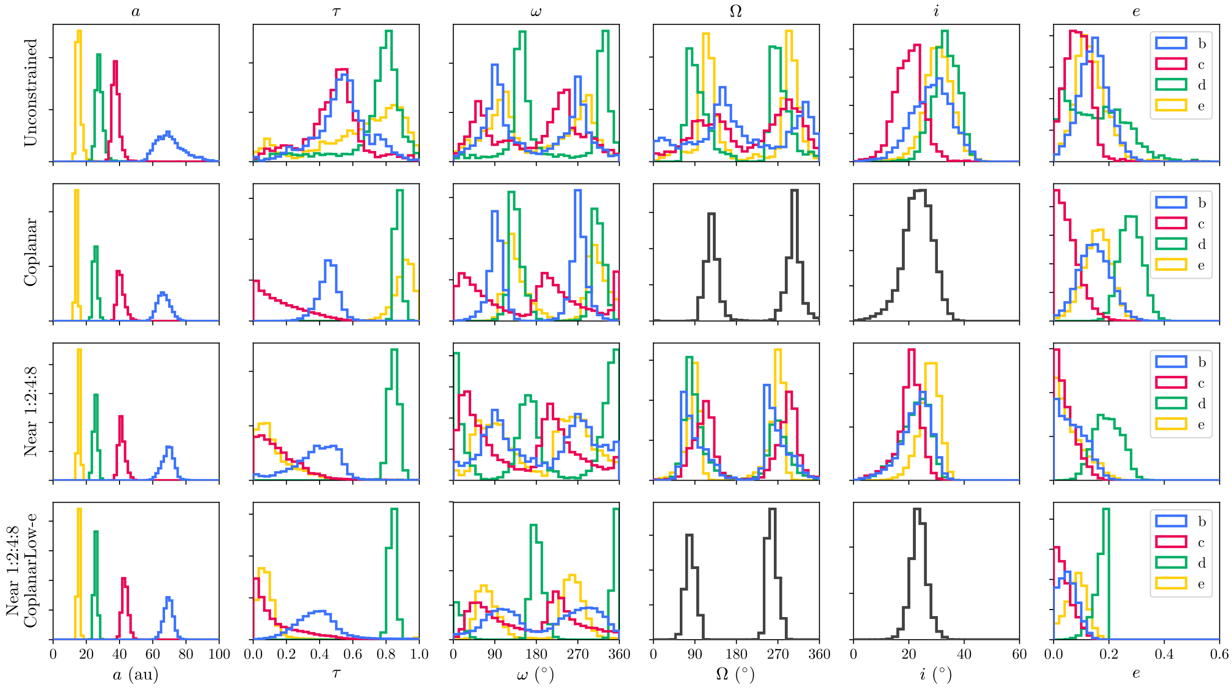


Figure 5.1: The posteriors of each planet's orbital parameters for each of the four different models considered in Section 5.3. Each row contains the four planet's posteriors (color coded by planet) for one model. For the coplanar models, the planets have the same Ω and i, so only one is plotted.

additional 10% uncertainty to account for systematic model errors as was done in Konopacky et al. (2016a). This case covers the full range of orbital parameters that are consistent with the data; the three following orbit fits will explore subsets of this parameter space. Due to the high dimensionality of the orbital parameters (26 in total), it will be incredibly difficult to find the dynamically stable orbits if they reside in a very small subspace. Regardless, this orbital fit is an important fiducial case to be used as a baseline model with minimal assumptions. We will refer to this orbital fit as the “Unconstrained" fit.

We generally used uniform priors on our orbital parameters. For each planet, the prior on a was uniform in $log(a)$ between 1 and 100 au; the prior on τ was uniform between 0 and 1; the priors on ω and Ω were uniform from 0 to 2π; the prior on i is the geometric $\sin(i)$ prior; and the prior on e was uniform between 0.000001 to 0.999.

We used the parallel-tempered affine-invariant sampler (Goodman & Weare 2010) implemented in `emcee` (Foreman-Mackey et al. 2013) using 15 temperatures and 1500 walkers per temperature. To improve the speed of convergence of the orbit fit, we initialized the walkers by drawing from allowed orbital parameters of individual fits to each planet using the same process. We ran each walker for 125,000 steps, after an initial burn in of 95,000 steps. Convergence was assessed using the autocorrelation time and confirming by-eye that ω and Ω had symmetric peaks. On a 32 core machine with AMD Opteron 6378 processors clocked at 2.3 GHz, this took seven days to complete, although we note that we did not make an attempt to optimize the code. We then thinned the chains by a factor of 75 to mitigate any correlation in the Markov chains. Taking only the lowest temperature walkers, we then were left with 2,499,000 samples of the posterior distribution. The posterior distributions are plotted in Figure 5.1 and reported in Table 5.3.

Following similar analyses from previous orbit fitting studies (e.g., Konopacky et al. 2016a; Wertz et al. 2017), we investigate the mutual inclination of the planets' orbits by plotting in Figure 5.2 Ω and i, the two orbital elements that describe the orientation of the orbital plane. We see that the 1σ contours for the four planets do overlap near $i \sim 30^\circ$ and $\Omega \sim 100^\circ$, indicating coplanar orbital solutions exist. This result agrees with the assessment of coplanarity by Konopacky et al. (2016a) using similar arguments, although they preferred a different Ω. However, we note that only 0.005% of our sampled orbits have all four planets being mutually inclined by $< 10^\circ$. This result likely indicates that without any constraints on the orientation of the orbital planes, it is extremely inefficient to sample coplanar orbits in large quantities. This is not surprising since the near-coplanar solutions are just a small subset of an eight-dimensional space in which we have chosen uniform, uncorrelated priors on each parameter. To more rigorously test coplanar orbits, we will fit directly for them (Section 5.3.2) and assess the fits (Section 5.3.5).

5.3.2 Coplanar Orbits

As planets form from the circumstellar disk, it would not be surprising to find the planets residing in coplanar orbits. The posteriors from the fit without constraints are consistent with coplanarity, but does not strongly favor it. Here we will explicitly fit for coplanar orbits,

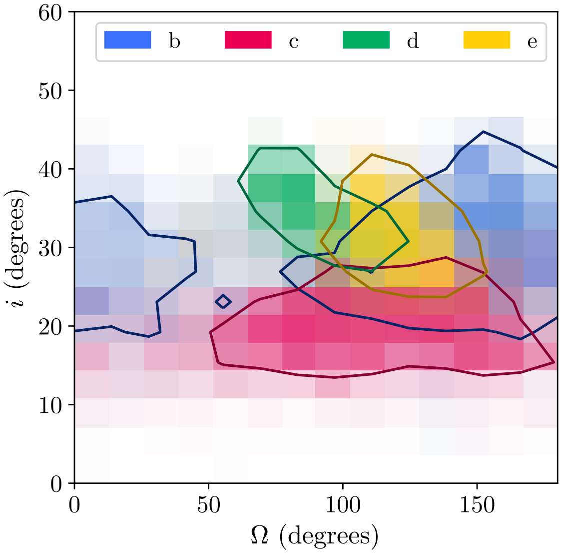


Figure 5.2: The posteriors of each planet's angles Ω and i for the Unconstrained fit. Blue, magenta, green, and yellow correspond to planets b, c, d, and e respectively. The 1σ contour is plotted on top of each planet's histogram. Overlapping regions indicate where coplanar orbits reside. Note that Ω is wrapped to only consider angles between 0° and 180° as the posterior is identical as between 180° and 360° by construction.

and in Section 5.3.5, we will assess if this approach fits the data as well as the unconstrained one. We modify our fit so that all four planets share the same values of Ω and i, reducing the fit to 20 orbital parameters. We will refer to this orbital fit as the "Coplanar" fit.

We used a parallel-tempered sampler with 15 temperatures and 1500 walkers per temperature. We ran each walker for 87,500 steps, after an initial burn in of 132,500 steps. Convergence was assessed in the same way as in Section 5.3.1. We again thinned the chains by a factor of 75, and formed our posterior distribution from the lowest temperature chains. Our posterior distribution has 1,749,000 samples. The posteriors are plotted in Figure 5.1 and reported in Table 5.3.

From the posteriors in Figure 5.1, we see that the angles Ω and i that define the orientation of the orbital plane are consistent with the orbital planes of the four planets of the Unconstrained fit. The Coplanar orbits favor inclinations between 20° and 30°, which is ∼10° more face-on than the solutions from Konopacky et al. (2016a) with just the Keck data alone. We still find $\Omega > 90°$, which is not preferred for coplanar orbits in Konopacky et al. (2016a). We also note that forcing the system to be coplanar causes the eccentricity of planet d to be much higher, with $< 2\%$ of the allowed orbits having $e < 0.2$. This was due to nearly all of planet d's low eccentricity orbits from the Unconstrained fit lying outside of the range of allowed orbital planes from the Coplanar fit. As Ω and i are constrained by the other three planets, raising e_d provided a way to obtain the best fits to the data.

5.3.3 Near 1:2:4:8 Period Ratio Orbits

We then investigated resonant orbits, focusing in particular on the 1:2:4:8 resonance, where consecutive pairs of planets are in 2:1 period resonance. We will first choose to be agnostic about the four planets' mutual inclinations. Because these planets are not massless, even if they are in resonance, they do not necessarily reside at the exact period commensurabilities. Additionally, precession of the planets' longitude of periastrons can further offset the observed period ratios from exact integer values. We note that previous orbit fitting work has assumed exact period commensurabilities when assessing if the fits were consistent with certain resonances.

At high planet masses like the HR 8799 planets, stable period ratios for the 2:1 two-body resonance tend to be larger than 2 due to resonance overlaps at smaller period ratios causing instability (Morrison & Kratter 2016). Thus, instead of fixing the period ratio of the planet pairs, we use a parameter that gives each period ratio room to float. We picked our priors empirically from our own preliminary analysis of where the stable orbits existed. Our prior on the period ratio between b:c and c:d is a uniform distribution between 1.8 and 2.4. For the d:e period ratio, we choose a narrower uniform prior between 1.8 and 2.2, because we found all of the dynamically stable orbits were in this more narrow range and limiting it as such improved the efficiency of finding dynamically stable orbits (Section 5.4). We will show in Section 5.4.2 that our choices for our priors did not exclude stable orbits. We note that we effectively replaced the parameters for the semi-major axes of the outer planets with their period ratios, so we did not reduce the number of parameters in our MCMC fit even though the parameter space has shrunk. We will refer to this orbital fit as the "Near 1:2:4:8" fit, which as the naming implies, only places the period ratios near resonance and does not guarantee the planets are indeed in resonance at all.

We initialized the walkers using coplanar solutions, which delayed convergence and caused the walkers to take a considerable amount of time to fully explore all of the allowed parameter space. We ran our parallel-tempered sampler with 15 temperatures and 1500 walkers per temperature for 75,000 steps, after a burn in of 495,000 steps that was chosen using the same metric for convergence as Section 5.3.1. We performed the same thinning of the chains by a factor of 75. The resulting posterior was taken from the lowest temperature walkers and has 1,500,000 samples. The posteriors are plotted in Figure 5.1 and reported in Table 5.3.

5.3.4 Near 1:2:4:8 Period Ratio Coplanar Low-e Orbits

Lastly, we looked at coplanar resonant orbits. We applied both the coplanarity and period ratio constraints from Sections 5.3.2 and 5.3.3. We also applied an additional constraint that the eccentricity of all of the orbits have to be less than 0.2. In Section 5.3.2, we found that $< 2\%$ of the coplanar orbits have $e_d < 0.2$. From preliminary analysis done concurrently with the orbit fits, we could only find stable orbits when all planets had $e < 0.2$. This fact will be further reinforced by the analysis in Section 5.4.2. Thus, we do not believe we lost stable orbits by applying this additional constraint, and merely improved the efficiency of finding stable orbits. We will refer to this orbital fit as the “Near 1:2:4:8 Coplanar" fit, which, like the Near 1:2:4:8 fit, does not guarantee the planets are actually in resonance.

For this 20-parameter orbit fit, we used a parallel-tempered sampler with 15 temperatures and 1500 walkers per temperature. We ran each walker for 125,000 steps, after an initial burn in of 95,000 steps. Convergence was confirmed using the metrics defined in Section 5.3.1. We again thinned the chains by a factor of 75, and formed our posterior distribution from the lowest temperature chains. This resulted in 2,499,000 samples of the posterior. The posteriors are plotted in Figure 5.1 and reported in Table 5.3.

As we expected, the posterior for the eccentricity of planet d runs up right against our prior bounds. Without stability constraints, higher eccentricity orbits are favored. Just like in the Coplanar orbit fit, we find an orbital inclination for the system in the 20° and 30° range. However, Ω is now in agreement with that found in Konopacky et al. (2016a) for coplanar orbits, unlike our previous orbit fits. It is likely this was a small family of orbits that were not represented in the 1σ range of our previous analyses.

5.3.5 Goodness of Fit

We used the reduced chi-squared (χ^2_ν) statistic to measure the goodness of fit of a model. Since the highest likelihood model often does not represent the whole posterior of possible orbital configurations, we compute χ^2_ν on 1000 randomly drawn allowed orbits for each model. We list the 16th, 50th, and 84th percentiles in Table 5.3. For the Unconstrained model, we found $\chi^2_\nu \approx 1$, indicating the unconstrained Keplerian orbits can suitably describe the data as one might expect if the uncertainties are estimated accurately, given it is a physical model. The other three models have χ^2_ν similarly close to unity, showing they also fit the data well.

We also investigated if more-restrictive models with additional, dynamically-motivated constraints better describe the data than the fiducial Unconstrained case. We calculated the Bayesian Information Criteria (BIC; Schwarz 1978; Liddle 2007) as a simplified alternative to full Bayesian model comparison. The BIC assesses how well a model fits the data and penalizes models that have more free parameters. Models with lower BIC are preferred. We define the ΔBIC as the difference between the BICs of a more restrictive model and the median BIC of the Unconstrained model. We also calculate ΔBIC using the same 1000 randomly drawn orbits for each model, and list the 16th, 50th, and 84th percentiles for this value in Table 5.3.

We find the ΔBIC is negative for the other three models relative to the Unconstrained fit. This indicates that adding constraints that tend the data towards what we believe are stable orbits makes the fits better, as we discard some parameter space containing likely unstable orbits that do not reflect reality. We also note that χ^2_ν and ΔBIC are not perfect metrics as they only consider the number of free parameters in the models, and not the total parameter space being considered. In particular, when we limit the period ratios, this does not decrease the number of free parameters while significantly limiting the space of possible orbits. Thus, we see these goodness of fit metrics favor the coplanar solutions as they explicitly reduce the number of parameters in the model. It would be better to have computed the Bayes factor between each pair of models to more rigorously compare models, but the Bayes factor is computationally difficult to calculate with a high-dimensional problem like this and our MCMC samplers were only set up to perform parameter estimation. Because of this, we do not think it is valid to conclude from solely these two metrics that coplanar orbits are favored. However, we can assert that adding constraints to the orbit fit does not worsen the fit from the fiducial case, and thus the constraints are reasonable given the current astrometric data. This conclusion agrees with the analysis from Konopacky et al. (2016a), who found that coplanar orbits and orbits near 1:2:4:8 period ratios were fully consistent with the Keck astrometry. While the metrics we have employed cannot decide which orbit model should be favored, the stability constraints to the system that are investigated in the following section will clearly show what the realistic orbits are.

5.4 Dynamical Constraints

Keplerian motion is not the only constraint on the orbits of the planets. We also know that these four planets must also have survived from their formation up to this point. The star is 42^{+6}_{-4} Myr old (Bell et al. 2015), so these orbits must have also been stable for roughly the same time, since giant planets likely formed quickly before the gas disk dispersed in the first few Myr (Williams & Cieza 2011). As the gas disk is difficult to model and exists for only a short period of the system's lifetime, we do not simulate the time between planet formation and gas disk dispersal. Thus, we investigated which orbital configurations allowed by astrometric measurements are also stable if we simulate the four planets' orbits backwards in time for 40 Myr. In this section, we will apply this dynamical constraint on each of the four orbit fits (Unconstrained, Coplanar, Near 1:2:4:8, Near 1:2:4:8 Coplanar), and investigate the family of stable orbits that arise.

5.4.1 Stability of Orbital Models

We used the `REBOUND` N-body simulation package (Rein & Liu 2012) with the WHFast integrator (Rein & Tamayo 2015). To set up a simulation, we added particles for planets e, d, c, b in that order, using a chosen set of orbital parameters from our fits and placing them at the predicted location on MJD 56609, the date of the first GPI epoch. We drew masses for each of the planets in a process described in the following paragraph, and set the primary mass to be the stellar mass from our orbit fits. We then reversed the present velocities of the planets and integrated the system for 40 Myr to simulate the past dynamical history of the system, using fixed timesteps equal to 1% of planet e's initial orbital period. We considered a configuration unstable if two planets passed too close, or if one planet was ejected from the system. We considered an encounter too close if any two planets passed with a distance less than the initial mutual Hill radius of planets d and e, which we approximated as

$$R_{Hd,e} = a_e \left(\frac{M_e + M_d}{3M_\star} \right)^{1/3}, \tag{5.1}$$

where $M_e, M_d, M_\star$ refer to the masses of planet e, planet d, and the primary respectively. We considered a planet to be ejected if it moved further than 500 au from the star. Any orbit that survived for 40 Myr without encountering either condition is considered stable.

To assess the dynamical stability of allowed orbital configurations for each model, we performed rejection sampling to assess which orbit models contained significant amounts of stable orbits. We drew one million random orbital configurations from each of the four model posteriors. The orbit fits do not specify the mass of the planets so we needed to add additional parameters for them. For simplicity, we set planets c, d, and e to be equal in mass as we would expect due to their similar luminosities (Marois et al. 2010b). We drew the mass of these planets, M_{cde}, from a uniform prior between 4 and 11 $M_{\rm Jup}$ to encompass the uncertainty on the luminosity-derived masses from Marois et al. (2010b). We drew the mass of planet b, M_b, from a uniform prior between 3 $M_{\rm Jup}$ and M_{cde} to account for its lower luminosity. We will discuss using a more informative prior based on the planets' luminosities in Section 5.4.4. We ran each of the configurations through the `REBOUND` setup described previously. Our dynamical stability prior sets the probability of an orbit to be 0 if the system is not stable, and 1 if it is stable, discarding the unstable orbits in our rejection sampling. In Table 5.3, we record the number of stable orbits from each of the configurations.

We found that the Unconstrained and Coplanar orbit solutions did not yield any stable orbits after one million draws. Especially for the Unconstrained case, the lack of stable draws does not mean that these models are inconsistent with stable orbits, but rather that the islands of stability in this high dimensional space are small and were not sampled even after millions of MCMC draws. Simply, these models do not currently allow for a practical search of stable orbits.

Both models that assume that the planets' orbital periods are near the 1:2:4:8 period ratio do yield stable orbits, with the model not assuming coplanarity, the Near 1:2:4:8 model, resulting in just one stable orbit after one million draws. This model encompasses all of the parameter space explored by the Near 1:2:4:8 Coplanar model so it is the more general model. We explored this model further by running twenty million `REBOUND` simulations in total, leading to 50 stable orbits.

We find that for masses of the inner three planets greater than 5 M_{Jup}, the maximum mutual inclination between any pair of planets in a stable system is $< 8°$, although we only have few samples in this regime (14 stable orbits spanning 8 unique present-day orbital configurations). We also do not find stable orbits above 6 $M_{\rm Jup}$, which likely reiterates the difficulty of finding stable noncoplanar orbits due to the high-dimensionality of the problem. Thus, with the limited orbital arcs we have so far, looking for noncoplanar stable orbits is impractical. Since in Section 5.3 we found that our astrometry is consistent with the system being coplanar, we will focus on those orbits since we can find many stable orbits with this assumption (hundreds per million tries). We will leave the thorough exploration

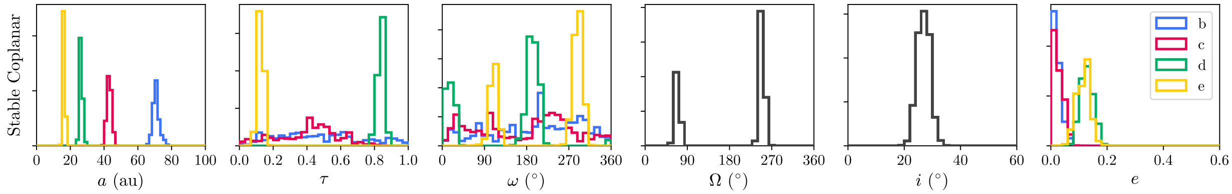


Figure 5.3: Posterior of stable orbital elements for coplanar configurations of the four planets. These posteriors show all stable orbits with $M_{cde} > 4\ M_{\rm Jup}$ and $3\ M_{\rm Jup} < M_b < M_{cde}$. As discussed in Section 5.4.2, solutions with higher planet masses lie within a smaller region of this space.

of orbits with mutual inclinations for future work with longer astrometric baselines and more computation time. However, in our preliminary analysis, it seems that the mutual inclinations are probably small in order for the system to be stable.

5.4.2 Stable Coplanar Orbital Solutions

For the rest of the analysis, we focus on the orbital parameters from the Near 1:2:4:8 Coplanar fit. We increase the number of N-body simulations from one million to twenty-two million, obtaining 9792 stable orbital configurations. We plot the initial osculating orbital elements (i.e., those on MJD 56609) of these stable orbits in Figure 5.3 and list them in Table 5.4.

We find that the posteriors have tightened significantly after applying the dynamical stability constraint. Figure 5.4 visually compares the spread of possible orbits on the 2-D sky plane for the orbit fits with increasing constraints placed on them. The stable coplanar orbits appear as a well defined ellipse with minimal uncertainty for each planet's orbit. This is also reflected visually and numerically in the posterior percentiles. The middle 68%, the difference between the 84th and 16th percentiles, of the semi-major axes of the planets decreased by 1.5 to 4.5 times when compared to the Unconstrained case, and by a factor of 1.17 to 1.50 when compared to the Near 1:2:4:8 Coplanar fit that the stable orbits were drawn from. Similarly, the middle 68% of the eccentricities also decreased by a factor between 2.2 and 4.7 compared to the Unconstrained fits. In fact, the fractional uncertainty on the semi-major axes is about the same as the fractional uncertainty of the Gaia DR1 parallax of the system (≈2%). The inclusion of the parallax from Gaia Data Release 2, released after this analysis was completed, should reduce its contribution to the semi-major axis and total system mass uncertainties by a factor of 7 (Gaia Collaboration et al. 2018). We have chosen not to rerun our analysis since the conclusions in this paper do not strongly depend on the exact semi-major axes of the orbits, and we will leave this for a future work.

The stable orbits, despite being much more restrictive, are good fits to the data. Aggregating 1000 random orbits, we find a $\chi^2_\nu = 1.01^{+0.06}_{-0.05}$ that is just as good as the fiducial Unconstrained model. The ΔBIC is similarly comparable to the other models. Although, once again we note that BIC does not account for the narrower parameter space due to the

Table 5.4: Stable Coplanar Orbital Parameters of HR 8799 bcde

Body	Orbital Element	Stable Coplanar
b	a_b (au)	$70.8^{+0.19}_{-0.18}$
	τ_b	$0.46^{+0.31}_{-0.26}$
	ω_b ($^\circ$)	87 ± 58
	Ω_b ($^\circ$)	$67.9^{+5.9}_{-5.2}$
	i_b ($^\circ$)	26.8 ± 2.3
	e_b	$0.018^{+0.018}_{-0.013}$
c	a_c (au)	$43.1^{+1.3}_{-1.4}$
	τ_c	$0.43^{+0.15}_{-0.24}$
	ω_c ($^\circ$)	67^{+59}_{-39}
	Ω_c ($^\circ$)	$67.9^{+5.9}_{-5.2}$
	i_c ($^\circ$)	26.8 ± 2.3
	e_c	$0.022^{+0.023}_{-0.017}$
d	a_d (au)	$26.2^{+0.9}_{-0.7}$
	τ_d	$0.839^{+0.020}_{-0.017}$
	ω_d ($^\circ$)	17^{+12}_{-11}
	Ω_d ($^\circ$)	$67.9^{+5.9}_{-5.2}$
	i_d ($^\circ$)	26.8 ± 2.3
	e_d	$0.129^{+0.022}_{-0.025}$
e	a_e (au)	16.2 ± 0.5
	τ_e	$0.124^{+0.019}_{-0.013}$
	ω_e ($^\circ$)	110 ± 9
	Ω_e ($^\circ$)	$67.9^{+5.9}_{-5.2}$
	i_e ($^\circ$)	26.8 ± 2.3
	e_e	$0.118^{+0.019}_{-0.028}$
A	Parallax ($''$)	$24.30^{+0.49}_{-0.69}$
	$M_\star$ ($M_\odot$)	$1.47^{+0.11}_{-0.08}$
	χ^2_ν	$1.01^{+0.06}_{-0.05}$
	ΔBIC	-22^{+6}_{-4}

Note. — Values are reported in the same way as Table 5.3

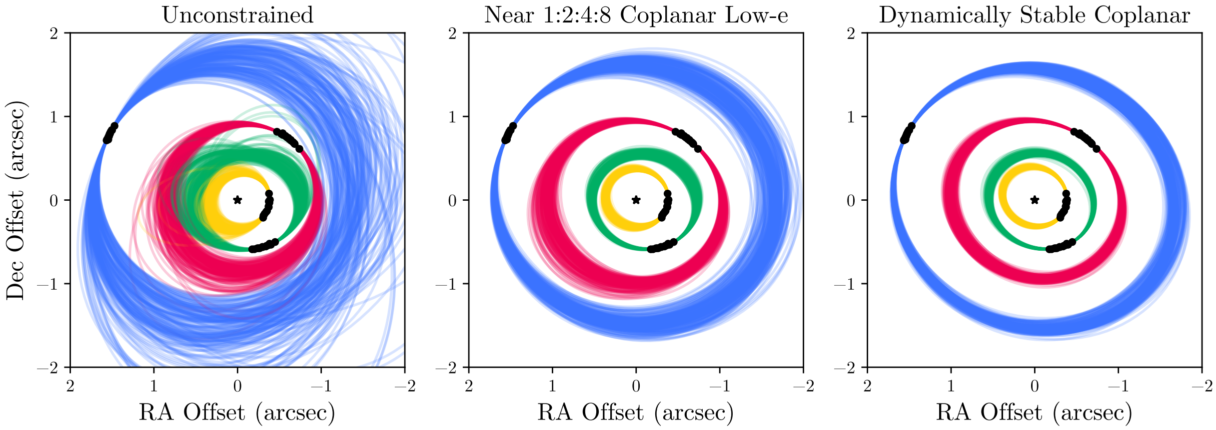


Figure 5.4: A comparison of 200 allowed orbits from the Unconstrained (Section 5.3.1), Near 1:2:4:8 Coplanar (Section 5.3.4), and dynamically stable coplanar solutions (Section 5.4.2) projected onto the sky plane. The black star in the middle represents the location of the star, the black circles are the measured astrometry (uncertainties too small to show on this scale), and the current orbit for each planet is colored in the same way as Figures 5.1 and 5.3 (i.e., planet b is blue, c is red, d is green, and e is yellow).

additional stability constraint. We conclude that these stable orbits are a small, but allowed part of a much larger space that we have explored through our Bayesian analysis.

The masses of the stable configurations are plotted in Figure 5.5. We will discuss mass constraints in Section 5.4.4 in detail. Briefly here, we can see that stable orbits exist with the mass of the inner three planets at almost 9 M_{Jup}, and separately with the mass of planet b to be nearly 7 M_{Jup}. Also, the majority of stable orbits we found are low mass. 95.6% of the orbits have $M_{cde} < 6$ M_{Jup} and 73.0% of the orbits have $M_{cde} < 5$ M_{Jup}. This highlights the difficulty in finding stable high mass solutions when starting with our current orbit fits.

With these stable orbits, we can look at the mass dependence on the orbital parameters to justify our choices of prior constraints in doing the Near 1:2:4:8 Coplanar orbit fit. In Figure 5.6, we plot the range of period ratios and eccentricities of stable orbits as a function of M_{cde}. We see that for $M_{cde} > 6$ M_{Jup}, none of the period ratios or the eccentricities are close to the bounds set by our priors. Below 6 M_{Jup}, the period ratio of planet d to e as well as the eccentricities of d and e are near the upper bound in the extreme case, indicating our prior may be excluding some low-mass stable orbits. Since the interquartile range of these parameters is far away from these bounds, only a few extreme low-mass cases have been excluded, so the effect should be minimal. As the masses increase, we see the range in the allowed parameter space decreases, indicating that the highest mass stable orbits reside in a subspace of the parameters we are exploring. Thus, we conclude that we are not unnecessarily excluding stable orbit configurations with our choice of priors that were designed to improve the efficiency of finding stable orbits.

There are several notable features in our posteriors of stable orbital configurations. The

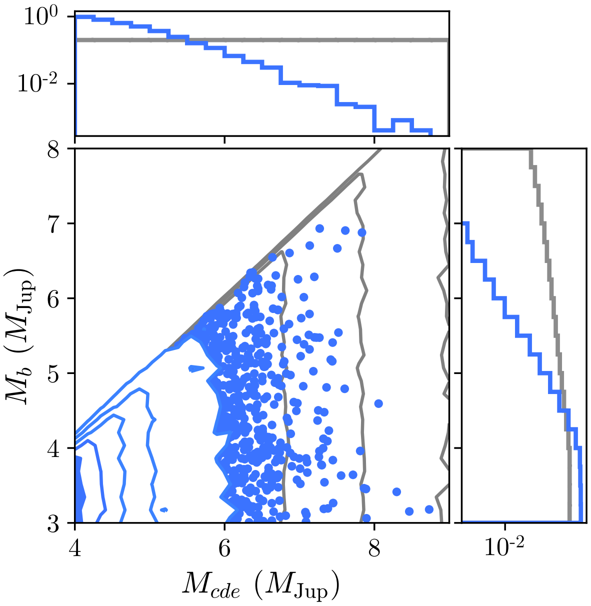


Figure 5.5: The distribution of masses of the stable orbits from Section 5.4.2 (blue) and comparison to the priors from which the masses were drawn (gray). The main plot in the bottom left shows the 2-D distribution of masses. The contour lines represent 15th, 35th, 55th, 75th, and 95th percentiles of the distribution, with everything outside the 95th percentile plotted individually as points. The top and right panels show 1-D histograms for M_{cde} and M_b respectively, with the frequency in each bin plotted on a logarithmic scale to highlight the high mass bins. The gray priors are plotted in the same fashion as the blue posteriors.

bimodality of the eccentricity posteriors is clear. The outer planets b and c have $e \sim 0$ while the inner planets d and e have $e \sim 0.1$. These eccentricities agree well with what was found by Goździewski & Migaszewski (2014) who migrated planets into resonance lock, rotated the orientations to match the astrometry, and selected orbital configurations with a χ^2_ν cutoff. Given that this conclusion was reached by two completely different analysis methods, the fact the inner two planets have slightly eccentric orbits while the outer two planets are in near-circular orbits is a notable result that seems to be required for most stable orbital configurations that are consistent with the measured astrometry. The increased eccentricities of planets d and e, and the proximity of all four planets to 1:2:4:8 period commensurabilities, are consistent with an early evolutionary period of convergent inward migration of all four planets, trapping of planet pairs d & e and c & d into 2:1 resonances, and pumping of the orbital eccentricities of d and e by continued migration while in resonance lock (e.g., Yu & Tremaine 2001; see also section 5.4.3).

Also, comparing our stable orbits with those of Goździewski & Migaszewski (2014) and Gozdziewski & Migaszewski (2018), we note that most orbital parameters agree fairly well except for the semi-major axes of the planets, which we find to be significantly larger.

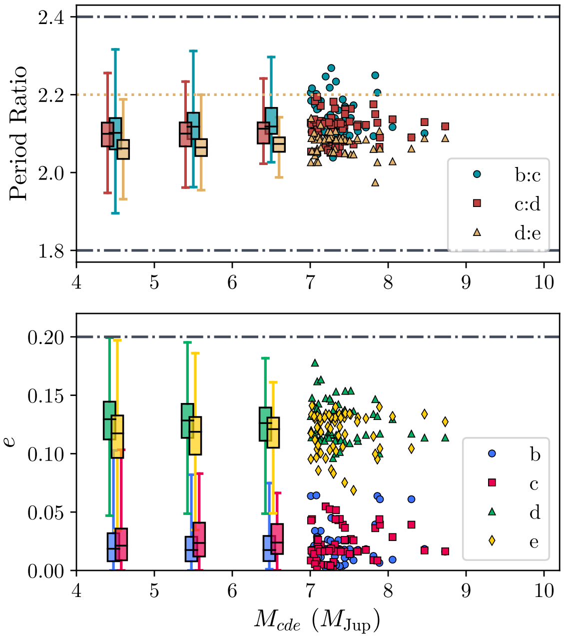


Figure 5.6: The distribution of current period ratios and eccentricities as a function of mass of the inner three planets for stable orbits. For $M_{cde} < 7\ M_{\rm Jup}$, the data is binned into one box plot per $M_{\rm Jup}$. Each box shows the 25th, 50th, and 75th percentiles of the given distribution, while the whiskers show the extrema. Above 7 $M_{\rm Jup}$, points are plotted individually as they are sparse enough. The bounds of the priors are plotted as gray dot-dashed lines, except for the upper bound of the d:e period ratio which is the yellow dotted line. These plots show how the range of allowed period ratios and eccentricities decrease as planet mass increases. Above $M_{cde} > 6\ M_{\rm Jup}$, the full range of stable orbits are not near the prior bounds.

For example, only 0.11% of our orbital solutions have $a_c \leq 39.4$ au, the best fit solution of Goździewski & Migaszewski (2014). Our uncertainties in parallax and stellar mass are consistent with the fixed parallax and stellar mass they used, so the difference in a is not just a result of different system parameters. We also generally have larger uncertainties on our values, which can be due to a combination of allowing lower mass orbits, not strictly enforcing 1:2:4:8 resonance lock, and a more systematic exploration of parameter space.

As the orbits are coplanar, each planet's argument of periastron can be used to measure the relative orientation of the planets' orbits. As both planet b and c have extremely low eccentricities, ω is basically unconstrained for these planets and is not notable since their orbits are near circular. The significant nonzero eccentricity of planets d and e however correspond to sharp peaks in ω for both planets. While a broader peak was already seen in ω_d in the Near 1:2:4:8 Coplanar fits before enforcing stability constraints, the addition of dynamical stability has disallowed circular orbits of planet e, giving rise to a sharp peak in ω_e. Interestingly, the orientation of the orbits of planets d and e are not aligned, with $\omega_e - \omega_d = 94^{+11}_{-9}$ degrees, essentially perpendicular to being aligned. We note that Goździewski & Migaszewski (2014) also found a similar result.

The period ratios of the planets shown in Figure 5.6 heavily favor period ratios above the nominal 2:1. We find period ratios $P_b/P_c = 2.11 \pm 0.06$, $P_c/P_d = 2.10^{+0.04}_{-0.05}$, and $P_d/P_e = 2.06^{+0.03}_{-0.04}$. For P_c/P_d and P_d/P_e, the data favors period ratios above two. This can been seen by computing the period ratios using the median a for the Near 1:2:4:8 Coplanar fits from Table 5.3 that do not have a dynamical prior applied. It seems that these period ratios are at these high values to satisfy the astrometry. For P_b/P_c, the data allows both period ratios above and below two, so having it strongly favor values above 2 (only 2% of stable configurations have $P_b/P_c < 2$) indicates that spacing the two planets slightly further apart enhances stability. The period ratios driven by the astrometry could be indicative of a primordial period ratio. In particular, the planets could have experienced eccentricity damping while in resonance and were repelled to period ratios greater than 2 while still maintaining resonance lock (Lithwick & Wu 2012; Batygin & Morbidelli 2013). As disk gas is the primary mechanism for eccentricity dissipation of Jupiter-mass planets at large-separations,[1] this may indicate that the planets were in or near their current location during the gas disk stage (Dong & Dawson 2016).

We found the system has an inclination of $i = 26.^\circ8 \pm 2.^\circ3$ and longitude of ascending node of $\Omega = 68.^\circ0^{+5.9}_{-5.3}$, consistent with the work by Konopacky et al. (2016a) fitting coplanar orbits, but a few times more precise. Both i and Ω match the debris disk inclination of $26^\circ \pm 3^\circ$ and position angle of $62^\circ \pm 3^\circ$ derived from far-infrared *Herschel* observations (Matthews et al. 2014). While the inclination is also consistent with the millimeter observations of the debris disk by the Submillimeter Array (SMA) and the Atacama Large Millimeter/submillimeter Array (ALMA), Ω is higher than the position angle of the disk of $35.^\circ6^{+9.4}_{-10.1}$ in the millimeter (Wilner et al. 2018). This implies the disk in the millimeter is mutually inclined from the planetary orbital plane by 16^{+22}_{-11} degrees. If this offset is real, we would be observing a process that decouples the millimeter planetesimals from the planets and smaller dust grains probed by *Herschel*. However, it is not clear what could cause that. As reported in Section 5.4.1, we do not find high mass solutions that are mutually inclined by over 8°, so it seems unlikely that planet b is torquing the disk, although we cannot definitely exclude planet b being inclined from the rest of the planets. Still, it would not explain why the smaller dust seen by *Herschel* are indeed coplanar with the planets. Deeper observations of the debris disk are needed to determine if the debris disk is coplanar to the planets.

Previous works have considered the need for an additional planet to carve the inner edge of the outer belt of debris (Booth et al. 2016; Read et al. 2018). Here we investigated whether

---

[1] Note, however, that damping of planetary eccentricity by gas dynamical friction does not conserve the planet's orbital angular momentum, contrary to damping of eccentricity by tidal dissipation in the planet; the latter, not the former, is considered by Lithwick & Wu (2012) and Batygin & Morbidelli (2013). However, the general mechanism of resonant repulsion also occurs for eccentricity damping by gas dynamical friction (e.g., as simulated for giant planets at wide sparations by Dong & Dawson 2016).

planet b is consistent with sculpting the inner edge of the outer belt, assuming the planets are coplanar or nearly so to the disk; the possible millimeter-wave offset of 16^{+22}_{-11} deg cited above is assumed negligible in this regard. Since planet b's orbit is likely near-circular, with 95% of the allowed stable orbits having $e_b < 0.05$, we can compute the clearing zone of planet b using the following equation from Morrison & Malhotra (2015) for the outer edge of a planet's chaotic zone, validated for high-mass planets like HR 8799 b:

$$R_{in} = a_p + 1.7a_p(M_p/M_\star)^{0.31}. \tag{5.2}$$

Here R_{in} is the inner radius of the disk and corresponds to the outer edge of a planet's clearing zone, a_p is the semi-major axis of the planet, M_p is the mass of the planet, and $M_\star$ is the mass of the star. Plugging in the numbers from our dynamically stable orbits, we find $R_{in} = 89^{+3}_{-2}$ au when considering all stable orbits, and $R_{in} = 93^{+3}_{-2}$ au when considering only stable orbits with $M_b > 5\ M_{\mathrm{Jup}}$. When compared to the inner edge of 104^{+8}_{-12} au derived by Wilner et al. (2018), our median value of the inner edge when only considering $M_b > 5\ M_{\mathrm{Jup}}$ is consistent with their middle 68% credible interval, while our median value when considering all of our stable solutions equally is slightly below this credible interval. However, given that our quoted numbers on the inner disk edge depends on our priors on the mass of the planet and given the uncertainty in the inner disk edge (Booth et al. 2016; Wilner et al. 2018), a 1σ disagreement is not significant. Our orbit fits place planet b at a location consistent with sculpting the inner edge of the debris disk, although finer studies of the dynamical interactions of system and more refined system parameters will help clarify the picture.

5.4.3 Orbital Resonances

Having stable orbits near integer period ratios does not guarantee resonance. To explore possible orbital resonances in our stable configurations, we saved the state of each stable system every 200 years using the *SimulationArchive* feature of `REBOUND` (Rein & Tamayo 2017). We looked at the resonant angles of each system as a function of time to infer the resonant nature of the system: planets in resonance will have a corresponding resonant angle that is librating, but the angle will circulate if the planets are not in resonance. For this four planet system, we looked at nine resonant angles. The first six angles are two-body resonant angles that look at whether consecutive planets are in 2:1 resonance. These 2:1 resonant angles are defined by

$$\theta_{i:j,k} = \lambda_j - 2\lambda_i + \varpi_k, \quad k \in \{i, j\}. \tag{5.3}$$

Here, $\varpi = \Omega + \omega$ is the longitude of periastron, and $\lambda = \varpi + M$ is the mean longitude where M is the mean anomaly. The labels i and j refer to the labels of a consecutive pair of planets with the inner planet being j, and k refers to either i or j, resulting in two resonant angles per pair of planets and thus six two-body resonant angles in total for the three consecutive pairs of planets. For example, the 2:1 resonant angle for planets c and d using planet d's ϖ would be written as $\theta_{c:d,d}$. In a similar notation, the three-body 1:2:4 Laplace resonance can

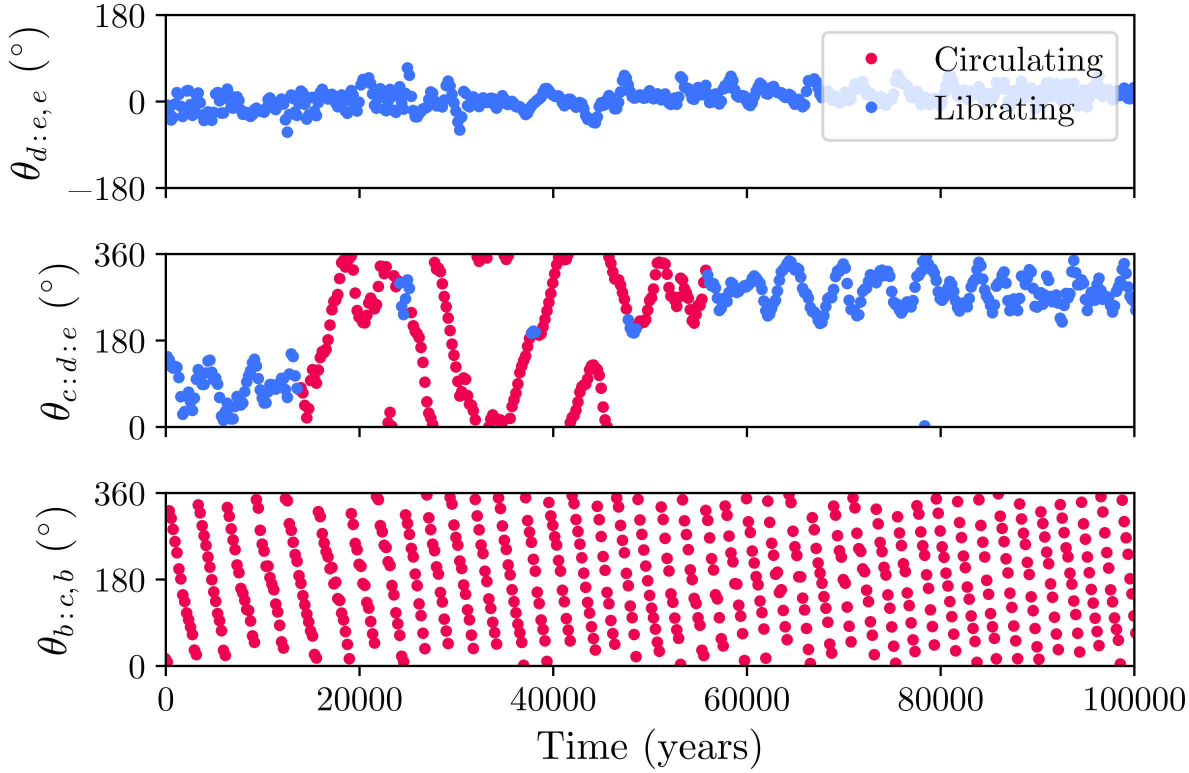


Figure 5.7: Demonstration of the algorithm to identify librating and circulating segments of each critical angle. The top, middle, and bottom plots show a purely librating, transitioning, and purely circulating angle respectively. The points colored red show regions identified as circulating, and the points colored blue show regions identified as librating. In this example, $\theta_{d:e,e}$ is always librating with a librating center of 0° and a libration amplitude of 36°. $\theta_{c:d:e}$ is librating only 33% of the time and $\theta_{b:c,b}$ is circulating 98% of the time so librating amplitudes and centers are not well defined.

be written as

$$\theta_{i:j:k} = \lambda_k - 3\lambda_j + 2\lambda_i, \tag{5.4}$$

where the innermost planet is k, and the outermost planet is i. Lastly, we looked at the same four-body 1:2:4:8 resonant angle as Goździewski & Migaszewski (2014):

$$\theta_{b:c:d:e} = \lambda_e - 2\lambda_d - \lambda_c + 2\lambda_b. \tag{5.5}$$

In our simulations, we found that these resonant angles varied in behavior, with some continuously librating (i.e., locked in resonance for 40 Myr), some continuously circulating, and some transitioning between the two over the 40 Myr orbit integration. To analyze all of the simulations uniformly, we developed an algorithm to identify libration and compute the fraction of time a resonant angle is librating or circulating over the course of a simulation.

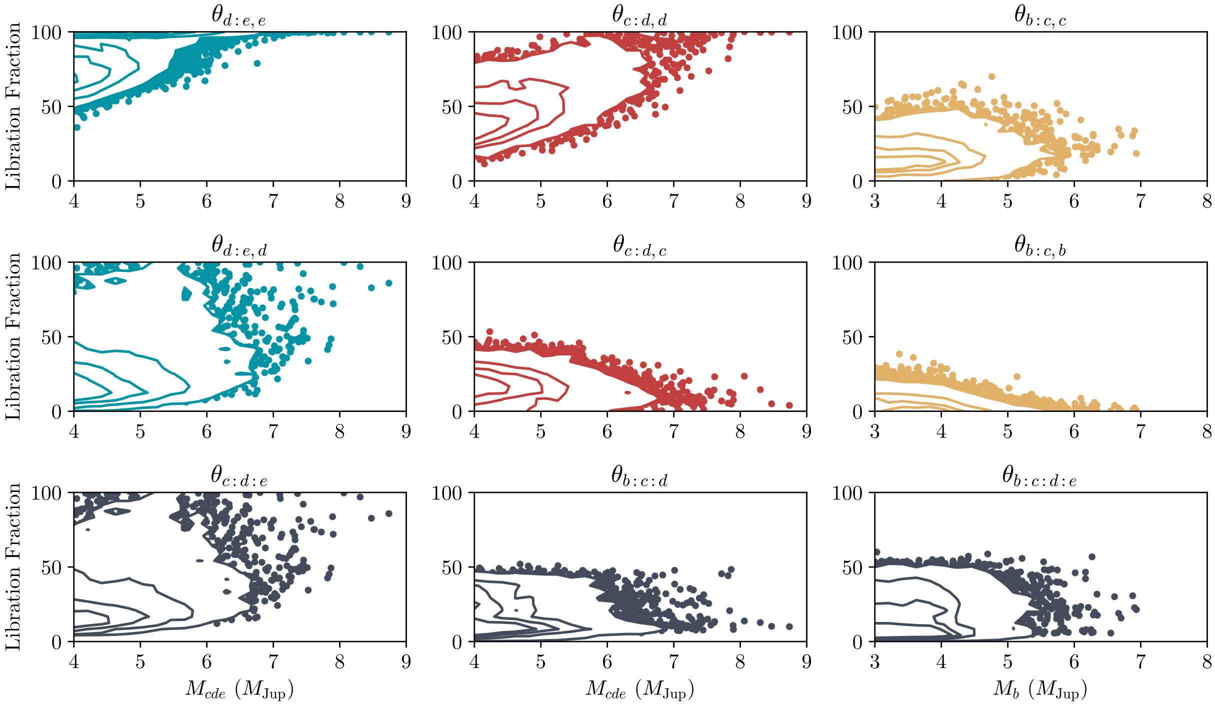


Figure 5.8: Distribution of systems in the space of libration fraction versus planet mass. Contours are 25th, 50th, 75th, and 97th percentiles (e.g., 97% of systems lie within the 97th percentile contour). Above that, individual points that correspond to particular stable orbital configurations are plotted. Libration fractions of 100% indicate resonance lock, while libration fractions $< 5\%$ indicate the planets are probably never in resonance.

The algorithm takes advantage of the fact that librations oscillate around a fixed value while circulating angles are monotonically changing. Briefly, the algorithm uses a Fourier transform to identify the periodicity of the data, smooths it on that scale, and computes the time derivative of the smoothed angle over the time series. Any sections of the time derivative with significant deviations from zero are deemed circulating and the rest are deemed librating. Figure 5.8 shows an example of this algorithm classifying librating and circulating sections of a few resonant angles. We note that this method is not perfect, and requires a subjective threshold to determine when a deviation is significant. However, inspecting the results from several resonant angles by eye, the algorithm seems comparable to by-eye identification, and seems to accurately identify resonant angles that are librating continuously (i.e., librating 100% of the time). For systems which by eye are transitioning quickly or which are always circulating, we estimate that we misidentified ∼5% of the time-series. This error is small, and the gain in having an automated algorithm to uniformly analyze all of these time-series is large.

We apply this algorithm to all nine resonant angles for each of our simulations. We plot the libration fraction, the fraction of time in the last 40 Myr during which that angle is librating, for each angle in Figure 5.8. At low masses ($M_{cde} < 6 M_{\mathrm{Jup}}$), the scatter in the libration fraction is high for all angles, indicating resonance lock for any subset of the planets is not necessary for stable orbits at low masses. Several of the angles never reach 100% libration fraction at any mass, indicating that all of the stable orbital configurations we found do not have all four planets in resonant lock. While we do not see any four planet resonant chains, it might be possible they reside in a small island of parameter space that our MCMC did not sample. All resonance angles involving planet b never reach 100% libration fraction. Planet b may occasionally come into resonance with the inner planets, but does not remain there. It is not too surprising that planet b does not favor resonance, insofar as the magnitude of the resonant potential associated with $\theta_{b:j,b}$ is proportional to the planet's orbital eccentricity e_b (e.g., Murray & Dermott 1999), and its median eccentricity is the lowest among the four planets.

The inner three planets do favor resonance more than planet b. Above ∼6 M_{Jup}, a large majority of stable orbits have $\theta_{d:e,e}$ librating 100% of the time. Similarly, above ∼7 M_{Jup}, $\theta_{c:d,d}$ is always librating for most stable orbits. For all masses, these two angles are always librating some of the time. However, the other two resonant angles, $\theta_{d:e,d}$ and $\theta_{c:d,c}$, appear primarily transitioning between libration and circulation, with $\theta_{d:e,d}$ trending to libration at high masses. This bheavior is also reflected in the three-body resonant angle between the inner three planets, with only 1.2% of the stable orbits having this three-body angle librating for at least 90% of the time. Still, this behavior indicates that the inner planets being in a 1:2:4 three-body resonance is both consistent with the data and dynamically stable for masses up to 8 M_{Jup}. In these cases, the libration center of $\theta_{c:d:e}$ often jumps between $\sim 90^\circ$ and $\sim -90^\circ$, but typically keeping $|\varpi_d - \varpi_e| \sim 90^\circ$. However, the 1:2:4 three-body resonance is not required for stability, even at high masses.

When the three planets are not locked in resonance together, pairs of planets can be in resonance. We find, averaged across the ensemble of simulations, these two-body resonant angles librate around 0°. However, in a single simulation, the libration center can be offset from 0°, a phenomenon know as asymmetric libration (e.g., Murray-Clay & Chiang 2005) that is observed for two body angles in other resonant chain systems (e.g., Kepler-80, MacDonald et al. 2016) and is caused by the gravitational effect of a third planet. For the situation where $\theta_{d:e,e}$ librates but $\theta_{d:e,d}$ does not, conjunction of planets d and e always occurs at the periastron of planet e's orbit, but is completely uncorrelated planet d's orbit. In the case where both $\theta_{d:e,e}$ and $\theta_{c:d,d}$ librate but the three-body angle $\theta_{c:d:e}$ does not, planet e's orbit orients itself so that it is lined up to the conjunction with planet d, while planet d's longitude of periastron is driven by the conjunction with planet c, which is not locked in with planet e. In this case, consecutive planet pairs appear to be locked in resonance for 40 Myr, but three-body resonance lock does not exist.

Having planets d and e and planets c and d locked in two-body resonances fits well with the picture that they were locked in resonance quickly after formation, before the gaseous protoplanetary disk disappeared. After the planets migrated into resonance lock, the eccentricities of e and d were both amplified by their resonant migration in the gas disk and damped by the gas, pushing the planets to larger period ratios. After the gas dispersed, the planets maintained their primordial eccentricities and period ratios, with these parameters only oscillating as the planets exchange energy and angular momentum in resonance.

5.4.4 Dynamical Mass Limits

As we only have a short orbital arc of data, we are limited on the mass constraints we can place based solely on dynamical considerations. We have not yet measured the perturbations of the planets' orbits by each other. Without seeing a significant effect, we cannot place a lower bound on the masses of the planets dynamically. Impractically, we may need to measure the change in the orbital elements of the planets after many orbits, akin to the masses derived from transit timing variations (Agol et al. 2005; Holman & Murray 2005). Thus, dynamical constraints based on short orbital arcs cannot fully constrain the masses alone.

Looking at Figure 5.5 again, our stable orbits heavily favor low masses, since we cannot place a lower bound on the masses and we used a uniform prior. For high masses, $M_{cde} > 8\ M_{\rm Jup}$, we have only a few stable orbits, all with the mass of the planet b near the lower-bound of what we would expect, indicating a possible upper limit to the masses of the planets. However, we cannot verify this is because of a lack of sampling of stable orbits. That is, the probability of drawing a stable orbital configuration with $M_{cde} > 8\ M_{\rm Jup}$ and $M_b > 5\ M_{\rm Jup}$ might be so small that we do not expect to find one with our current sampling. We are potentially limited by the fact that 2.5 million samples of the posterior are not sufficient when the posterior has 20 dimensions and the islands of stability at extremely high masses are extremely small. Thus, there is no indication of a sharp drop-off that would point to a firm upper limit on the mass based on dynamical considerations. In fact, Goździewski & Migaszewski (2014) and Gozdziewski & Migaszewski (2018) found stable orbits above 9 $M_{\rm Jup}$ for the inner three planets, indicating higher-mass stable orbits exist if one forces the entire system to be in a 4-planet resonance lock. It might be that we missed those systems since we did not enforce such a global lock and therefore had a larger parameter space of allowed stable orbits. If we acknowledge that we are dependent on our choice of priors for the mass of the planets, we can say that 99.9% of the orbits that are dynamically stable for the last 40 Myr have $M_{cde} < 7.6\ M_{\rm Jup}$ and $M_b < 6.3\ M_{\rm Jup}$ for this particular choice of prior. These mass upper limits are consistent with the luminosity derived masses of $M_{cde} = 7\ M_{\rm Jup}$ and $M_b = 5\ M_{\rm Jup}$ based on hot-start evolutionary models (Marois et al. 2008, 2010b).

We also investigated using a mass prior based on the measured luminosity and hot-start evolutionary models. Since it is computationally intensive to rerun all 22 million N-body simulations, we instead weighted each sample drawn from the previous prior distribution with the relative change in probability due to switching to a luminosity-based prior, resulting in a down-weighting of the lowest mass configurations which are disfavored by the measured luminosities. To accomplish this, we reevaluated the luminosity-derived masses with newest age estimates for the Columba moving group from Bell et al. (2015) which we approximate as a Gaussian distribution of 42 ± 5 Myr. This age has better stated uncertainties than

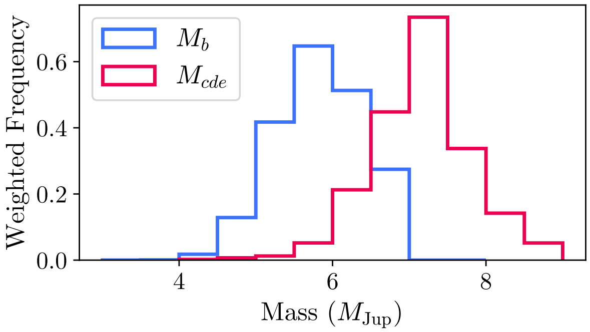


Figure 5.9: The histogram of dynamically stable masses after being weighted by the luminosity prior on the planet masses assuming hot-start evolutionary tracks. Planet b is in blue, and planets c, d, and e are assumed to follow the same red histogram.

the 30-60 Myr range given in Marois et al. (2010b). We still used the same luminosities measured by Marois et al. (2008) and Marois et al. (2010b), as there has not been an update to them with stated uncertainties. Using the Baraffe et al. (2003) hot-start cooling tracks, we get model-dependent masses of 6.0 ± 0.7, 8.7 ± 1.0, 8.7 ± 1.0, and 8.7 ± 1.7 $M_{\rm Jup}$ for planets b, c, d, and e respectively. To stay self-consistent with our simulations that fix the masses of the inner three planets to be the same, we choose to use the 8.7 ± 1.0 $M_{\rm Jup}$ of planets c and d for the mass of the planet e also. Otherwise, planet e will tend toward lower masses, and be inconsistent with our simulation assumption of equal masses for the inner three planets. Indeed, spectrophotometric measurements from the latest generation of high-contrast imagers confirm that planet e has similar near-infrared fluxes to planet c and d, and not planet b (Zurlo et al. 2016; Greenbaum et al. 2018), so our assumption should be robust. Altogether, when combining the luminosity and dynamical constraints for the mass assuming hot-start evolutionary tracks, we get a mass of 5.8 ± 0.5 $M_{\rm Jup}$ for planet b and $7.2^{+0.6}_{-0.7}$ $M_{\rm Jup}$ for planets c, d, and e. These mass distributions are plotted in Figure 5.9. Note that our mass estimates depend on the resonances we found the planets to be in, and that higher masses can be achieved by assuming four planet resonance lock (Goździewski & Migaszewski 2014; Gozdziewski & Migaszewski 2018).

It is uncertain exactly how bright planets are during the first 100 Myr as it depends on uncertain formation mechanisms. Planet cooling tracks are instead parameterized by a quantity like the initial entropy of the material that formed the planet (Spiegel & Burrows 2012). Since the hot-start models really are the high-entropy upper limit with regards to planet formation models, the masses estimated assuming these tracks are the lowest masses for the planets. Thus, we can use the hot-start model to quote a lower limit on the mass. Combining the dynamical constraints with the luminosity prior on the masses from the hot-start model, we find 95% of the stable systems have $M_b > 4.9$ $M_{\rm Jup}$ and $M_{cde} > 6.1$ $M_{\rm Jup}$. Alternatively, we can use the upper limits on the masses from dynamical stability alone to constrain the initial conditions of the cooling tracks. Using the same stated luminosities and

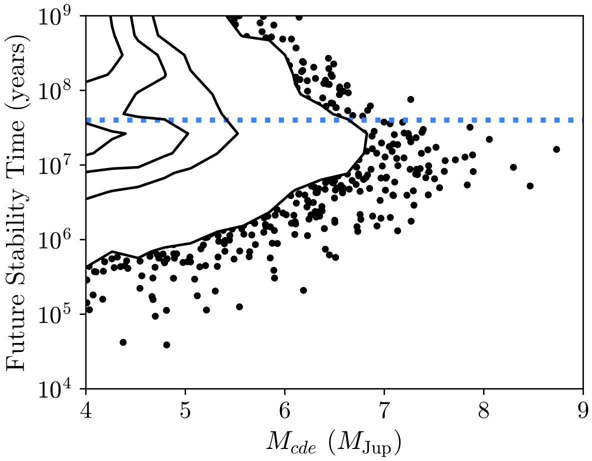


Figure 5.10: The stability time of the dynamically stable orbits integrating forwards in time for up to 1 Gyr. Contours are 25th, 50th, 75th, and 97th percentiles. The horizontal blue dotted line indicates 40 Myr, the time we integrated the systems backwards as reference.

age of the system and using the Spiegel & Burrows (2012) warm-start models, our upper limits of $M_{cde} < 7.6$ $M_{\rm Jup}$ and $M_b < 6.3$ $M_{\rm Jup}$ would then correspond to a lower limit on the initial entropy of 9 k_B per baryon.

5.4.5 Long-Term Dynamical Stability

Even though the system has been stable for ∼40 Myr, we investigate whether the system we see today is reflective of the final state of system, or whether this configuration is a transient one. We used the saved *SimulationArchive* of each stable orbit, reversed the velocities again, and now integrated them forwards in time for up to 1 Gyr. We used the same criteria to assess stability as before. We plot the amount of time in the future each system is stable as a function of the mass of the inner three planets in Figure 5.10.

We find orbits that are stable for 1 Gyr up to 6 $M_{\rm Jup}$. Above 7 $M_{\rm Jup}$, none of our orbits are stable for over 100 Myr. Thus, if these planets have masses above 7 $M_{\rm Jup}$ as is favored by our combined dynamical and luminosity constraint, then the system is not likely stable. As there is evidence in our own Solar System of dynamical upheaval of our less-tightly-packed gas giants early on (Tsiganis et al. 2005), it is not surprising to find that the HR 8799 system will become unstable.

Certain resonances do seem to improve stability. In particular, in systems where the three-body angle $\theta_{c:d:e}$ is librating at least 50% of the time, systems are 10 times less likely to go unstable in 1 Myr and 4 times less likely to become unstable in the 10 Myr than systems with $\theta_{c:d:e}$ librating less than 50% of the time. There does not appear to be strong correlation between the two-body angles and stability except for $\theta_{b:c,c}$. Systems for which $\theta_{b:c,c}$ is libration more than 30% of the time are 5 and 15 times more likely to be stable for at least 1 Myr and 10 Myr respectively than systems where this angle librates less than 30% of the time. While certain resonances seem to prevent systems from short-term instabilities, there is no indication that spending more time in resonance or achieving resonance lock improves stability at the 1 Gyr level. This perhaps is due to the fact we did not find systems with the four planets locked in resonance, which could have improved longer term stability.

5.5 Conclusion

This paper has aimed to explore the dynamically stable orbits of the HR 8799 system. In the first part of the paper, we continued to demonstrate the precise astrometry that can be achieved by GPI and explored various assumptions on the orbits of the planets.

- Using GPI IFS data from 2014-2016, we measured the astrometry of the HR 8799 with one milliarcsecond precision using the open-source `pyKLIP` package.

- We utilized MCMC methods to fully explore the 20+ dimensional space of orbital configurations using Bayesian parameter estimation and different assumptions on the coplanarity and resonant nature of the system.

- We found that assuming the system is coplanar or the system is near the 1:2:4:8 period ratio resonance does not significantly worsen the fit to the data, and in fact might make it better. We find including both assumptions provide adequate fits to the data, agreeing well with the conclusions of Konopacky et al. (2016a).

In the second half of this paper, we have presented the first attempt to rigorously fold in dynamical constraints into orbit fits of a directly imaged system, and demonstrated the power of including a dynamical prior.

- We performed rejection sampling on our posteriors of orbit fits to apply our prior of dynamical stability. Using the `REBOUND` N-body integrator, we ran orbits backwards in time for 40 Myr, the age of the system, varying the masses of the planets and looked for the stable orbital configurations.

- We find that coplanar orbits near the 1:2:4:8 resonance produces orders of magnitude more stable orbits than any other scenario. We find a few orbits near 1:2:4:8 resonance with some mutual inclinations that are stable, but the inefficiency of finding them makes studying that family of orbits impractical with current astrometric data.

- As demonstrated by Figure 5.4, the stable coplanar orbits lie within a small fraction of the allowed orbital space. In this subspace, we find the outer two planets have near zero eccentricity, while the inner two planets have $e \sim 0.1$.

- Our orbits are consistent with being coplanar with the *Herschel*-derived debris disk plane, but misaligned with the plane derived from SMA and ALMA by 16^{+22}_{-11} degrees.

Our fitted orbit for planet b is consistent with that planet sculpting the inner edge of the debris disk in the millimeter, assuming the orbits are close enough to coplanar.

- If $M_{cde} \gtrsim 6\ M_{\rm Jup}$, planet e needs to be locked in resonance with planet d in order for the system to be stable. Likewise, if $M_{cde} \gtrsim 7\ M_{\rm Jup}$, planet d is likely in resonance lock with planet c. Although we find stable configurations where the inner three planets are in a 1:2:4 Laplace resonance, such a 3-body resonance is not required, as we found many stable configurations where only pairs of planets are in resonance. Planet b does not need to be in resonance for this system to be stable so far.
- Using a uniform prior on M_{cde} and a slightly low-mass-favored prior on M_b, we find 99.9% of our stable orbits have $M_b < 6.3\ M_{\rm Jup}$ and $M_{cde} < 7.6\ M_{\rm Jup}$. Folding in the mass constraints from the planet luminosities and hot-start evolutionary models, we find $M_b = 5.8 \pm 0.5\ M_{\rm Jup}$ and $M_{cde} = 7.2^{+0.6}_{-0.7}\ M_{\rm Jup}$. Either way, our mass constraints are consistent with hot-start evolutionary tracks.
- We do not find systems with the inner planets above 7 $M_{\rm Jup}$ that are stable for the next 1 Gyr.

In the future, as more of the orbital arcs are traced out with precise astrometry, it will become clearer where in the 20+ dimensional space the planets' true orbital configuration lie. In the meantime using more computational power, we can attempt search for stable orbits that are not forced to be exactly coplanar, are stable at higher masses, or are stable at other resonances. Additionally, the new parallax value from Gaia Data Release 2 (Gaia Collaboration et al. 2018) should help tighten the constraints on the semi-major axes of the planets and total mass of the system for the stable coplanar solutions. Even a decade since its discovery, the HR 8799 planetary system is one of the most unique and interesting systems that we know, and combining both detailed dynamical studies with atmospheric characterization will help us understand how these planets formed and how they will interact.

Our technique of performing rejection sampling to apply dynamical constraints after MCMC sampling of the orbital parameters can also be applied to other directly-imaged multi-planet systems to better constrain the orbits with just a short orbital arc. Here we have shown that even with < 15% of the full orbit covered, we can constrain orbital parameters to a few percent. This can remove the orbital uncertainty that comes with exoplanets discovered through imaging alone, where we typically need to wait for long-period planets to complete a orbital revolution before fully-constraining its orbit. This is also potentially valuable for future space-based imaging missions that search for exo-Earths in multi-planet systems, since this method can estimate the mass and orbit with a small orbital arc. This can allow for the mission to better prioritize which exo-Earth candidates are observed with expensive spectroscopic observations by determining which are most likely Earth-mass and orbiting at a favorable distance from the star.

5.6 Acknowledgements

Simulations in this paper made use of the REBOUND code which can be downloaded freely at http://github.com/hannorein/rebound. We thank Dan Tamayo for offering REBOUND tutorials and help with setting up the simulations. We also thank Eve Lee for helpful discussions on dynamics. J.J.W., J.R.G., and P.K. were supported in part by NSF AST-1518332, NASA NNX15AC89G and NNX15AD95G. This work benefited from NASA's Nexus for Exoplanet System Science (NExSS) research coordination network sponsored by NASA's Science Mission Directorate. R.I.D. gratefully acknowledges support from NASA XRP 80NSSC18K0355. The Center for Exoplanets and Habitable Worlds is supported by the Pennsylvania State University, the Eberly College of Science, and the Pennsylvania Space Grant Consortium. Portions of this work were performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

The GPI project has been supported by Gemini Observatory, which is operated by AURA, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the NSF (USA), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), MCTI (Brazil) and MINCYT (Argentina). The Gemini Observatory is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), Ministério da Ciência, Tecnologia e Inovação (Brazil), and Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina). This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France.

Software: Gemini Planet Imager Data Reduction Pipeline (Perrin et al. 2014, 2016), pyKLIP (Wang et al. 2015), emcee (Foreman-Mackey et al. 2013), REBOUND (Rein & Liu 2012), Astropy (Astropy Collaboration et al. 2013b), matplotlib (Hunter 2007)

Facility: Gemini:South (GPI)

Bibliography

Agol, E., Steffen, J., Sari, R., & Clarkson, W. 2005, MNRAS, 359, 567

Alzner, A., & Argyle, R. W. 2012, The Orbital Elements of a Visual Binary Star, 71

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013a, A&A, 558, A33

—. 2013b, A&A, 558, A33

Bailey, V. P., Poyneer, L. A., Macintosh, B. A., et al. 2016, in Proc. SPIE, Vol. 9909, Adaptive Optics Systems V, 99090V

Baines, E. K., White, R. J., Huber, D., et al. 2012, ApJ, 761, 57

Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, A&A, 402, 701

Batygin, K., & Brown, M. E. 2016, AJ, 151, 22

Batygin, K., & Morbidelli, A. 2013, AJ, 145, 1

Beichman, C. A., Krist, J., Trauger, J. T., et al. 2010, PASP, 122, 162

Bell, C. P. M., Mamajek, E. E., & Naylor, T. 2015, MNRAS, 454, 593

Beust, H., & Morbidelli, A. 2000, Icarus, 143, 170

Binks, A. S., & Jeffries, R. D. 2016, MNRAS, 455, 3345

Birney, D., Gonzalez, G., & Oesper, D. 2006, Observational Astronomy (Cambridge University Press)

Booth, M., Jordán, A., Casassus, S., et al. 2016, MNRAS, 460, L10

Bottom, M., Kuhn, J., Mennesson, B., et al. 2015, ApJ, 809, 11

Bowler, B. P. 2016, PASP, 128, 102001

Brandt, T. D., McElwain, M. W., Turner, E. L., et al. 2013, ApJ, 764, 183

Burrows, C. J., Krist, J. E., Stapelfeldt, K. R., & WFPC2 Investigation Definition Team. 1995, in American Astronomical Society Meeting Abstracts, Vol. 187, 32.05

Cantalloube, F., Mouillet, D., Mugnier, L. M., et al. 2015, A&A, 582, A89

Chauvin, G., Lagrange, A. M., Beust, H., et al. 2012, A&A, 542, A41

Chilcote, J., Barman, T., Fitzgerald, M. P., et al. 2015, ApJ, 798, L3

Chilcote, J. K., Larkin, J. E., Maire, J., et al. 2012, in Proc. SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84468W

Crifo, F., Vidal-Madjar, A., Lallement, R., Ferlet, R., & Gerbaldi, M. 1997, A&A, 320, L29

Crill, B. P., & Siegler, N. 2017, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10398, 103980H

Currie, T., Fukagawa, M., Thalmann, C., Matsumura, S., & Plavchan, P. 2012, ApJ, 755, L34

Czekala, I., Andrews, S. M., Mandel, K. S., Hogg, D. W., & Green, G. M. 2015, ApJ, 812, 128

Dawson, R. I., Murray-Clay, R. A., & Fabrycky, D. C. 2011, ApJ, 743, L17

de Mooij, E. J. W., Watson, C. A., & Kenworthy, M. A. 2017, MNRAS, 472, 2713

De Rosa, R. J., Nielsen, E. L., Blunt, S. C., et al. 2015, ApJ, 814, L3

Deck, K. M., Holman, M. J., Agol, E., et al. 2012, ApJ, 755, L21

Defrère, D., Lebreton, J., Le Bouquin, J. B., et al. 2012, A&A, 546, L9

Dong, R., & Dawson, R. 2016, ApJ, 825, 77

Esposito, S., Mesa, D., Skemer, A., et al. 2013, A&A, 549, A52

Esposito, T. M., Fitzgerald, M. P., Graham, J. R., & Kalas, P. 2014, ApJ, 780, 25

Fabrycky, D. C., & Murray-Clay, R. A. 2010, ApJ, 710, 1408

Fitzgerald, M. P., & Graham, J. R. 2006, ApJ, 637, 541

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publications of the Astronomical Society of the Pacific, 125, 306

Foreman-Mackey, D., Vousden, W., Price-Whelan, A., et al. 2016, corner.py: corner.py v1.0.2, Zenodo Software Release

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, ArXiv e-prints, arXiv:1804.09365

—. 2016, A&A, 595, A2

Galicher, R., & Marois, C. 2011, in Second International Conference on Adaptive Optics for Extremely Large Telescopes. Online at http://ao4elt2.lesia.obspm.fr, P25

Galle, J. G. 1846, MNRAS, 7, 153

Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Götberg, Y., Davies, M. B., Mustill, A. J., Johansen, A., & Church, R. P. 2016, A&A, 592, A147

Goździewski, K., & Migaszewski, C. 2014, MNRAS, 440, 3140

Gozdziewski, K., & Migaszewski, C. 2018, ArXiv e-prints, arXiv:1807.05898 [astro-ph.EP]

Graham, J. R. 2009, Gemini Planet Imager Operational Concept Definition Document V4.1

Graham, J. R., Macintosh, B., Doyon, R., et al. 2007, ArXiv e-prints, arXiv:0704.1454

Green, R. 1985, Spherical Astronomy (Cambridge University Press)

Greenbaum, A. Z., Pueyo, L., Ruffio, J.-B., et al. 2018, AJ, 155, 226

Hamilton, D. P., & Burns, J. A. 1992, Icarus, 96, 43

Heap, S. R., Lindler, D. J., Lanz, T. M., et al. 2000, ApJ, 539, 435

Herschel, W., & Watson, D. 1781, Philosophical Transactions of the Royal Society of London Series I, 71, 492

Hibon, P., Thomas, S., Dunn, J., et al. 2014, in Ground-based and Airborne Instrumentation for Astronomy V, Vol. 9147, 91474U

Holman, M. J., & Murray, N. W. 2005, Science, 307, 1288

Hughes, A. M., Wilner, D. J., Andrews, S. M., et al. 2011, ApJ, 740, 38

Hung, L.-W., Bruzzone, S., Millar-Blanchaer, M. A., et al. 2016, in Proc. SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 99083A

Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90

Ingraham, P., Perrin, M. D., Sadakuni, N., et al. 2014, in Ground-based and Airborne Instrumentation for Astronomy V, Vol. 9147, 91477O

Kalas, P., Graham, J. R., Fitzgerald, M. P., & Clampin, M. 2013, ApJ, 775, 56

Kalas, P., Graham, J. R., Chiang, E., et al. 2008, Science, 322, 1345

Kenworthy, M. A., & Mamajek, E. E. 2015, ApJ, 800, 126

Koen, C., Balona, L. A., Khadaroo, K., et al. 2003, MNRAS, 344, 1250

Konopacky, Q. M., Barman, T. S., Macintosh, B. A., & Marois, C. 2013, Science, 339, 1398

Konopacky, Q. M., Marois, C., Macintosh, B. A., et al. 2016a, AJ, 152, 28

Konopacky, Q. M., Thomas, S. J., Macintosh, B. A., et al. 2014, in Ground-based and Airborne Instrumentation for Astronomy V, Vol. 9147, 914784

Konopacky, Q. M., Rameau, J., Duchêne, G., et al. 2016b, ApJ, 829, L4

Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, ApJ, 660, 770

Lagrange, A. M., Gratadour, D., Chauvin, G., et al. 2009, A&A, 493, L21

Lagrange, A. M., Bonnefoy, M., Chauvin, G., et al. 2010, Science, 329, 57

Lagrange, A. M., Boccaletti, A., Milli, J., et al. 2012, A&A, 542, A40

Larkin, J. E., Chilcote, J. K., Aliado, T., et al. 2014, in Proc. SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 91471K

Lavie, B., Mendonça, J. M., Mordasini, C., et al. 2017, AJ, 154, 91

Lecavelier des Etangs, A., & Vidal-Madjar, A. 2016, A&A, 588, A60

Lecavelier Des Etangs, A., Vidal-Madjar, A., Burki, G., et al. 1997, A&A, 328, 311

Leonard, F. C. 1930, Leaflet of the Astronomical Society of the Pacific, 1, 121

Liddle, A. R. 2007, MNRAS, 377, L74

Lithwick, Y., & Wu, Y. 2012, ApJ, 756, L11

MacDonald, M. G., Ragozzine, D., Fabrycky, D. C., et al. 2016, AJ, 152, 105

Macintosh, B., Graham, J. R., Ingraham, P., et al. 2014, Proceedings of the National Academy of Science, 111, 12661

Maire, A.-L., Skemer, A. J., Hinz, P. M., et al. 2015, A&A, 576, A133

Maire, J., Ingraham, P. J., De Rosa, R. J., et al. 2014, in Ground-based and Airborne Instrumentation for Astronomy V, Vol. 9147, 914785

Mamajek, E. E., & Bell, C. P. M. 2014, MNRAS, 445, 2169

Mamajek, E. E., Quillen, A. C., Pecaut, M. J., et al. 2012, AJ, 143, 72

Marcy, G. W., Butler, R. P., Fischer, D., et al. 2001, ApJ, 556, 296

Marley, M. S., Fortney, J. J., Hubickyj, O., Bodenheimer, P., & Lissauer, J. J. 2007, ApJ, 655, 541

Marois, C., Correia, C., Galicher, R., et al. 2014, in Proc. SPIE, Vol. 9148, Adaptive Optics Systems IV, 91480U

Marois, C., Doyon, R., Racine, R., & Nadeau, D. 2000, Publications of the Astronomical Society of the Pacific, 112, 91

Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006a, ApJ, 641, 556

Marois, C., Lafrenière, D., Macintosh, B., & Doyon, R. 2006b, ApJ, 647, 612

Marois, C., Macintosh, B., Barman, T., et al. 2008, Science, 322, 1348

Marois, C., Macintosh, B., & Véran, J.-P. 2010a, in Adaptive Optics Systems II, Vol. 7736, 77361J

Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., & Barman, T. 2010b, Nature, 468, 1080

Matthews, B., Kennedy, G., Sibthorpe, B., et al. 2014, ApJ, 780, 97

Mawet, D., Milli, J., Wahhaj, Z., et al. 2014, ApJ, 792, 97

McBride, J., Graham, J. R., Macintosh, B., et al. 2011, Publications of the Astronomical Society of the Pacific, 123, 692

Mékarnia, D., Chapellier, E., Guillot, T., et al. 2017, A&A, 608, L6

Millar-Blanchaer, M. A., Esposito, T. M., Fitzgerald, M. P., et al. 2017, in Proc. SPIE, Vol. 10407, Polarization Science and Remote Sensing VIII

Millar-Blanchaer, M. A., Graham, J. R., Pueyo, L., et al. 2015, ApJ, 811, 18

Millar-Blanchaer, M. A., Perrin, M. D., Hung, L.-W., et al. 2016, in Ground-based and Airborne Instrumentation for Astronomy VI, Vol. 9908, 990836

Morrison, S., & Malhotra, R. 2015, ApJ, 799, 41

Morrison, S. J., & Kratter, K. M. 2016, ApJ, 823, 118

Morzinski, K. M., Males, J. R., Skemer, A. J., et al. 2015, ApJ, 815, 108

Mouillet, D., Larwood, J. D., Papaloizou, J. C. B., & Lagrange, A. M. 1997, MNRAS, 292, 896

Murray, C. D., & Dermott, S. F. 1999, Solar system dynamics

Murray-Clay, R. A., & Chiang, E. I. 2005, ApJ, 619, 623

Nielsen, E. L., Liu, M. C., Wahhaj, Z., et al. 2014, ApJ, 794, 158

Nielsen, E. L., De Rosa, R. J., Rameau, J., et al. 2017, ArXiv e-prints, arXiv:1705.06851 [astro-ph.EP]

Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, ApJ, 743, L16

Perrin, M. D., Maire, J., Ingraham, P., et al. 2014, in Ground-based and Airborne Instrumentation for Astronomy V, Vol. 9147, 91473J

Perrin, M. D., Duchene, G., Millar-Blanchaer, M., et al. 2015, ApJ, 799, 182

Perrin, M. D., Ingraham, P., Follette, K. B., et al. 2016, in Proc. SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 990837

Perryman, M., Hartman, J., Bakos, G. Á., & Lindegren, L. 2014, ApJ, 797, 14

Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, Icarus, 124, 62

Poyneer, L. A., Palmer, D. W., Macintosh, B., et al. 2016, Appl. Opt., 55, 323

Pueyo, L. 2016, ApJ, 824, 117

Pueyo, L., Soummer, R., Hoffmann, J., et al. 2015, ApJ, 803, 31

Quarles, B., Quintana, E. V., Lopez, E., Schlieder, J. E., & Barclay, T. 2017, ApJ, 842, L5

Rameau, J., Nielsen, E. L., De Rosa, R. J., et al. 2016, ApJ, 822, L29

Read, M. J., Wyatt, M. C., Marino, S., & Kennedy, G. M. 2018, MNRAS, 475, 4953

Rein, H., & Liu, S.-F. 2012, A&A, 537, A128

Rein, H., & Tamayo, D. 2015, MNRAS, 452, 376

—. 2017, MNRAS, 467, 2377

Rivera, E. J., Laughlin, G., Butler, R. P., et al. 2010, ApJ, 719, 890

Ruffio, J.-B., Macintosh, B., Wang, J. J., et al. 2017, ApJ, 842, 14

Savransky, D., Macintosh, B. A., Graham, J., & Konopacky, Q. M. 2013, in Proceedings of the International Astronomical Union, Vol. 8 (Cambridge Univ Press), 68

Savransky, D., Thomas, S. J., Poyneer, L. A., & Macintosh, B. A. 2013, Applied Optics, 52, 3394

Schmid, H. M., Joos, F., & Tschan, D. 2006, A&A, 452, 657

Schwarz, G. 1978, Annals of Statistics, 6, 461

Shen, Y., & Tremaine, S. 2008, AJ, 136, 2453

Sivaramakrishnan, A., & Oppenheimer, B. R. 2006, ApJ, 647, 620

Smith, B. A., & Terrile, R. J. 1984, Science, 226, 1421

Snellen, I. A. G., Brandl, B. R., de Kok, R. J., et al. 2014, Nature, 509, 63

Soummer, R., Hagan, J. B., Pueyo, L., et al. 2011a, ApJ, 741, 55

Soummer, R., Pueyo, L., & Larkin, J. 2012, ApJ, 755, L28

Soummer, R., Sivaramakrishnan, A., Pueyo, L., Macintosh, B., & Oppenheimer, B. R. 2011b, ApJ, 729, 144

Sozzetti, A., Bonavita, M., Desidera, S., Gratton, R., & Lattanzi, M. G. 2016, in IAU Symposium, Vol. 314, Young Stars & Planets Near the Sun, ed. J. H. Kastner, B. Stelzer, & S. A. Metchev, 264

Spiegel, D. S., & Burrows, A. 2012, ApJ, 745, 174

Stuik, R., Bailey, J. I., Dorval, P., et al. 2017, A&A, 607, A45

Su, K. Y. L., Rieke, G. H., Stapelfeldt, K. R., et al. 2009, ApJ, 705, 314

Sudol, J. J., & Haghighipour, N. 2012, ApJ, 755, 38

Tamayo, D., Rein, H., Petrovich, C., & Murray, N. 2017, ApJ, 840, L19

Thébault, P., & Beust, H. 2001, A&A, 376, 621

Towns, J., Cockerill, T., Dahan, M., et al. 2014, Computing in Science & Engineering, 16, 62

Tsiganis, K., Gomes, R., Morbidelli, A., & Levison, H. F. 2005, Nature, 435, 459

van Leeuwen, F. 2007, A&A, 474, 653

Wang, J. J., Ruffio, J.-B., De Rosa, R. J., et al. 2015, pyKLIP: PSF Subtraction for Exoplanets and Disks, Astrophysics Source Code Library, ascl:1506.001

Wang, J. J., Rajan, A., Graham, J. R., et al. 2014, in Ground-based and Airborne Instrumentation for Astronomy V, Vol. 9147, 914755

Wang, J. J., Graham, J. R., Pueyo, L., et al. 2016, AJ, 152, 97

Wang, J. J., Perrin, M. D., Savransky, D., et al. 2018, Journal of Astronomical Telescopes, Instruments, and Systems, 4, 018002

Wertz, O., Absil, O., Gómez González, C. A., et al. 2017, A&A, 598, A83

Williams, J. P., & Cieza, L. A. 2011, Annual Review of Astronomy and Astrophysics, 49, 67

Wilner, D. J., MacGregor, M. A., Andrews, S. M., et al. 2018, ApJ, 855, 56

Wolff, S. G., Perrin, M. D., Maire, J., et al. 2014, in Ground-based and Airborne Instrumentation for Astronomy V, Vol. 9147, 91477H

Yu, Q., & Tremaine, S. 2001, AJ, 121, 1736

Zanazzi, J. J., & Lai, D. 2017, MNRAS, 464, 3945
Zuckerman, B., Rhee, J. H., Song, I., & Bessell, M. S. 2011, ApJ, 732, 61
Zurlo, A., Vigan, A., Galicher, R., et al. 2016, A&A, 587, A57

Appendix A

Appendix for Chapter 4

A.1 Analytic Forward Modeling of the Planet PSF after Stellar PSF Subtraction

Using the notation of Pueyo (2016), we can analytically forward model the PSF of the planet through the process of stellar PSF subtraction. Let us denote the target image as $T_{\lambda_p,t_p}(\mathbf{x})$, which is taken at wavelength λ_p and time t_p and contains the pixels $\mathbf{x}$ from which we want to subtract the stellar PSF.

To forward model the PSF of the planet, we need a model of the PSF of the planet prior to stellar PSF subtraction. For our GPI data, we use the satellite spots to generate a realization of an unocculted and unprocessed point source as a function of wavelength. Using this model, we can generate $A_{\lambda_p,t_p}(\mathbf{x})$, a frame that consists solely of the unprocessed PSF of the planet in the target frame. The initial estimate for the planet's position in the frame is determined using an input separation and PA of the planet and the location of the star. Similarly, we can also generate an array of frames, each containing the unprocessed PSF of the planet for each image in the reference library, accounting for the fact that the planet position changes due to ADI field rotation and due to SDI rescaling to align speckles. In these frames, the PSF shape only depends on the wavelength of the frame and the position of the PSF depends on the apparent movement of the planet in the data due to ADI and SDI. This array of unprocessed PSFs of the planet will be used to calculate the perturbation of the KL modes due to the existence of the planet.

Assuming a spectrum, f, of the planet in the case of IFS data (for imaging data, f is just a scalar), we can compute $\Delta Z_k(\mathbf{x})$, the perturbation on the kth KL mode, $Z_k(\mathbf{x})$, due to the existence of a planet, using equation E18 or E20 of Pueyo (2016). We will not write out explicitly the exact formula to compute ΔZ_k here since we have not made any modifications to it. We note that we use Equation E18 from Pueyo (2016) to compute ΔZ_k as it is computationally faster than Equation E20 for a fixed input spectrum. We calculate the perturbed ΔZ_k for the first k_{Klip} KL modes, where k_{Klip} is the number of KL modes we choose to use in our PSF subtraction.

From here, the forward-modeled PSF of the planet for this frame is computed as

$$\begin{aligned}F_{\lambda_p,t_p}(\mathbf{x}) = A_{\lambda_p,t_p}(\mathbf{x})\\ &- \left[\sum_{k=1}^{K_{Klip}} < A_{\lambda_p,t_p}(\mathbf{x}), Z_k(\mathbf{x}) > Z_k(\mathbf{x})\right]\\ &- \left[\sum_{k=1}^{K_{Klip}} \big(< T_{\lambda_p,t_p}(\mathbf{x}), Z_k(\mathbf{x}) > \Delta Z_k(\mathbf{x})\right.\\ &\left.+ < T_{\lambda_p,t_p}(\mathbf{x}), \Delta Z_k(\mathbf{x}) > Z_k(\mathbf{x})\big)\right], \end{aligned} \tag{A.1}$$

where $< \bullet, \bullet >$ is the inner product. Equation A.1 is very similar to Eq. F7 in Pueyo (2016), but is focused on generating the forward model with a fixed input spectrum and not concerned with keeping the planet's spectrum as a free parameter for spectral extraction. As mentioned in Pueyo (2016), the term in the first square bracket is the over-subtraction term that is due to the projection of the KL modes on data with a planet in it, and the terms in the second square bracket are the self-subtraction terms due to the presence of a planet in the reference images influencing the KL modes.

After generating forward models for each frame, we take the mean of all the forward models in time to give us a single forward-modeled PSF cube, $F_{\lambda_p}(\mathbf{x})$ that still has spectral information in the third dimension. In principle, this PSF cube can be used to retrieve spectral information from the data, but it is outside of the scope of this paper. For our astrometry purposes, we also take the mean in wavelength, generating a single 2-D forward-modeled PSF, which we call F, for each dataset.

A.2 BKA Fit Residuals

In Figure A.1, we plot the residuals to the best-fit model for each epoch using BKA. As β Pic b moves closer to its host star, the magnitude of speckle noise increases relative to the signal of β Pic b. This is especially true in the last epoch when β Pic b was closest to its star. However, the fit was still accurate as the bright positive core of the PSF of the planet was successfully modeled and not seen in the residuals, which are consistent with noise.

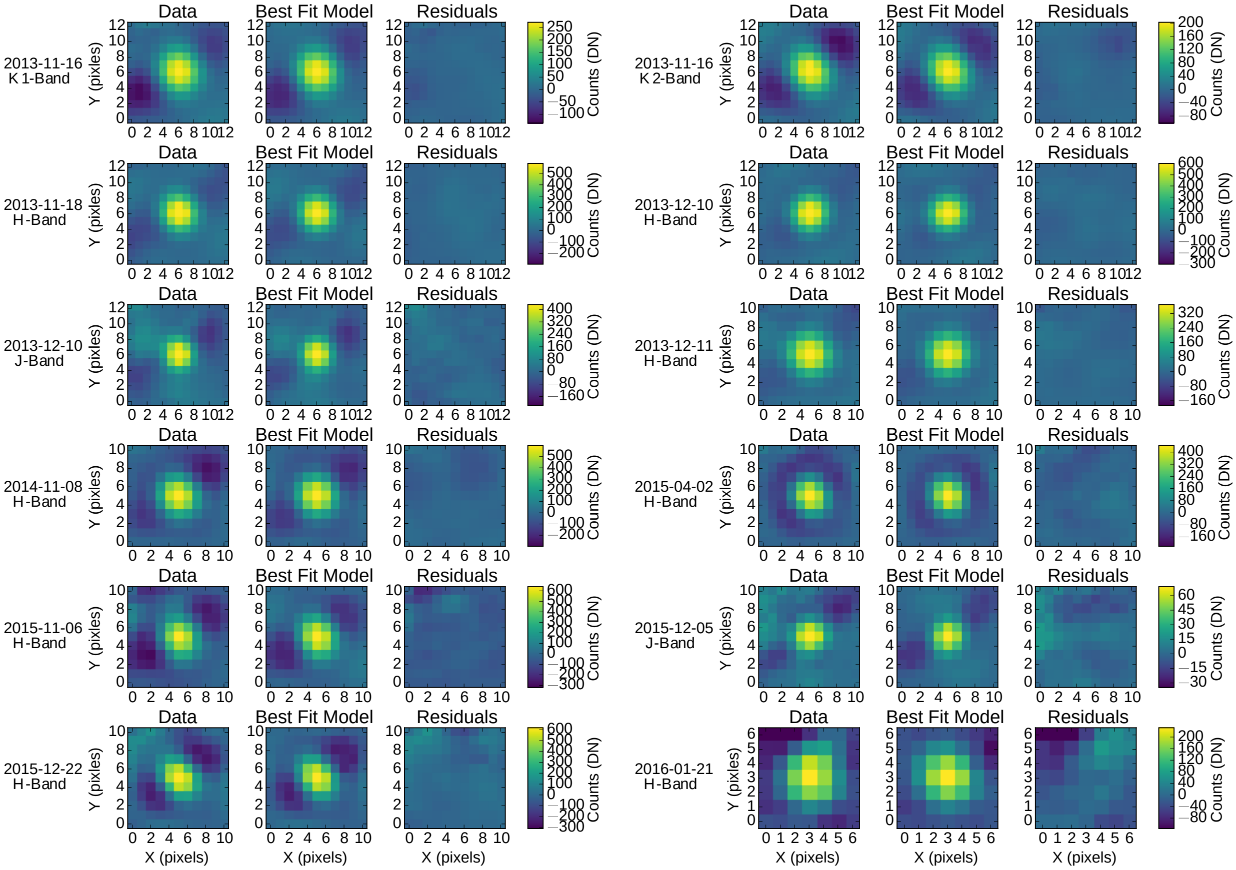


Figure A.1: The data, best fit forward model, and residual map after the model as been subtracted from the data for each of the twelve datasets. In each row, we plot two datasets. For each dataset, we plot the data (left), best fit forward model (center), and residual map (right) on the same color scale. While the scale of each dataset is different, zero is mapped to the same color for all the datasets.

Appendix B

Astrometry with Forward Modeled PSFs in a Frequentist Framework

In Chapter 4, we introduced a procedure to fit the location of a directly-imaged planet using a forward-modeled PSF of the planet in a Bayesian framework. The choice of using a Bayesian framework was improper in this context. While we could obtain the posterior distribution of the planet's location in the image, converting the measurement to an astrophysical quantity, the angular distance between the the planet and star, requires using the astrometric calibration of GPI and associated uncertainties. As the uncertainties in the astrometric calibration of GPI was derived with a frequentist methodology (Konopacky et al. 2014), it is not commensurable to simply combine the Bayesian uncertainties on the location of the planet and the frequentist uncertainties in the astrometric calibration. Repeating the astrometric calibration in a Bayesian framework is possible in order to be compatible with the Bayesian planet astrometry, but would require a significant amount of work.

Here, we present a frequentist framework to measure the position of the planet that preserves the forward modeling and covariance estimation that was developed for the Bayesian technique. We will use the same notation as Section 4.2.5. Using `pyKLIP`, we are able to produce a forward-model planet PSF, F, alongside of the data, D, which is the image of the planet with residual speckle noise due to the imperfect subtraction of the stellar PSF. We want to estimate the location of the planet (x_p, y_p) in the 2-D image and a scale factor α that matches the brightness of the forward model to the planet. We still want to address the covariances in our images, which are dominated by spatially-correlated speckle noise (photon noise is generally negligible for GPI spectroscopy data). Speckle noise is due to both the changing atmosphere and thermal drifts in the instrument. At any location in the image, the brightness of speckles follows a modified Rice distribution with a positive tail, and any realization of speckle noise is coherent in time for less than a second (Fitzgerald & Graham 2006). Thus, in each of our datacubes ($\sim$60 s exposures), each location contains a sum of many speckles. Each individual speckle also has the same spatial size as a point source ($\sim$3 pixels for GPI data). As we observe using ADI, the sky rotates in our images at a rate governed by the Earth's rotation. To sum all of our datacubes together, we must rotate each frame to compensate for the apparent rotation of the sky. As the net rotation for each of our sequences is much greater than the size of a speckle, when we are combining our datacubes, we are adding many independent speckles, where each sample is a sum of many realizations of that speckle over time. By the central limit theorem, the noise in the final combination of data is approximately Gaussian. Thus, we will perform maximum likelihood estimation using a Gaussian likelihood function.

We can write our likelihood function in the same functional form as Section 4.2.5:

$$\begin{aligned} P(D|x_p, y_p, \alpha, l) &= \ln \mathcal{L}(x_p, y_p, \alpha, l) \\ &= -\frac{1}{2}(R^T C_l^{-1} R + \ln(\det C_l) + N_{\text{pix}} \ln(2\pi)). \end{aligned} \tag{B.1}$$

In the same notation as before, $R \equiv R(x_p, y_p, \alpha) = (D - \alpha F(x_p, y_p))_{\mathcal{F}}$, where $\mathcal{F}$ is the fitting region defined by a fixed rectangular box centered at the approximate location of the planet in the data with N_{pix} datapoints in it. R is a $N_{\text{pix}} \times 1$ matrix and C_l is the covariance matrix of size $N_{\text{pix}} \times N_{\text{pix}}$ with a parametrized characteristic spatial correlation size l that we fit for. We use the same Matérn covariance function with $\nu = 3/2$ as in Section 4.2.5, where the only free parameter is the characteristic correlation length scale l. The noise of each pixel is estimated as the standard deviation of pixels in an annulus that is centered on the star with a mean radius corresponding to the distance between that pixel and the star and a width of 6 pixels. We use these noise estimates to fix the amplitude of each element of C_l.

To find the parameters of our model that maximize the probability of obtaining the data we have recorded, we use the open-source `scipy.optimize.minimize` routine available in version 1.0 of `scipy` to minimize the negative of our likelihood function as an equivalent optimization to maximizing the likelihood function. We perform this optimization twice. First, we use the keyword `method='Nelder-Mead'` to use a Simplex algorithm that was the most robust in finding the global minimum. Using the solution from the Simplex algorithm as the initial guess, we run it again using the keyword `method='BFGS'` which uses a quasi-Newton method that also approximates the Hessian inverse during the optimization.

We take the solution from the BFGS method as our estimate for the position of the planet in the data. We use the Hessian inverse estimated at the solution to approximate the covariance of the fitted parameters. We find that the off-diagonal terms are two-orders of magnitude smaller than the diagonal entries, indicating the parameters x_p and y_p are mostly independent. Thus, we estimate the uncertainties on the position of the planet as the square root of the diagonal entries of the Hessian inverse that correspond to x_p and y_p. We convert the measured location of the planet in the image to a relative separation from the star using the astrometric calibration numbers from De Rosa et al. (2015). Uncertainties on the star center, plate scale, and north angle were assumed to be uncorrelated and added in quadrature with the estimated uncertainty on the planet position to obtain the final uncertainty on the relative offset between the planet and the star.

We demonstrate this technique on the 2013 November 16 data and obtain the following measurements: ΔRA $= -228.5 \pm 1.3$ mas, ΔDec $= -366.2 \pm 1.1$, Sep $= 431.7 \pm 1.0$, and PA $= 212.0 \pm 0.2$ degrees. These values are 2-3 times more precise than those reported in Millar-Blanchaer et al. (2015), and demonstrate the ability to reach one milliarcsecond precision with this technique.