

APF-50: A ROBOTIC SEARCH FOR EARTH'S NEAREST NEIGHBORS

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAI'I AT MĀNOA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

ASTRONOMY

MAY 2017

By

Benjamin J. Fulton

Thesis Committee:

Andrew W. Howard, Chairperson
Rolf-Peter Kudritzki
Michael C. Liu
Jessica R. Lu
Jonathan P. Williams
Alexander N. Krot

Keywords: astronomy, exoplanets

Copyright © 2017 by
Benjamin J. Fulton

ACKNOWLEDGMENTS

B.J.F thanks all members of his thesis committee for their expert guidance over the course of his graduate school career. We thank Josh Winn, Jason Rowe, Eric Lopez, Jeff Valenti, Daniel Huber, and Leslie Rogers for contributing insight during many helpful conversations. We thank Kyle Lanclos, Matt Radovan, Will Deich and the rest of the UCO Lick staff for their invaluable help shepherding, planning, and executing observations, in addition to writing the low-level software that made the automated APF observations possible. We acknowledge R. Paul Butler and S. S. Vogt for many years of contributing to the Keck/HIRES data presented here. Chapter 3 and the corresponding publication is dedicated to the memory of our partner and collaborator, Giorgio Corfini, who contributed to the discovery and passed away in 2014. Some of the data presented here were determined directly from observations at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and NASA. We are grateful to the time assignment committees of the University of Hawaii, the University of California, the California Institute of Technology, and NASA for their generous allocations of observing time that enabled these projects. We are very grateful for the donations of the Levy family that helped facilitate the construction of the Levy spectrograph on APF. Without their support the APF would not be contributing to the discovery of planets like those presented in this work. In their honor we informally name the HD 7924 system “The Levy Planetary System”. B. J. F. acknowledges that this material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 2014184874. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The Robo-AO system was developed by collaborating partner institutions, the California Institute of Technology and the Inter-University Centre for Astronomy and Astrophysics, and with the support of the National Science Foundation under Grant Nos. AST-0906060, AST-0960343 and AST-1207891, the Mt. Cuba Astronomical Foundation and by a gift from Samuel Oschin. Ongoing science operation support of Robo-AO is provided by the California Institute of Technology and the University of Hawai‘i. We thank Jason Eastman for providing the EXOFAST code to community. This research made use of the Exoplanet Orbit Database and the Exoplanet Data Explorer at exoplanets.org. This work made use of the SIMBAD database (operated at CDS, Strasbourg, France), NASA’s Astrophysics Data System Bibliographic Services, and the NASA Star and Exoplanet Database (NStED). This research has made use of the Washington Double Star Catalog maintained at the U.S. Naval Observatory. We thank Debra Fischer, Jason Wright, John Johnson and many others for their many nights of observing that contributed to the data presented in this work. We thank Gail Schaefer for her help with the calculations related to the CHARA interferometric observations. The Pan-STARRS1 Surveys (PS1) have been made possible through contributions of the Institute

for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation under Grant No. AST-1238877, the University of Maryland, and Eotvos Lorand University (ELTE). Support for this work was provided by National Science Foundation grant AST-1009749. Based on observations obtained at the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnologa e Innovacin Productiva (Argentina), and Ministrio da Cincia, Tecnologa e Inovao (Brazil). We gratefully acknowledge the efforts and dedication of the Keck Observatory staff, especially Scott Dahm, Greg Doppman, Hien Tran, and Grant Hill for support of HIRES and Greg Wirth for support of remote observing. Finally, the authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawai‘ian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

ABSTRACT

The discovery of thousands of exoplanets during the past decade opened the door to detailed studies of exoplanet demographics. Now we are able to group planets into different categories based on observable characteristics and study population-level properties. Patterns and trends in the known population of planets are emerging which provide insight into the processes that drive the formation and evolution of exoplanets. In this work we develop the tools necessary to discover and accurately characterize a statistically useful sample of exoplanets. We use these tools to discover several new planets and examine the mass function of small planets orbiting bright, nearby stars. By leveraging the fully-robotic Automated Planet Finder telescope, we conduct the “APF-50” Doppler survey which provides greater sensitivity to low-mass planets than was previously possible with classically-scheduled instruments. We study the planet population orbiting stars similar to our sun and also the ultimate fate of these planetary systems by searching for planets orbiting white dwarfs. To date, the statistical power of NASA’s *Kepler* mission remains unmatched due to the shear number of planet detections and unprecedented sensitivity to small planets. We utilize the *Kepler* dataset combined with high-resolution spectroscopy from Keck Observatory to re-examine the radius function of small planets in fine detail. We find that planets between the size of Earth and Neptune typically fall into one of two distinct size groups. We discuss the implications of these findings by comparing to the mass function of small planets measured by the APF-50 survey and find that we are only just beginning to scratch the surface of the population of small planets that *Kepler* found to be so prevalent. This discovery supports the emerging picture that close-in planets smaller than Neptune are composed of rocky cores measuring $1.5 R_{\oplus}$ or smaller with varying amounts of low-density gas that determines their total sizes.

TABLE OF CONTENTS

Acknowledgments	iii
Abstract	v
List of Tables	xiii
List of Figures	xv
1 Introduction and Motivation	1
1.1 Exoplanet Demographics	1
1.2 Mass-Radius Relation	2
1.3 Planet Formation Theories	3
1.4 The APF-50 Project Components	5
2 The stellar obliquity and the long-period planet in the HAT-P-17 exoplanetary system	6
2.1 Introduction	6
2.2 Observations	7
2.2.1 Keck/HIRES Spectroscopy	7
2.2.2 KECK/NIRC2 adaptive optics imaging	8
2.3 Analysis	8
2.3.1 Radial velocities	8
2.3.2 Spectroscopic transit	12
2.3.3 Additional test for misalignment	17
2.3.4 Adaptive optics imaging	18
2.4 Discussion	19

3 KELT-8b: A highly inflated transiting hot Jupiter and a new technique for extracting high-precision radial velocities from noisy spectra	23
3.1 Introduction	23
3.2 Observations	24
3.2.1 KELT Photometry	24
3.2.2 Follow-up Photometry	25
3.2.3 High-resolution Spectroscopy	28
3.2.4 Adaptive Optics Imaging	29
3.3 Stellar Properties	30
3.3.1 Keck/HIRES Spectroscopy	30
3.3.2 Spectral Energy Distribution	31
3.3.3 Evolutionary State	31
3.3.4 Unrelated visual companion	35
3.3.5 UVW Space motion	35
3.4 Synthetic Template Radial Velocities	35
3.5 Planetary Characterization	38
3.5.1 Global Fit	38
3.5.2 False-Positive Assessment	42
3.6 Discussion	45
3.6.1 Prospects for further characterization of the host star	45
3.6.2 Comparative Planetology	45
3.6.3 Mass-Radius predictions	47
3.6.4 Prospects for atmospheric characterziation	47

3.6.5	Irradiation History	47
3.7	Summary	50
4	Three Super-Earths Orbiting HD 7924	53
4.1	Introduction	53
4.2	Stellar Properties	54
4.3	Measurements	55
4.3.1	Keck/HIRES Spectroscopy	55
4.3.2	APF/Levy Spectroscopy	56
4.4	Keplerian Analysis	57
4.4.1	Discovery	57
4.4.2	Characterization	59
4.4.3	False Alarm Assessment	64
4.4.4	Searching for Period Aliases	64
4.4.5	Chromospheric Activity	67
4.5	Photometry	73
4.6	Spitzer Transit search	77
4.7	Architecture and Stability	80
4.7.1	Compact Multi-planet Systems	80
4.7.2	Dynamical Stability	82
4.8	APF vs. Keck	82
4.9	Discussion & Summary	84
5	Three Temperate Neptunes Orbiting Nearby Stars	88
5.1	Introduction	88

5.2	Radial Velocity Measurements	90
5.3	Stellar Properties	92
5.3.1	HD 42618	92
5.3.2	HD 164922	93
5.3.3	HD 143761	95
5.4	Keplerian Analysis	96
5.4.1	Discovery	96
5.4.2	Characterization	100
5.4.3	Bootstrap False Alarm Assessment	102
5.4.4	Brown Dwarf Companion to HD 143761?	106
5.4.5	Additional Planet Candidates	109
5.4.6	Chromospheric Activity	110
5.5	Photometry	110
5.5.1	CoRoT Photometry of HD 42618	110
5.5.2	APT Photometry	117
5.6	Discussion & Summary	121
6	A search for planetary eclipses of white dwarfs in the Pan-STARRS1 medium-deep fields	127
6.1	Introduction	127
6.2	Methods	128
6.2.1	WD sample	128
6.2.2	Control sample	129
6.2.3	Light curves	129

6.2.4	Eclipse detection	130
6.3	Analysis	133
6.3.1	Theoretical eclipse probabilities	133
6.3.2	Occurrence constraints	135
6.4	Discussion	135
6.4.1	Future survey design	137
6.4.2	Pan-STARRS1 3 π	137
6.5	Conclusions	138
7	Mass distribution of low-mass planets in the Solar neighborhood	146
7.1	Introduction	146
7.2	Observations	146
7.2.1	Sample Selection	146
7.2.2	Radial Velocities	148
7.3	Analysis	152
7.3.1	Automated planet search	152
7.3.2	Sensitivity to injected signals	153
7.3.3	Completeness Corrections	154
7.4	Results	158
7.4.1	HD 1461	158
7.4.2	HD 26965	158
7.4.3	HD 32147	158
7.4.4	HD 55575	159
7.4.5	HD 141004	159

7.4.6	HD 164922	159
7.4.7	HD 168009	159
7.4.8	HD 190360	159
7.5	Discussion	160
7.5.1	Candidates and Survey Sensitivity	160
7.5.2	Mass Distribution	161
7.5.3	Comparison with Radius Distribution	161
8	A Gap in the Radius Distribution of Small Planets	170
8.1	Introduction	170
8.2	Sample of Planets	172
8.2.1	California-Kepler Survey	172
8.2.2	Sample Selection	172
8.3	Completeness Corrections	174
8.3.1	Pipeline Efficiency	175
8.3.2	Survey Sensitivity	176
8.3.3	Occurrence Calculation	177
8.4	The Planet Radius Gap	177
8.4.1	Comparison with Log-Uniform Distribution	178
8.4.2	Dip Test of Multimodality	179
8.4.3	Spline Model	179
8.4.4	Relative Frequency of Super-Earths and Sub-Neptunes	180
8.4.5	Two-Dimensional Weighted Kernel Density Estimation	180
8.5	Discussion	182

8.5.1	Previous Studies of the Radius Distribution	182
8.5.2	Rocky to Gaseous Transition	184
8.5.3	Potential Explanations for the Gap	185
8.5.4	Core Mass Distribution	186
8.6	Conclusion	187
9	Conclusion	200
A	Supplementary Information for Chapter 8	203
A.1	Non-cummulative Filters	203
A.2	Weighted Kernel Density Estimation	203
A.3	Validation of the Completeness Corrections	205
	Bibliography	210

LIST OF TABLES

2.1	Radial velocity data ^a	8
2.2	Radial velocity MCMC results	13
2.3	Rossiter-McLaughlin MCMC results	15
3.1	Summary of Photometric Observations	28
3.2	Radial Velocities of KELT-8	29
3.3	Adopted Stellar Properties of KELT-8	34
3.4	Median values and 68% confidence intervals for KELT-8b.	43
3.5	Transit Times for KELT-8b.	44
4.1	Adopted Stellar Properties of HD 7924	55
4.2	Radial Velocities of HD 7924 [†]	58
4.3	Orbital Parameters	62
4.4	Derived Properties	63
4.5	Summary of APT Photometric Observations	80
4.6	APF vs. Keck Radial Velocity Precision	85
5.1	Radial Velocities ^a	92
5.2	Adopted Stellar Properties	93
5.3	Orbital Parameters for HD 42618	102
5.4	Orbital Parameters for HD 164922	103
5.5	Orbital Parameters for HD 143761	104

5.6	Derived Planet Properties	105
6.1	Detection Statistics	132
7.1	APF-50 Target Selection Filters	148
7.2	APF-50 Targets	149
7.3	Planet Detections	160
8.1	Depth of the Gap	173
8.2	Planet Detection Statistics	176
8.3	Planet Occurrence	178
8.4	Spline Fit	179
8.5	Occurrence Rate Comparison	184
A.1	Bin Uncertainty Scaling Factors	208

LIST OF FIGURES

1.1	Contours of search completeness from the Eta-Earth survey at Keck (Howard et al., 2010a). Planet detections are over plotted as green circles and planet candidates are plotted as yellow triangles.	2
1.2	Mass distribution of planets detected by the HARPS-south search for low mass planets (Mayor et al., 2011) before (black) and after (red) completeness corrections for all planets out to orbital periods of 10 years. Smaller planets are clearly more common, but the precise location of the sharp rise in occurrence is unclear as well as the details of the apparently bimodal mass distribution for planets below $\approx 20 M_{\oplus}$	3
1.3	2D map of planet occurrence rates derived from <i>Kepler</i> planet detections. This plot is similar to that of Howard et al. (2012b), but updated with more recent <i>Kepler</i> data containing more planet candidates (Batalha et al., 2013), and extended out to longer orbital periods (250 days vs. 50 days in Howard et al. (2012b)).	4
2.1	<i>Top:</i> Keck/HIRES RV measurements for HAT-P-17 as a function of BJD _{TDB} with the best-fitting two-planet model and associated residuals found by χ^2 minimization of the DE-MCMC chains. A stellar “jitter” term (see Table 5.1) has been added in quadrature to the measurement errors. Data taken during the transit of HAT-P-17b for the purpose of measuring the RM effect was excluded from the RV fit and are not included in this plot. Data to the right of the vertical dashed line are new to this work, and data to the left are from H12. <i>Bottom:</i> same RV measurements phase-folded to the orbital ephemerides of planets b (upper) and c (lower). Phase 0 corresponds to the time of mid-transit (or hypothetical transit). In each case the orbit of the other planet and an arbitrary center of mass velocity relative to a template spectrum (γ) has been removed.	9
2.2	Probability distribution of eccentricity vs. minimum mass ($M \sin i$) in Jupiter masses for planet c from the DE-MCMC analysis. The dashed lines are 68%, 95%, and 99% confidence intervals and the white dot is the median value. The median value is offset from the mode (most likely value) due to the asymmetric posterior distributions. 10	10

- 2.3 *Left:* RV variation during the transit of HAT-P-17b due to the RM effect with best-fit model found from χ^2 minimization overplotted. The upper panel includes the RV variation due to the orbital motion of HAT-P-17b, the middle panel shows the data and model with the orbital motion removed, and the bottom panel shows the residuals to the model. The solid blue line represents our adopted model including all line broadening effects and the convective blueshift, the dashed green line shows the model without the convective blueshift, and the dot-dashed magenta line shows an idealized model in which the line profiles are described only by rotational broadening, similar to the approach of Ohta et al. (2005). Note the asymmetry in the RM curve caused by a combination of the convective blueshift and a slight misalignment. Three data points to the right of the x-axis limit were included in the modeling, but omitted from the plot for clarity. *Right:* Posterior distribution of $v \sin i$ vs. λ from the DEMCMC analysis of the spectroscopic transit. The dashed lines are 68%, 95%, and 99% confidence intervals and the white dot is the median value of the distribution. The dot-dashed line in the right histogram shows the $0.3 \pm 1.5 \text{ m s}^{-1}$ prior on $v \sin i$ from the SME analysis of H12. 14
- 2.4 *Left:* Keck/NIRC2 adaptive optics image. *Right:* contrast achieved based on the final reduced AO image. Our diffraction-limited observations rule out the presence of companions with $\Delta m > 7$ mags for separations beyond $\approx 0.7''$ 19
- 2.5 *Top:* We convert the measured contrast to a mass-sensitivity curve using the Baraffe et al. 2003 evolutionary models. With an age of 7.8 Gyr, we could have detected any stellar companions ($M > 80M_J$) associated with HAT-P-17 at projected separations beyond ≈ 60 AU. *Bottom:* Same as *top* converted into a predicted RV slope using Equation 2.6. The region below the line is allowed by the data. The 3σ slope constraint ($|\dot{\gamma}| \leq 19 \text{ m s}^{-1}\text{yr}^{-1}$) from the RV analysis is just below the lower y-axis limit. 20
- 2.6 *Top:* measured projected obliquity as a function of the alignment timescale calibrated from binary studies (Albrecht et al., 2012b). Stars with temperatures higher than 6250 K are shown with red filled symbols. Blue open symbols show stars with temperatures lower than 6250 K. Stars for which measured effective temperatures include 6250 K in their $1-\sigma$ interval are shown by split symbols. We are computing the relative tidal dissipation rates as a function of stellar type, planet-to-star mass ratio, and orbital distance, using the scaling relations presented by Albrecht et al. (2012). Note that both timescales have been divided by 5×10^9 21
- 3.1 Discovery light curve of KELT-8b from the KELT-North telescope. The photometry is phase-folded to the ephemeris listed in Table 4.3. The red line is the same data binned in intervals of width 0.01 in orbital phase (≈ 46.7 minutes). 25

3.2	<i>Top:</i> Follow-up photometry of KELT-8b primary transits. The source, date, and filter for each transit is annotated. The best-fit models are shown in red. <i>Bottom:</i> All follow-up transit light curves combined and binned in 5 minute intervals. This light curve is not used in the analysis and is shown simply to demonstrate the combined photometric precision of the follow-up light-curves.	26
3.3	<i>Left:</i> Robo-AO image of KELT-8 showing the background giant star 8''8 to the SW of KELT-8. <i>Right:</i> Five σ contrast limit near KELT-8 derived from the Robo-AO image.	30
3.4	SED fit to KELT-8. The red crosses are the photometry measurements. The vertical errors are the 1σ photometric uncertainties, and the horizontal error bars are the effective widths of the passbands. The leftmost point (GALEX NUV) was omitted from the fit due to possible contamination from the neighboring star at 8''8 separation. The black line is the best-fit stellar atmosphere model and the blue points are the predicted passband-integrated fluxes of the model corresponding to our observed photometric bands.	32
3.5	Yonsei-Yale isochrone fit for KELT-8. The solid black line is the isochrone for a star with the mass and metallicity of KELT-8 listed in Table 4.3. The red cross is the position of KELT-8, and the blue points label various ages along the evolutionary track. The green cross is the position of KELT-8 using the spectroscopically-determined $\log g$. The “blue hook” is the kink in the track where the star’s T_{eff} increases while $\log g$ decreases for a short amount of time between an age of 4.9 and 5.8 Gyr.	33
3.6	Astrometry measurements of KELT-8. The green squares were measured using astrometric eyepieces or photographic plates. The black points were measured using CCD data and the most recent point is from our Robo-AO observation. The blue lines are the predicted change in separation (upper panel) and position angle (lower panel) assuming the star 8''8 to the SW of KELT-8 is in the background with negligible proper motion and KELT-8 has a constant proper motion with the values given in Table 5.2. The dashed black line is a model assuming that the two components are physically bound and the orbital motion of the pair is negligible during the span of observations.	36
3.7	Synthetic template RV performance as a function of S/N ratio of the iodine observations. Artificial Gaussian noise was injected into the spectra of a well-known RV standard star (HD 9407) before running them through the RV extraction pipeline for all data points excluding the one at S/N=200. The synthetic template technique shows predictable and consistent performance down to extremely low S/N ratios and a model-limited noise floor of $\approx 10 \text{ m s}^{-1}$	39

3.8	Synthetic template RV performance for three RV standard stars. Artificial Gaussian noise was injected into all spectra for each star so that S/N=50 for each spectrum. This is approximately the same S/N ratio that was collected for each KELT-8 observation. The RV RMS is about 10 m s^{-1} for each star.	40
3.9	<i>Top:</i> APF RVs and residuals for KELT-8. The best-fit model is shown in red. <i>Bottom:</i> RVs and bisector span measurements phase-folded to the best-fit linear ephemeris. The best-fit model is shown in red. The predicted Rossiter-McLaughlin effect assumes perfect spin-orbit alignment and it is not constrained by our data.	42
3.10	Residuals of the transit times from the best-fit ephemeris for the followup light curves. The source of each transit time is denoted in the upper right. The dashed lines indicate the 1σ uncertainty on the linear ephemeris.	44
3.11	BS measurements for the APF spectra used for radial velocity measurements. We find no significant correlation between RV and BS.	46
3.12	48
3.13	HR diagram showing well-characterized host stars of hot Jupiters. The red squares indicate stars that host planets with radii greater than $1.5 R_J$. The green star marks the position of KELT-8, and the black circles are all other systems.	49
3.14	<i>Top:</i> Irradiation history of KELT-8b. The insolation received by the planet is well above the empirical inflation irradiation threshold (Demory & Seager, 2011) for the entire main-sequence existence of the star except in the case of $\log Q'_* = 5$ in the very early stages of stellar evolution. <i>Bottom:</i> Orbital semi-major axis history of KELT-8b. The planet's semi-major axis is rapidly decreasing as the star evolves off the main sequence. It appears unlikely that KELT-8b will survive past the star's current subgiant phase.	51

4.1	Best-fit 3-planet Keplerian orbital model plus one additional long-period Keplerian to model the stellar magnetic activity cycle. The model plotted is the one that produces the lowest χ^2 while the orbital parameters annotated and listed in Tables 4.3 and 5.6 are the median values of the posterior distributions. a) Full binned RV time series. Open black squares indicate pre-upgrade Keck/HIRES data (see §5.2), open black circles are post-upgrade Keck/HIRES data, and filled green diamonds are APF data. The thin blue line is the best fit 3-planet plus stellar activity model. We add in quadrature the RV jitter term listed in Table 4.3 with the measurement uncertainties for all RVs. b) Residuals to the best fit 3-planet plus stellar activity model. c) Binned RVs phase-folded to the ephemeris of planet b. The two other planets and the long-period stellar activity signal have been subtracted. The small point colors and symbols are the same as in panel a . For visual clarity, we also bin the velocities in 0.05 units of orbital phase (red circles). The phase-folded model for planet b is shown as the blue line. d) 2DKLS periodogram comparing a 2-planet plus activity model to the full 3-planet fit when planet b is included. Panels e) and f) , and panels g) and h) are the same as panels c) and d) but for planets HD 7924 c and HD 7924 d respectively.	60
4.2	Velocity-activity correlation. <i>Top</i> : Binned RV time series of the post-upgrade Keck data with planets b, c, and d subtracted. <i>Middle</i> : Binned S_{HK} time series of the post-upgrade Keck data only. Note the similarities between the variability in the top and middle panels. <i>Bottom</i> : Spearman rank correlation test of the velocities with S_{HK} values (Spearman, 1904).	61
4.3	<i>Top</i> : Distribution of maximum periodogram peak heights for 1000 2DKLS periodograms of scrambled RV time series. For each periodogram planets c, d, and the magnetic activity cycle were subtracted before scrambling the data set. The vertical dashed blue line marks the height of the original peak for planet b which is clearly separated from the distribution of peaks caused by random fluctuations. <i>Middle</i> : Same as the top panel for planet c. <i>Bottom</i> : Same as the top panel for planet d.	65
4.4	The window function of the Keck and APF RV time series. While the APF window function has some power at frequency multiples of one day, it is flat otherwise, whereas the Keck window function has power at low frequencies (corresponding to long-period aliases) and power in broader swaths around the frequency multiples of one day.	67
4.5	Left: LS periodograms of the data associated with each planet identified, from top to bottom: planet b, planet c, and planet d. In each panel, signals from the other planets and stellar activity have been subtracted. The phase of the frequency associated with the peak is given in radians. Right: LS periodograms of the best Keplerian model for each of the planets, from top to bottom, planets b, c, and d. The periodogram of each Keplerian model reproduces the peak period and alias structure of the data.	68

- 4.6 Top: a candidate periodic signal at 40.8 or 17.1 days emerges in the periodogram of the residuals to the 3-planet plus long-term stellar activity model. These two periods are related to each other by the one synodic month alias. Either 40.8 days or 17.1 days could correspond to the rotation period of the star, although the 40.8 day period is more prominent in the periodogram of the S_{HK} values. To test both periods, we model the best-fit Keplerian at 40.83 days (center) and 17.1 days (bottom) and show their periodograms, complete with phase information. 69
- 4.7 *Top*: Periodogram of Keck HIRES RV measurements \pm 100-days versus time (JD-2450000) with our Keplerian model of planet b ($P=5.4$ days) removed. Each colored, single-pixel column is a unique periodogram of all RV measurements within a 200-day window centered on the time indicated on the horizontal axis. Dark grey boxes indicate times of observation having fewer than 24 measurements of both RV and S_{HK} within \pm 100 days. Each periodogram has been normalized independently by dividing all periodogram powers by the maximum periodgram value. *Middle*: Same as top, with planets b and c removed. *Bottom*: Same as top, with planets b, c, and d removed. 74
- 4.8 *Top*: Same as Figure 4.7 but for S_{HK} (stellar activity) data. *Bottom*: Same as Figure 4.7 bottom panel but for a synthetic RV time-series created by superposing a 1.46 m s^{-1} , 40.8-day sinusoid with the best 3-planet model to the real Keck RVs and 2.36 m s^{-1} Gaussian noise, then fitting for and removing the best 3-planet model ($P=5.4, 15.3, 24.5$ days). Time sampling matches the real-data. Dark grey boxes indicate times of observation having fewer than 24 measurements of both RV and S_{HK} within \pm 100 days. Epoch-to-epoch fluctuation in the recovered period of the injected 40.8-day signal is an artifact of time-sampling, Keplerian fitting noise, and injected Gaussian noise. 75
- 4.9 *Top*: Running periodogram of the entire RV time-series, where only the subset of RV measurements that precede each time have been decorrelated with stellar activity. Keplerian models of planets b ($P=5.4$ days), c ($P=15.3$ days) and the long-term-stellar-magnetic-activity have been removed. Dark grey boxes indicate times of observation having fewer than 24 measurements of both RV and S_{HK} within \pm 100 days. All periodogram powers have been normalized to the same color scale. *Bottom*: Periodogram of entire RV time-series before and after decorrelation with stellar activity. The periods of peaks attributed to planets and stellar activity are labeled accordingly by dashed vertical lines. 76

- 4.10 Differential photometry of HD 7924 from APT. *Top*: Relative flux time-series of HD 7924 in the combined Strömgren ($b + y$)/2 passband. Seasonal offsets are removed by dividing by the mean within each season. The standard deviation of the photometric time series is 2.3 parts per thousand (ppt). *Middle*: LS periodogram of the photometric time series. *Bottom*: Differential photometry from APT phase-folded to the 16.9 day peak found in the LS periodogram and binned to widths of 0.1 phase units (red circles). The light grey points and open red circles show the same data wrapped by one period. The peak-to-peak amplitude of the 16.9 day photometric signal is 1.1 ppt. 78
- 4.11 Differential photometry from APT of HD 7924 phase-folded to the ephemerides of the three planets. *Left*: Differential photometry phase-folded to the orbital period of planet b (top), c (middle), and d (bottom). Red circles are the photometry measurements grouped in bins of width 0.04 units of orbital phase. The grey shaded region spans the uncertainty in mid-transit time. The semi-amplitude of a sine wave least squares fit to the data is annotated at the bottom of each panel. *Right*: Same as the left panels zoomed in around the phase of center transit. The individual measurements are omitted and the red circles correspond to binned photometry data in bins of width 0.005 units of orbital phase. The three curves are Mandel & Agol (2002) transit models for planets with the masses listed in Table 5.6 and densities of iron (black), water (blue), and hydrogen (green). 79
- 4.12 Comparison of the transit search time window covered by the *Spitzer* observations of Kammer et al. (2014) to the transit times as calculated from the updated ephemeris in this work. The thick black line shows the posterior probability density function (PDF) of the mid-transit time for the transit that was targeted by Kammer et al. (2014). The blue shaded region shows the time window covered by their observations, and the red dashed lines show the most likely ingress and egress times. 81
- 4.13 The architecture of *Kepler* multi-planet systems compared with the architecture of the HD 7924 system. We plot the *Kepler* systems with exactly three currently known planets, a largest orbital period of 30 days or less, and a maximum planetary radius of $4 R_{\oplus}$. The HD 7924 planets are plotted as filled black circles. The systems are sorted by the semi-major axis of the shortest period planet in the system. The size of the circles is proportional to planet radius. The radii of the HD 7924 planets have been calculated as $\frac{R_p}{R_{\oplus}} = 0.371 \left(\frac{M_p}{M_{\oplus}} \right)^{1.08}$ (Weiss & Marcy, 2014). 83
- 4.14 A comparison between RV performance at Keck and APF for two well known RV standard stars. These stars are some of the most well-observed stars at both Keck and APF. In the first 1.5 years of APF operations we have already collected nearly half of the data that has been collected at Keck over the last \sim 15 years. Stellar jitter has been added in quadrature with the binned velocities such that the χ^2 of the velocities with respect to their median value is 1.0. 86

5.1	Observed squared visibility versus spatial frequency for HD 164922 (blue diamonds). The red line shows the best fit limb-darkened model. See Section 5.3.2 for a discussion of the implications of this figure.	94
5.2	One-planet Keplerian orbital model plus one additional long-period Keplerian to model the stellar magnetic activity cycle for HD 42618. The most likely model is plotted but the orbital parameters annotated on the figure and listed in Tables 5.3 and 5.6 are the median values of the posterior distributions. The process used to find the orbital solution is described in 5.4.2. a) Full binned RV time series. Open black squares indicate pre-upgrade Keck/HIRES data (see §5.2), open black circles are post-upgrade Keck/HIRES data, and filled green diamonds are APF measurements. The thin blue line is the most probable 1-planet plus stellar activity model. We add in quadrature the RV jitter term listed in Table 5.3 with the measurement uncertainties for all RVs. b) Residuals to the most probable 1-planet plus stellar activity model. c) Binned RVs phase-folded to the ephemeris of planet b. The long-period stellar activity signal has been subtracted. The small point colors and symbols are the same as in panel a . For visual clarity, we also bin the velocities in 0.08 units of orbital phase (red circles). The phase-folded model for planet b is shown as the blue line. d) 2DKLS periodogram showing the improvement to χ^2 for a model including the long period activity signal and a single planet compared to a model that only includes the activity signal.	98
5.3	Two-planet Keplerian orbital model posterior distributions for HD 164922. The most likely model from the posterior distribution is plotted while the orbital parameters annotated on the figure and listed in Tables 5.4 and 5.6 are the median values of the posterior distributions. The process used to find the orbital solution is described in 5.4.2. a) Full binned RV time series. Open black squares indicate pre-upgrade Keck/HIRES data (see §5.2), open black circles are post-upgrade Keck/HIRES data, and filled green diamonds are APF measurements. The thin blue line is the most probable 2-planet model. We add in quadrature the RV jitter term listed in Table 5.4 with the measurement uncertainties for all RVs. b) Residuals to the most probable 2-planet model. c) Binned RVs phase-folded to the ephemeris of planet b. The Keplerian orbital model for planet c has been subtracted. The small point colors and symbols are the same as in panel a . For visual clarity, we also bin the velocities in 0.08 units of orbital phase (red circles). The phase-folded model for planet b is shown as the blue line. d) 2DKLS periodogram comparing a model including only the long period planet to the two planet model. Panels e) and f) are the same as panels c) and d) but for planet HD 164922 c. The shaded region of panel a is re-plotted in Figure 5.4.	99
5.4	Recent RVs for HD 164922 highlighted in the grey box of Figure 5.3 panel a) featuring the high cadence APF observations collected during the most recent observational season.	100
5.5	Same as Figure 5.3 but for planets HD 143761 b and c.	101

5.6	Graphical representation of the bootstrap false alarm tests described in Section 5.4.3. <i>Top:</i> Distribution of maximum periodogram peak heights for 1000 2DKLS periodograms of scrambled RV time series for HD 42618. The long period activity signal was subtracted before scrambling the data set. The vertical dashed blue line marks the height of the original periodogram peak for planet b which is clearly separated from the distribution of peaks caused by random fluctuations. <i>Middle:</i> Same as the top panel for planet HD 164922 c. <i>Bottom:</i> Same as the top panel for planet HD 143761 c. In each case, the scrambled RVs generate peaks with significantly lower power than the power observed from the new planetary signals.	107
5.7	Limiting magnitude as a function of separation from the ρ CrB. Also shown are a cubic spline interpolation of the 5σ detection limit (dashed line) and limiting magnitudes for $0''.1$ and $0''.2$.	108
5.8	Velocity-activity correlation for HD 42618. A discussion of the chromospheric activity of each of the three stars can be found in Section 5.4.6. <i>Top:</i> Binned RV time series of the post-upgrade Keck data with planet b. <i>Middle:</i> Binned S_{HK} time series of the post-upgrade Keck data only. Note the similarities between the variability in the top and middle panels. <i>Bottom:</i> Spearman rank correlation test of the velocities with S_{HK} values (Spearman, 1904). We do not subtract this correlation from the RVs of HD 42618 but instead model the magnetic activity cycle as an additional long-period Keplerian (see Section 5.4.2).	111
5.9	Lomb-Scargle periodograms of S_{HK} chromospheric activity. In each panel the period of the planet announced in this work is marked by the blue dashed line and the power corresponding to an analytical false alarm probability of 1% is marked by the red dotted line (Schwarzenberg-Czerny, 1998). S_{HK} values measured from spectra with $SNR < 40$ per pixel or exposure times $> 25\%$ longer than the median exposure time (due to clouds and/or seeing) can be badly contaminated by the solar spectrum and cause our S_{HK} extraction pipeline to produce large outliers. These measurements were excluded before calculating the periodograms. No significant periodicity is detected in any of the stars at the orbital periods of the new planets. <i>Top:</i> Periodogram of S_{HK} values for HD 42618. <i>Middle:</i> Periodogram of S_{HK} values for HD 143761. <i>Bottom:</i> Periodogram of S_{HK} values for HD 164922.	112
5.10	CoRoT photometry of HD 42618 discussed in Section 5.5.1. <i>Top:</i> Detrended light curve. <i>Middle:</i> Lomb-scargle periodogram of the light curve. <i>Bottom:</i> Photometry phase-folded to the period corresponding to the highest peak in the Lomb-Scargle periodogram (16.9 days). We also bin the photometry with bin widths of 0.05 units of phase (red circles).	114

- 5.11 CoRoT photometry of HD 42618 phase-folded to the orbital period of planet b. The transit search for HD 42618 b is discussed in Section 5.5.1. *Top:* Photometry over the full orbital phase of planet b. The red circles are binned photometric measurements with bin widths of 0.04 units of orbital phase. The grey shaded region shows the 1σ uncertainty on the time of inferior conjunction derived from the RV modeling. *Bottom:* Same as top panel with the x-axis zoomed-in near the time of inferior conjunction. In this panel we only plot the measurements binned with bin widths of 0.002 units of orbital phase. Again, the shaded region represents the 1σ uncertainty on the time of inferior conjunction. The black transit model shows the predicted transit depth for a solid iron planet using the mass-radius relation of Weiss & Marcy (2014). 115
- 5.12 Smoothed one-dimensional power spectrum of HD 42618 from the CoRoT data. The comb of peaks in the power spectrum near a frequency of 3 mHz is the signature of solar-like asteroseismic oscillations. Our asteroseismic analysis of HD 42618 is described in Section 5.5.1. 116
- 5.13 Long-term photometric observations of the planetary candidate host stars HD 42618 (top), HD 143761 (middle), and HD 164922 (bottom) acquired with TSU’s T11 0.80 m, T4 0.75 m, and T12 0.80 m APTs at Fairborn Observatory in southern Arizona. All three stars are plotted with identical x and y scales. The horizontal line in each panel marks the mean of each data set. The APT photometry and analysis are described in Section 5.5.2. 118
- 5.14 *Top:* Fifteen years of photometric observations of HD 42618 from the top panel of Fig. 5.13 plotted against the 149.6-day planetary orbital period and time of conjunction derived from the radial velocity observations. A least-squares sine fit on the radial velocity period gives a semi-amplitude of just 0.000037 ± 0.000033 mag, firmly establishing the lack of stellar activity on the radial velocity period and thus confirming the presence of stellar reflex motion caused by an orbiting planet. *Bottom:* Closeup of the observations near the time of planetary conjunction at phase 0.0. The solid line shows a toy model transit of a sphere of constant 1.0 g cm^{-3} density and radius determined by the relation of Weiss & Marcy (2014). The vertical lines mark the uncertainty in the predicted transit times. Our current photometric observations provide no evidence for transits. 119

5.15 <i>Top:</i> Yearly means of the Mt. Wilson-calibrated S_{HK} values acquired along with the radial velocity measurements. <i>Bottom three panels:</i> Yearly means of HD 42618's $P - C1$, $P - C2$, and $C2 - C1$ differential magnitudes. The horizontal dotted lines designate the grand means of the observations while the numbers in the lower-left and lower-right give the total range and standard deviation of each data set, respectively. It is evident that we have resolved low-level brightness variability in HD 42618 compared to the two comparison stars, $C1$ and $C2$. Low-amplitude cycles of roughly 0.001 mag over 5 years are seen in both the $P - C1$ and $P - C2$ light curves. There appears to be little or no correlation of S_{HK} values with photometric brightness.	120
5.16 Same as Figure 5.14 but for HD 164922 b and c.	122
5.17 Same as Figure 5.15 but for HD 164922. In this case we see a positive correlation of the brightness of HD 164922 with the S_{HK} index.	123
5.18 Same as Figure 5.14 but for HD 143761 b and c.	124
5.19 Same as Figure 5.15 but for HD 143761. In this case we again see a positive correlation of the brightness of HD 143761 with the S_{HK} index.	125
6.1 Astrometrically-selected WDs (blue) and control sample stars (red). The small black points are all detections from the deep stacks that were not selected for either the control or WD samples.	130
6.2 The color-selected WDs (blue points) are identified by the narrow tail of extremely blue stars in the $(g_{P1} - r_{P1})$ vs. $(r_{P1} - i_{P1})$ color plane. The small black points are all detections from the deep stacks not selected for either the control or WD samples.	131
6.3 <i>Top:</i> Distribution of relative flux measurements for all WD and control sample stars. The solid blue line is the distribution for the WDs and the dashed red line is for the control sample stars. <i>Bottom:</i> Distribution of relative flux measurements divided by the measurement uncertainties corrected by adding in quadrature the reported measurement uncertainties with the standard deviation of the light curve (by filter). Measurements with $\Delta F/\sigma_{lc} \geq 5$ are considered eclipse candidates. As in the top panel, the solid blue line is the distribution for the WDs and the dashed red line comes from the control sample stars.	140

- 6.4 Illustration of the filtering process for a light curve of a typical $g' = 18.9$ WD in medium-deep field 3. The total number of measurements and the number of dropouts ($\Delta F/\sigma_{lc} \geq 5$) are shown the the lower-right of each panel. Dropout candidates are plotted as triangles. *a)* Raw light curve before any filtering. Error bars are equivalent to the reported measurement uncertainties. Notice the large number (1977) of dropout candidates. *b)* Light curve after applying the chip location-based filters described in §6.2.4. *c)* Light curve after removing measurements in which neighboring stars show large deviations from the median flux level or large scatter. *d)* Light curve after de-correlating against the neighboring star relative flux and re-scaling the measurement uncertainties by adding the reported uncertainties in quadrature with the standard deviation of the light curve in each filter. This tends to inflate the error bars and pushes the vast majority of dropout events below the 5-sigma cutoff. *e)* Light curve after the final level of photometry-based filtering. In this stage we compare the CCD pixel positions of the stars during dropout events with known masked regions of the CCD array. Two dropout events remain after all photometry-based filters. Postage stamp images are downloaded and visually inspected for the remaining dropout events. 141
- 6.5 Candidate z-band eclipse with a reported depth of 53% that was not filtered by the automated filtering techniques described in §6.2.4. *a)* 5.8 hour stack of the $2.5' \times 2.5'$ region centered on the target. The target WD is circled. *b)* The same field of view as panel *a* from the single exposure corresponding to the reported 53% deep eclipse. *c)* A difference image of the stack in panel *a* convolved and scaled to match the PSF and subtracted from the dropout image in panel *b*. Notice that all stars – including the target – show no detectable residual flux. *d)* Same as panel *c* with a synthetic 53% eclipse injected onto the target before the image subtraction. The negative residuals on the target are clearly evident. 142
- 6.6 *Left:* probability of measuring an eclipse with depth ΔF during a single 240 s exposure of a random WD that hosts a single companion with the orbital parameters shown. p is the planet to star radius ratio, R_p is the radius of the planet in Earth radii, a/R_{WD} is the orbital semi-major axis scaled to the radius of the WD, and a is the semi-major axis in AU. *Right:* Model eclipse light curves for the planet parameters shown on the left panel and an impact parameter 1.0. The red circle is the mean flux for an exposure centered on the mid-eclipse time. The bar extending from the red circle shows the length of the exposure time. This is the largest signal that we could expect to find for planets with these parameters. This corresponds to the maximum $\langle \Delta F \rangle$ bin with a probability greater than zero in the left panel. 142

6.7	Expected detectable eclipse rate per million exposures of the medium-deep survey. An eclipse is deemed detectable if the depth is greater or equal to five times the measurement uncertainty. The measurement uncertainty is calculated by adding the reported uncertainty in quadrature with the standard deviation of the light curve on a per filter basis. The dashed line marks the point at which the eclipse duration is equal to the integration time. Eclipses caused by objects with parameters that fall in the region above and to the right of the dashed line will have eclipses that may span multiple adjacent exposures. Our assumption that each light curve measurement is independent is invalid in this regime and our expected eclipse rate will be slightly overestimated.	143
6.8	<i>Top:</i> Maximum planet occurrence rate compatible with the observations at 95% confidence. <i>Bottom:</i> Maximum planet occurrence rate compatible with the observations at 68% confidence. In both panels dashed line is the same as in Figure 6.7. The maximum occurrence rates will be slightly underestimated in the region to the upper right of this dashed line.	144
6.9	Maximum planet occurrence rate consistent with our data as a function of planet radius at a semi-major axis of $a = 0.01$ AU for confidence levels of 95% (solid) and 68% (dashed). Shaded regions are disfavored by our data. This plot represents a slice through Figure 6.8 at $a = 0.01$ AU.	145
6.10	Expected detectable eclipse rates calculated as described in §6.3.1 for hypothetical surveys using the Pan-STARRS1-like throughput with different exposure times. The numbers within the dashed box indicate the mean eclipse rate in that region of parameter space. Shorter exposure times give increased eclipse detectability for the shortest-period objects within ~ 0.03 AU but planets orbiting this close to their host WD would likely be ripped apart by tidal forces. Although the mean eclipse rate in the region of interest goes up with longer exposure times this is reversed if you consider a fixed total survey exposure time (take twice as many 60 second exposures as 120 second exposures, etc.). However, the eclipse rates remain nearly constant indicating that the best way to increase sensitivity in this regime is to increase the number of epochs observed (larger number of WDs and/or higher cadence).	145
7.1	Hipparcos H-R diagram showing the APF-50 targets. The large colored points are the APF-50 targets. The point size indicates apparent magnitude with larger points being brighter stars. The color of the point reflects stellar activity as measured from the Ca II H & K index. The light grey points are other Hipparcos stars within 50 pc that were not selected for the APF-50 survey.	147
7.2	Number distributions of RV measurements.	150

7.3	<i>Top:</i> Distributions of velocity uncertainties for the APF-50 survey. The median velocity uncertainty for Keck and APF is annotated. The bimodal distribution is driven by differences in observing strategies. The measurements with smaller uncertainties consist of three consecutive measurements binned together, while the larger uncertainties consist of only a single spectrum. <i>Bottom:</i> Distribution of the RMS of the velocity residuals after subtracting all known signals. The medians of both distributions are annotated.	151
7.4	<i>Top:</i> Mean survey completeness as a function of velocity semi-amplitude and orbital period. Planet detections are plotted as green circles. <i>Bottom:</i> Mean survey completeness as a function of $M \sin i$ and semi-major axis.	155
7.5	<i>Top:</i> 2DKLS periodogram and RV time-series of the quiet star HD 221354. The solid red line is our cutoff for a positive detection for an eFAP=1%. The dotted vertical lines mark the 1-year and 1-month period aliases of the highest period peak (203.7 days). The periodogram shows that no signals fall above or near our detection threshold. <i>Bottom:</i> Results of our injection recovery tests for HD 221354. The blue dots indicate the position of injected planets that were successfully recovered by the pipeline, and red dots indicate injections that were not recovered.	156
7.6	Example candidate vetting plots for HD 141004 b produced by the automated planet detection algorithm. A similar plot is produced at each iteration of the planet search. <i>Top:</i> 2DKLS periodogram of the RV time-series comparing the χ^2 of a single-planet Keplerian fit to that of the null hypothesis (flat line). The red line is our cutoff for a positive detection for an eFAP=1%. The dotted vertical lines mark the 1-year and 1-month period aliases of the highest period peak (15.5 days). <i>Top-middle:</i> the complete RV time-series for HD 141004 that spans \sim 10 years with the best-fitting single-planet Keplerian fit over-plotted in blue. <i>Bottom-middle:</i> residuals to the single-planet Keplerian fit. <i>Bottom:</i> the RV data phase-folded to the period corresponding to the highest peak in the periodogram and binned in orbital phase (large red points).	157
7.7	Same as Figure 7.6 for 1461 d.	162
7.8	Same as Figure 7.6 for 26965 b.	163
7.9	Same as Figure 7.6 for 32147 b.	164
7.10	Same as Figure 7.6 for 55575 b.	165
7.11	Same as Figure 7.6 for 164922 d.	166
7.12	Same as Figure 7.6 for 168009 d.	167
7.13	Same as Figure 7.6 for 190360 d.	168

7.14 Planet occurrence as a function of planet mass in the APF-50 sample for planets with orbital periods less than 1000 days. The dotted grey line shows the actual planet detections and the solid black line is the completeness-corrected mass distribution. To within uncertainties the occurrence rate as a function of planet mass rises gradually between 30 and 7 M_{\oplus} with a possible flattening or fall-off in planet occurrence below 7 M_{\oplus}	169
8.1 <i>Top:</i> HR diagram of the sample of stars selected for analysis. The full Paper II sample is plotted in light grey points and the sample selected for analysis after applying the filters discussed in Section 8.2.2 are plotted as blue squares. Giant planet hosting stars that fall above the dashed line given by Equation 8.1 are omitted from the final sample. <i>Bottom:</i> Stellar radius of CKS stars as a function of <i>Kepler</i> magnitude (K_p). We note that stars fainter than 14.2 do not follow the same stellar radius distribution. We omit stars fainter than $K_p = 14.2$ to avoid biasing our planet radius distribution. The point colors are the same as in the <i>top</i> panel.	188
8.2 (a) Size distribution of all planet candidates in the CKS planet sample. Panels (b)–(g) show the radius distribution after applying several successive cuts to (b): remove known false positives, (c): keep candidates orbiting bright stars ($K_p < 14.2$), (d): retain candidates with low impact parameters ($b < 0.7$), (e): keep candidates with orbital periods shorter than 100 days, (f): remove candidates orbiting giant host stars, and (g): include only candidates orbiting stars within our adopted T_{eff} range (4700 K $< T_{\text{eff}} <$ 6500 K). The number of planets remaining after applying each successive filter is annotated in the upper right portion of each panel. Our filters produce a reliable sample of accurate planet radii and accentuate the deficit of planets at 1.8 R_{\oplus}	189
8.3 Histograms of planet radii broken up into the three magnitude ranges annotated in each panel. All of the filters have been applied to the sample as described in §8.2.2. The gap is apparent in all magnitude ranges. The distribution of planet radii in the two brightest magnitude ranges are indistinguishable (p-value = 0.6). However, the planets orbiting stars with $K_p > 14.2$ are statistically different (p-value = 0.0004) when compared to the $K_p = 13.5$ –14.2 magnitude range. This is expected due to the non-systematic nature of the target selection for CKS and KIC stars fainter than $K_p = 14.2$. This motivates our removal of planets with hosts fainter than $K_p = 14.2$. 190	
8.4 <i>Top:</i> Histogram of stellar radii derived in Paper II and used to update planet radii in this work after the filters described in Section 8.2.2 are applied. <i>Bottom:</i> Histogram of planet-to-star radii ratios for the stars remaining after the filters described in Section 8.2.2 are applied to the full Paper II sample of planet candidates. In both cases, the median measurement uncertainties are plotted in the upper right. Neither of these two histograms shows the same bimodal feature that is observed in the planet radius distribution, which demonstrates that the feature is not an artifact of our stellar sample or transit fitting.	191

8.5	Fraction of injected transit signals recovered as a function of signal to noise ratio (m_i , Equation 8.2) in our subsample of the <i>Kepler</i> target stars using the injection recovery tests from Christiansen et al. (2015). We fit a Γ CDF (Equation 8.3) and plot the best-fit model in green.	192
8.6	<i>Top:</i> Mean survey completeness for transiting planets orbiting the stars in our sample (p_{det}). <i>Bottom:</i> Mean survey completeness for all planets orbiting stars in our sample ($p_{\text{det}} \cdot p_{\text{tr}}$).	193
8.7	<i>Top:</i> Completeness-corrected histogram of planet radii for planets with orbital periods shorter than 100 days. Uncertainties in the bin amplitudes are calculated using the suite of simulated surveys described in Section A.3. The light gray region of the histogram for radii smaller than $1.14 R_{\oplus}$ suffers from low completeness. The histogram plotted in the dotted grey line is the same distribution of planet radii uncorrected for completeness. The median radius uncertainty is plotted in the upper right portion of the plot. <i>Bottom:</i> Same as top panel with the best-fit spline model over-plotted in the solid dark red line. The region of the histogram plotted in light grey is not included in the fit due to low completeness. Lightly shaded regions encompass our definitions of “super-Earths” (light red) and “sub-Neptunes” (light cyan). The dashed cyan line is a plausible model for the underlying occurrence distribution after removing the smearing caused by uncertainties on the planet radii measurements. The cyan circles on the dashed cyan line mark the node positions and values from the spline fit described in §8.4.3.	194
8.8	<i>Top:</i> Two-dimensional planet radius distribution as a function of orbital period using stellar parameters from the Q16 catalog. <i>Bottom:</i> Two-dimensional planet radius distribution as a function of orbital period using updated planet parameters from Paper II. In both cases the median uncertainty is plotted in the upper left. Individual planet detections are plotted as black points. The contours are corrected for completeness using the wKDE technique.	195
8.9	Two-dimensional planet radius distribution as a function of stellar radius using updated planet parameters from Paper II. The median uncertainty is plotted in the upper left. Individual planet detections are plotted as black points. The underlying contours are not corrected for completeness. The bifurcation of planet radii is independent of the size of the host star.	196

8.10 <i>Top:</i> Two-dimensional distribution of planet size and incident stellar flux. The median uncertainty is plotted in the upper left. There are at least two peaks in the distribution. One class of planets has typical radii of $\sim 1.3 R_{\oplus}$ and generally orbit in environments with $S_{\text{inc}} > 100 S_{\oplus}$, while another class of slightly larger planets with typical radii of $\sim 2.4 R_{\oplus}$ orbit in less irradiated environments with $S_{\text{inc}} < 200 S_{\oplus}$. <i>Bottom:</i> Same as top panel with individual planet detection points removed, annotations added, and vertical axis scaling changed. The region enclosed by the dashed blue lines marks the photoevaporation desert, or hot-Super Earth desert as defined by Lundkvist et al. (2016). The shaded region in the lower right indicates low completeness. Pipeline completeness in this region is less than 25%. The purple and black lines show the scaling relations for the photoevaporation valley predicted by Lopez & Rice (2016) for scenarios where these planets are the remnant cores of photoevaporated Neptune size planets (dashed purple line) or that these planets are formed at late times in a gas-poor disk (dotted black line).	197
8.11 Histograms of planet radii broken up into the ranges of incident flux (S_{inc}) annotated in the upper right region of each panel. Planets orbiting in environments of higher S_{inc} tend to be smaller than those in low S_{inc} environments. Regions of the histograms plotted in light grey are highly uncertain due to pipeline completeness (<25%).	198
8.12 An empirical fit to planet radius and mass measurements from (Weiss et al., 2016) over-plotted on the completeness-corrected planet radius distribution derived in this work. The maximum in the planet density fit peaks near the gap in the planet radius distribution.	199
8.13 Model for planet radius as a function of envelope size from Lopez & Fortney (2014). The final planet radius is plotted for a simulated planet with a $2 M_{\oplus}$ core mass that has been irradiated by 32 times the incident flux received by Earth for a period of 5 Gyr. A bare $2 M_{\oplus}$ core has a radius of $1.2 R_{\oplus}$. Adding an envelope of H/He which is less than 0.2% of the planet's mass inflates the planet to over $1.6 R_{\oplus}$. An additional 0.7% envelope by mass inflates the planet to $2 R_{\oplus}$	199
A.1 (a) size distribution of planet candidates from the CKS sample. (b) planets removed by the specified filter. Panels (c)–(n) show the radius distribution and planets removed from the full sample after applying only a single cut removing known false positives (c), planets orbiting faint host stars (e), planets with grazing transits (g), planets with orbital periods longer than 100 days (i), planets orbiting giant host stars (k), and planets orbiting host stars cooler than 4700 K or hotter than 6500 K (m). No completeness corrections have been applied. The $b < 0.7$ cut and the $K_p < 14.2$ cut remove the most planet candidates, but no filter preferentially removes planets in the gap (between blue dotted lines).	204

A.2	Bin-free view of the planet radius distribution calculated using wKDE (Equation A.1). The 1-sigma uncertainty region is shaded in red and calculated using a suite of simulated transit surveys as described in Appendix A.2.	205
A.3	Radius vs. period distribution for simulated sample of planets. For plotting clarity and speed we plot only 1,000 randomly chosen simulated planets out of the 45,000 simulated planets.	206
A.4	<i>Top:</i> Results from simulating 100 transit surveys with a known input distribution of planets. The input distribution of simulated planets is plotted in blue, and the simulated detections are plotted in a red dashed line. The completeness-corrected distributions measured from each of the simulations are plotted as thin grey lines and the median of those recovered distributions is plotted in a thick black line. The thick black error bars are the standard deviation of all of the simulations in each bin and the thin red error bars are poisson uncertainties on the number of detections in each bin scaled by the completeness correction for that bin. <i>Bottom:</i> Same as <i>top</i> panel but calculated using the wKDE technique described in §8.3.3. The shaded red area encompasses the standard deviation of the resulting wKDEs over all 100 simulations. We adopt this fractional uncertainty for the one-dimensional KDE plotted in Figure A.2.	207

CHAPTER 1

INTRODUCTION AND MOTIVATION

One of the most pressing questions in astronomy is the existence and frequency of Earth-like planets that have the potential to harbor life. We have recently determined that exoplanets orbit a majority of stars (Mayor et al., 2011), small exoplanets are particularly abundant (Howard et al., 2010a), and Earth-size planets in the Habitable Zone are prevalent (Petigura et al., 2013a). The *Kepler* mission has been tremendously successful in accomplishing its goal of determining the frequency of planets around Sun-like stars. While the *Kepler* mission has opened the door to the field of exoplanet demographics by giving us detailed knowledge of the frequency of planets of various sizes there are still some major questions that are critical to informing our understanding of our context in the galaxy and the processes that dictate planet formation. The mass of a planet is likely one of the most important properties of a planet in the context of it's formation and evolution. However, we still do not know the frequency of planets as a function of planet mass in high enough detail to inform models of planet formation. Although we recently discovered a gap in the radius distribution of planets which gives us a way to estimate the bulk composition of planet with only a precise radius measurement (Fulton et al., 2017), we still need to measure the underlying the mass function of planets to confirm this interpretation.

1.1 Exoplanet Demographics

The NASA-UC Eta-Earth program was one of the first attempts to measure the frequency of Earth-mass planets in the habitable zones of nearby stars (Howard et al., 2010a). They conducted a 6-year radial velocity survey using Keck/HIRES to monitor 235 nearby F, G, K, and M dwarf stars. They discovered dozens of planets over the course of the survey (Howard et al., 2009, 2011a,c, 2014), but the key result was the discovery that planets with masses between 3 and 10 M_{\oplus} are the most common type of planet orbiting solar type stars in the local neighborhood. They found that these planets with periods less than 50 days are present around $\approx 12\%$ of the stars in their survey, and that these planets are more than an order of magnitude more common than Jupiter-mass planets. However, while this survey was a tremendous success, they could only make a crude measurement of the planet mass function by fitting a power law to large bins in $M \sin i$, they were subject to large and somewhat uncertain completeness corrections in their lowest mass bins, and many of the stars that are the most amenable to the detection of small planets were not exhaustively searched due to observing time constraints. Figure 1.1 shows the cumulative search completeness and planets detected by the Eta-Earth survey.

At the time, a parallel effort was being undertaken for the Southern sky by Michel Mayor's group in Geneva using the HARPS-south spectrograph (Mayor et al., 2011). They also combined the new HARPS data with data from a lower-precision spectrograph that had been operating for much

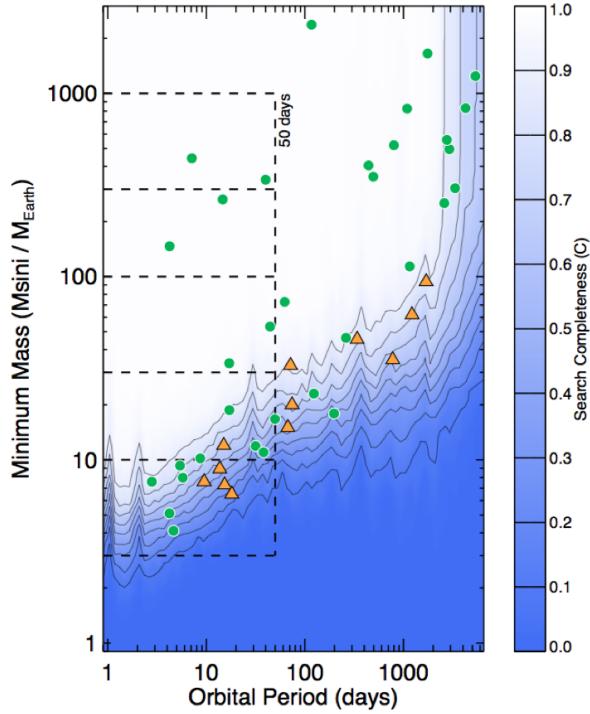


Figure 1.1 Contours of search completeness from the Eta-Earth survey at Keck (Howard et al., 2010a). Planet detections are over plotted as green circles and planet candidates are plotted as yellow triangles.

longer (CORALIE; Baranne et al., 1996) to extend their baseline and sensitivity to massive long-period planets. The final result of this survey was posted to arXiv in 2011 (Mayor et al., 2011), but never published in a refereed journal. In this paper, the authors announced several new planets and performed a statistical analysis of the planets detected in their combined sample. They also found that super-Earth to Neptune mass planets are extremely common, occurring around $\approx 50\%$ of stars out to orbital periods of 10 years (although they were sensitive to these planets only inside ≈ 100 days), and that these planets are much more common than their Jupiter mass counterparts (see Figure 1.2). Our work here builds on these results by locating the low mass planets in the Northern hemisphere and increasing the sensitivity to small planets by collecting many radial velocity (RV) measurements at high cadence.

1.2 Mass-Radius Relation

With a large sample of exoplanets we can look beyond simply trying to measure how many stars have exoplanets, but begin to understand the typical compositions of these planets. We would

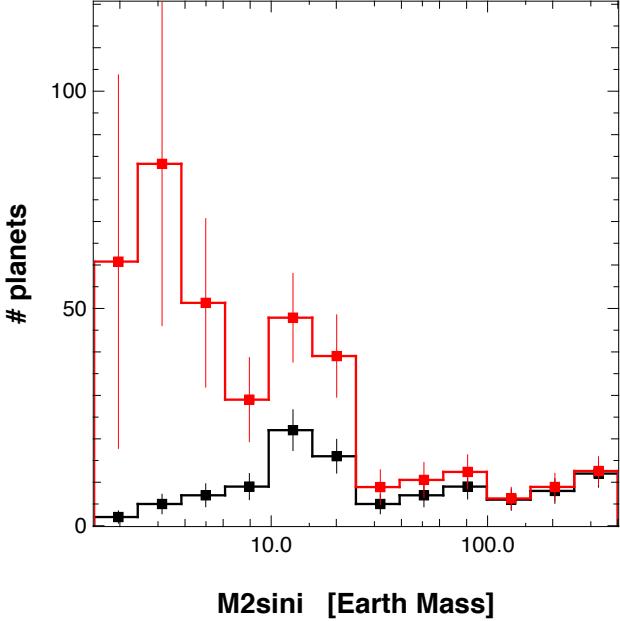


Figure 1.2 Mass distribution of planets detected by the HARPS-south search for low mass planets (Mayor et al., 2011) before (black) and after (red) completeness corrections for all planets out to orbital periods of 10 years. Smaller planets are clearly more common, but the precise location of the sharp rise in occurrence is unclear as well as the details of the apparently bimodal mass distribution for planets below $\approx 20 M_{\oplus}$.

especially like to understand, in a statistical sense, which planets have thick gaseous envelopes, and which planets have solid surfaces. We think that the radii of small planets orbiting relatively close to their host stars can be used to estimate their bulk compositions since the radius of a rocky planet with a hydrogen/helium envelope can be used to distinguish between planets with rocky surfaces and those enshrouded by relatively thick envelopes (Lopez, 2014). Understanding the mass-radius relation for small exoplanets is key to understanding the bulk compositions of super-Earth and Neptune mass planets. Only a few dozen planets smaller than Neptune have measured masses and radii (Marcy et al., 2014) because each mass measurement is observationally very expensive. In this work we constrain the mass-radius relation in a statistical sense by qualitatively comparing our new mass distribution to the new radius distribution that we derive in Chapter 8.

1.3 Planet Formation Theories

The discovery of a large population of super-Earths orbiting close to their host stars was a surprise. Population synthesis models of planet formation had predicted that such systems would be rare (Ida & Lin, 2004; Mordasini et al., 2009a). Planet cores were expected to mostly form beyond the

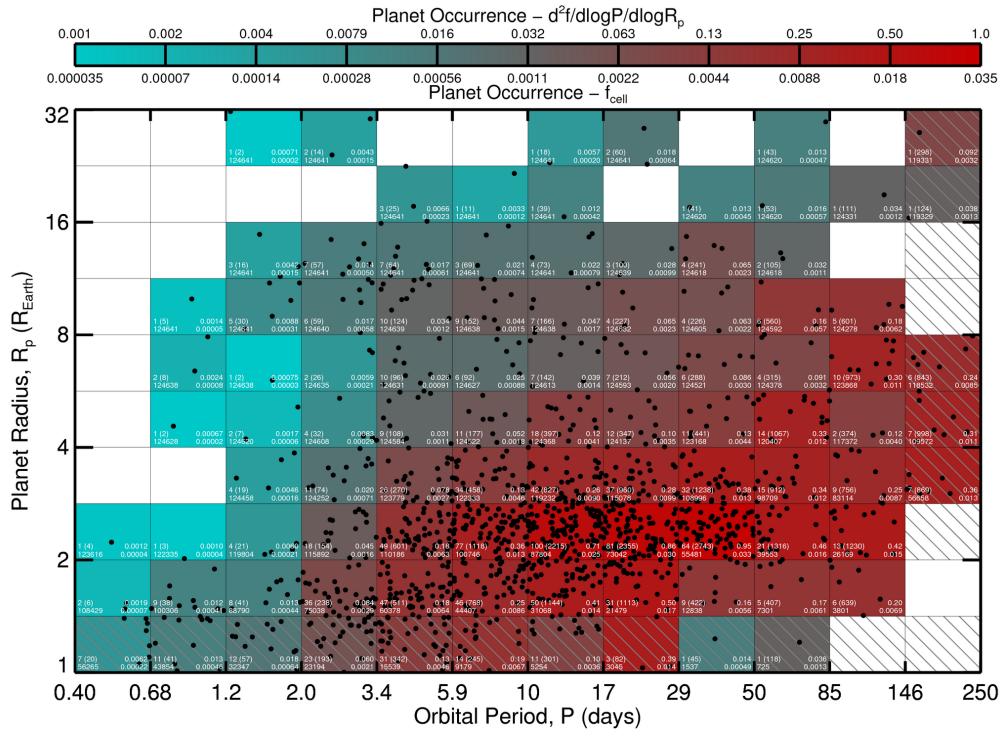


Figure 1.3 2D map of planet occurrence rates derived from *Kepler* planet detections. This plot is similar to that of Howard et al. (2012b), but updated with more recent *Kepler* data containing more planet candidates (Batalha et al., 2013), and extended out to longer orbital periods (250 days vs. 50 days in Howard et al. (2012b)).

ice line and rarely migrate to close orbits unless they first grew to become gas giants. Nevertheless, close-in, low-mass planets are common and often appear in compact multi-planet systems (Lissauer et al., 2011; Fang & Margot, 2012). Theoretical models are catching up, with refinements to the disk migration and multi-planet dynamics in the population synthesis family of models (Ida & Lin, 2010; Alibert et al., 2013; Schlichting, 2014; Lee et al., 2014). A new class of “in situ” formation models have also been proposed in which systems of super-Earths and Neptunes emerge naturally from massive disks (Hansen & Murray, 2012; Chiang & Laughlin, 2013). All planet formation models are currently driven by observations. Since planet masses are dominated by the mass in the core, the planet occurrence rates as a function of planet mass provides strong constraints to the core mass distribution for planet formation models of both flavors.

1.4 The APF-50 Project Components

The APF-50 survey evolved into a project with many various components. We first needed to develop the tools to discover planets and measure precise planet properties in Chapters 2 and 3. We applied those tools to discover several planets orbiting nearby stars using RV data autonomously-collected using the software that we developed for the Automated Planet Finder (APF) telescope in Chapters 4 and 5. We then developed our statistical techniques to measure the frequency of planets for planets and studied the population of planets orbiting white dwarf stars in Chapter 6. We measure the mass function of small planets orbiting bright, nearby stars in Chapter 7. The thesis is culminated in Chapter 8 by the discovery that the radius distribution of small planets in the *Kepler* field is bimodal. Planets smaller than Neptune typically fall into one of two distinct size classes, which has significant implications for a population of planets which we now know is likely to be heavily influenced by photoevaporation (Lopez & Rice, 2016; Owen & Wu, 2017; Jin & Mordasini, 2017). We conclude with some closing remarks in Chapter 9.

CHAPTER 2

THE STELLAR OBLIQUITY AND THE LONG-PERIOD PLANET IN THE HAT-P-17 EXOPLANETARY SYSTEM

This chapter is a reproduction of Fulton et al. (2013) included with permission from AAS journals.

2.1 Introduction

HAT-P-17 is an early K dwarf star that hosts a transiting Saturn-mass planet (planet b) on a 10.3 day orbit and a more massive outer companion (planet c) on a long-period orbit (Howard et al., 2012a, hereafter H12). Transits of planet b were discovered in 2010 by the Hungarian-made Automated Telescope Network (HATNet, Bakos et al., 2004). Followup Keck/HIRES RVs were used to measure the mass of planet b and enabled the discovery of planet c. More than 150 hot Jupiters have been discovered, but it appears that hot Jupiters tend to lack additional short-period giant planet companions (Steffen et al., 2012). HAT-P-17 is one of only six of systems with a transiting Jovian-sized planet and an additional substellar companion. The five other systems include HAT-P-13 (Bakos et al., 2009), HAT-P-31 (Kipping et al., 2011), Kepler-9 (Holman et al., 2010), Kepler-30 (Fabrycky et al., 2012), and KOI-94 (Hirano et al., 2012). These rare multi-planet transiting Jovian systems provide important insight into the formation and evolution of hot Jupiters.

Current theory suggests that Jovian planets form at orbital distances of $\gtrsim 1$ AU where additional protoplanetary solids (ice) augment their formation. They then migrate inwards to become hot Jupiters. Popular theories that attempt to explain their resulting close-in orbits involve a 3rd body (in addition to the Jovian planet and its host star) that perturbs the orbit of the soon-to-be hot Jupiter and excites high eccentricities through either the Kozai mechanism or planet-planet scattering. This highly eccentric orbit then decays through tidal interactions into a close-in circular orbit (Nagasawa et al., 2008; Fabrycky & Tremaine, 2007; Naoz et al., 2011). This scenario would produce hot Jupiters with a large range of orbital obliquities. Others suggest that hot Jupiters migrate within the circumstellar disk from which they formed through interactions with the disk (Lin et al., 1996). In this case we expect that all of the bodies would lie in coplanar orbits that are all well-aligned with the stellar spin axis. If the orbit of planet b is aligned to host star's spin, it would suggest that this system was formed by migration rather than perturbation if the two planets are coplanar. A coplanar and apsidally locked geometry would also allow for a precise measurement of the interior density structure of planet b (Batygin et al., 2009; Mardling, 2010).

An emerging trend suggests that hot Jupiters around cool stars ($T_{\text{eff}} \lesssim 6250$ K) with large convective envelopes tend to be better aligned with their host star's rotation axis (Albrecht et al., 2012b). Tidal energy is most efficiently dissipated by turbulent eddies in the convective regions of

stars (Zahn, 2008). As a result, the rate of tidal dissipation depends on the mass of the convective envelope. Strong tidal interactions with the convective envelope force the system into alignment in a relatively short time. Stars hotter than 6250 K have small or no convective envelopes, and it takes much longer for the system to align (Winn et al., 2010b). HAT-P-17 is a cool star with $T_{\text{eff}} \sim 5200$ K, but planet b's orbital distance is relatively large making tidal interactions weak. According to the tidal figure of merit devised by Albrecht et al. (2012b), we would expect the tidal dissipation rate for this system to be too slow to cause obliquity damping, despite the star's thick convective envelope. This makes HAT-P-17 an interesting test case.

In this work we revisit the orbital parameters of planet c with new Keck/HIRES RV and Keck/NIRC2 adaptive optics images, and present a measurement of the sky-projected orbital obliquity of the star relative to planet b. In §2.2 we discuss our observational techniques. We discuss our RV and RM modeling and results in §2.3, and in §2.4 we interpret and summarize our findings.

2.2 Observations

2.2.1 Keck/HIRES Spectroscopy

Since the publication of Howard et al. (2012a), we have measured the RV of HAT-P-17 ($V=10.54$) for an additional 1.8 years using HIRES (Vogt et al., 1994) on the Keck I telescope. We adopted the same observing strategy and Doppler analysis techniques described in Section 2.3 of H12. In brief, we observed HAT-P-17 through a cell of gaseous iodine and measured the subtle Doppler shifts of the stellar lines with respect to the reference iodine lines using a forward modeling analysis (Butler et al., 1996a).

Our observations were designed to measure the Keplerian orbits of HAT-P-17b and c and also to measure the obliquity of HAT-P-17. For the latter, we observed a transit of HAT-P-17b on UT 26 August 2012. Our observing sequence lasted nearly six hours and bracketed the 3.2 hour long transit. We made 42 observations of \sim 500 second duration separated by 45 second detector reads. To constrain the Keplerian slope, we made three additional observations on the same night approximately 3.8 hours after transit egress.

Julian dates of the photon-weighted exposure mid-times were recorded during the observations, and then later converted to BJD_{TDB} using the tools described in Eastman et al. (2010)¹. The photon-weighted exposure times are only accurate to \sim 1 second due to internal limitations of the exposure meter.

The complete set of RV measurements and their uncertainties are listed in Table 2.2.1. These 100 RVs include 42 RVs from Howard et al. (2012a), 45 new RVs taken on the night of UT 26 August 2012 to measure the RM effect, and 13 additional RVs taken sporadically between 2010 April and 2012 December to measure the orbit of HAT-P-17c.

¹IDL tools for time systems conversion; <http://astroutils.astronomy.ohio-state.edu/time/>

Table 2.1. Radial velocity data^a

Time BJD _{TDB} – 2440000	RV m s ⁻¹	σ_{RV} m s ⁻¹
14396.8272772	-5.25	1.62
14397.7946382	-32.79	1.60
14427.7815123	-10.63	1.58
14429.8199962	-68.92	1.78
14430.8485952	-97.84	1.90
14454.7162454	9.70	2.66
14455.7078194	14.89	1.90

^aThis table is available in its entirety in machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.

2.2.2 KECK/NIRC2 adaptive optics imaging

In order to search for additional companions and sources of possible photometric dilution, we obtained high spatial resolution images of HAT-P-17 using NIRC2 (instrument PI: Keith Matthews) at the Keck II telescope on 2012-05-07 UT. Photometric dilution would affect the radius of HAT-P-17b measured by H12, and the presence of a physically associated companion would put constraints on our radial velocity fit. In addition, a statistical sample of the wide companions to exoplanet host stars may help our understanding of planetary formation mechanisms. Our observations consist of dithered images acquired using the K filter (central wavelength = 2.12 μm). We used the small camera setting to provide fine spatial sampling of the instrument point spread function. The total on-source integration time was 16.2 seconds. Images were processed using standard techniques to flat-field the array, replace hot pixels, subtract the thermal background, and align and co-add individual frames.

2.3 Analysis

2.3.1 Radial velocities

With 1.8 years of new radial velocities we revisited the orbital parameters of the outer companion in the HAT-P-17 system (planet c). We analyzed the entire dataset with a custom version of EXOFAST² (Eastman et al., 2013) ported to Python (ExoPy hereafter). ExoPy utilizes the

²IDL code available at; <http://astroutils.astronomy.ohio-state.edu/exofast/>

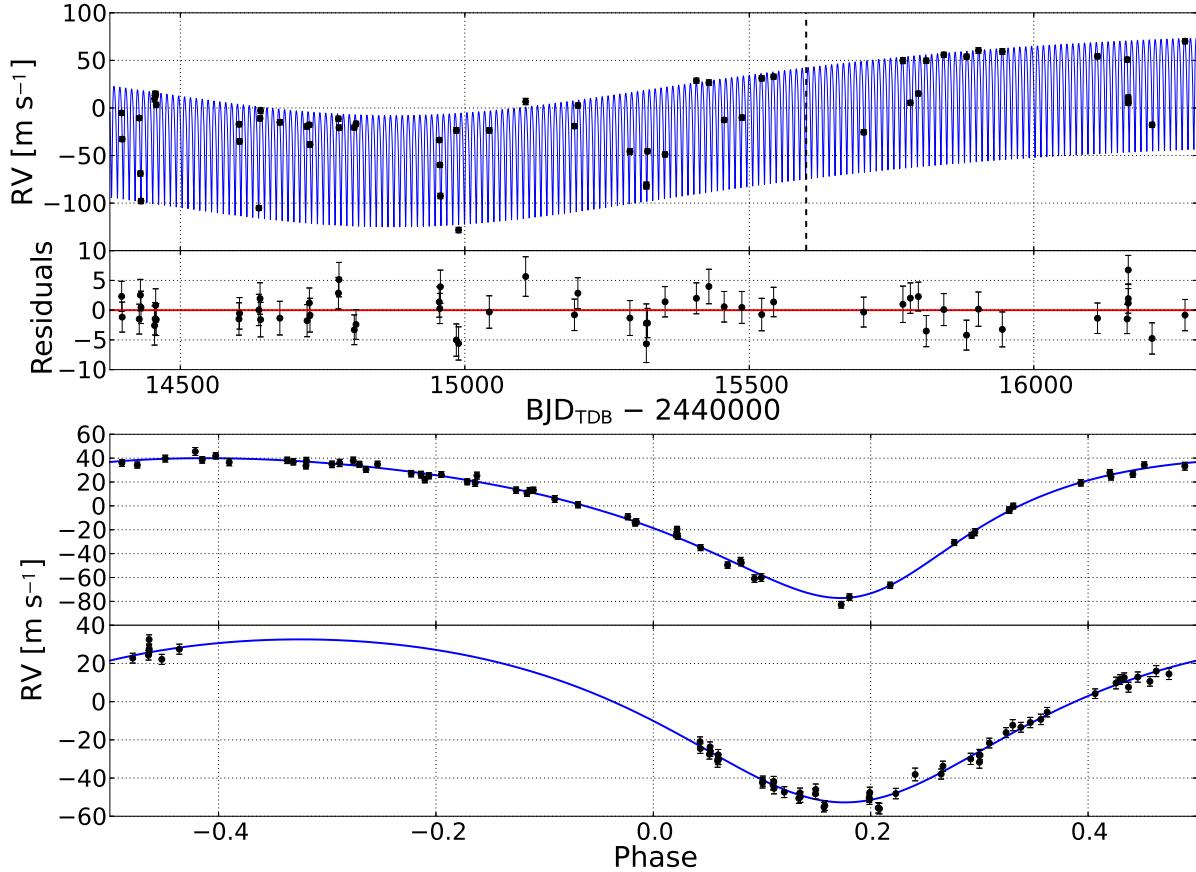


Figure 2.1 *Top:* Keck/HIRES RV measurements for HAT-P-17 as a function of BJD_{TDB} with the best-fitting two-planet model and associated residuals found by χ^2 minimization of the DE-MCMC chains. A stellar “jitter” term (see Table 5.1) has been added in quadrature to the measurement errors. Data taken during the transit of HAT-P-17b for the purpose of measuring the RM effect was excluded from the RV fit and are not included in this plot. Data to the right of the vertical dashed line are new to this work, and data to the left are from H12. *Bottom:* same RV measurements phase-folded to the orbital ephemerides of planets b (upper) and c (lower). Phase 0 corresponds to the time of mid-transit (or hypothetical transit). In each case the orbit of the other planet and an arbitrary center of mass velocity relative to a template spectrum (γ) has been removed.

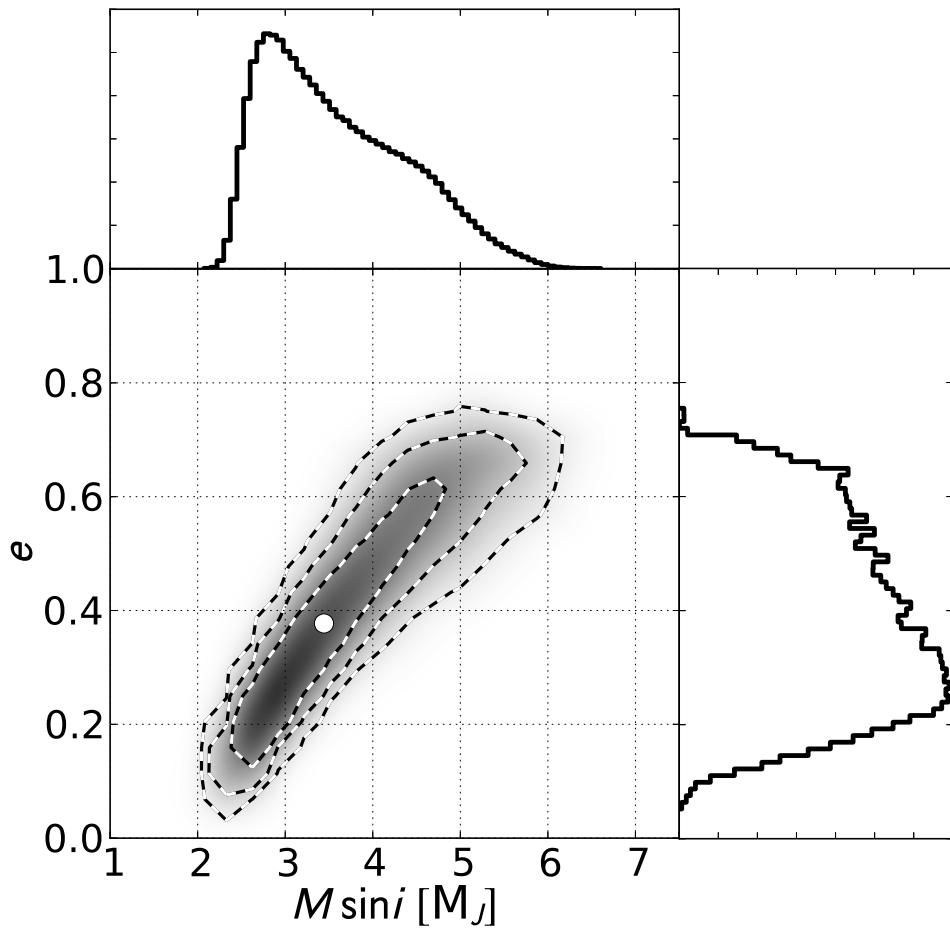


Figure 2.2 Probability distribution of eccentricity vs. minimum mass ($M \sin i$) in Jupiter masses for planet c from the DE-MCMC analysis. The dashed lines are 68%, 95%, and 99% confidence intervals and the white dot is the median value. The median value is offset from the mode (most likely value) due to the asymmetric posterior distributions.

Differential-Evolution Markov Chain Monte Carlo (DE-MCMC) technique (Ter Braak, 2006) to find the best-fitting parameters and their associated uncertainties. We also ported a subset of the RVLIN³ (Wright et al., 2009) package to Python for quick calculation of the Keplerian orbit model. We fit a 12 parameter model to the radial velocity data that included the period (P_b), time of transit ($T_{\text{tra},b}$), eccentricity (e_b), argument of periastron of the star’s orbit (ω_b), radial velocity semi-amplitude (K_b) of planet b, the same parameters for planet c (P_c , $T_{\text{conj},c}$, e_c , ω_c , K_c), the center of mass velocity of the system normalized to an arbitrary reference spectrum (γ), and a stellar “jitter” term.

We computed 24 DE-MCMC chains in parallel, continuously checking for convergence following the prescription of Eastman et al. (2013). We considered the chains well-mixed and halted the DE-MCMC run when the number of independent draws (T_z , as defined in Ford, 2006) was greater than 1000 and the Gelman-Rubin statistic (Gelman et al., 2003; Ford, 2006; Holman et al., 2006) was within 1% of unity for all parameters. In order to speed convergence, ensure that all parameter space was adequately explored, and minimize biases in parameters that physically must be finite and positive, we step in the modified and/or combinations of parameters shown in Table 5.1. Namely, due to the highly correlated uncertainties of e with ω and $v \sin i$ with λ we step in $\sqrt{e} \cos \omega$, $\sqrt{e} \sin \omega$, $\sqrt{v \sin i} \cos \lambda$, and $\sqrt{v \sin i} \sin \lambda$ (Eastman et al., 2013; Albrecht et al., 2012b).

We assigned Gaussian priors to P_b and T_{tra} from the values given in the HAT-P-17b,c discovery paper ($P_b = 10.338523 \pm 9 \times 10^{-6}$ days, $T_{\text{tra},b} = 2454801.16943 \pm 2 \times 10^{-4}$ BJD_{TDB}, H12) that came from the highly constraining photometric transit data, and we assigned uniform priors to all other step parameters. We ignored any variation in transit times due to perturbations caused by planet c, but these are expected to be at least an order of magnitude smaller than the propagated uncertainty on $T_{\text{tra},b}$ (H12). We excluded the spectroscopic transit data from the Keplerian radial velocity fit because the higher density of data on the one night could bias the results in the presence of short-term (∼hours) trends. The best fitting values and upper and lower “1 σ ” errors for each parameter were determined by taking the median, 85.1, and 15.9 percentile values, respectively, of the resulting posterior distributions.

The results of the RV analysis are presented in Table 5.1. All of the parameters for planet b are consistent with the values from H12. However, with the new RV data we can now see that the period of planet c is much longer than the initially reported period of 1620 ± 20 days. Our RV time-series span a total timespan of 1869 days. We still have not seen a complete orbit of planet c, and thus the fit is quite poorly constrained. Figure 2.2 shows the probability distributions of $M_c \sin i_c$ and eccentricity, and indicates that the allowed mass range (3σ) for planet c is $2\text{--}6 M_J$.

We also explored the possibility of a 4th body causing a linear trend ($\dot{\gamma}$) in the radial velocities in addition to the signal from planet c. Fits that included $\dot{\gamma}$ as a free parameter preferred a slope consistent with zero ($\dot{\gamma} = -3.7^{+5.5}_{-8.0} \text{ m s}^{-1}\text{yr}^{-1}$) with a 3σ limit of $|\dot{\gamma}| \leq 19 \text{ m s}^{-1}\text{yr}^{-1}$. To assess

³IDL code available at; <http://exoplanets.org/code/>

the validity of adding one more free parameter to our model we calculate the bayesian information criterion (BIC):

$$\text{BIC} \equiv \chi^2 + k \ln n \quad (2.1)$$

where k is the number of degrees of freedom, and n is the number of data points in the fit (Liddle, 2007). The BIC increased when $\dot{\gamma}$ was a free parameter (104 vs. 100 with or without $\dot{\gamma}$ as a free parameter respectively). The BIC increase, our model fit favoring $\dot{\gamma} = 0.0$, and the AO image (see §3.2.4) all indicate that the data are better described by a model with $\dot{\gamma}$ fixed at zero.

2.3.2 Spectroscopic transit

At first glance the RM data follow the typical redshift then blueshift pattern of a spin-orbit-aligned system. However, the data do not cross zero until slightly after the time of mid-transit. The small asymmetry in the RM curve (Figure 2.3) suggests a slight misalignment. We also used ExoPy to analyze the spectroscopic transit data. Our model of the RM effect takes the form of

$$RM_{\text{net}}(t) = \Delta v(t) + V_{CB}(t) + S(t - T_{\text{tra}}) + v_{CM} \quad (2.2)$$

where $\Delta v(t)$ is given by equation 16 of Hirano et al. (2011) and is discussed in more detail in section 2.3.2 below. $V_{CB}(t)$ is the anomalous radial velocity shift due to the convective blueshift (Shporer & Brown, 2011, discussed in section 2.3.2 below), $\dot{\gamma}_{RM}$ is the radial velocity slope observed during transit due to the orbital motion of HAT-P-17b+c, t are the flux-weighted exposure mid-times of the observations in BJD_{TDB}, T_{tra} is the BJD_{TDB} of mid-transit, and v_{CM} is an arbitrary additive constant velocity. P_b is constrained to the value obtained in the RV analysis, and $T_{\text{tra},b}$ is constrained by propagating the error on P_b and $T_{\text{tra},b}$ found from the Keplerian analysis to the transit epoch of the night of 2012 Aug 26 ($T_{\text{tra},b} = 2456165.8553 \pm 0.0012$ BJD_{TDB}). The amplitude of the HJD_{UTC} to BJD_{TDB} correction applied to the RV data was ~ 67 seconds, or about half of the propagated uncertainty on the mid-transit time which highlights the importance of working in a standardized and consistent time system. The same stellar jitter that contributes to the scatter in the residuals to our Keplerian orbital fit can be seen as systematic trends on shorter timescales, and allowing $\dot{\gamma}_{RM}$ and v_{CM} to be free parameters in the fit prevents these trends from biasing the obliquity measurement. We refer the reader to Albrecht et al. (2012a) for a detailed discussion of the effect of stellar jitter on obliquity measurements via the RM effect.

Semi-analytical Rossiter-McLaughlin model

The shape and amplitude of Δv depends on nine parameters. Five describe the decrease in flux as the planet transits its host star; the planet to star radius ratio (R_p/R_\star), the semi-major axis of the orbit in units of stellar radii (a/R_\star), the inclination of the orbit relative to our line-of-sight (i), and two quadratic limb darkening coefficients (u_1, u_2). We assigned Gaussian priors to P_b , $T_{\text{tra},b}$,

Table 2.2. Radial velocity MCMC results

Parameter	Value	Units
RV Step Parameters:		
$\log(P_b)$	$1.0144585 \pm 3.7e - 07$	$\log(\text{days})$
$T_{\text{tra,b}}$	2454801.1702 ± 0.0003	BJD_{TDB}
$\sqrt{e_b} \cos \omega_b$	-0.5442 ± 0.0052	
$\sqrt{e_b} \sin \omega_b$	-0.214 ± 0.016	
$\log(K_b)$	1.7678 ± 0.0051	m s^{-1}
$\log(P_c)$	3.75 ± 0.38	$\log(\text{days})$
$T_{\text{conj,c}}$	2454146 ± 100	BJD_{TDB}
$\sqrt{e_c} \cos \omega_c$	-0.63 ± 0.15	
$\sqrt{e_c} \sin \omega_c$	-0.017 ± 0.068	
$\log(K_c)$	1.689 ± 0.08	m s^{-1}
γ	20 ± 27	m s^{-1}
$\dot{\gamma}$	$\equiv 0.0$	$\text{m s}^{-1}\text{day}^{-1}$
$\ddot{\gamma}$	$\equiv 0.0$	$\text{m s}^{-1}\text{day}^{-2}$
$\log(\text{jitter})$	0.312 ± 0.076	$\log(\text{m s}^{-1})$
RV Model Parameters:		
P_b	$10.338523 \pm 8.8e - 06$	days
$T_{\text{tra,b}}$	2454801.1702 ± 0.0003	BJD_{TDB}
$T_{\text{peri,b}}$	2454803.24 ± 0.05	BJD_{TDB}
e_b	0.3422 ± 0.0046	
ω_b	201.5 ± 1.6	degrees
K_b	58.58 ± 0.69	m s^{-1}
P_c	5584 ± 7700	days
$T_{\text{conj,c}}$	2454146 ± 100	BJD_{TDB}
$T_{\text{peri,c}}$	2454885 ± 45	BJD_{TDB}
e_c	0.39 ± 0.23	
ω_c	181.5 ± 5.3	degrees
K_c	48.8 ± 9.9	m s^{-1}
γ	20 ± 27	m s^{-1}
$\dot{\gamma}$	$\equiv 0.0$	$\text{m s}^{-1}\text{day}^{-1}$
$\ddot{\gamma}$	$\equiv 0.0$	$\text{m s}^{-1}\text{day}^{-2}$
jitter	2.05 ± 0.39	m s^{-1}
RV Derived Parameters:		
M_b	0.532 ± 0.018	M_J
a_b	0.0882 ± 0.0014	AU
$M_c \sin i_c$	3.4 ± 1.1	M_J
a_c	5.6 ± 3.5	AU

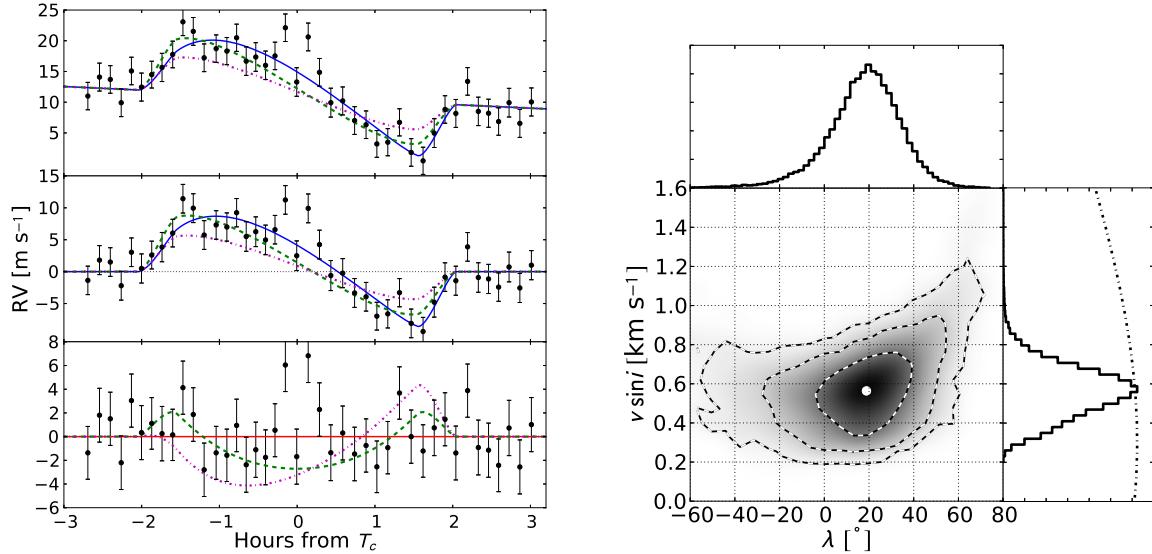


Figure 2.3 *Left:* RV variation during the transit of HAT-P-17b due to the RM effect with best-fit model found from χ^2 minimization overplotted. The upper panel includes the RV variation due to the orbital motion of HAT-P-17b, the middle panel shows the data and model with the orbital motion removed, and the bottom panel shows the residuals to the model. The solid blue line represents our adopted model including all line broadening effects and the convective blueshift, the dashed green line shows the model without the convective blueshift, and the dot-dashed magenta line shows an idealized model in which the line profiles are described only by rotational broadening, similar to the approach of Ohta et al. (2005). Note the asymmetry in the RM curve caused by a combination of the convective blueshift and a slight misalignment. Three data points to the right of the x-axis limit were included in the modeling, but omitted from the plot for clarity. *Right:* Posterior distribution of $v \sin i$ vs. λ from the DE-MCMC analysis of the spectroscopic transit. The dashed lines are 68%, 95%, and 99% confidence intervals and the white dot is the median value of the distribution. The dot-dashed line in the right histogram shows the $0.3 \pm 1.5 \text{ m s}^{-1}$ prior on $v \sin i$ from the SME analysis of H12.

Table 2.3. Rossiter-McLaughlin MCMC results

Parameter	Value	Units
RM Step Parameters:		
$\log(P)$	$1.0144585 \pm 3.8e - 07$	$\log(\text{days})$
T_{tra}	2456165.8551 ± 0.0011	BJD_{TDB}
$\sqrt{e} \cos \omega$	$-0.544^{+0.007}_{-0.0068}$	
$\sqrt{e} \sin \omega$	-0.214 ± 0.015	
$\log(R_p/R_\star)$	-0.9073 ± 0.0032	
$\log(a/R_\star)$	$1.3531^{+0.0082}_{-0.0085}$	
$\cos i$	$0.0123^{+0.0029}_{-0.0032}$	
$u_1 + u_2$	$0.736^{+0.097}_{-0.095}$	
$\sqrt{v \sin i} \cos \lambda$	$0.687^{+0.076}_{-0.1}$	$\sqrt{\text{km s}^{-1}}$
$\sqrt{v \sin i} \sin \lambda$	$0.24^{+0.19}_{-0.2}$	$\sqrt{\text{km s}^{-1}}$
β	$4.1^{+2.6}_{-2.3}$	m s^{-1}
γ_H	$\equiv 0.9$	m s^{-1}
ζ	$\equiv 4.8$	m s^{-1}
v_{CB}	-0.65 ± 0.23	$\log(\text{m s}^{-1})$
v_{CM}	$11.13^{+0.63}_{-0.6}$	m s^{-1}
$\dot{\gamma}_{RM}$	$-17.5^{+4.8}_{-4.9}$	$\text{m s}^{-1}\text{day}^{-1}$
$\log(\text{jitter})$	$0.289^{+0.077}_{-0.083}$	$\log(\text{m s}^{-1})$
RM Model Parameters:		
P	$10.3385231^{+9.1e-06}_{-9.2e-06}$	days
T_{tra}	2456165.8551 ± 0.0011	BJD_{TDB}
e	$0.342^{+0.0046}_{-0.0047}$	
ω	$201.5^{+1.5}_{-1.6}$	degrees
R_p/R_\star	$0.12378^{+0.00092}_{-0.00091}$	
a/R_\star	$22.55^{+0.43}_{-0.44}$	
i	$89.3^{+0.18}_{-0.17}$	degrees
u_1	$0.575^{+0.048}_{-0.047}$	
$v \sin i$	$0.56^{+0.12}_{-0.14}$	m s^{-1}
λ	19^{+14}_{-16}	degrees
β	$4.1^{+2.6}_{-2.3}$	m s^{-1}
γ_H	$\equiv 0.9$	m s^{-1}
ζ	$\equiv 4.8$	m s^{-1}
v_{CB}	-0.65 ± 0.23	m s^{-1}
v_{CM}	$11.13^{+0.63}_{-0.6}$	m s^{-1}
$\dot{\gamma}_{RM}$	$-17.5^{+4.8}_{-4.9}$	$\text{m s}^{-1}\text{day}^{-1}$
jitter	$1.95^{+0.38}_{-0.34}$	m s^{-1}

R_p/R_\star , a/R_\star , and i from the values given in H12 as these are poorly constrained by the RM data alone.

Two more geometrical parameters contribute to the shape of the spectroscopic transit; the rotational velocity of the star projected onto the plane of the sky ($v \sin i_\star$), and the angle between the rotational axis of the star projected onto the plane of the sky and the planet's orbital angular momentum vector (λ). We adopt a value of $v \sin i_\star = 0.3 \pm 1.5 \text{ m s}^{-1}$ as a Gaussian prior that was obtained from the Spectrocospy Made Easy (SME, Valenti & Piskunov, 1996) analysis performed in H12.

Some of the orbital parameters from the RV analysis also have a small effect on the timing and duration of the transit. We assigned Gaussian priors to e_b and ω_b from the results of the RV analysis.

The semi-analytical model of Hirano et al. (2011) also includes three parameters that describe the sources of line broadening (β , γ_H , and ζ). Together with the rotational broadening of the star, these parameters provide a realistic analytical description of the observed line profiles in the spectra. β includes both the Gaussian instrumental line profile and the Gaussian dispersion from micro-turbulence. We adopted a fixed value of 3.0 m s^{-1} for β that is the result of summing in quadrature the width of the HIRES PSF (2.2 m s^{-1}) and 2.0 m s^{-1} micro-turbulence broadening profile (Albrecht et al., 2012b). γ_H is the Lorentzian dispersion of the spectral lines primarily due to pressure broadening. We adopted a fixed value of 0.9 m s^{-1} that was found to be a good match to the HIRES spectra of several stars (Hirano et al., 2011). The most significant of the line profile parameters is the macroturbulence broadening (ζ). We used equation 2.3 from Valenti & Fischer (2005),

$$\zeta = \left(3.98 - \frac{T_{\text{eff}} - 5770 \text{ K}}{650 \text{ K}} \right) \text{ km s}^{-1} \quad (2.3)$$

and $T_{\text{eff}} = 5246 \text{ K}$ (Howard et al., 2012a) to calculate a value of 4.8 m s^{-1} and assigned a conservative Gaussian prior of 3.0 m s^{-1} in the DE-MCMC analysis. We found that changing the prior centers on β , γ_H , and ζ had little effect on the resulting posterior distributions of λ , and for this reason we also could not remove the Gaussian prior on ζ without the DE-MCMC chains wandering into unphysical regions of parameter space.

Convective blueshift

The convective blueshift ($V_{CB}(t)$) is caused by the net convective motion of the stellar photosphere. Hotter material from below the photosphere rises upward towards the observer due to convection and is only partially canceled by downwelling cold material, causing a net blueshift of order 1 m s^{-1} . Since we are only interested in relative radial velocities this net blueshift is unimportant. However, because the convective blueshift is strongest near the center of the star and weaker near the limbs, the transiting planet occults areas of the star that have different contributions to the

net convective blueshift. This causes a time-varying component of the convective blueshift during the spectroscopic transit of order 2 m s^{-1} . We refer the reader to Shporer & Brown (2011) for a more detailed discussion of the convective blueshift, and its influence on the measurement of λ . Since the $v \sin i$ of HAT-P-17 is relatively low, the amplitude of the spectroscopic transit signal is only about 7 m s^{-1} and thus the convective blueshift is a significant effect and must be included in our model. We found that adding the convective blueshift changes the measurement of λ by $\sim 1 \sigma$, pushing it towards zero when the $V_{CB}(t)$ is included.

We used a numerical model based on the work by Shporer & Brown (2011) similar to the approach used by (Albrecht et al., 2012a). We made an initial assumption that the convective blueshift is similar to that of the sun to create a model grid for a range of R_p/R_\star and impact parameters. We interpolated this grid at each step in the DE-MCMC chains. We left the velocity of the photosphere (v_{CB}) as a free parameter to account for the differences between HAT-P-17 and the sun. By definition we expect v_{CB} to be negative and for this reason we rejected models with positive v_{CB} in the DE-MCMC chains. Note the difference between the time-dependent RV signal caused by the convective blueshift ($V_{CB}(t)$) and the fitted scaling factor (v_{CB}).

Results

The results of the RM modeling are presented in Table 2.3.1. Figure 2.3 shows the spectroscopic transit data with the best-fitting model overplotted and the resulting posterior distributions of $v \sin i_\star$ and λ . We measure the sky-projected angle between the orbital angular momentum vector and the stellar rotation axis to be $\lambda = 19^{+14}_{-16}$ degrees. This indicates that planet b's orbit is misaligned with the stellar rotation at a confidence level of only 1.2σ . Our value of $v \sin i = 0.54 \pm 0.15 \text{ m s}^{-1}$ is slightly larger than the value reported in H12 ($v \sin i = 0.3 \pm 0.5 \text{ m s}^{-1}$), but well within the 1σ uncertainty from the SME analysis.

We experimented with fixing $v \sin i$, v_{CB} , and the transit parameters and saw no significant changes in the resulting posterior distribution of λ . When we neglect the convective blueshift in our model we measure a much more significant (presumably artificial) misalignment with $\lambda = 37 \pm 12$ degrees. We also examined the diagnostics from the Doppler analysis of the two outliers on either side of the mid-transit. We found no evidence of systematic errors, poor fits, or other reasons to doubt the integrity of these model outlier points. Removing them from our fit did not change the results other than decreasing the reduced χ^2 .

2.3.3 Additional test for misalignment

We also used the method of Schlaufman (2010) to check for consistency with our RM modeling. This approach compares the measured $v \sin i$ to an empirical estimate of the expected value of $v = 2\pi R_\star/P_{\text{rot}}$, where P_{rot} is the rotation period of the star based on the mass-age-rotation relations established from observations of the Hyades and Praesepe clusters summarized by Irwin & Bouvier

(2009). If the sky-projected inclination of the stellar rotation (i_\star) is close to 90° we would expect the measured $v \sin i$ to closely correspond to v . Since we know that the orbit of planet b is viewed nearly edge on ($i = 89.3^{+0.18}_{-0.17}$ degrees), an observed $v \sin i$ significantly different than v would suggest spin-orbit misalignment. We use

$$P_\star(M_\star, \tau_\star) = P_{\star,0}(M_\star) \left(\frac{\tau_\star}{650 \text{ Myr}} \right)^{1/2} \quad (2.4)$$

from Schlaufman (2010) to calculate the expected rotation period of HAT-P-17 at the age given by H12. In the above equation, $P_{\star,0}(M_\star)$ is the rotation period of the star as a function of mass at an age of 650 Myr (12 days), and $\tau_\star = 7.8 \pm 3.3$ Gyr is the current age of the star. Assuming that the uncertainty in the age of HAT-P-17 is the dominant source of uncertainty, we calculate $P_\star(0.857 M_\odot, 7.8 \pm 3.3 \text{ Gyr}) = 42^{+8}_{-10}$ days. For $i_\star = 90^\circ$ we calculate $v = 1.0^{+0.4}_{-0.2} \text{ m s}^{-1}$.

We compared our measured $v \sin i = 0.56^{+0.12}_{-0.14}$ from the RM analysis to v by calculating the rotation statistic (Θ) from Schlaufman (2010) defined as:

$$\Theta \equiv \frac{v - v \sin i_{\text{obs}}}{\sqrt{\sigma_v^2 + \sigma_{\text{obs}}^2}} \quad (2.5)$$

where $v \sin i_{\text{obs}}$ is measured from the RM analysis, and σ_v and σ_{obs} are the uncertainties on v and $v \sin i$ respectively. The difference between $v \sin i$ and v may suggest that the stellar rotation is inclined with respect to our line of sight. However, the value of $\Theta = 1.9$ (equivalent to 1.9σ) is below the threshold for misalignment as defined by Schlaufman ($\Theta > 2.9$). This threshold accounts for the scatter in the empirical mass-age-rotation calibration which makes our determination of the initial rotation period of HAT-P-17 uncertain. The Schlaufman method provides weak, independent evidence of spin-orbit misalignment that is consistent with our obliquity measurement from the RM analysis. However, these two low significance measurements (each less than 2σ) do not conclusively show that the system is misaligned.

2.3.4 Adaptive optics imaging

We carried out high resolution and high contrast imaging with adaptive optics to check for near-by companions in the context of understanding the architecture of the HAT-P-17 planetary system. Such companions are important in understanding the orbital evolution of the system. We find no evidence for off-axis sources in the immediate vicinity of HAT-P-17. To estimate our sensitivity to faint companions, we calculated the average contrast level achieved as a function of angular separation. Specifically, we compared the peak stellar intensity to the standard deviation (σ) in scattered light within a square box of width 3 FWHM, where FWHM is the PSF full-width at half-maximum (also the size of a speckle). The standard deviation is evaluated at numerous locations and the results are azimuthally averaged to create a contrast radial profile.

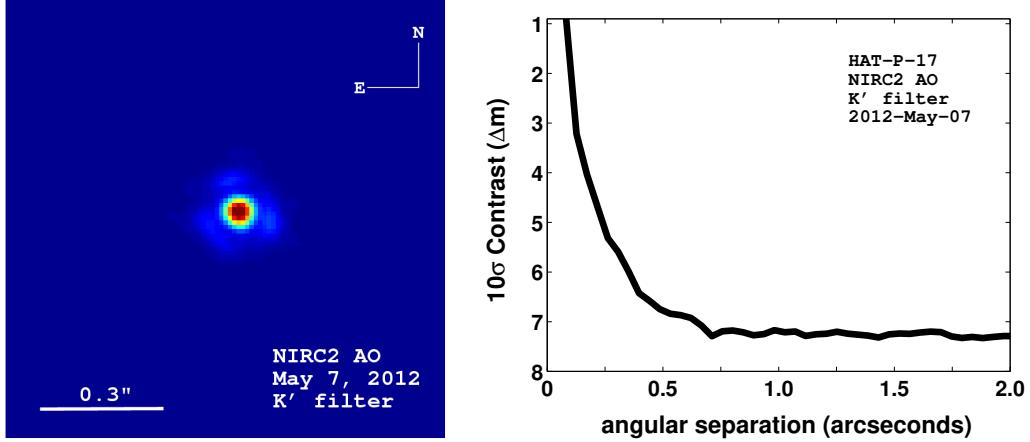


Figure 2.4 *Left*: Keck/NIRC2 adaptive optics image. *Right*: contrast achieved based on the final reduced AO image. Our diffraction-limited observations rule out the presence of companions with $\Delta m > 7$ mags for separations beyond $\approx 0.7''$.

We converted the contrast curve into a minimum detectable mass as a function of projected orbital separation (figure 2.5, left panel) by interpolating the models of Baraffe et al. (2002) at the age and distance of HAT-P-17 from the analysis of H12. Assuming a circular orbit and $M_P \ll M_*$, an order-of-magnitude approximation for the maximum RV slope caused by a fourth body in the system is given by (Winn et al., 2009):

$$\dot{\gamma} \approx \frac{GM_c \sin i_c}{a_c^2}. \quad (2.6)$$

We used this approximation to find the RV slope that would be produced by a planet at the minimum detectable mass as a function of projected orbital separation (figure 2.5, right panel). The minimum detectable mass at large separations is $\sim 80 M_J$ (coincident with the hydrogen-burning limit), far larger than the range of masses that are allowed by our RV fit and does not provide a good constraint on the orbit of planet c or a fourth companion. However, the RV data could still allow for a long-period companion in a nearly face-on orbit or one that is currently near apsis which would minimize the radial velocity slope. The AO data help us rule out these scenarios for stellar/brown dwarf companions outside ~ 50 AU.

2.4 Discussion

HAT-P-17 is a rare planetary system with a transiting hot Jupiter and a long-period companion (HAT-P-17c). We have shown that the orbit of planet c is poorly constrained with the current RV data. We will not be able to conclusively measure the orbital parameters until a significant portion of the orbit has been observed. We find no evidence to suggest the presence of a massive 4th body.

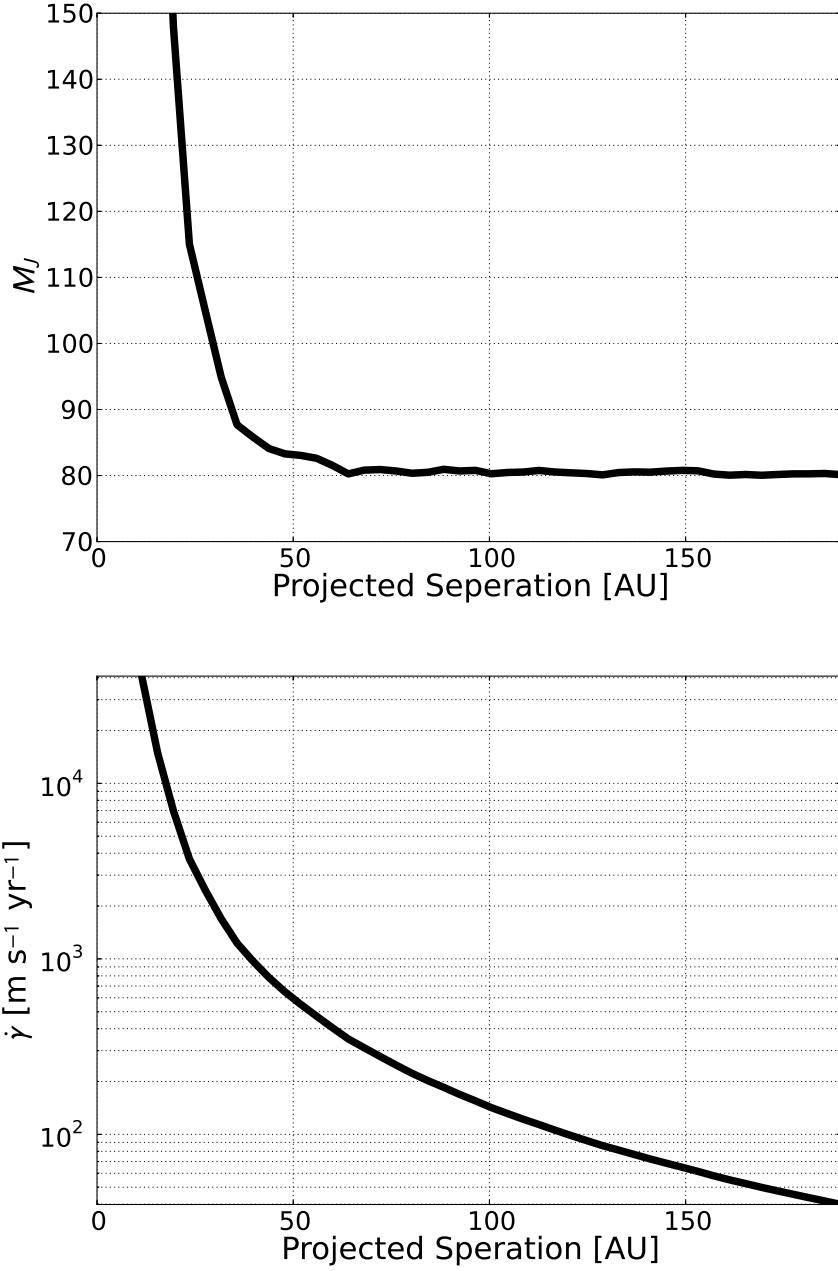


Figure 2.5 *Top*: We convert the measured contrast to a mass-sensitivity curve using the Baraffe et al. 2003 evolutionary models. With an age of 7.8 Gyr, we could have detected any stellar companions ($M > 80M_J$) associated with HAT-P-17 at projected separations beyond ≈ 60 AU. *Bottom*: Same as *top* converted into a predicted RV slope using Equation 2.6. The region below the line is allowed by the data. The 3σ slope constraint ($|\dot{\gamma}| \leq 19 \text{ m s}^{-1} \text{yr}^{-1}$) from the RV analysis is just below the lower y-axis limit.

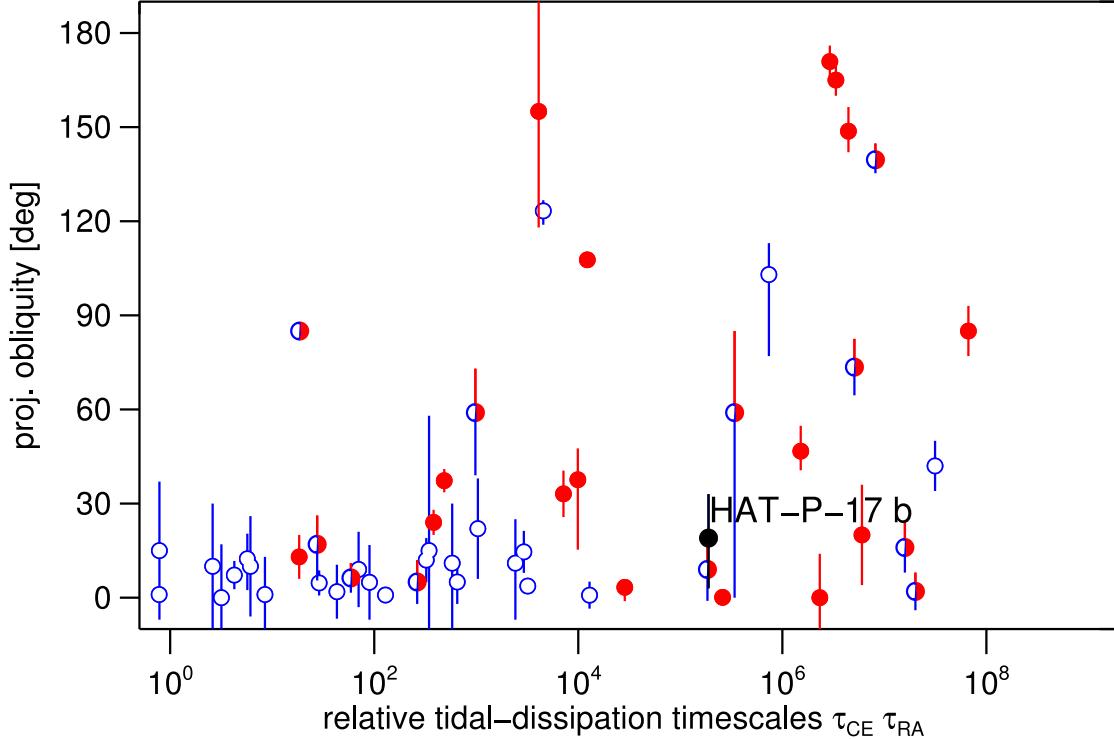


Figure 2.6 *Top*: measured projected obliquity as a function of the alignment timescale calibrated from binary studies (Albrecht et al., 2012b). Stars with temperatures higher than 6250 K are shown with red filled symbols. Blue open symbols show stars with temperatures lower than 6250 K. Stars for which measured effective temperatures include 6250 K in their $1-\sigma$ interval are shown by split symbols. We are computing the relative tidal dissipation rates as a function of stellar type, planet-to-star mass ratio, and orbital distance, using the scaling relations presented by Albrecht et al. (2012). Note that both timescales have been divided by 5×10^9 .

We modeled the RM effect of planet b and measure a possible misalignment of the projected plane of the orbit and the rotation axis of the host star.

Our constraints on a long-term RV trend (in addition to the two planet model) give an upper limit to the mass of a 4th companion of $M_d \sin i_d \left(\frac{a_d}{10 \text{AU}} \right)^2 < 10M_J$ with the assumptions that the potential 4th companion is currently near a time of conjunction in a circular orbit. The lack of companions seen in the adaptive optics image provides complimentary evidence against the presence of a 4th body more massive than $\sim 80M_J$ at separations larger than ~ 50 AU for a wide range of orbital configurations.

Given that the period of planet c reported by H12 was underestimated we do not want to overinterpret any of planet c's parameters. Instead we urge the community to continue observing this interesting system in the coming years. We will be able to assess our measured 16.8 year orbital period in ~ 5 years when the RVs start to decrease rapidly as planet c approaches periaxis, and we will be able to start ruling out the short end of our estimate ($P_c \sim 10$ years) in ~ 3 years.

If it can be shown that the system is coplanar (this would require a spectacular observational effort by searching for transits of planet c) and apsidially locked then HAT-P-17 will be of further interest because it will give us a rare opportunity to probe the interior structure of an exoplanet by measuring the tidal Love number and quality factor through dynamical modeling (Batygin et al., 2009; Mardling, 2010).

We measure the sky-projected angle between the stellar spin axis and orbital angular momentum of the inner planet (stellar obliquity) by modeling the RM effect in Keck/HIRES RV data. The RM analysis suggests a slight spin-orbit misalignment of planet b with $\sim 1.2 \sigma$ confidence ($\lambda = 19^{+14}_{-16}$ degrees). The Schlaufman method provides additional evidence for spin-orbit misalignment, but due to the dependence on somewhat uncertain stellar evolution models and the unknown initial angular momentum of HAT-P-17 we believe that the result from our detailed RM modeling to be more robust. However, the two low-confidence measurements do not allow us to distinguish between a well-aligned system or one with a small, but non-zero, spin-orbit misalignment.

Winn et al. (2010a) noted an emerging trend in the orbital obliquities of hot Jupiter hosts cooler than 6250 K being generally well-aligned, and hot Jupiters around stars hotter than 6250 K having a wide distribution of λ . HAT-P-17 is a cool star around which we would normally expect to find spin-orbit-aligned hot Jupiters, but due to the relatively wide orbit of HAT-P-17b we do not expect that the planet would have had time to align itself with the spin of the host star (Winn et al., 2010a; Albrecht et al., 2012b) if it were perturbed into a misaligned orbit. This is consistent with our findings of a marginally non-zero λ , however a more significant spin-orbit misalignment would provide stronger evidence against the disk-migration formation scenario. Figure 2.6 compares the timescales for alignment of this system with the systems studied in Albrecht et al. (2012b). HAT-P-17b lies in a region of Figure 2.6 that shows large scatter in λ due to the longer realignment timescales. We expect that the planets in this region of the plot retain their spin-orbit angle from the time shortly after their migration because the tidal interactions are too small to force a realignment over the age of the star. If misaligned, this is the first multi-planet system in which a spin-orbit misalignment has been measured. We also note that our measurement of λ is only slightly less likely to be consistent with zero, and coplanarity of planet c's orbit would provide strong evidence that the system migrated quiescently in the protoplanetary disk.

For a low-amplitude RM system like HAT-P-17, we find that the convective blueshift is an important effect that must be included in our model for an accurate measurement of λ . With 2-3 additional measurements of the RM effect we should be able to conclusively ($\sim 3 \sigma$) determine if the system is misaligned which will help us understand the formation of the HAT-P-17 system and other similar systems.

CHAPTER 3

KELT-8b: A HIGHLY INFLATED TRANSITING HOT JUPITER AND A NEW TECHNIQUE FOR EXTRACTING HIGH-PRECISION RADIAL VELOCITIES FROM NOISY SPECTRA

This chapter is a reproduction of Fulton et al. (2015a) included with permission from AAS journals.

3.1 Introduction

Ground-based surveys for transiting exoplanets have been extraordinarily successful in detecting Jupiter-size planets orbiting very close to their host stars (hot Jupiters). These planets are some of the easiest to detect and characterize but also some of the most intrinsically rare, occurring around only 0.3 to 1.5% of Sun-like stars (Gould et al., 2006; Cumming et al., 2008; Bayliss & Sackett, 2011; Wright et al., 2012; Howard et al., 2012b). Both radial velocity (RV) surveys and transit surveys are heavily biased to detect these massive close-in objects yet we still know of only 162 examples¹ of hot Jupiters out of the thousands of currently known exoplanets.

Due to their short orbital periods and large transit and/or radial velocity (RV) signals, these planets make excellent laboratories to study planet formation and migration theories (Hansen & Murray, 2012; Ida & Lin, 2008; Mordasini et al., 2009a), atmospheric properties and composition (Zhao et al., 2014; Lockwood et al., 2014; Sing et al., 2015), and even the rotational velocities of giant planets (Snellen et al., 2014). Several mysteries remain unsolved for this population of exoplanets including the differences between the orbital distance distribution of close-in Jupiters orbiting metal rich and metal poor stars (Dawson & Murray-Clay, 2013), the frequency of long-period companions in hot Jupiter systems (Knutson et al., 2014; Ngo et al., 2015), the source of large misalignments between the orbital axis and stellar spin axis (Albrecht et al., 2012b; Dawson, 2014), and the reason that many hot Jupiters are inflated to extremely large radii (Batygin et al., 2011; Ginzburg & Sari, 2015). Each hot Jupiter discovery enhances our ability to explore these phenomena and study individual systems in exquisite detail.

RV surveys have discovered many non-transiting hot Jupiters around very bright stars, while the target stars of transit surveys tend to be much fainter. This is primarily due to the fact that only a narrow range of orbital parameters cause a planet to transit and many tens of thousands of stars must be observed in order to detect a single transiting planet. Since bright stars are distributed across a large area of the sky it is generally more efficient to observe small areas of the sky to a greater depth for transits of fainter stars. The Kilodegree Extremely Little Telescope survey

¹Planets with orbital periods shorter than 10 days and $M \sin i \geq 0.5 M_J$ based on a 2015 Mar 3 query of exoplanets.org

(KELT, Pepper et al., 2007) serves to bridge the gap in host star brightness between RV-detected and transit-detected planets by using extremely small aperture telescopes with large fields of view to observe nearly the entire sky. The survey is optimized to detect planets orbiting stars with V-band magnitudes between 8 and 10 but it has also been successful in detecting planets around slightly fainter stars as well (Collins et al., 2014).

We report the discovery of a highly inflated ($R_p = 1.86^{+0.18}_{-0.16} R_J$, $M_p = 0.867^{+0.065}_{-0.061} M_J$) hot Jupiter orbiting the moderately bright $V=10.84$ G dwarf; HD 343246 (KELT-8 hereafter). KELT-8b has the 2nd largest radius and 7th lowest density among all transiting exoplanets, with only WASP-17b being larger (Southworth et al., 2012; Bento et al., 2014). The planet lies well above the Seager et al. (2007) mass-radius relation for a pure hydrogen composition. In order to facilitate rapid and efficient RV confirmation of the planet we developed a new technique that saves significant telescope resources when collecting RV measurements using the iodine technique (Butler et al., 1996b). In §3.2 we describe our discovery and follow-up observations. We describe our spectroscopic analysis of the host star and summarize the inferred stellar properties in §5.3. A close companion stellar object that was ultimately found to be a background contaminant is described in §3.3.4. We outline our new technique for RV extraction in §3.4. A global analysis of the photometric and RV data is presented in §5.4.2, and we conclude with a discussion in §8.5 and summary in §3.7.

3.2 Observations

3.2.1 KELT Photometry

The KELT-North survey telescope consists of an Apogee AP16E imager (4K x 4K CCD with $9\mu\text{m}$ pixels) and an 80mm Mamiya 645 camera lens (42 mm diameter, f/1.9) behind a Kodak Wratten #8 red-pass filter. This setup achieves a $26^\circ \times 26^\circ$ field of view with roughly $23'' \text{ pixel}^{-1}$. For a more complete description of hardware and operations, see Pepper et al. (2007).

Raw KELT science images are dark-subtracted and flat-fielded using standard methods and then reduced using a heavily-modified implementation of the ISIS image subtraction package (Alard & Lupton, 1998; Alard, 2000) coupled with our own high-performance background-removal routines. Stars are identified for extraction using the standalone DAOPHOT II PSF-fitting software package (Stetson, 1987, 1990). To reduce systematic errors, extracted light curves are processed with the trend filtering algorithm (TFA; Kovács et al., 2005) prior to period search. A more complete description of the data reduction pipeline and candidate selection process are available in Siverd et al. (2012).

KELT-8 resides in KELT-North field 11, centered on ($\alpha = 19^h 26^m 48^s$, $\delta = +31^\circ 39' 56''$; J2000). The data set consists of 5978 images acquired between 30 May, 2007 and 14 June, 2013. The full discovery light curve, phased to the KELT-8b transit ephemeris, is shown in Figure 3.1.

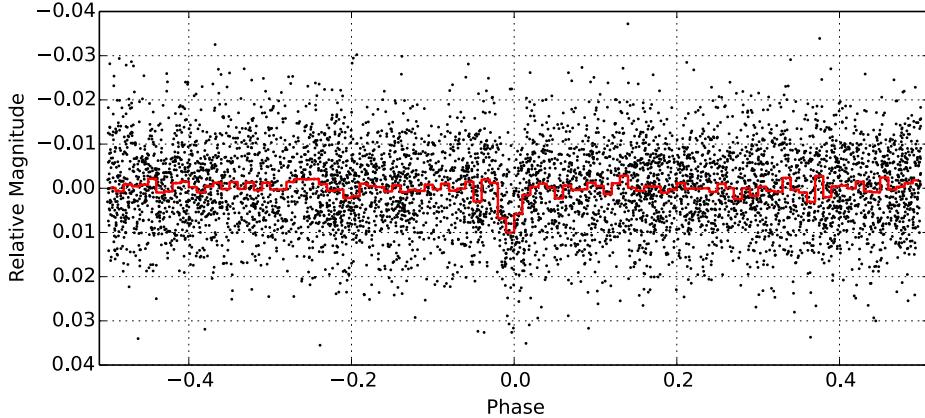


Figure 3.1 Discovery light curve of KELT-8b from the KELT-North telescope. The photometry is phase-folded to the ephemeris listed in Table 4.3. The red line is the same data binned in intervals of width 0.01 in orbital phase (≈ 46.7 minutes).

3.2.2 Follow-up Photometry

We acquired follow-up time-series photometry of KELT-8 to better determine the transit shape and to check for a consistent transit depth across the optical filter bands. To schedule follow-up photometry, we used the `Tapir` software package (Jensen, 2013). We obtained full primary transits in multiple bands between July 2014 and September 2014. Figure 3.2 shows all the primary transit follow-up light curves assembled. A summary of the follow-up photometric observations is shown in Table 3.1. We find consistent R_p/R_\star ratios in the g' , r' , i' , V , and I_C filters after compensating for blending with a nearby neighbor in some observations (see below), helping to rule out false positives due to blended eclipsing binaries. The lower panel of Figure 3.2 shows all primary transit follow-up light curves from Figure 3.2 combined and binned in 5 minute intervals. This combined and binned light curve is not used for analysis, but rather to show the overall statistical power of the follow-up photometry.

`AstroImageJ`² (AIJ) (K. A. Collins & J. F. Kielkopf 2015, in preparation) was used to calculate the differential photometric data from all calibrated image sequences. AIJ is a general purpose image processing package, but is optimized for processing time-series astronomical image sequences.

MORC

We observed one complete KELT-8b transit on UT 2014-08-04 from the 0.6 m Moore Observatory RCOS (MORC) telescope, operated by the University of Louisville. MORC is equipped with an Apogee U16M $4\text{K} \times 4\text{K}$ CCD giving a $26' \times 26'$ field of view and $0.39 \text{ arcseconds pixel}^{-1}$. The transit was observed pseudo-simultaneously in two filters by alternating between the g' and i' filters

²<http://www.astro.louisville.edu/software/astroimagej/>

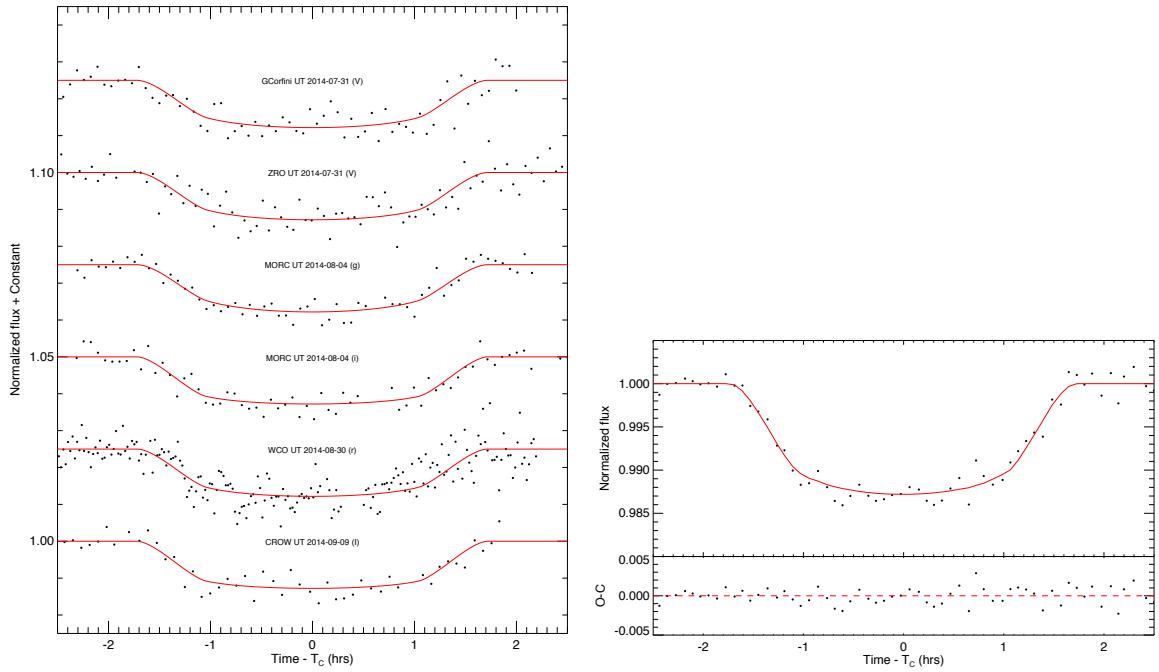


Figure 3.2 *Top:* Follow-up photometry of KELT-8b primary transits. The source, date, and filter for each transit is annotated. The best-fit models are shown in red. *Bottom:* All follow-up transit light curves combined and binned in 5 minute intervals. This light curve is not used in the analysis and is shown simply to demonstrate the combined photometric precision of the follow-up light-curves.

from one exposure to the next. The filter change is made during CCD read-out to minimize loss of time on the the sky. Since KELT-8b has a nearby neighbor, telescope defocusing was minimal to provide better separation of the flux from the two stars on the detector.

GCorfini

We observed one full transit on UT 2014-07-31 from Giorgio Corfini’s home observatory (GCorfini) in Lucca, Italy. The observations were obtained in V-band using a 0.2 m Newtonian telescope and a SBIG STT-6303 ME CCD 1536×1024 pixel camera, giving a $59' \times 39'$ field of view. No defocusing was applied, but the target and neighbor star were severely blended on the detector due to seeing. A photometric aperture was chosen that included both KELT-8 and its neighbor. The resulting light curve was corrected for blending as described in §5.4.2.

ZRO

We observed one full transit in the V_C filter on UT 2014-07-31 at Roberto Zambelli’s personal observatory (ZRO). The telescope was not defocused and both stars were cleanly resolved from each other. The observations were obtained using a Meade LX 200 ACF GPS 12” telescope, a f/6.3 focal reducer, and an SBIG ST8 XME 1530×1020 CCD with a pixel scale of 0.92 arc seconds pixel $^{-1}$ and a $23' \times 15'$ field of view.

WCO

We observed a full transit in the r' band at Westminster College Observatory (WCO) in Pennsylvania on UT 2014-08-30. The observations were obtained using a Celestron 0.35 m C14 telescope with an SBIG STL-6303E 3072×2048 CCD, giving a $24' \times 16'$ field of view and 0.5 arcseconds pixel $^{-1}$. The data were collected with the telescope in focus and the two stars were resolved from each other.

CROW

We observed one full transit in the I_C band at Canela’s Robotic Observatory (CROW) in Portugal on UT 2014-09-09. The observations were obtained using a 0.3 m LX200 telescope with an SBIG ST-8XME 1530×1020 CCD, giving a $28' \times 19'$ field of view and 1.11 arcseconds pixel $^{-1}$. No telescope defocus was applied, but KELT-8 and the neighbor were blended on the detector due to seeing. A photometric aperture was selected to encircle the flux from both stars. The resulting light curve was corrected for blending as described in §5.4.2.

Table 3.1. Summary of Photometric Observations

Telescope	UT (2014)	# Obs	Filter Band	Cycle ^a (sec)	RMS ^b (10^{-3})	PNR ^c ($\frac{10^{-3}}{\text{min}}$)	α^d
GCorfini	7-31	73	V	243	3.0	6.0	1.1
ZRO	7-31	110	V	216	3.5	6.6	1.9
MORC	8-04	65	g'	127	2.1	3.0	2.2
MORC	8-04	62	i'	127	2.3	3.3	2.6
WCO	8-30	209	r'	81	3.8	4.4	3.1
CROW	9-09	49	I_C	332	2.3	5.4	1.3

^aCycle time in seconds, calculated as the mean of exposure time plus dead time during periods of back-to-back exposures. The MORC g' and i' exposures were alternating, so cycle time is calculated for exposures in both filters combined.

^bRMS of residuals from the best fit model in units of 10^{-3} .

^cPhotometric noise rate in units of $10^{-3} \text{ minute}^{-1}$, calculated as $\text{RMS}/\sqrt{\Gamma}$, where RMS is the scatter in the light curve residuals and Γ is the mean number of cycles (exposure time and dead time) per minute during periods of back-to-back exposures (adapted from Fulton et al. 2011).

^dScaling factor applied to the uncertainties to ensure that the best fitting model has $\chi^2=1$ (see §5.4.2).

3.2.3 High-resolution Spectroscopy

We collected a total of 13 RV measurements of KELT-8 using the Levy high-resolution optical spectrograph mounted at the Nasmyth foci of the Automated Planet Finder Telescope (APF) at Lick observatory (Vogt et al., 2014a; Radovan et al., 2014; Burt et al., 2014). The measurements were collected between UT 2014 August 15 and UT 2014 November 9 and are presented in Table 5.1. We observed this star using the $2 \times 8''$ slit for a spectral resolution of $R \approx 80,000$. We pass the starlight through a cell of gaseous iodine which serves as a simultaneous wavelength and point spread function (PSF) reference (Marcy & Butler, 1992). Relative radial velocities are calculated by tracing the doppler shift of the stellar spectrum with respect to the dense forest of iodine lines using a forward modeling technique described in Butler et al. (1996b). Traditionally, a high signal-to-noise iodine-free observation of the same star is deconvolved with the instrumental PSF and used as the stellar template in the forward modeling process. However, in this case the star is too faint to collect the signal-to-noise needed for reliable deconvolution in a reasonable amount of time on the APF. Instead we simulate this observation by using the SpecMatch software (Petigura et al. 2015 in prep.) to construct a synthetic template from the Coelho (2014) models and best fit stellar parameters. This process is described in more detail in Section 3.4.

Table 3.2. Radial Velocities of KELT-8

BJD_{TDB} (- 2440000)	RV (m s^{-1})	σ_{RV} (m s^{-1})	BS (m s^{-1})	σ_{BS} (m s^{-1})
16884.7240398	-52.700	13.698	18.194	14.832
16886.7223760	16.089	11.487	60.348	22.517
16888.6894588	54.770	11.598	21.460	21.508
16890.6920118	-85.923	11.536	104.118	56.715
16891.6849308	39.139	11.894	-18.528	26.524
16892.6862628	120.081	10.732	15.611	12.552
16894.6776168	-30.453	11.729	-18.657	35.822
16895.6940738	98.475	11.258	18.095	18.990
16900.7203146	-99.117	11.290	-3.809	25.208
16903.7467735	-104.461	12.930	67.604	41.392
16905.6816215	91.985	13.023	6.130	15.736
16906.7226164	-90.377	12.995	117.227	77.109
16970.6342444	55.520	20.228	73.214	34.466

3.2.4 Adaptive Optics Imaging

We acquired visible-light adaptive optics images of KELT-8 using the Robo-AO system (Baranec et al., 2013, 2014) on the 60-inch Telescope at Palomar Observatory. On UT 2015 March 8, we observed KELT-8 in the Sloan- i' filter as a sequence of full-frame-transfer detector readouts at the maximum rate of 8.6 Hz for a total of 120 s of integration time. The individual images are corrected for detector bias and flat-fielding effects before being combined using post-facto shift-and-add processing using KELT-8 as the tip-tilt star with 100% frame selection to synthesize a long-exposure image (Law et al., 2014a).

We calculate the 5σ contrast limit as a function of angular separation by first determining the background noise level for concentric rings of width equal to the FWHM moving outward from the primary star. A simulated companion (a dimmed cutout of the primary star) is inserted into the closest ring, with a random PA. An auto-companion detection code then searches for the simulated companion. If it is found, the companion is dimmed further and reinserted. This is continued until the companion is not found with a confidence greater than 5σ . We repeat this process for each annulus, and fit the sparse measurements with a function of the form $a - b/(r - c)$, where a , b , and c are free parameters in the fit and r is the radius from the target star. We convert contrast limits to mass limits using the models of Baraffe et al. (2002). The resulting contrast and mass limits are presented in Figure 3.3

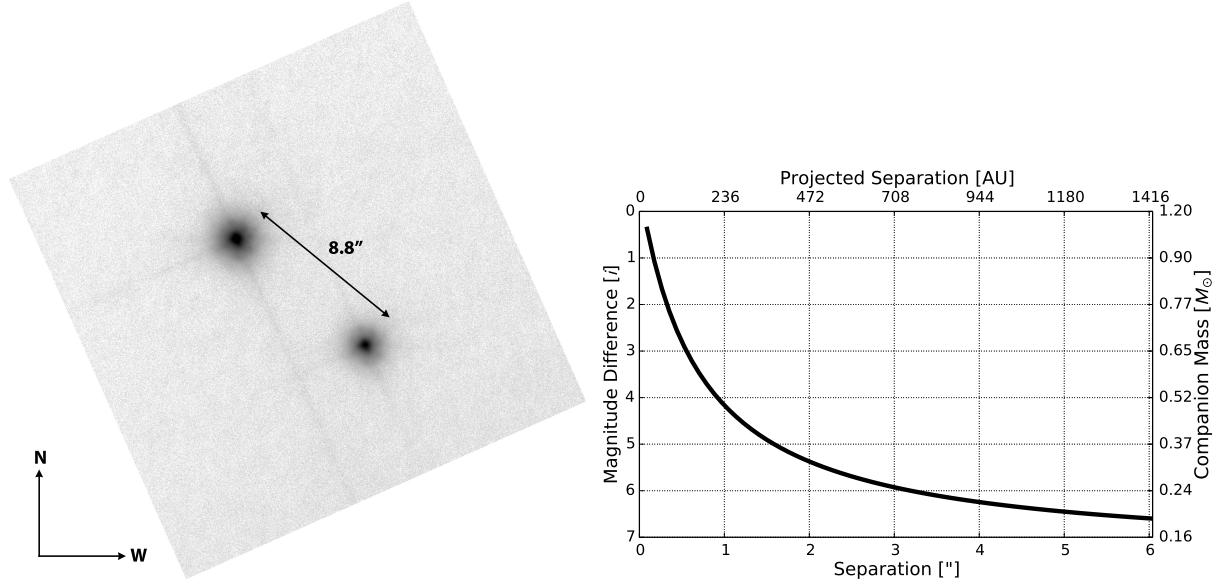


Figure 3.3 *Left:* Robo-AO image of KELT-8 showing the background giant star 8''.8 to the SW of KELT-8. *Right:* Five σ contrast limit near KELT-8 derived from the Robo-AO image.

3.3 Stellar Properties

3.3.1 Keck/HIRES Spectroscopy

In order to obtain precise values for the stellar parameters we collected a moderate signal-to-noise iodine-free observation using the HIRES spectrograph on the Keck I telescope (Vogt et al., 1994). We measured the effective temperature (T_{eff}), surface gravity ($\log g$), iron abundance ([Fe/H]), and rotational velocity of the star ($v \sin i$) using the tools available in the SpecMatch software package (Petigura et al. 2015 in prep). We first correct the observed wavelengths to be in the observer's rest frame by cross-correlating a solar model with the observed spectrum. Then we fit for T_{eff} , $\log g$, [Fe/H], $v \sin i$, and the instrumental PSF using the underlying differential-evolution Markov Chain Monte Carlo (DE-MCMC, Ter Braak, 2006) machinery of ExoPy (Fulton et al., 2013). At each point in the MCMC chains a synthetic spectrum is created by interpolating the Coelho (2014) grid of stellar models for a set of T_{eff} , $\log g$, [Fe/H] values and solar alpha abundance. We convolve this synthetic spectrum with a rotational plus mactroturbulence broadening kernel using the prescriptions of Valenti & Fischer (2005) and Hirano et al. (2011). Finally, we perform another convolution with a gaussian kernel to account for the instrumental PSF and the synthetic spectrum is compared with the observed spectrum using χ^2 to assess the goodness of fit. The priors are uniform in T_{eff} , $\log g$, and [Fe/H], but we assign a gaussian prior to the instrumental PSF that encompasses the typical variability in the PSF width caused by seeing changes and guiding errors. Five echelle orders of the spectrum are fit separately and the resulting posterior distributions are combined before taking the median values for each parameter.

Parameter uncertainties are estimated as the scatter in spectroscopic parameters given by SpecMatch relative to the values in Valenti & Fischer (2005) for 352 of the stars in their sample and 76 stars in the Huber et al. (2012) asteroseismic sample. Systematic trends in SpecMatch values as a function of T_{eff} , $\log g$, and [Fe/H] relative to these benchmark samples are fit for and removed in the final quoted parameter values. More details about the benchmark comparison and calibrations will be available in Petigura et al. (2015, in prep). These spectroscopic parameters are then used as input priors for the global fit discussed in §5.4.2. The best fit spectroscopically-determined stellar parameters are $\log g = 4.23 \pm 0.08$, $[\text{Fe}/\text{H}] = 0.272 \pm 0.038$, and $T_{\text{eff}} = 5754^{+54}_{-55}$ K.

The $\log g$ derived from spectroscopy is only marginally consistent with the $\log g$ derived from the global fit ($4.078^{+0.049}_{-0.054}$). The $\log g$ value measured from the transit light curves is likely more reliable because $\log g$ can be a difficult quantity to measure from high-resolution spectroscopy alone. We perform a global analysis both with and without the spectroscopic $\log g$ as a prior (see §5.4.2).

We also collected a single exposure with the iodine cell in the light path with a slightly higher S/N ratio to establish a long term baseline for possible followup of this target with Keck in the future. We used this observation to extract the $\log R'_{\text{HK}}$ stellar activity metric. We find that $\log R'_{\text{HK}} = -5.108$ which indicates that the star very chromospherically quiet.

3.3.2 Spectral Energy Distribution

We estimate the distance and reddening to KELT-8 by fitting the B-band through WISE W4-band spectral energy distribution (SED) to the Kurucz (1979) stellar atmosphere models. We also found a GALEX NUV measurement in the literature (Bianchi et al., 2011), however this measurement shows a large excess in flux relative to the SED compatible with all other measurements. Since the PSF of GALEX in the NUV band is $4''.9$ we suspect the GALEX measurement is contaminated by the visual companion $8''.8$ away and we omit this measurement from the fit. The WISE measurements also have a large PSF relative to the separation between the two components, but we do not find any evidence for significant contamination. We fix T_{eff} , $\log g$, and [Fe/H] to the best fit values in Table 4.3 and leave distance (d) and reddening (A_V) as free parameters. We find best-fit parameters of $A_V = 0.15 \pm 0.06$ magnitudes and $d = 236 \pm 9$ pc with a reduced $\chi^2 = 3.7$ (considering only the flux uncertainty and not the uncertainty in the model). This agrees quite well with the distance of 233 pc estimated by Pickles & Depagne (2010).

3.3.3 Evolutionary State

We estimate the age of KELT-8 by fitting Yonsei-Yale isochrones to the values of T_{eff} , $\log g$, and [Fe/H] given in Table 4.3. We fix the stellar mass to the value of $M_* = 1.211 M_\odot$ listed in Table 4.3. Our best fit stellar parameters indicate that the star happens to fall on a very rapid part of the evolutionary track, the so-called “blue hook” just prior to crossing the Hertzsprung gap (see Figure 3.5). In principle, this would allow for a very precise measurement of the stellar age, however, the

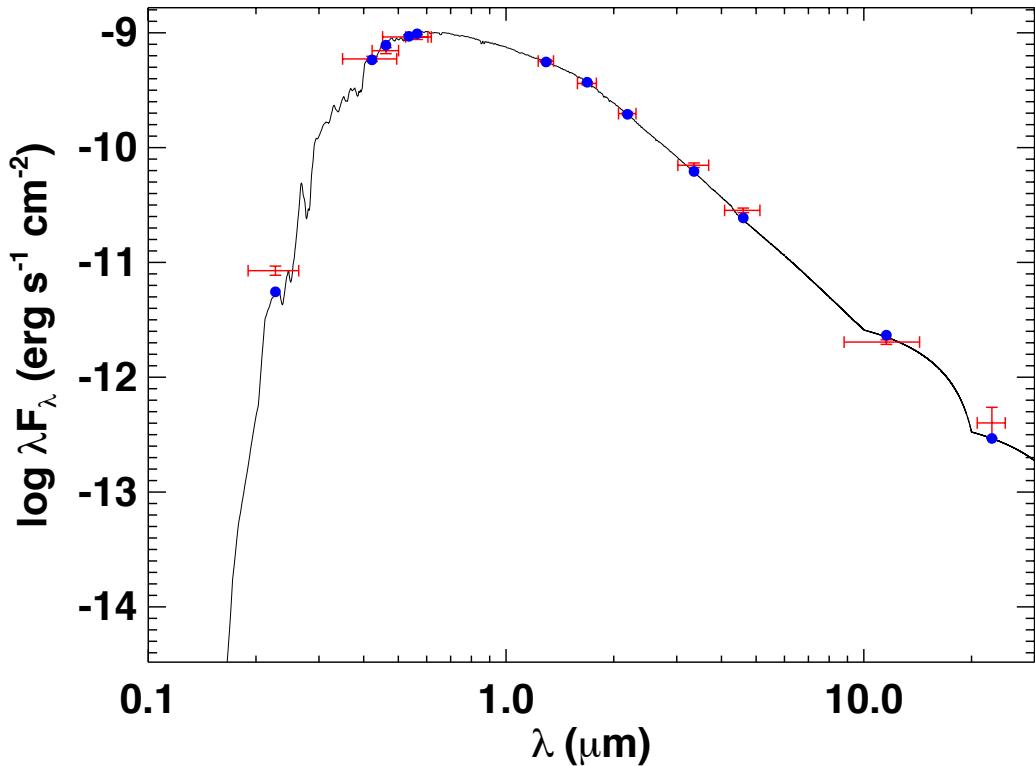


Figure 3.4 SED fit to KELT-8. The red crosses are the photometry measurements. The vertical errors are the 1σ photometric uncertainties, and the horizontal error bars are the effective widths of the passbands. The leftmost point (GALEX NUV) was omitted from the fit due to possible contamination from the neighboring star at $8''8$ separation. The black line is the best-fit stellar atmosphere model and the blue points are the predicted passband-integrated fluxes of the model corresponding to our observed photometric bands.

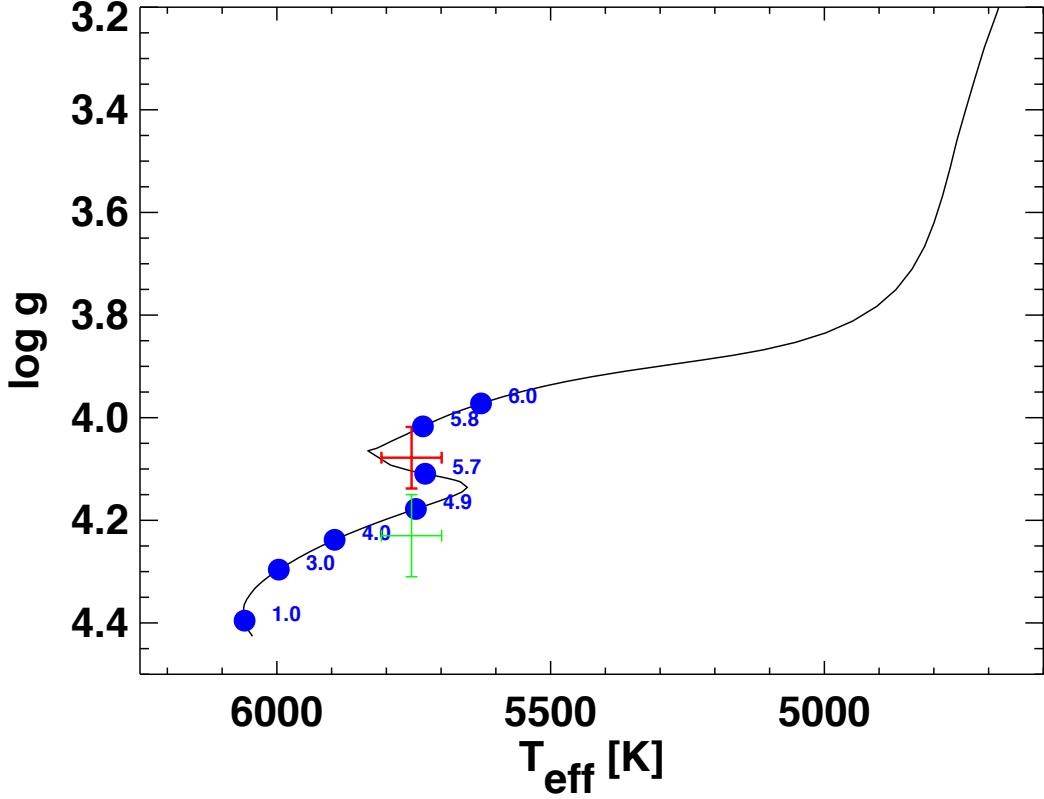


Figure 3.5 Yonsei-Yale isochrone fit for KELT-8. The solid black line is the isochrone for a star with the mass and metallicity of KELT-8 listed in Table 4.3. The red cross is the position of KELT-8, and the blue points label various ages along the evolutionary track. The green cross is the position of KELT-8 using the spectroscopically-determined $\log g$. The “blue hook” is the kink in the track where the star’s T_{eff} increases while $\log g$ decreases for a short amount of time between an age of 4.9 and 5.8 Gyr.

measurement errors on T_{eff} and $\log g$ don’t allow us to locate the star’s exact position within the blue hook. We conservatively quote an age range of 4.9–5.8 Gyr that spans the entire blue hook. We note that the uncertainty on the age likely does not follow a Gaussian distribution because the rapid evolution through the blue hook means that stars on the blue hook should be rare. It is much more likely that it falls on one side or the other of the blue hook simply because stars spend such a small fraction of the lifetimes on the blue hook. A more detailed evolutionary analysis that includes rotation, chemical composition, and a full consideration of all relevant priors is needed to determine the exact evolutionary state of KELT-8.

Table 3.3. Adopted Stellar Properties of KELT-8

Parameter (units)	Value	Source	Reference
Names	HD 343246 GSC 02109-00049 2MASS J18531332+2407385 TYC 2109-49-1 WDS 18533+2407	SIMBAD SIMBAD SIMBAD SIMBAD WDS	Hartkopf et al. (2013a)
Spectral type	G2V	this work	
RA (J2000)	18 53 13.3216	Tycho-2	Høg et al. (2000)
DEC (J2000)	+24 07 38.603	Tycho-2	Høg et al. (2000)
$U^a(m\ s^{-1})$	20.6 ± 1.7	this work	
$V (m\ s^{-1})$	-32.6 ± 1.2	this work	
$W (m\ s^{-1})$	0.7 ± 1.5	this work	
RA proper motion (mas/yr)	-13.7 ± 1.4	Tycho-2	Høg et al. (2000)
DEC proper motion (mas/yr)	-29.5 ± 1.3	Tycho-2	Høg et al. (2000)
Systemic velocity ($m\ s^{-1}$)	-31.2 ± 0.3	this work	
NUV (mag)	16.884 ± 0.1^b	GALEX	Bianchi et al. (2011)
B_T (mag)	11.713 ± 0.057	Tycho-2	Høg et al. (2000)
V_T (mag)	10.925 ± 0.048	Tycho-2	Høg et al. (2000)
B (mag)	11.545 ± 0.065	ASCC	Kharchenko (2001)
V (mag)	10.833 ± 0.054	ASCC	Kharchenko (2001)
J (mag)	9.586 ± 0.026	2MASS	Cutri et al. (2003)
H (mag)	9.269 ± 0.032	2MASS	Cutri et al. (2003)
K (mag)	9.177 ± 0.021	2MASS	Cutri et al. (2003)
$W1$ (mag)	11.664 ± 0.05	WISE	Cutri et al. (2014)
$W2$ (mag)	12.302 ± 0.05	WISE	Cutri et al. (2014)
$W3$ (mag)	14.17 ± 0.05	WISE	Cutri et al. (2014)
$W4$ (mag)	15.193 ± 0.339	WISE	Cutri et al. (2014)
$v \sin i(km\ s^{-1})$	3.7 ± 1.5	this work	
Age (Gyr)	$5.4^{+0.4}_{-0.5}$	this work	
Distance (pc)	236 ± 9	this work	
A_V	0.15 ± 0.06	this work	
$\log R'_{HK}$	-5.108	this work	

^aPositive U is in the direction of the Galactic Center.

^bLikely contaminated by the neighbor star 8''8 away

3.3.4 Unrelated visual companion

KELT-8 was identified as a visual binary in 1905 by French astronomer Abel Pourteau (Pourteau, 1933). Figure 3.3 shows KELT-8 and the fainter visual companion $8''8$ to the SW clearly resolved. In addition to the discovery astrometry, further astrometric measurements were made in 1951 (Gellera, 1984, 1990), 2000 (Cutri et al., 2003), and 2001 (Hartkopf et al., 2013b). These measurements all show a trend of decreasing separation and increasing position angle, although the scatter in the pre-CCD measurements is large (see Figure 3.6). We also measured astrometry from our Robo-AO data and find that it is consistent with this trend, especially if only the CCD measurements are considered. This is consistent with KELT-8 moving with the constant proper motions given in Table 5.2 and the companion being stationary in the background.

We also collected an APF spectrum of the secondary component and used the SpecMatch pipeline to extract stellar parameters. We cross-correlated the best-fitting model with the observed spectrum after correcting for the Earth’s barycentric motion to determine an absolute systemic velocity. SpecMatch returned $T_{\text{eff}} = 4702 \pm 60$ K, $\log g = 2.77 \pm 0.08$, and $[\text{Fe}/\text{H}] = 0.01 \pm 0.04$ which is consistent with the companion being a G or K giant with an intrinsic luminosity much greater than that of KELT-8. Since this component appears approximately one magnitude fainter than KELT-8 in V band it must be in the background. In addition, the systemic RV of the secondary of 34 ± 1 m s $^{-1}$ is highly discrepant with the velocity of KELT-8 (-31.2 ± 0.3 km s $^{-1}$). These lines of evidence all suggest that the two stars are physically unrelated.

3.3.5 UVW Space motion

We evaluate the motion of KELT-8 through the galaxy. The absolute radial velocity measured from the APF spectra of KELT-8 is -31.2 ± 0.3 km s $^{-1}$, and the proper motions from Tycho-2 are -13.7 ± 1.4 and -29.5 ± 1.3 milli-arcseconds per year in RA and DEC respectively. These values transform to U,V,W space motions of 20.6 ± 1.7 , -32.6 ± 1.2 , 0.7 ± 1.5 km s $^{-1}$ respectively, making KELT-8 an unambiguous member of the thin disk stellar population (Bensby et al., 2003). Since it is not a halo star this further supports our adoption of the stellar $\log g$ derived from the transit fit which gives stellar parameters (particularly metallicity) that are much more common for thin disk stars (Ivezić et al., 2012).

3.4 Synthetic Template Radial Velocities

Iodine-based radial velocities are traditionally derived using a forward modeling process that requires three primary ingredients; a template spectrum of the star with the iodine out of the light path, an ultra-high resolution spectrum of the iodine absorption for the particular cell in use, and a description of the instrumental PSF (Butler et al., 1996b). The stellar template and the iodine spectrum are multiplied together then convolved with the PSF to match the observed spectrum

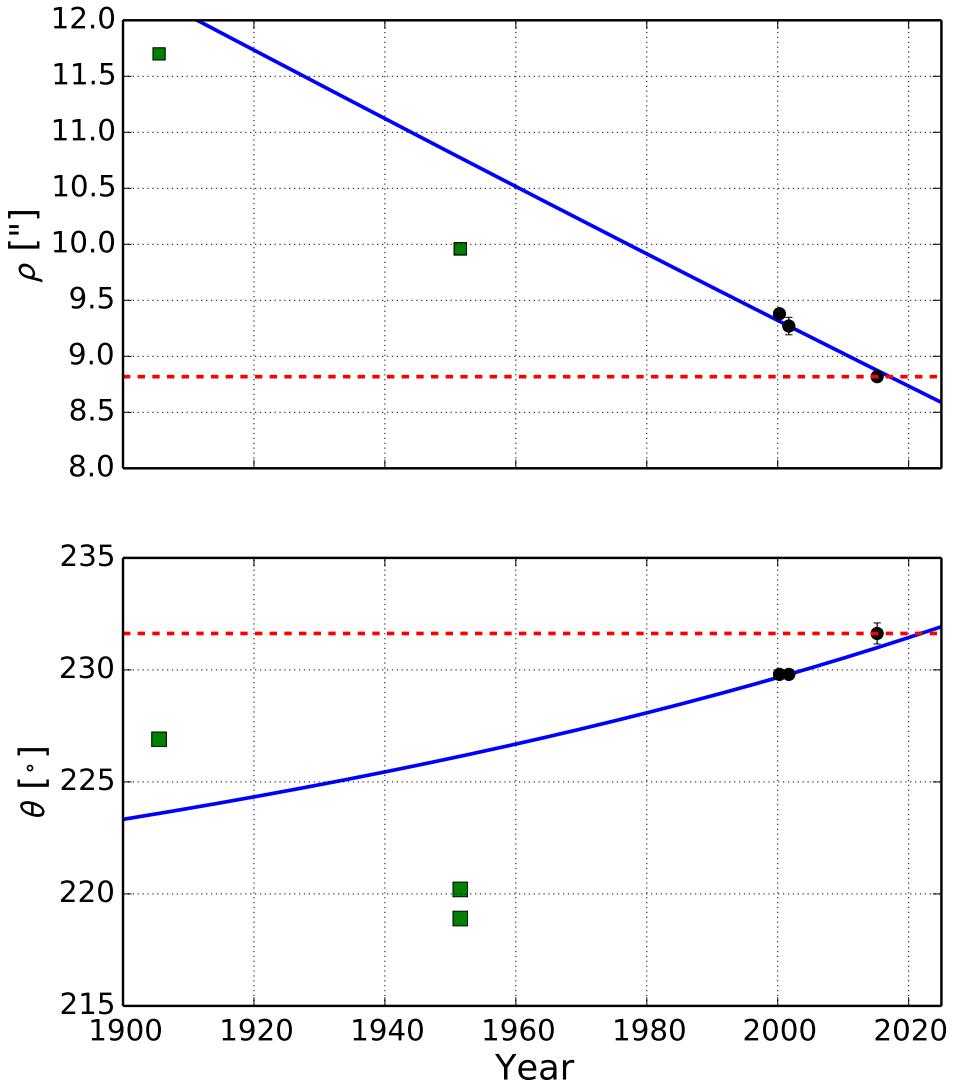


Figure 3.6 Astrometry measurements of KELT-8. The green squares were measured using astrometric eyepieces or photographic plates. The black points were measured using CCD data and the most recent point is from our Robo-AO observation. The blue lines are the predicted change in separation (upper panel) and position angle (lower panel) assuming the star 8"8 to the SW of KELT-8 is in the background with negligible proper motion and KELT-8 has a constant proper motion with the values given in Table 5.2. The dashed black line is a model assuming that the two components are physically bound and the orbital motion of the pair is negligible during the span of observations.

taken through iodine. The stellar template is usually an observation of the same star taken at 2-3 times higher signal-to-noise (S/N) ratio than the iodine observations and using a narrower slit for higher resolution. Since the PSF is constantly changing on a slit-fed spectrograph due to changes in the slit illumination caused by variable seeing and guiding patterns, we need to deconvolve this stellar template to remove distortions in the line profile. This deconvolution process also requires a measurement of the instrumental PSF. By construction, we do not have a simultaneous PSF calibration for the template observations as we do for the iodine observations. We rely on bracketing observations of rapidly rotating B stars through iodine, but the PSF is never exactly the same for any two observations and this non-simultaneous calibration introduces errors. In addition, the deconvolution process is inherently unstable because we are essentially trying to create a view of the star at infinite spectral resolution using data taken at finite spectral resolution. Collecting as much signal as possible for the template observation helps ensure that this component of the model is not the limiting factor on the final RV precision.

For faint stars ($V > 12$), iodine exposures on Keck/HIRES can take 45-60 min each. Collecting a template at 2-3 times higher S/N ratio would take a significant fraction of a night. At APF the situation is even worse due to its much smaller aperture (2.4 m vs. 10 m). When collecting reconnaissance RVs for a new candidate from a transit survey to look for various false positive scenarios only 2-3 RVs may be needed before a false alarm scenario can be confirmed or eliminated. Collecting a high S/N ratio template can take as much time as the entire set of RV observations, effectively doubling the observational time needed in these scenarios.

In order to circumvent these problems we developed a technique that uses a synthetic template using the Coelho (2014) models and spectral synthesis code from SpecMatch. We first determine the spectral parameters from the best available information. Approximate values for T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$ can be obtained using spectral regions outside of the iodine absorption regime, photometry, or low resolution spectroscopy. We normally use a single echelle order redward of the iodine regime to estimate $v \sin i$. T_{eff} and $v \sin i$ have the most drastic impact on the resulting RV, but fortunately T_{eff} is the easiest parameter to estimate in the absence of a high resolution spectrum and $v \sin i$ can be estimated with very few spectral lines. Once the spectral parameters are established we synthesize the full spectrum within the iodine region setting the instrumental PSF width to zero. This provides an extremely high-resolution, noise-free stellar template to use as one of the inputs to the normal RV extraction code. This foregoes the need for deconvolution and saves valuable telescope time.

This technique is not perfect due to inaccuracies in the model spectra. Spectral lines are missing in some regions, blended lines may only be accounted for as a single line, or the central wavelengths of lines may be slightly inaccurate. This means that using these synthetic templates we can never achieve the $2-3 \text{ m s}^{-1}$ precision that we can obtain using the traditional technique for bright stars. Figure 3.7 shows velocity RMS for the RV standard star (HD 9407) as a function of

S/N ratio of the iodine observations using the synthetic template technique. We artificially inject Gaussian noise into the spectra before sending them through the normal RV extraction pipeline. We use an exposure meter to expose to S/N=200 for all normal observations and the highest S/N ratio datapoint shown in Figure 3.7 has no artificial noise injected. The RV RMS monotonically decreases down to extremely low S/N ratio levels with a noise floor at just below 10 m s^{-1} for high S/N ratio. Two other RV standard stars were also run through the synthetic template pipeline with S/N ratio artificially decreased to match that of our KELT-8 observations. In each case the synthetic template RV RMS was about 10 m s^{-1} (see Figure 3.8). We use this technique to extract the RVs from the APF data for KELT-8 and achieve a median per-measurement precision of 11.8 m s^{-1} .

While our technique is unique for iodine-based RVs in the optical, several other very similar techniques have previously been developed to extract high precision RVs. Bailey et al. (2012) and Tanner et al. (2012) use interpolated model atmospheres for their templates, but instead of an iodine cell they use telluric lines for their PSF and wavelength calibration. They were able to achieve $\sim 50 \text{ m s}^{-1}$ precision in the near infrared with NIRSPEC on Keck. Other telluric-calibrated techniques have proven successful in the optical with HIRES to obtain velocities to 100 m s^{-1} precision (Chubak et al., 2012). Johnson et al. (2006) perturb an existing empirical template of a similar star in order to avoid the need to collect a unique template for each star observed. They were able to achieve a precision of $3\text{--}5 \text{ m s}^{-1}$ using these perturbed empirical templates.

3.5 Planetary Characterization

3.5.1 Global Fit

We determine the system's orbital and planetary properties by performing a global analysis on the RVs, spectroscopic stellar parameters, and follow-up light curves using a custom version of EXOFAST (Eastman et al., 2013). Our global analysis technique is very similar to that used for the previous KELT discoveries (Siverd et al., 2012; Beatty et al., 2012; Pepper et al., 2013; Collins et al., 2014; Bieryla et al., 2015). The EXOFAST package uses a DE-MCMC fitting routing to extract the median parameter values and 68% confidence intervals using all available data and these values are presented in Table 4.3. Constraints on the stellar parameters M_* and R_* are also included using the Yonsei-Yale stellar models and spectroscopically derived values for T_{eff} and $[\text{Fe}/\text{H}]$ as input priors. Before running the MCMC chains, a best-fit solution is found using an AMOEBA χ^2 minimization routine and the photometric and RV uncertainties are scaled such that $\chi^2=1$. The scaling factors for the photometric data are listed in Table 3.1 and the scaling factor applied to the APF RV uncertainties is 1.44. The scaled uncertainties include both instrumental and astrophysical noise sources and this ensures that the widths of final posterior properly represent the uncertainties in the model parameters.

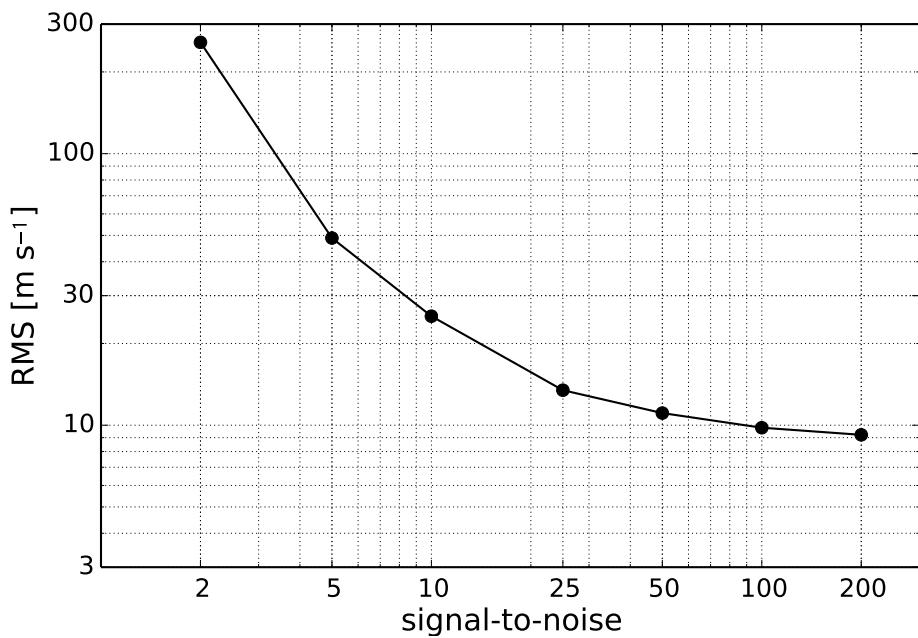


Figure 3.7 Synthetic template RV performance as a function of S/N ratio of the iodine observations. Artificial Gaussian noise was injected into the spectra of a well-known RV standard star (HD 9407) before running them through the RV extraction pipeline for all data points excluding the one at S/N=200. The synthetic template technique shows predictable and consistent performance down to extremely low S/N ratios and a model-limited noise floor of $\approx 10 \text{ m s}^{-1}$.

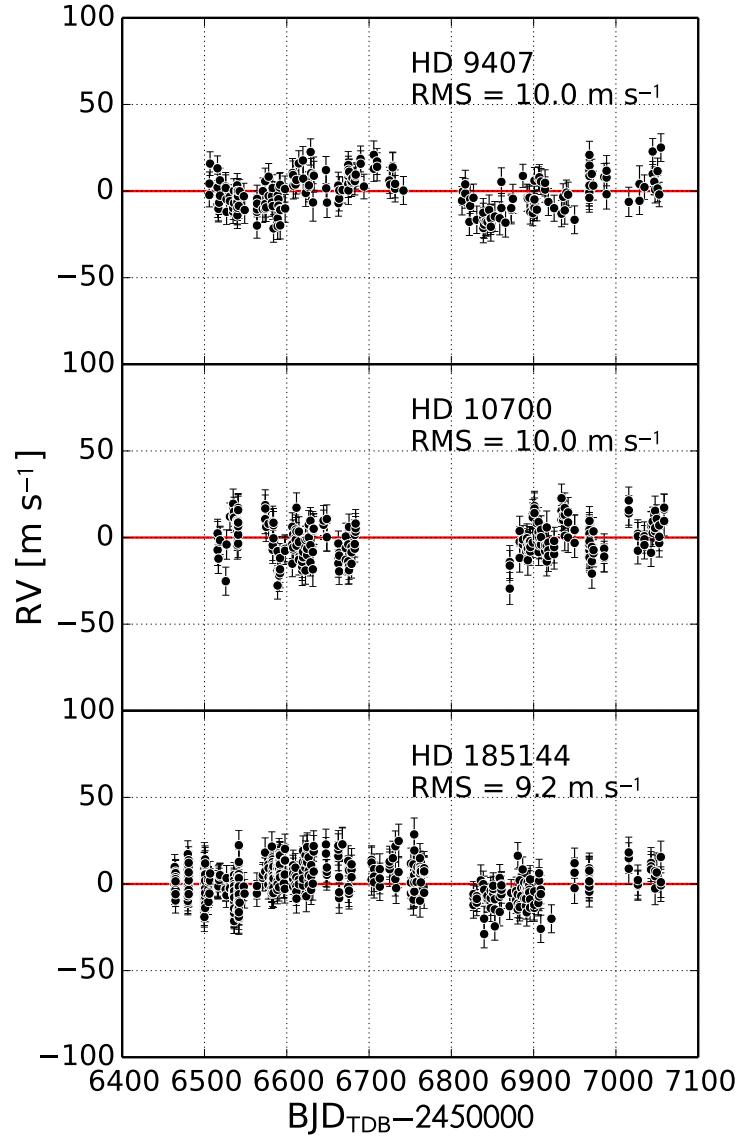


Figure 3.8 Synthetic template RV performance for three RV standard stars. Artificial Gaussian noise was injected into all spectra for each star so that S/N=50 for each spectrum. This is approximately the same S/N ratio that was collected for each KELT-8 observation. The RV RMS is about 10 m s⁻¹ for each star.

As described in Collins et al. (2014), we identify trend datasets that best improve the individual light curve fits using AIJ. These trend datasets are included as input to EXOFAST, along with the undetrended light curve data. As part of the global fit, EXOFAST simultaneously finds the best transit model fit and removes linear trends that are correlated with different combinations of airmass, time, FWHM, and CCD position.

The light from the visual companion to KELT-8 was included in the photometric aperture of the GCorfini, and ZRO light curves. We corrected for this by measuring the flux ratio of the primary to secondary component in the well resolved MORC data. The flux ratios are 1.24 and 1.44 in g' and i' band respectively. We then interpolated/extrapolated these flux ratios to the central wavelengths of the V and Ic GCorfini and ZRO light curves and corrected the input light curves for the dilution before fitting them in EXOFAST. We also ran the global fit excluding these two transits and found that all parameters are consistent to well within 1σ . This gives us confidence that these blended light curves are not biasing the inferred parameters.

We ran four permutations of the global fit using slightly different input priors and constraints. First, we assert constraints on the stellar mass and radius from the Yonsei-Yale isochrones and a uniform prior on $\log g$. We adopt the parameter values from this fit for all analysis and interpretation. Second, we performed the same fit using constraints from the Torres et al. (2010) relations instead. These two fits produced nearly identical results. Then we performed these same two global fits but imposed a prior on $\log g$ from the high-resolution spectroscopy. The $\log g$ derived from spectroscopy is slightly higher and only marginally consistent with the $\log g$ derived from transits. The results from all four fits are presented in Table 4.3. The differences in parameter values from these four fits give an estimate of our systematic uncertainty which appears to be similar to the statistical uncertainties for most parameters. This gives us confidence that our parameter uncertainties are properly estimated and that our interpretations of physical parameters (such as the large planetary radius) are robust.

We measure an eccentricity for KELT-8b of $e = 0.035^{+0.050}_{-0.025}$ which deviates from zero by only 1.4σ . The measurement of eccentricity is easily biased to artificially larger values and a significance of $\geq 2.5\sigma$ is generally accepted as being required to claim a non-circular orbit (Lucy & Sweeney, 1971). Although we do not find any significant eccentricity in the system, moderate eccentricity is not unprecedented in other hot Jupiters (e.g. HAT-P-2b, XO-3b, WASP14b, HAT-P-31b Bakos et al., 2007; Winn et al., 2009; Joshi et al., 2009; Kipping et al., 2011). We allow for eccentricity to vary as a free parameter in order to ensure that the errors on all other parameters are not underestimated.

We searched for possible transit timing variations (TTVs) in the system by allowing the transit times for each of the follow-up light curves to vary. The ephemeris is constrained by the RV data and a prior imposed from the KELT discovery data. We find no evidence for significant transit timing variations. The individual transit times are presented in Table 3.5 and Figure 3.10.

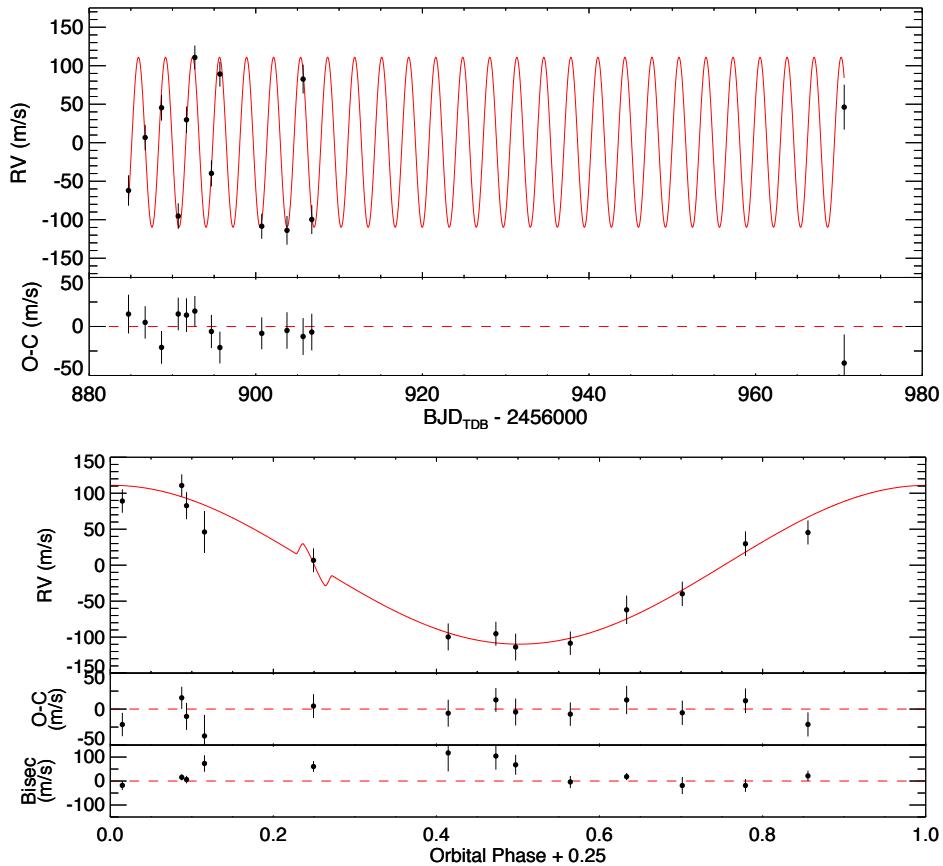


Figure 3.9 *Top:* APF RVs and residuals for KELT-8. The best-fit model is shown in red. *Bottom:* RVs and bisector span measurements phase-folded to the best-fit linear ephemeris. The best-fit model is shown in red. The predicted Rossiter-McLaughlin effect assumes perfect spin-orbit alignment and it is not constrained by our data.

After the completion of the global fit we then derive a linear ephemeris. The period is constrained by the RV data and comes from the global fit. We fit for T_0 using the transit times from the followup light curves. The best-fit linear ephemeris is listed in Table 4.3. We find $P = 3.24406 \pm 0.00016$ days and $T_0 = 2456883.4803 \pm 0.0007$ BJD_{TDB}. The χ^2 for the linear ephemeris fit is 7.4 with 4 degrees of freedom which indicates that a linear ephemeris adequately describes the observed transit times.

3.5.2 False-Positive Assessment

We calculate bisector spans (BS) for each of the APF spectra in order to investigate possible blend scenarios in which the measured RV variations are caused by line distortions that arise from an unresolved eclipsing binary instead of true reflex motion of the star. In the case that the measured Doppler shifts are caused by a second set of incompletely resolved stellar lines from a luminous

Table 3.4. Median values and 68% confidence intervals for KELT-8b.

Parameter	Units	Adopted Value (YY)	Value (Torres)	Value (YY+log g prior)	Value (Torres+log g prior)
Stellar Parameters:					
M_*	Mass (M_\odot)	$1.211^{+0.078}_{-0.066}$	$1.233^{+0.066}_{-0.063}$	$1.188^{+0.067}_{-0.063}$	$1.211^{+0.060}_{-0.059}$
R_*	Radius (R_\odot)	$1.67^{+0.14}_{-0.12}$	$1.69^{+0.15}_{-0.12}$	$1.570^{+0.099}_{-0.097}$	1.577 ± 0.096
L_*	Luminosity (L_\odot)	$2.74^{+0.51}_{-0.40}$	$2.79^{+0.54}_{-0.41}$	$2.42^{+0.35}_{-0.32}$	$2.44^{+0.34}_{-0.31}$
ρ_*	Density (cgs)	$0.369^{+0.073}_{-0.067}$	$0.363^{+0.078}_{-0.072}$	$0.433^{+0.076}_{-0.061}$	$0.436^{+0.078}_{-0.062}$
$\log g_*$	Surface gravity (cgs)	$4.078^{+0.049}_{-0.054}$	$4.075^{+0.054}_{-0.061}$	$4.121^{+0.043}_{-0.041}$	$4.125^{+0.045}_{-0.042}$
T_{eff}	Effective temperature (K)	5754^{+54}_{-55}	5749 ± 59	5748 ± 53	5750 ± 59
[Fe/H]	Metallicity	0.272 ± 0.038	0.270 ± 0.039	$0.270^{+0.038}_{-0.039}$	$0.271^{+0.040}_{-0.039}$
Planetary Parameters:					
e	Eccentricity	$0.035^{+0.050}_{-0.025}$	$0.037^{+0.057}_{-0.027}$	$0.028^{+0.033}_{-0.019}$	$0.027^{+0.033}_{-0.020}$
ω_*	Argument of periastron (degrees)	85^{+87}_{-97}	85^{+82}_{-95}	-80^{+120}_{-130}	-80^{+120}_{-130}
a	Semi-major axis (AU)	$0.04571^{+0.00096}_{-0.00084}$	$0.04599^{+0.00081}_{-0.00079}$	$0.04542^{+0.00084}_{-0.00082}$	$0.04571^{+0.0074}_{-0.0075}$
M_P	Mass (M_J)	$0.867^{+0.065}_{-0.061}$	$0.876^{+0.062}_{-0.061}$	$0.855^{+0.062}_{-0.061}$	$0.869^{+0.061}_{-0.060}$
R_P	Radius (R_J)	$1.86^{+0.18}_{-0.16}$	$1.88^{+0.19}_{-0.16}$	1.73 ± 0.13	1.74 ± 0.13
ρ_P	Density (cgs)	$0.167^{+0.047}_{-0.038}$	$0.163^{+0.049}_{-0.039}$	$0.205^{+0.051}_{-0.039}$	$0.205^{+0.052}_{-0.040}$
$\log g_P$	Surface gravity	$2.793^{+0.072}_{-0.075}$	$2.786^{+0.076}_{-0.080}$	$2.850^{+0.065}_{-0.063}$	$2.853^{+0.066}_{-0.064}$
T_{eq}	Equilibrium temperature (K)	1675^{+61}_{-55}	1679^{+56}_{-57}	1629^{+44}_{-48}	1628^{+46}_{-47}
Θ	Safronov number	$0.0351^{+0.0040}_{-0.0037}$	$0.0346^{+0.0041}_{-0.0038}$	$0.0378^{+0.0039}_{-0.0035}$	$0.0377^{+0.0039}_{-0.0035}$
$\langle F \rangle$	Incident flux ($10^9 \text{ erg s}^{-1} \text{ cm}^{-2}$)	$1.78^{+0.27}_{-0.22}$	$1.80^{+0.29}_{-0.23}$	$1.60^{+0.19}_{-0.18}$	$1.59^{+0.19}_{-0.18}$
RV Parameters:					
T_P	Time of periastron (BJD _{TDB} -2450000)	$6870.47^{+0.75}_{-0.86}$	$6870.47^{+0.70}_{-0.83}$	$6869.0^{+1.1}_{-1.2}$	$6869.0^{+1.1}_{-1.2}$
K	RV semi-amplitude (m s^{-1})	104.0 ± 6.4	104.0 ± 6.4	104.1 ± 6.5	104.4 ± 6.5
M_P/M_* ..	Mass ratio	$0.000683^{+0.000045}_{-0.000044}$	$0.000679^{+0.000044}_{-0.000043}$	$0.000687^{+0.000045}_{-0.000044}$	0.000685 ± 0.000044
γ_{APF}	m s^{-1}	6.1 ± 4.9	6.0 ± 5.0	6.7 ± 4.9	$6.7^{+4.9}_{-5.0}$
$e \cos \omega_*$	$0.001^{+0.024}_{-0.021}$	$0.001^{+0.025}_{-0.021}$	$0.001^{+0.022}_{-0.019}$	$0.001^{+0.022}_{-0.018}$
$e \sin \omega_*$	$0.012^{+0.063}_{-0.025}$	$0.014^{+0.071}_{-0.027}$	$-0.001^{+0.029}_{-0.032}$	$-0.001^{+0.028}_{-0.034}$
Primary Transit Parameters:					
R_P/R_* ..	Planet to star radius ratio	0.1145 ± 0.0026	0.1146 ± 0.0026	0.1133 ± 0.0025	0.1132 ± 0.0025
a/R_* ..	Semi-major axis in stellar radii	$5.90^{+0.37}_{-0.38}$	$5.87^{+0.39}_{-0.42}$	$6.22^{+0.34}_{-0.31}$	$6.23^{+0.35}_{-0.31}$
i'	Inclination (degrees)	$82.65^{+0.81}_{-1.0}$	$82.56^{+0.88}_{-1.2}$	83.36 ± 0.69	$83.38^{+0.69}_{-0.70}$
b	Impact parameter	$0.741^{+0.027}_{-0.033}$	$0.743^{+0.027}_{-0.033}$	$0.722^{+0.028}_{-0.036}$	$0.722^{+0.029}_{-0.037}$
δ	Transit depth	0.01311 ± 0.00059	$0.01313^{+0.00061}_{-0.00059}$	$0.01283^{+0.00058}_{-0.00057}$	0.01282 ± 0.00057
τ	Ingress/egress duration (days)	$0.0303^{+0.0035}_{-0.0034}$	$0.0305^{+0.0039}_{-0.0036}$	$0.0280^{+0.0033}_{-0.0032}$	$0.0280^{+0.0034}_{-0.0032}$
T_{14}	Total duration (days)	$0.1444^{+0.0033}_{-0.0033}$	0.1446 ± 0.0034	$0.1423^{+0.0031}_{-0.0031}$	0.1423 ± 0.0032
u_{1I}	Linear Limb-darkening	$0.2961^{+0.0092}_{-0.0088}$	$0.2966^{+0.0096}_{-0.0093}$	$0.2977^{+0.0089}_{-0.0087}$	$0.2975^{+0.0096}_{-0.0092}$
u_{2I}	Quadratic Limb-darkening	$0.2793^{+0.0044}_{-0.0049}$	$0.2790^{+0.0047}_{-0.0052}$	$0.2782^{+0.0042}_{-0.0047}$	$0.2783^{+0.0046}_{-0.0051}$
$u_{1Sloang'}$	Linear Limb-darkening	0.607 ± 0.015	$0.607^{+0.016}_{-0.015}$	0.608 ± 0.015	$0.608^{+0.016}_{-0.015}$
$u_{2Sloang'}$	Quadratic Limb-darkening	0.184 ± 0.010	$0.184^{+0.011}_{-0.012}$	0.183 ± 0.010	0.184 ± 0.010
$u_{1Sloani'}$	Linear Limb-darkening	$0.3178^{+0.0097}_{-0.0092}$	$0.3183^{+0.0101}_{-0.0097}$	$0.3193^{+0.0094}_{-0.0091}$	$0.3192^{+0.0096}_{-0.0096}$
$u_{2Sloani'}$	Quadratic Limb-darkening	$0.2789^{+0.0046}_{-0.0052}$	$0.2786^{+0.0050}_{-0.0056}$	$0.2779^{+0.0045}_{-0.0051}$	$0.2779^{+0.0049}_{-0.0055}$
$u_{1Sloanr'}$	Linear Limb-darkening	0.410 ± 0.012	0.410 ± 0.012	$0.411^{+0.012}_{-0.011}$	0.411 ± 0.012
$u_{2Sloanr'}$	Quadratic Limb-darkening	$0.2725^{+0.0062}_{-0.0070}$	$0.2722^{+0.0066}_{-0.0074}$	$0.2715^{+0.0066}_{-0.0078}$	$0.2717^{+0.0066}_{-0.0074}$
u_{1V}	Linear Limb-darkening	$0.487^{+0.012}_{-0.013}$	$0.488^{+0.013}_{-0.013}$	$0.488^{+0.012}_{-0.012}$	$0.488^{+0.013}_{-0.013}$
u_{2V}	Quadratic Limb-darkening	$0.2446^{+0.0077}_{-0.0088}$	$0.2442^{+0.0082}_{-0.0093}$	$0.2435^{+0.0078}_{-0.0087}$	$0.2437^{+0.0082}_{-0.0093}$
Linear Ephemeris from Follow-up Transits:					
P	Period (days)	3.24406 ± 0.00016	3.24406 ± 0.00016	3.24406 ± 0.00016	3.24406 ± 0.00016
T_0	Time of transit (BJD _{TDB} -2450000)	6883.4803 ± 0.0007	6883.4803 ± 0.0007	6883.4804 ± 0.0007	6883.4804 ± 0.0007
Predicted Secondary Eclipse Parameters:					
T_S	Time of eclipse (BJD _{TDB} -2450000)	$6868.886^{+0.049}_{-0.043}$	$6868.887^{+0.050}_{-0.044}$	$6868.886^{+0.044}_{-0.039}$	$6872.130^{+0.044}_{-0.038}$
b_S	Impact parameter	$0.763^{+0.097}_{-0.057}$	$0.769^{+0.11}_{-0.061}$	$0.718^{+0.051}_{-0.054}$	$0.716^{+0.051}_{-0.055}$
τ_S	Ingress/egress duration (days)	$0.0327^{+0.016}_{-0.0059}$	$0.0332^{+0.019}_{-0.0065}$	$0.0277^{+0.0054}_{-0.0042}$	$0.0276^{+0.0053}_{-0.0042}$
$T_{S,14}$	Total duration (days)	$0.1436^{+0.0034}_{-0.0049}$	$0.1435^{+0.0036}_{-0.0060}$	$0.1419^{+0.0033}_{-0.0041}$	$0.1418^{+0.0034}_{-0.0042}$

Table 3.5. Transit Times for KELT-8b.

Epoch	T_C (BJD _{TDB})	σ_{T_C} (s)	O-C (s)	O-C (σ_{T_C})	Telescope
-4	2456870.502251	200	-156.3	-0.78	ZRO
-4	2456870.509188	203	443.1	2.18	GCorfini
-3	2456873.746945	114	-101.5	-0.89	MORC
-3	2456873.749450	137	114.9	0.83	MORC
5	2456899.699249	138	-116.7	-0.85	WCO
8	2456909.433338	177	48.2	0.27	CROW

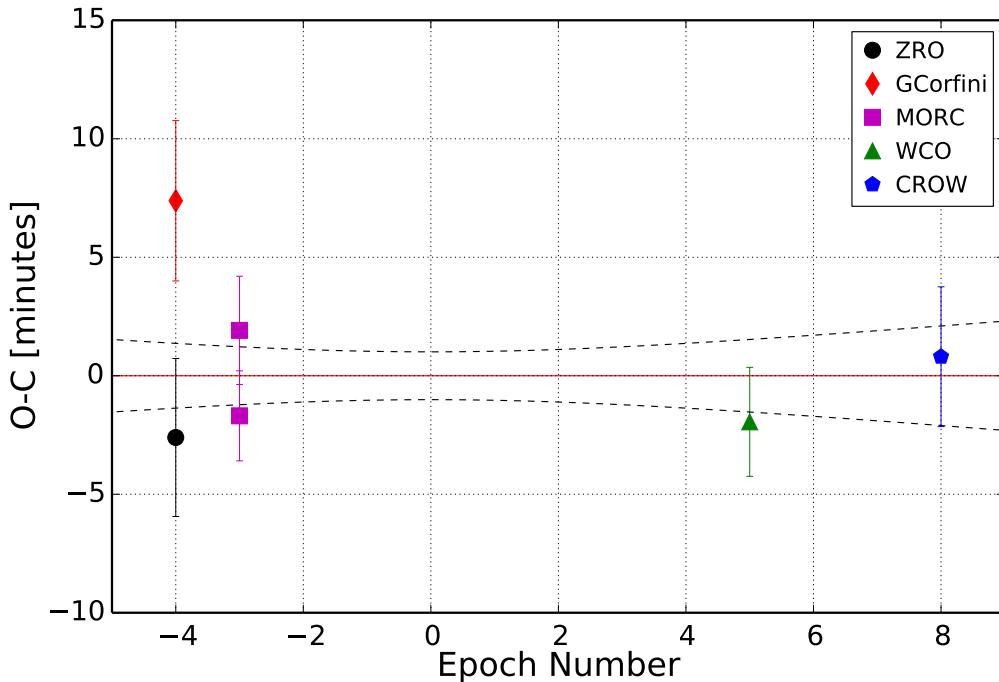


Figure 3.10 Residuals of the transit times from the best-fit ephemeris for the followup light curves. The source of each transit time is denoted in the upper right. The dashed lines indicate the 1 σ uncertainty on the linear ephemeris.

companion we would expect to see BSs that vary in phase with the orbital signal and a correlation in BS vs. RV (Santos et al., 2002; Torres et al., 2005). The amplitude would be similar to that of the velocities (Queloz et al., 2001; Mandushev et al., 2005).

Following the prescription of Torres et al. (2005) and Torres et al. (2007), each APF spectrum was cross-correlated with a synthetic spectrum derived using the stellar properties determined from the high S/N ratio Keck spectrum (see Section 5.3). We analyzed the cross-correlation function for each of the echelle orders between 4260 and 5000 Å in order to avoid the iodine lines, and telluric lines. By restricting the BS analysis to blue orders we reduce the effect of instrumental PSF variations caused by guiding errors because the seeing is degraded toward the blue which means that the slit is more evenly illuminated. We measure the asymmetry in the spectral line profile for each order by calculating the velocity at the midpoint of lines connecting the CCF at many different fractional levels of CCF peak. The BS is then the difference in velocity between the 65th and 95th percentile levels of the CCF. The bisector spans and error corresponding to a given observation are the mean and standard deviation on the mean over the 15 spectral orders analyzed.

Our BS measurements and errors are listed in Table 5.1 and shown in Figure 3.11. Figure 3.9 also shows the BS measurements as a function of orbital phase. We find no statistically significant correlation of BS with RV. The Spearman rank correlation coefficient is -0.35 ($p=0.24$, Spearman, 1904).

The lack of in-phase BS variations, and the fact that the stellar $\log g$ derived from spectroscopy and $\log g$ derived from the transit light curves are consistent to within $1-\sigma$ lead us to conclude that the RV variations and transit signals are caused by a highly-inflated Jovian planet.

3.6 Discussion

3.6.1 Prospects for further characterization of the host star

Further stellar characterization of KELT-8 is needed to better constrain the mass and radius of the star which will improve our knowledge of the planetary mass and radius. Several more light curves with higher precision in which the two components are well-resolved would help to better constrain the stellar density. A more detailed analysis of higher S/N ratio spectra may help to nail down the spectroscopically derived parameters and allow for chemical abundance measurements. GAIA (de Bruijne, 2012) should provide a precise parallax to constrain the stellar luminosity and thus radius, and TESS and/or PLATO (Rauer et al., 2014) may provide sufficient photometric precision to measure the stellar density using asteroseismology.

3.6.2 Comparative Planetology

KELT-8b has the 2nd largest radius of all known exoplanets, but has a mass slightly less than that of Jupiter. It lies well above the theoretical mass-radius curve for pure hydrogen (see Figure 3.12).

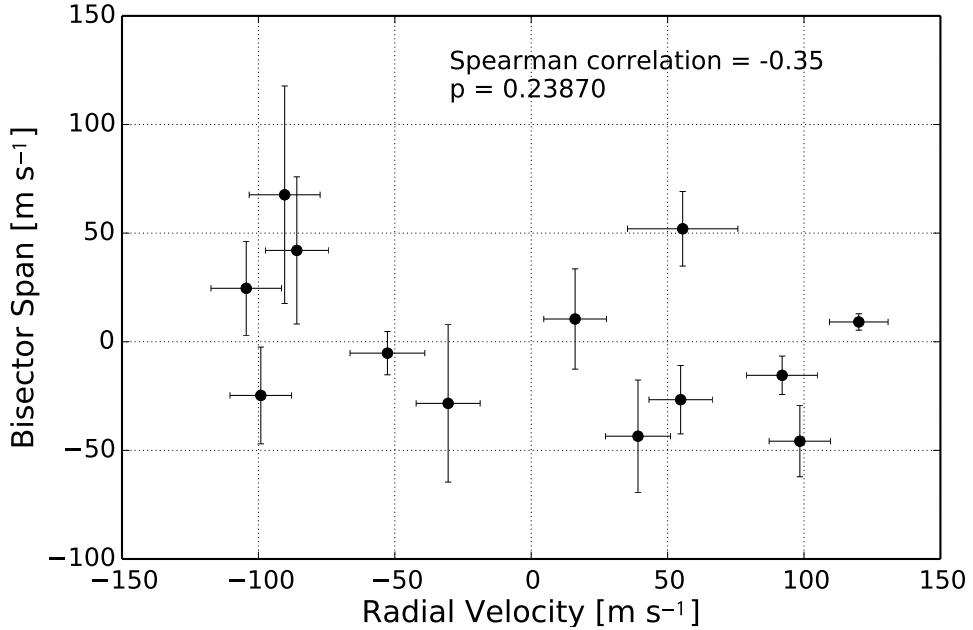


Figure 3.11 BS measurements for the APF spectra used for radial velocity measurements. We find no significant correlation between RV and BS.

KELT-8b is similar to WASP-1b (Collier Cameron et al., 2007), and HAT-P-41b (Hartman et al., 2012). It orbits a star more massive and metal rich than the Sun ($M_\star \geq 1 M_\odot$ and $[Fe/H] \geq 0.1$), the planet is highly irradiated ($a \leq 0.05$ AU), and it is less massive than Jupiter but extremely inflated ($R_p \geq 1.5 R_J$).

Next, we compare the stellar parameters of KELT-8 with other well-characterized hot Jupiter host stars. We use the exoplanets.org catalogue to select the stars that host planets with orbital periods shorter than 10 days, and masses between 0.4 and 1.1 M_J . We plot these stars on an HR diagram in Figure 8.1. We split this sample into two groups; one group containing the stars that host planets with radii greater than $1.5 R_J$ and the other group containing stars that host planets smaller than $1.5 R_J$. KELT-8 is the coolest star to host a highly inflated hot Jupiter ($T_{\text{eff}} = 5754$ K), but we find that all other stars that host similar highly inflated planets tend to be hotter than ≈ 5900 K. However, there are also plenty of stars hotter than 5900 K that host non-inflated hot Jupiters. We also compare stellar metallically, $v \sin i$, flux received by the planet, orbital eccentricity, stellar multiplicity, and spin-orbit misalignment, but find no clear differences between the hot stars that host “normal” hot Jupiters to those that host the highly inflated hot Jupiters with radii $\geq 1.5 R_J$. However, we note that several of these host star parameters are poorly constrained for many of the systems. Albedo would be another interesting parameter to compare for these planets but it has not been measured for the majority of planets. It is possible that the

highly inflated planets have abnormally low albedos and absorb more energy into their atmospheres at a given stellar insolation. TESS (Ricker et al., 2014) should provide light curves of sufficient precision to measure secondary eclipses in reflected light for most or all of these systems.

3.6.3 Mass-Radius predictions

Several authors have identified empirical relations that attempt to predict the radius of a planet based on a combination of stellar and planetary properties. Enoch et al. (2012) established a relation that predicts planetary radius based only on the equilibrium temperature and orbital semi-major axis. KELT-8 falls into the Jupiter mass bin as defined by Enoch et al. (2012) and the relation for that mass regime predicts a radius of $1.43 R_J$ which is far too small relative to our measured value. The most recent empirical relation was presented by Weiss et al. (2013), which uses the incident flux and planetary mass as dependent variables. The Weiss et al. (2013) relation predicts a radius of $1.63 R_J$ for KELT-8. If one takes into account the scatter in the Weiss et al. (2013) relation ($0.1 R_J$), and the error on our radius measurement the predicted and measured values agree to within 1.2σ .

3.6.4 Prospects for atmospheric characterization

KELT-8 joins a rare breed of highly inflated hot Jupiters orbiting relatively bright stars. The extended atmosphere, combined with the large transit depth and bright apparent magnitude makes KELT-8 one of the best targets for transmission spectroscopy. In addition, the large planetary radius and high equilibrium temperature give rise to a significant secondary eclipse depth. We calculate an expected eclipse depth of ≈ 1.3 mmag at 3.6 microns by taking the ratio of the blackbody emission from the planet to that of the star and multiplying by $(R_P/R_*)^2$. A secondary eclipse of this depth would be easily accessible by *Spitzer*. Combining $T_{eq} = 1675$ K from Table 4.3 with the planetary surface gravity of $\log g_P = 2.793$, we find that the scale height of a H₂-dominated atmosphere ($\mu_m = 2$) on KELT-8 would be $H = 1113$ km. This large scale height implies that the amplitude of a transmission spectroscopy signal would be $\frac{2R_P H}{R_*^2} = 0.22$ mmag (Winn, 2010, assuming $H \ll R_P$). All of these calculations assume perfect heat redistribution and zero albedo for the planet. Studies of the atmospheric composition or temperature profile of KELT-8 may provide an explanation for its highly inflated radius.

3.6.5 Irradiation History

We have shown above that KELT-8b is a highly inflated planet, joining the ranks of other hot Jupiters that manifest radii much larger than predicted by standard models for non-irradiated objects with Jovian masses. Several authors (e.g., Demory & Seager, 2011) have suggested an

³based on a 2/17/2015 query of exoplanets.org

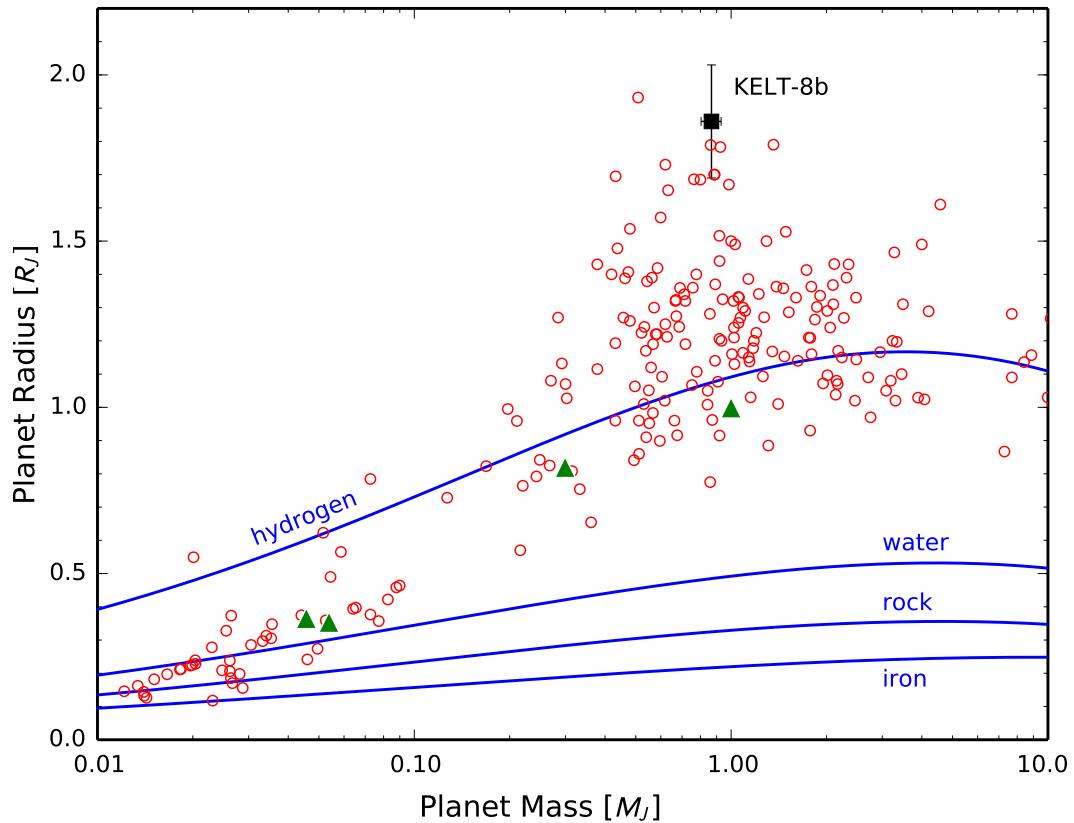


Figure 3.12 Mass radius diagram of all confirmed transiting exoplanets³(red circles).

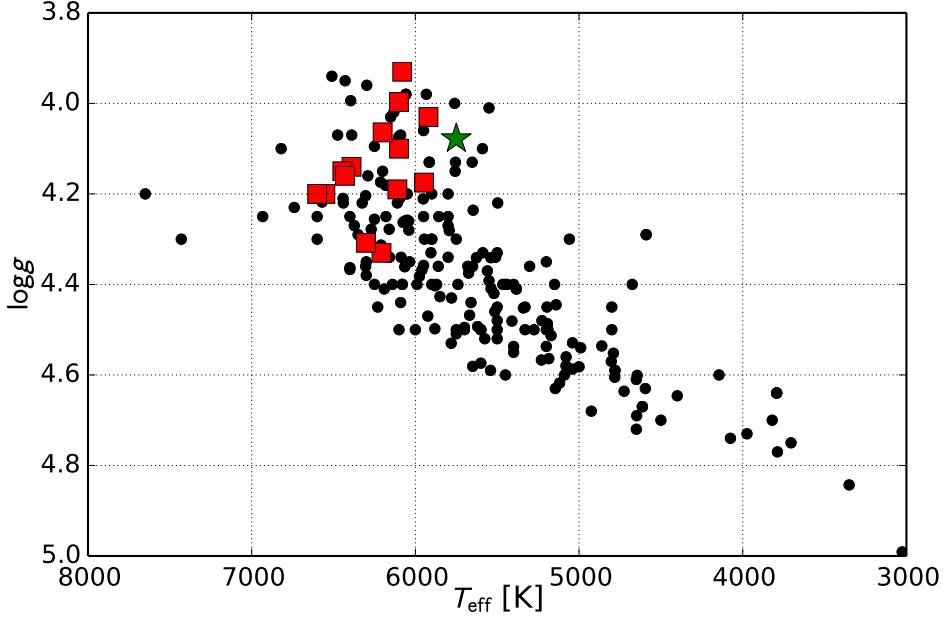


Figure 3.13 HR diagram showing well-characterized host stars of hot Jupiters. The red squares indicate stars that host planets with radii greater than $1.5 R_J$. The green star marks the position of KELT-8, and the black circles are all other systems.

empirical insolation threshold ($\approx 2 \times 10^8 \text{ erg s}^{-1} \text{ cm}^{-2}$) above which hot Jupiters exhibit increasing amounts of radius inflation. KELT-8b clearly lies above this threshold, with a current estimated insolation of $1.78^{+0.27}_{-0.22} 10^9 \text{ erg s}^{-1} \text{ cm}^{-2}$, and therefore its currently large inflated radius is not surprising. At the same time, the KELT-8 host star is found to currently be in a very rapid state of evolution, such that its radius is rapidly expanding as the star crosses the Hertzsprung gap toward the red giant branch. This means that the star's surface is rapidly encroaching on the planet, which presumably is rapidly driving up the planet's insolation and also the rate of any tidal interactions between the planet and the star.

Therefore it is interesting to consider two questions. First, has KELT-8b's incident radiation from its host star been below the empirical radius inflation threshold in the past? If KELT-8b's insolation only recently exceeded the inflation threshold, the system could then serve as an empirical testbed for the different timescales predicted by different inflation mechanisms (see, e.g., Assef et al., 2009; Spiegel & Madhusudhan, 2012). Second, what is the expected fate of the KELT-8b planet given the increasingly strong tidal interactions it is experiencing with its encroaching host star?

To investigate these questions, we follow Penev et al. (2014) to simulate the reverse and forward evolution of the star-planet system, using the measured parameters listed in Tables 5.2 and 4.3 as the present-day boundary conditions. This analysis is not intended to examine any type of planet-planet or planet-disk migration effects. Rather, it is a way to investigate (1) the change in insolation of the

planet over time due to the changing luminosity of the star and changing star-planet separation, and (2) the change in the planet’s orbital semi-major axis due to the changing tidal torque as the star-planet separation changes with the evolving stellar radius. We include the evolution of the star, assumed to follow the Yonsei-Yale stellar model with mass and metallicity as in Table 4.3. For simplicity we assume that the stellar rotation is negligible and treat the star as a solid body. We also assume a circular orbit aligned with the stellar equator throughout the full analysis. The results of our simulations are shown in Figure 3.14. We tested a range of values for the tidal quality factor of the star Q'_* , from $\log Q'_* = 5$ to $\log Q'_* = 7$ (assuming a constant phase lag between the tidal bulge and the star-planet direction). Q'_* is defined as the tidal quality factor divided by the Love number ($Q'_* = Q_*/k_2$). We find that although for certain values of Q'_* the planet has moved substantially closer to its host during the past Gyr, in all cases the planet has always received more than enough flux from its host to keep the planet irradiated beyond the insolation threshold identified by Demory & Seager (2011), except perhaps during the pre-main-sequence (prior to an age of ~ 100 Myr).

Interestingly, the currently rapid evolution of the star suggests a concomitant rapid in-spiral of the planet over the next few 100 Myr, unless the stellar Q'_* is large. This planet therefore does not appear destined to survive beyond the star’s current subgiant phase. As additional systems like KELT-8b are discovered and their evolution investigated in detail, it will be interesting to examine the statistics of planet survival and to compare these to predictions such as those shown here in Figure 3.14 to constrain mechanisms of planet-star interaction generally and the values of Q'_* specifically.

3.7 Summary

We announce the discovery of the highly inflated hot Jupiter, KELT-8b. This planet was initially discovered in KELT photometry, then confirmed via high-precision followup light curves and RVs. We also present adaptive optics imaging of KELT-8 from the Robo-AO system. Astrometry measurements over the past 110 years combined with high resolution spectra of both components firmly establishes the visual companion at $8.^{\circ}8$ separation as an unrelated background star.

We develop a new technique to extract reliable RVs from noisy data that saves a significant amount of telescope time. In the case of RV follow-up for the confirmation or mass measurements of transiting planets where the ephemeris is known and only a few well-timed RV measurements are needed this technique can save 30-50% of the time by avoiding the need to collect a high quality iodine-free template observation. In addition, lower S/N ratio is required for the RV measurement observations ($S/N \approx 50$) through the iodine cell because one component of our model is now completely noise free. Normally, we collect $S/N=200$ for our precision RV exposures using the standard observed template technique.

KELT-8b has one of the largest radii of any known transiting planet with $R_P = 1.86_{-0.16}^{+0.18} R_J$.

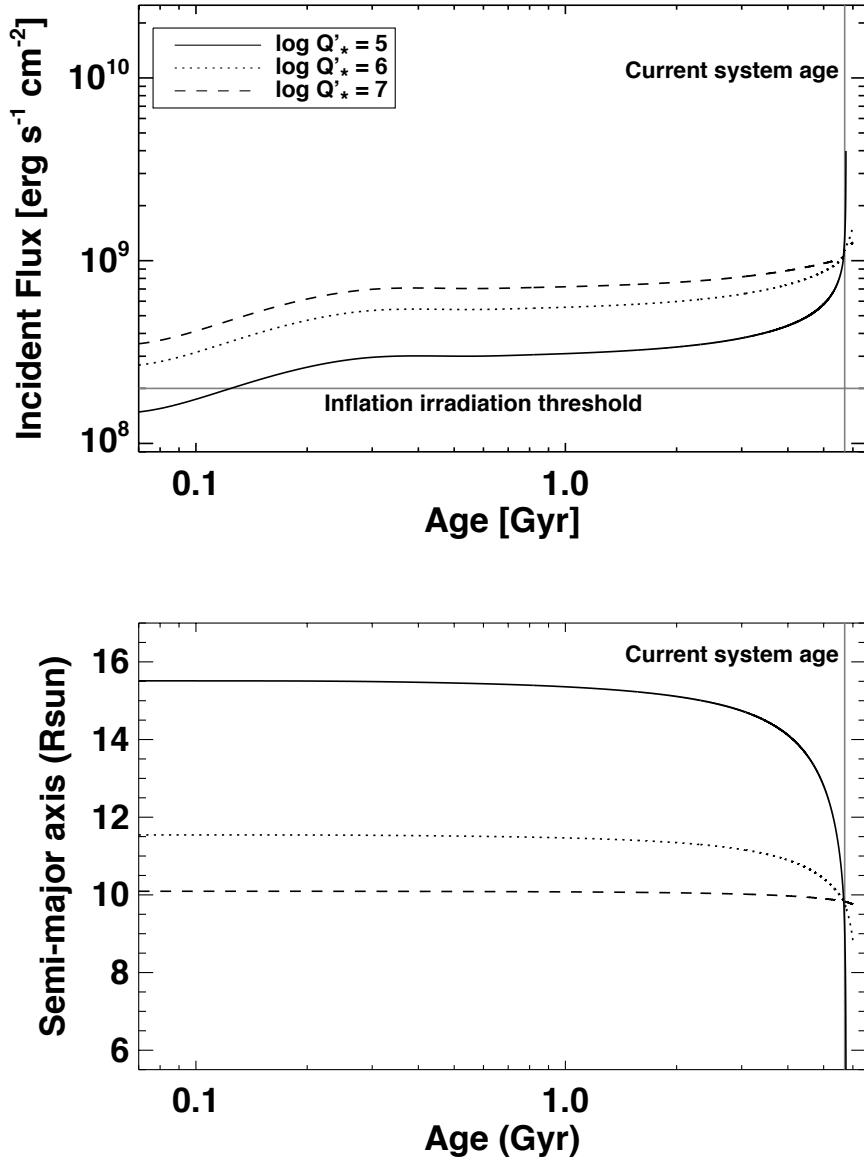


Figure 3.14 *Top:* Irradiation history of KELT-8b. The insolation received by the planet is well above the empirical inflation irradiation threshold (Demory & Seager, 2011) for the entire main-sequence existence of the star except in the case of $\log Q'_\star = 5$ in the very early stages of stellar evolution. *Bottom:* Orbital semi-major axis history of KELT-8b. The planet's semi-major axis is rapidly decreasing as the star evolves off the main sequence. It appears unlikely that KELT-8b will survive past the star's current subgiant phase.

It joins a small and interesting class of highly inflated hot Jupiters orbiting stars slightly more massive than the Sun.

CHAPTER 4

THREE SUPER-EARTHS ORBITING HD 7924

This chapter is a reproduction of Fulton et al. (2015b) included with permission from AAS journals.

4.1 Introduction

The archetypal planets of our Solar System—Jupiter the gas giant, Neptune the “ice” giant, and Earth the terrestrial planet—represent an incomplete inventory of the planet types in our galaxy. We are locally impoverished in “super-Earths,” the broad category of planets intermediate in size and mass between Earth and Neptune. Doppler searches of nearby stars showed that super-Earth planets in close-in orbits are plentiful (Howard et al., 2010a; Mayor et al., 2011). Results from the Kepler mission confirmed and refined our knowledge of the size and orbital period distribution of these and other planets (Howard et al., 2012b; Petigura et al., 2013a; Fressin et al., 2013). These planets have a wide range of bulk densities (Marcy et al., 2014), suggesting a diversity of compositions spanning rocky planets with negligible atmospheres (Howard et al., 2013; Pepe et al., 2013) to puffy planets with thick gas envelopes (Kipping et al., 2014). Intermediate planets with densities of $\sim 3 \text{ g cm}^{-3}$ are consistent with a broad range of interior structures and atmosphere sizes. Planets smaller than ~ 1.6 Earth radii (R_{\oplus}) are more likely to have a high density and presumed rocky composition (Weiss & Marcy, 2014; Rogers, 2014).

The large population of super-Earths orbiting close to their host stars was a surprise. Population synthesis models of planet formation had predicted that such systems would be rare (Ida & Lin, 2004; Mordasini et al., 2009a). Planet cores were expected to mostly form beyond the ice line and rarely migrate to close orbits unless they first grew to become gas giants. Nevertheless, close-in, low-mass planets are common and often appear in compact multi-planet systems (Lissauer et al., 2011; Fang & Margot, 2012). Theoretical models are catching up, with refinements to the disk migration and multi-planet dynamics in the population synthesis family of models (Ida & Lin, 2010; Alibert et al., 2013; Schlichting, 2014; Lee et al., 2014). A new class of “*in situ*” formation models have also been proposed in which systems of super-Earths and Neptunes emerge naturally from massive disks (Hansen & Murray, 2012; Chiang & Laughlin, 2013).

The Eta-Earth Survey (Howard et al., 2010a) at Keck Observatory played a important role in the discovery that super-Earths are abundant. Using the HIRES spectrometer, our team searched for planets in a volume-limited sample of 166 nearby G and K dwarf stars. Our search yielded new planets (Howard et al., 2009, 2011a,c, 2014) and detection limits for each star. Putting these together, we measured the prevalence of planets in close-in orbits as a function of planet mass ($M \sin i$). This mass function rises steeply with decreasing mass: planets in the mass range 3–10 M_{\oplus} are about twice as common as 10–30 M_{\oplus} planets.

The first low-mass planet discovered in the Eta-Earth Survey was HD 7924b (Howard et al., 2009, H09 hereafter), a super-Earth with a mass of $8.7 M_{\oplus}$ and an orbital period of 5.4 days. We have continued to observe HD 7924 and other Eta-Earth Survey stars using HIRES. These additional measurements probe smaller masses and larger star-planet separations. We have also started observing a subset of the Eta-Earth Survey stars with the Automated Planet Finder (APF; Vogt et al., 2014b), a new telescope at Lick Observatory. APF is a robotic 2.4-m telescope designed exclusively for Doppler discovery of exoplanets. It feeds the high-resolution Levy Spectrometer (Radovan et al., 2010) that uses an iodine reference spectrum to calibrate the wavelength scale and point spread function (Butler et al., 1996b), achieving a Doppler precision similar to HIRES while running without human intervention during an observing night. APF exploits high measurement cadence (nearly nightly) to disentangle the complicated low-amplitude signals of multi-planet systems in the face of stellar activity.

In this paper we announce two additional super-Earths orbiting HD 7924 based on RVs from the Keck-HIRES and APF-Levy spectrometers. We describe the properties of the star HD 7924 in Sec. 5.3 and our Doppler measurements from APF/Levy and Keck/HIRES in Sec. 5.2. Our analysis of the RV data, including discovery of the three Keplerian signals and consideration of false alarm probabilities, alias periods, and chromospheric activity, is described in Sec. 4.4. We conclude with a discussion and summary in Sec. 8.5.

4.2 Stellar Properties

HD 7924, also known as HIP 6379 or GJ 56.5, is a nearby (16.82 pc; van Leeuwen, 2007) and bright K0.5V dwarf star (von Braun et al., 2014). It is slightly metal poor relative to the Sun and hosts one previously known planet. It is relatively inactive with $\log R'_{HK} = -4.89$ (Isaacson & Fischer, 2010), but we do detect some low-level chromospheric activity nonetheless (see Section 5.4.6). We list our adopted stellar parameters in Table 5.2.

Most of our spectroscopically-derived stellar parameters are adopted from H09, which were originally derived by Valenti & Fischer (2005) using the Spectroscopy Made Easy (SME) LTE spectral synthesis code. However, HD 7924 has been the focus of several studies since the discovery of HD 7924b. Santos et al. (2013) performed a uniform analysis of 48 planet-hosting stars. They find $T_{\text{eff}} = 5133 \pm 68$ K, $\log g = 4.46 \pm 0.12$, and $[\text{Fe}/\text{H}] = -0.22 \pm 0.04$ for this star. von Braun et al. (2014) use the empirical relations of Boyajian et al. (2012) to calculate a mass of $0.81 M_{\odot}$ for HD 7924 with an error estimate of 30%. We computed the stellar mass and radius from T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$ using the Torres et al. (2010) relations and found $0.81 \pm 0.02 M_{\odot}$ and $0.75 \pm 0.03 R_{\odot}$. All of these values are within $1-\sigma$ of our adopted values.

HD 7924 was observed by von Braun et al. (2014) using long-baseline interferometry on the Georgia State University Center for High Angular Resolution Astronomy (CHARA) Array (ten Brummelaar et al., 2005a). We adopt their value for T_{eff} of 5075 ± 83 K that they obtained

Table 4.1. Adopted Stellar Properties of HD 7924

Parameter	HD 7924	Source
Spectral type	K0.5V	von Braun et al. (2014)
$B - V$ (mag)	0.826	H09
V (mag)	7.185	H09
J (mag)	5.618 ± 0.026	Cutri et al. (2003)
H (mag)	5.231 ± 0.033	Cutri et al. (2003)
K (mag)	5.159 ± 0.020	Cutri et al. (2003)
Distance (pc)	16.82 ± 0.13	van Leeuwen (2007)
T_{eff} (K)	5075 ± 83	von Braun et al. (2014)
$\log g$ (cgs)	$4.59^{+0.02}_{-0.03}$	H09
[Fe/H] (dex)	-0.15 ± 0.03	H09
$v \sin i$ (km s^{-1})	1.35 ± 0.5	H09
L_{\star} (L_{\odot})	0.3648 ± 0.0077	von Braun et al. (2014)
M_{\star} (M_{\odot})	$0.832^{+0.022}_{-0.036}$	Takeda et al. (2007)
R_{\star} (R_{\odot})	0.7821 ± 0.0258	von Braun et al. (2014)
$\log R'_{\text{HK}}$	-4.89	Isaacson & Fischer (2010)
S_{HK}	0.20	Isaacson & Fischer (2010)

by fitting the spectral energy distribution from optical through infrared wavelengths. They also obtained precise values for the luminosity and radius that we list in Table 5.2. These values are consistent within $1 - \sigma$ with the values adopted in H09 that were based on fitting stellar evolution models with spectroscopic parameters. Mason et al. (2011) conducted high-resolution speckle imaging of HD 7924 and detected no companions within three V magnitudes of HD 7924 with separations between $0''.03$ and $1''.5$.

4.3 Measurements

4.3.1 Keck/HIRES Spectroscopy

We collected 599 new RV measurements for HD 7924 over the last 5 years since the discovery of HD 7924b (H09) using the HIRES spectrograph on the Keck I telescope (Vogt et al., 1994). When this new data is combined with the data from H09 and APF we have over 10 years of observational baseline (see Figure 4.1a). Our data collection and reduction techniques are described in detail in H09. On Keck/HIRES we observe the star through a cell of gaseous iodine in order to simultaneously forward model the instrumental line broadening function (PSF) and the subtle shifts of the stellar lines relative to the forest of iodine lines. The HIRES detector was upgraded in August of 2004 and the RV zero-point between the pre-and post-upgrade data may not necessarily be the same. For this reason, we allow for separate RV zero-points for the pre-upgrade, post-upgrade, and APF

data sets. All RV measurements and associated Ca 2 H & K S_{HK} activity indices (the ratio of the flux in the cores of the H & K lines to neighboring continuum levels) are listed in Table 5.1.

4.3.2 APF/Levy Spectroscopy

The Automated Planet Finder is a 2.4 m f/15 Cassegrain telescope built by Electro-Optical Systems Technologies housed in an IceStorm-2 dome located at Lick Observatory atop Mount Hamilton, 20 miles east of San Jose, California. The telescope operates completely unattended using a collection of Python, Tcl, bash, and csh scripts that interact with the lower level software operating on the Keck Task Library keyword system (Lupton & Conrad, 1993). Every scheduled observing night, our automation software queries an online Google spreadsheet (that serves as our target database) and creates two observing plans: one for good conditions and the other for poor. Weather permitting, shortly after sunset the observatory opens automatically and determines whether to use the observing plan designed for good or poor conditions by monitoring a bright star. Once the observing plan is started, the high-level software continues to monitor the conditions and adjusts the schedule or switches observing plans if necessary. Observations of our targets (primarily for radial velocity measurements) continue throughout the night until conditions deteriorate too much to remain open or the until morning 9-degree twilight.

The Levy Spectrograph is a high-resolution slit-fed optical echelle spectrograph mounted at one of the two Nasmyth foci of the APF designed specifically for the detection and characterization of exoplanets (Vogt et al., 2014a; Radovan et al., 2014; Burt et al., 2014). Each spectrum covers a continuous wavelength range from 3740 Å to 9700 Å. We observed HD 7924 using a 1''.0 wide decker for an approximate spectral resolution of $R = 100,000$. Starlight passes through a cell of gaseous iodine that serves as a simultaneous calibration source for the instrumental point spread function (PSF) and wavelength reference. In addition, we collected a high signal-to-noise spectrum through the 0''.5 wide decker ($R = 150,000$) with the iodine cell out of the light path. This spectrum serves as a template from which we measure the relative doppler shifts of the stellar absorption lines with respect to the iodine lines while simultaneously modeling the PSF and wavelength scale of each spectrum. APF guides on an image of the star before the slit using a slanted, uncoated glass plate that deflects 4% of the light to a guide camera. Photon-weighted times of mid-exposure are recorded using a software-based exposure meter based on guide camera images that monitors the sky-subtracted light entering the slit during an exposure (Kibrick et al., 2006).

We measure relative radial velocities (RVs) using a Doppler pipeline descended from the iodine technique in Butler et al. (1996). For the APF, we forward-model 848 segments of each spectrum between 5000-6200 Å. The model consists of a stellar template spectrum, an ultra high-resolution Fourier transform spectrum of the iodine absorption of the Levy cell, a spatially-variable PSF, a wavelength solution, and RV. We estimate RV uncertainties (Table 5.1) as the uncertainty on the mean RV from the large number of spectral segments. Well-established RV standard stars

have a RV scatter of \sim 2–3 m s $^{-1}$ based on APF measurements. This long-term scatter represents a combination of photo-limited uncertainties, stellar jitter, and instrumental systematics. We collected a total of 109 RV measurements of HD 7924 on 80 separate nights (post outlier rejection) over a baseline of 1.3 years with a typical per-measurement uncertainty of 2.0 m s $^{-1}$.

For both the APF and Keck data, Julian dates of the photon-weighted exposure mid-times were recorded during the observations, then later converted to Barycentric Julian date in the dynamical time system (BJD_{TDB}) using the tools of Eastman et al. (2010)¹.

We rejected a handful of low signal-to-noise ratio spectra ($S/N < 70$ per pixel) and measurements with uncertainties greater than nine times the median absolute deviation of all measurement uncertainties relative to the median uncertainty for each instrument. This removed a total of 11 RVs out of the 906 total measurements in the combined data set (Keck plus APF).

After outlier rejection we bin together any velocities taken less than 0.5 days apart on a single telescope. Since data taken in short succession are likely affected by the same systematic errors (e.g. spectrograph defocus), these measurements are not truly independent. Binning helps to reduce the effects of time-correlated noise by preventing multiple measurements plagued by the same systematic errors from being given too much weight in the Keplerian analysis. When the data are binned together and the uncertainty for the resulting data point is added in quadrature with the stellar jitter (an additional error term that accounts for both stellar and instrumental systematic noise), the binned data point receives only as much weight as a single measurement. While this likely reduces our sensitivity slightly we accept this as a tradeoff for more well-behaved errors and smoother χ^2 surfaces. An independent analysis using the unbinned data in Section 4.4.4 finds three planets having the same orbital periods, eccentricities, and masses within $1-\sigma$ as those discovered by analyzing the binned data.

4.4 Keplerian Analysis

4.4.1 Discovery

We identify significant periodic signals in the RVs using an iterative multi-planet detection algorithm based on the two-dimensional Keplerian Lomb-Scargle (2DKLS) periodogram (O’Toole et al., 2009). Instead of fitting sinusoidal functions to the RV time series (Lomb, 1976; Scargle, 1982), we create a periodogram by fitting the RV data with Keplerian orbits at many different starting points on a 2D grid over orbital period and eccentricity. This technique allows for relative offsets and uncertainties between different data sets to be incorporated directly into the periodogram and enhances the sensitivity to moderate and high eccentricity planets.

We fit the Keplerian models using the Levenberg-Marquardt (L-M) χ^2 -minimization routine in the RVLIN IDL package (Wright & Howard, 2009). Multi-planet models are sums of single-planet

¹IDL tools for time systems conversion; <http://astroutils.astronomy.ohio-state.edu/time/>.

Table 4.2. Radial Velocities of HD 7924[†]

BJD _{TDB} (– 2440000)	RV (m s ^{−1})	Uncertainty (m s ^{−1})	Instrument ¹	<i>S</i> _{HK}
12307.77162	5.23	1.31	k	...
12535.95639	3.24	1.16	k	...
13239.08220	-3.00	0.91	j	0.218
13338.79766	3.83	1.08	j	0.227
16505.99731	-7.97	1.46	a	...
16515.90636	3.88	1.63	a	...

[†](This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

¹k = pre-upgrade Keck/HIRES, j = post-upgrade Keck/HIRES,
a = APF

models, with planet-planet gravitational interactions neglected. Such an approximation is valid since the interaction terms are expected to be $\ll 1$ m s^{−1} for non-resonant, small planets. We define a grid of search periods following the prescription of Horne & Baliunas (1986) and at each period we seed an L-M fit at five evenly-spaced eccentricity values between 0.05 and 0.7. Period and eccentricity are constrained to intervals that allow them to vary only half the distance to adjacent search periods and eccentricities. All other model parameters are free to vary, including the parameters of any previously identified planets. The period and eccentricity for previously identified planets are constrained to be within $\pm 5\%$ and $^{+5}_{-10}\%$, respectively, of their initial values but all other parameters are unconstrained. This prevents slightly incorrect fits of the first detected planets from injecting periodic residuals that could mimic further planetary signals. The 2DKLS periodogram power at each point in the grid is

$$Z(P, e) = \frac{\chi^2 - \chi_B^2}{\chi_B^2}, \quad (4.1)$$

where χ^2 is the sum of the squared residuals to the current $N + 1$ planet fit, and χ_B^2 is the sum of the squared residuals to the best N -planet fit. In the first iteration of the planet search (comparing a 1-planet model to a 0-planet model), χ_B^2 is the sum of the squared error-normalized residuals to the mean (the *B* subscript stands for baseline). The 2D periodogram is collapsed into a 1D periodogram as a function of period by taking the maximal *Z* for each period searched (i.e. the best fit eccentricity for every period).

We start the iterative planet search by comparing a 0-planet model (flat line) to a grid of 1-planet models. A strong signal at a period of 5.4 days is detected at very high significance in the

first iteration. This planet was initially discovered by H09 using only 22% of the data presented in this work. When we calculate two vs. one planet and three vs. two planet periodograms, we find two additional highly significant signals with periods of 15.3 days and 24.5 days respectively (see Figure 4.1). These two periodic signals are best fit by Keplerian orbital models with semi-amplitudes of 2.3 m s^{-1} and 1.7 m s^{-1} , respectively, with no significant eccentricity. See Tables 4.3 and 5.6 for the full orbital solution.

We find a fourth, significant signal at ~ 2400 days, but we also find a very similar signal (approximately sinusoidal with the same period and phase) upon inspection of the time series of S_{HK} stellar activity measurements (see Figure 5.8). Although we fit for this periodic signal as an additional Keplerian we do not interpret this as an additional planet. Instead we interpret this as the signature of the stellar magnetic activity cycle. We searched for a fifth periodic signal and found two additional marginally significant peaks at 17.1 days and 40.8 days. Due to their marginal strengths and the fact that the two periods are related by the synodic month alias ($1/17.1 \approx 1/40.8 + 1/29.5$) we are especially cautious in the interpretation of these signals. We scrutinize these candidate signals in depth in Sections 4.4.4 and 5.4.6. We also see a peak at around 40 days in a periodogram of the S_{HK} time series after the long-period signal from the stellar magnetic activity cycle is removed that is presumably the signature of rotationally modulated star spots. We conclude that the 17.1 and 40.8 day signals are most likely caused by rotational modulation of starspots, but the 5.4, 15.3, and 24.5 day signals are caused by three planetary companions with minimum ($M_p \sin i_p$) masses of $8.7 M_{\oplus}$, $7.9 M_{\oplus}$, and $6.4 M_{\oplus}$.

4.4.2 Characterization

We determine the orbital parameters and associated uncertainties for the three planet system using the ExoPy Differential-Evolution Markov Chain Monte Carlo (DE-MCMC, Ter Braak, 2006) engine described in Fulton et al. (2013) and Knutson et al. (2014). We treat the total RV model as a sum of Keplerian orbits each parameterized by orbital period (P_i), time of inferior conjunction ($T_{\text{conj},i}$), eccentricity (e_i), argument of periastron of the star’s orbit (ω_i), and velocity semi-amplitude (K_i) where i is an index corresponding to each planet (b – d). An RV “jitter” term (σ_{jitt}) is added in quadrature with the measurement uncertainties at each step in the MCMC chains. We also fit for independent RV zero-points for the APF, pre-upgrade Keck, and post-upgrade Keck data. The long-period RV signal presumably caused by the stellar magnetic activity cycle is treated as an additional Keplerian orbit with the same free parameters as for each of the three planets. In order to speed convergence and avoid biasing parameters that must physically be finite and positive we step in the transformed and/or combinations of parameters listed in Table 4.3. We add a χ^2 penalty for large jitter values of the following form

$$\chi^2_{\text{new}} = \chi^2 + 2 \sum_n \ln \sqrt{2\pi(\sigma_{\text{vel},n}^2 + \sigma_{\text{jitt}}^2)}, \quad (4.2)$$

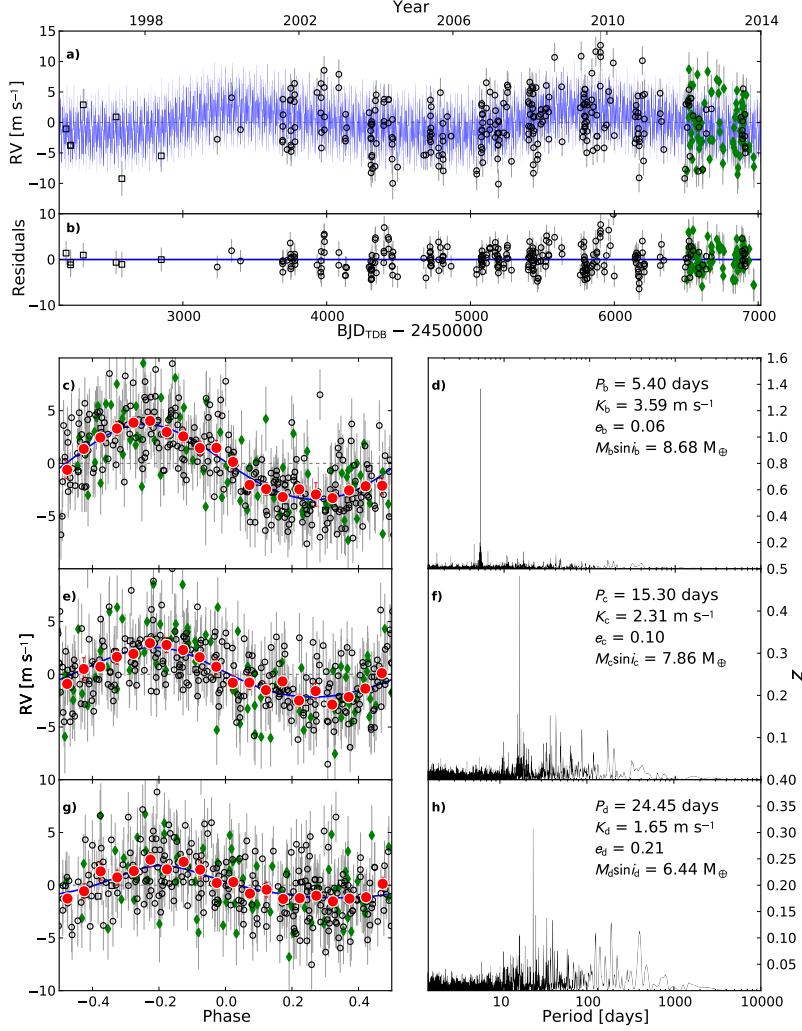


Figure 4.1 Best-fit 3-planet Keplerian orbital model plus one additional long-period Keplerian to model the stellar magnetic activity cycle. The model plotted is the one that produces the lowest χ^2 while the orbital parameters annotated and listed in Tables 4.3 and 5.6 are the median values of the posterior distributions. **a)** Full binned RV time series. Open black squares indicate pre-upgrade Keck/HIRES data (see §5.2), open black circles are post-upgrade Keck/HIRES data, and filled green diamonds are APF data. The thin blue line is the best fit 3-planet plus stellar activity model. We add in quadrature the RV jitter term listed in Table 4.3 with the measurement uncertainties for all RVs. **b)** Residuals to the best fit 3-planet plus stellar activity model. **c)** Binned RVs phase-folded to the ephemeris of planet b. The two other planets and the long-period stellar activity signal have been subtracted. The small point colors and symbols are the same as in panel **a**. For visual clarity, we also bin the velocities in 0.05 units of orbital phase (red circles). The phase-folded model for planet b is shown as the blue line. **d)** 2DKLS periodogram comparing a 2-planet plus activity model to the full 3-planet fit when planet b is included. Panels **e)** and **f)**, and panels **g)** and **h)** are the same as panels **c)** and **d)** but for planets HD 7924 c and HD 7924 d respectively.

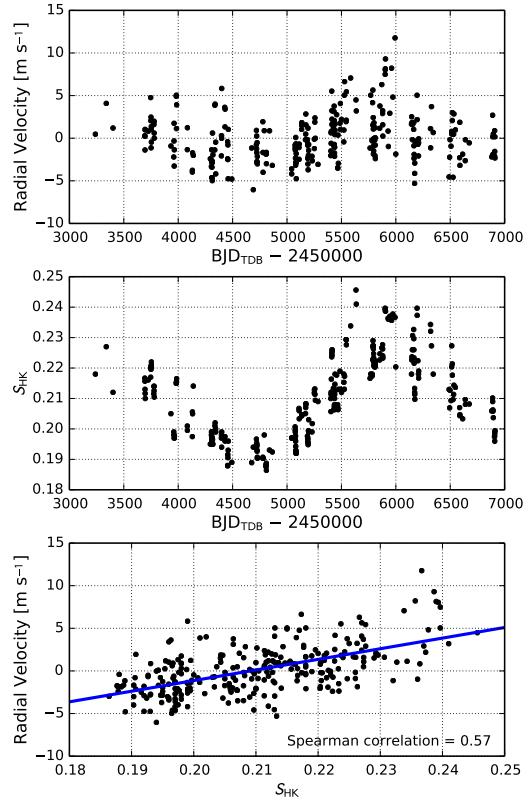


Figure 4.2 Velocity-activity correlation. *Top*: Binned RV time series of the post-upgrade Keck data with planets b, c, and d subtracted. *Middle*: Binned S_{HK} time series of the post-upgrade Keck data only. Note the similarities between the variability in the top and middle panels. *Bottom*: Spearman rank correlation test of the velocities with S_{HK} values (Spearman, 1904).

Table 4.3. Orbital Parameters

Parameter	Value	Units
Modified DE-MCMC Step Parameters¹		
$\log(P_b)$	$0.73223 \pm 2e - 05$	$\log(\text{days})$
$\sqrt{e_b} \cos \omega_b$	$0.15^{+0.13}_{-0.17}$	
$\sqrt{e_b} \sin \omega_b$	$-0.09^{+0.18}_{-0.15}$	
$\log(K_b)$	$0.555^{+0.023}_{-0.024}$	m s^{-1}
$\log(P_c)$	$1.184663^{+9.2e-05}_{-9.3e-05}$	$\log(\text{days})$
$\sqrt{e_c} \cos \omega_c$	$0.20^{+0.17}_{-0.24}$	
$\sqrt{e_c} \sin \omega_c$	$0.11^{+0.17}_{-0.20}$	
$\log(K_c)$	$0.364^{+0.037}_{-0.040}$	m s^{-1}
$\log(P_d)$	$1.3883^{+0.00027}_{-0.00031}$	$\log(\text{days})$
$\sqrt{e_d} \cos \omega_d$	$0.31^{+0.16}_{-0.24}$	
$\sqrt{e_d} \sin \omega_d$	$0.02^{+0.30}_{-0.37}$	
$\log(K_d)$	$0.219^{+0.052}_{-0.057}$	m s^{-1}
Model Parameters		
P_b	5.39792 ± 0.00025	days
$T_{\text{conj},b}$	$2455586.38^{+0.086}_{-0.110}$	BJD_{TDB}
e_b	$0.058^{+0.056}_{-0.040}$	
ω_b	332^{+71}_{-50}	degrees
K_b	$3.59^{+0.20}_{-0.19}$	m s^{-1}
P_c	$15.299^{+0.0032}_{-0.0033}$	days
$T_{\text{conj},c}$	$2455586.29^{+0.40}_{-0.47}$	BJD_{TDB}
e_c	$0.098^{+0.096}_{-0.069}$	
ω_c	27^{+52}_{-60}	degrees
K_c	$2.31^{+0.21}_{-0.20}$	m s^{-1}
P_d	$24.451^{+0.015}_{-0.017}$	days
$T_{\text{conj},d}$	$2455579.1^{+1.0}_{-0.9}$	BJD_{TDB}
e_d	$0.21^{+0.13}_{-0.12}$	
ω_d	119^{+210}_{-97}	degrees
K_d	1.65 ± 0.21	m s^{-1}
$\gamma_{\text{post-upgrade Keck}}$	-0.19 ± 0.16	m s^{-1}
$\gamma_{\text{pre-upgrade Keck}}$	$2.0^{+1.1}_{-1.2}$	m s^{-1}
γ_{APF}	$0.28^{+0.46}_{-0.47}$	m s^{-1}
σ_{jitt}	$2.41^{+0.11}_{-0.10}$	m s^{-1}

¹MCMC jump parameters that were modified from the physical parameters in order to speed convergence and avoid biasing parameters that must physically be finite and positive.

Table 4.4. Derived Properties

Parameter	Value	Units
$e_b \cos \omega_b$	0.0073 $^{+0.0210}_{-0.0076}$	
$e_b \sin \omega_b$	-0.0031 $^{+0.0063}_{-0.0190}$	
a_b	0.05664 $^{+0.00067}_{-0.00069}$	AU
$M_b \sin i_b$	8.68 $^{+0.52}_{-0.51}$	M_{\oplus}
S_b^*	113.7 $^{+3.7}_{-3.6}$	S_{\oplus}
$T_{eq,b}^{**}$	825.9 $^{+6.6}_{-6.5}$	K
$e_c \cos \omega_c$	0.017 $^{+0.051}_{-0.018}$	
$e_c \sin \omega_c$	0.008 $^{+0.037}_{-0.013}$	
a_c	0.1134 $^{+0.0013}_{-0.0014}$	AU
$M_c \sin i_c$	7.86 $^{+0.73}_{-0.71}$	M_{\oplus}
S_c^*	28.35 $^{+0.92}_{-0.89}$	S_{\oplus}
$T_{eq,c}^{**}$	583.6 $^{+4.7}_{-4.6}$	K
$e_d \cos \omega_d$	0.059 $^{+0.084}_{-0.054}$	
$e_d \sin \omega_d$	0.001 $^{+0.076}_{-0.074}$	
a_d	0.1551 $^{+0.0018}_{-0.0019}$	AU
$M_d \sin i_d$	6.44 $^{+0.79}_{-0.78}$	M_{\oplus}
S_d^*	15.17 $^{+0.49}_{-0.48}$	S_{\oplus}
$T_{eq,d}^{**}$	499 ± 4	K

*Stellar irradiance received at the planet relative to the Earth.

**Assuming a bond albedo of 0.32; the mean total albedo of super-Earth size planets (Demory, 2014).

where $\sigma_{\text{vel},n}$ is the velocity uncertainty for each of the n measurements (Johnson et al., 2011).

We fit for the 23 free parameters by running 46 chains in parallel continuously checking for convergence using the prescription of Eastman et al. (2013). When the number of independent draws (T_z as defined by Ford, 2006) is greater than 1000 and the Gelman-Rubin statistic (Gelman et al., 2003; Holman et al., 2006) is within 1% of unity for all free step parameters we halt the fitting process and compile the results in Table 4.3. We also list additional derived properties of the system in Table 5.6 that depend on the stellar properties listed in Table 5.2.

4.4.3 False Alarm Assessment

We attempted to empirically determine the probability that Gaussian random noise in the data could conspire to produce an apparent periodic signal with similar significance to the periodogram peaks corresponding to each of the planets. We calculated 1000 2DKLS periodograms, each time scrambling the velocities in a random order drawn from a uniform distribution. We located and measured the height of the global maxima of each periodogram and compare the distribution of these maxima to the periodogram peak heights at the periods of the three planets in the original periodograms. The power for each periodogram within the set of 1000 is a $\Delta\chi^2$ between a 2-planet plus activity model to a 3-planet model assuming the other two planets as “known”. Figure 5.6 shows that the distribution of maxima from the periodograms of the scrambled data are clearly separated from the original peaks. None of the trials produce periodogram peaks anywhere near the heights of the original peaks corresponding to the three planets. Because the 2DKLS periodogram allows eccentric solutions, we explore the scrambled RVs for non-sinusoidal solutions and are therefore sensitive to a wide variety of false alarm signals. We conclude that the false alarm probabilities for all three planets are < 0.001 .

4.4.4 Searching for Period Aliases

The high cadence of the APF data set allows us to explore short-period orbital solutions. Traditional observations on large telescopes such as Keck often yield only a few nights of data per year, making it difficult to determine whether a short orbital period is an alias of a longer period, or a true physical signal. Although planets with orbital periods shorter than one day are uncommon in the galaxy (around $0.83 \pm 0.18\%$ of K dwarfs; Sanchis-Ojeda et al., 2014), eleven ultra-short period planets have been found² due to their high detectability. To search as carefully as possible for short period signals, we use the unbinned data sets, which consist of 797 RVs from Keck and 109 RVs from the APF. Our use of the unbinned data in this section also provides independent confirmation of the results obtained with the binned data above.

Dawson & Fabrycky (2010) outline a rigorous procedure to distinguish between physical and alias periods. Our method for finding the orbital periods and distinguishing aliases is as follows:

²Based on a 2014 Nov 20 query of exoplanets.org (Wright et al., 2011; Han et al., 2014)

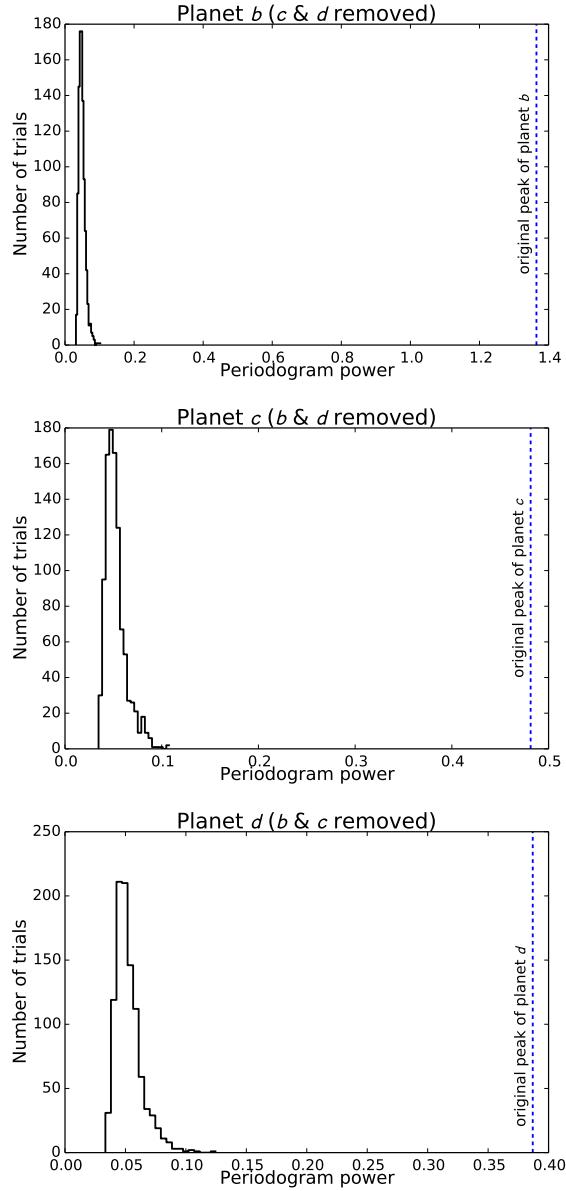


Figure 4.3 *Top:* Distribution of maximum periodogram peak heights for 1000 2DKLS periodograms of scrambled RV time series. For each periodogram planets c, d, and the magnetic activity cycle were subtracted before scrambling the data set. The vertical dashed blue line marks the height of the original peak for planet b which is clearly separated from the distribution of peaks caused by random fluctuations. *Middle:* Same as the top panel for planet c. *Bottom:* Same as the top panel for planet d.

1. Determine the window function of the data to understand which aliases are likely to appear.
2. Compute the periodogram of the data, determining the power and phase at each input frequency.
3. If there is a strong peak in the periodogram, fit an N-planet Keplerian (starting with N=1), using the periodogram peak as the trial period.
4. Subtract the N-planet Keplerian from the data.
5. Compute the periodogram of the Nth planet in the model and compare it to the periodogram of the Nth planet in the data minus the model of the other planets.
6. If a second peak in the periodogram has similar height to the tallest peak and is located at an alias period, repeat steps 3-5 using that trial period.
7. If you explored an alias period, choose the model that minimizes χ^2 and best reproduces the observed periodogram. Subtract this model from the RVs.
8. Treat the residuals as the new data set and go back to step 2. Examine the residuals from the N-planet fit for additional planets, and continue until there are no more signals in the periodogram.

The window functions of the individual and combined Keck and APF RV time series are shown in Figure 4.4. The window function is given by

$$W(\nu) = \frac{1}{N} \sum_{j=1}^N \exp(-2\pi i \nu t_j), \quad (4.3)$$

where ν is the frequency in units of days $^{-1}$ and t_j is the time of the j th observation. The data sets are complementary: 109 RVs from APF over the past year are well-distributed over the months and the year, and so are only susceptible to the daily aliases, whereas the 797 RVs from Keck over the last decade are distributed in a way that gives some power to daily aliases as well as longer-period aliases. The power in the combined window function illustrates that we might be susceptible to daily aliases, and weak signals in the periodogram might even be susceptible to monthly or yearly aliases.

To take the periodogram of the time series, we use a version of `fasper` (Press & Rybicki, 1989) written for Python. We find the same peaks in the periodogram of the data and residuals at 5.4, 15.3, and 24.5 days that we interpret as planets. The periodograms of the data and periodograms of the Keplerian models associated with these periods are shown in Figure 4.5. Again, we recover a fourth peak at 2570 days that we attribute to long-term stellar activity due to the correlation

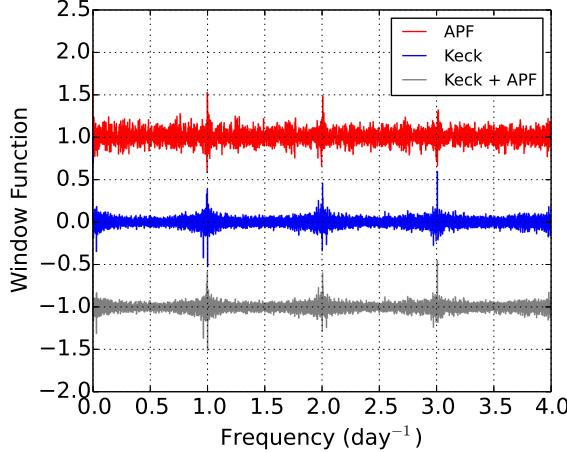


Figure 4.4 The window function of the Keck and APF RV time series. While the APF window function has some power at frequency multiples of one day, it is flat otherwise, whereas the Keck window function has power at low frequencies (corresponding to long-period aliases) and power in broader swaths around the frequency multiples of one day.

between the RVs and the S_{HK} values. The 3 planet plus stellar activity model found with the alias-search method yields periods, eccentricities, and planet masses within $1-\sigma$ of the results quoted in Table 4.3 for all three planets. ω_b is consistent within $1-\sigma$, but inconsistent for planets c and d. However, the arguments of periastron for all three planets are poorly constrained due to their nearly circular orbits. We find a fifth peak at 40.8 days, which is also prominent in the periodogram of the stellar activity and is likely the rotation period of the star. The 40.8-day signal has a strong alias at 17.1 days and so we test Keplerian models at both 40.8 days and 17.1 days to discriminate which is the true signal and which is the alias (see Figure 4.6). We find that the periodogram of the model 40.8 day signal better matches the alias structure in the periodogram of the data, and so we prefer 40.8 days as a candidate stellar rotation period.

4.4.5 Chromospheric Activity

Although this star is relatively inactive with R'_{HK} values in the literature between -4.89 (Isaacson & Fischer, 2010) and -4.85 (Canto Martins et al., 2011), some low-level chromospheric activity is detectable with our high-precision RVs. The most obvious feature is the long period signal with a period of ≈ 6.6 years and an amplitude of $\approx 5 \text{ m s}^{-1}$. Although this signal looks promising as a long-period sub-Jupiter mass planet candidate, upon inspection of the S_{HK} time series we notice that this activity indicator is highly correlated with this long-period signal in the velocities (see Figure 5.8). We interpret this signal as the signature of the stellar magnetic activity cycle of which we have observed nearly two full cycles. If we subtract this long-period signal from the S_{HK} values and make a periodogram of the residuals we find a marginally significant peak at ≈ 41 days. This

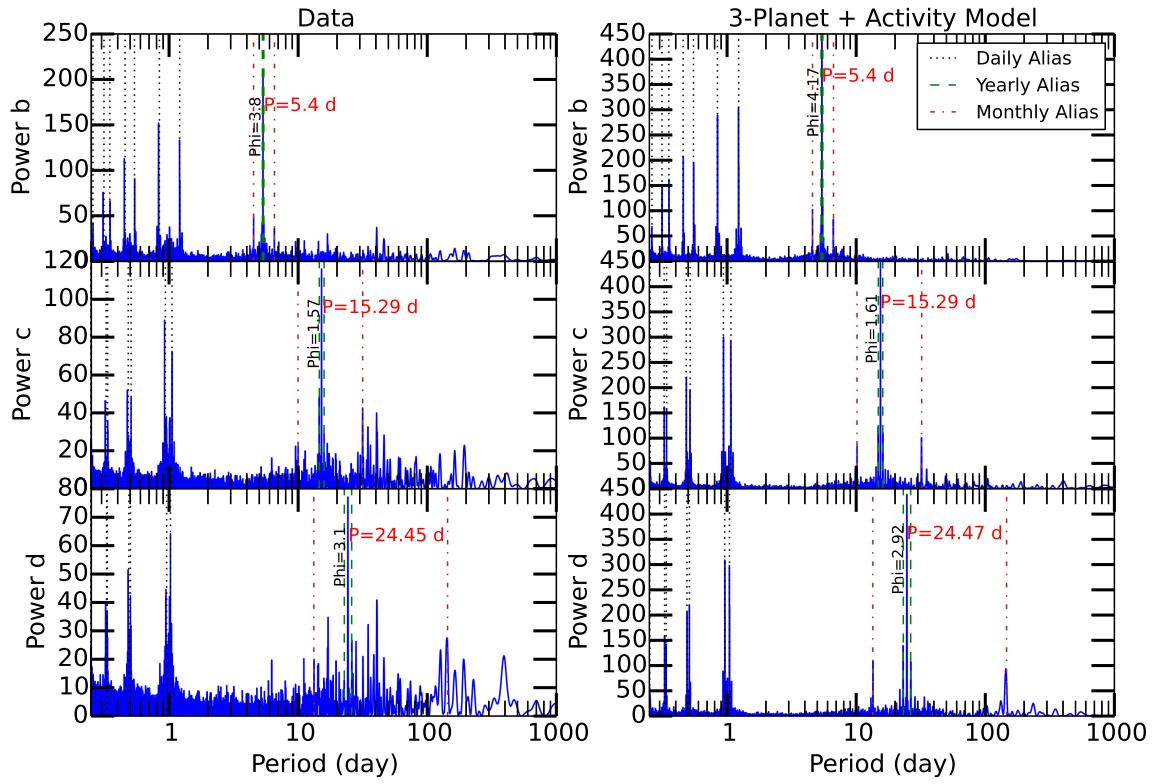


Figure 4.5 Left: LS periodograms of the data associated with each planet identified, from top to bottom: planet b, planet c, and planet d. In each panel, signals from the other planets and stellar activity have been subtracted. The phase of the frequency associated with the peak is given in radians. Right: LS periodograms of the best Keplerian model for each of the planets, from top to bottom, planets b, c, and d. The periodogram of each Keplerian model reproduces the peak period and alias structure of the data.

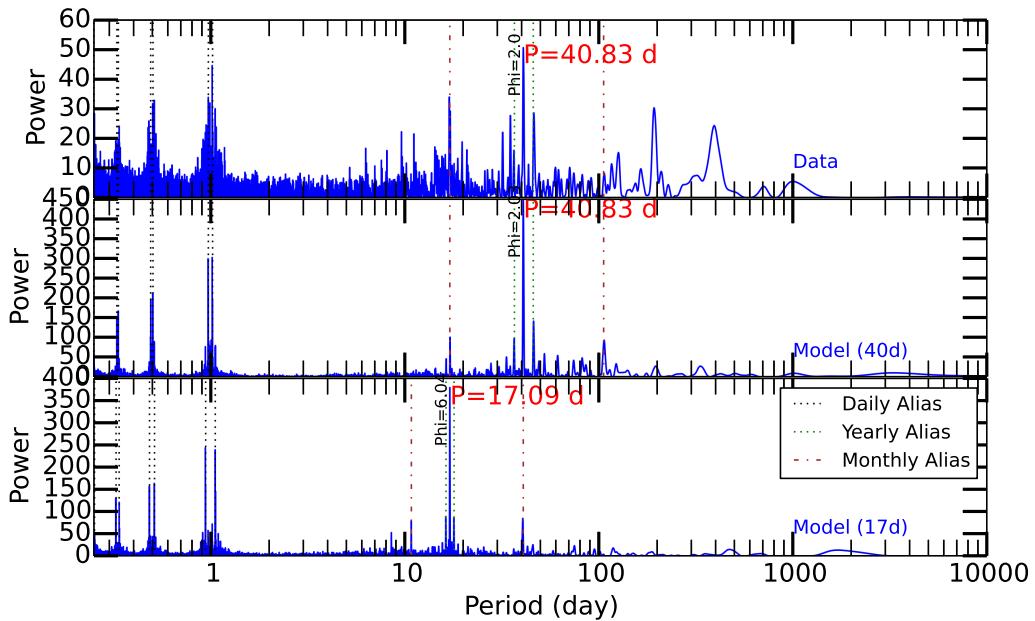


Figure 4.6 Top: a candidate periodic signal at 40.8 or 17.1 days emerges in the periodogram of the residuals to the 3-planet plus long-term stellar activity model. These two periods are related to each other by the one synodic month alias. Either 40.8 days or 17.1 days could correspond to the rotation period of the star, although the 40.8 day period is more prominent in the periodogram of the S_{HK} values. To test both periods, we model the best-fit Keplerian at 40.83 days (center) and 17.1 days (bottom) and show their periodograms, complete with phase information.

is most likely caused by rotational modulation of starspots, because 41 days is near the expected rotation period (38 days, Isaacson & Fischer, 2010) for a star of this spectral type and age.

To gain confidence in the source of these Doppler signals (activity or planets) we employed a time-dependent RV- S_{HK} decorrelation that probes evolution of the RV periodogram on timescales of the stellar rotation. Simple linear decorrelation is limited by the natural phase offset between RV and activity signals. At any given time, the measured RV and activity both depend on the flux-weighted fractional coverage of the magnetically active region on the stellar disk. The measured S_{HK} peaks when the magnetically active regions are closest to the disk center, where its projected area is largest and the star would otherwise be brightest. However, the RV approaches zero at the center of the stellar disk. The activity-induced RV shift is maximized when the active region is closer to the limb, $\sim 1/8$ phase from center (e.g. Queloz et al., 2001; Aigrain et al., 2012). In addition, the magnetically active regions continuously evolve on timescales of several stellar rotation periods weakening the RV- S_{HK} correlations over long timescales.

Motivated by these limitations, our alternative RV- S_{HK} decorrelation method is effective on timescales of both the long-term magnetic activity cycle and the stellar rotation period. All APF data are excluded from this analysis because activity measurements are not presently available.

Starting with the same outlier-removed time-series as described above, we bin together any radial velocity and S_{HK} measurements taken on the same night. The period of the long-term stellar magnetic activity cycle is then identified from a Lomb-Scargle (LS) periodogram of the activity time-series. That period serves as the initial guess in a subsequent single-Keplerian fit to the RV time-series using `rv_mp_fit`. Next, planet candidates are identified using an iterative approach in which the number of fitted planets N increases by one starting at $N = 1$. $N + 1$ Keplerians are fit at each step, one accounting for the long-term stellar magnetic activity cycle:

1. **Determine period of N_{th} planet candidate:** Identify the period, P_N , of the highest peak in the LS periodogram of the residuals to the N -planet Keplerian fit to the RV time-series.
2. **Obtain best N -planet solution:** Perform an $(N + 1)$ -planet Keplerian fit with initial guesses being the N Keplerians from the previous fit plus a Keplerian with initial period guess P_N .
3. **Repeat:** Steps 1-2 for $N = N + 1$.

We wish to subsequently remove residual RV signatures that can be confidently attributed to magnetically active regions rotating around the stellar surface. Since the lifespan of such regions is typically a few rotation periods, each measured RV was decorrelated with S_{HK} activity measurements within ± 100 days. We chose ± 100 day windows because such windows were long enough to allow us to detect the 41 day rotation period in each window, and short enough to capture long-term variability in the power spectrum on time-scales of hundreds of days. However, before discussing

our decorrelation technique, we investigate the degree to which our RV and S_{HK} measurements are correlated and the reliability of our planet candidate signals.

Figure 4.7 shows a time-series of LS periodograms corresponding to all RVs measured within ± 100 days. Each panel contains more than 100 individual columns corresponding to each periodogram derived from all data within ± 100 days of the times for each column. These periodograms are only displayed at times when a minimum data cadence requirement (discussed below) is met and are grayed out elsewhere. Individual periodograms have been normalized independently to facilitate inter-epoch comparison of the relative distribution of power across all periods. In each panel from top to bottom, an additional planet signal has been removed from the data. The long-term magnetic activity cycle is removed from the data in all three panels. The fitted longterm magnetic activity-induced component has also been removed in all cases. After removing planet b (top panel), the majority of epochs/columns show significant power at 15.3 days, corresponding to planet c. The few exceptions are epochs with the lowest cadence. The higher cadence epochs also have significant power near 24.5 days, corresponding to planet d. The 24.5-day signal is more apparent after planet c is also removed (middle panel). The presence of 15.3-day and 24.5-day peaks across all high-cadence epochs is consistent with the coherent RV signature of a planet.

Figure 4.7 also reveals significant RV power at ~ 30 – 50 -day periods which can be attributed to magnetically active regions (i.e. spots and plages) rotating around the stellar surface at the ~ 40.8 -day stellar rotation period. The activity-induced RV shifts dominate the periodograms after the planets b, c and d have been removed (bottom panel). A number of epochs also show substantial power at ~ 17.1 days. This is unlikely to be the signature of a fourth planet as the period corresponds to the monthly synodic alias of the 40.8-day rotation period. Moreover, a system of planets with 15.3-day and 17.1-day periods is unlikely to be dynamically stable.

The direct correlation between RV and stellar activity on stellar rotation timescales is further highlighted in Figure 4.8. The top panel is the S_{HK} -analog of Figure 4.7 in that each column is the running periodogram of all S_{HK} measurements within ± 100 days. The suite of 30–50-day peaks mirror those in the RV periodograms of Figure 4.7, indicating that magnetic activity is responsible for RV variation on these timescales.

The observed ~ 20 -day range in the period of these activity-induced peaks might seem surprisingly large, given that differential stellar rotation is likely of order several days (e.g. the observed difference in rotation period at the Sun’s equator and 60° latitude is ~ 5 days). However, some of this variation could also be statistical noise, simply an artifact of the limited sampling of ~ 40 -day signals over a relatively small time baseline. To test this hypothesis, we generated a synthetic RV time series of the exact same cadence as our real Keck observations and attempted to recover an injected signal of period 40.8 days. We first injected our three-planet model, adding Gaussian noise with standard deviation 2.36 m s^{-1} , the median velocity error in the post-upgrade Keck dataset (see Table 4.6). We then superposed the 40.8-day sinusoid of semi-amplitude 1.46 m s^{-1} . The cho-

sen semi-amplitude is the median of semi-amplitudes of all Keplerian fits to activity-induced RVs as determined by our decorrelation algorithm (see below). We applied the same planet-detection algorithm described above to these simulated data and subtracted three Keplerians from the best 4-Keplerian model, leaving only the 40.8-day signal.

The bottom plot in Figure 4.8 shows the resulting 200-day-running-window periodogram, akin to the bottom plot in Figure 4.7. The \sim 10–15 day fluctuation in the recovered period of the injected 40.8-day signal is consistent with activity-induced features in our real S_{HK} and RV data. We conclude that the varying power of the RV signal with periods of 30–50 days observed in the real data is a limitation of the observing cadence and not intrinsic to the star.

The strong correlation observed between RV and stellar activity indicates tremendous potential for decorrelation to validate our planet candidates. We apply a decorrelation algorithm that iterates through each RV measurement $v(t_i)$ where t_i is the time of the i^{th} measurement:

1. Define subsets S_i and V_i , consisting of all S_{HK} and RV measurements within 100 days of time t_i .
2. Define n_S and n_V as the number of datapoints in S_i and V_i respectively. If n_S and $n_V > n_{\min}$ then proceed to next step, otherwise revert to step 1 for $i = i + 1$.
3. Identify the period, P_i , of the highest peak in the LS periodogram of S_i .
4. Perform a single-Keplerian fit to V_i with fixed period P_i .
5. Subtract Keplerian fit from $v(t_i)$.
6. Repeat for $i = i + 1$.

As step 2 indicates, we only decorrelated an RV measurement with S_{HK} activity if both the number of S_{HK} and RV measurements within ± 100 days, n_S and n_V respectively, exceeded some minimum number, n_{\min} . This was done to avoid removal of spurious correlations. n_{\min} was chosen by examining the dependence of correlation significance on n_S and n_V . More specifically, for each RV measurement, we computed the Pearson product-moment correlation coefficient, r , between RV and S_{HK} periodograms from all data within ± 100 days. In cases of n_S and n_V between \sim 10–20, r values were widely distributed between 0–1, suggesting correlations were spurious. Supporting this notion, the best-fitting Keplerians to the activity-correlated RV noise had unusually large semi-amplitudes, as high as 10 m s^{-1} , even in cases where $r > 0.5$. In contrast, when n_S and n_V exceeded 23, r ranged from 0.34–0.76, with a median of 0.51 and semi-amplitudes were a more reasonable $1\text{--}3 \text{ m s}^{-1}$. We therefore adopted $n_{\min} = 23$ resulting in activity decorrelation of 115 of 281 RV measurements.

The top panel of Figure 4.9, shows the evolution of the periodogram of the entire RV time series as the activity decorrelation scheme is cumulatively applied to data points in chronological

order. That is, the columnar periodogram at each given time, t , corresponds to that of the entire RV time series, with activity decorrelation of all RV measurements before and including time t . Planets b and c have been removed as well as the long term magnetic activity. The bottom panel displays the RV periodograms before and after activity decorrelation, corresponding to the leftmost and rightmost periodograms in the top panel respectively. After decorrelation the periodogram has significant power at 24.5, 35.1, and 40.8 days and several other more moderate peaks. However, after decorrelation (rightmost column), the 24.5 signal dominates, in support of its planetary origin, while most other peaks, including those at 40.8 days and 35.1 days, have been drastically reduced, consistent with manifestations of stellar activity. The single exception is the peak at 17.1 days. While it is unclear why this peak does not vanish, we remind the reader that is the monthly synodic alias of the 40.8 day stellar rotation period, and a planet with such an orbital period would likely be dynamically unstable with planet c ($P=15.3$ days). The 3-planet result of this analysis is consistent with the other analyses previously described. This method of decorrelating stellar noise from RV measurements will be useful for distinguishing planets from stellar activity in other high-cadence RV planet searches.

4.5 Photometry

Photometric observations of planet host stars can be used to measure stellar rotation, detect planetary transits, and sense false planet detections in RV signals. We collected 1855 differential photometry measurements of HD 7924 using the T8 0.8 m automated photometric telescope (APT) at Fairborn observatory in Arizona. The measurements were collected between 31 December, 2006 and 1 December, 2014. The instrument uses two photomultiplier tubes to measure flux in Strömgren b and y filters simultaneously. The b and y measurements are later combined into a single $(b+y)/2$ passband to improve signal to noise. The telescope nods between the target star and several comparison stars. A more detailed description of the observing procedure and data reduction can be found in H09 and Henry et al. (2013). We noticed season to season offsets in the photometry of HD 7924 with a maximum amplitude of ~ 1.5 mmag. These offsets are well correlated with the S_{HK} values during the same time. To better characterize the photometric variability on short timescales we remove these offsets for the subsequent analysis.

The 1855 photometric observations are plotted in the top panel of Figure 4.10 and summarized in Table 4.5, where we have removed yearly offsets by dividing the observations in each season by the seasonal mean. We find a significant periodic signal in the full photometric dataset with a period of 16.9 days and peak-to-peak amplitude of 1.9 ppt. This is similar to the 17.1 day period found in the RVs and most likely a signature of the stellar rotation. Spots may form on opposite hemispheres and cause photometric fluctuations at one half of the orbital period. When the data are broken up into the individual observing seasons we find that the strongest photometric periodicity happens during the 2012-2013 season and has a period of 41.5 days. This season corresponds to

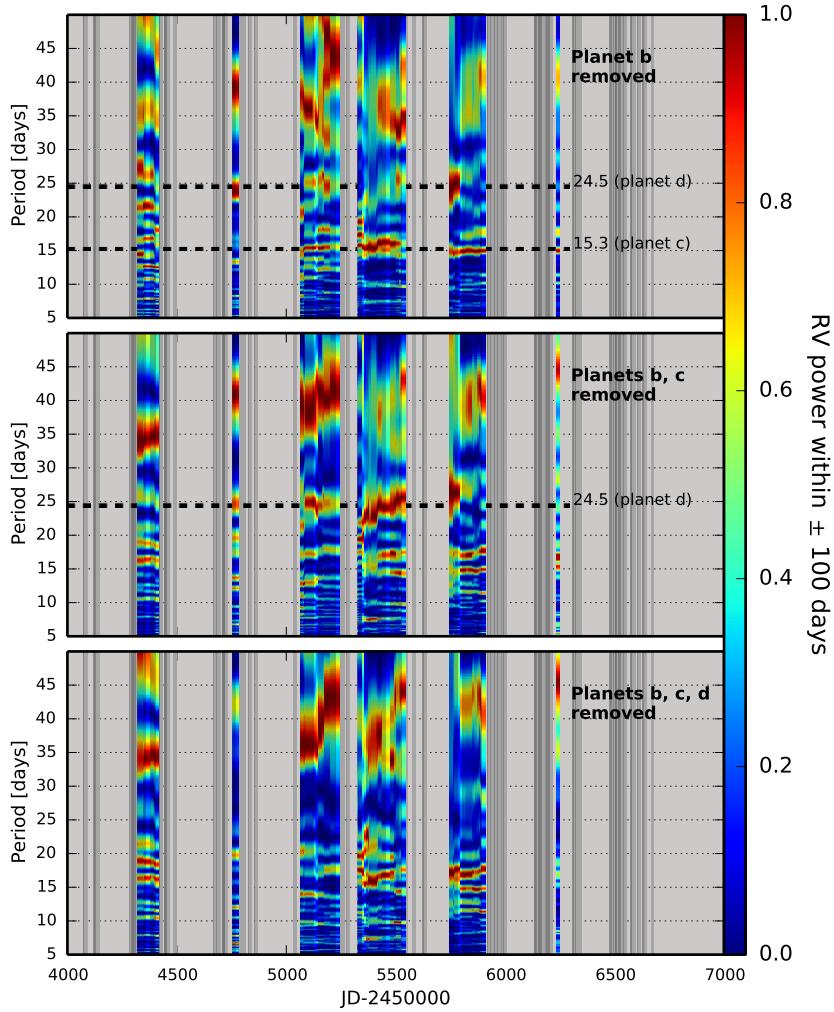


Figure 4.7 *Top*: Periodogram of Keck HIRES RV measurements ± 100 -days versus time (JD-2450000) with our Keplerian model of planet b ($P=5.4$ days) removed. Each colored, single-pixel column is a unique periodogram of all RV measurements within a 200-day window centered on the time indicated on the horizontal axis. Dark grey boxes indicate times of observation having fewer than 24 measurements of both RV and S_{HK} within ± 100 days. Each periodogram has been normalized independently by dividing all periodogram powers by the maximum periodogram value. *Middle*: Same as top, with planets b and c removed. *Bottom*: Same as top, with planets b, c, and d removed.

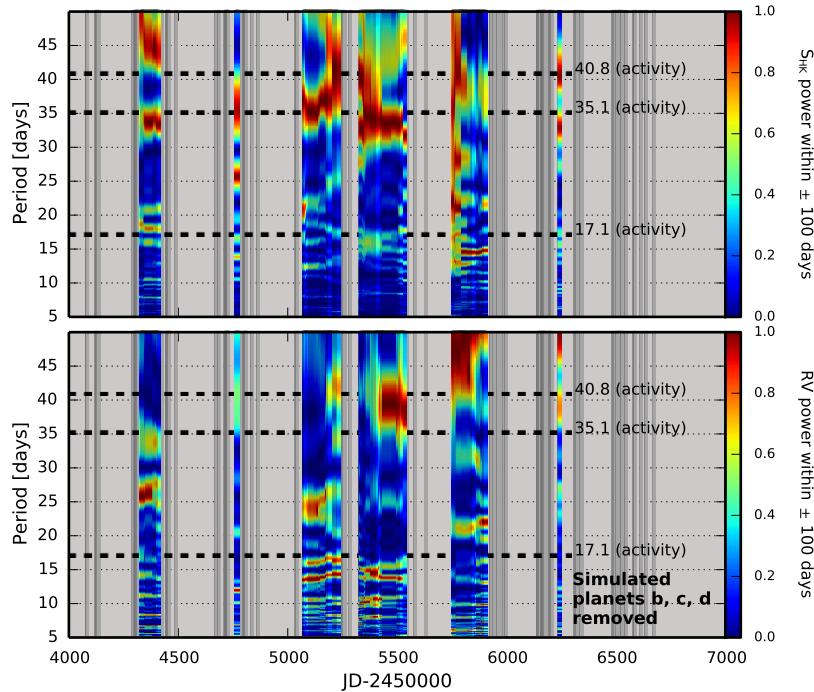


Figure 4.8 *Top*: Same as Figure 4.7 but for S_{HK} (stellar activity) data. *Bottom*: Same as Figure 4.7 bottom panel but for a synthetic RV time-series created by superposing a 1.46 m s^{-1} , 40.8-day sinusoid with the best 3-planet model to the real Keck RVs and 2.36 m s^{-1} Gaussian noise, then fitting for and removing the best 3-planet model ($P=5.4, 15.3, 24.5$ days). Time sampling matches the real-data. Dark grey boxes indicate times of observation having fewer than 24 measurements of both RV and S_{HK} within ± 100 days. Epoch-to-epoch fluctuation in the recovered period of the injected 40.8-day signal is an artifact of time-sampling, Keplerian fitting noise, and injected Gaussian noise.

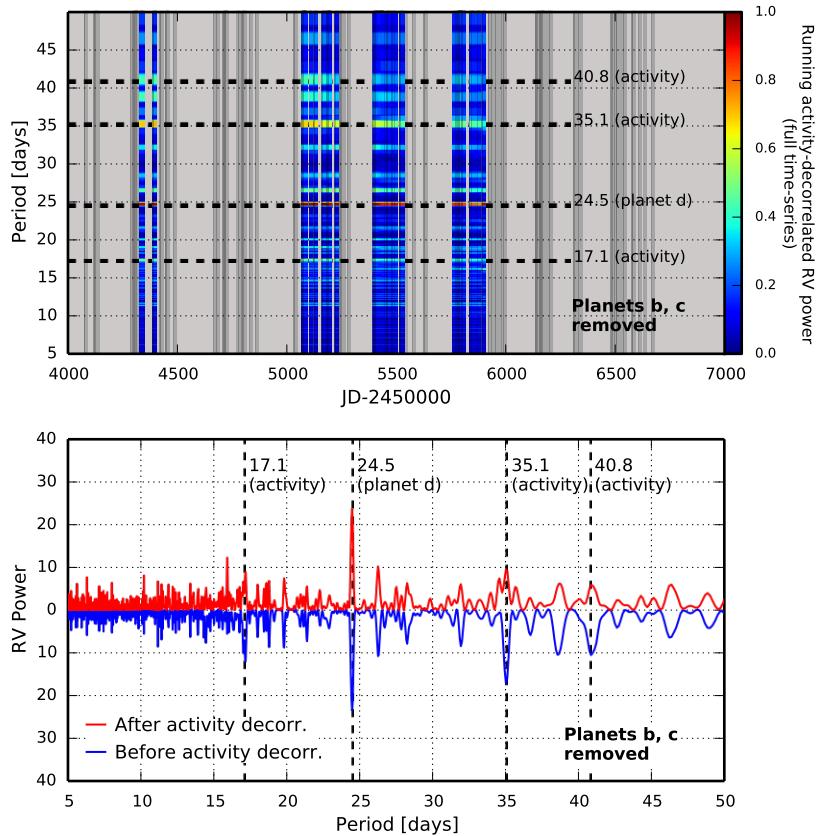


Figure 4.9 *Top*: Running periodogram of the entire RV time-series, where only the subset of RV measurements that precede each time have been decorrelated with stellar activity. Keplerian models of planets b ($P=5.4$ days), c ($P=15.3$ days) and the long-term-stellar-magnetic-activity have been removed. Dark grey boxes indicate times of observation having fewer than 24 measurements of both RV and S_{HK} within ± 100 days. All periodogram powers have been normalized to the same color scale. *Bottom*: Periodogram of entire RV time-series before and after decorrelation with stellar activity. The periods of peaks attributed to planets and stellar activity are labeled accordingly by dashed vertical lines.

the time of maximal S_{HK} activity index. Starspots during this time likely dominate the rotationally induced signal in the RV data.

No significant periodic variability is found in the photometry at the periods of the three planets. A least-squares sine wave fit to the data gives semi-amplitudes of 0.11 ppt, 0.25 ppt, and 0.17 ppt at the orbital periods of planets b, c, and d respectively. Figure 4.11 shows the photometric data phase-folded to the periods of each of the three planets. This gives further evidence that the 5.4, 15.3, and 24.5 day signals are planetary in nature and not caused by stellar activity that would be visible in the photometry at those periods.

With the APT data we are able to rule out transits deeper than ≈ 2 ppt for planet b and 5-8 ppt for planets c and d. We binned the data in bins of width 0.005 units of orbital phase around the time of center transit for each of the planets. For planet b, the lowest binned measurement within the uncertainty window of the transit time is 2.3 ± 1.2 below the mean suggesting that transits deeper than 3.5 ppt do not occur. With a mass of $8.7 M_{\oplus}$ we can rule out non-grazing transits of a pure hydrogen planet, but transits of a denser planet would not be detectable in our data (see Figure 4.11). Our constraints on possible transits of planets c and d are not as strong because their ephemerides have not yet been subject to intense campaigns to search for transits. However, the lowest binned measurements within the transit windows for planets c and d are 6.23 ± 1.5 ppt and 3.4 ± 1.5 ppt, respectively, and suggest that transits deeper than 7.7 ppt for planet c and 4.9 ppt for planet d are not present in the data. This rules out transits of a pure hydrogen planet d, but the lowest binned measurement for planet c is consistent with the depth of a transit caused by a pure hydrogen planet. However, if this were a real event we would expect the neighboring measurements to also be slightly lower than the mean since the transit duration is longer than the width of the bins but this is not the case. Both of the neighboring measurements are higher than the mean flux level. We conclude that transits of planets with compositions dominated by hydrogen for any of the three planets are unlikely, but more and/or higher precision observations are needed to exclude transits of rocky planets.

4.6 Spitzer Transit search

A photometric campaign to look for transits of HD 7924b using the *Spitzer* Space Telescope showed no evidence of transiting planets larger than $1.16 R_{\oplus}$ ($2 - \sigma$ confidence, Kammer et al., 2014). However, they assumed a circular orbit which caused them to underestimate the uncertainty on the time of transit. We find an ephemeris that is inconsistent with that of Kammer et al. (2014) by $\sim 3 \sigma$, but nearly identical to that of H09. With the new ephemeris listed in Table 4.3 we calculate a time of transit for the epoch during which the *Spitzer* observations were collected of $2455867.07^{+0.09}_{-0.11}$. The Kammer et al. (2014) observations would have covered 70% of the predicted ingress or egress times assuming a perfectly edge-on viewing angle for the orbit. The a priori transit probability for planet b is 6.4%. Assuming the Kammer et al. (2014) observations are of sufficient precision

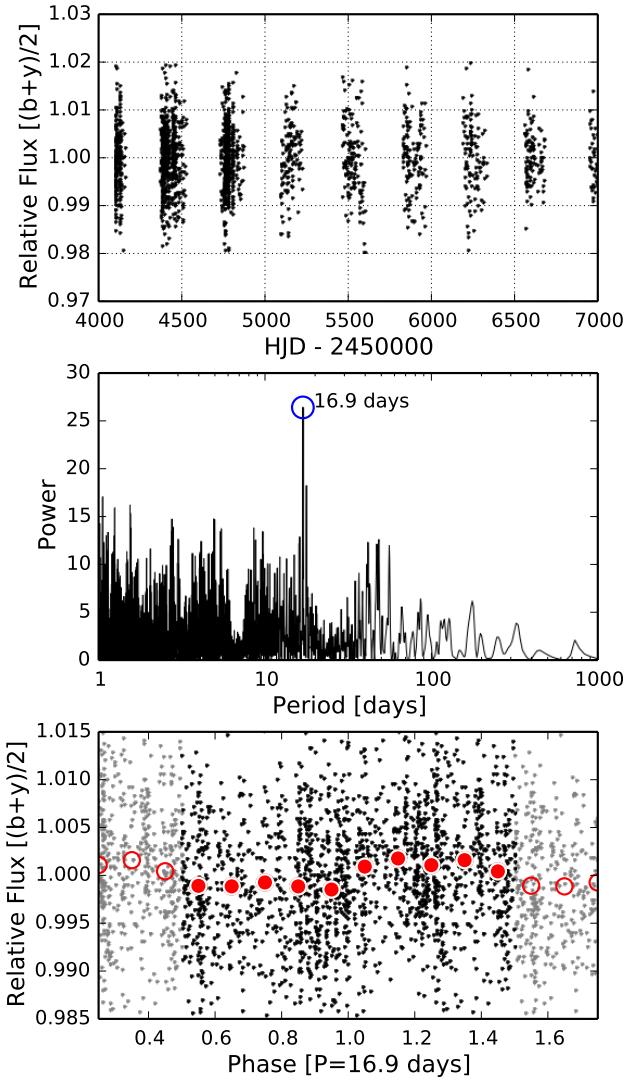


Figure 4.10 Differential photometry of HD 7924 from APT. *Top:* Relative flux time-series of HD 7924 in the combined Strömgren $(b + y)/2$ passband. Seasonal offsets are removed by dividing by the mean within each season. The standard deviation of the photometric time series is 2.3 parts per thousand (ppt). *Middle:* LS periodogram of the photometric time series. *Bottom:* Differential photometry from APT phase-folded to the 16.9 day peak found in the LS periodogram and binned to widths of 0.1 phase units (red circles). The light grey points and open red circles show the same data wrapped by one period. The peak-to-peak amplitude of the 16.9 day photometric signal is 1.1 ppt.

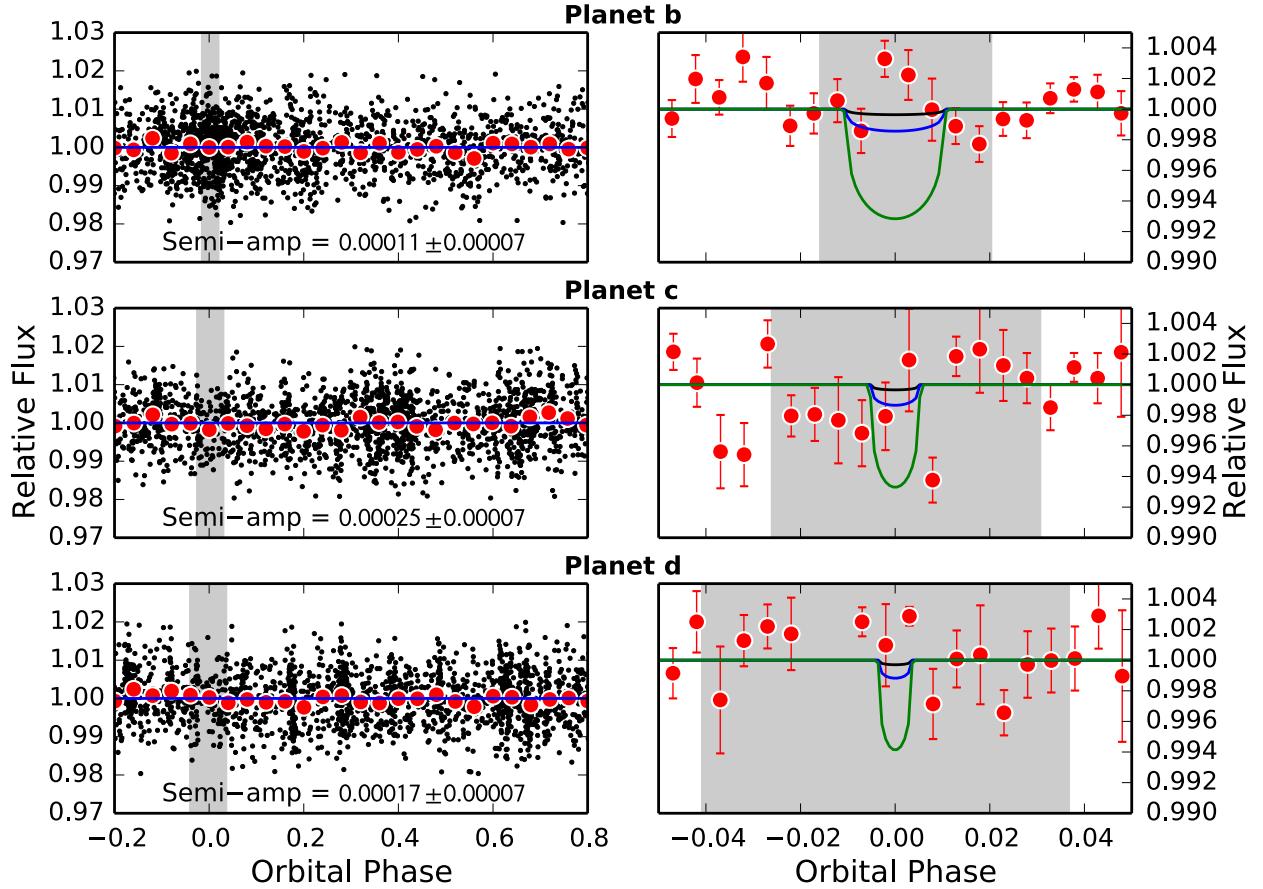


Figure 4.11 Differential photometry from APT of HD 7924 phase-folded to the ephemerides of the three planets. *Left:* Differential photometry phase-folded to the orbital period of planet b (top), c (middle), and d (bottom). Red circles are the photometry measurements grouped in bins of width 0.04 units of orbital phase. The grey shaded region spans the uncertainty in mid-transit time. The semi-amplitude of a sine wave least squares fit to the data is annotated at the bottom of each panel. *Right:* Same as the left panels zoomed in around the phase of center transit. The individual measurements are omitted and the red circles correspond to binned photometry data in bins of width 0.005 units of orbital phase. The three curves are Mandel & Agol (2002) transit models for planets with the masses listed in Table 5.6 and densities of iron (black), water (blue), and hydrogen (green).

Table 4.5. Summary of APT Photometric Observations

Observing Season	N_{obs}	Julian Date Range (HJD – 2,440,000)	Sigma (mag)	P_{rot}^1 (days)	Full Amplitude ² (mag)
2006–2007	231	14100–14158	0.00212	17.1 ± 0.2	0.0021 ± 0.0004
2007–2008	524	14370–14523	0.00215	17.1 ± 0.1	0.0011 ± 0.0003
2008–2009	464	14728–14867	0.00208	45.7 ± 0.5	0.0013 ± 0.0003
2009–2010	123	15092–15222	0.00212
2010–2011	140	15459–15598	0.00239	18.1 ± 0.1	0.0024 ± 0.0005
2011–2012	125	15823–15963	0.00222	16.8 ± 0.1	0.0021 ± 0.0005
2012–2013	109	16185–16330	0.00235	41.5 ± 0.8	0.0027 ± 0.0006
2013–2014	100	16555–16674	0.00168	25.1 ± 0.3	0.0022 ± 0.0004
2006–2014	1816	14100–16674	0.00197	16.8922 ± 0.0014	0.00098 ± 0.00013

¹Period of most significant peak in a periodogram analysis.

²Amplitude of best-fit sine function with the period fixed to P_{rot} .

to detect any ingress or egress that would have occurred within their observational window, the a posteriori transit probability is 2.0%. Figure 4.12 compares the observational window of the *Spitzer* observations to the predicted time of transit based on the ephemeris of this work.

4.7 Architecture and Stability

4.7.1 Compact Multi-planet Systems

In order to compare the architecture of the HD 7924 system with other compact multi-planet systems we compiled a catalogue of similar multi-planet systems from *Kepler*. We restricted the *Kepler* multi-planet systems to those with exactly three currently known transiting planets in the system³, a largest orbital period less than or equal to 30 days, and all three planets must have radii smaller than $4 R_{\oplus}$. This left a total of 31 systems from the sample of confirmed *Kepler* systems (Borucki et al., 2011; Batalha et al., 2013; Marcy et al., 2014; Rowe et al., 2014). Their architectures are presented in Figure 4.13.

We notice nothing unusual about the architecture of HD 7924 when compared with the *Kepler* systems. Most of these multi-planet systems contain three planets with masses between 5–10 M_{\oplus} and semi-major axis between 0.05 and 0.3 AU. The systems of *Kepler*-194, *Kepler*-124, *Kepler*-219, *Kepler*-372, *Kepler*-310, *Kepler*-127, and *Kepler*-342 all host one inner planet and two outer planets that are closer to each other than to the innermost planet as in the HD 7924 system. The *Kepler*-372 system is particularly similar with one planet orbiting at 0.07 AU and a pair of outer planets

³As returned by a 2014 Nov 24 query of exoplanets.org

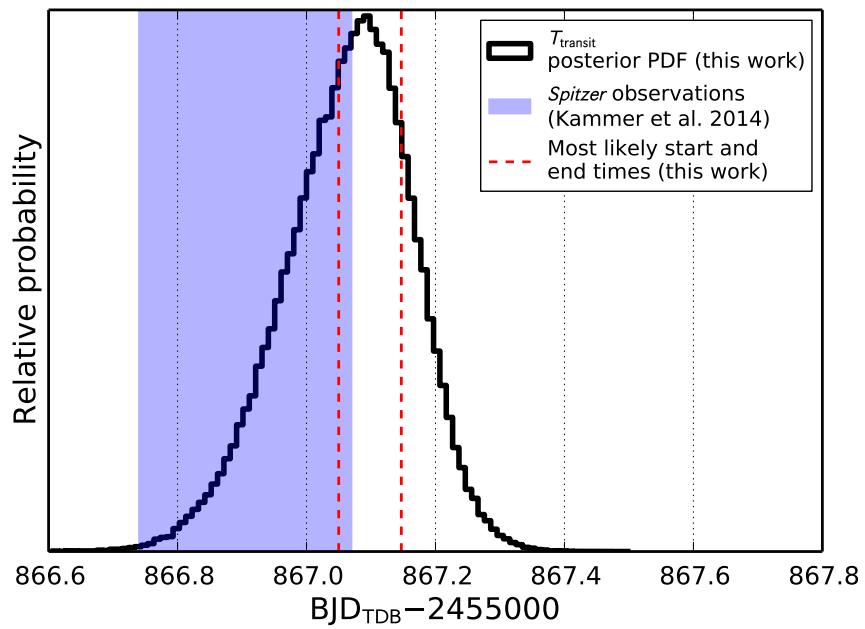


Figure 4.12 Comparison of the transit search time window covered by the *Spitzer* observations of Kammer et al. (2014) to the transit times as calculated from the updated ephemeris in this work. The thick black line shows the posterior probability density function (PDF) of the mid-transit time for the transit that was targeted by Kammer et al. (2014). The blue shaded region shows the time window covered by their observations, and the red dashed lines show the most likely ingress and egress times.

orbiting at 0.14 and 0.19 AU. This demonstrates that the HD 7924 system is not abnormal in the context of other known compact multi-planet systems. While this is not a direct demonstration of dynamical stability, the fact that many other systems exist with very similar architectures gives strong empirical evidence that the HD 7924 system is a stable planetary configuration.

The HD 7924 planetary system is one of very few RV-detected systems hosting three super Earths. Only HD 40307 (Mayor et al., 2009a), and HD 20794 (Pepe et al., 2011) are known to host three planets with masses all below $10 M_{\oplus}$. Both of these stars are unobservable from most northern hemisphere observatories. If we expand the mass limit to include systems with three or more planets with masses below $25 M_{\oplus}$ we find four additional systems that match these criteria, HD 69830 (Lovis et al., 2006), Gl 581 (Mayor et al., 2009b), HD 10180 (Lovis et al., 2011), and 61 Virginis (Vogt et al., 2010). These systems are difficult to detect given their small and complex RV signals and HD 7924 is the first discovery of such a system since 2011.

4.7.2 Dynamical Stability

In order to check that the HD 7924 system is dynamically stable for many orbital periods we ran a numerical integration of the three planet system using the *MERCURY* code (Chambers, 1999). We started the simulation using the median orbital elements presented in Table 4.3 and let it proceed 10^5 years into the future. We assume that the system is perfectly coplanar with zero mutual inclination and we assume that the masses of the planets are equal to their minimum masses ($\sin i = 0$). No close passages between any of the three planets were found to occur (≤ 1 Hill radius) during the entire simulation, suggesting that the system is stable in the current configuration.

Following the arguments of Fabrycky et al. (2014) we calculated the sum of separations between the planets (b to c, plus c to d) in units of their Hill radii. Fabrycky et al. (2014) found that the sum of the separations for the vast majority of *Kepler* multi-planet systems is larger than 18 Hill radii. The sum of the separations between the planets of the HD 7924 system is 36 Hill radii, providing further evidence that this system is stable and not abnormally compact relative to the many *Kepler* multi-planet systems.

4.8 APF vs. Keck

Since this is the first publication from our group utilizing APF data we perform a comparison between the two data sets to assess their relative performance. First, we use the automated planet detection algorithm described in §4.4.1 on each data set independently. When the APF and Keck data are analyzed independently the 5.4, 15.3, and 24.5 day signals are the first to be identified in both cases (albeit with much lower significance in the APF data alone) making it extremely unlikely that these signals could be caused by instrumental systematic noise.

The binned measurement uncertainties for the APF data are generally higher than those of Keck

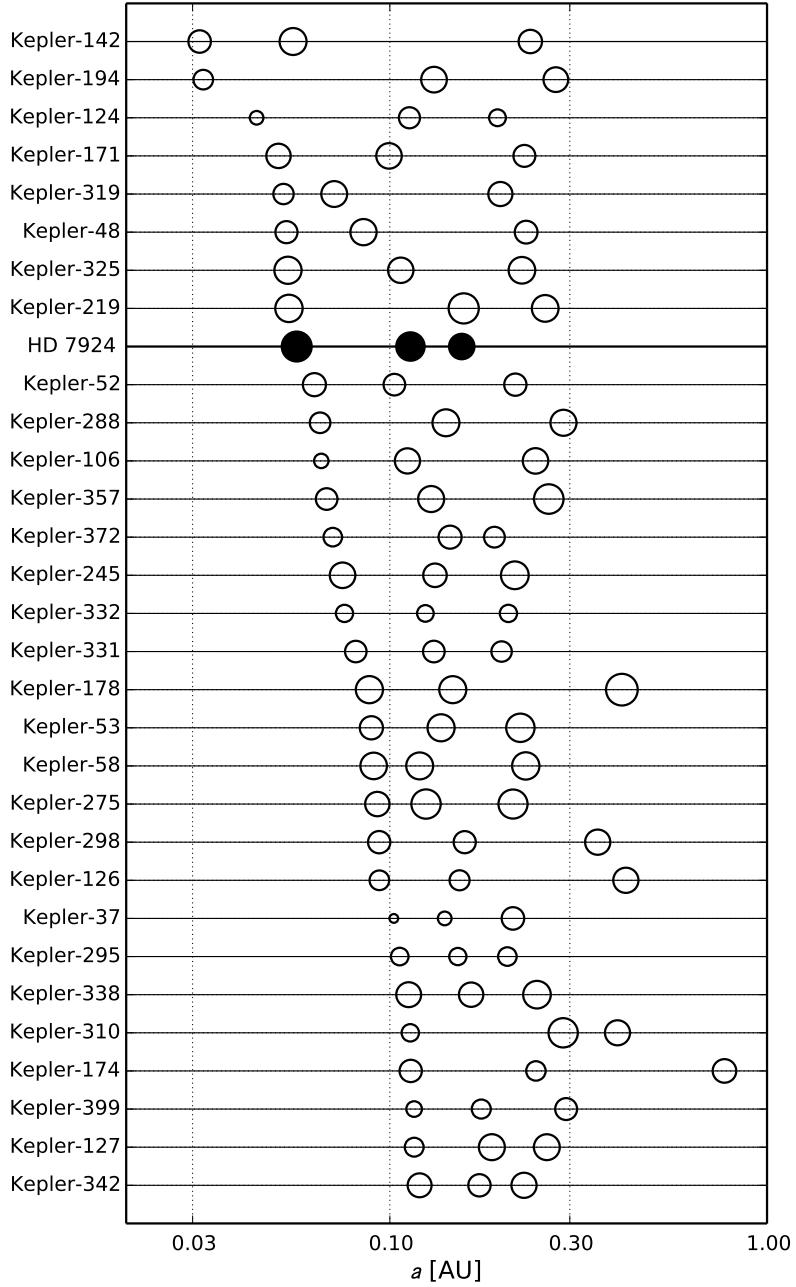


Figure 4.13 The architecture of *Kepler* multi-planet systems compared with the architecture of the HD 7924 system. We plot the *Kepler* systems with exactly three currently known planets, a largest orbital period of 30 days or less, and a maximum planetary radius of $4 R_{\oplus}$. The HD 7924 planets are plotted as filled black circles. The systems are sorted by the semi-major axis of the shortest period planet. The size of the circles is proportional to planet radius. The radii of the HD 7924 planets have been calculated as $\frac{R_p}{R_{\oplus}} = 0.371 \left(\frac{M_p}{M_{\oplus}} \right)^{1.08}$ (Weiss & Marcy, 2014).

(before stellar jitter is accounted for). This is not surprising since the APF data is usually collected at lower signal to noise than the Keck data. However, since we are allocated many more nights on APF we can collect more measurements over a shorter amount of time. It took \sim 13 years to collect 281 binned measurements with Keck, but only 1.5 years to collect 80 binned measurements with the APF. In addition, HD 7924’s circumpolar track at Lick Observatory and the long nights during the winter months in Northern California both increase the observability of this target. The standard deviation of the residuals to the best fit model are similar in the two cases at 2.8 m s^{-1} for APF and 2.5 m s^{-1} for the post-upgrade Keck data. In order to compare the tradeoffs between higher cadence and higher precision we define a metric similar to the “photometric noise rate” as used by Fulton et al. (2011) to compare the relative performance between photometric data sets. We define a “velocity noise rate” as

$$\text{VNR} = \frac{\sigma}{\sqrt{\Gamma}}, \quad (4.4)$$

where σ is the RMS of the velocities and Γ is the mean number of observations per year. We assume that no further signals are present in the data and/or they contribute negligibly to the RMS. The VNR gives the approximate K in m s^{-1} that would be detectable with $S/N=1$ after one year of observing at typical cadence. Of course, we generally require $S/N \gg 1$ in order to consider a signal a viable planet candidate, but the VNR still gives a good reference point for comparison. We list the VNR and other performance characteristics for Keck and APF data for HD 7924 and two well-known RV standard stars (HD 10700, and HD 9407) in Table 4.6. The velocities for the three stars are compared side-by-side in Figure 4.14. The VNR for APF data is 25–50% smaller than that of Keck indicating that we will be sensitive to smaller planets once we have observed the stars for a comparable amount of time or, in other words, our APF data will be of comparable sensitivity to the existing Keck RVs in 1/2 to 3/4 of the length of time it took to reach that sensitivity at Keck.

We are actively tuning and perfecting our Doppler analysis pipeline for APF. Recently, we implemented an experimental technique to correct the velocities for any correlations of the RVs with environmental parameters (e.g. atmospheric pressure, CCD temperature). This technique reduces the RV RMS for Keck data slightly (5–10%), but significantly reduces the RMS for APF velocities in most cases and up to a factor of two in some cases. The RV RMS for HD 9407 and HD 10700 reduce to 2.24 m s^{-1} and 1.67 m s^{-1} respectively when correlations with non-astrophysical variables are removed.

4.9 Discussion & Summary

We present the discovery of two additional super-Earth mass planets orbiting the bright K0.5 dwarf HD 7924. These planets join a previously-known planet in a system with at least three super-Earth mass planets. The two new planets have minimum masses of $M_c \sin i_c = 7.8 M_{\oplus}$ and

Table 4.6. APF vs. Keck Radial Velocity Precision

Instrument	RMS (m s ⁻¹)	Median Uncertainty ¹ (m s ⁻¹)	N_{obs}^2	Mean Cadence (days)	VNR ³ m s ⁻¹ y ⁻¹
HD 7924					
Pre-upgrade Keck/HIRES	1.20	2.64	7	110.4	0.66
Post-upgrade Keck/HIRES	2.50	2.36	281	13.5	0.48
APF/Levy	2.80	2.86	80	5.9	0.36
HD 10700 (τ Ceti)					
Pre-upgrade Keck/HIRES	2.87	2.84	84	17.5	0.63
Post-upgrade Keck/HIRES	2.32	2.22	190	19.4	0.53
APF/Levy	2.16	2.12	66	7.0	0.30
HD 9407					
Pre-upgrade Keck/HIRES	1.89	1.94	12	162.8	1.26
Post-upgrade Keck/HIRES	1.89	1.86	202	18.3	0.52
APF/Levy	2.34	2.30	104	4.5	0.26

¹Stellar jitter has been added in quadrature with the binned velocities such that the χ^2 of the velocities with respect to their median value is 1.0.

²Binned in units of 0.5 days.

³See Equation 4.4.

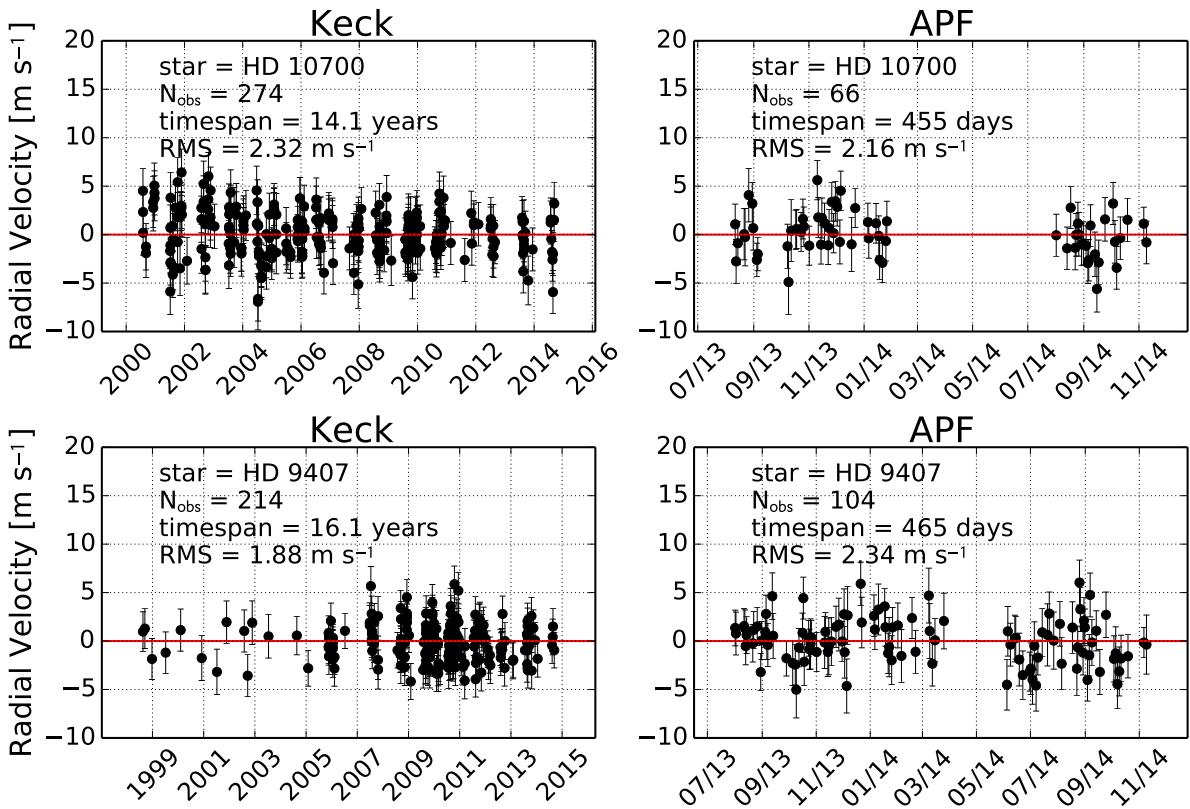


Figure 4.14 A comparison between RV performance at Keck and APF for two well known RV standard stars. These stars are some of the most well-observed stars at both Keck and APF. In the first 1.5 years of APF operations we have already collected nearly half of the data that has been collected at Keck over the last ~ 15 years. Stellar jitter has been added in quadrature with the binned velocities such that the χ^2 of the velocities with respect to their median value is 1.0.

$M_d \sin i_d = 6.3 M_{\oplus}$, and orbit HD 7924 with semi-major axis of $a_c = 0.113$ AU and $a_d = 0.155$ AU. Both planets receive far too much radiation from HD 7924 to be within the habitable zone as defined by Kopparapu et al. (2013) with incident stellar irradiation values 114, 28, and 15 times that received from the Sun by Earth for planets b, c, and d respectively. Assuming that these planets have bond albedos similar to the mean total albedos of super-Earths ($A_t = 0.32$, Demory, 2014) their equilibrium temperatures are 826, 584, and 499 K.

The stellar magnetic activity cycle is clearly visible in our long-baseline RV time series and we observe nearly two complete cycles. We simultaneously model the RV shift due to the magnetic cycle and the three planets in order to extract accurate planetary parameters. A tentative RV signal from rotationally-modulated starspots is also found and we perform a rigorous analysis to determine that the planetary signals are distinct from the stellar activity signals.

With the largely expanded data set we are able to refine the ephemeris of planet b and show that additional transit monitoring is needed. Since HD 7924 is near the ecliptic pole, it will be near the continuous viewing zone for the James Webb Space Telescope (Gardner et al., 2006) and a transiting planet would make an excellent candidate for space-based transmission spectroscopy. The Transiting Exoplanet Survey Telescope (Ricker et al., 2014) will observe HD 7924 once launched and will be able to conclusively determine if any of the three planets are transiting.

This system is a good example of a compact system of short-period planets for which high-cadence observations are incredibly valuable to determine the coherence of signals and detect the true physical periods as opposed to aliases. Since we know that short-period super-Earth size planets are common (Howard et al., 2010a), we expect our continued nearly nightly observations with the APF to uncover many more systems like HD 7924 from which we will build a comprehensive census of the small planets in our local solar neighborhood.

CHAPTER 5

THREE TEMPERATE NEPTUNES ORBITING NEARBY STARS

This chapter is a reproduction of Fulton et al. (2016) included with permission from AAS journals.

5.1 Introduction

The mass function of extrasolar planets potentially offers rich clues to the processes that shape their growth and evolution. This mass function is known to rise with decreasing planet mass based on the discovery and characterization of planets orbiting nearby stars (Howard et al., 2010a; Mayor et al., 2011). Although the *Kepler* mission discovered thousands of transiting exoplanets allowing for detailed characterization of the planet radius distribution (Howard et al., 2012b; Fressin et al., 2013; Petigura et al., 2013b), we have a much sparser sample of planets orbiting nearby stars. We only know of 17 confirmed planets with measured minimum masses $M \sin i < 30 M_{\oplus}$ and orbital periods $P > 75$ days¹. These low-mass, temperate planets reside in an region of parameter space that must be explored in order to understand the formation of the extremely abundant population of close-in super-Earths.

The discovery that the occurrence rate of Jovian planets increases at orbital distances of 1–3 AU (Cumming et al., 2008) has been suggested to be a sign that the ice line is important to the formation of Jovian planets. The increased abundance of solids in the protoplanetary disk beyond the ice-line is expected to speed up the coagulation of planetesimals and $\sim 10 M_{\oplus}$ cores that can undergo runaway gas accretion before the gas in the disk is dissipated (Ida & Lin, 2008).

The formation of close-in intermediate mass planets known as super-Earths or mini-Neptunes presents some challenges for planet formation theory. It was initially thought that such planets should either remain small terrestrial planets, or if they grow to large cores then they should quickly accrete substantial nebular gas and grow to be gas giants (Ida & Lin, 2004; Mordasini et al., 2009a). Alternatively, planets that form *in situ* likely require either a protoplanetary disk that is much more massive than the minimum mass solar nebula (Hansen & Murray, 2012; Chiang & Laughlin, 2013), an extremely metal rich disk, or fine tuning of the formation timescales (Lee et al., 2014). New models featuring gas-drag driven “pebble accretion” which offers a mechanism to transport solids in the disc (Chatterjee & Tan, 2014), or the delayed formation of super-Earth cores until the gas disk begins to dissipate (Lee & Chiang, 2015) could facilitate *in situ* formation of super-Earths and/or mini Neptunes.

Alternatively, migration from further out in the disk where there is plenty of material to form massive planet cores or super-Earths could explain the presence of these planets near their host

¹Based on a query of exoplanets.org on 28 June 2016 (Han et al., 2014)

stars. We would expect to see multi-planet systems in resonant chains if a slow, smooth migration were the dominant mechanism and this is not observed (Veras & Ford, 2012; Fabrycky et al., 2014). Another indication that migration is at play would be an increased occurrence rate of super-Earths at large orbital separations. The occurrence rates of super-Earths as a function of orbital period appears fairly flat to periods as long as \sim 200 days (Petigura et al., 2013b), however the sensitivity to long-period planets from transit surveys is very low. In addition, the planet radius distribution can be influenced by a variety of different factors including stellar irradiation, and thermal evolution (Lopez, 2014). Only a very small percentage by mass of volatiles can significantly inflate a planet’s radius and hide the fundamental properties of the planet that encode information about the formation mechanism (Lopez & Fortney, 2014). RV surveys are now starting to discover a statistically useful sample of super-Earth to Neptune mass planets at larger orbital separations and lower stellar irradiance that will help to map out the details of the mass function for long-period super-Earth to Neptune mass planets.

The Eta-Earth RV survey of nearby stars (Howard et al., 2010a) was conducted using the HIRES spectrograph at Keck observatory (Vogt et al., 1994). They searched for planets in a volume-limited sample of 166 nearby G and K dwarfs. With the catalogue of planets detected in this survey (Howard et al., 2009, 2011a,c, 2014), and the completeness limits calculated for each star they were able to measure the occurrence rate of small, short-period planets and show that planets with masses of 3-30 M_{\oplus} are much more common than planets larger than 30 M_{\oplus} . Although these low-mass planets are common, they are still very difficult to detect, requiring > 100 measurements per star with $\lesssim 2$ m s $^{-1}$ precision. The number of nearby stars for which large, high-precision radial velocity datasets exist is increasing thanks to the proliferation of dedicated and robotic radial velocity facilities such as the APF and MINERVA (Swift et al., 2015) and the ongoing long-term surveys from Keck, HARPS-N (Motalebi et al., 2015), and HARPS (Pepe et al., 2004).

We are currently using the APF telescope to conduct a RV survey of 51 of the brightest, and least chromospherically active, stars from the Eta-Earth survey. We capitalize on the robotic nature of the telescope to monitor the stars at high cadence for the entire four year duration of the survey. This survey builds on the Eta-Earth Survey, but with improved Doppler precision due to the high observing cadence and larger number of measurements. We will measure the occurrence rate and mass function of small planets in our local neighborhood using the new planets discovered by the APF-50 survey and the set of planets already known to orbit stars in our sample. With a larger sample of planets with measured masses in the 3-30 M_{\oplus} range we will measure the mass function for small planets with higher mass resolution. Combining the mass function from this survey with the size distribution from *Kepler*, we will probe the density and core mass properties of super-Earths to inform formation theories of the galaxy’s most abundant planets.

Nearby G and K dwarf stars are observationally advantageous. High SNR spectroscopy can often be obtained with relatively short exposure times, facilitating many precise time series RV

measurements with sensitivity to planets with small Doppler amplitudes. The stars can be characterized precisely using spectroscopy and asteroseismology (when available) that improve estimates of the star and planet properties. The advantages of nearby, bright targets are critical for characterization of the planets atmospheres transmission, emission, and direct spectroscopy.

Here we present the discovery of three roughly Neptune mass planets orbiting bright stars within 25 pc. This paper is structured as follows. Our observational setup and RV measurements are described in Section 5.2. In Section 5.3 we discuss our derived stellar properties for each of the three stars and compare with previous literature studies. We describe our methods used to discover these planets in Section 5.4. We describe our modeling procedure used to obtain the final adopted parameters and their associated uncertainties and our various tests to ensure that the signals are planetary in nature in Section 5.4.2. We analyze photometry of each of the three systems in Section 5.5, and conclude with a summary and discussion in Section 8.5.

5.2 Radial Velocity Measurements

We collected 571 RV measurements of HD 42618 using Keck/HIRES (Vogt et al., 1994) and 35 measurements using the Levy spectrograph on the Automated Planet Finder (APF, Vogt et al., 2014b; Radovan et al., 2014) over the past 19 years starting in 1996. For each RV measurement the starlight is passed through a cell of gaseous iodine that imprints a dense forest of molecular absorption lines onto the stellar spectrum and serves as both a wavelength and point spread function (PSF) reference. We also collected a single set of iodine-free observations of this star that was deconvolved with the instrumental PSF and used as a model of the intrinsic stellar spectrum. Each observation was forward modeled as the intrinsic stellar spectrum doppler shifted by an arbitrary amount, then multiplied by the transmission of iodine, and convolved with the instrumental PSF modeled as a sum of 13 Gaussians with fixed widths and positions but heights free to vary (Butler et al., 1996b). The Levy slit-fed spectrograph also relies on an iodine cell for precise RVs and our observational setup is described in detail in Fulton et al. (2015b).

Our setup was identical for the three stars. We collected 328 Keck/HIRES measurements and 73 APF measurements for HD 164922 over the past 19 years. All of the APF measurements and 244 of the Keck measurements are new since the publication of Butler et al. (2006). For HD 143761 we obtained 519 RV measurements using Keck/HIRES and 157 measurements using the Levy spectrograph on the APF. The Keck observations of HD 143761 started in 2006 and the total observational baseline is 8 years. We do not include the Lick 3.0 m data from Noyes et al. (1997) and Butler et al. (2006) in our analysis. We find that including the Lick data does not significantly improve the uncertainties on any of the orbital parameters and it adds an additional source of systematic uncertainty that is less well characterized and understood.

Our Doppler pipeline has been tuned in small ways over the years to improve RV precision. Here we describe a new pipeline improvement that decorrelates the measured RVs with nuisance

parameters from the model and spectrometer state parameters. This decorrelation offers modest improvements in Doppler precision ($\sim 1 \text{ m s}^{-1}$) and is only applied to time series RV of stars for which the number of spectra greatly exceeds the number of potential decorrelation parameters. The nuisance parameters in Doppler analysis include descriptions of the point spread function (PSF) over the spectral format and the wavelength solution. The PSF is parameterized as a sum Gaussians with fixed widths and centers, but variable amplitudes. We also have a wealth of information about the weather and environment inside and outside the spectrograph extracted from the FITS headers of the raw APF spectra. Some environmental information is available in the FITS headers for the Keck data, but we have not yet implemented a system to extract these values for the tens of thousands of Keck spectra taken over the last 20 years. For Keck data we only include the nuisance parameters that are part of the forward modeling process. We clean the RVs of systematic trends by removing any correlations that these parameters show with the final RVs. We search for significant correlations of the RVs with all of the PSF parameters by calculating the Spearman rank correlation coefficient (Spearman, 1904) and flag any parameters that show correlation coefficients greater than 0.1. The flagged parameters are included in a multivariate ordinary least squares linear regression using the `STATSMODELS`² package in Python. This model for RV as a function of all parameters included in the fit is then subtracted from the raw RVs. This process is done blindly in the planet discovery/identification phase but once planet candidates are identified in a given dataset we first model the system to find the best fit N-planet Keplerian model then perform the detrending procedure on the residuals to this model. The detection of all newly discovered planets in this work does not depend on this detrending and they are easily identified at high significance in either the detrended or non-detrended datasets.

We reject measurements with low signal-to-noise ratios ($\text{SNR} < 60$ per pixel) and/or uncertainties more than 9σ larger than the median uncertainty, which results in the omission of $< 1\%$ of the data for each of the three stars. Since these stars are exceptionally bright we almost always collect three consecutive measurements in order to average out RV shifts caused by p-mode oscillations (Dumusque et al., 2011). The three measurements are then binned together before the stellar jitter is added in quadrature during the modeling process (see Section 5.4.2). This effectively reduces the weight of the three measurements to that of a single measurement, but averages out some of the astrophysical noise in the process and prevents time-correlated instrumental systematic noise from biasing the results. We also extract the Ca 2 H & K activity index (S_{HK}) using the technique of Isaacson & Fischer (2010) for every RV measurement on both Keck and APF, however there may be an arbitrary zero point offset in the S_{HK} values between the Keck and APF values. The uncertainties for the S_{HK} measurements are systematically limited to 0.002 for Keck and 0.004 for APF. This was estimated by measuring the standard deviation of all measurements of the extremely chromospherically quiet star, HD 10700. All RV measurements and the associated S_{HK} values can

²<https://pypi.python.org/pypi/statsmodels>

Table 5.1. Radial Velocities^a

HD	BJD _{TDB} (− 2440000)	RV ¹ (m s ^{−1})	Unc. (m s ^{−1})	Inst. ²	<i>S</i> _{HK} ³
42618	2450366.126333	+2.85	1.12	k	...
42618	2453694.093412	+1.17	1.13	j	0.161
164922	2454777.744397	−9.19	1.04	j	0.152
164922	2457267.662041	−4.43	1.90	a	0.145
143761	2455455.762546	−41.98	1.27	j	0.149
143761	2457292.685768	−8.08	2.92	a	0.135

^a(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

¹Zero point offsets between instruments have not been removed and must be fit as free parameters when analyzing this dataset

²k = pre-upgrade Keck/HIRES, j = post-upgrade Keck/HIRES, a = APF

³Uncertainties on *S*_{HK} are 0.002 for all Keck measurements and 0.004 for all APF measurements

be found in Table 5.1. We include only the detrended velocities in Table 5.1 but the full set of environmental and PSF parameters for each observation along with the non-detrended velocities can be downloaded from https://github.com/bjfultn/three_neptunes.

5.3 Stellar Properties

5.3.1 HD 42618

HD 42618, also known as HIP 29432 and Gl 3387, is a well studied solar analogue located at a distance of 23.5 pc (van Leeuwen, 2007). The star was not previously known to host any exoplanets. We analyzed 5 high SNR Keck-HIRES spectra (described below) using **SpecMatch** (Petigura, 2015) to obtain the mean spectroscopic parameters listed in Table 5.2. **SpecMatch** uses trilinear interpolation to synthesize high resolution model spectra from the Coelho (2014) grid of models for any set of arbitrary stellar parameters (T_{eff} , $\log g$, [Fe/H], and $v \sin i$) that are contained within the limits of the model grid. The interpolated models are then compared to the observed spectrum. We maximize the likelihood ($\mathcal{L} = e^{-\chi^2/2}$) to determine the optimal stellar parameters, where χ^2 is summed over the extracted spectral pixels and normalized by the flux uncertainties.

Our spectral analysis is consistent with the results of Valenti & Fischer (2005) who extracted spectroscopic parameters using Spectroscopy Made Easy (SME) and found $T_{\text{eff}} = 5747 \pm 44$ K, $\log g = 4.43 \pm 0.06$, and [Fe/H] = -0.11 ± 0.03 . HD 42618 is chromospherically quiet with $\log R'_{\text{HK}} =$

Table 5.2. Adopted Stellar Properties

Parameter	HD 42618	HD 164922	HD 143761
Spectral type	G4V ¹	G9V ⁶	G0V ⁸
$B - V$ (mag)	0.657 ²	0.800 ⁷	0.600
V (mag)	6.839 ²	6.99 ⁷	5.41 ⁸
J (mag)	5.701 \pm 0.023 ³	5.553 \pm 0.026 ³	4.09 ³
H (mag)	5.385 \pm 0.024 ³	5.203 \pm 0.017 ³	3.99 ³
K (mag)	5.301 \pm 0.020 ³	5.113 \pm 0.020 ³	3.89 \pm 0.05 ⁸
Distance (pc)	23.50 \pm 0.30 ⁴	22.13 \pm 0.27 ⁴	17.236 \pm 0.024 ⁴
T_{eff} (K)	5727 \pm 60	5293 \pm 32	5627 \pm 54 ¹¹
$\log g$ (cgs)	4.44 \pm 0.07	4.387 \pm 0.014	4.121 \pm 0.018
[Fe/H] (dex)	-0.09 \pm 0.04	+0.16 \pm 0.05	-0.31 \pm 0.05
$v \sin i$ (km s ⁻¹)	\leq 2	\leq 2	\leq 2
L_{\star} (L_{\odot})	0.98 \pm 0.17	0.703 \pm 0.017	1.706 \pm 0.042 ¹¹
M_{\star} (M_{\odot})	1.015 \pm 0.061 ¹²	0.874 \pm 0.012	0.889 \pm 0.030
R_{\star} (R_{\odot})	0.999 \pm 0.087 ¹²	0.999 \pm 0.017	1.3617 \pm 0.0262 ¹¹
$\log R'_{\text{HK}}$	-5.01 ⁵	-5.06 ⁵	-5.05 ⁵
S_{HK}	0.157 ⁵	0.154 ⁵	0.150 ⁵

-- References - (1) Medhi et al. (2007); (2) Koen et al. (2010); (3) Cutri et al. (2003); (4) van Leeuwen (2007); (5) Isaacson & Fischer (2010); (6) Gray et al. (2003); (7) Butler et al. (2006); (8) van Belle & von Braun (2009); (9) Noyes et al. (1997); (10) van Belle & von Braun (2009); (11) von Braun et al. (2014); (12) Torres et al. (2010)

-5.01 (Isaacson & Fischer, 2010). It was deemed to be a good solar-analog based on a very similar chemical abundance pattern to the Sun (Morel et al., 2013a). Those authors also derive $T_{\text{eff}} = 5765 \pm 17$ K, and $\log g = 4.48 \pm 0.04$, consistent with our SpecMatch results. Ramírez et al. (2014) measured the fundamental parameters of HD 42618 differentially relative to the Sun which allowed them to obtain highly precise values for $T_{\text{eff}} = 5758 \pm 5$ K, $\log g = 4.44 \pm 0.01$, and $[\text{Fe}/\text{H}] = -0.096 \pm 0.005$ that show good agreement with our results. It has also been noted that HD 42618 shows a low lithium abundance of $A_{\text{Li}} = 1.22$ (Ramírez et al., 2012) similar to that of the Sun ($A_{\text{Li}} = 1.05 \pm 0.1$ Asplund et al., 2009). Our adopted stellar mass and radius for HD 42618 are based on the relations of Torres et al. (2010) using our spectroscopic constraints on T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$. HD 42618 was also target of the CoRoT mission (Baglin et al., 2009), with a preliminary detection of solar-like oscillations presented by Barban et al. (2013). We performed an independent asteroseismic analysis of the CoRoT photometry (see Section 5.5.1), which yielded a mass and radius in agreement with our adopted values.

5.3.2 HD 164922

HD 164922, also known as HIP 88348 and Gl 9613, is a bright, chromospherically inactive ($\log R'_{\text{HK}} = -5.06$, Isaacson & Fischer, 2010) G9 V dwarf located 22.1 pc away (van Leeuwen, 2007). It was

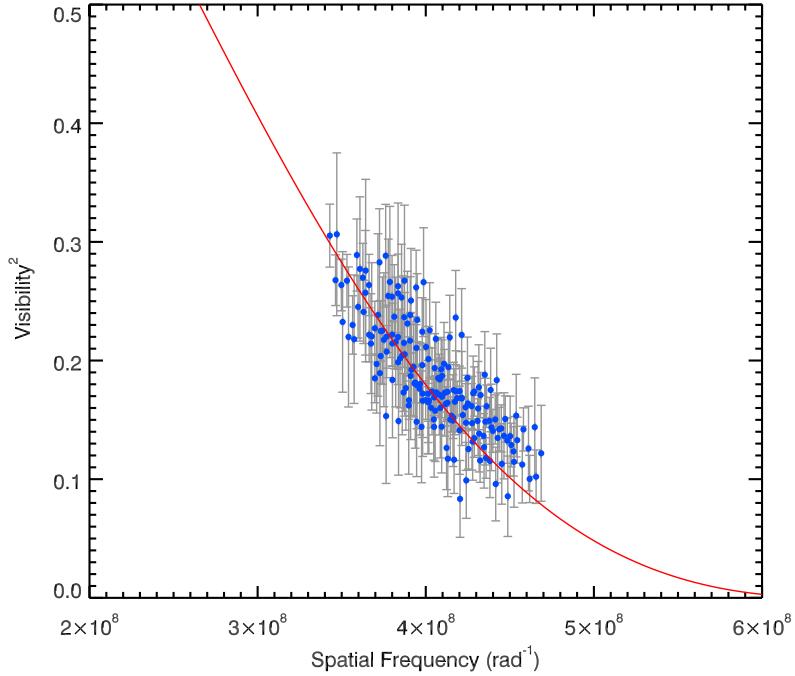


Figure 5.1 Observed squared visibility versus spatial frequency for HD 164922 (blue diamonds). The red line shows the best fit limb-darkened model. See Section 5.3.2 for a discussion of the implications of this figure.

previously known to host a single Saturn-mass planet orbiting with a semi-major axis of 2.1 AU (Butler et al., 2006). This target was one of several selected for more intensive long-term RV monitoring by Keck/HIRES based on both the stellar properties, and the mass and orbit of the previously detected planet making the system particularly well-suited for detecting additional low-mass planets. It was also on the Eta-Earth target list as part of a deep Doppler survey for low-mass planets (Howard et al., 2010a).

We measured the stellar radius for HD 164922 using the CHARA Array. Interferometric observations of HD 164922 were taken on 2012 May 13 and 14 using the Pavo beam combiner (ten Brummelaar et al., 2005b; Ireland et al., 2008). Observations of the science target were interleaved with the calibrator stars HD 164900, HD 161019, and HD 165373 (Bonneau et al., 2006, 2011). The data were reduced and calibrated using the standard data reduction pipeline (for details see White et al. 2013). We use the R -band limb-darkening coefficient from Claret & Bloemen (2011), $\mu_R = 0.633$, to determine a limb-darkened angular diameter $\theta_{UD} = \text{mas}$ (Figure 5.1).

Combining the angular diameter with the parallax yields a stellar radius of $0.999 \pm 0.017 R_\odot$. We determine a stellar bolometric flux of $F_{\text{Bol}} = 4.61 \pm 0.03 \text{ erg/s/cm}^2$ by fitting a spectral template from Pickles (1998) to flux calibrated photometry after applying revised filter profiles from Mann

& von Braun (2015). This translates to a luminosity $L = 0.703 \pm 0.017 L_\odot$. Lastly, we use our measured angular diameter with the star’s bolometric flux to derive an empirical effective temperature $T_{\text{eff}} = 5293 \pm 32$ K.

We then use the interferometrically determined parameters to inform a `SpecMatch` analysis of a stack of 5 high SNR APF spectra using an iterative technique. An initial uninformed `SpecMatch` analysis of the APF spectra (all priors uniform) gives $T_{\text{eff}} = 5318 \pm 70$ K, $\log g = 4.36 \pm 0.08$ and $[\text{Fe}/\text{H}] = 0.17 \pm 0.05$. We use this spectroscopically measured $[\text{Fe}/\text{H}]$ combined with the T_{eff} and R_\star determined from the CHARA data along with existing J, H , and K photometry (Cutri et al., 2003) as Gaussian priors in a fit to Dartmouth isochrones (Dotter et al., 2008) using the `isochrones` package (Morton, 2015)³. This gives the first estimate of the full set of stellar parameters. We then re-run `SpecMatch` with Gaussian priors applied to T_{eff} and $\log g$ from the `isochrones` output. The full likelihood with the Gaussian priors is

$$\mathcal{L} = \exp \left[-\frac{1}{2} \left(\chi^2 + \left(\frac{T_{\text{eff}} - 5293 \text{ K}}{32 \text{ K}} \right)^2 + \left(\frac{\log g - 4.387}{0.014} \right)^2 \right) \right]. \quad (5.1)$$

In this case the changes to stellar parameters become negligible after two iterations and the process is halted. The resulting stellar parameters are listed in Table 5.2.

Since the star is bright and already a known planet host it has been the subject of many spectroscopic studies. Santos et al. (2013) find $T_{\text{eff}} = 5356 \pm 45$ K, $\log g = 4.34 \pm 0.08$, and $[\text{Fe}/\text{H}] = +0.14 \pm 0.03$, all consistent with our analysis to within 1σ . Valenti & Fischer (2005) find a significantly higher value for $\log g = 4.51 \pm 0.06$ and $T_{\text{eff}} = 5385 \pm 44$ K but a consistent metallicity value of $[\text{Fe}/\text{H}] = +0.17 \pm 0.03$. Ghezzi et al. (2010) measure $T_{\text{eff}} = 5378 \pm 50$ K, $\log g = 4.30 \pm 0.22$, and $[\text{Fe}/\text{H}] = +0.21 \pm 0.03$ also using high-resolution spectroscopy which, except for T_{eff} , is also consistent with our analysis to within 1σ . As seen by the range of $\log g$ values obtained by these various studies, it can be difficult to pin down the stellar gravity from high resolution spectra alone. Since our $\log g$ value is constrained via the direct measurement of the stellar radii from interferometry we adopt the new slightly lower value for $\log g$.

5.3.3 HD 143761

HD 143761 is the closest and brightest star of the three studied in this work. The star is also known as ρ Corona Borealis, HIP 78459, and Gl 9537. It is a slightly evolved naked eye ($V = 5.41$) G0 V star (van Belle & von Braun, 2009) located at a distance of 17.236 pc (van Leeuwen, 2007). It was previously known to host a warm Jupiter-mass planet with an orbital period of 39 days (Noyes et al., 1997). This star was also part of the Eta-Earth survey and was independently selected for intensive long-term RV monitoring based on both the stellar properties, and the mass and orbit of the previously detected planet. Like HD 42618 and HD 164922 this star is chromosherically quiet

³<https://github.com/timothydmorton/isochrones>

with $R'_{\text{HK}} = -5.05$ (Isaacson & Fischer, 2010). As with HD 164922 we performed an iterative interferometric+spectroscopic analysis using the CHARA results from von Braun et al. (2014) and a stack of 5 high SNR APF spectra. The likelihood was essentially the same as Equation 5.1 but with the T_{eff} and $\log g$ values for HD 143761 substituted in the last two terms.

Valenti & Fischer (2005) measure $T_{\text{eff}} = 5822 \pm 44$ K, and Fuhrmann et al. (1998) measure $T_{\text{eff}} = 5821 \pm 20$ K both using high resolution spectroscopy. Both of these values are significantly hotter than our adopted value of $T_{\text{eff}} = 5627 \pm 54$ K from von Braun et al. (2014). We chose to adopt the value from von Braun et al. (2014) in order to maintain self-consistency with the interferometrically measured stellar radius and luminosity. Our metallicity value is also consistent within 1σ to that of Fuhrmann et al. (1998) and but is significantly lower than that of Valenti & Fischer (2005). Valenti & Fischer (2005) again measure a significantly higher value for $\log g = 4.36 \pm 0.06$, but our $\log g = 4.121 \pm 0.018$ value is consistent with $\log g = 4.12 \pm 0.1$ from Fuhrmann et al. (1998).

5.4 Keplerian Analysis

5.4.1 Discovery

We discovered each of the three new planets and re-discovered the previously known planets using a technique essentially identical to that of Fulton et al. (2015b). In brief, we calculate a two-dimensional Lomb-Scargle periodogram (2DKLS, O’Toole et al., 2009) to look for significant periodic signals that are well fit by a Keplerian orbital model. Our implementation of the 2DKLS periodogram incorporates arbitrary zero-point offsets between each instrument. The periodogram power (Z) represents the improvement to the χ^2 statistic relative to that of a baseline fit. When searching for the first planet in a system the baseline fit is simply a flat line or linear trend. If any significant signals are found after the first iteration the baseline model then becomes the single planet Keplerian model and we calculate the improvement to χ^2 when a second planet is added, without subtracting the first. We repeat this process until no more significant peaks are found in the 2DKLS periodogram. We start the search assuming no known signals in order to ensure that the previously published planets can be automatically detected using our pipeline. An initial jitter term of 2.0 m s^{-1} is added in quadrature with the RVs before starting the 2DKLS search in order to ensure fair weighting between the Keck and APF data sets. We expect both data sets to be limited by instrumental/astrophysical systematics rather than photon noise.

The discovery pipeline is completely automated in order to facilitate injection recovery tests that will allow us to characterize the pipeline completeness for future occurrence analysis(Howard & Fulton, 2016). We calculate an empirical periodogram false alarm probability (eFAP) by fitting a power law to the distribution of periodogram values between the 50th and 97th quartiles. This fit provides an estimate of the significance of periodogram peaks of a given value. When multiplied

by the number of independent test periods, the fit gives the approximate probability that we would find a peak of a given value within any particular periodogram. Any periodogram peak with an eFAP below 0.1% is automatically considered a viable candidate and the search is continued until no more periodogram peaks fall above the 0.1% eFAP threshold. Further details of the automated planet detection pipeline can be found in Howard & Fulton (2016). We note that the eFAP metric is used simply to automatically identify candidates. The significance of the corresponding periodogram peaks are checked using the bootstrapping technique described in Section 5.4.3. Each of the previously known planets were re-discovered with eFAPs much less than those of the new planets announced in this work.

We discover two significant signals in the RV time series of HD 42618. One with a long period of \sim 4850 days, and a second at a period of 149.6 days. Upon inspection of the Ca 2 H & K activity index time series we notice that this index shows a periodicity with a period very similar to that of the long period RV signal. The period of 4850 days is also very similar to the period of the sun's magnetic activity cycle. We conclude that this is likely the signature of the stellar magnetic activity cycle and not the signature of an orbiting planet. We include this long-period signal as an additional eccentric Keplerian in all further modeling. HD 42618 b is easily detected in the Keck data alone and the combined Keck+APF dataset but we don't yet have enough APF measurements to detect it in the APF data alone. Figure 5.2 shows the most likely model from the posterior of the two-Keplerian model, the 2DKLS periodogram used to discover HD 42618 b, and the RVs phased to the orbital period of planet b.

Wright et al. (2007) mentioned a candidate planetary signal with a period of 75.8 d and $K = 3 \text{ m s}^{-1}$ orbiting HD 164922 but did not have sufficient data to claim a significant detection. With seven years of additional Keck data, and 2 years of APF data we can firmly establish this signal as being coherent and persistent as expected for the Doppler motion caused by an orbiting planet. The short period planet is easily detected in either the APF or Keck data alone. The long-period planet can only be detected in the long baseline Keck data but we do observe a linear RV trend that emerged during the most recent APF observing season which is a result of the massive outer planet. Figures 5.3 and 5.4 show the most likely model from the posteriors for the two planet Keplerian model and the 2DKLS periodograms used to discover/re-discover each of the two planets.

We discover a super-Neptune mass planet orbiting HD 143761 exterior to the known Jupiter mass planet that has an orbital period of 39 days. The new planet has an orbital period of 102 days and a semi-amplitude of 3.7 m s^{-1} . Each planet is discovered with very high significance in both the Keck and Keck+APF datasets individually. The most likely model from the posteriors for the two planet model and the 2DKLS detection periodograms are shown in Figure 5.5.

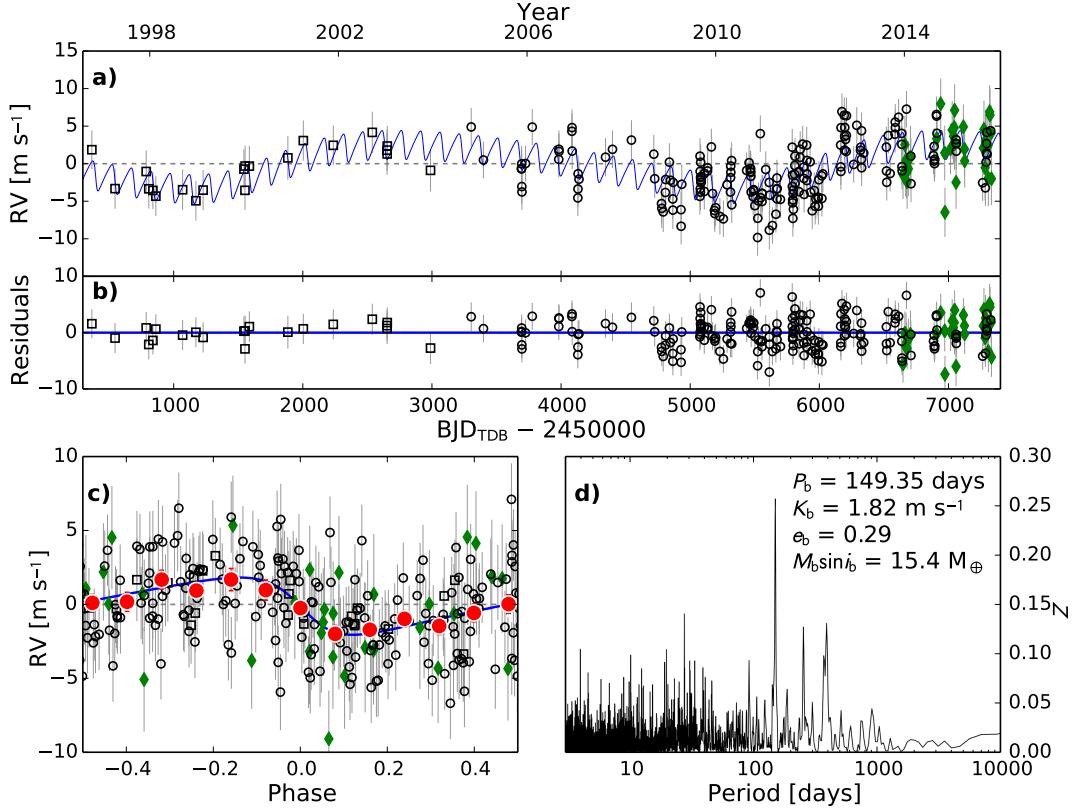


Figure 5.2 One-planet Keplerian orbital model plus one additional long-period Keplerian to model the stellar magnetic activity cycle for HD 42618. The most likely model is plotted but the orbital parameters annotated on the figure and listed in Tables 5.3 and 5.6 are the median values of the posterior distributions. The process used to find the orbital solution is described in 5.4.2. **a)** Full binned RV time series. Open black squares indicate pre-upgrade Keck/HIRES data (see §5.2), open black circles are post-upgrade Keck/HIRES data, and filled green diamonds are APF measurements. The thin blue line is the most probable 1-planet plus stellar activity model. We add in quadrature the RV jitter term listed in Table 5.3 with the measurement uncertainties for all RVs. **b)** Residuals to the most probable 1-planet plus stellar activity model. **c)** Binned RVs phase-folded to the ephemeris of planet b. The long-period stellar activity signal has been subtracted. The small point colors and symbols are the same as in panel **a**. For visual clarity, we also bin the velocities in 0.08 units of orbital phase (red circles). The phase-folded model for planet b is shown as the blue line. **d)** 2DKLS periodogram showing the improvement to χ^2 for a model including the long period activity signal and a single planet compared to a model that only includes the activity signal.

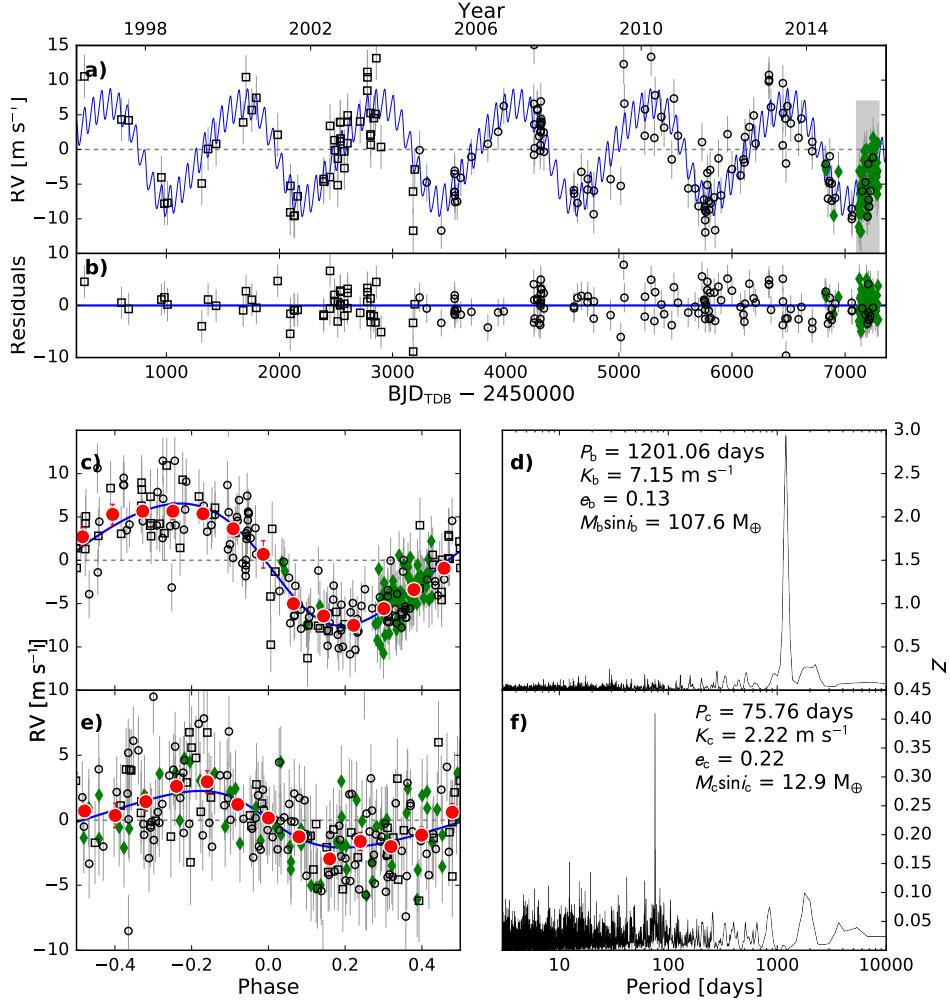


Figure 5.3 Two-planet Keplerian orbital model posterior distributions for HD 164922. The most likely model from the posterior distribution is plotted while the orbital parameters annotated on the figure and listed in Tables 5.4 and 5.6 are the median values of the posterior distributions. The process used to find the orbital solution is described in 5.4.2. **a)** Full binned RV time series. Open black squares indicate pre-upgrade Keck/HIRES data (see §5.2), open black circles are post-upgrade Keck/HIRES data, and filled green diamonds are APF measurements. The thin blue line is the most probable 2-planet model. We add in quadrature the RV jitter term listed in Table 5.4 with the measurement uncertainties for all RVs. **b)** Residuals to the most probable 2-planet model. **c)** Binned RVs phase-folded to the ephemeris of planet b. The Keplerian orbital model for planet c has been subtracted. The small point colors and symbols are the same as in panel **a**. For visual clarity, we also bin the velocities in 0.08 units of orbital phase (red circles). The phase-folded model for planet b is shown as the blue line. **d)** 2DKLS periodogram comparing a model including only the long period planet to the two planet model. Panels **e)** and **f)** are the same as panels **c)** and **d)** but for planet HD 164922 c. The shaded region of panel **a** is re-plotted in Figure 5.4.

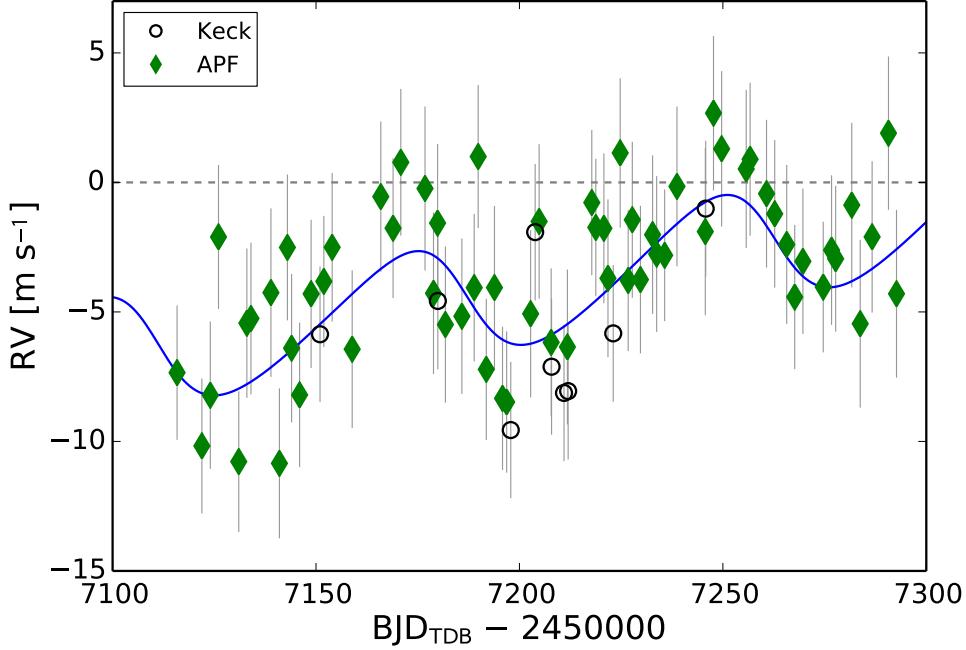


Figure 5.4 Recent RVs for HD 164922 highlighted in the grey box of Figure 5.3 panel a) featuring the high cadence APF observations collected during the most recent observational season.

5.4.2 Characterization

We estimated orbital parameters and their associated uncertainties using the ExoPy Differential Evolution Markov Chain Monte Carlo (DE-MCMC, Ter Braak, 2006) modeling code and a technique identical to that of Fulton et al. (2013), Knutson et al. (2014), and Fulton et al. (2015b).

Our multi-planet RV model is a sum of Keplerian single-planet models over all planets in the system. For each single-planet Keplerian model (indexed by i) we compute posteriors for the orbital period (P_i), time of inferior conjunction ($T_{\text{conj},i}$), eccentricity (e_i), argument of periastron of the star's orbit (ω_i), velocity semi-amplitude (K_i). We also compute posteriors for the offsets between pre-upgrade, post-upgrade, and the APF datasets (γ), and an RV jitter term with the specific prior described in Fulton et al. (2015b) and Johnson et al. (2011). In order to speed convergence, we choose to re-parameterize some of the physical parameters as $\sqrt{e} \cos \omega$, $\sqrt{e} \sin \omega$, and $\log K$. The full likelihood for this Keplerian model is

$$\mathcal{L} = \exp \left(-\frac{1}{2} \sum_n \left[\frac{(v_n - M_n)^2}{\sigma_{v,n}^2 + \sigma_{\text{jitt}}^2} + 2 \ln (\sqrt{2\pi}(\sigma_{v,n}^2 + \sigma_{\text{jitt}}^2)^{1/4}) \right] \right) \quad (5.2)$$

summed over n RV measurements (v_n) with associated uncertainties $\sigma_{v,n}$. M_n is the Keplerian model for observation n . We assign uniform priors to $\log P$, T_{conj} , $\sqrt{e} \cos \omega$, $\sqrt{e} \sin \omega$, $\log K$, and γ for each instrument. We follow the prescription of Eastman et al. (2013) checking for convergence at

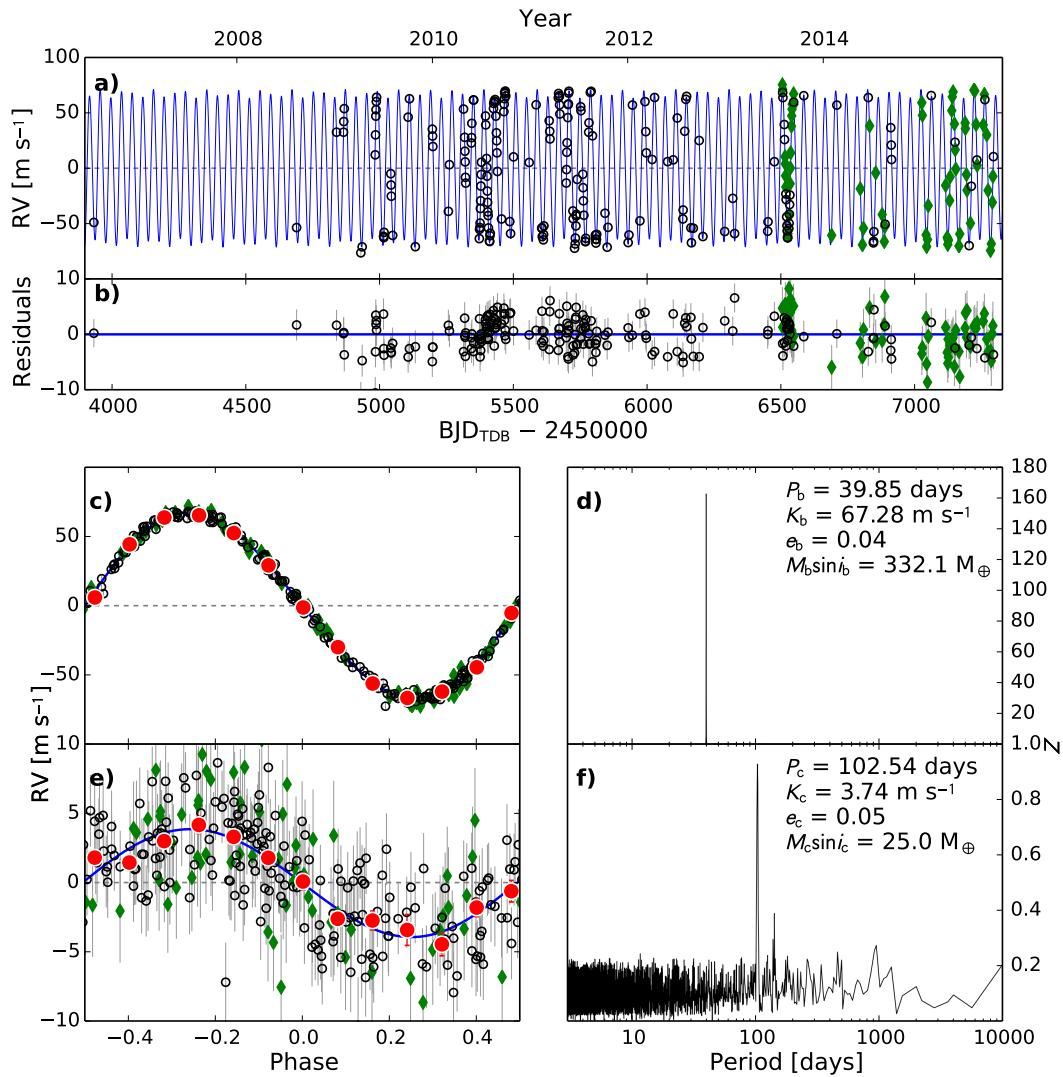


Figure 5.5 Same as Figure 5.3 but for planets HD 143761 b and c.

Table 5.3. Orbital Parameters for HD
42618

Parameter	Value	Units
Modified DE-MCMC Step Parameters¹		
$\log(P_b)$	$2.17497^{+0.0011}_{-0.00098}$	log(days)
$\sqrt{e_b} \cos \omega_b$	-0.03 ± 0.24	
$\sqrt{e_b} \sin \omega_b$	$+0.34^{+0.19}_{-0.31}$	
$\log(K_b)$	$0.276^{+0.061}_{-0.071}$	log(m s ⁻¹)
Model Parameters		
P_b	$149.61^{+0.37}_{-0.34}$	days
$T_{\text{conj},b}$	$2456670.2^{+6.1}_{-5.6}$	BJD _{TDB}
e_b	$0.19^{+0.15}_{-0.12}$	
ω_b	101^{+69}_{-39}	degrees
K_b	$1.89^{+0.29}_{-0.28}$	m s ⁻¹
$\gamma_{\text{post-upgrade Keck}}$	0.61 ± 0.21	m s ⁻¹
$\gamma_{\text{pre-upgrade Keck}}$	$0.71^{+0.64}_{-0.65}$	m s ⁻¹
$\dot{\gamma}$	$\equiv 0.0 \pm 0.0$	m s ⁻¹ day ⁻¹
$\ddot{\gamma}$	$\equiv 0.0 \pm 0.0$	m s ⁻¹ day ⁻²
σ_{jitt}	$2.34^{+0.14}_{-0.12}$	m s ⁻¹

¹MCMC jump parameters that were modified from the physical parameters in order to speed convergence and avoid biasing parameters that must physically be finite and positive.

regular intervals during the MCMC runs by calculating the Gelmin-Rubin statistic and the number of independent draws (T_Z , Ford, 2006). We consider the chains well mixed and halt the MCMC run when the Gelmin-Rubin statistic is within 1% of unity and $T_Z > 1000$ for all free parameters. All of the adopted median values and 68% confidence intervals of the posterior distributions are listed in Tables 5.3–5.5.

5.4.3 Bootstrap False Alarm Assessment

We conduct a bootstrap false alarm assessment to verify and double check that the periodogram peaks with low eFAPs are indeed statistically significant periodic signals and not caused by random fluctuations of noise. For all three stars we scramble the RV time series 1000 times and recalculate the 2DKLS periodogram searching for N+1 planets where N is the number of previously published planets in the system. We record the highest periodogram value from each trial and plot the distribution of periodogram peak heights relative to the periodogram peak values corresponding to

Table 5.4. Orbital Parameters for HD
164922

Parameter	Value	Units
Modified DE-MCMC Step Parameters¹		
$\log(P_b)$	3.08 ± 0.002	log(days)
$\sqrt{e_b} \cos \omega_b$	$-0.214^{+0.100}_{-0.081}$	
$\sqrt{e_b} \sin \omega_b$	$+0.264^{+0.096}_{-0.130}$	
$\log(K_b)$	0.854 ± 0.019	m s^{-1}
$\log(P_c)$	$1.87947^{+0.00033}_{-0.00032}$	log(days)
$\sqrt{e_c} \cos \omega_c$	$+0.04^{+0.29}_{-0.30}$	
$\sqrt{e_c} \sin \omega_c$	$+0.37^{+0.14}_{-0.23}$	
$\log(K_c)$	$0.346^{+0.054}_{-0.062}$	log(m s^{-1})
Model Parameters		
P_b	$1201.1^{+5.6}_{-5.5}$	days
$T_{\text{conj},b}$	2456778^{+18}_{-19}	BJD _{TDB}
e_b	$0.126^{+0.049}_{-0.050}$	
ω_b	129^{+24}_{-20}	degrees
K_b	7.15 ± 0.31	m s^{-1}
P_c	$75.765^{+0.058}_{-0.056}$	days
$T_{\text{conj},c}$	2456277.6 ± 2.7	BJD _{TDB}
e_c	0.22 ± 0.13	
ω_c	81^{+45}_{-49}	degrees
K_c	$2.22^{+0.30}_{-0.29}$	m s^{-1}
$\gamma_{\text{post-upgrade Keck}}$	0.23 ± 0.27	m s^{-1}
$\gamma_{\text{pre-upgrade Keck}}$	1.02 ± 0.54	m s^{-1}
γ_{APF}	$0.2^{+0.47}_{-0.48}$	m s^{-1}
$\dot{\gamma}$	$\equiv 0.0 \pm 0.0$	$\text{m s}^{-1}\text{day}^{-1}$
$\ddot{\gamma}$	$\equiv 0.0 \pm 0.0$	$\text{m s}^{-1}\text{day}^{-2}$
σ_{jitt}	$2.63^{+0.15}_{-0.14}$	m s^{-1}

¹MCMC jump parameters that were modified from the physical parameters in order to speed convergence and avoid biasing parameters that must physically be finite and positive.

Table 5.5. Orbital Parameters for HD 143761

Parameter	Value	Units
Modified DE-MCMC Step Parameters¹		
$\log(P_b)$	$1.600382 \pm 1.6e - 05$	log(days)
$\sqrt{e_b} \cos \omega_b$	$+0.002 \pm 0.02$	
$\sqrt{e_b} \sin \omega_b$	$-0.192^{+0.011}_{-0.010}$	
$\log(K_b)$	1.8279 ± 0.0016	m s^{-1}
$\log(P_c)$	$2.01091^{+0.00073}_{-0.00070}$	log(days)
$\sqrt{e_c} \cos \omega_c$	$+0.01 \pm 0.19$	
$\sqrt{e_c} \sin \omega_c$	-0.01 ± 0.19	
$\log(K_c)$	$0.573^{+0.032}_{-0.034}$	$\log(\text{m s}^{-1})$
Model Parameters		
P_b	$39.8458^{+0.0015}_{-0.0014}$	days
$T_{\text{conj},b}$	2455759.091 ± 0.056	BJD _{TDB}
e_b	$0.0373^{+0.0040}_{-0.0039}$	
ω_b	$270.6^{+5.9}_{-5.8}$	degrees
K_b	67.28 ± 0.25	m s^{-1}
P_c	102.54 ± 0.17	days
$T_{\text{conj},c}$	2455822 ± 2	BJD _{TDB}
e_c	$0.052^{+0.061}_{-0.037}$	
ω_c	190^{+110}_{-140}	degrees
K_c	3.74 ± 0.28	m s^{-1}
$\gamma_{\text{post-upgrade Keck}}$	-0.6 ± 0.2	m s^{-1}
γ_{APP}	$-0.7^{+0.50}_{-0.49}$	m s^{-1}
$\dot{\gamma}$	$\equiv 0.0 \pm 0.0$	$\text{m s}^{-1}\text{day}^{-1}$
$\ddot{\gamma}$	$\equiv 0.0 \pm 0.0$	$\text{m s}^{-1}\text{day}^{-2}$
σ_{jitt}	$2.57^{+0.14}_{-0.13}$	m s^{-1}

¹MCMC jump parameters that were modified from the physical parameters in order to speed convergence and avoid biasing parameters that must physically be finite and positive.

Table 5.6. Derived Planet Properties

Parameter	Value	Units
HD 42618		
$e_b \cos \omega_b$	-0.009 $^{+0.06}_{-0.076}$	
$e_b \sin \omega_b$	0.14 $^{+0.13}_{-0.11}$	
a_b	0.554 ± 0.011	AU
a_b/R_\star	119.1 $^{+12.0}_{-9.8}$	
$M_b \sin i_b$	14.4 $^{+2.5}_{-2.4}$	M_\oplus
S_b^*	3.16 $^{+0.61}_{-0.55}$	S_\oplus
$T_{eq,b}^{**}$	337 $^{+15}_{-16}$	K
HD 164922		
$e_b \cos \omega_b$	-0.025 $^{+0.016}_{-0.019}$	
$e_b \sin \omega_b$	0.032 $^{+0.028}_{-0.022}$	
a_b	2.115 ± 0.012	AU
a_b/R_\star	454.9 $^{+8.3}_{-7.9}$	
$M_b \sin i_b$	107.6 $^{+4.9}_{-4.8}$	M_\oplus
S_b^*	0.1578 $^{+0.0069}_{-0.0067}$	S_\oplus
$T_{eq,b}^{**}$	159.4 ± 1.7	K
$e_c \cos \omega_c$	0.003 $^{+0.073}_{-0.063}$	
$e_c \sin \omega_c$	0.079 $^{+0.089}_{-0.066}$	
a_c	0.3351 ± 0.0015	AU
a_c/R_\star	72.1 $^{+1.3}_{-1.2}$	
$M_c \sin i_c$	12.9 ± 1.6	M_\oplus
S_c^*	6.29 $^{+0.27}_{-0.26}$	S_\oplus
$T_{eq,c}^{**}$	400.5 ± 4.3	K
HD 143761		
$e_b \cos \omega_b$	7e-05 $^{+0.00072}_{-0.00073}$	
$e_b \sin \omega_b$	-0.0072 $^{+0.0011}_{-0.0012}$	
a_b	0.2196 $^{+0.0024}_{-0.0025}$	AU
a_b/R_\star	34.66 $^{+0.78}_{-0.76}$	
$M_b \sin i_b$	332.1 $^{+7.5}_{-7.6}$	M_\oplus
S_b^*	34.7 $^{+2.1}_{-2.0}$	S_\oplus
$T_{eq,b}^{**}$	614.0 $^{+9.1}_{-9.0}$	K
$e_c \cos \omega_c$	0.0001 $^{+0.013}_{-0.011}$	
$e_c \sin \omega_c$	-0.0001 $^{+0.011}_{-0.015}$	
a_c	0.4123 $^{+0.0046}_{-0.0047}$	AU
a_c/R_\star	65.1 $^{+1.5}_{-1.4}$	
$M_c \sin i_c$	25 ± 2	M_\oplus
S_c^*	9.85 $^{+0.6}_{-0.56}$	S_\oplus
$T_{eq,c}^{**}$	448.1 ± 6.6	K

*Stellar irradiance received at the planet relative to the Earth.

**Assuming a bond albedo of 0.32 (Demory, 2014).

the newly discovered planets. These distributions are plotted in Figure 5.6. The periodogram peak heights corresponding to each of the new planets are well separated from the distribution of peaks in the scrambled RV trials. This indicates that the probability that random noise could conspire to create the periodogram peaks used to detect the planets are $< 0.1\%$. However, a visual inspection of the distribution of periodogram peak heights in Figure 5.6 suggests that the FAPs are likely much lower.

5.4.4 Brown Dwarf Companion to HD 143761?

HD 143761 b was one of the first exoplanets discovered (Noyes et al., 1997). Gatewood et al. (2001) later reported that they detected the signature of HD 143761 b in astrometric data from the Multichannel Astrometric Photometer and Hipparcos. Zucker & Mazeh (2001) were quick to point out that the statistical significance of this astrometric detection is only 2σ . Over a decade after the discovery of HD 143761 b, Reffert & Quirrenbach (2011) claimed to detect the same astrometric signal of the warm Jupiter in a re-reduction of Hipparcos data. The astrometric orbit suggests that the system is nearly face on with an inclination between 0.4 and 0.7 degrees. This would imply that, after correcting for the viewing angle, HD 143761 b is not a planet but instead a low mass M star with $100 < M_b < 200 \text{ M}_J$.

Interferometry of HD 143761

Long-baseline interferometry is sensitive to some stellar binaries. HD 143761 was observed interferometrically with the CHARA Array (von Braun et al., 2014). If we assume a face-on orbit of HD 143761 b at a distance of 17.2 pc, then the angular distance between it and the principal component is 14 milliarcseconds independent of phase angle. This is detectable as a separated fringe packet at H -band with CHARA (Farrington et al., 2010), provided the brightness contrast is not larger than $\Delta H \sim 2$ (Farrington et al., 2010; Raghavan et al., 2012).

Using the mass constraints from the astrometric orbit, a comparison with the Dartmouth isochrones for metallicities spanning the 1σ [Fe/H] uncertainty of HD 143761 from Table 5.2 corresponds to an apparent H magnitude at a distance of 17.2 pc of around 10. This implies a $\Delta H \simeq 6$, which is significantly below what could be detected as a separate component in CHARA data (cf. equation A5 in Boyajian et al., 2008). Since component c's apparent H magnitude is much fainter than component b's, its detection is impossible in the CHARA data. As expected, an inspection of the CHARA data used in von Braun et al. (2014) did not yield any indication of additional fringe packets.

We also obtained imaging observations of ρ CrB using the Differential Speckle Survey Instrument (DSSI) on Gemini-North during the nights of July 19, 24, and 25 in 2014. The DSSI camera is a dual-channel speckle imaging system, expounded upon in more detail by Horch et al. (2009, 2011). Observations were acquired using red and near-infrared filters centered on 692 nm and 880

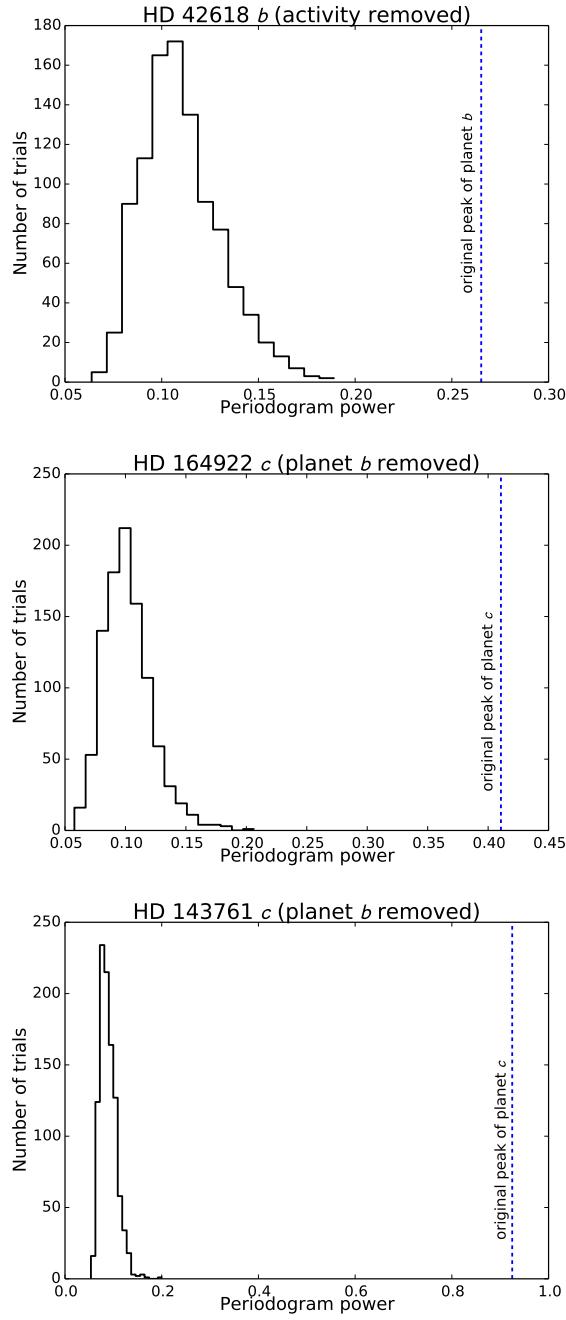


Figure 5.6 Graphical representation of the bootstrap false alarm tests described in Section 5.4.3. *Top:* Distribution of maximum periodogram peak heights for 1000 2DKLS periodograms of scrambled RV time series for HD 42618. The long period activity signal was subtracted before scrambling the data set. The vertical dashed blue line marks the height of the original periodogram peak for planet *b* which is clearly separated from the distribution of peaks caused by random fluctuations. *Middle:* Same as the top panel for planet HD 164922 *c*. *Bottom:* Same as the top panel for planet HD 143761 *c*. In each case, the scrambled RVs generate peaks with significantly lower power than the power observed from the new planetary signals.

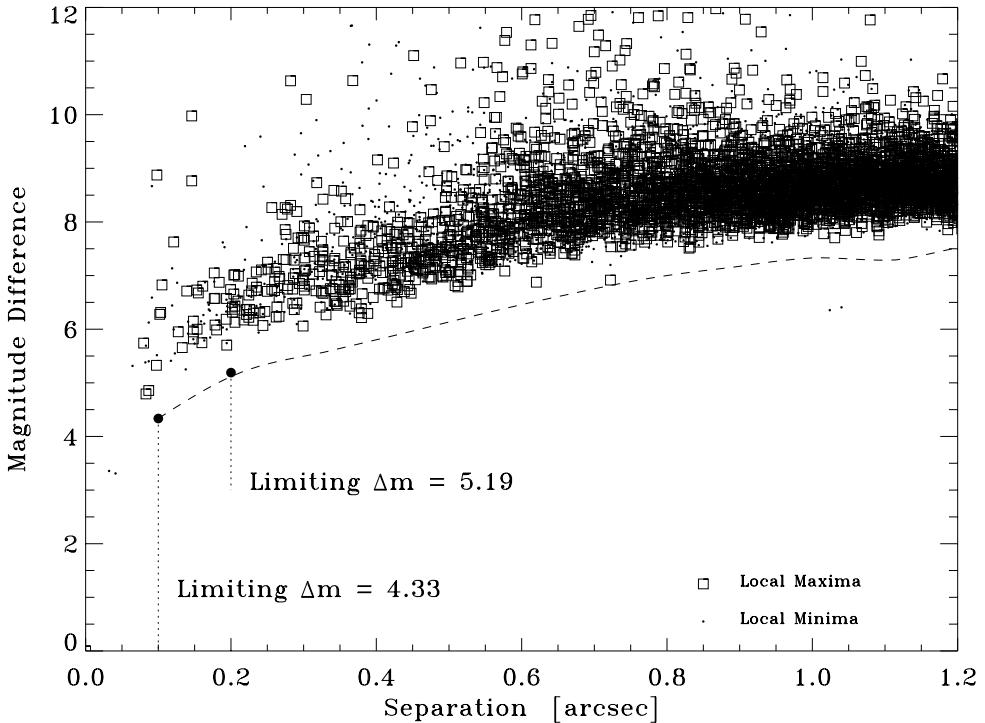


Figure 5.7 Limiting magnitude as a function of separation from the ρ CrB. Also shown are a cubic spline interpolation of the 5σ detection limit (dashed line) and limiting magnitudes for $0''.1$ and $0''.2$.

nm respectively. Our instrument setup is the same as that described in Horch et al. (2012) and our analysis methodology is outlined by Kane et al. (2014). Briefly, we estimate the limiting magnitude Δm (difference between local image maxima and minima) as a function of target separation resulting in a 5σ detection curve. More details on the derivation of the DSSI detection limits can be found in Howell et al. (2011). All of our DSSI ρ CrB observations show no evidence of a stellar companion to the host star. Figure 5.7 shows the detection curve from the 880 nm image acquired for ρ CrB on the night of July 25 2014. The dashed curve is the cubic spline interpolation of the 5σ detection limit from $0''.1$ to $1''.2$. The results exclude companions with $\Delta m \sim 5.2$ and $\Delta m \sim 7.5$ at separations of $0''.1$ and $1''.4$ respectively. Given the distance of ρ CrB of 17.236 pc, these angular separations correspond to a physical exclusion range of 1.7–24.1 AU. We can thus rule out stellar companions in close proximity to the host star, supporting the evidence that the system is not a face-on triple star system, but instead a multi-planet system viewed at moderate to high inclination.

Stability of the HD 143761 System

Two planet systems on circular orbits are likely to be unstable if $\Delta < 2\sqrt{3}$ Gladman (1993) for all mutual inclinations, where $\Delta = \frac{a_{in} - a_{out}}{R_H}$ and R_H is the Hill radius,

$$R_H = \left(\frac{M_{in} + M_{out}}{3M_*} \right)^{1/3} \frac{a_{in} + a_{out}}{2}. \quad (5.3)$$

Sky-projected inclinations smaller than 4 degrees combined with the $M \sin i$ constraints listed in Table 5.6 imply large companion masses and push Δ below the $2\sqrt{3}$ stability threshold. If we assume that the system is dynamically stable then the presence of HD 143761 c rules out the low inclination orbit found by Gatewood et al. (2001), and Reffert & Quirrenbach (2011).

HD 164922 b and c are widely separated and intuitively we would expect them to be in a stable configuration. For completeness, we calculate $\Delta = 26.3$ for the minimum masses which suggests that this system configuration is likely to be dynamically stable for a long time.

5.4.5 Additional Planet Candidates

There is an additional significant periodic signal in the RV data for HD 42618 at a period of 388 days and a velocity semi-amplitude of 2 m s^{-1} . This would be a $M \sin i = 22 M_\oplus$ planet orbiting just outside 1 AU. The periodogram peak for this candidate falls above the eFAP threshold automatically calculated by our discovery pipeline. However, we do not believe that we have enough evidence to claim a concrete detection of a bona fide planet due to the proximity of this period to 1 year and incomplete phase coverage of the orbit. Telluric contamination of the template or problems with the barycentric correction could inject a false signal with a period near 1 year (Wright & Eastman, 2014; Fischer et al., 2016, S. Wang personal communication). The amplitude and periodicity of this signal depends on the outcomes of modeling HD 42618 b, and MCMC runs for models including this candidate fail to converge. We will continue to monitor this star intensively with both Keck and the APF to confirm or refute this planet candidate in the upcoming years.

We find a candidate periodicity in the HD 164922 system with a period of 41.7 d and an amplitude of 1.9 m s^{-1} . The eFAP of this 2DKLS periodogram peak is 0.00098 and falls just above our 0.1% eFAP threshold. However, we do not consider this to be a viable planet candidate due to its marginal detection and proximity to the expected rotation period for this star (44 days, Isaacson & Fischer, 2010). Further monitoring and a detailed analysis that includes the effects of rotational modulation of starspots is needed to determine the nature of this signal.

There is no evidence for significant periodic signals from other candidates in the periodograms for HD 143761. However, visually there appears to be some long-period structure in the residuals to our most probable model (see Figure 5.5). This marginal variability, if real, likely has a period of ≥ 10 years and an amplitude of only a few m s^{-1} and it appears to be at a shorter period than

the stellar magnetic activity cycle as seen in the S_{HK} values for this star. Long-term monitoring of this target is required to determine if this signal is real and the signature of a planetary companion.

5.4.6 Chromospheric Activity

These stars were all selected to be part of the APF-50 survey of nearby stars due, in part, to their extremely low mean chromospheric activity of $R'_{HK} \leq -4.95$. However, in the case of HD 42618 we do detect significant long-period variability in the S_{HK} values that is strongly correlated with the RVs (see Figure 5.8) that is likely the signature of the stellar magnetic activity cycle. We do not find any significant periodic signals in the S_{HK} values after removal of this long-period trend that might be the signature of rotation. However, we clearly identify the rotation period of the star to be 16.9 days in CoRoT photometry (see Section 5.5.1). We account for the activity cycle in the RV data of HD 42618 by including an additional long-period Keplerian signal in the model.

HD 164922 shows only a linear trend in the S_{HK} values but we do not detect the effect of this change in chromospheric activity in the residuals to the two planet fit. There is also a very weak peak in the Lomb-Scargle periodogram (L-S, Lomb, 1976; Scargle, 1982) of the S_{HK} values at 37.8 days (see Figure 5.9). This may be the signature of stellar rotation since this is near the expected period for a star of this type and age (Isaacson & Fischer, 2010). However, this rotation period is well separated from the orbital periods of the two planets and does not influence our two planet fits.

We do not detect any long term variability in the S_{HK} values of HD 143761 but we see a clear peak in a periodogram of the S_{HK} values at 18.5 days that is likely caused by the rotational modulation of star spots (see Figure 5.9). Since the rotation period is well separated from the periods of either of the planets orbiting HD 143761 this does not affect our Keplerian modeling and is likely absorbed into the stellar jitter term.

5.5 Photometry

5.5.1 CoRoT Photometry of HD 42618

HD 42618 was the target of high cadence, high precision, continuous photometric monitoring for ≈ 0.5 years with the purpose of detecting solar like oscillations (Baglin et al., 2012). We perform a simple polynomial detrending of the space-based photometry. After removing large ramp-shaped features at the start of two long observing campaigns we then fit an 8th order polynomial to all continuous segments of the data. These segments are 2-20 days in length.

We use the detrended photometry to look for periodic photometric variability that might be caused by rotationally modulated star spots. We detect significant variability with a period of 16.9 days but with a broad distribution of periodogram power around the highest peak (see Figure 5.10). This is a clear signature of stellar rotation with slightly changing phase and/or differential

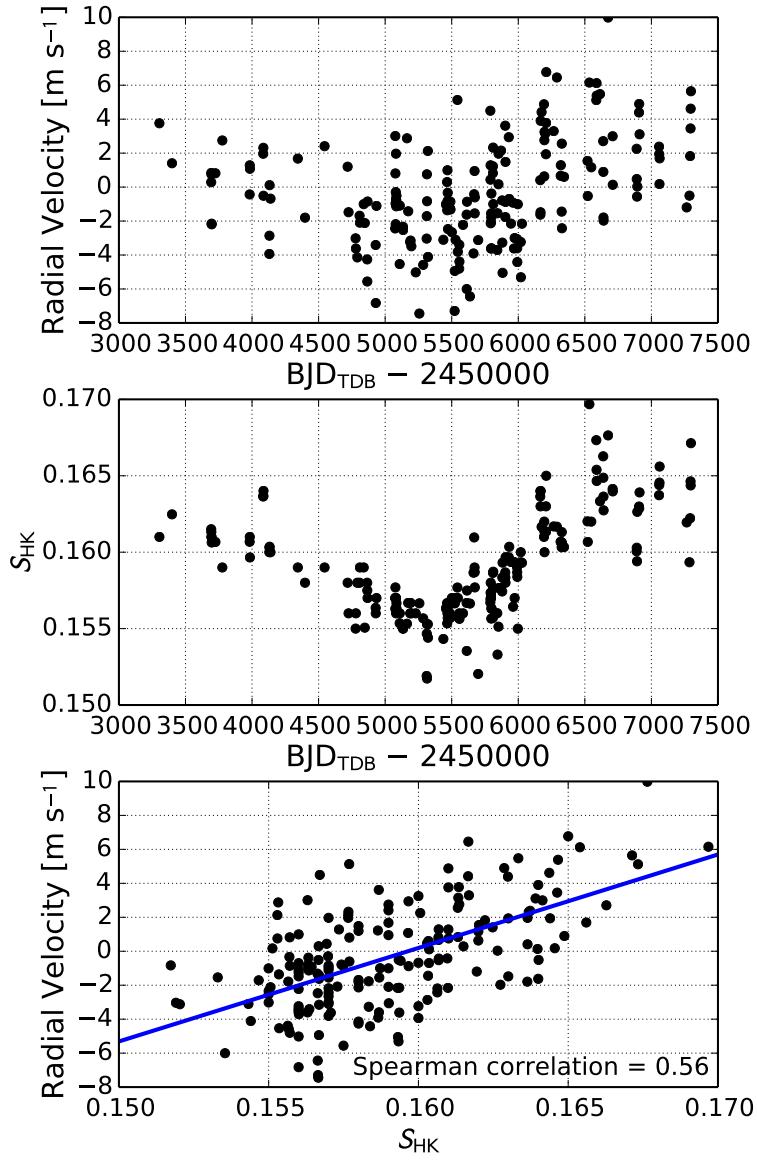


Figure 5.8 Velocity-activity correlation for HD 42618. A discussion of the chromospheric activity of each of the three stars can be found in Section 5.4.6. *Top:* Binned RV time series of the post-upgrade Keck data with planet b. *Middle:* Binned S_{HK} time series of the post-upgrade Keck data only. Note the similarities between the variability in the top and middle panels. *Bottom:* Spearman rank correlation test of the velocities with S_{HK} values (Spearman, 1904). We do not subtract this correlation from the RVs of HD 42618 but instead model the magnetic activity cycle as an additional long-period Keplerian (see Section 5.4.2).

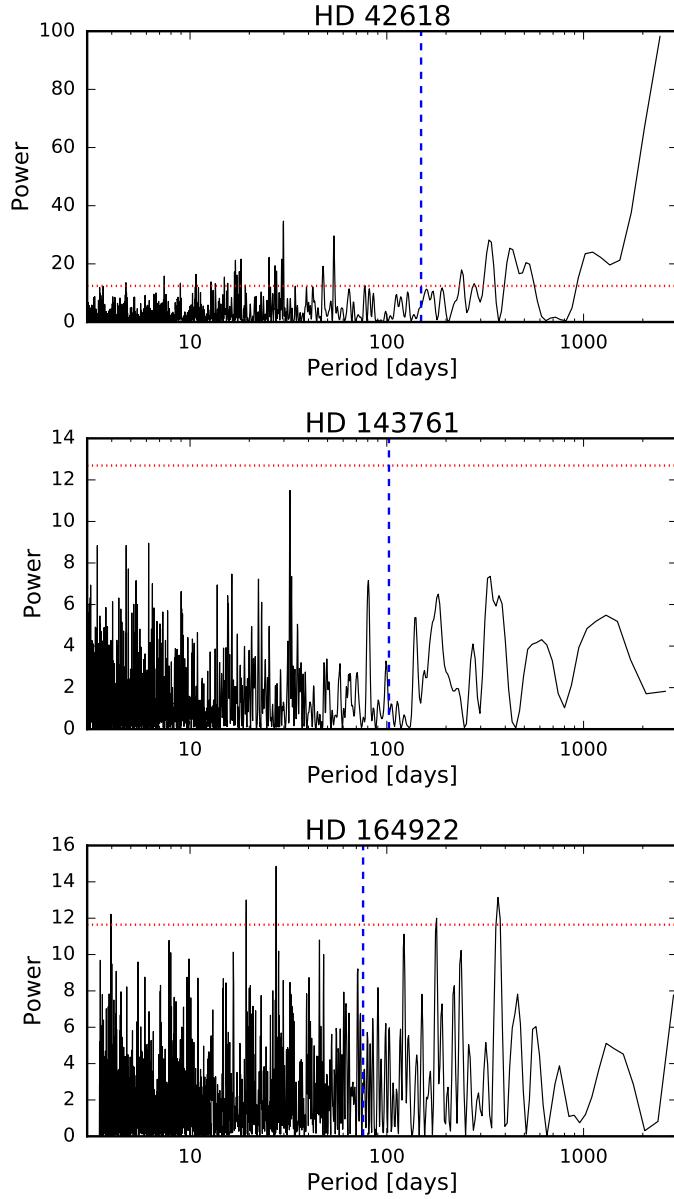


Figure 5.9 Lomb-Scargle periodograms of S_{HK} chromospheric activity. In each panel the period of the planet announced in this work is marked by the blue dashed line and the power corresponding to an analytical false alarm probability of 1% is marked by the red dotted line (Schwarzenberg-Czerny, 1998). S_{HK} values measured from spectra with $\text{SNR} < 40$ per pixel or exposure times $> 25\%$ longer than the median exposure time (due to clouds and/or seeing) can be badly contaminated by the solar spectrum and cause our S_{HK} extraction pipeline to produce large outliers. These measurements were excluded before calculating the periodograms. No significant periodicity is detected in any of the stars at the orbital periods of the new planets. *Top:* Periodogram of S_{HK} values for HD 42618. *Middle:* Periodogram of S_{HK} values for HD 143761. *Bottom:* Periodogram of S_{HK} values for HD 164922.

rotation which creates a broad distribution of increased power in Fourier space near the true rotation period. This star is very similar to the Sun in mass, age, and chemical abundance so the fact that the rotation period is also similar to that of the Sun (26 days) is not surprising. However, we note that the precise location of the highest peak in the Lomb-Scargle periodogram of HD 42618 depends on the polynomial order used to detrend the CoRoT photometry. We also tried high-pass filtering the CoRoT photometry using running median filters with window widths of 20-50 days and only analyzing continuous segments of data longer than 20 days. We found that the period of highest power is somewhat variable but always falls between 12-18 days. Since the periodogram period is dependent on the detrending algorithm we can't determine the rotation period of HD 42618 precisely, but we estimate that it falls within the range of 12-18 days. We do not detect any significant periodic signal in the RV data near the photometric period.

We searched through the detrended CoRoT light curve using the TERRA planet detection algorithm (Petigura et al., 2013b,c). We did not find any periodic box-shaped dimmings with signal-to-noise ratios (SNR) greater than 7. Searches of Kepler photometry commonly require $\text{SNR} > 7$ (Jenkins et al., 2010) or $\text{SNR} > 12$ (Petigura et al., 2013c), though the SNR threshold depends on the noise structure of the photometry. We conclude that there are no transiting planets having periods between 0.5 and 60 days with transits that are detectable above Poisson, stellar, and instrumental noise. Given the photometric noise properties of HD 42618, we can rule out planets with transits deeper than ≈ 150 ppm ($\approx 1.3 R_{\oplus}$) at ≈ 5 day orbital periods and transits deeper than ≈ 300 ppm ($\approx 1.9 R_{\oplus}$) for ≈ 50 day orbital periods. The *a priori* transit probability for HD 42618 b is only 0.8% so it is not surprising that we do not detect transits.

Asteroseismic Mass Determination

Convection in the outer layers of a star excites stochastic oscillations, which can be observed on the stellar surface. In the case of main sequence stars, these oscillations manifest themselves as periodic variations on the order of cm s^{-1} in radial velocity data or $\sim \text{ppm}$ in photometric data. Photometric space telescopes such as CoRoT proved to be quite effective for measuring and characterizing these oscillations, which can be used to derive global stellar properties (such as radius, mass and age), as well as to constrain the stellar interior (e.g. Michel et al., 2008; Chaplin & Miglio, 2013).

We measured the mass of HD 42618 from CoRoT photometry obtained during two long observing runs spanning 79 and 94 days, respectively. Through a Fourier analysis of the CoRoT lightcurve, we produced the power spectral density function shown in Figure 5.12. We then stacked the power spectrum in equally sized pieces to create an echelle diagram, revealing the distinct $l=0$, 1, and 2 latitudinal modes of oscillation. We then collapsed this echelle diagram, effectively creating a binned power spectrum, and fit the power excess with a Gaussian to measure a maximum oscillation power frequency ν_{max} of 3.16 ± 0.10 mHz. We then collapsed the echelle diagram along the perpendicular axis to preserve the frequency spacing, computed its autocorrelation, and fit the

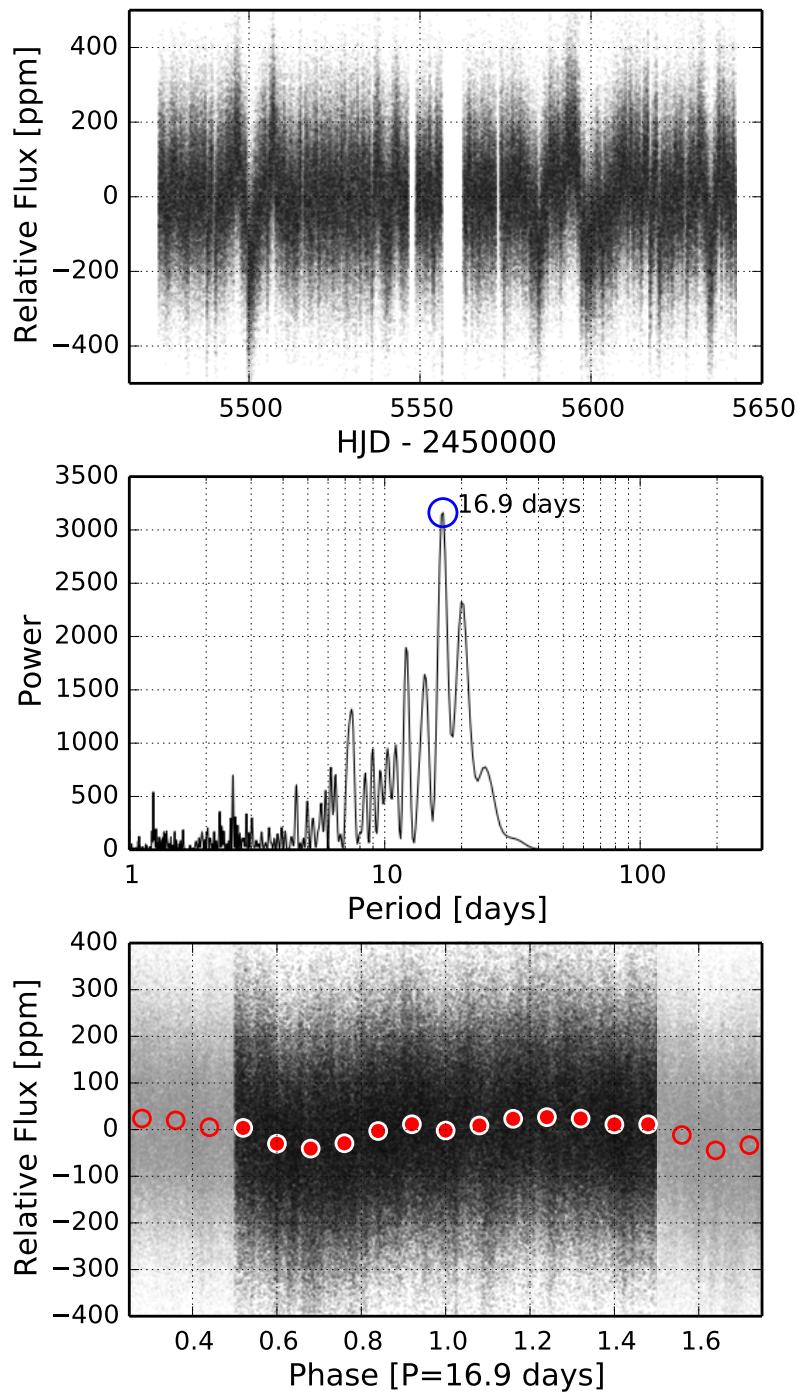


Figure 5.10 CoRoT photometry of HD 42618 discussed in Section 5.5.1. *Top:* Detrended light curve. *Middle:* Lomb-scargle periodogram of the light curve. *Bottom:* Photometry phase-folded to the period corresponding to the highest peak in the Lomb-Scargle periodogram (16.9 days). We also bin the photometry with bin widths of 0.05 units of phase (red circles).

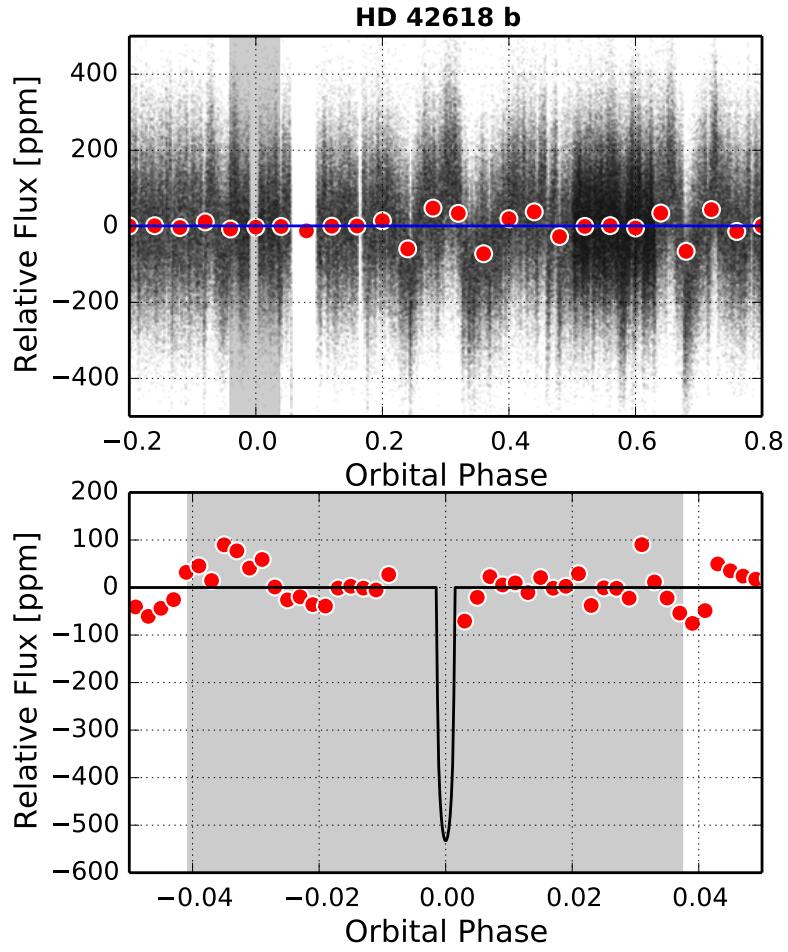


Figure 5.11 CoRoT photometry of HD 42618 phase-folded to the orbital period of planet b. The transit search for HD 42618 b is discussed in Section 5.5.1. *Top:* Photometry over the full orbital phase of planet b. The red circles are binned photometric measurements with bin widths of 0.04 units of orbital phase. The grey shaded region shows the 1σ uncertainty on the time of inferior conjunction derived from the RV modeling. *Bottom:* Same as top panel with the x-axis zoomed-in near the time of inferior conjunction. In this panel we only plot the measurements binned with bin widths of 0.002 units of orbital phase. Again, the shaded region represents the 1σ uncertainty on the time of inferior conjunction. The black transit model shows the predicted transit depth for a solid iron planet using the mass-radius relation of Weiss & Marcy (2014).

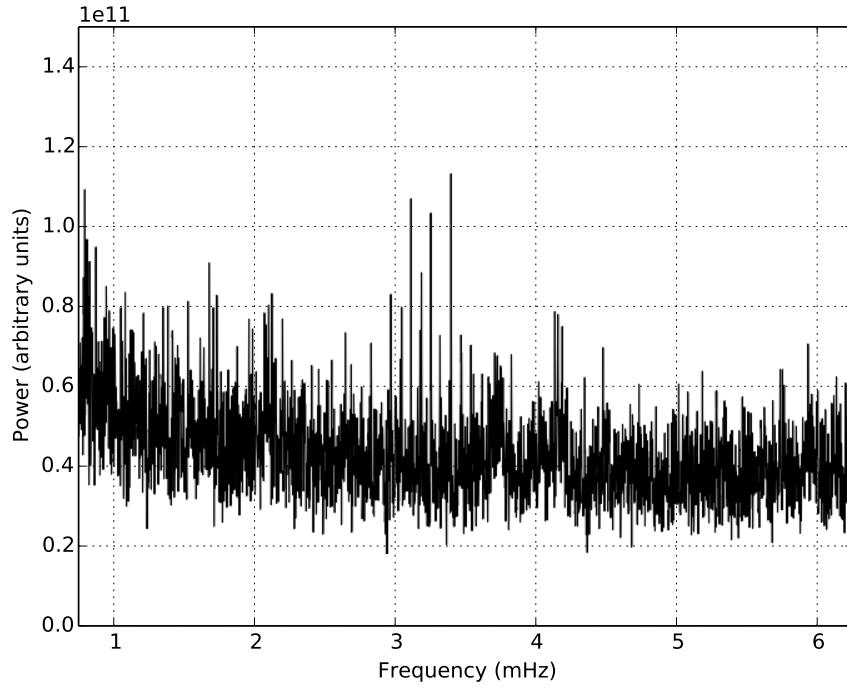


Figure 5.12 Smoothed one-dimensional power spectrum of HD 42618 from the CoRoT data. The comb of peaks in the power spectrum near a frequency of 3 mHz is the signature of solar-like asteroseismic oscillations. Our asteroseismic analysis of HD 42618 is described in Section 5.5.1.

autocorrelation with a Gaussian to measure a large oscillation frequency spacing $\Delta\nu$ of $141.6 \pm 0.8 \mu\text{Hz}$. Using scaling relations (Christensen-Dalsgaard & Frandsen, 1983; Kjeldsen & Bedding, 1995; Kallinger et al., 2010; Huber et al., 2011), solar parameters taken from Huber et al. (2011), and an effective temperature equivalent to the Sun's within errors (Morel et al., 2013b) we measure an asteroseismic radius of $0.95 \pm 0.05 R_\odot$, and an asteroseismic mass of $0.93 \pm 0.13 M_\odot$. This is in agreement with a previous CoRoT asteroseismic analysis (Barban et al., 2013) and our estimate of the stellar mass and radius of HD 42618 using our spectroscopic constraints and the Torres et al. (2010) relations. The precision on the asteroseismic mass is lower compared to our spectroscopic+isochrone mass but it is much less model-dependent. If we were to fit the spectroscopic parameters to isochrones derived using different input physics we may find that the error on the spectroscopic mass is much larger. We adopt the higher precision, spectroscopic mass for all calculations of planet minimum masses and orbital separations.

5.5.2 APT Photometry

Long-term photometric observations of HD 42618, HD 143761, and HD 164922 were collected with Tennessee State University’s T11 0.80 m, T4 0.75 m, and T12 0.80 m APTs at Fairborn Observatory. These three stars are among a collection of more than 300 being observed by the APTs to study magnetic cycles in solar-type stars (e.g., Lockwood et al. (2013) and references therein) and have APT observational histories between 15 and 23 years. At the beginning of the APF survey, the vast majority of the target stars were already being observed by the APTs. The remaining few have been added so that all 51 stars in the APF survey are also being observed nightly by the APTs.

The APTs are equipped with two-channel precision photometers that use a dichroic filter and two EMI 9124QB bi-alkali photomultiplier tubes to measure the Strömgren b and y pass bands simultaneously. The APTs are programmed to make differential brightness measurements of a program star with respect to three comparison stars. For the APF project, we use the two best comparison stars (C1 and C2) and compute the differential magnitudes $P - C1$, $P - C2$, and $C2 - C1$, correct them for atmospheric extinction, and transform them to the Strömgren system. To maximize the precision of the nightly observations, we combine the differential b and y observations into a single $(b+y)/2$ “passband” and also compute the differential magnitudes of the program star against the mean brightness of the two comparison stars. The resulting precision of the individual $P - (C1+C2)/2_{by}$ differential magnitudes ranges between ~ 0.0010 mag and ~ 0.0015 mag on good nights. Further details of our automatic telescopes, precision photometers, and observing and data reduction procedures can be found in Henry (1999), Eaton et al. (2003), and Henry et al. (2013).

APT Photometry of HD 42618

We collected 2241 relative flux measurements of HD 42618 over the past 15 years. We search for photometric variability on short timescales by first subtracting the mean magnitude from each observing season to remove seasonal offsets. This removes all astrophysical and systematic instrumental variability on timescales longer than one year. A L-S period search returns a very weak periodicity with a period of 16.5 days and an amplitude of 0.3 mmag. This may be the same signature of stellar rotation as detected in the CoRoT data but it is too close to the precision limit of the ground-based dataset to be certain.

Photometric variability on the timescale of the orbital period may indicate that the RV fluctuations are the result of rotational modulation of star spots (Queloz et al., 2001). We find no evidence of photometric variability at the orbital period of the planet to the limit of our photometric precision. A least-squares sine fit on the orbital period of HD 42618 gives a semi-amplitude of just 0.000037 mag, showing the complete absence of any surface activity that could affect the radial velocities. Figure 5.13 shows the full photometric dataset and Figure 5.14 shows the photometry phase folded to the orbital period of planet b. The lack of variability at the orbital period is consistent with the results of the CoRoT analysis and strengthens our claim that the RV fluctuations

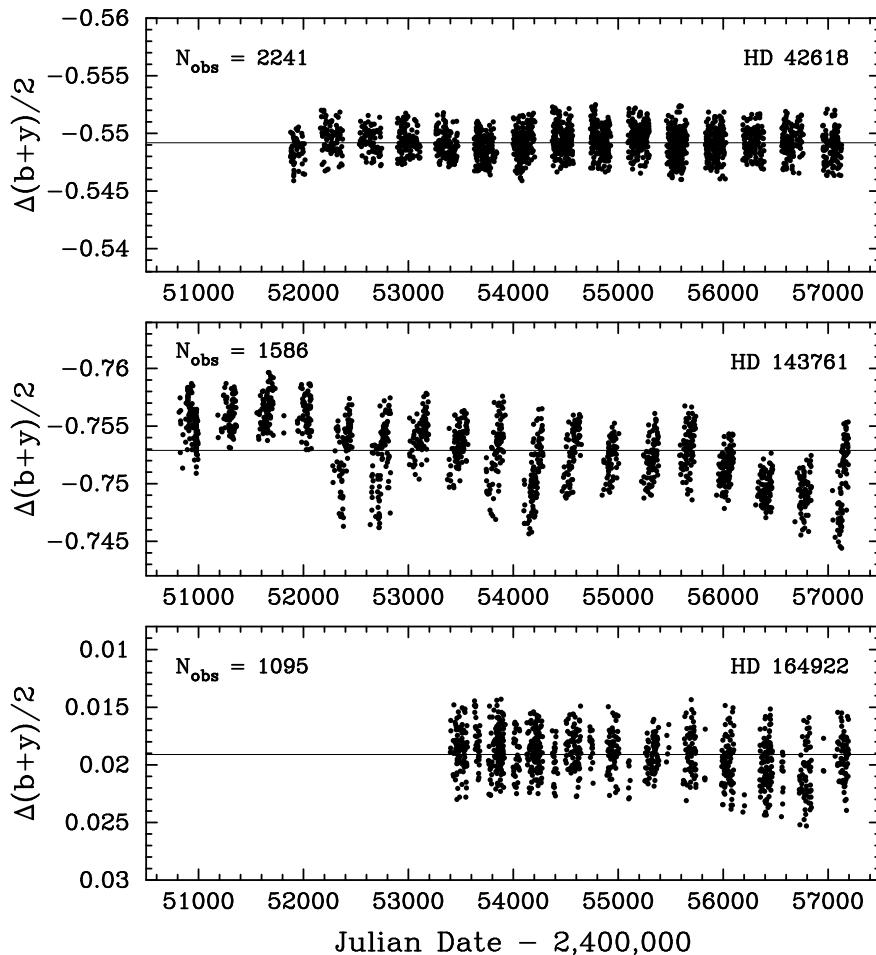


Figure 5.13 Long-term photometric observations of the planetary candidate host stars HD 42618 (top), HD 143761 (middle), and HD 164922 (bottom) acquired with TSU's T11 0.80 m, T4 0.75 m, and T12 0.80 m APTs at Fairborn Observatory in southern Arizona. All three stars are plotted with identical x and y scales. The horizontal line in each panel marks the mean of each data set. The APT photometry and analysis are described in Section 5.5.2.

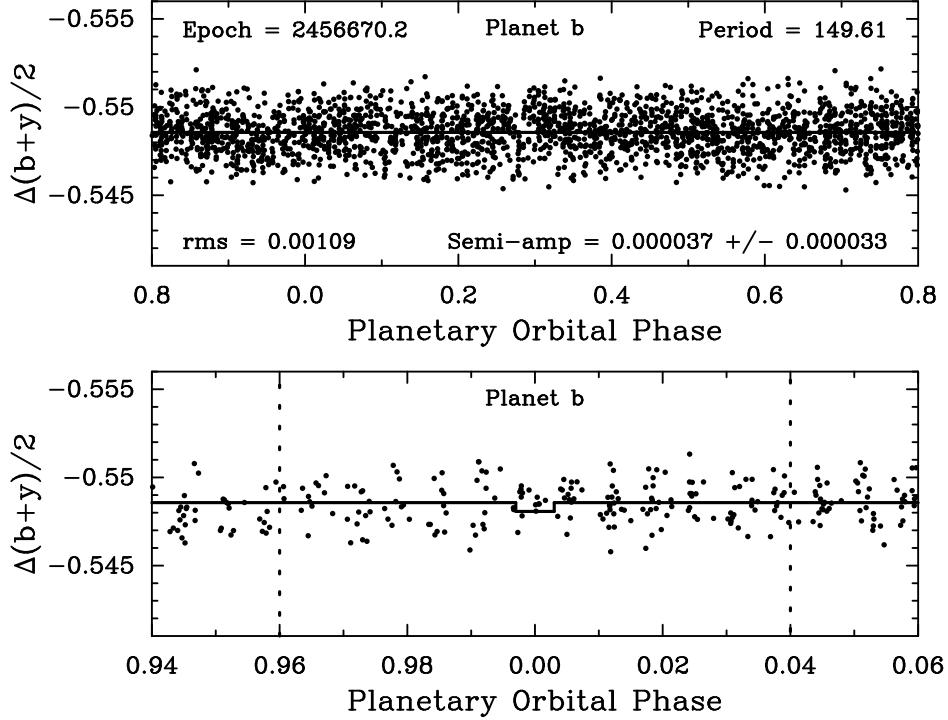


Figure 5.14 *Top*: Fifteen years of photometric observations of HD 42618 from the top panel of Fig. 5.13 plotted against the 149.6-day planetary orbital period and time of conjunction derived from the radial velocity observations. A least-squares sine fit on the radial velocity period gives a semi-amplitude of just 0.000037 ± 0.000033 mag, firmly establishing the lack of stellar activity on the radial velocity period and thus confirming the presence of stellar reflex motion caused by an orbiting planet. *Bottom*: Closeup of the observations near the time of planetary conjunction at phase 0.0. The solid line shows a toy model transit of a sphere of constant 1.0 g cm^{-3} density and radius determined by the relation of Weiss & Marcy (2014). The vertical lines mark the uncertainty in the predicted transit times. Our current photometric observations provide no evidence for transits.

are caused by a Neptune-mass planet orbiting HD 42618. We also find no evidence of the transit of HD 42618 b in the APT data.

Figure 5.15 shows the mean S_{HK} values from Keck and APF and the long-term photometric variability of HD 42618 by plotting the seasonal means of both the S_{HK} values and APT photometry. For old solar type stars we expect a positive correlation of chromospheric activity as measured by the S_{HK} values with the mean brightness since the number of bright faculae regions on the star increases during more active periods. However, in this case we see no correlation of mean brightness with S_{HK} . Young stars typically show a negative correlation of brightness with S_{HK} because their photometric variations are spot-dominated instead of faculae dominated. While somewhat unusual, this behavior is not unprecedented among similar stars (Hall et al., 2009).

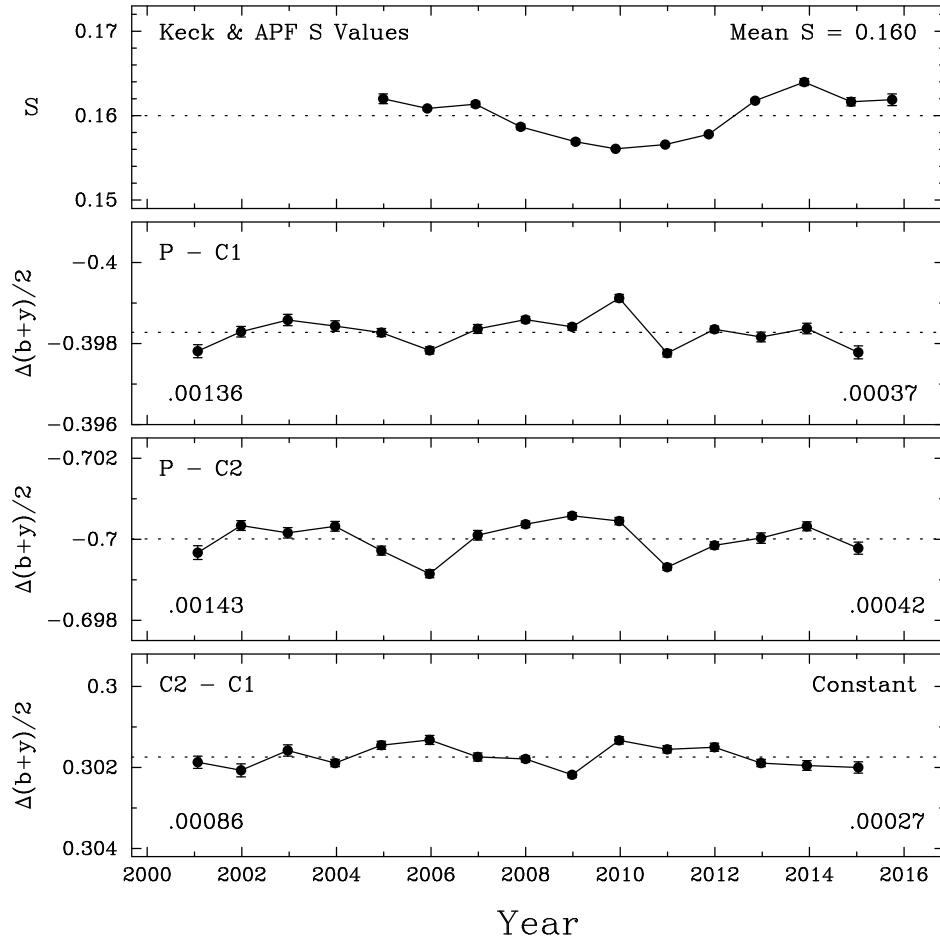


Figure 5.15 *Top:* Yearly means of the Mt. Wilson-calibrated S_{HK} values acquired along with the radial velocity measurements. *Bottom three panels:* Yearly means of HD 42618's $P - C1$, $P - C2$, and $C2 - C1$ differential magnitudes. The horizontal dotted lines designate the grand means of the observations while the numbers in the lower-left and lower-right give the total range and standard deviation of each data set, respectively. It is evident that we have resolved low-level brightness variability in HD 42618 compared to the two comparison stars, $C1$ and $C2$. Low-amplitude cycles of roughly 0.001 mag over 5 years are seen in both the $P - C1$ and $P - C2$ light curves. There appears to be little or no correlation of S_{HK} values with photometric brightness.

APT Photometry of HD 164922

We collected a total of 1095 photometric measurements for HD 164922 over the past 11 observing seasons from 2005 to 2015. As with HD 42618 we remove seasonal offsets from the photometry to search for short period variability and search for transits of HD 164922 b and c. We find no significant periodic variability with a period between 1 and 100 days and do not detect the rotation period of the star. We also find no evidence of transits for either planet b or c or periodic photometric variability at the orbital period of either planet (Figure 5.16).

We study the long-term photometric variability of HD 164922 by comparing the mean brightness of the star to the S_{HK} activity index (Figure 5.17). In contrast to the results for HD 42618, in this case we see a clear positive correlation of the brightness of HD 164922 with the S_{HK} index. It is interesting that we do not see a RV vs. S_{HK} correlation for HD 164922, but we do find that the RV is strongly correlated with S_{HK} for HD 42618 where the photometry is not. In other cases we have seen a correlation in both the photometry and RV data (e.g. Fulton et al., 2015b).

APT Photometry of HD 143761

We collected 1586 photometric measurements of HD 143761 over the past 18 observing seasons from 1997 to 2015 (Figure 5.13). Our reduction and analysis techniques are the same as for HD 42618 and HD 164922 discussed in the previous two sections. We find no evidence of the photometric signature of rotationally modulated star spots or photometric variability at the orbital periods of HD 143761 b or c. There is no evidence of transits of either planet b or c (Figure 5.18), however shallow transits of a rocky planet c can not be ruled out by this dataset.

The mean photometric brightness binned by observing season is well correlated with the S_{HK} values measured using Keck and APF as expected for an old solar type star. As with HD 164922, we do not see a correlation of S_{HK} with RV but there is a positive correlation of S_{HK} with mean brightness (Figure 5.19).

5.6 Discussion & Summary

We present the discovery of three approximately Neptune mass planets orbiting three bright, nearby stars. The planet orbiting HD 42618 has a minimum mass of $M \sin i = 15.4 \pm 2.4 M_{\oplus}$ and is the first discovered to orbit this star. There has been some discussion in the literature that stellar abundance patterns similar to the Sun might be evidence of the formation of terrestrial planets similar to those that exist in our solar system (e.g. Meléndez et al., 2009; González Hernández et al., 2010). While we can not rule out the existence of terrestrial planets in the HD 42618 system, the presence of a temperate Neptune mass planet orbiting at 0.554 AU with an orbital period of 149 days shows that, at the present time, this system is not a close analogue to our own solar system.

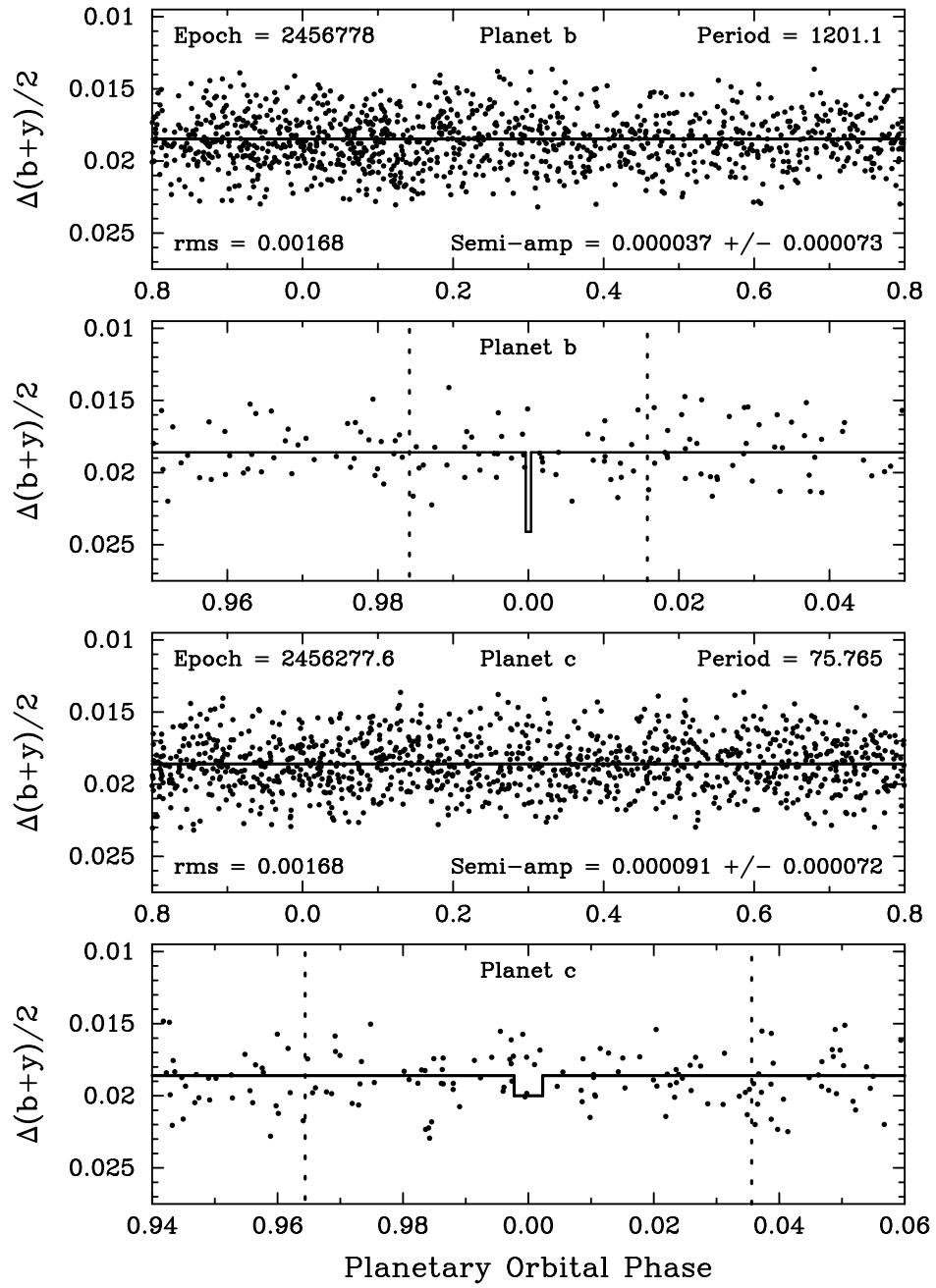


Figure 5.16 Same as Figure 5.14 but for HD 164922 b and c.

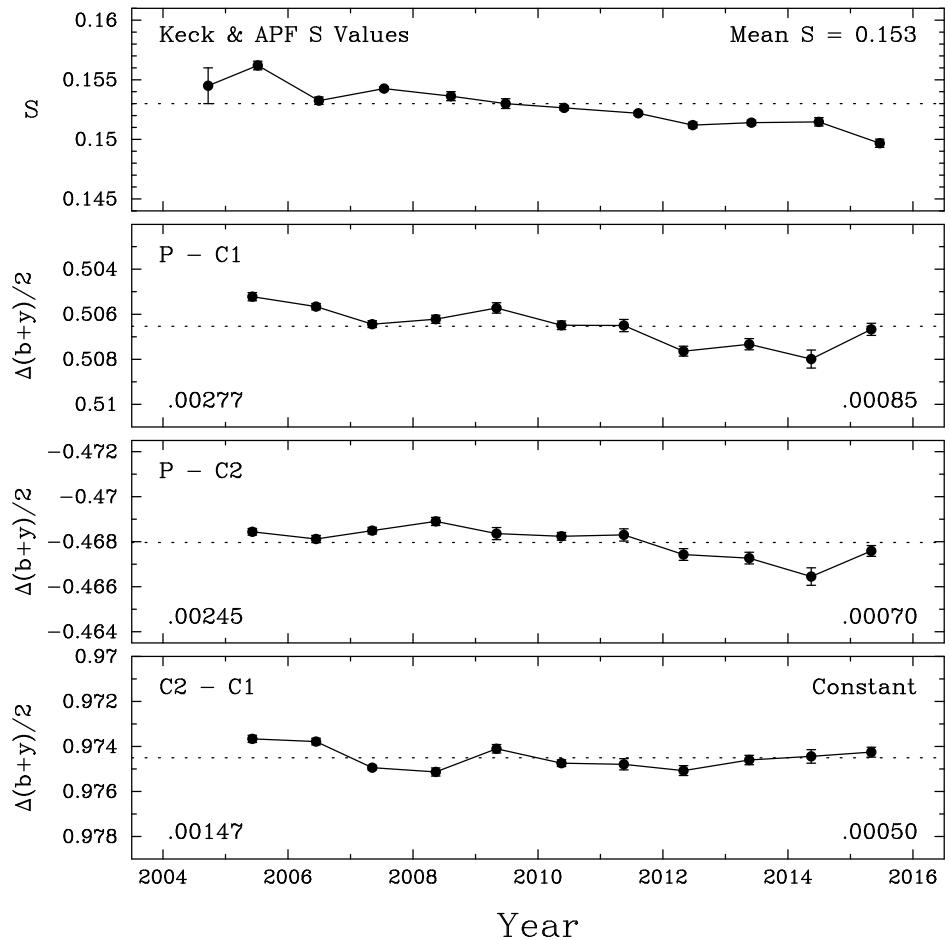


Figure 5.17 Same as Figure 5.15 but for HD 164922. In this case we see a positive correlation of the brightness of HD 164922 with the S_{HK} index.

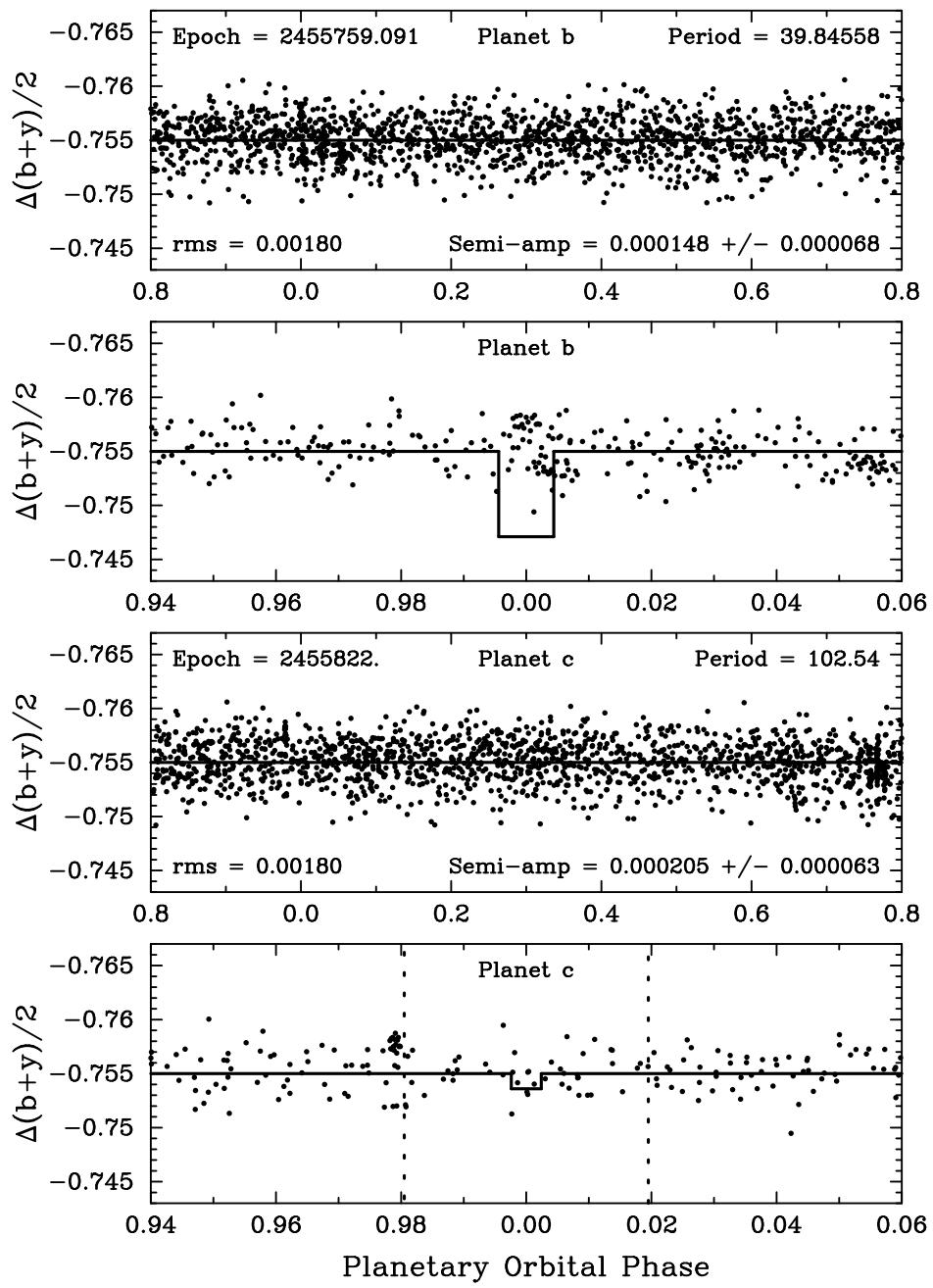


Figure 5.18 Same as Figure 5.14 but for HD 143761 b and c.

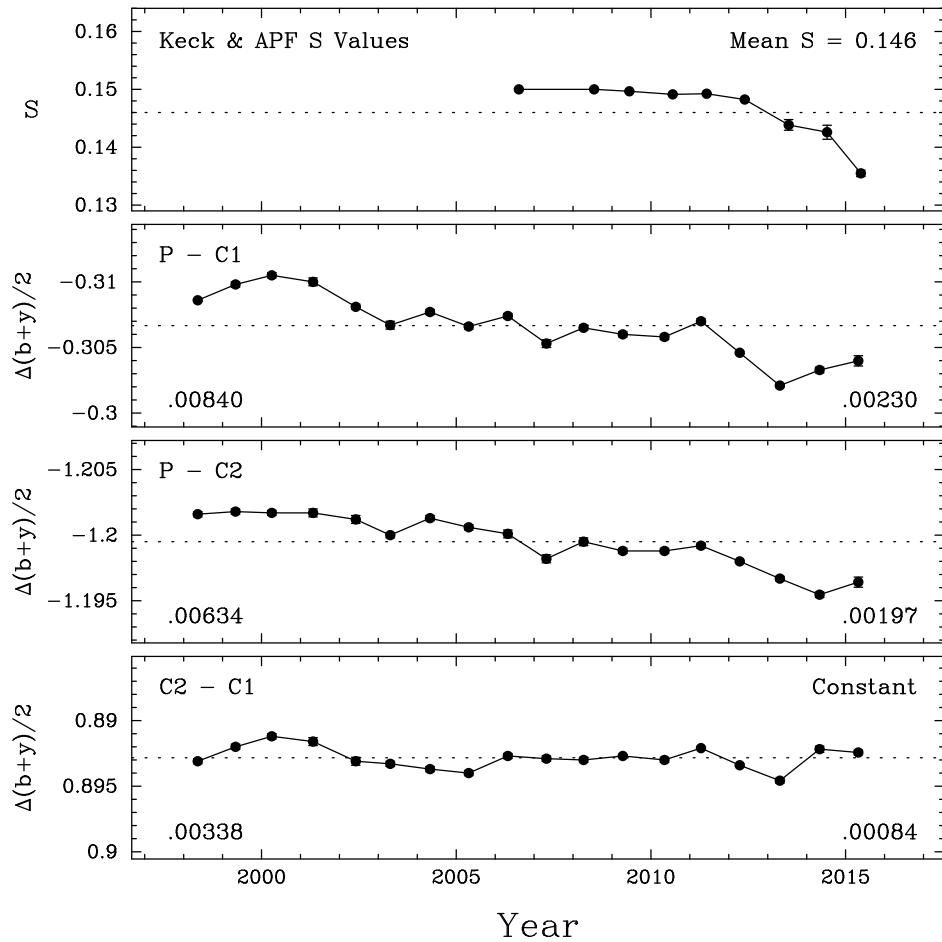


Figure 5.19 Same as Figure 5.15 but for HD 143761. In this case we again see a positive correlation of the brightness of HD 143761 with the S_{HK} index.

We cannot determine the initial planetary architecture of this system because migration may have played an important role to sculpt the current configuration.

We also detect the signature of the stellar magnetic activity cycle with a period of \sim 12 years. This activity cycle manifests as a 3.1 m s^{-1} amplitude signal in the RV time series. We identify the rotation period of the star to be \approx 17 days using public data from the CoRoT space telescope. Transits of HD 42618 b are expected to be extremely unlikely and we do not find any evidence for transits of this planet in the CoRoT data. This is a temperate planet receiving only 3.1 times the radiation that the Earth receives from the Sun. The planet's equilibrium temperature, assuming a bond albedo of 0.32 (Demory, 2014), is 337 K. We perform an asteroseismic study of HD 42618 to detect solar like oscillations and measure a precise stellar radius and mass.

HD 164922 c is the second planet in a system previously known to host one Jupiter mass planet orbiting at 2.1 AU. The new planet announced in this work is a sub-Neptune mass planet with $M \sin i = 12.9 \pm 1.6 M_{\oplus}$ orbiting at a distance of $a = 0.34$ AU and an orbital period of 75 days. This planet is also temperate with an equilibrium temperature of 401 K and receiving 6.3 times the flux received by the Earth from the Sun.

HD 143761 c is the second planet in a system previously known to host a warm Jupiter mass planet orbiting with a period of 39 days. The new planet is a super-Neptune with $M \sin i = 25 \pm 2 M_{\oplus}$ orbiting with a period of 102 days. This planet is the warmest of the three with a stellar irradiance 9.6 times that of the Earth-Sun system and an equilibrium temperature of 445 K. We find that the previous low inclination orbit for HD 143761 detected in Hipparcos astrometry can not be stable with the presence of HD 143761 c.

These three planets are some of the nearest long period Neptune mass planets yet discovered. They demonstrate the capabilities of the combined Keck+APF-50 survey and are the beginning of a complete census of small planets in the local neighborhood.

CHAPTER 6

A SEARCH FOR PLANETARY ECLIPSES OF WHITE DWARFS IN THE PAN-STARRS1 MEDIUM-DEEP FIELDS

This chapter is a reproduction of Fulton et al. (2014) included with permission from AAS journals.

6.1 Introduction

Searches for planets outside our solar system have focused primarily on hydrogen-burning main-sequence stars similar to our Sun (e.g. Bakos et al., 2004; Howard et al., 2010; Borucki et al., 2010). As we discovered that planets are nearly ubiquitous in our Solar neighborhood (Howard et al., 2010b) and in the *Kepler* field (Petigura et al., 2013a) searches around M-dwarfs gained popularity (e.g. Nutzman & Charbonneau, 2008). Studies of M-dwarfs enjoy a boost in sensitivity to small planets because transits block a larger fraction of the stellar disk and induce a larger amplitude reflex motion of the star around the barycenter due to their low mass. Some studies have also searched for and explored the planet occurrence rates as a function of stellar mass from M-dwarfs to intermediate-mass subgiants (Johnson et al., 2007). Microlensing campaigns survey stars of many types and are sensitive to planets around all massive hosts regardless of their stage in stellar evolution (Gaudi, 2012) but followup characterization of these planets is impossible. However, there have been few dedicated searches for planets around white dwarfs (WDs).

Many studies including Mullally (2007), Farihi et al. (2008), and Kilic et al. (2009) searched for infrared-excess indicative of planetary companions to WDs. They detected several brown dwarf companions (Zuckerman & Becklin, 1992; Farihi et al., 2005; Steele et al., 2009) but no planetary-mass objects. Mullally (2007) also searched for companions using pulsations of WDs to look for periodic deviations in the pulse arrival times caused by an orbiting companion. They find evidence of a $2.4 M_J$ companion in a 4.6 year orbit. Hogan et al. (2009), and Debes et al. (2005) conducted high contrast imaging surveys of nearby WDs to search for low-mass companions at large separations. Burleigh et al. (2006) found a brown dwarf in the near-IR spectrum of WD 0137-349 with an orbital period of only 2 hours. This object may have survived the common-envelope phase or migrated from larger orbital distances after the formation of the WD. Faedi et al. (2011) conduct a transit search for a sample of 174 WDs using SuperWASP data (Pollacco et al., 2006) and find no eclipsing companions but can put only weak constraints on the planet occurrence rates due to their small sample size (<10% for Jupiter-size planets). Drake et al. (2010) search for eclipses of $\sim 12,000$ color-selected WDs using Catalina Sky Survey photometry and Sloan Digital Sky Survey spectroscopy. They find 20 eclipsing systems and three of them have radii consistent with substellar objects and no detectable flux in the spectra.

WDs have radii only $\sim 1\%$ of the Sun, or about the same size as the Earth. This implies that

an Earth-sized object transiting the WD with an impact parameter of 1.0 would cause a complete occultation. Although these occultations are short-duration, they can be easily detected from small ground-based telescopes with short exposure times and relatively low photometric precision (Drake et al., 2010). In addition, the most common WDs are old and cool with surface temperatures of \sim 5000 K. Their small radii and low surface temperatures imply that their luminosity is low, with typical values of \sim 10 $^{-4}$ L $_{\odot}$, and the habitable zone is close-in ($a \sim$ 0.01 AU, Agol, 2011) giving rise to significant transit probabilities. This makes Earth-size planets orbiting in the habitable zones of old, cool WDs relatively easy to detect via the transit method.

Most main-sequence stars, including our Sun, will eventually end their lives slowly cooling as WDs. Since approximately 50% of main sequence stars host at least one planet (Mayor et al., 2011) it is interesting to consider their fate as the star evolves into a WD. It is unlikely that any planets inside \sim 1 AU would survive engulfment by their host stars as they expand onto the red giant branch but it is unclear what becomes of the planetary debris. Since WDs quietly cool for the age of the universe, it is conceivable that new planets could form out of the debris of a previous generation of planets. Migration of planets from outside of 1 AU is also plausible, but little theoretical work has been done on the formation or migration of planets hosted by WDs. Several studies have identified pollution by heavy elements on the surfaces of WDs (Zuckerman et al., 2010) and IR excess indicative of a debris disk (Debes et al., 2011). Extensive work has been done to identify the chemical composition of this pollution. Silicates and glasses were detected in the atmosphere of six WDs by Jura et al. (2009) and interpreted as signs of accretion of asteroid-like bodies onto the WD. A detailed study by Xu et al. (2014) using data from the Keck and Hubble Space Telescopes showed strong evidence that the composition of metals in the atmospheres of WDs G29-38 and GD 133 closely mirror the composition of the bulk Earth. Furthering the idea that close-in terrestrial planets orbit and eventually accrete onto WDs.

We present a systematic search for eclipses of WDs by planetary-size objects in the Pan-STARRS1 medium-deep fields (Tonry et al., 2012). We use a combination of astrometric and photometric selection techniques to identify 3179 WDs with a range of ages and temperatures. Each WD was observed on 1000-3000 epochs during the past 5 years for a total of 4.3 million measurements. Although we do not detect any substellar companions, this large number of observations allows us to place tight constraints on the occurrence rates of planets orbiting WDs.

6.2 Methods

6.2.1 WD sample

We analyze a total of 3179 WD candidates spread across the 10 medium-deep fields spanning 70 square degrees on the sky. Each field is observed on 1000-3000 epochs with four to eight consecutive 240 s exposures per night. Our sample of WDs is segregated into two categories. We identify 661

targets using their proper motions as described in Tonry et al. (2012) (astrometric sample hereafter). These objects have a high probability of being bona fide WDs and a very low contamination rate.

The remaining 2518 WDs were selected based on their photometric colors (color-selected sample hereafter). We use the following criteria to select the locus of hot, blue stars from the $(g_{P1} - r_{P1})$ vs. $(r_{P1} - i_{P1})$ color plane shown in Figure 6.2; $(g_{P1} - r_{P1}) < 0.18 + 1.4(r_{P1} - i_{P1})$, $(g_{P1} - r_{P1}) > 0.06 + 1.4(r_{P1} - i_{P1})$, $(g_{P1} - r_{P1}) < 0.25 - 1.25(r_{P1} - i_{P1})$, and $i_{P1} < 22$. This sample is restricted to hot WDs due to the requirement of blue colors and is likely contaminated by other hot stars.

To quantify the contamination rate of the color-selected sample we created a Besancon galactic simulation of the medium-deep fields (Robin et al., 2003). When we make the same color-cuts we find that 42% of the stars are bona fide WDs according to the model. The stars that are within this locus but not WDs are mostly distant A and B-type subdwarfs in the halo of the galaxy. Closer F-type subdwarfs would also fall into the locus, but are mostly far too bright to be included in our sample. We also find that the contamination rate is highly dependent on apparent magnitude with the fainter stars being much more likely to be WDs. We assume a 58% contamination rate for our color-selected sample for all further analysis. This reduces our total number of WDs to 1718.

6.2.2 Control sample

Our control sample consists of stars with similar magnitudes and colors to the astrometrically-selected WDs but with undetectable proper motions. These should be relatively hot stars with radii much larger than WDs around which we would not expect to see the very short-duration eclipses indicative of a planet occulting a WD. We can compare the number of potential eclipses found in the WD sample to the number that we find in the control sample to better understand the frequency of eclipse-like events caused by non-astrophysical effects.

We select the control sample by binning the astrometric sample of WDs in 2-dimensional color bins of r_{P1} vs. $(r_{P1} - i_{P1})$. For each bin that contains at least one WD we select two times the number of WDs in that bin from a sample of all stellar detections derived from deep stacks of the medium-deep fields excluding stars that are already part of the WD samples. Figure 6.1 shows our control sample and astrometric WD sample in the r_{P1} vs. $(r_{P1} - i_{P1})$ color plane. If fewer than three field stars are available in a particular bin we select all available stars. This produces a total of 1296 stars for the control sample which is later trimmed down to 1288 by removing RR-Lyrae, Delta-Scuti and other variable stars (see § 6.2.4).

6.2.3 Light curves

Light curves are extracted for each WD and control sample star by directly analyzing the first-level Pan-STARRS1 photometry product (SMF files). These SMF files consist of the raw photometry extracted from the calibrated images before a zero-point or precise world coordinate system (WCS) is established. Each camera exposure corresponds to a single SMF file. For each SMF file we first

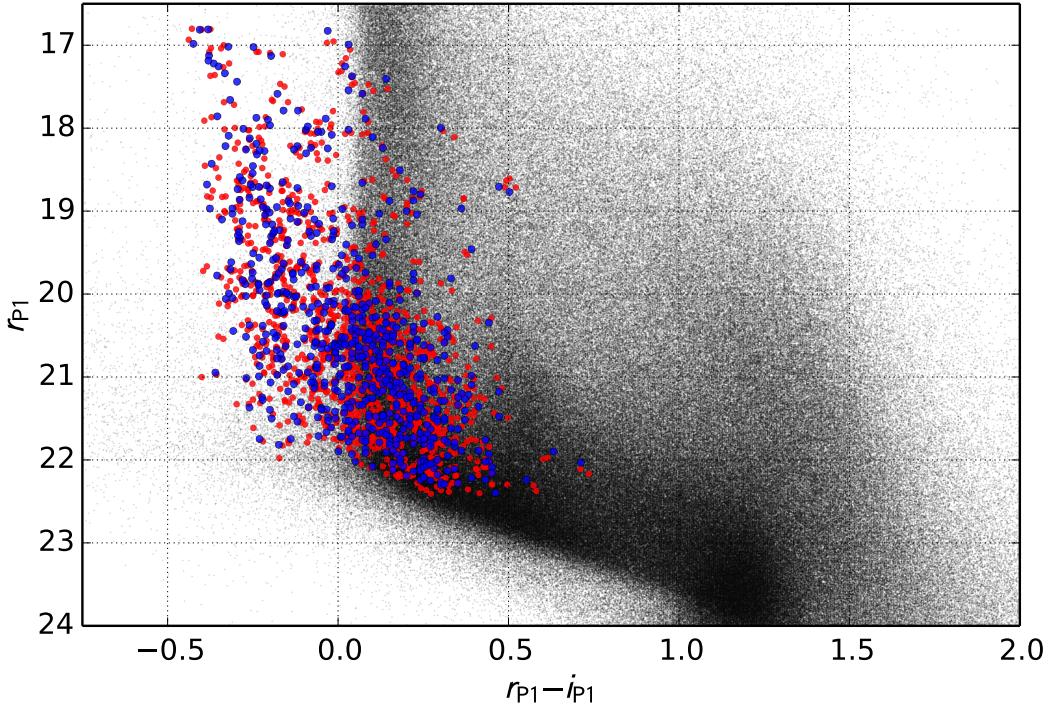


Figure 6.1 Astrometrically-selected WDs (blue) and control sample stars (red). The small black points are all detections from the deep stacks that were not selected for either the control or WD samples.

find the WCS solution in order to associate pixel locations with sky positions. We then associate the per-image detections with detections in deep stacks for each field and extract the PSF-fitted photometry to obtain raw instrumental magnitudes. We fit for the photometric zero-point using the technique described in (Schlafly et al., 2012). The instrumental magnitudes for all detections within 5 arcminutes of the target are also extracted and recorded along with the target instrumental magnitudes. All epochs for which a target could not be matched to a detection in the SMF file are carefully recorded and the neighboring star photometry is still extracted if available. This ensures that we are sensitive to large decreases in flux that may cause the target to fall below the detection threshold in a particular image and in some cases we can use the photometric statistics of the neighboring stars to explain the non-detection. We also record the pixel locations relative to the entire CCD array and particular chip for each epoch.

6.2.4 Eclipse detection

Since eclipses are rare and extremely short duration traditional periodic search algorithms such as the box-least-squares periodogram (BLS, Kovács et al., 2002) fail to recover such signals. BLS

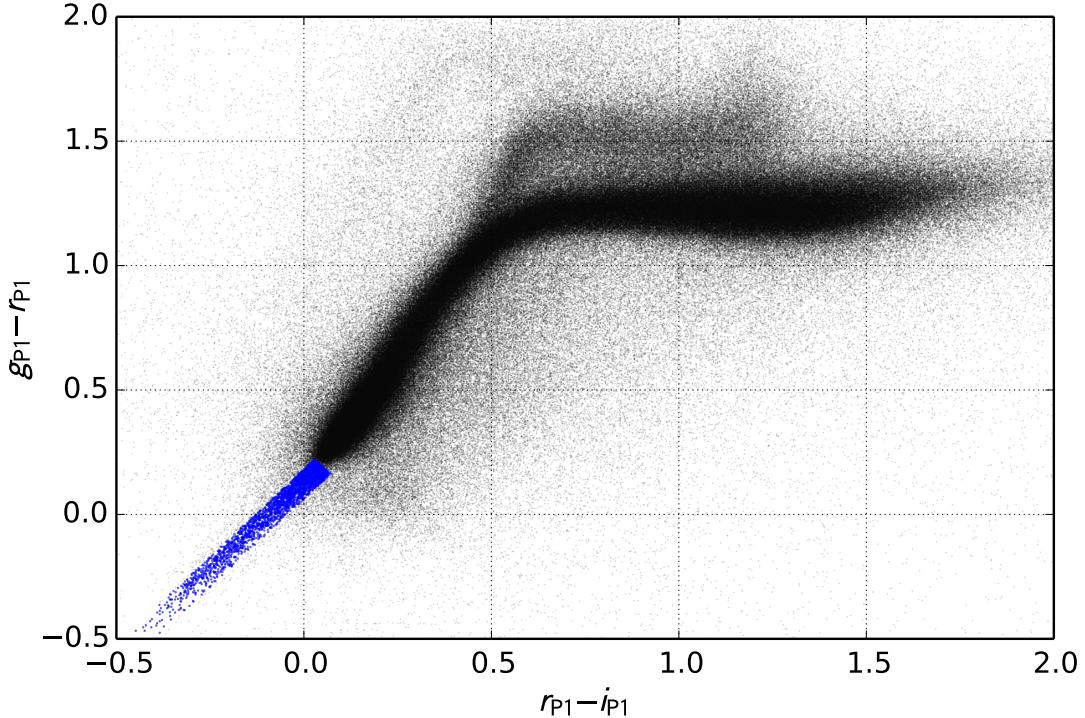


Figure 6.2 The color-selected WDs (blue points) are identified by the narrow tail of extremely blue stars in the $(g_{P1} - r_{P1})$ vs. $(r_{P1} - i_{P1})$ color plane. The small black points are all detections from the deep stacks not selected for either the control or WD samples.

excels at detecting signals in the regime of many transits with low single-event S/N but planetary eclipses of our target stars would produce very infrequent, but very deep, high S/N eclipses. Instead we employ an extremely simple eclipse detection technique. We look for low outliers in the light curves (dropouts) that are caused either by a complete non-detection or show a deficit of flux relative to the median flux level (ΔF) that is greater than five times the measurement uncertainty ($\Delta F/\sigma_{lc} \geq 5$). Figure 6.3 shows the distribution of ΔF and $\Delta F/\sigma_{lc}$ for all light curves.

The raw light curves are heavily contaminated with non-detections and large flux drops that could be indicative of an eclipse event or a variety of non-astrophysical scenarios. For every dropout we first check that the star did not fall off of, or too near the edge of a chip. We initially noticed that the dropout events were concentrated around the edges of the chips. This is likely caused by the PSF fit failing due to a strong gradient in the background region near the edges of the chips. This effect is worse at the corners of the chips near the readout electronics. For these reasons we remove all light curve measurements that fall within 10 pixels ($2''.5$) of an edge or within 100 pixels ($25''$) of a corner. We consider this filter unbiased with respect to eclipses because there is no reason to expect that real eclipses would preferentially occur when the stars fall near the edge of a chip.

Table 6.1. Detection Statistics

Filter	WD		Control	
	N_detections	N_dropouts	N_detections	N_dropouts
No filters	5,650,109	6,963,603	1,814,296	3,106,873
CCD location-based filters	4,757,706	1,771,860	1,523,212	622,154
Neighboring star filter	4,509,855	1,651,266	1,439,106	577,904
Re-calculate measurement errors	4,349,232	15,120	1,363,979	3,983
Remove masked CCD regions	4,343,011	9,000	1,362,535	2,570

Measurements with reported positions that fall between chip gaps or off the array are also excluded at this stage. All non-detections are removed with these chip location-based filters.

If the photometry of the neighboring stars also show a large decrease in flux at the same time of the target dropout, clouds or poor seeing is likely to blame. We exclude all measurements for which the median magnitude of the neighboring stars drops by more than 0.5 magnitudes or the standard deviation of the neighbor magnitudes is greater than one. We also de-correlate the target relative flux measurements against the median ΔF of the neighboring stars to reduce the effect of spatially-dependent extinction.

Now that we have removed most of the egregious outliers from the light curve we re-define the measurement errors. We sum in quadrature the reported measurement uncertainties with the median absolute deviation (MAD) of the full light curve in each filter. This processes always inflates the errors relative to the original measurement uncertainties and effectively removes many remaining candidate dropouts by decreasing the value of $\Delta F/\sigma_{lc}$.

At this stage we use the VARTOOLS package to create BLS and analysis-of-variance (AoV) periodograms (Hartman et al., 2008; Kovács et al., 2002; Schwarzenberg-Czerny, 1989; Devor, 2005) for all WD and control sample stars. We visually inspect these periodograms and the light curves phase-folded to the ephemeris that corresponds to the highest peak in each periodogram. Obvious periodic variable stars are removed from further analysis. Thirty-three RR-Lyrae and Delta-Scuti stars, one dwarf nova (IY Uma) and three variables of unknown type are identified and removed at this stage.

For the remaining dropouts we check their CCD locations against the regions of the array that are consistently masked by the Pan-STARRS1 Image Processing Pipeline (IPP). After applying all of the photometry-based tests we are left with 11570 potential dropout events and of a total of 4.3 million detections. 2570 of the dropout candidates are from the control sample and the remaining 9000 are from the merged WD samples. This photometric filtering process for a single representative case is illustrated in Figure 6.4, and the total number of detections and non-detections removed at each stage in the filtering process are listed in Table 6.1.

We download the corresponding postage stamp images for any dropouts that make it through all of these light curve-based tests for additional screening. In addition to the postage stamp

corresponding to the dropout we also download a deep stack around the target and the image that corresponds to the light curve measurement that is closest to the median value for that filter. We apply a few more automated filters before visually inspecting the remaining candidates. The images are automatically inspected for masking or CCD defects around the target that produce not-a-number (nan) values, very poor seeing, or clouds as indicated by a low zeropoint magnitude. We also perform aperture photometry on the three images and correct to an absolute apparent magnitude using the zeropoint magnitude provided in the image headers. Our photometry acts as check that the magnitude value reported by the IPP is in rough agreement with simple aperture photometry.

As a final step we use the HOTPANTS implementation of the ISIS image subtraction software (Alard, 2000) to produce a difference image using the deep stack as a template. We convolve the template to match the PSF and zero point of the dropout candidate image and subtract the convolved template from the candidate postage stamp. This difference image was used to aid the visual inspection of the 133 dropout events that could not be explained by any of the photometry or image-based filters. Figure 6.5 shows an example dropout candidate image and the image-differencing processes used for visual inspection. We find no eclipse with a duration compatible with an eclipse by a substellar object in any WD or control sample light curve.

6.3 Analysis

6.3.1 Theoretical eclipse probabilities

In order to asses the likelihood that an occultation would have occurred during our observing window, we calculate the probability of eclipse as a function of eclipse depth and then apply the noise properties and eclipse detection techniques that we used in our search. This tells us the number of eclipses we should have been able to detect as a function planet radius, orbital semi-major axis and the occurrence rate of planets around WDs (η).

The flux when a dark sphere eclipses a uniformly illuminated sphere is given by:

$$1 - F(p, b) = \begin{cases} 0 & 1 + p < b \\ \frac{1}{\pi} \left(p^2 \kappa_0 + \kappa_1 - \sqrt{\frac{4b^2 - (1+b^2-p^2)^2}{4}} \right) & |1 - p| < b \leq 1 + p \\ p^2 & b \leq 1 - p \\ 1 & b \leq p - 1, \end{cases} \quad (6.1)$$

$$b(t) \approx a \sqrt{\sin^2(\Omega + \omega t + \alpha_0) + \sin^2 \theta \cos^2(\Omega + \omega t + \alpha_0)}, \quad (6.2)$$

where $\kappa_1 = \cos^{-1}[(1 - p^2 + b^2)/2b]$, $\kappa_0 = \cos^{-1}[(p^2 + b^2 - 1)/2pb]$, and $p \equiv R_p/R_{WD}$ is the planet to white dwarf radius ratio (Mandel & Agol, 2002).

Equation 2 for $b(t)$ gives the sky-projected center to center distance between the star and planet as a function of time (t). Ω is the longitude of the ascending node of the planet's orbit, a is the semi-major axis of the orbit, ω is the angular frequency of the orbit, θ is the inclination of the planet's orbit, and α_0 is the phase of inferior conjunction. Minimizing Equation 2 leads to the smallest sky projected separation over the orbit, $b_0 = R_{WD} \cos \theta$.

In order to determine the likelihood that a particular ΔF could be caused by an eclipse of the WD we calculate the probability of eclipses as a function of eclipse depth. First, we make some assumptions for physical parameters that are mostly constant within the parameter region of interest. We assume that $M_{WD} = 0.6 M_\odot$, $R_{WD} = 0.01 R_\odot$, all theoretical companions are on circular orbits, no limb darkening, and 240 s as the integration time for every exposure. The probability of measuring an eclipse depth $\langle \Delta F(p, b) \rangle$ at time t averaged over an exposure time of Δt is

$$\langle \Delta F(p, b) \rangle = \frac{1}{\Delta t} \int_{t_0}^{t_0 + \Delta t} F(p, b) dt \quad (6.3)$$

Eclipses will only occur if $|b_0| < 1 + p$, therefore the probability that a randomly-oriented, circular orbit will eclipse is

$$P_{\text{eclipse}} = \frac{R_p + R_{WD}}{a}. \quad (6.4)$$

Although systems with $|b_0| < 1 + p$ will eclipse at some time during the orbit, the fraction of orbital phase covered during eclipse is small. The probability that any part of an eclipse will overlap with the integration time of our survey is

$$P_{\text{phase}} = \frac{T_{\text{dur}} + E}{P}, \quad (6.5)$$

where P is the orbital period and T_{dur} is the eclipse duration and E is the integration time.

For eclipses with durations shorter or equal to the exposure time the likelihood of any given measurement being in eclipse is then the sum of the probabilities for all possible orbital configurations that would produce an observed eclipse of depth m . For example, a measurement with $m = 0.1$ could be caused by a very small planet transiting slowly across the face of the star with a transit duration approximately equal to the exposure time. Alternatively, an eclipse of a much larger planet causing a complete occultation of the WD on a very short-period orbit would streak across the face of the star with a transit duration much shorter than the exposure time. The mean flux during the exposure may look identical in these two cases. Both of these cases and all other situations that could cause an observed eclipse depth m must give the appropriate weight in the final likelihood calculation. Figure 6.6 shows the eclipse depth probability distributions for a few hypothetical scenarios.

By the definition of our eclipse detection algorithm each exposure is sensitive to eclipses of depth $m \geq 5\Delta F/\sigma_{lc}$. By integrating over all scenarios that would cause an observed eclipse depth greater than or equal to $5\Delta F/\sigma_{lc}$ for every measurement we derive the probability that we could have detected an eclipse during each exposure if $\eta = 1$. The inverse of the summed probabilities over all exposures for all light curves gives a total number of expected eclipses for the survey as a Poisson expectation value for the rate of eclipses (Figure 6.7). We then compare this Poisson distribution for the expected number of eclipses with the lack of detected eclipses for many values of a, p , and η .

6.3.2 Occurrence constraints

If we treat the number of expected eclipses as a Poisson expectation value (λ), the probability that we should detect k eclipses is

$$P(k, a, p) = \frac{\lambda(\eta, a, p)^k \exp(-\lambda(\eta, a, p))}{k!}. \quad (6.6)$$

Since we have zero detected eclipses this can be simplified to $P(0, a, p) = \exp(-\lambda(\eta, a, p))$. By setting $P(0, a, p)$ equal to a confidence interval C and decomposing $\lambda(a, p)$ into the expectation value of eclipses if the planet occurrence rate is equal to 1 ($\lambda_1(a, p)$) multiplied by the actual planet occurrence rate (η) we derive the maximum planet occurrence rate that is compatible with the observations at a confidence level of C

$$\eta \leq \frac{\ln(1 - C)}{\lambda_1(a, p)}, \quad (6.7)$$

assuming the planet occurrence rate is constant as a function of a and p .

6.4 Discussion

Although we find no convincing detections of eclipses with durations consistent with substellar objects we are still able to put strong constraints on the WD-hosted planet occurrence rate. Figures 6.8 and 6.9 show the maximum occurrence rate that is consistent with our observations at 95% and 68% confidence levels assuming $R_{WD} = 0.01 R_\odot$ and $M_{WD} = 0.6 M_\odot$. This should be a relatively good approximation since the masses and radii of most WDs fall close to these values. For each reported occurrence rate (η) we first state the value corresponding to the maximum allowable occurrence rate averaged over the specified region of interest for the 95% confidence limit and then the 68% confidence limit immediately following in parenthesis. For example, our results suggest that less than 0.4% (0.2%) of WDs host planets with radii greater than ~ 2 Earth radii and semi-major axis between 0.002 and 0.01 AU. 0.4% is the maximum occurrence rate allowed by our data at 95% confidence and 0.2% is the same for a confidence level of 68%.

It is an interesting exercise to break up the two-dimensional occurrence limits into regions that correspond to classes of planets that we are more familiar with orbiting main-sequence stars. Other studies have shown similarities between the architectures of exoplanetary systems around low-mass M-dwarfs with the moons of Jupiter (Muirhead et al., 2012) and scaled-down versions of our solar system or exoplanetary systems around more massive stars. If we scale down the orbital distances of the known exoplanet population we can look at the occurrence limits in a few interesting regimes; hot Jupiters, hot super-Earths, and habitable-zone super-Earths.

The Roche limit for a fluid body with mean density ρ_p orbiting a WD with density ρ_{WD} and radius R_{WD} can be approximated as

$$L_R \approx 2.44R_{\text{WD}} \left(\frac{\rho_{\text{WD}}}{\rho_p} \right)^{1/3}. \quad (6.8)$$

For our assumed WD properties the Roche limit for a Jupiter-like planet is $L_R \approx 0.01$ AU. It is not surprising that we do not detect any Jupiter-sized objects inside 0.01 AU. However, we can equate a population of Jupiter-sized planets orbiting between 0.01 and 0.04 AU to the hot Jupiters observed orbiting very close to solar-type stars. In this regime an eclipse duration is slightly longer than the duration of a single exposure. Therefore our expectation value for eclipses is slightly overestimated, however we do not expect this to be the dominant source of error in the occurrence rate limits. The mean maximum occurrence rate for WD-hosted hot Jupiters ($R = 10 - 20R_{\oplus}$) is 0.5% (0.2%). Indicating that hot Jupiters around WDs are very rare or non-existent. This is in good agreement with the frequency of hot Jupiters around solar-type stars measured to be between 0.3% and 1.5% (Marcy et al., 2005; Gould et al., 2006; Cumming et al., 2008; Howard et al., 2011b; Mayor et al., 2011; Wright et al., 2012).

A rigid body can orbit slightly closer to the WD without being tidely disrupted. Planets with radii larger than $\sim 1.5 R_{\oplus}$ generally have densities lower than that of the Earth and likely have an extended gas-dominated atmosphere (Weiss & Marcy, 2014). However, some super-Earths with slightly larger radii have high densities consistent with a rocky composition, e.g. CoRoT-7b (Léger et al., 2009), Kepler-20b (Gautier et al., 2012), and Kepler-19b (Ballard et al., 2011). This class of planets may be the remaining cores of evaporated gas giant planets (Hébrard et al., 2004). Our results suggest that less than 1.5% (0.6%) of WDs host planets with radii between 2.0 and 5.0 R_{\oplus} orbiting with semi-major axis between 0.005 and 0.01 AU. Howard et al. (2012b) measure an occurrence rate of 13% for 2-4 R_{\oplus} planets with orbital periods shorter than 50 days. However, the occurrence rate drops with shorter orbital periods to 2.5% for periods shorter than 10 days. Our lack of detections indicate that hot super-Earths are almost certainly less common around WDs than they are around solar-type stars.

Perhaps the most interesting planets to consider are those that have an equilibrium temperature such that they could sustain liquid water on their surfaces. Since WDs cool and decrease in

luminosity as they age, the habitable zone (HZ) boundaries also change as a function of time. Agol (2011) define the WD continuous habitable zone (CHZ) as the range of semi-major axis that would be within the HZ for a minimum of 3 Gyr and also outside of the tidal destruction radius for an Earth-density planet. For a $0.6 M_{\odot}$ WD this corresponds semi-major axis between 0.005 and 0.02 AU. Our data show that planets in the CHZ with radii between $2-5 R_{\oplus}$ could be present around no more than 3.4% (1.3%) of WDs. This is significantly less than the predicted frequency of Earth-size planets in the habitable zone of solar-type stars ($\sim 22\%$, Petigura et al., 2013a).

A large population of short-period planets orbiting solar-type and M-dwarf stars has been observed. We might expect WDs to host similar planets if they can reform from a post-giant phase debris disk or migrate from larger orbital distances once the star becomes a WD. However, our observations are quite sensitive to planets larger than the Earth orbiting close to the WD, and the lack of any eclipses suggests that these processes are highly inefficient if they occur at all. There are very few planets in short-period orbits around WDs.

6.4.1 Future survey design

Since eclipse times are generally shorter than the 4 min exposure times for the medium-deep survey we explore the idea of designing a similar survey with shorter exposure times and decreased sensitivity to shallow eclipses. This would cause less dilution of the eclipse signals over the duration of the exposure. We re-calculate the expected eclipse rates for exposure times of 30, 60, and 120 seconds scaling the measured noise properties from our 240 s data. We use the mean eclipse rate for planets with radii between $1-5 R_{\oplus}$ orbiting between 0.005 and 0.02 AU as a metric for comparison. Figure 6.10 illustrates the result. We find that decreasing the exposure time gives a modest boost in sensitivity to these planets for a given total survey exposure time. The most dramatic increase in sensitivity when going to short exposure times is for the very short period planets orbiting interior to 0.003 AU. However, planets are not able to withstand the tidal forces this close to the WD so we would not expect planets to exist in this regime. The expected eclipse rate in our region of interest is dominated by signal-to-noise of the individual detections. Although the eclipses are diluted by long exposure times, this is balanced by the increased gain in sensitivity to these shallow, diluted eclipses due to the greater signal to noise obtained in longer exposures. This suggests that the best way to detect these Earth to Neptune size planets in the WD CHZ may be to increase the etendue of the survey to detect more WDs on a greater number of epochs by covering a large area of the sky at high cadence. The ATLAS (Tonry, 2011) and Large Synoptic Survey Telescope (Ivezic et al., 2008) surveys should be ideal for detecting these extremely rare events.

6.4.2 Pan-STARRS1 3π

The Pan-STARRS1 3π survey covers 30,000 square degrees with approximately 60 observational epochs per object (Kaiser et al., 2010; Magnier et al., 2013). The depth and cadence are inferior

to that of the medium deep fields, but the huge amount of sky observed makes it interesting to explore the contribution that this survey could make to the occurrence rate limits if we were to perform a similar analysis on a combined dataset.

We start with an order-of magnitude estimate of the number of WDs we would expect to find in the 3π survey data via reduced proper motion. The exposure times for the 3π survey are 60 seconds vs. 240 seconds for the medium deep fields, but let us assume that our ability to detect WDs is limited by the length of the observational baseline and not by signal to noise of the detections. Since the sky coverage is a factor of ~ 400 greater in the 3π survey it is reasonable to scale the number of astrometrically-selected WDs found in the medium-deep fields (661) by 400. Therefore, we expect to find $\sim 30,000$ WDs via reduced proper motion in the 3π data. Since each WD is observed 60 times this gives a total of 1.8 million measurements. The shorter exposure times increase our sensitivity to very short duration eclipses, however the largest gain in sensitivity is to planets orbiting well inside the tidal destruction radius (see Figure 6.10 and §6.4.1). Combining these 1.8 million epochs with the 4.3 million epochs from the medium-deep fields increases our total number of measurements by a factor of 1.4 and strengthens (decreases) our maximum occurrence constraints by this same factor. This $\sim \sqrt{2}$ improvement would not change our primary conclusion that planets around WDs are rare.

6.5 Conclusions

Our systematic search for eclipses of WDs in the Pan-STARRS1 medium-deep fields places strong constraints on the WD planet occurrence rates. We analyze a sample of ~ 3000 WDs selected via proper motion and color along with a control sample of ~ 1200 stars. These WDs were observed for 5 years on over 4.3 million epochs.

We search for potential eclipses by identifying low outliers in the light curves. A total of 133 candidate eclipses are identified after applying a series of photometry then image-based filters to remove outliers caused by weather, CCD artifacts, or an improperly modeled PSF. After visual inspection of all candidates we find none that is consistent with an eclipse or occultation by a substellar object.

We calculate the number of expected eclipses if every WD hosted at least one planet ($\eta = 1$) by convolving a trapezoidal transit model with the survey exposure time and integrating over all possible geometric orientations and many values of R_p and a . The expected number of eclipses are treated as a Poisson expectation value for the rate of events which are converted into 95% (68%) confidence intervals. We then invert these rates to obtain the maximum value of η that is consistent with our data.

Our results suggest that hot Jupiters around WDs are at least as rare as they are around solar-type stars, occurring around no more than 0.5% (0.2%) of WDs. Hot super-Earths occur around no more than 1.5% (0.6%) of stars, and super-Earths in the CHZ are present around no more than

3.4% (1.3%) of WDs. All evidence presented in this study indicate that short-period planets around WDs are significantly less abundant than short-period planets orbiting main-sequence stars.

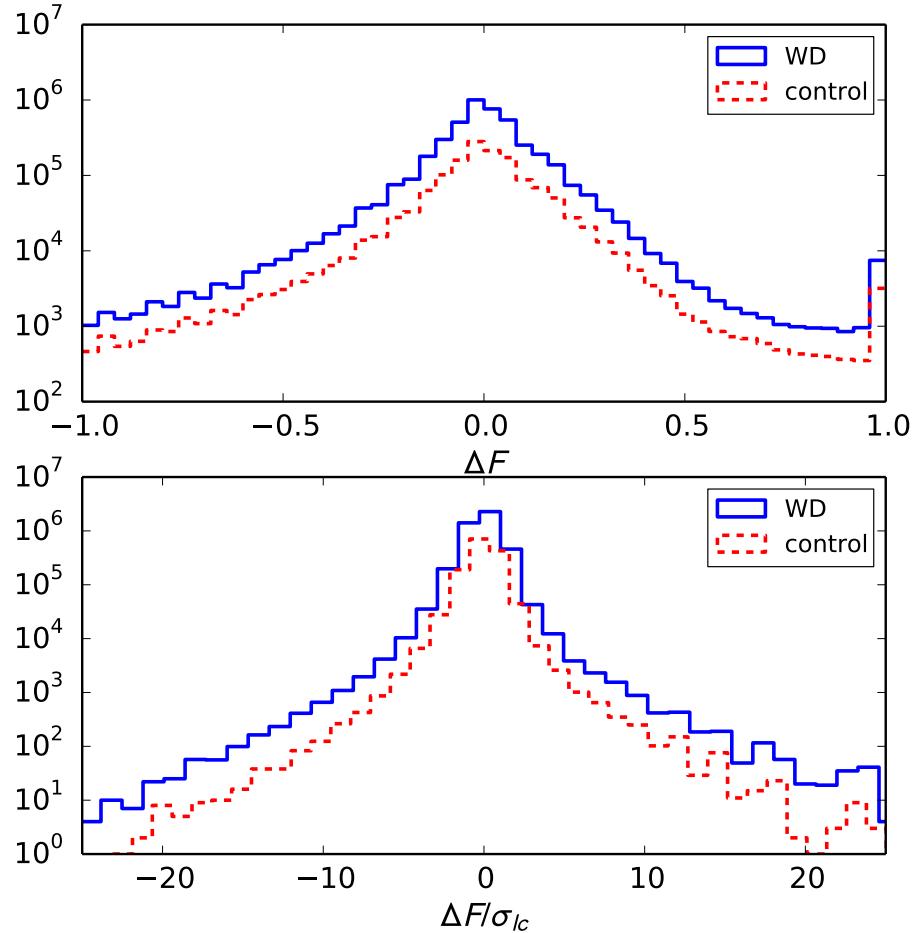


Figure 6.3 *Top:* Distribution of relative flux measurements for all WD and control sample stars. The solid blue line is the distribution for the WDs and the dashed red line is for the control sample stars. *Bottom:* Distribution of relative flux measurements divided by the measurement uncertainties corrected by adding in quadrature the reported measurement uncertainties with the standard deviation of the light curve (by filter). Measurements with $\Delta F/\sigma_{lc} \geq 5$ are considered eclipse candidates. As in the top panel, the solid blue line is the distribution for the WDs and the dashed red line comes from the control sample stars.

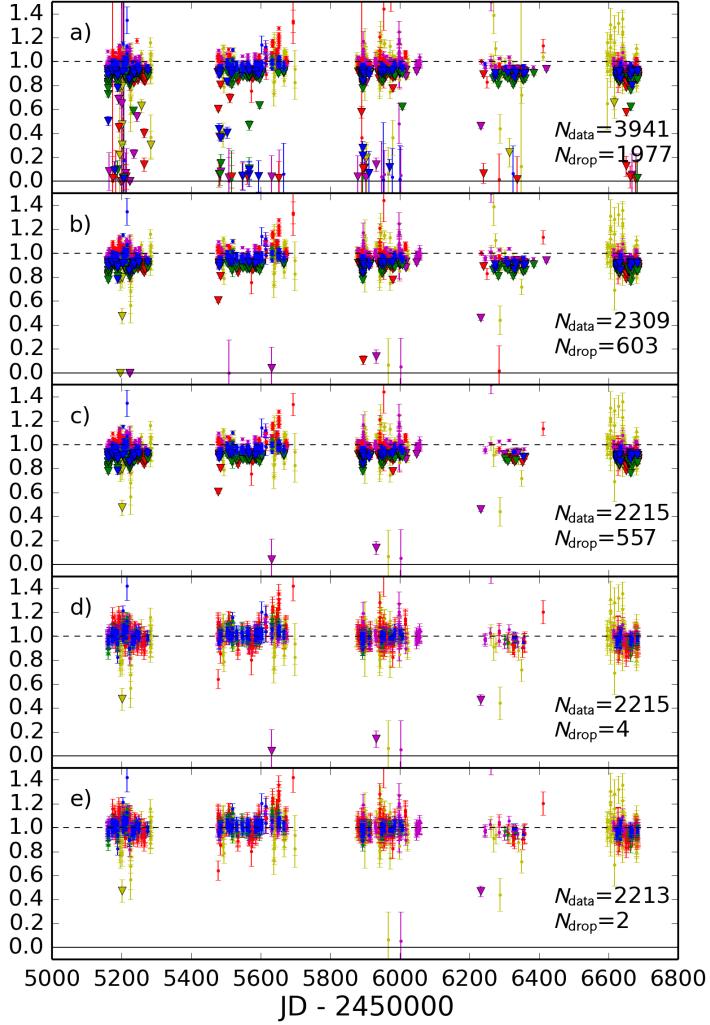


Figure 6.4 Illustration of the filtering process for a light curve of a typical $g' = 18.9$ WD in medium-deep field 3. The total number of measurements and the number of dropouts ($\Delta F/\sigma_{lc} \geq 5$) are shown in the lower-right of each panel. Dropout candidates are plotted as triangles. *a)* Raw light curve before any filtering. Error bars are equivalent to the reported measurement uncertainties. Notice the large number (1977) of dropout candidates. *b)* Light curve after applying the chip location-based filters described in §6.2.4. *c)* Light curve after removing measurements in which neighboring stars show large deviations from the median flux level or large scatter. *d)* Light curve after de-correlating against the neighboring star relative flux and re-scaling the measurement uncertainties by adding the reported uncertainties in quadrature with the standard deviation of the light curve in each filter. This tends to inflate the error bars and pushes the vast majority of dropout events below the 5-sigma cutoff. *e)* Light curve after the final level of photometry-based filtering. In this stage we compare the CCD pixel positions of the stars during dropout events with known masked regions of the CCD array. Two dropout events remain after all photometry-based filters. Postage stamp images are downloaded and visually inspected for the remaining dropout events.

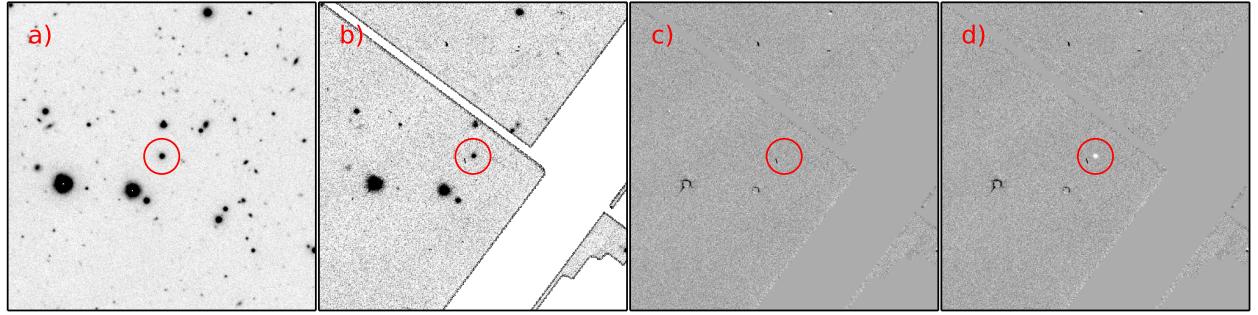


Figure 6.5 Candidate z-band eclipse with a reported depth of 53% that was not filtered by the automated filtering techniques described in §6.2.4. *a)* 5.8 hour stack of the $2.5' \times 2.5'$ region centered on the target. The target WD is circled. *b)* The same field of view as panel *a* from the single exposure corresponding to the reported 53% deep eclipse. *c)* A difference image of the stack in panel *a* convolved and scaled to match the PSF and subtracted from the dropout image in panel *b*. Notice that all stars – including the target – show no detectable residual flux. *d)* Same as panel *c* with a synthetic 53% eclipse injected onto the target before the image subtraction. The negative residuals on the target are clearly evident.

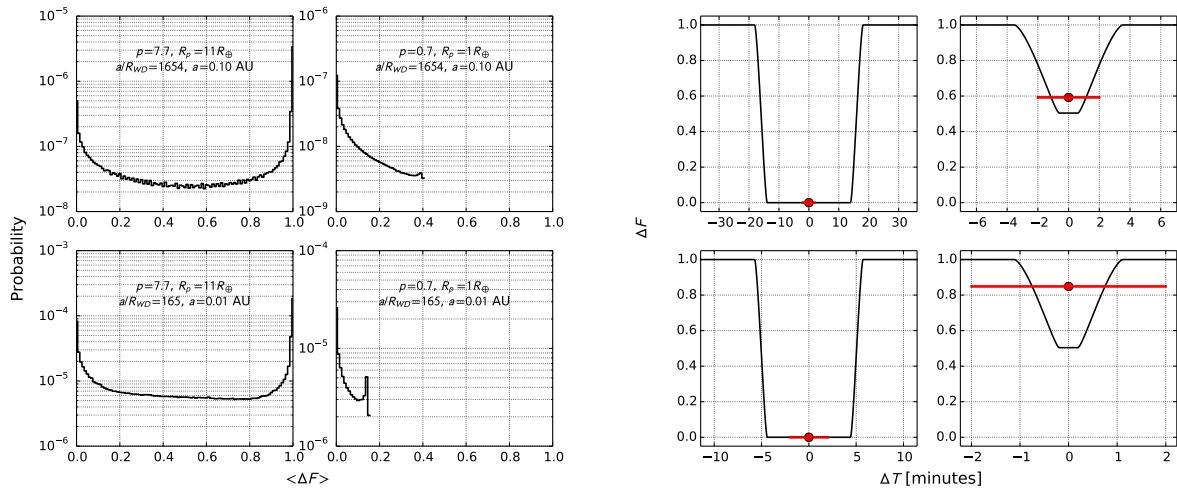


Figure 6.6 *Left:* probability of measuring an eclipse with depth ΔF during a single 240 s exposure of a random WD that hosts a single companion with the orbital parameters shown. p is the planet to star radius ratio, R_p is the radius of the planet in Earth radii, a/R_{WD} is the orbital semi-major axis scaled to the radius of the WD, and a is the semi-major axis in AU. *Right:* Model eclipse light curves for the planet parameters shown on the left panel and an impact parameter 1.0. The red circle is the mean flux for an exposure centered on the mid-eclipse time. The bar extending from the red circle shows the length of the exposure time. This is the largest signal that we could expect to find for planets with these parameters. This corresponds to the maximum $\langle \Delta F \rangle$ bin with a probability greater than zero in the left panel.

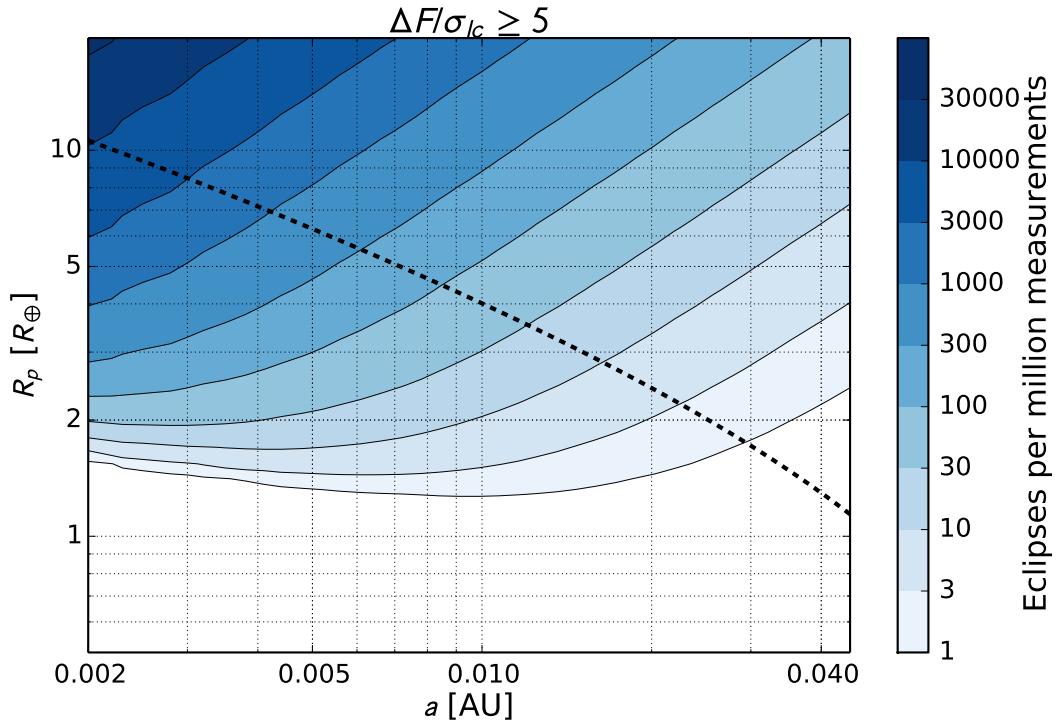


Figure 6.7 Expected detectable eclipse rate per million exposures of the medium-deep survey. An eclipse is deemed detectable if the depth is greater or equal to five times the measurement uncertainty. The measurement uncertainty is calculated by adding the reported uncertainty in quadrature with the standard deviation of the light curve on a per filter basis. The dashed line marks the point at which the eclipse duration is equal to the integration time. Eclipses caused by objects with parameters that fall in the region above and to the right of the dashed line will have eclipses that may span multiple adjacent exposures. Our assumption that each light curve measurement is independent is invalid in this regime and our expected eclipse rate will be slightly overestimated.

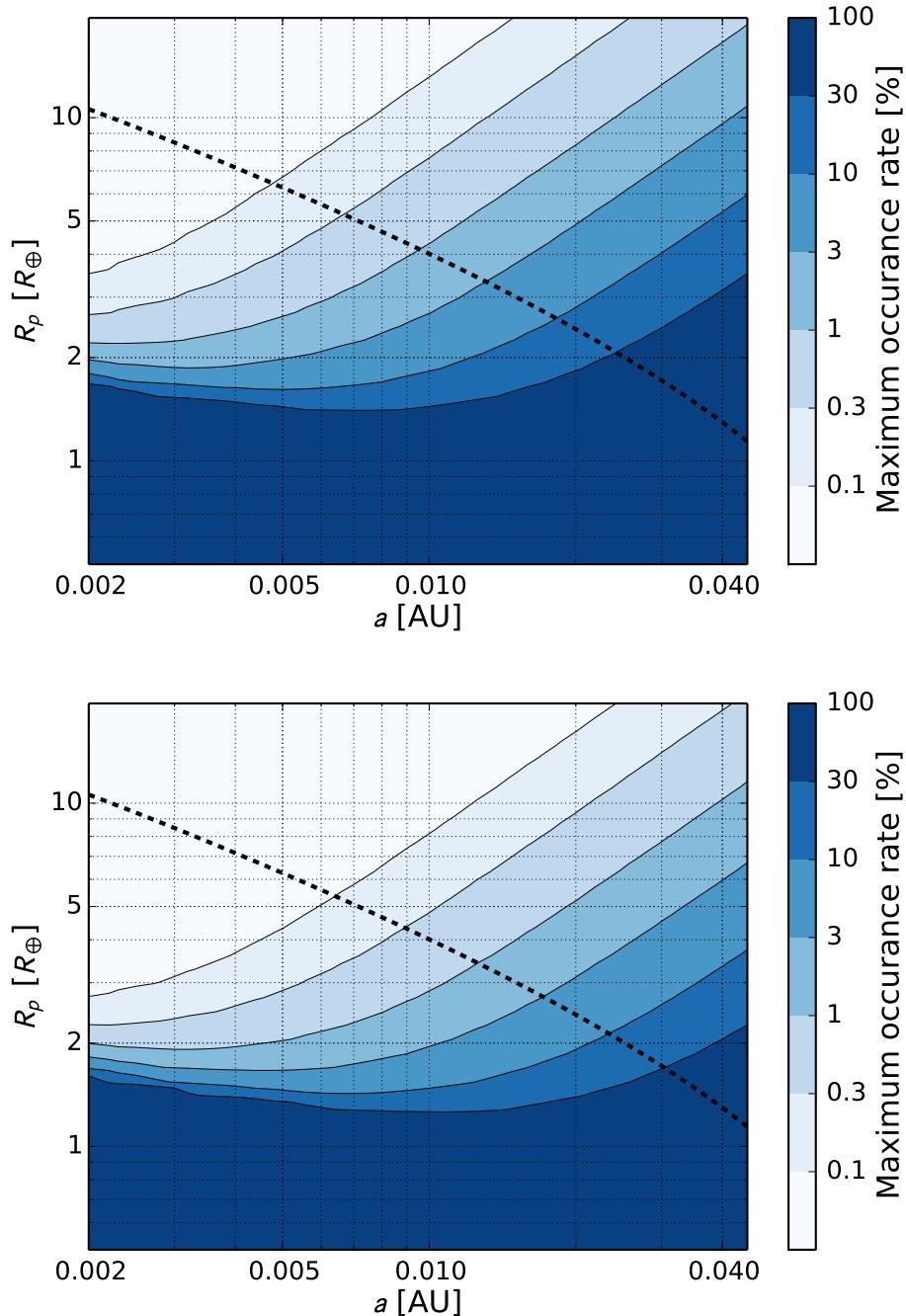


Figure 6.8 *Top*: Maximum planet occurrence rate compatible with the observations at 95% confidence. *Bottom*: Maximum planet occurrence rate compatible with the observations at 68% confidence. In both panels dashed line is the same as in Figure 6.7. The maximum occurrence rates will be slightly underestimated in the region to the upper right of this dashed line.

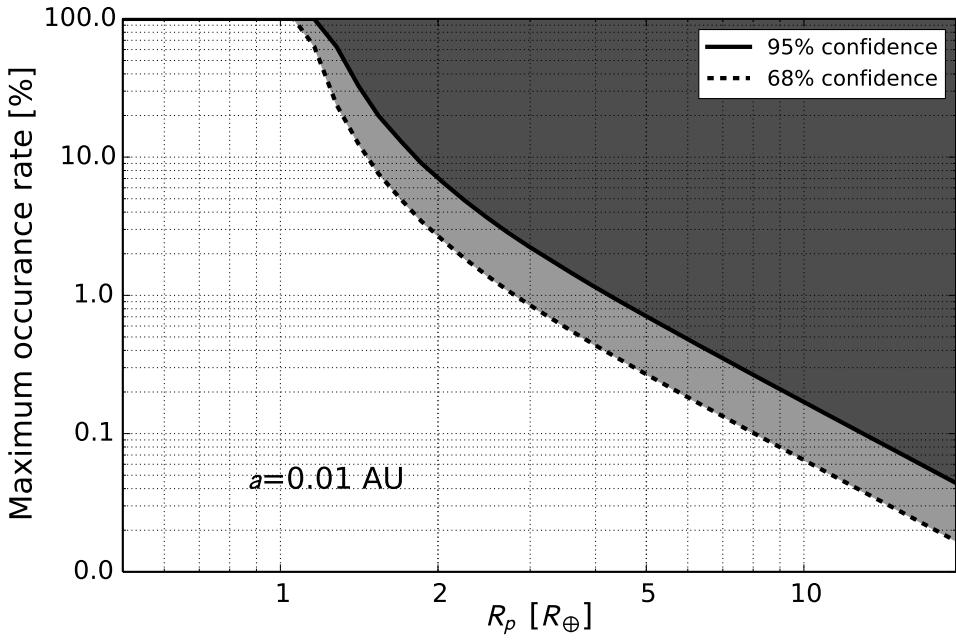


Figure 6.9 Maximum planet occurrence rate consistent with our data as a function of planet radius at a semi-major axis of $a = 0.01$ AU for confidence levels of 95% (solid) and 68% (dashed). Shaded regions are disfavored by our data. This plot represents a slice through Figure 6.8 at $a = 0.01$ AU.

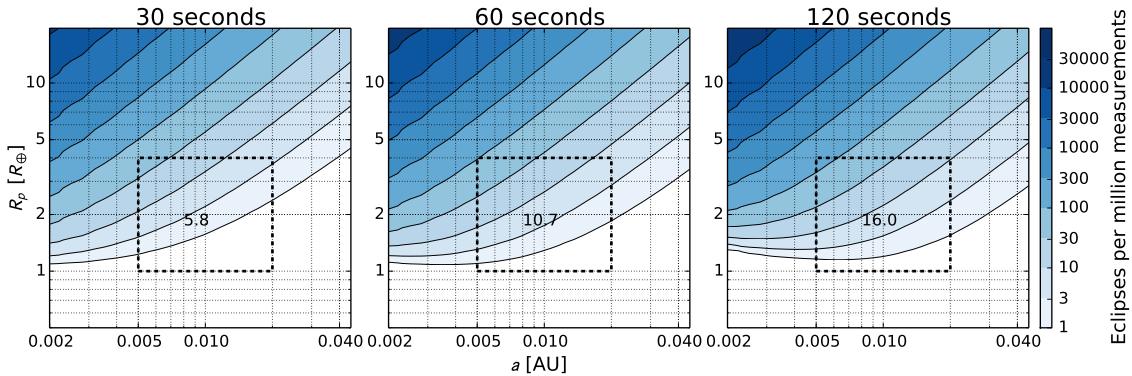


Figure 6.10 Expected detectable eclipse rates calculated as described in §6.3.1 for hypothetical surveys using the Pan-STARRS1-like throughput with different exposure times. The numbers within the dashed box indicate the mean eclipse rate in that region of parameter space. Shorter exposure times give increased eclipse detectability for the shortest-period objects within ~ 0.03 AU but planets orbiting this close to their host WD would likely be ripped apart by tidal forces. Although the mean eclipse rate in the region of interest goes up with longer exposure times this is reversed if you consider a fixed total survey exposure time (take twice as many 60 second exposures as 120 second exposures, etc.). However, the eclipse rates remain nearly constant indicating that the best way to increase sensitivity in this regime is to increase the number of epochs observed (larger number of WDs and/or higher cadence).

CHAPTER 7

MASS DISTRIBUTION OF LOW-MASS PLANETS IN THE SOLAR NEIGHBORHOOD

7.1 Introduction

The field of exoplanet demographics has exploded in the era of the *Kepler* space telescope. *Kepler* discovered thousands of planets and, for the first time, allowed for studies of exoplanets as populations instead of individual objects. One of the most impressive results from the *Kepler* mission was the discovery that small planets are extremely common (e.g. Howard et al., 2012b; Petigura et al., 2013a; Fressin et al., 2013). Fulton et al. (2017) also discovered that the distribution of those small, common planets is bimodal. There is a significant dearth of planets with radii between $1.5\text{--}2 R_{\oplus}$. However, Kepler surveyed only one small patch of sky and measured planet radii, not masses. The underlying mass distribution of those planets is important to understand the range of bulk compositions within this population. Previous radial velocity (RV) surveys have also shown that low-mass planets are common, but have not yet been able to resolve fine structure in the mass distribution which might encode important information about the formation of these planets (Howard et al., 2010a; Mayor et al., 2011).

We have conducted a survey of 51 nearby, Sun-like stars to search for close-in, sub-Neptune mass planets in order to measure the detailed shape of the mass function of small planets. This “APF-50” survey utilized the Automated Planet Finder (APF, Vogt et al., 2014b; Radovan et al., 2014) telescope to autonomously monitor the stars at high cadence for the duration of the survey. It builds on previous surveys, but with better precision due to high observing cadence and a larger number of measurements made possible by the automated and dedicated nature of the APF facility. We previously discovered three Neptune-mass planets as part of the survey (Fulton et al., 2016) and announce the detection of nine new planets with masses in the range $7.5\text{--}12.8 M_{\oplus}$. Using these planets and the larger set of previously-known planets that were re-detected by our survey, we measure the occurrence rate and mass distribution of low-mass, short-period planets in our local neighborhood.

7.2 Observations

7.2.1 Sample Selection

The APF-50 target list is a subset of the Eta-Earth sample of stars (Howard et al., 2010a), which consisted of 235 G, K and M-type dwarf stars selected from the Hipparcos catalogue (Perryman & ESA, 1997). They selected stars using cuts in brightness, distance, luminosity, chromospheric

activity ($\log R'_{\text{HK}}$), lack of stellar companions, and observability from Keck Observatory. We use these same metrics to select from the Hipparcos catalogue, but with more stringent requirements in brightness and $\log R'_{\text{HK}}$. We do not impose a $B - V$ color cut, but the luminosity and apparent magnitude filters ensure that the majority of the stars in the sample are G and K dwarfs. Our sample consists of the “best” Eta-Earth targets that have the smallest astrophysical sources of systematic noise (RV “jitter”). Survey simulations including targets from other programs showed that ~ 50 targets is the maximum number that can be observed with the APF every night that the stars are visible in the night sky. This requirement of ~ 50 stars drove the magnitude cut to include only stars brighter than 7th magnitude in V . The final target list consists of 51 stars. A comparison of the cuts used for this survey compared with those of the Eta-Earth survey can be found in Table 7.1 and Figure 7.1 shows the locations of the APF-50 targets on an H-R diagram.

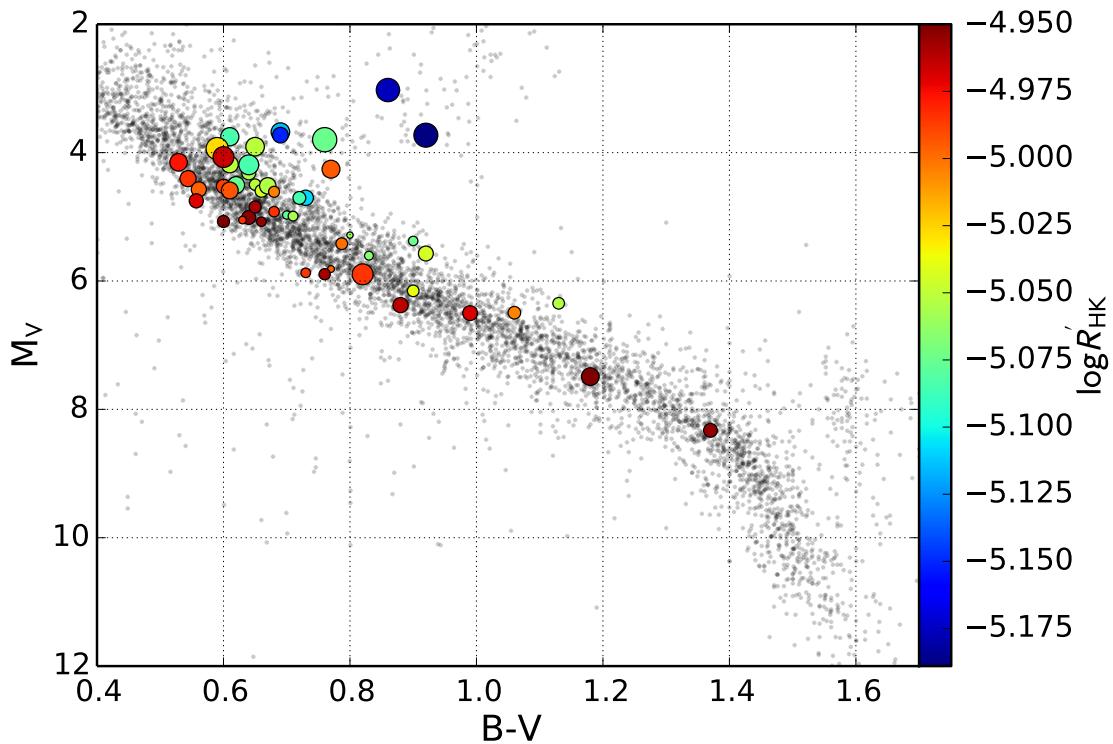


Figure 7.1 Hipparcos H-R diagram showing the APF-50 targets. The large colored points are the APF-50 targets. The point size indicates apparent magnitude with larger points being brighter stars. The color of the point reflects stellar activity as measured from the Ca II H & K index. The light grey points are other Hipparcos stars within 50 pc that were not selected for the APF-50 survey.

Table 7.1. APF-50 Target Selection Filters

Parameter	Eta-Earth	APF-50
apparent magnitude (V)	< 11	< 7
distance	< 25 pc	< 25 pc
luminosity (M_V)	> 3	> 3
chromospheric activity ($\log R'_{HK}$)	< -4.7	< -4.95
declination (degrees)	> -30	> -10

7.2.2 Radial Velocities

All stars were observed using the standard high-precision observing techniques described in Fulton et al. (2015b, 2016). The starlight is passed through a cell of gaseous iodine which superimposes a dense forest of molecular absorption lines to be used as a wavelength and point spread function (PSF) fiducial. We extract RVs by forward modeling the instrumental PSF, wavelength solution, and RV shift in \sim 800 small spectral regions (Butler & Marcy, 1996).

We aimed to acquire an average of 100 observations for each star in the survey. We exceeded that goal with a mean of 125 observations per star. Only 7 of the 51 stars were observed less than 100 times (Table 7.2) and every star was observed at least 82 times. Figure 7.2 shows the number of observations collected for all targets on both APF and Keck.

The median velocity uncertainties are about a factor of two higher for the APF observations relative to Keck/HIRES (Vogt et al., 1994) observations of the same stars. We generally collect $\text{SNR} \geq 200$ on Keck, but closer to $\text{SNR}=150$ on APF. The median velocity uncertainty at APF is still a factor of ~ 2 below the velocity RMS after subtracting any known planets or planet candidates (Figure 7.3). The RMS distribution for both the Keck and APF data match very well even with the higher SNR and smaller measurement uncertainties at Keck. This is a good indication that our precision is not limited by photons, but instead limited by astrophysical and instrumental sources of noise.

Table 7.2. APF-50 Targets

HD #	RA	Dec.	Distance (pc)	V (mag)	B–V (mag)	M _V (mag)	$\log R'_{\text{HK}}$	N _{APF} ¹ (nights)
1461	0	-8	23.4	6.46	0.68	4.61	-5.00	121
3651	1	21	11.1	5.80	0.92	5.57	-5.04	118
4628	1	5	7.5	5.75	0.88	6.38	-4.96	188
9407	2	69	21.0	6.53	0.68	4.92	-4.98	192
16160	3	7	7.2	5.82	0.99	6.50	-4.97	104
19373	3	50	10.5	4.05	0.59	3.94	-5.03	135
23249	4	-10	9.0	3.51	0.92	3.73	-5.19	110
26965	4	-8	5.0	4.41	0.82	5.90	-4.99	116
32147	5	-6	8.8	6.22	1.06	6.49	-5.00	99
32923	5	19	15.9	5.01	0.65	3.91	-5.05	100
34411	5	40	12.6	4.70	0.64	4.19	-5.08	111
38858	6	-4	15.6	5.97	0.64	5.01	-4.95	88
40397	6	-5	23.2	6.80	0.70	4.97	-5.08	82
42618	6	7	23.1	6.87	0.63	5.05	-4.98	89
48682	7	44	16.5	5.25	0.53	4.15	-4.98	129
50692	7	25	17.3	5.76	0.56	4.57	-5.00	91
52711	7	29	19.1	5.93	0.60	4.53	-4.99	111
55575	7	47	16.9	5.55	0.54	4.41	-4.99	102
65583	8	29	16.8	6.94	0.77	5.81	-5.00	99
84737	10	46	18.4	5.10	0.61	3.75	-5.08	115
86728	10	32	14.9	5.40	0.62	4.50	-5.08	107
89269	10	44	20.6	6.65	0.66	5.08	-4.96	83
110897	13	39	17.4	5.95	0.56	4.75	-4.97	117
117176	13	14	18.1	5.00	0.69	3.68	-5.12	119
122064	14	61	10.1	6.37	1.13	6.35	-5.05	101
126053	14	1	17.6	6.30	0.60	5.07	-4.95	126
127334	14	42	23.6	6.40	0.65	4.50	-5.05	119
141004	16	7	11.8	4.43	0.60	4.07	-4.96	173
143761	16	33	17.4	5.40	0.61	4.18	-5.05	132
144579	16	39	14.4	6.66	0.73	5.87	-4.99	111
145675	16	44	18.1	6.67	0.90	5.38	-5.08	114
157214	17	32	14.4	5.40	0.61	4.59	-4.99	119
157347	17	-2	19.5	6.29	0.65	4.85	-4.96	112
158633	17	67	12.8	6.43	0.76	5.89	-4.96	120
161797	18	28	8.4	3.41	0.76	3.80	-5.08	151
164922	18	26	21.9	6.99	0.80	5.29	-5.06	157
166620	18	38	11.1	6.37	0.90	6.15	-5.04	114
168009	18	45	22.7	6.31	0.60	4.52	-5.00	162
182488	19	33	15.5	6.36	0.79	5.42	-5.00	145
182572	19	12	15.1	5.16	0.77	4.26	-4.99	121
186408	20	51	21.6	5.96	0.64	4.32	-5.06	144
186427	20	51	21.4	6.20	0.66	4.60	-5.04	170
188512	20	6	13.7	3.71	0.86	3.03	-5.18	138
190360	20	30	15.9	5.71	0.73	4.70	-5.11	195
195564	21	-10	24.2	5.65	0.69	3.73	-5.15	121
201091	21	39	3.5	5.21	1.18	7.49	-4.95	136
201092	21	39	3.5	6.03	1.37	8.33	-4.95	121
210277	22	-8	21.3	6.63	0.71	4.99	-5.06	118
217014	23	21	15.4	5.49	0.67	4.52	-5.05	125
217107	23	-2	19.7	6.18	0.72	4.71	-5.08	139
221354	24	59	16.9	6.74	0.83	5.61	-5.07	119

¹Number of nights observed on APF as of Jun 14 2017

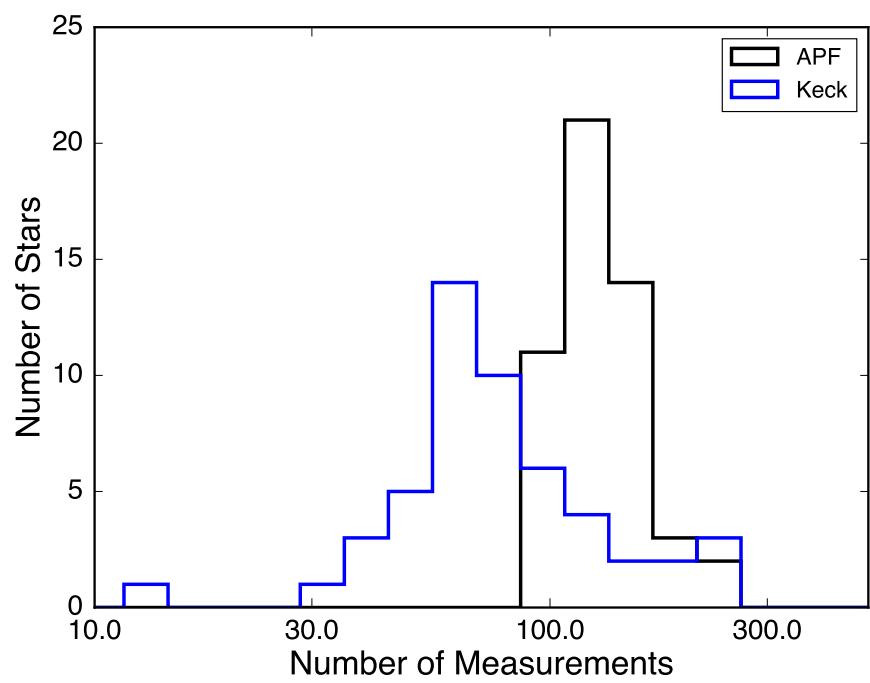


Figure 7.2 Number distributions of RV measurements.

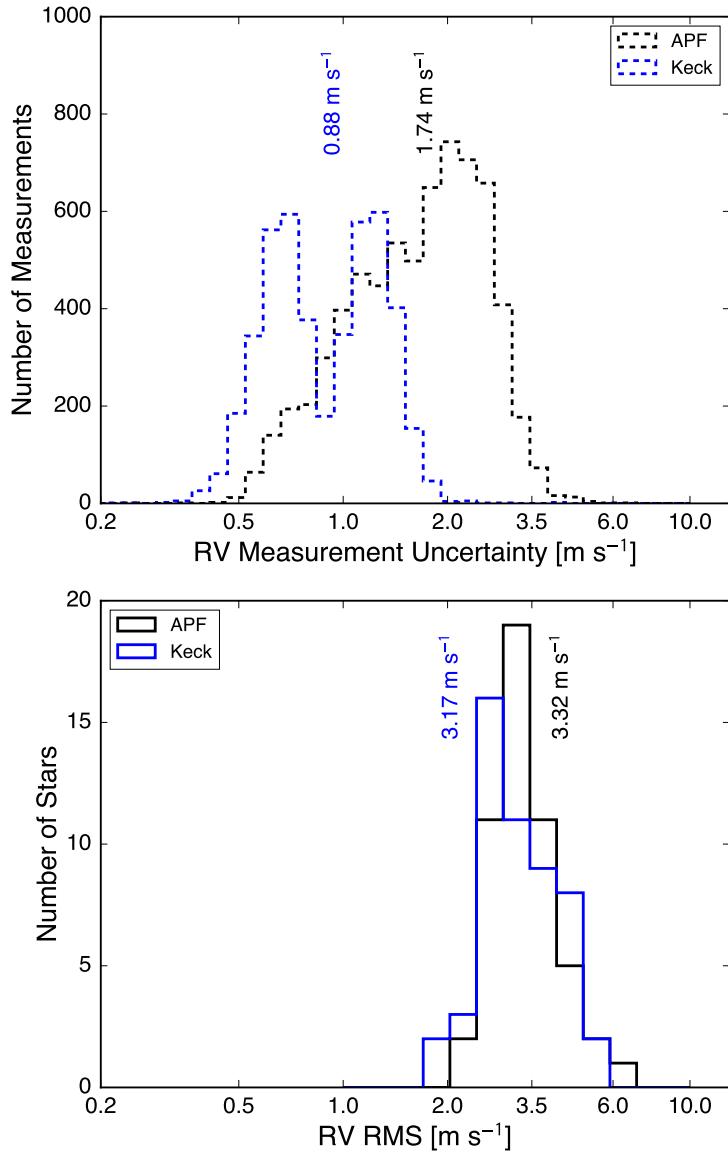


Figure 7.3 *Top*: Distributions of velocity uncertainties for the APF-50 survey. The median velocity uncertainty for Keck and APF is annotated. The bimodal distribution is driven by differences in observing strategies. The measurements with smaller uncertainties consist of three consecutive measurements binned together, while the larger uncertainties consist of only a single spectrum. *Bottom*: Distribution of the RMS of the velocity residuals after subtracting all known signals. The medians of both distributions are annotated.

7.3 Analysis

7.3.1 Automated planet search

Traditionally planet search algorithms involve searching for periodic signals in the data by fitting sinusoids on a finely spaced grid of frequencies and looking for periods at which the χ^2 is significantly reduced. This method falls under the broader formalism of Least-Squares spectral analysis and is similar to Fourier analysis. For RV planet searches with highly non-uniform observing cadence Lomb-Scargle is the periodogram of choice due to its simplicity, availability, and speed (Lomb, 1976; Scargle, 1982).

We search for planets in the RV data using an iterative multi-planet detection algorithm that is based on the two-dimensional Keplerian Lomb-Scargle periodogram (2DKLS, O'Toole et al., 2009). This periodogram is created by fitting a Keplerian RV model to the dataset at many starting points on a two-dimensional grid of period and eccentricity. The primary advantages of the 2DKLS periodogram over the Lomb-Scargle periodogram is that it is more sensitive to eccentric planets, and measurement errors and offsets between data from different instruments can be incorporated into the fit. Our model is parameterized by the orbital period (P), time of periastron (T_P), eccentricity (e), the argument of periastron of the star's orbit (ω_*), the velocity semi-amplitude (K), an arbitrary RV zero-point (γ), and in cases where there is a significant linear trend observed a RV slope ($\dot{\gamma}$). We fit the model to the data using the Levenberg-Marquardt-based RV fitting package (RVLIN, Wright & Howard, 2009). The total RV model then becomes a sum of the single planet models, but with γ (and $\dot{\gamma}$ if present) being shared by all planets. We define our grid of search periods following the prescription of Horne & Baliunas (1986) for optimal frequency sampling, and at each period we start an L-M fit at five eccentricities between 0 and 0.7. All parameters are free to vary in each fit, but period and eccentricity are constrained to intervals that allow them to vary only half the distance to the immediately adjacent search periods and eccentricities. All parameters for any previously detected planets are completely unconstrained so that slightly incorrect fits (sometimes caused by the presence of other planets) for any previous planets do not lead to the false detection of additional planets. The periodogram power is defined as:

$$Z(P, e) = \frac{\chi^2 - \chi_B^2}{\chi_B^2} \quad (7.1)$$

where χ^2 is the sum of the squared residuals to the current N -planet Keplerian fit, and χ_B^2 is the sum of the squared residuals to the best $N-1$ -planet fit. In the first iteration of the planet search (comparing 1-planet models to 0-planet models) χ_B^2 is simply the squared residuals to the mean or a linear fit.

We start the iterative planet search by fitting for any known planet candidates in the system using catalogued orbital parameters as initial guesses for RVLIN. We then look for a significant

linear trend in the RV time series by fitting a line to the RVs and checking if the total change in RV due to the fitted line is greater than 10 times the median of the individual measurement errors. If a significant trend is detected we allow $\dot{\gamma}$ to vary at each point in the 2DKLS periodogram, or if no trend is detected using this criterion we fix $\dot{\gamma}$ to 0. If there are known planets or candidates in the system we start by creating a periodogram checking that an $N+1$ planet model is a better fit, otherwise we start by checking for 1 planet and comparing to the null hypothesis (no planets exist around this star).

Since the statistical errors on the RV measurements are generally underestimated due to the presence of systematic and/or poorly-understood astrophysical noise, the periodogram power can not be directly converted into a significance estimate using traditional χ^2 statistics. Instead we derive an empirical false alarm probability (eFAP) by fitting a histogram of the periodogram values higher than the median power value to a linear function in $\log(n)$ vs. $\log(Z)$. This provides an estimate of the number of peaks that should fall above a given value, and when multiplied by the number of independent test periods, the approximate probability that we would find a peak of a given value within the particular periodogram. We use this to define a threshold of eFAP=1% above which if a spike in the periodogram is detected we mark it as a planet candidate and continue searching for the next potential planet in the system. If no significant signals are detected we stop the search. We adopt 1% as our cutoff FAP by making sure that it is low enough to detect all of the previously published planets in our sample, but high enough that any signals falling above this threshold are believable with a high level of confidence upon visual examination.

7.3.2 Sensitivity to injected signals

In order to determine our sensitivity to planets as a function of $M \sin i$ and period we inject synthetic planetary RV signals into the real RV data for each star and use our iterative automated planet search algorithm to attempt to recover the injected signals. We inject planets on circular orbits uniformly distributed in $\log(K)$ and $\log(P)$ centered around the sensitivity threshold line determined from the method of Howard et al. (2010b). We inject 3000 synthetic planets into each of the 51 stars in our sample. The synthetic planets are injected on top of any known planets in the system and our search algorithm starts by searching for one additional planet while simultaneously fitting for all known planets in the system. An injected planet is considered recovered if the highest peak in the periodogram is above our detection threshold, the period of that peak is within 25% of the injected period, and the phase of the recovered orbit is within $\pi/6$ of the injected phase. In some cases multiple planets are recovered when only a single planet is injected. We still consider these cases good recoveries if any of the detected periods are within 25% of the injected period. We then sum all of the injections and recoveries for all of the stars in our sample in order to derive 2-dimensional completeness contours for the survey (Figure 7.4).

7.3.3 Completeness Corrections

We account for the planets we may have missed to due lack of completeness following the technique of Fulton et al. (2017). For each planet detection there are a number of similar planets that would not have been detected due to a lack of sensitivity as characterized by the injection/recovery tests described in Section 7.3.2. To compensate, we weighted each planet detection by the inverse of these probabilities,

$$w_i = \frac{1}{p_{\text{det}}}, \quad (7.2)$$

where p_{det} is the fraction of stars in our sample where a planet with a given period (P) and velocity semi-amplitude (k) axis could be detected:

$$p_{\text{det}} = \frac{1}{N_\star} \sum_i^{N_\star} C(P, k). \quad (7.3)$$

$C(P, k)$ is measured directly from the injection/recovery tests. The recovery contours (e.g. as plotted in Figure 7.5) are interpolated to the precise values of P and k for each detected planet. $N_\star = 51$ is the total number of stars in our sample.

The average planet occurrence rate (number of planets per star) for any discrete bin in planet radius or orbital period is the sum of these weights divided by the total number of stars in the sample (N_\star):

$$f_{\text{bin}} = \frac{1}{N_\star} \sum_{i=1}^{n_{\text{pl,bin}}} w_i. \quad (7.4)$$

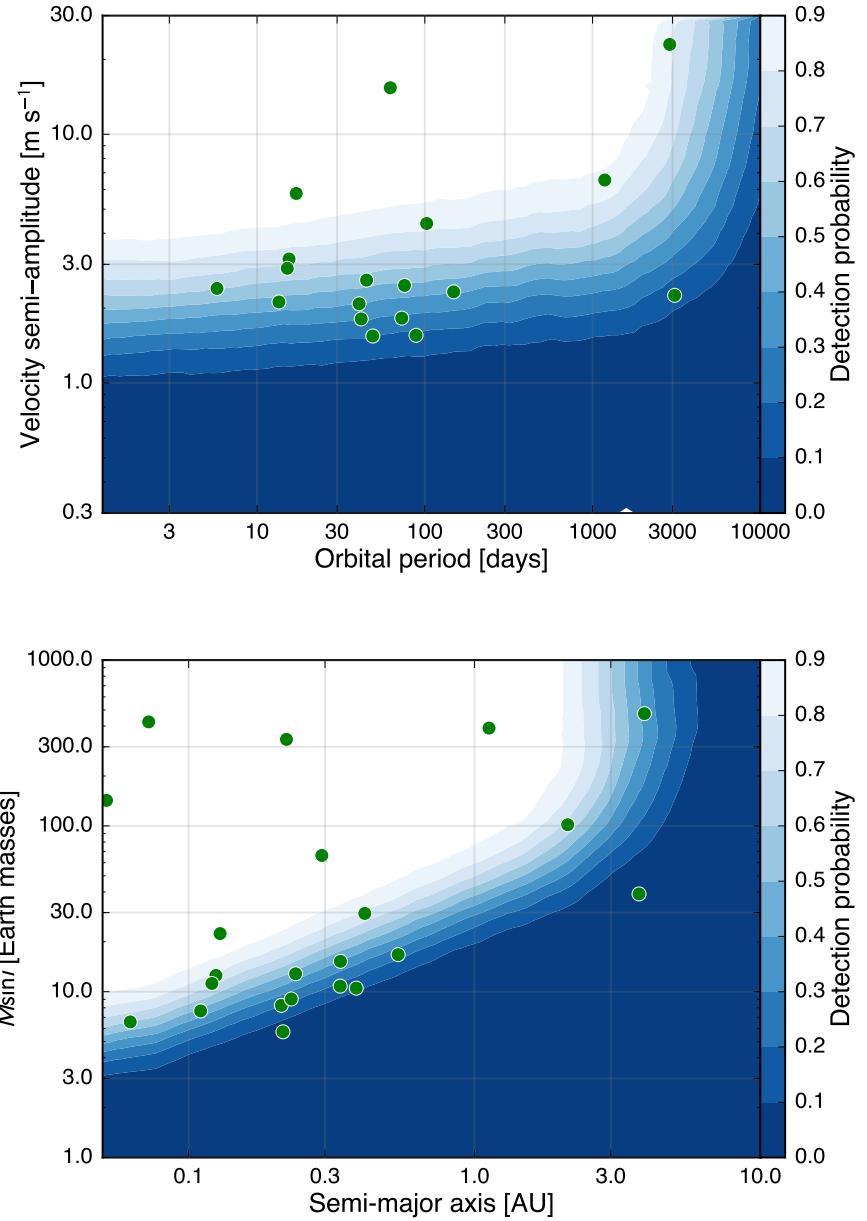


Figure 7.4 *Top*: Mean survey completeness as a function of velocity semi-amplitude and orbital period. Planet detections are plotted as green circles. *Bottom*: Mean survey completeness as a function of $M \sin i$ and semi-major axis.

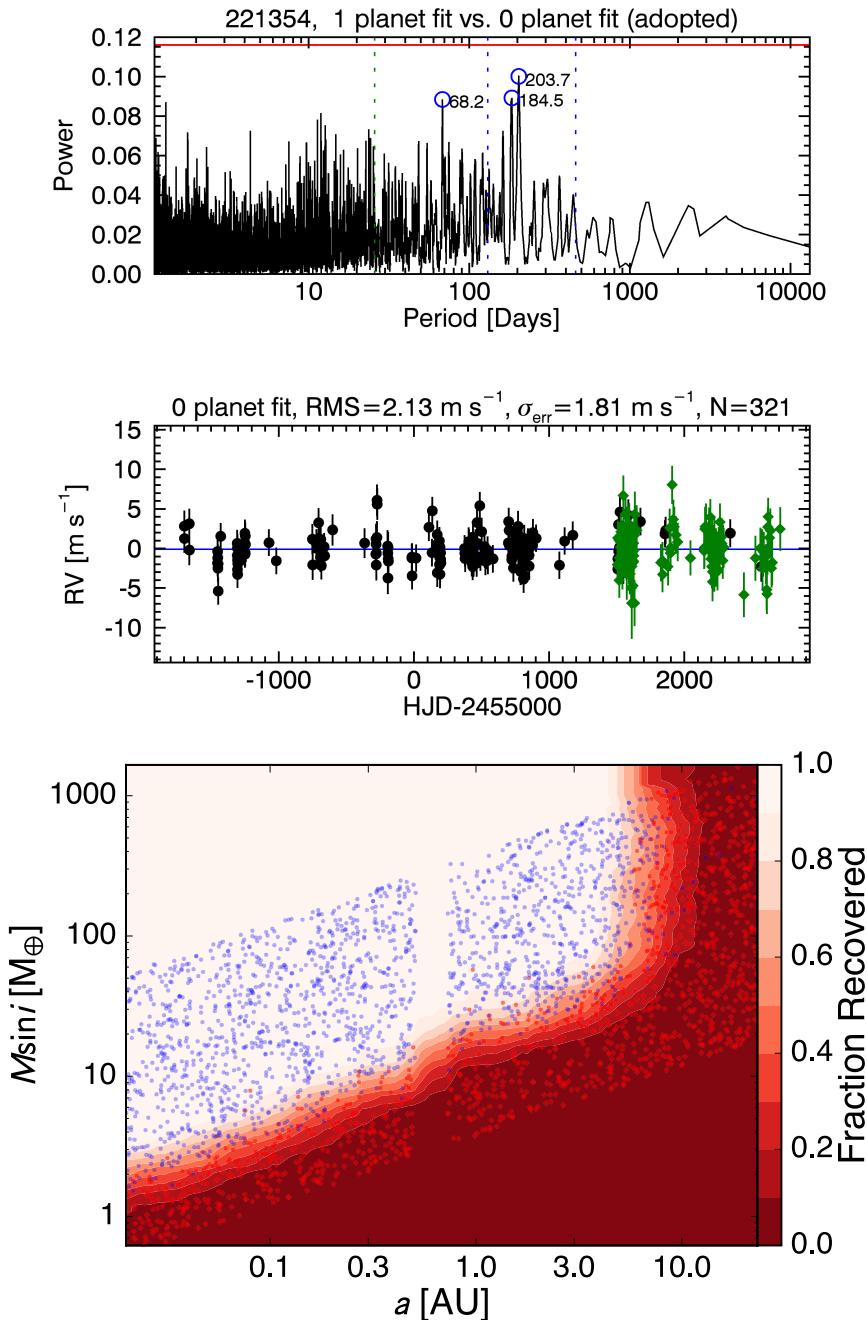


Figure 7.5 *Top:* 2DKLS periodogram and RV time-series of the quiet star HD 221354. The solid red line is our cutoff for a positive detection for an eFAP=1%. The dotted vertical lines mark the 1-year and 1-month period aliases of the highest period peak (203.7 days). The periodogram shows that no signals fall above or near our detection threshold. *Bottom:* Results of our injection recovery tests for HD 221354. The blue dots indicate the position of injected planets that were successfully recovered by the pipeline, and red dots indicate injections that were not recovered.

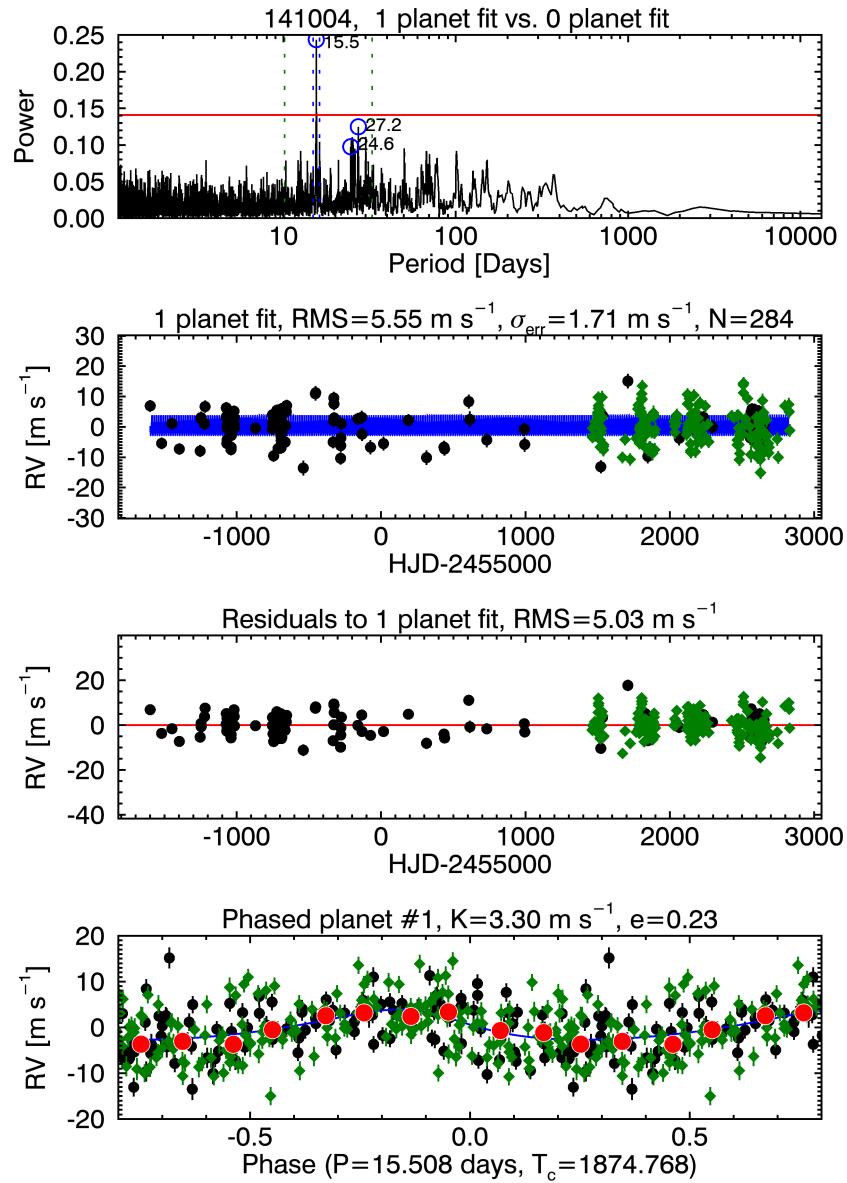


Figure 7.6 Example candidate vetting plots for HD 141004 b produced by the automated planet detection algorithm. A similar plot is produced at each iteration of the planet search. *Top:* 2DKLS periodogram of the RV time-series comparing the χ^2 of a single-planet Keplerian fit to that of the null hypothesis (flat line). The red line is our cutoff for a positive detection for an eFAP=1%. The dotted vertical lines mark the 1-year and 1-month period aliases of the highest period peak (15.5 days). *Top-middle:* the complete RV time-series for HD 141004 that spans \sim 10 years with the best-fitting single-planet Keplerian fit over-plotted in blue. *Bottom-middle:* residuals to the single-planet Keplerian fit. *Bottom:* the RV data phase-folded to the period corresponding to the highest peak in the periodogram and binned in orbital phase (large red points).

7.4 Results

Figure 7.4 shows the periods, amplitudes, masses and semi-major axes of all new candidates and previously-known planets that were redetected by the automated detection pipeline. We also show the mean sensitivity contours for the survey. Next we briefly discuss all new planet candidates.

7.4.1 HD 1461

HD 1461 was previously known to host two planets with periods of 5.8 and 13.5 days (Rivera et al., 2010; Díaz et al., 2016). We detect a third planet with an orbital period of 73 days and a mass of $11 M_{\oplus}$ (Figure 7.7). We detect a long-period trend in the S_{HK} values with a period of ~ 10 years but do not detect any corresponding long-period variability in the RV time series. There is no evidence of short-period modulation of the S_{HK} values at the rotation period. We also find a signal with a period of 372 days that we attribute to systematic noise correlated with barycentric corrections.

7.4.2 HD 26965

We detect a single low-mass planet orbiting HD 26965 with a period of 40 days (Figure 7.8). We unambiguously detect the rotation period in the S_{HK} measurements at a period of 37 days, but the radial velocity variations do not phase up at this period. We detect a second signal at 24 days that falls just below the 1% eFAP threshold. We calculate a series of periodograms, starting with only the first 20 measurements, we progressively add more measurements and recalculate the 2DKLS periodograms at each step. We then plot the periodogram peak height at the period of the peak found in the original periodogram (we call this technique a “running periodogram”). If the source of the periodogram peak is coherent over the entire observational baseline we would expect to see a monotonic increase in the periodogram peak height. We find that the peak height grows nearly monotonically for both the 40-day and 24-day signals.

7.4.3 HD 32147

We detect a new, Neptune-mass planet orbiting HD 32147 with a period of 3085 days and a semi-major axis of 3.8 AU (Figure 7.9). This is near the period of the magnetic activity period which we detect in the S_{HK} values with a period of ~ 10 years, but the RVs are not in phase with the S_{HK} measurements as is usually the case when we see RV signals at the same period as the magnetic activity cycle. We see tentative evidence of the rotation period at 27 days in the S_{HK} measurements, but no further evidence of any signals in the RV time series.

7.4.4 HD 55575

We detect a relatively strong signal at a period of 45 days with an amplitude of 2.6 m s^{-1} (Figure 7.10). We see no evidence of any periodic signal in the S_{HK} measurements or any additional signals in the RV time series.

7.4.5 HD 141004

We detect strong signals at 15.5 days, and 25.1 days in the RV time series of HD 141004 (see Figure 7.6). However, we also detect a strong peak in a periodogram of the S_{HK} values at 25.3 days which phases up reasonably well with the RV measurements. We conclude that the 25.1 day signal is caused by rotational modulation of starspots and we do not include it in our final planet catalogue (Fulton et al., 2016).

7.4.6 HD 164922

This star was previously known to host two planets. A Jupiter-mass planet orbiting at 2 AU was detected by Butler et al. (2006), and a temperate Neptune-mass planet was discovered as part of this survey with an orbital period of 75 days (Fulton et al., 2016). Here we detect another significant, periodic signal with a period of 41 days (Figure 7.11). This signal was marginal at the time Fulton et al. (2016) was published, but it's significance has continued to increase with time. The candidate is now detected with an eFAP of 0.1%. There is no significant periodicity in the S_{HK} values at any period. We detect another signal with a period of 12 days at lower significance which will require continued monitoring to establish it's nature.

7.4.7 HD 168009

HD 168009 is a late F star with properties very similar to HD 141004. Like HD 141004 b we also detect a sub-Neptune-mass planet orbiting with a period of 15 days (Figure 7.12). We clearly detect the stellar rotation period at 29.3 days in both the S_{HK} values and the RV timeseries. We also find a periodic signal with a period near 1 year which we attribute to systematic problems correlated with barycentric corrections.

7.4.8 HD 190360

HD 190360 is a well-studied star hosting two previously-discovered planets. HD 190360 b is a Jupiter-mass planet orbiting at 4 AU and HD 190360 c is a short-period Neptune-mass planet with a period of 17 days. Here we detect a third planet with a mass of $11 M_{\oplus}$ and a period of 89 days (Figure 7.13). We see no long-term or short-term variability in the S_{HK} values.

Table 7.3. Planet Detections

Planet Name	Period (days)	K (m s $^{-1}$)	$M_p \sin i_p$ (M_{\oplus})	a (AU)	Discovery Reference
HD 1461 b	5.77161 ± 0.00039	2.4 ± 0.26	6.58 ± 0.76	$0.0624995 \pm 2.8e - 06$	Rivera et al. (2010)
HD 1461 c	13.5047 ± 0.0026	2.12 ± 0.26	7.65 ± 0.97	$0.110155 \pm 1.4e - 05$	Diaz et al. (2016)
HD 1461 d	72.91 ± 0.078	1.82 ± 0.32	10.8 ± 1.7	0.33902 ± 0.00028	this work
HD 3651 b	62.257 ± 0.013	15.4 ± 0.6	66.3 ± 2.6	$0.292403 \pm 3.6e - 05$	Fischer et al. (2003)
HD 26965 b	40.611 ± 0.019	2.08 ± 0.38	8.3 ± 1.2	$0.211032 \pm 6.7e - 05$	this work
HD 32147 b	3085 ± 230	2.25 ± 0.35	39 ± 6	3.76 ± 0.17	this work
HD 42618 b	148.49 ± 0.25	2.32 ± 0.56	16.7 ± 2.3	0.54085 ± 0.00063	Fulton et al. (2016)
HD 55575 b	44.975 ± 0.044	2.59 ± 0.42	12.8 ± 2.1	0.23666 ± 0.00016	this work
70 Vir b / HD 117176 b	116.6946 ± 0.0021	315.7 ± 0.6	2264 ± 130	$0.4726554 \pm 5.3e - 06$	Marcy & Butler (1996)
HD 141004 b	15.51 ± 0.0044	3.15 ± 0.46	12.5 ± 1.8	$0.124605 \pm 2.2e - 05$	this work
rho CrB b / HD 143761 b	39.8438 ± 0.0011	67.1 ± 0.25	333 ± 11	$0.2199101 \pm 3.6e - 06$	Noyes et al. (1997)
rho CrB c / HD 143761 c	102.61 ± 0.12	4.38 ± 0.26	30 ± 2	0.4132 ± 0.0003	Fulton et al. (2016)
14 Her b / HD 145675 b	1773 ± 2	89.5 ± 0.7	1575 ± 45	2.863 ± 0.002	Butler et al. (2003)
HD 164922 b	1182.5 ± 6.7	6.55 ± 0.27	101.7 ± 5.1	2.1224 ± 0.0081	Butler et al. (2006)
HD 164922 c	75.7 ± 0.08	2.47 ± 0.26	15.2 ± 1.6	0.33967 ± 0.00024	Fulton et al. (2016)
HD 164922 d	41.774 ± 0.034	1.81 ± 0.28	9.0 ± 1.4	0.22855 ± 0.00012	this work
HD 168009 b	15.1396 ± 0.0055	2.89 ± 0.41	11.2 ± 1.6	$0.120606 \pm 2.7e - 05$	this work
16 Cyg B b / HD 186427 b	800.25 ± 0.36	53.03 ± 0.37	554 ± 17	1.67302 ± 0.00051	Cochran et al. (1997)
HD 190360 b	2884 ± 14	22.98 ± 0.35	476 ± 17	3.932 ± 0.013	Naef et al. (2003)
HD 190360 c	17.1163 ± 0.0013	5.78 ± 0.22	22.4 ± 1.1	$0.12882 \pm 6e - 06$	Vogt et al. (2005)
HD 190360 d	88.69 ± 0.13	1.55 ± 0.23	10.5 ± 1.6	0.38572 ± 0.00041	this work
HD 201092 b	48.973 ± 0.085	1.5 ± 0.3	5.7 ± 1.1	0.21376 ± 0.00035	this work
HD 210277 b	442.86 ± 0.22	37.96 ± 0.36	389 ± 13	1.12448 ± 0.00033	Marcy et al. (1999)
51 Peg b / HD 217014 b	$4.230785 \pm 1.7e - 05$	55.41 ± 0.27	142.5 ± 4.1	$0.05159415 \pm 1.3e - 07$	Mayor & Queloz (1995)
HD 217107 b	$7.126898 \pm 1.9e - 05$	141.88 ± 0.36	424 ± 14	$0.07247633 \pm 1.2e - 07$	Fischer et al. (1999)
HD 217107 c	4938 ± 330	52 ± 1	1287 ± 53	5.68 ± 0.27	Vogt et al. (2005)

7.5 Discussion

7.5.1 Candidates and Survey Sensitivity

Figure 7.4 and Table 7.3 show all of the planets detected by the APF-50 survey (including previously-known planets). We detect a total of 26 planets orbiting 18 stars. Nine of these planets have not been previously published.

There is a noticeable dearth of planets with masses of $\sim 30 M_{\oplus}$ and orbital periods between 120 and 1000 days. This is caused by the fact that we commonly find periodic signals in the RV time series with periods very near 1 year, 1/2 year, or 1/3 year. We attribute this to systematics in the data caused by the large motion of the stellar absorption lines on the CCD caused by Earth's barycentric motion. The largest component of Earth's barycentric motion is its orbit around the sun (± 30 km s $^{-1}$), which has a period of 1 year. This can be aliased down to 1/2 or 1/3 year due to seasonal sampling. Potential candidates near these periods are treated with more scrutiny which is not well-characterized by our injection/recovery tests. For this reason we choose to limit our occurrence rate analysis to planets with periods shorter than 100 days. This also makes for a cleaner comparison to the most recent occurrence rate results from *Kepler* (Fulton et al., 2017) where they were also limited to orbital periods shorter than 100 days.

7.5.2 Mass Distribution

Figure 7.14 shows the completeness-corrected mass distribution from the APF-50 survey. There is a clear peak in the occurrence rate of nearby, low-mass planets at about $10 M_{\oplus}$ and a fall-off in the occurrence of slightly lower mass planets. Mayor et al. (2011) saw a similar feature in the mass distribution with lower significance. Howard et al. (2010a) did not observe a peak in the mass distribution, but would not have had the mass resolution to see this feature. Planet population synthesis models also predict a peak in the mass function at around $10\text{--}20 M_{\oplus}$ (Mordasini et al., 2009b). This evidence suggests that the typical core masses for the planets discovered in the APF-50 survey is $\sim 10 M_{\oplus}$.

7.5.3 Comparison with Radius Distribution

Fulton et al. (2017) recently discovered that the distribution of planet radii as observed by *Kepler* was bimodal with peaks at $\sim 1.3 R_{\oplus}$ and $2.4 R_{\oplus}$. Owen & Wu (2017) suggest that the observed radius distribution must be caused by two populations of planets, one with core masses $\leq 3 M_{\oplus}$ that are born with very little hydrogen/helium, and a second population with core masses $\geq 3 M_{\oplus}$ that are born with atmospheres of at least a few percent by mass hydrogen and helium. A rocky core with a mass of $10 M_{\oplus}$ and no envelope would be about $1.8 R_{\oplus}$ (Lopez & Fortney, 2013) and fall right in the middle of the observed gap in the radius distribution. We suspect that the observed peak at $10 M_{\oplus}$ in the mass distribution is the source of the peak of larger planets at $2.4 R_{\oplus}$ observed by *Kepler*. These planets are massive enough to easily retain $\sim 1\%$ by mass atmospheres to be inflated to $\sim 2.1\text{--}4 R_{\oplus}$.

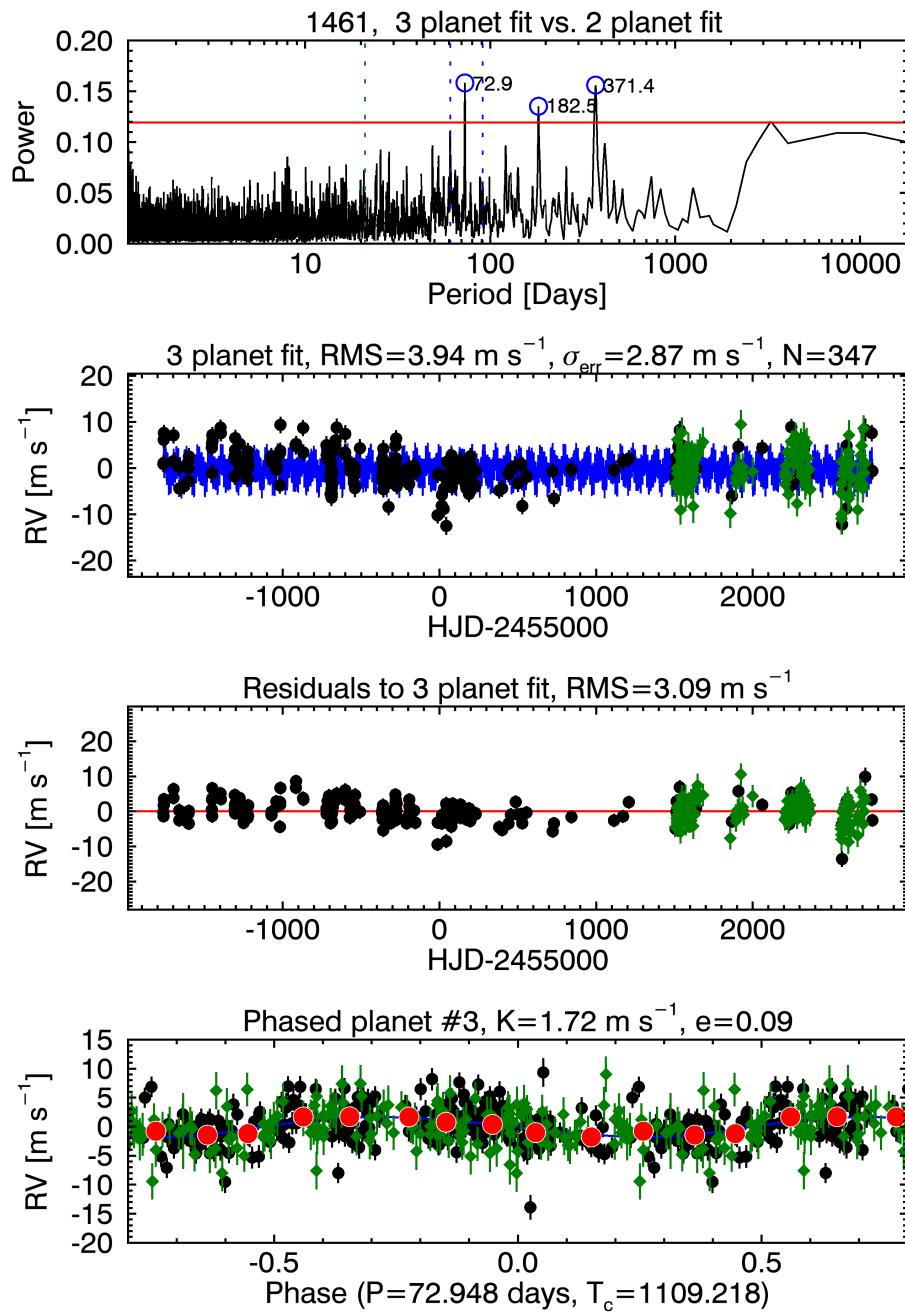


Figure 7.7 Same as Figure 7.6 for 1461 d.

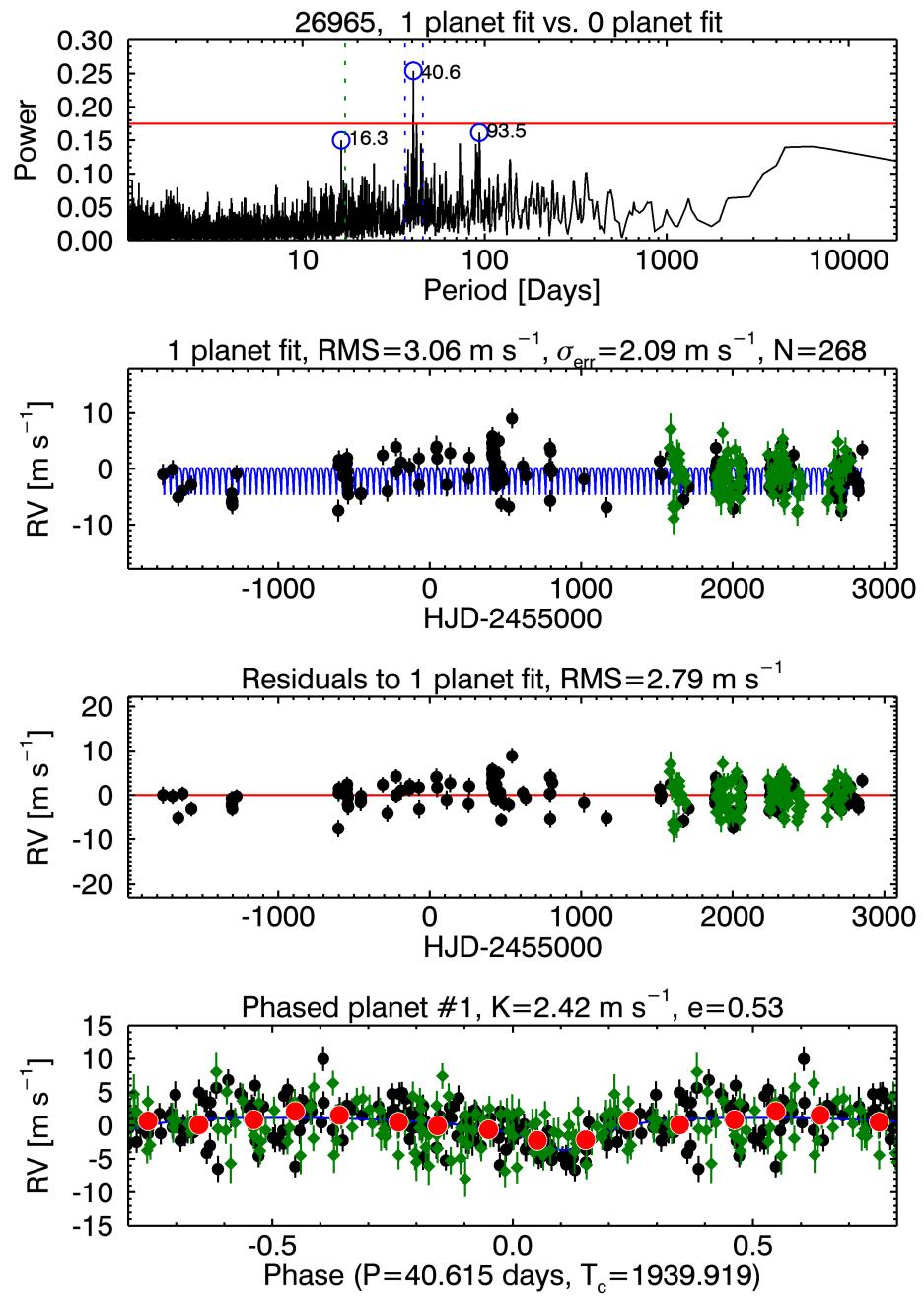


Figure 7.8 Same as Figure 7.6 for 26965 b.

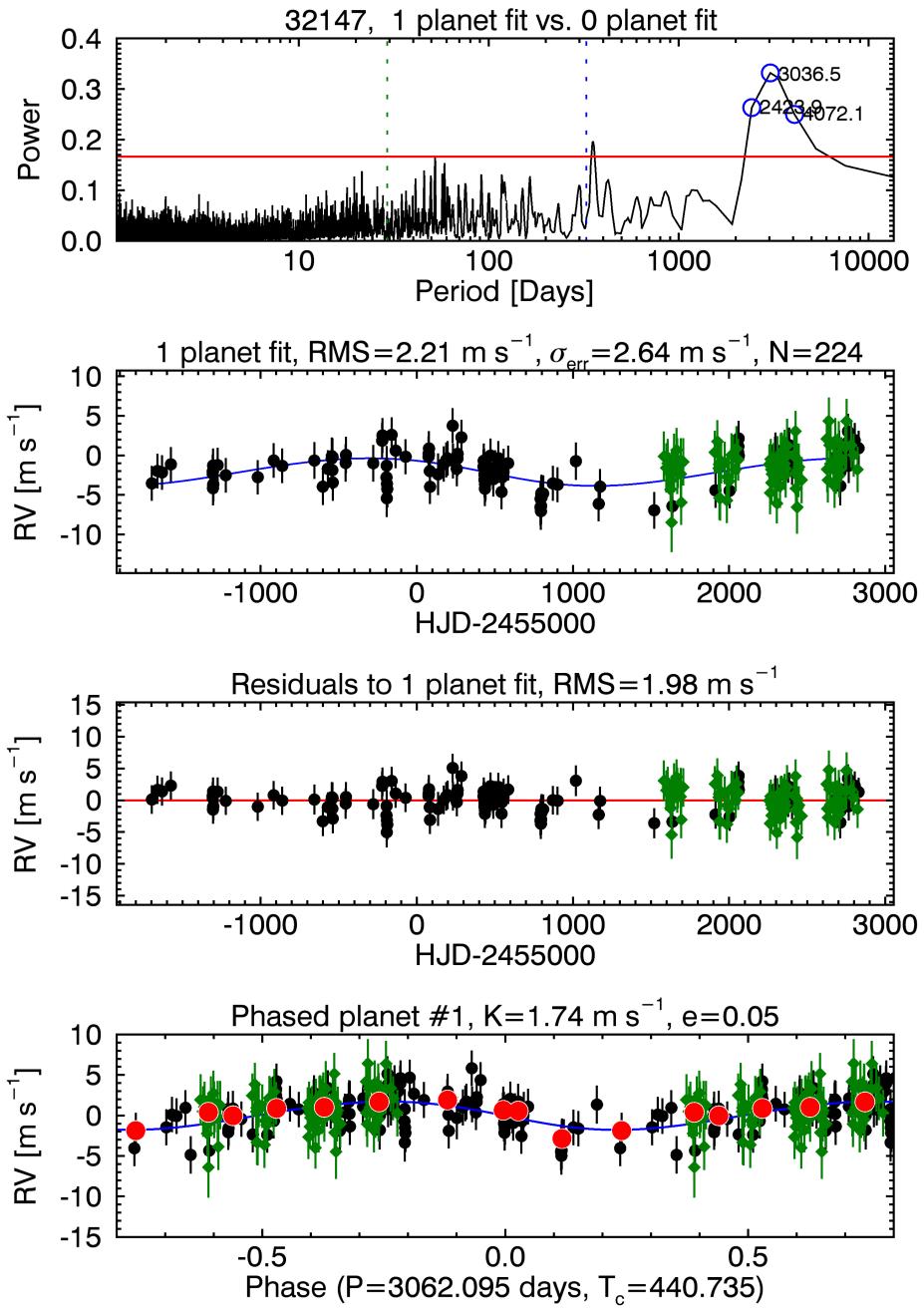


Figure 7.9 Same as Figure 7.6 for 32147 b.

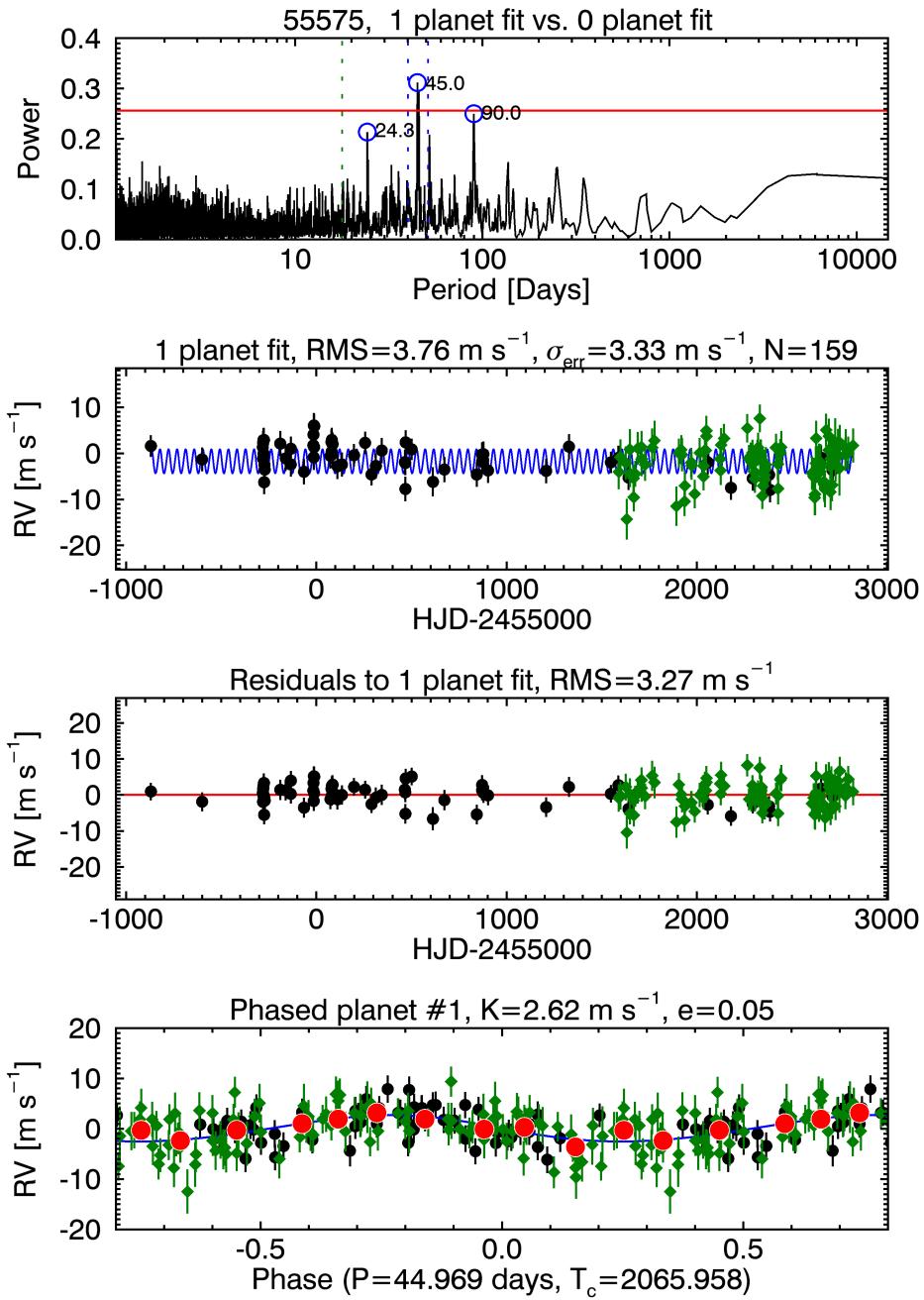


Figure 7.10 Same as Figure 7.6 for 55575 b.

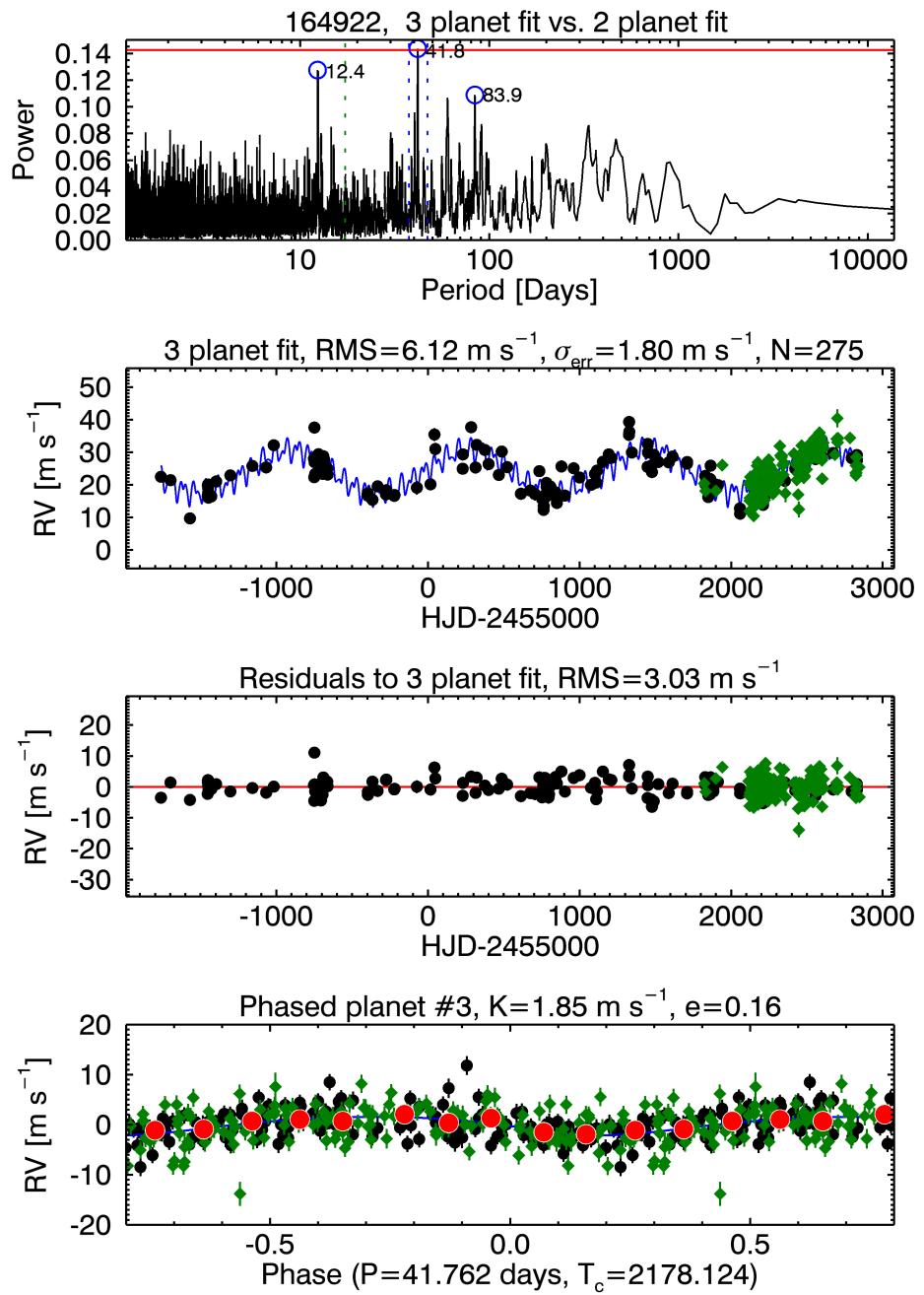


Figure 7.11 Same as Figure 7.6 for 164922 d.

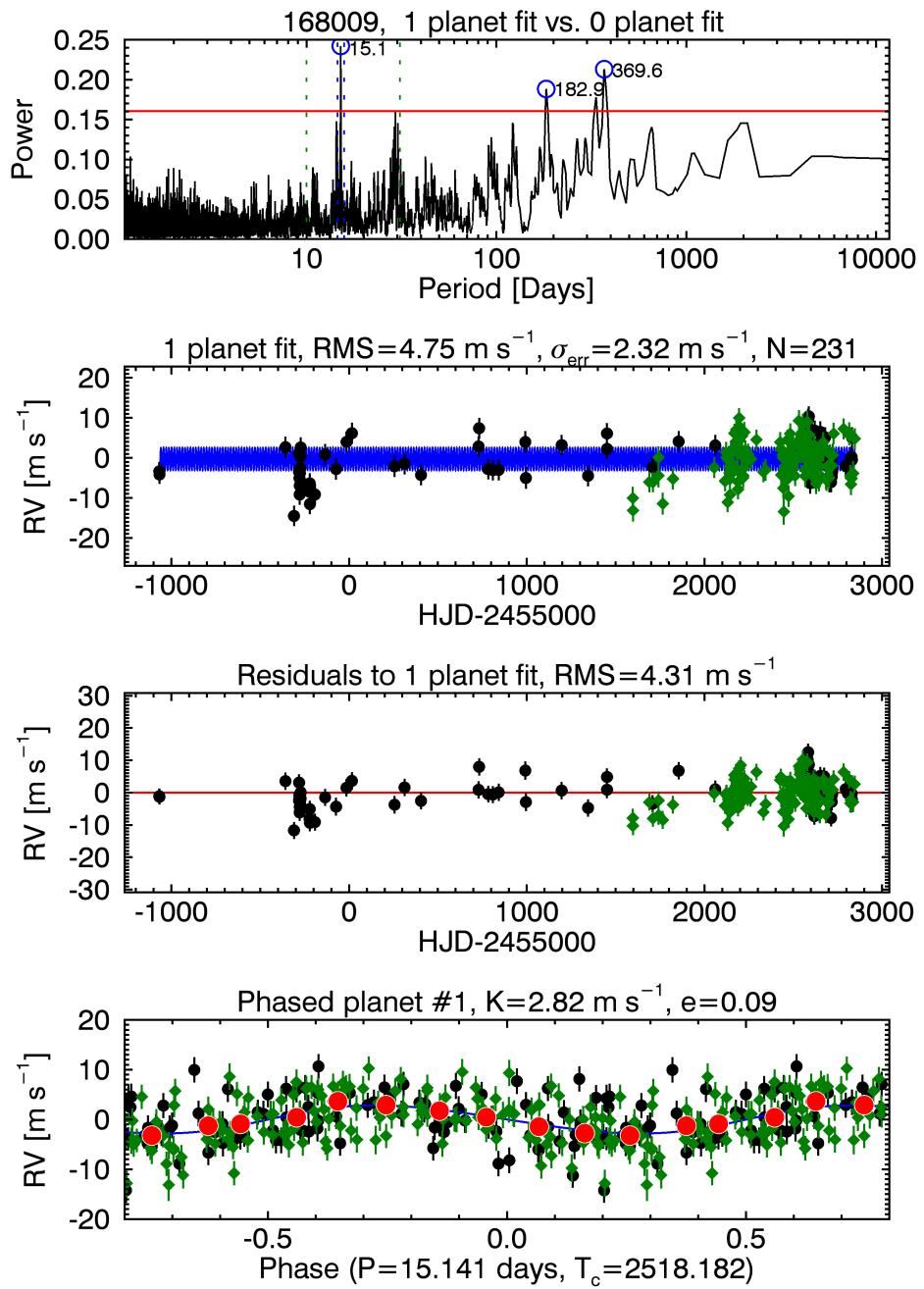


Figure 7.12 Same as Figure 7.6 for 168009 d.

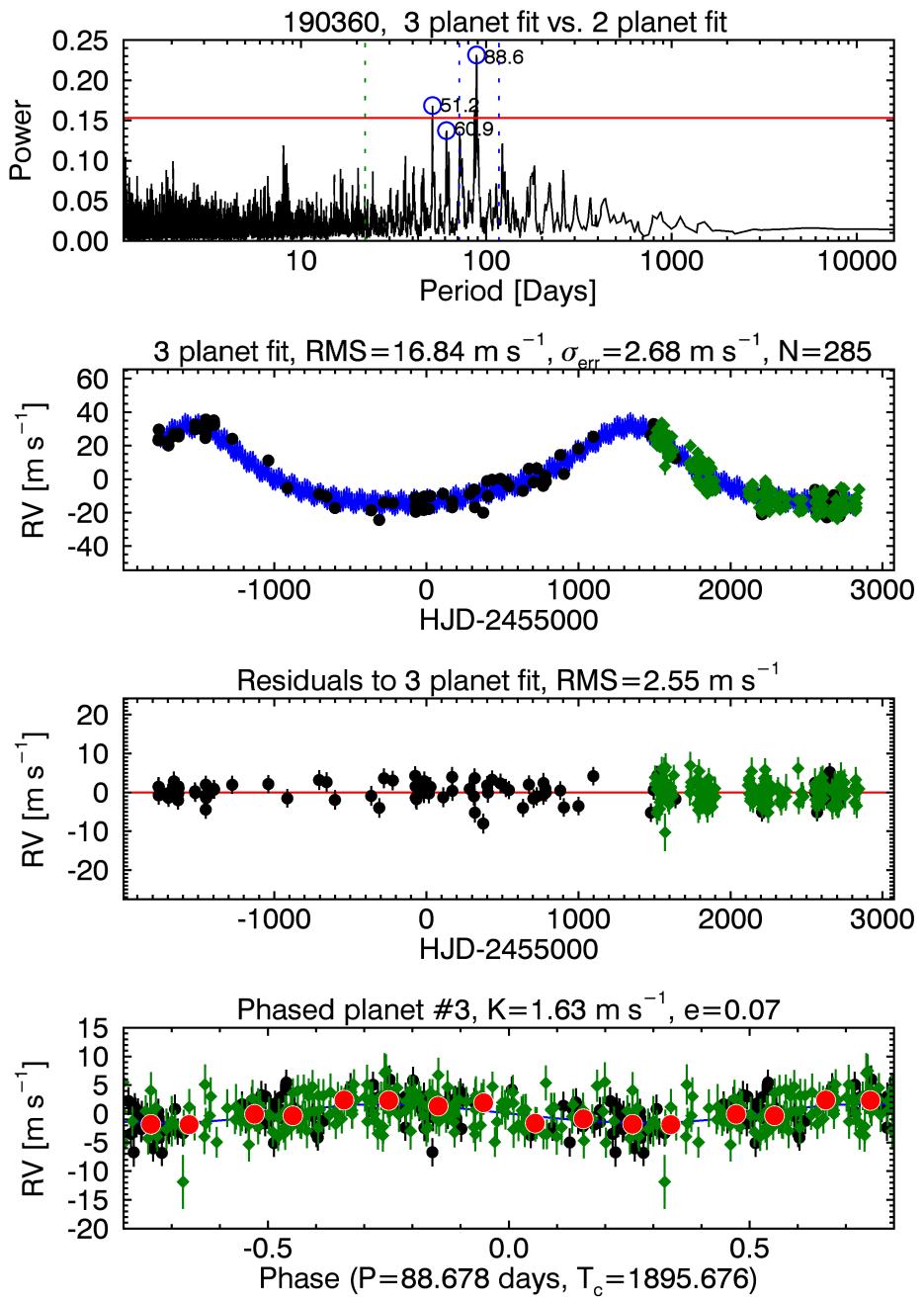


Figure 7.13 Same as Figure 7.6 for 190360 d.

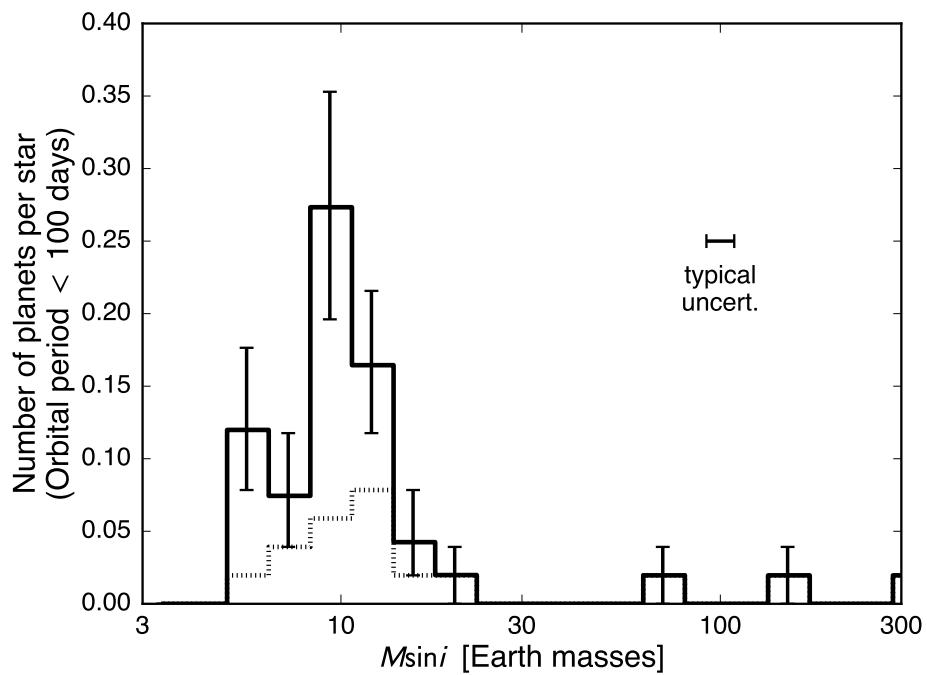


Figure 7.14 Planet occurrence as a function of planet mass in the APF-50 sample for planets with orbital periods less than 1000 days. The dotted grey line shows the actual planet detections and the solid black line is the completeness-corrected mass distribution. To within uncertainties the occurrence rate as a function of planet mass rises gradually between 30 and 7 M_{\oplus} with a possible flattening or fall-off in planet occurrence below 7 M_{\oplus} .

CHAPTER 8

A GAP IN THE RADIUS DISTRIBUTION OF SMALL PLANETS

This chapter is a reproduction of Fulton et al. (2017) included with permission from AAS journals.

8.1 Introduction

NASA’s *Kepler* space telescope enabled the discovery of over 4000 transiting planet candidates^{1,2} opened the door to detailed studies of exoplanet demographics. One of the first surprises to arise from studies of the newly revealed sample of planets was the multitude of planets with radii smaller than Neptune but larger than Earth ($R_P=1.0\text{--}3.9 R_{\oplus}$, Batalha et al., 2013). Our solar system has no example of these intermediate planets, yet they are by far the most common in the *Kepler* sample (Howard et al., 2012b; Fressin et al., 2013; Petigura et al., 2013c; Youdin, 2011; Christiansen et al., 2015; Dressing & Charbonneau, 2015; Morton & Swift, 2014).

A key early question of the *Kepler* mission was whether these sub-Neptune-size planets are predominantly rocky or possess low-density envelopes that contribute significantly to the planet’s overall size. The radial velocity (RV) follow-up effort of the *Kepler* project focused on 22 stars hosting one or more sub-Neptunes (Marcy et al., 2014). In addition, detailed modeling of transit timing variations (TTVs) provided mass constraints for a large number of systems in specific architectures (e.g., Wu & Lithwick, 2013; Hadden & Lithwick, 2014, 2016). The resulting mass measurements revealed that most planets larger than $1.6 R_{\oplus}$ have low densities that were inconsistent with purely rocky compositions, and instead required gaseous envelopes (Weiss & Marcy, 2014; Rogers, 2015).

The distinction between rocky and gaseous planets reflects the typical core sizes of planets as well as the physical mechanisms by which planets acquire (and lose) gaseous envelopes. The densities of planets with radii smaller than $\sim 1.6 R_{\oplus}$ are generally consistent with a purely rocky composition (Weiss & Marcy, 2014; Rogers, 2015) and their radius distribution likely reflects their initial core sizes. However, a small amount of H/He gas added to a roughly Earth-size rocky core can substantially increase planet size, without significantly increasing planet mass. For this reason, it has been suggested that the radii of sub-Neptune-size planets, along with knowledge of the irradiation history, would be sufficient to estimate bulk composition without additional information (Lopez & Fortney, 2013; Wolfgang & Lopez, 2015).

The large number of planets smaller than Neptune discovered by the *Kepler* mission was unexpected given prevailing theories of planet formation, which were developed to explain the distribution of giant planets (Ida & Lin, 2004; Mordasini et al., 2009a). These theories predicted

¹NASA Exoplanet Archive, 2/27/2017

²The false positive probability for the majority of the *Kepler* candidates is 5–10% (Morton & Johnson, 2011).

that planets should either fail to accrete enough material to become super-Earths, or they would grow quickly, accreting all of the gas in their feeding zones growing to massive, gas-rich giant planets. Modern formation models are now able to reproduce the observed population of super-Earths (Hansen & Murray, 2012; Mordasini et al., 2012; Alibert et al., 2013; Chiang & Laughlin, 2013; Lee et al., 2014; Chatterjee & Tan, 2014; Coleman & Nelson, 2014; Raymond & Cossou, 2014; Lee & Chiang, 2016). Many of these new models can be corroborated by measuring the bulk properties of individual planets and the typical properties of the population.

As formation models continue to be refined, the role of atmospheric erosion on these short-period planets is becoming more apparent. Several authors have predicted the existence of a “photoevaporation valley” in the distribution of planet radii (e.g., Owen & Wu, 2013; Lopez & Fortney, 2014; Jin et al., 2014; Chen & Rogers, 2016; Lopez & Rice, 2016).

Photoevaporation models predict that there should be a dearth of intermediate sub-Neptune size planets orbiting in highly irradiated environments. The mass of H/He in the envelope must be finely tuned to produce a planet in this intermediate size range. Planets with too little gas in their envelopes are stripped to bare, rocky cores by the radiation from their host stars. In general, the radii of bare, rocky cores versus planets with a few percent by mass H/He envelopes depend on many uncertain variables such as the initial core mass distribution and the insolation flux received by the planet. A rift in the distribution of small planet radii is a common result of the planet formation models that include photoevaporation.

Owen & Wu (2013) provided tentative observational evidence for such a feature in the radius distribution of *Kepler* planets. They observed a bimodal structure in the planet radius distribution, particularly when the planet sample was split into subsamples with low and high integrated X-ray exposure histories. However, the relatively large planet radius uncertainties in Owen & Wu (2013) diluted the gap and reduced its statistical significance. Their study also considered the number distribution of planets, and was not corrected for completeness as we do below. Such corrections mitigate sample bias and allow for the recovery of the underlying planet distribution from the observed one.

Here, we examine a sample of planets orbiting stars with precisely measured radii from the California-Kepler Survey (CKS; see Petigura et al. (2017) and Johnson et al. (2017)). We use the precise stellar radii to update the planet radii, bringing the distribution of planet radii into sharper focus and revealing a gap between 1.5 and $2.0 R_{\oplus}$.

This paper is structured as follows. In §8.2 we discuss our stellar and planetary samples. We describe our methods for correcting for pipeline search sensitivity and transit probabilities in §8.3. In §8.4 we examine the one-dimensional marginalized radius distribution and also two-dimensional distributions of planet radius as a function of orbital period, stellar radius, and insolation flux. We discuss potential explanations for the observed planet radius gap in §8.5 and finish with some concluding remarks in §8.6.

8.2 Sample of Planets

8.2.1 California-Kepler Survey

For this work we adopt the stellar sample and the measured stellar parameters from the CKS program (Petigura et al., 2017, hereafter Paper I). The measured values of T_{eff} , $\log g$, and [Fe/H] are based on a detailed spectroscopic characterization of 1305 *Kepler* Object of Interest (KOI) host stars using observations from Keck/HIRES (Vogt et al., 1994). In Johnson et al. (2017, hereafter Paper II), we associated those stellar parameters from Paper I to Dartmouth isochrones (Dotter et al., 2008) to derive improved stellar radii and masses, allowing us to recalculate planetary radii using the light curve parameters from Mullally et al. (2015), hereafter “Q16”. Median uncertainties in stellar radius improve from 25% (Huber et al., 2014) to 11% after our CKS spectroscopic analysis. Stellar mass uncertainties improve from 14% to 4% in the Paper II catalog. This leads to median uncertainties in planet radii of 12% which enable the detection of finer structures in the planet radius distribution.

8.2.2 Sample Selection

The CKS stellar sample was constructed to address a variety of science topics (Paper I). The core sample is a magnitude-limited set of KOIs ($K_p < 14.2$). Additional fainter stars were added to include habitable zone planets, ultra-short-period planets, and multi-planet systems. Here, we enumerate a list of cuts in parameter space designed to create a sample of planets with well-measured radii and with well-quantified detection completeness. The primary goal is to determine anew the occurrence of planets as a function of planet radius, with greater reliability than was previously possible.

We start by removing planet candidates deemed false positives in Paper I. The Paper I false positive designations were determined using the false positive probabilities calculated by Morton & Johnson (2011); Morton (2012); Morton et al. (2016), the *Kepler* team’s designation available on the NASA Exoplanet Archive, and a search for secondary lines in the HIRES spectra (Kolbl et al., 2015) as well as any other information available in the literature for individual KOIs. Next, we restrict our sample to only the magnitude-limited portion of the larger CKS sample ($K_p < 14.2$).

The planet-to-star radius ratio (R_P/R_{\star}) becomes uncertain at high impact parameters (b) due to degeneracies with limb-darkening. We excluded KOIs with $b > 0.7$ to minimize the impact of grazing geometries. We experimented other thresholds in b and found that our results are relatively insensitive to $b < 0.6$, 0.7, or 0.8, with the trade-off of smaller sample size with decreasing threshold in b .

We removed planets with orbital periods longer than 100 days in order to avoid domains of low completeness (especially for planets smaller than about $4 R_{\oplus}$) and low transit probability.

We also excised planets orbiting evolved stars since they have somewhat lower detectability

Table 8.1. Depth of the Gap

Filter	V_A
Full CKS sample	0.746
False positives removed	0.742
$K_p < 14.2$	0.686
$b < 0.7$	0.572
$P < 100$ d	0.498
Giant stars removed	0.507
$T_{\text{eff}} = 4700\text{--}6500$ K	0.483

and less certain radii. This was implemented using an *ad hoc* temperature-dependent stellar radius filter,

$$\frac{R_{\star}}{R_{\odot}} > 10^{0.00025(T_{\text{eff}}/\text{K} - 5500) + 0.20}, \quad (8.1)$$

which is plotted in Figure 8.1. We also restricted our sample to planets orbiting stars within the temperature range where we can extract precise stellar parameters from our high resolution optical spectra (6500–4700 K). Finally, we accounted for uncertainties in the completeness corrections caused by systematic and random measurement errors in the simulations, described in Appendix A.3.

The multiple filters purify the CKS sample of stars and planets and are summarized in Figure 8.2. We assessed the impact of filters on the depth of the planet radius valley using an *ad hoc* metric V_A . This quantity is defined as the ratio of the number of planets with radii of 1.64–1.97 R_{\oplus} (the bottom of the valley) to the average number of planets with radii of 1.2–1.44 R_{\oplus} or 2.16–2.62 R_{\oplus} (the peaks of the distribution immediately outside of the valley). The radius limits for the calculation of V_A were chosen so that $V_A = 1$ for a log-uniform distribution of planets with radii between 1.2 R_{\oplus} and 2.62 R_{\oplus} . Smaller values of V_A denote a deeper valley. The values of V_A after applying each successive filter are tabulated in Table 8.1.

Furlan et al. (2017) compiled a catalog of KOI host stars that were observed using a collection of high-resolution imaging facilities (Lillo-Box et al., 2012, 2014; Horch et al., 2012, 2014; Everett et al., 2015; Gilliland et al., 2015; Cartier et al., 2015; Wang et al., 2015a,b; Adams et al., 2012, 2013; Dressing et al., 2014; Law et al., 2014b; Baranec et al., 2016; Howell et al., 2011). Many of the 1902 KOIs in the Furlan et al. (2017) catalog also appear in our sample. We investigated removing KOI hosts with known companions or large dilution corrections but found no significant changes to the shape of the distribution. Since only a subset of our KOIs were observed by Furlan et al. (2017) and it is difficult to determine the binarity of the parent stellar population for occurrence calculations, we chose not to filter our planet catalog using the results of high-resolution imaging.

However, many of these stars may have already been identified as false positives in the Paper I catalog and therefore removed from our final sample of planets.

We investigated the impact of our apparent magnitude cut by examining the size distribution for three ranges of K_p (Figure 8.3). For these tests we applied all of the filters described in this section except the $K_p < 14.2$ magnitude cut. We found that the planet radius distribution for $K_p < 13.5$ is statistically indistinguishable from the radius distribution for planets orbiting stars with $13.5 < K_p \leq 14.2$. An Anderson-Darling test (Anderson & Darling, 1952; Scholz & Stephens, 1987) predicts that the two distributions were drawn from the same parent population with a p-value of 0.6. However, the radius distribution of planets orbiting host stars with $K_p \geq 14.2$ is visually and statistically different (p-value < 0.0004). This is somewhat expected given the non-systematic target selection for both the initial *Kepler* target stars and the stars observed in the CKS survey. Stars with $K_p > 14.2$ were only observed in the CKS program because they were hosts to multi-planet systems, habitable-zone candidates, ultra-short period planets, or other special cases. Targets fainter than $K_p > 14.0$ were observed by *Kepler* only if their stellar and noise properties indicated that there was a high probability of the detection of small planets (Batalha et al., 2010). These non-uniform *Kepler* target selection effects motivate our choice to exclude faint stars. The final distributions of planet radii do not depend on the $K_p < 14.2$ or $K_p < 14.0$ (p-value > 0.95) choice. But there are 153 planet candidates with $14.0 < K_p < 14.2$ so we choose to include those additional candidates to maximize the statistical power of the final sample.

The two distinct peaks separated by a valley (Figure 8.2) are apparent in the initial number distribution of planet radii and the final distribution after the filters are applied. The depth of the valley increases as we apply these filters, suggesting that the purity of the planet sample improves with filter application. Note that the filters act on the stellar characteristics and are agnostic to planet radius.

Figure 8.4 shows histograms of the stellar radii and planet-to-star radius ratios (R_P/R_\star) for the filtered sample stars. These two distributions are both unimodal. This demonstrates that the bimodality of the planet radius distribution is not an artifact of the stellar sample or the light curve fitting used to measure R_P/R_\star .

8.3 Completeness Corrections

To recover the underlying planet radius distribution from the observed distribution we made completeness corrections to compensate for decreasing detectability of planets with small radii and/or long orbital periods.

An additional complication associated with the completeness corrections in this work is that the stellar properties of the planet-hosting stars come from a different source and have higher precision than the stellar properties for the full set of *Kepler* target stars. We explore the additional uncertainties introduced by this fact by running a suite of simulated transit surveys described in

Appendix A.3. We inflate the uncertainties on the histogram bin heights by the scaling factors listed in Table A.1 to account for these effects.

8.3.1 Pipeline Efficiency

We followed the procedure described in Christiansen et al. (2016) using the results from their injection-recovery experiments (Christiansen et al., 2015). They injected about ten-thousand transit signals into the raw pixel data and processed the results with version 9.1 of the official *Kepler* pipeline (Jenkins et al., 2010). These completeness tests were used to identify combinations of transit light curve parameters that could be recovered by the *Kepler* pipeline for a given sample of target stars. They injected signals onto both target stars and neighboring pixels to quantify the pipeline’s ability to identify astrophysical false positives. We assumed that our sample is free of the vast majority of false positives so we only considered injections of transits onto the target stars. We only considered injections on stars that would have been included in the CKS sample and would not be removed by the filters described in §8.2.2. Namely, we considered injected impact parameters less than 0.7, injected periods shorter than 100 days, $Kp \leq 14.2$, $4700 \text{ K} < T_{\text{eff}} < 6500 \text{ K}$, and stellar radii compatible with Equation 8.1 based on the values in the Stellar17 catalog³ prepared by the *Kepler* stellar parameters working group (Mathur et al., 2016). This leaves a total of 3840 synthetic transit signals injected onto the target pixels of 3840 stars observed by *Kepler*. We also apply these same filters to the stars in the Stellar17 catalog. The number of stars remaining after the filters are applied is the number of stars observed by *Kepler* that could have led to detections of planets that would be present in our filtered planet catalog ($N_{\star} = 36,075$). We calculated the fraction of injected signals recovered as a function of injected signal-to-noise as

$$m_i = \left(\frac{R_P}{R_{\star,i}} \right)^2 \sqrt{\frac{T_{\text{obs},i}}{P}} \left(\frac{1}{\text{CDPP}_{\text{dur},i}} \right), \quad (8.2)$$

where R_P and P are the radius and period of the particular injected planet. $R_{\star,i}$ is the stellar radius for the i^{th} star in the Stellar17 catalog, $T_{\text{obs},i}$ is the amount of time that the particular star was observed, and $\text{CDPP}_{\text{dur},i}$ is the Combined Differential Photometric Precision (CDDP, Koch et al., 2010) value for each star extrapolated to the transit duration for each injection. We fit a 2nd order polynomial in $1/\sqrt{d}$ to the $d = 3, 6$, and 12-hour CDPP values for each star to perform the extrapolation (Sinukoff et al., 2013).

We fit a Γ cumulative distribution function (CDF) to the recovery fraction vs. injected (m_i) of the form

$$C(m_i; k, \theta, l) = \Gamma(k) \int_0^{\frac{m_i - l}{\theta}} t^{k-1} e^{-t} dt, \quad (8.3)$$

to derive the average pipeline efficiency. $C(m_i)$ is the probability that a signal with a given value

³<https://archive.stsci.edu/kepler/stellar17/search.php>

Table 8.2. Planet Detection Statistics

Planet candidate	P d	R_P R_\oplus	SNR m_i	Detection probability p_{det}	Transit probability p_{det}	Weight $1/w_i$
K00002.01	2.20	13.41	750.22	1.00	0.14	6.94
K00003.01	4.89	5.11	877.10	1.00	0.05	20.14
K00007.01	3.21	4.13	146.38	1.00	0.11	8.88
K00010.01	3.52	13.39	914.62	1.00	0.09	11.06
K00017.01	3.23	15.04	1212.38	1.00	0.11	9.40
K00018.01	3.55	13.94	820.96	1.00	0.10	9.58
K00020.01	4.44	21.41	1469.42	1.00	0.10	10.15
K00022.01	7.89	14.20	1085.97	1.00	0.06	17.98
K00041.01	12.82	2.37	37.15	0.98	0.05	22.37
K00041.02	6.89	1.35	15.04	0.91	0.07	15.98

Note. — Table 8.2 is available in its entirety in machine-readable format, which also includes period and radius uncertainties. A portion is shown here for guidance regarding its form and content. Refer to Paper II for the CKS stellar parameters associated with each KOI. This table contains only the subset of planet detections that passed the filters described in §8.2.2. The full sample of planet candidates orbiting CKS target stars can be found in Paper II.

of m_i would actually be detected by the *Kepler* transit search pipeline. In practice we used the `scipy.stats.gammacdf(t, k, l, theta)` function in SciPy version 0.18.1. Using the `lmfit` Python package (Newville et al., 2014) to minimize the residuals we found best-fit values of $k = 17.56$, $l = 1.00$ (fixed), and $\theta = 0.49$. Figure 8.5 shows the fraction of injections recovered as a function of m_i and our model for pipeline efficiency.

Our pipeline efficiency curve is \sim 15–25% lower than the efficiency as a function of the *Kepler* multi-event statistic (MES) derived in (Christiansen et al., 2015) for their FGK subsample. The difference can be explained by the fact that the MES is estimated in the *Kepler* pipeline during a multidimensional grid search. In most cases, the search grid is not fine enough to find the exact period and transit time for a given planet candidate. Since the grid search doesn’t find the best-fit transit model it generally underestimates the SNR (m_i) by a factor of \sim 25% (Petigura et al., in preparation).

8.3.2 Survey Sensitivity

For each planet detection there are a number of similar planets that would not have been detected due to a lack of sensitivity or unfavorable geometric transit probability. To compensate, we weighted each planet detection by the inverse of these probabilities,

$$w_i = \frac{1}{(p_{\text{det}} \cdot p_{\text{tr}})}, \quad (8.4)$$

where p_{det} is the fraction of stars in our sample where a transiting planet with a given signal to noise ratio given by Equation 8.2 could be detected:

$$p_{\text{det}} = \frac{1}{N_\star} \sum_i^{N_\star} C(m_i). \quad (8.5)$$

The geometric transit probability is $p_{\text{tr}} = 0.7R_\star/a$. The factor of 0.7 compensates for our omission of planet detections with $b > 0.7$ from the planet catalog. Figure 8.6 shows the mean pipeline completeness (p_{det}) and mean total search completeness ($1/w_i$) as a function of planet radius and orbital period for the filtered Stellar17 sample of *Kepler* target stars. The detection probabilities, transit probabilities, and weights (w_i) for each planet in our final catalog are listed in Table 8.2.

8.3.3 Occurrence Calculation

Following the definitions in Petigura et al. (2013b), the average planet occurrence rate (number of planets per star) for any discrete bin in planet radius or orbital period is the sum of these weights divided by the total number of stars in the sample (N_\star):

$$f_{\text{bin}} = \frac{1}{N_\star} \sum_{i=1}^{n_{\text{pl,bin}}} w_i. \quad (8.6)$$

Again, $N_\star = 36,075$ is the total number of dwarf stars in the Stellar17 catalog that pass the same filters on stellar parameters that were applied to the planet catalog: no giant stars (selected using Equation 8.1), $4700 \text{ K} < T_{\text{eff}} < 6500 \text{ K}$, and $K_p \leq 14.2$.

8.4 The Planet Radius Gap

Figure 8.7 shows the completeness-corrected distribution of planet radii for the filtered sample of 900 planets and the corresponding occurrence values are tabulated in Table 8.3. Uncertainties on the bin heights are calculated using Poisson statistics on the number of detections within the bin, scaled by the size of the completeness correction in each bin, and scaled again by a correction factor determined from a collection of simulated transit surveys as described in Section A.3. The completeness corrections are generally small. We are sensitive to $> 80\%$ of $2.0 R_\oplus$ planets out to orbital periods of 100 days, and $> 50\%$ of $1.0 R_\oplus$ planets out to 30 days (Figure 8.6). The transit probability term in Equation 8.4 dominates the corrections in most of the parameter space explored. Somewhat surprisingly, the larger, sub-Neptunes receive a completeness boost that is larger than the boost received by the smaller, super-Earths (compare the dotted grey line in Figure 8.7 to the solid black line) because the sub-Neptunes tend to orbit at larger orbital distances where transit

Table 8.3. Planet Occurrence

Radius bin R_{\oplus}	Number of planets per star f_{bin} for $P < 100$ d
1.16–1.29	0.078 ± 0.017
1.29–1.43	0.08 ± 0.013
1.43–1.59	0.053 ± 0.011
1.59–1.77	0.0334 ± 0.0092
1.77–1.97	0.05 ± 0.01
1.97–2.19	0.086 ± 0.016
2.19–2.43	0.098 ± 0.016
2.43–2.70	0.077 ± 0.016
2.70–3.00	0.053 ± 0.012
3.00–3.33	0.0316 ± 0.0089
3.33–3.70	0.0242 ± 0.0066
3.70–4.12	0.0094 ± 0.0057
4.12–4.57	0.0056 ± 0.0034
4.57–5.08	0.0037 ± 0.0031
5.08–5.65	0.0066 ± 0.0048
5.65–6.27	0.005 ± 0.003
6.27–6.97	$0.0 \pm \text{inf}$
6.97–7.75	0.0019 ± 0.0029
7.75–8.61	0.0044 ± 0.0034
8.61–9.56	0.00022 ± 0.00032
9.56–10.63	0.001 ± 0.0015
10.63–11.81	0.00035 ± 0.00053
11.81–13.12	0.00104 ± 0.00094
13.12–14.58	0.0038 ± 0.0021
14.58–16.20	0.00084 ± 0.00066
16.20–18.00	0.0003 ± 0.0004

probabilities are smaller. The mean transit probability (p_{tr}) for planets with radii of 1.0 – $1.75 R_{\oplus}$ in our sample is 6% while the transit probability for planets with radii of 1.75 – $3.5 R_{\oplus}$ is a factor of two lower (3%). However, the mean detectability (p_{det}) for those same two classes of planets are both very high at 86% and 96% respectively.

8.4.1 Comparison with Log-Uniform Distribution

We performed several tests to quantify the significance of the gap in the planet radius distribution. First, we performed a two-sided Kolmogorov-Smirnov (K-S, Kolmogorov, 1933; Smirnov, 1948) test to assess the probability that the planet radius number distribution for radii in the range 1 – $3 R_{\oplus}$ is drawn from a log-uniform distribution. This test returns a probability of 0.003 that the planet radii between 1 – $3 R_{\oplus}$ are drawn from a log-uniform distribution. However, we note that blind interpretation of p-values from K-S tests can often lead to overestimates of significance (Babu & Feigelson, 2006). Similarly, an Anderson-Darling test also rejects the hypothesis that the planet

Table 8.4. Spline Fit

Node Location R_{\oplus}	Best-fit Value (f_{bin})	1 σ Credible Interval (f_{bin})
1.3	0.078	fixed
1.5	0.051	0.05 ± 0.02
1.9	0.030	0.03 ± 0.02
2.4	0.116	0.11 ± 0.01
3.0	0.043	0.044 ± 0.005
4.5	0.0050	0.005 ± 0.002
11.0	0.00050	0.0005 ± 0.0003

radii between $1\text{--}3 R_{\oplus}$ were drawn from a log-uniform distribution with a p-value of 0.012.

8.4.2 Dip Test of Multimodality

Hartigan's dip test is a statistical tool used to estimate the probability that a sample was drawn from a unimodal distribution or a multi-modal distribution with ≥ 2 modes (Hartigan & Hartigan, 1985). It is similar to the K-S statistic in that it measures the maximum distance between an empirical distribution and a unimodal distribution. Applying this test to the number distribution of $\log R_P$ for planet radii in the range $1\text{--}3 R_{\oplus}$ returns a p-value of 1.4×10^{-3} that the distribution was drawn from a unimodal distribution. This strongly suggests that the planet radius distribution is multi-modal.

8.4.3 Spline Model

Modeling the planet radius distribution with splines having nodes at fixed values gives a good fit for a range of planet sizes. Virtues of this model are the small number of free parameters and model flexibility, particularly in asymptotic regions where others models (e.g. Gaussians) force the distribution to zero. We fit a second-order spline with seven node points fixed at specific radii to the weighted histogram of planet occurrence. We excluded from the fit bins for radii smaller than $1.14 R_{\oplus}$ where the pipeline completeness at $P = 100$ days is less than 25%. The model was adjusted by varying the amplitudes of the spline nodes, then convolving with a Gaussian kernel whose width is the median fractional planet radius uncertainty (12%). The convolved model is averaged over each of the histogram bins before performing the χ^2 comparison. This allows us to separate the smearing of the observed distribution due to measurement uncertainties from a “deconvolved” view of the underlying distribution. Again we found the best-fit solution using the `lmfit` package to minimize the normalized residuals of the histogram bins relative to the convolved model. We used

the `emcee` (Foreman-Mackey et al., 2013) interface built into `lmfit` to estimate the uncertainties on the node values. We performed the fits working in $\log(R_P)$ and required positive occurrence values for the deconvolved model. For radii outside of the range spanned by our node locations, we extrapolated assuming constant (log-uniform) occurrence.

Deconvolution is an inherently unstable process and we caution against over-interpretation of the deconvolved model. Our best-fit deconvolved model is not the only solution that could produce an equivalent convolved model. The deconvolved model is also somewhat sensitive to the choice of the node locations, while the convolved model is insensitive to those choices. However, the deconvolved model suggests that the gap is likely deeper than observed. This motivates detailed follow-up and characterization of the planets that fall within the gap. The best-fit model (red line) and deconvolved model (dashed cyan line) are both over-plotted on the completeness-corrected planet radius distribution in Figure 8.7. Table 8.4 lists the locations, best-fit values, and 1σ credible intervals for the spline nodes.

8.4.4 Relative Frequency of Super-Earths and Sub-Neptunes

Many authors use the terms “super-Earth” and “sub-Neptune” interchangeably, or draw arbitrary distinctions in mass or radius between these two classes. The observed gap in the radius distribution of small planets suggests a less arbitrary division. In the text below we define a “super-Earth” as a planet with a radius of $1\text{--}1.75 R_\oplus$, and a “sub-Neptune” as having a radius of $1.75\text{--}3.5 R_\oplus$.

We calculated the occurrence ratio of super-Earths to sub-Neptunes to be 0.8 ± 0.2 . The uncertainty is determined using a suite of simulated surveys described in Appendix A.3. The nearly equal occurrence of super-Earths and sub-Neptunes with $P < 100$ days provides an important constraint for planet formation models. This is likely a lower limit on this ratio since the super-Earth domain likely extends to sizes smaller than $1.1 R_\oplus$.

8.4.5 Two-Dimensional Weighted Kernel Density Estimation

In the following subsections we present and discuss several contour plots. The contours were derived using the Weighted Kernel Density Estimation (wKDE) technique described in Appendix A.2 and have all been corrected for completeness (with the exception of Figure 8.9). We calculated bivariate Gaussians for each pair of planet parameters over a fixed high-resolution grid in the two parameters, sum these Gaussians over all planets, and divide by the total number of stars in the sample ($N_\star=36,075$). Each bi-variate Gaussian is normalized to have a maximum value of 1.0, then multiplied by the weight associated with the given planet detection (w_i , Equation 8.4). The points plotted are the CKS parameters.

Planet Radius vs. Orbital Period

We first look at the distribution of planet radii as a function of orbital period (P). Figure 8.8 shows the distribution of planet radii as a function of orbital period for planet and stellar parameters from the Q16 catalog (*top panel*). It also offers a comparison with the same distribution derived from the CKS parameters (*bottom panel*).

There is a declining number of small planet detections going toward long orbital periods. However the underlying completeness-corrected contours suggest that the occurrence rate of these planets does not fall off with the number of detections. Instead, the lack of detections is likely an artifact of decreasing transit detectability and probability.

Figure 8.8 shows that small planets are significantly more common than large planets. The fact that planets smaller than Neptune ($4 R_{\oplus}$) are much more common than Jovian-size planets has been well documented in the literature (e.g. Howard et al. (2010a); Mayor et al. (2011); Howard et al. (2012b); Fressin et al. (2013); Dong & Zhu (2013); Petigura et al. (2013b); Dressing & Charbonneau (2015); Burke et al. (2015)). However, the increase in occurrence with decreasing planet size is evidently more rapid than was apparent in previous studies.

There is another feature in the R_P vs. P occurrence distribution that motivates a closer examination of the planet radius distribution along other axes. There are very few planets larger than $2 R_{\oplus}$ with orbital periods shorter than about 10 days while planets with radii smaller than $1.8 R_{\oplus}$ remain quite common down to orbital periods of about 3 days. A sharp decline in the occurrence rate of planets larger than approximately $1.6 R_{\oplus}$ with orbital periods shorter than 10 days has been previously observed (Howard et al., 2012b; Dong & Zhu, 2013; Sanchis-Ojeda et al., 2014).

Planet Radius vs. Stellar Radius

Figure 8.9 shows the distribution of planet size as a function of host star size. This distribution shows two distinct populations of planets with a gap separating them. Planets appear to preferentially fall into two classes, one with radii of $\sim 2.4 R_{\oplus}$ and another with radii of $\sim 1.3 R_{\oplus}$. Planets with intermediate radii of $1.5\text{--}2.0 R_{\oplus}$ are comparatively rare. The gap occurs at the same planet radius for all stellar sizes in our sample. The bimodal planet size distribution holds true for planets orbiting stars with radii ranging from $0.7 R_{\odot}$ to $2.0 R_{\odot}$.

Planet Radius vs. Incident Flux

Figure 8.10 shows the planet radius distribution as a function of incident flux. The two planet populations shear apart in this domain. There is a dearth of sub-Neptunes orbiting in high incident flux environments. This trend is also visible in one-dimensional histograms of planet radii when broken up into groups based on S_{inc} (Figure 8.11). Most of the planets that contribute to the peak

in the marginalized radius distribution at $1.3 R_{\oplus}$ are orbiting in environments with $S_{\text{inc}} > 200 S_{\oplus}$, while the planets that contribute to the peak at $2.4 R_{\oplus}$ experience $S_{\text{inc}} < 80 S_{\oplus}$. It is clear that the gap is present even at low incident fluxes and the two-dimensional S_{inc} and period distributions show a potential deepening and/or widening of the gap toward lower incident fluxes. However, we can not determine if the gap radius is dependent on incident flux, or if the break radius is constant as a function of incident flux due to lack of completeness for small planets orbiting in cool environments.

There is also an upper envelope of planet size which decreases as a function of incident flux. Although there are a few exceptions, there is a clear dearth of planets in the upper left quadrant of Figure 8.10. These should be some of the easiest planets to detect yet they do not appear in our sample of planets. This feature has been previously observed (e.g., Howard et al., 2012b; Mazeh et al., 2016; Lundkvist et al., 2016) but our larger sample of planets with high-precision host star properties sharpens the boundary. The lack of planet detections in the lower right region of Figure 8.10 is the result of low survey completeness for small, long-period planets.

8.5 Discussion

We have provided observational evidence that the distribution of planet sizes is not smooth (Figure 8.7). Small planets have characteristic sizes of $\sim 1.3 R_{\oplus}$ (super-Earths) and $\sim 2.4 R_{\oplus}$ (sub-Neptunes). These two planet populations each have intrinsic widths in their size distributions, but there is a gap that separates them. Intermediate-size planets with radii of $\sim 1.5\text{--}2.0 R_{\oplus}$ are comparatively rare.

8.5.1 Previous Studies of the Radius Distribution

Many studies have examined the planet radius distribution using the *Kepler* sample. To date, none have shown statistically significant evidence for a gap in the distribution at $1.5\text{--}2.0 R_{\oplus}$.

The pioneering study of Owen & Wu (2013) pointed out a marginally-significant gap at $\sim 1.5\text{--}2 R_{\oplus}$ in the observed radius distribution and interpreted it as connected to the high-energy irradiation history of the planets. They did not have a large set of accurate planet radii and they did not perform the completeness corrections necessary to confirm the feature. Here, we firmly detect a gap in the planet radius distribution between two peaks at $2.4 R_{\oplus}$ and $\leq 1.3 R_{\oplus}$.

Based on the initial *Kepler* planet catalog, Howard et al. (2012b) investigated the domain of planets with $R_P > 2 R_{\oplus}$ and $P < 50$ days. They demonstrated that small planets are common. However, they did not examine the detailed shape of the small planet occurrence function, due to the severe lack of completeness to small planets with the early *Kepler* data releases, and large uncertainties in the planetary radii. At that time, the planetary radii were based on the relatively coarse estimates of the stellar radii from the KIC.

Follow-on studies (Youdin, 2011; Catanzarite & Shao, 2011; Traub, 2012) were similarly limited. Dong & Zhu (2013) benefited from a larger dataset. They focused on the orbital period distribution, with large (factor of two) bins in planet radius. Petigura et al. (2013b) utilized a much longer photometric time series (lasting 15 of 17 *Kepler* quarters), and a custom planet detection pipeline enabling completeness corrections, but the sample was only large enough to allow for three bins in the radius range 1.0–2.8 R_{\oplus} . Silburt et al. (2015) measured occurrence for planets with radii between 1.0 and 4.0 R_{\oplus} and orbital periods between 20 and 200 days. They found a peak in the distribution near 2.4 R_{\oplus} and a slight decline in the frequency of smaller planets. More recently, Burke et al. (2015) studied the occurrence of small, long-period planets. With 1σ significance, they observed a diminution in planet occurrence in the 1.5–2.0 R_{\oplus} interval for planets having $P = 300$ –700 days.

Occurrence Rate Comparisons

Table 5 compares the occurrence rates measured in this work to those of several touchstone studies from the literature: Howard et al. (2012b, H12), Petigura et al. (2013c, P13), Fressin et al. (2013, F13), and Mulders et al. (2015, M15). These works all analyzed Kepler planets, but used catalogs constructed from different amounts of Kepler photometry. In addition, these studies applied different treatments of pipeline completeness, adopted different false positive rates, analyzed different sub-samples of Kepler stars, and accounted for multi-planet systems in different ways. All of these differences can significantly affect the derived occurrence (Burke et al., 2015). However, the relative occurrence rates between bins are insensitive to most of these issues and potential discrepancies in the absolute occurrence rates do not affect the presence or shape of the gap in the radius distribution.

We choose to closely compare our occurrence values in this work to those of P13, because they used a nearly complete photometric dataset (43/48 months)⁴ and corrected for pipeline completeness through direct injection and recovery. Our occurrence rates are typically 50% higher than those of P13. However, P13 (and H12) measured the fraction of stars with planets as opposed to the number of stars per planet measured in this work (and in F13 and M15). The number of planets per star will always be larger than the fraction of stars with planets due to multi-planet systems. P13 estimated that their occurrence rates would have been 25–45% higher if they had included multi-planet systems (depending on period and radius limits), which can reconcile much of the differences between the two studies.

In comparing to previous results, we find that 2–2.8 R_{\oplus} are more common than 1.4–2 R_{\oplus} planets, in a relative sense. For example, we find that P13 found that 18.6% of stars had a 2–2.8 R_{\oplus} planet with $P < 100$ d vs. 14.2% of stars with a 1.4–2 R_{\oplus} planet in the same period range. This corresponds to a ratio of $18.6\%/14.2\% = 1.3$. In this work, that ratio is $16.1\% / 27.0\% = 0.6$. We

⁴H12, F13, P13, and M15 used 4, 16, 43, and 22 months of photometry, respectively.

Table 8.5. Occurrence Rate Comparison

Radius Interval R_{\oplus}	Period Interval (days)	This Work ¹ (f_{bin} %)	H12 ^{2,6,7} (f_{bin} %)	P13 ^{3,6} (f_{bin} %)	F13 ⁴ (f_{bin} %)	M15 ⁵ (f_{bin} %)
1.4–2.8	< 100	43.1 ± 2.2	...	32.8 ± 1.4	35.0 ± 2.8^8	26.7 ± 1.7^8
2–2.8	< 50	19.4 ± 1.4	9.0 ± 1.5	18.6 ± 1.6	17.5 ± 1.6	12.8 ± 0.5
2–4	< 50	25.4 ± 1.6	13.0 ± 0.8	16.6 ± 1.8	18.3 ± 1.3	18.6 ± 0.6
2–4	< 100	36.6 ± 2.2	...	24.1 ± 2.3	24.0 ± 2.2^8	22.9 ± 0.8^8

¹Uncertainties do not include the scaling factors derived in Appendix A.3

²Howard et al. (2012b)

³Petigura et al. (2013c)

⁴Fressin et al. (2013)

⁵Mulders et al. (2015)

⁶Measured fraction of stars with planets instead of number of planets per star

⁷Only studied planets with periods shorter than 50 days and larger than $2 R_{\oplus}$

⁸Periods shorter than 85 days

Note. — Each occurrence rate study focused on different stellar samples, planet detection pipelines, period limits, etc. This table is not meant to be an exact comparison of the results from each study, but instead a rough comparison to show general agreement or highlight large disagreements.

can understand this difference in terms of the gap between 1.5 and 2.0 R_{\oplus} and the peak between 2.0 and 2.4 R_{\oplus} that emerged after we refined the host star radii through spectroscopy. Planets with true sizes between 2.0 and 2.8 R_{\oplus} were often scattered to the 1.4–2.0 R_{\oplus} bin due to the 40% radius uncertainties from photometry. Thus the peak from 2.0–2.8 R_{\oplus} was diminished, while the gap from 1.4–2.0 R_{\oplus} was filled in. In summary, the integrated occurrence rates presented are largely consistent with previous works, with differences in the detailed radius distribution, owing to improved stellar radii.

8.5.2 Rocky to Gaseous Transition

Studies of the relationship between planet density and radius suggest that planet core sizes reach a maximum of about $1.6 R_{\oplus}$. Planets with larger radii and measured masses are mostly low-density and require an extended atmosphere to simultaneously explain their masses and radii (Marcy et al., 2014; Weiss & Marcy, 2014; Rogers, 2015; Wolfgang & Lopez, 2015). Figure 8.12 shows the radius distribution derived in this work and an empirical fit to the densities and radii of small planets (Weiss et al., 2016). This fit to a sample of planets with measured densities peaks near our observed gap in the planet radius distribution. This suggests that the majority of planets smaller than the

minimum in the occurrence distribution are rocky while larger planets likely contain enough volatiles to contribute significantly to the planets' radii.

Additionally, ultra-short-period planets (USPs, having $P < 1$ day) present a clean sample of stripped, rocky planet cores. It is unlikely that H/He atmospheres could survive on small planets bathed in the intense irradiation experienced by USPs. These planets must be bare, rocky cores, stripped of any significant atmosphere. Sanchis-Ojeda et al. (2014) found that the occurrence of ultra-short period planets falls off sharply for $R_P > 1.6 R_{\oplus}$. The apparent lack of rocky cores larger than $1.6 R_{\oplus}$ also suggests that planets larger than that must have non-negligible volatile envelopes.

8.5.3 Potential Explanations for the Gap

Photoevaporation

Photoevaporation provides a possible mechanism to produce a gap in the radius distribution, even if the initial radius distribution was continuous (Owen & Wu, 2013). Lopez & Rice (2016) modeled the masses and radii of planets with various gas envelope fractions. A bare, rocky planet (no envelope) with a mass of $2 M_{\oplus}$ has a radius of $1.2 R_{\oplus}$ in their models. Adding an H/He envelope with a mass of $0.002 M_{\oplus}$ (0.1% mass fraction) increases the planet size to $1.5 R_{\oplus}$, a large change in size for a small change in mass. Adding an additional 0.7% by mass of H/He swells the planet to $2.0 R_{\oplus}$ (see Figure 8.13). This non-linear mass-radius dependence on volatile fraction has two effects. First, making a planet with a thin atmosphere requires a finely tuned amount of H/He. Second, photoevaporating a planet's envelope significantly changes its size. Our observation of two peaks in the planet size distribution is consistent with super-Earths being rocky planets with atmospheres that contribute negligibly to their size, while sub-Neptunes are planets that retain envelopes with mass fractions of a few percent.

Gas-poor formation

Accretion of a modest gas envelope poses a theoretical challenge because fine-tuning is required to end up with an appreciable atmosphere that does not trigger runaway gas accretion and giant planet formation. Lee et al. (2014) proposed a mechanism that produces small planets with low envelope fractions by delaying gas accretion until the gas in the protoplanetary disk is nearly dissipated. They also proposed that small planets could form in very metal-rich disks where high opacity slows cooling and accretion.

In addition, a few-percent-by-mass secondary atmosphere can be outgassed during planet formation and evolution (Adams et al., 2008). Our observed gap in the planet radius distribution could be explained by a mechanism that causes the creation of a secondary atmosphere during the formation of only $\sim 50\%$ of terrestrial planets.

Impact Erosion

Impacts can also provide a way to sculpt the atmospheric properties of small planets and strip large primordial envelopes down to a few percent by mass (e.g., Schlichting et al., 2015; Liu et al., 2015; Inamdar & Schlichting, 2016). It is unclear whether a gap in the radius distribution could arise from impacts alone since impact erosion is a highly stochastic process. However, the atmospheric heating initiated by an impact can cause the envelope to expand, making it more susceptible to photoevaporation.

Signatures of Atmospheric Sculpting

Lopez & Rice (2016) considered two scenarios for the formation of sub-Neptunes/super-Earths. In one scenario, super-Earths are the remnant cores of photoevaporated, Neptune-size planets. In the other scenario super-Earths form late in the evolution of the protoplanetary disk, just as the gas dissipates (Lee et al., 2014). They predict that the transition radius between these two populations (the gap that we observed) should be a function of semi-major axis. If super-Earths are evaporated cores then the transition radius should be larger at lower incident flux. However, if super-Earths form in a gas-poor disk, or lose gas during the late stages of formation due to giant impacts, then the transition radius should decrease with increasing orbital distance. The distribution of planet radii as a function of insolation flux (Figure 8.10) does not show a clearly increasing or decreasing transition radius.

If photoevaporation is the dominant mechanism driving the distribution of planet sizes at short orbital periods, then we might expect that closely-spaced planets within multi-planet systems which experience similar irradiation histories would have similar sizes. Kepler-36 is one example to the contrary with both a sub-Neptune and super-Earth orbiting the same star at very similar orbital distances (Carter et al., 2012). A detailed analysis of the statistical properties of multi-planet systems utilizing the CKS stellar parameters is currently ongoing (Weiss et al. (in preparation)).

8.5.4 Core Mass Distribution

The masses of planets smaller than Neptune are dominated by the solid core. Thus, measuring the distribution of core masses provides a valuable constraint on their formation histories. The precise location and depth of the photo-evaporation valley likely depends on the underlying core mass distribution. Planet masses can be constrained using TTVs (Holman & Murray, 2005; Agol et al., 2005), but only in specific architectures that may probe different underlying populations. Most of the *Kepler* systems studied in this work are faint and out of reach of the current generation of RV instruments. And the number of RV mass measurements for small planets is too small to map out the core mass distribution in fine detail (Howard et al., 2010a; Mayor et al., 2011). Teasing out this distribution will require a large sample of low-mass planets amenable to mass measurements.

Ongoing and upcoming surveys such as the APF-50 survey (Fulton et al., 2016), the HARPS-N rocky planet search (Motalebi et al., 2015), MINERVA (Swift et al., 2015), and TESS (Ricker et al., 2014) are working to achieve this goal.

8.6 Conclusion

Using precise planet radii for 2025 *Kepler* planets from the CKS Survey, we examined the planet radius distribution at high-resolution. We find evidence for a bimodal distribution of small planet sizes. Sub-Neptunes and super-Earths appear to be two distinct planet classes. Planets tend to prefer radii of either $\sim 1.3 R_{\oplus}$ or $\sim 2.4 R_{\oplus}$, with relatively few planets having radii of 1.5–2.0 R_{\oplus} . Planets in the gap have the maximum size for a rocky core, as seen in previous studies of bulk planet density and of ultra-short period planets. We posit that the bimodal planet radius distribution stems from differences in the envelope masses of small planets. While our current dataset is insufficient to distinguish between theoretical models that produce the gap, it charts a path forward to unraveling further details of the properties of the galaxy’s most abundant planets.

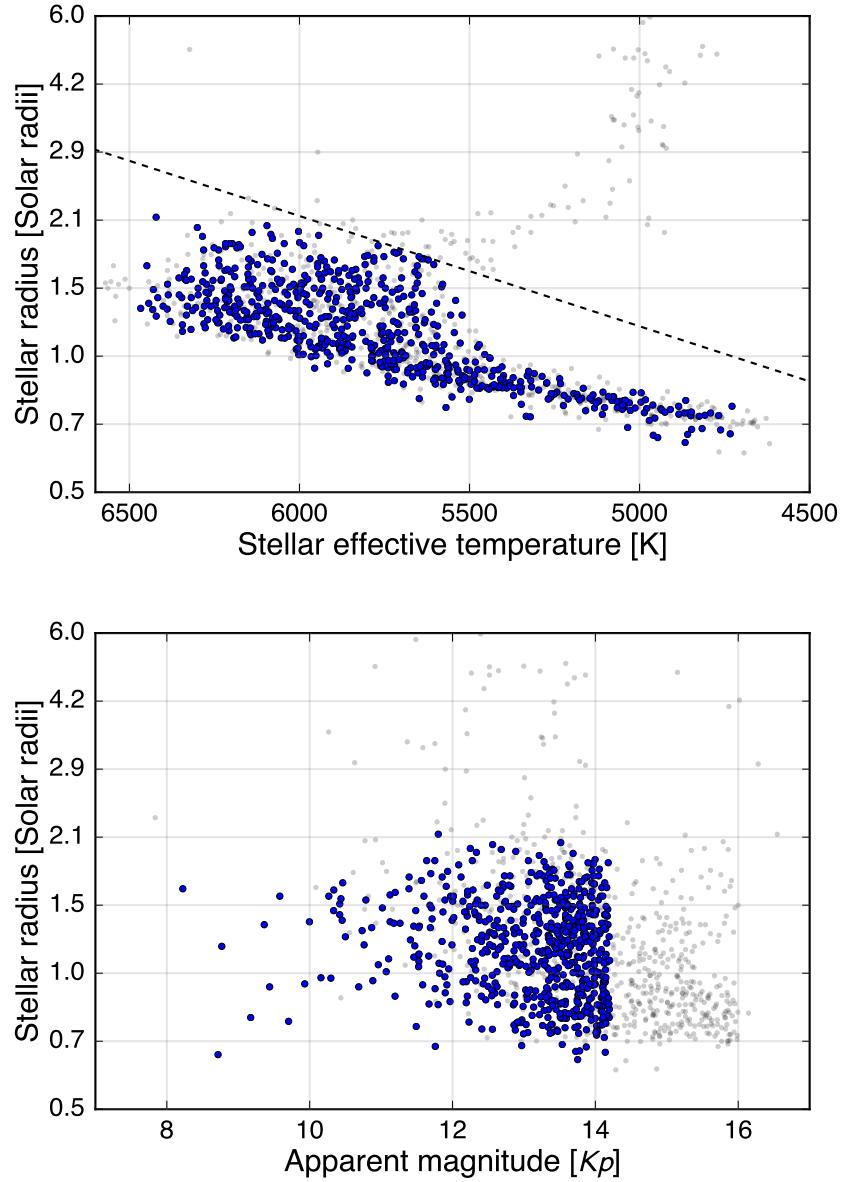


Figure 8.1 *Top*: HR diagram of the sample of stars selected for analysis. The full Paper II sample is plotted in light grey points and the sample selected for analysis after applying the filters discussed in Section 8.2.2 are plotted as blue squares. Giant planet hosting stars that fall above the dashed line given by Equation 8.1 are omitted from the final sample. *Bottom*: Stellar radius of CKS stars as a function of *Kepler* magnitude (K_p). We note that stars fainter than 14.2 do not follow the same stellar radius distribution. We omit stars fainter than $K_p = 14.2$ to avoid biasing our planet radius distribution. The point colors are the same as in the *top* panel.

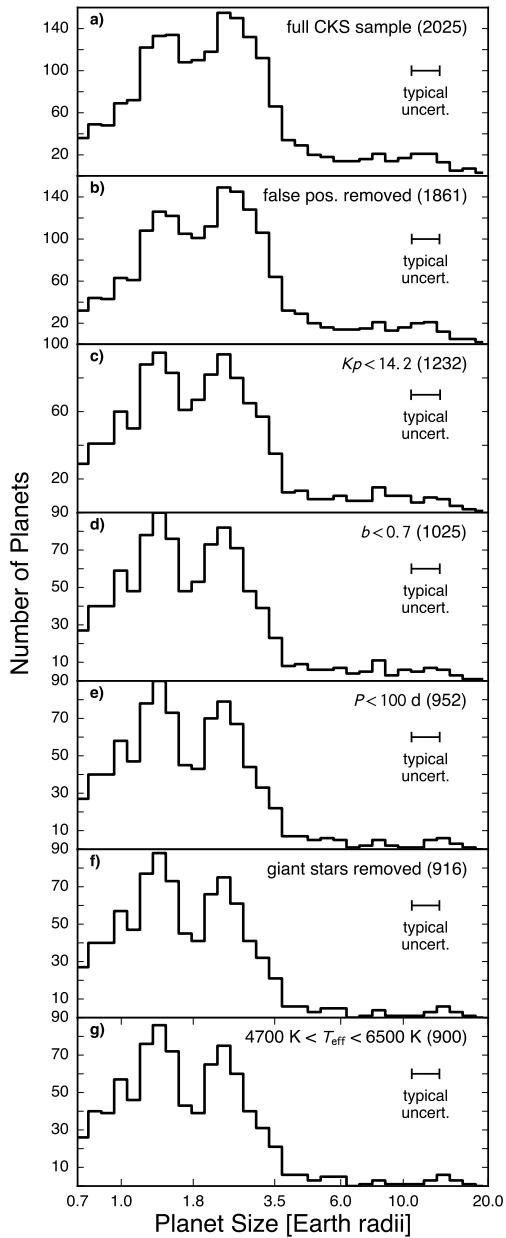


Figure 8.2 (a) Size distribution of all planet candidates in the CKS planet sample. Panels (b)–(g) show the radius distribution after applying several successive cuts to (b): remove known false positives, (c): keep candidates orbiting bright stars ($K_p < 14.2$), (d): retain candidates with low impact parameters ($b < 0.7$), (e): keep candidates with orbital periods shorter than 100 days, (f): remove candidates orbiting giant host stars, and (g): include only candidates orbiting stars within our adopted T_{eff} range ($4700 \text{ K} < T_{\text{eff}} < 6500 \text{ K}$). The number of planets remaining after applying each successive filter is annotated in the upper right portion of each panel. Our filters produce a reliable sample of accurate planet radii and accentuate the deficit of planets at $1.8 R_{\oplus}$.

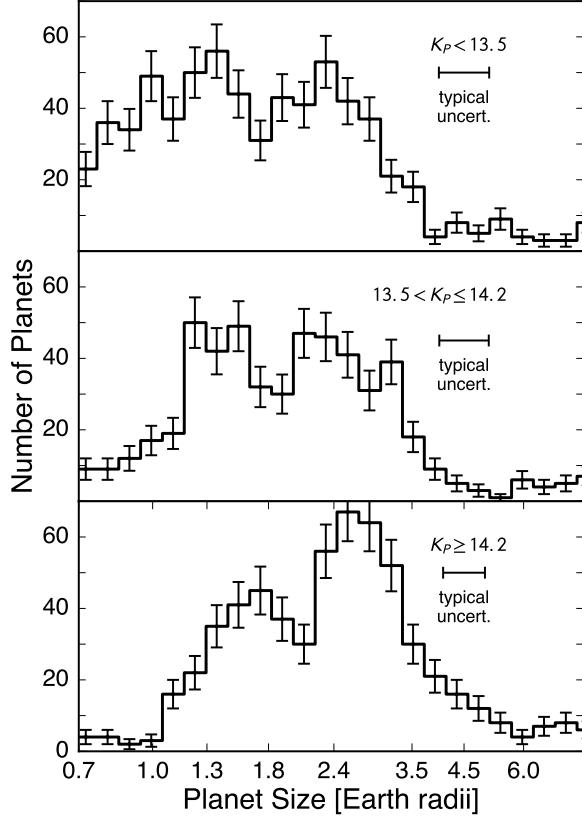


Figure 8.3 Histograms of planet radii broken up into the three magnitude ranges annotated in each panel. All of the filters have been applied to the sample as described in §8.2.2. The gap is apparent in all magnitude ranges. The distribution of planet radii in the two brightest magnitude ranges are indistinguishable ($p\text{-value} = 0.6$). However, the planets orbiting stars with $K_p > 14.2$ are statistically different ($p\text{-value} = 0.0004$) when compared to the $K_p = 13.5\text{--}14.2$ magnitude range. This is expected due to the non-systematic nature of the target selection for CKS and KIC stars fainter than $K_p = 14.2$. This motivates our removal of planets with hosts fainter than $K_p = 14.2$.

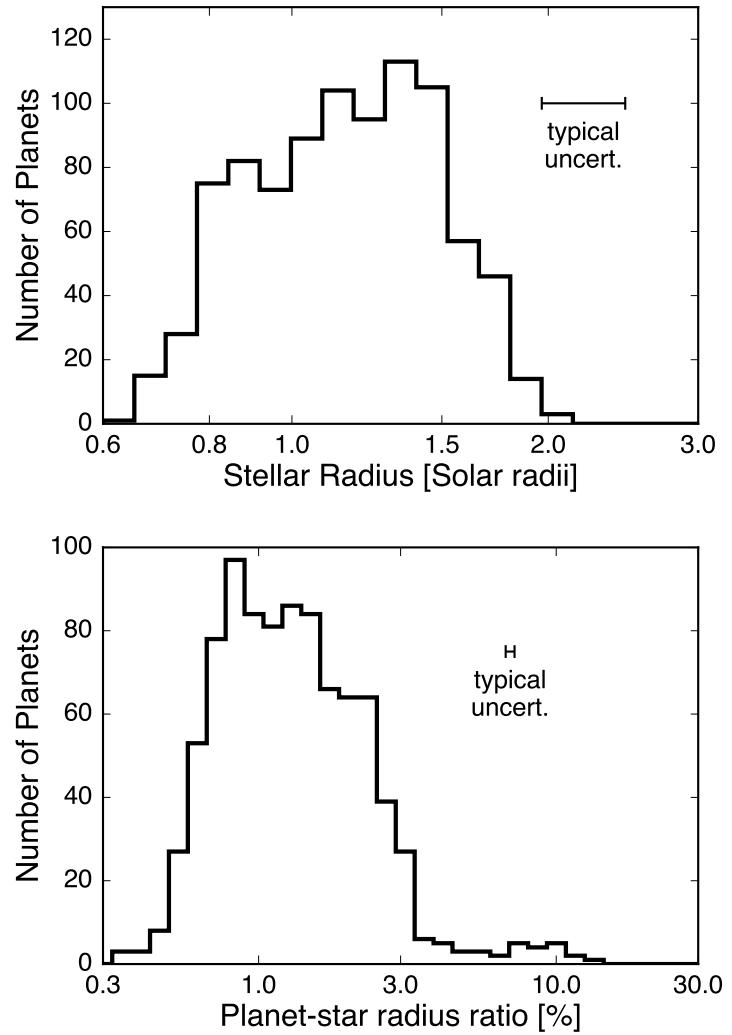


Figure 8.4 *Top*: Histogram of stellar radii derived in Paper II and used to update planet radii in this work after the filters described in Section 8.2.2 are applied. *Bottom*: Histogram of planet-to-star radii ratios for the stars remaining after the filters described in Section 8.2.2 are applied to the full Paper II sample of planet candidates. In both cases, the median measurement uncertainties are plotted in the upper right. Neither of these two histograms shows the same bimodal feature that is observed in the planet radius distribution, which demonstrates that the feature is not an artifact of our stellar sample or transit fitting.

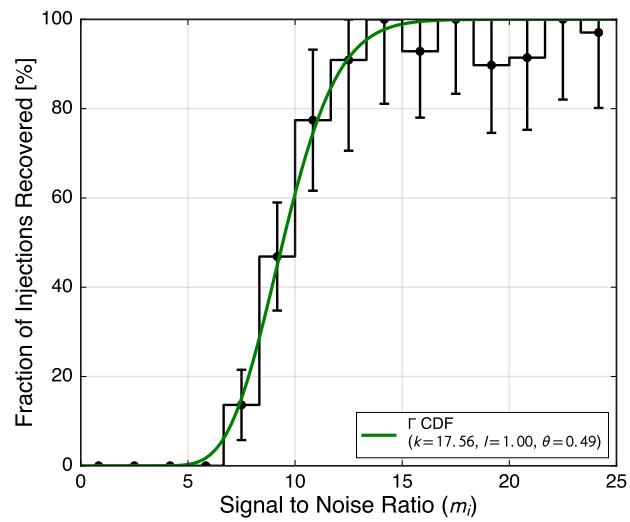


Figure 8.5 Fraction of injected transit signals recovered as a function of signal to noise ratio (m_i , Equation 8.2) in our subsample of the *Kepler* target stars using the injection recovery tests from Christiansen et al. (2015). We fit a Γ CDF (Equation 8.3) and plot the best-fit model in green.

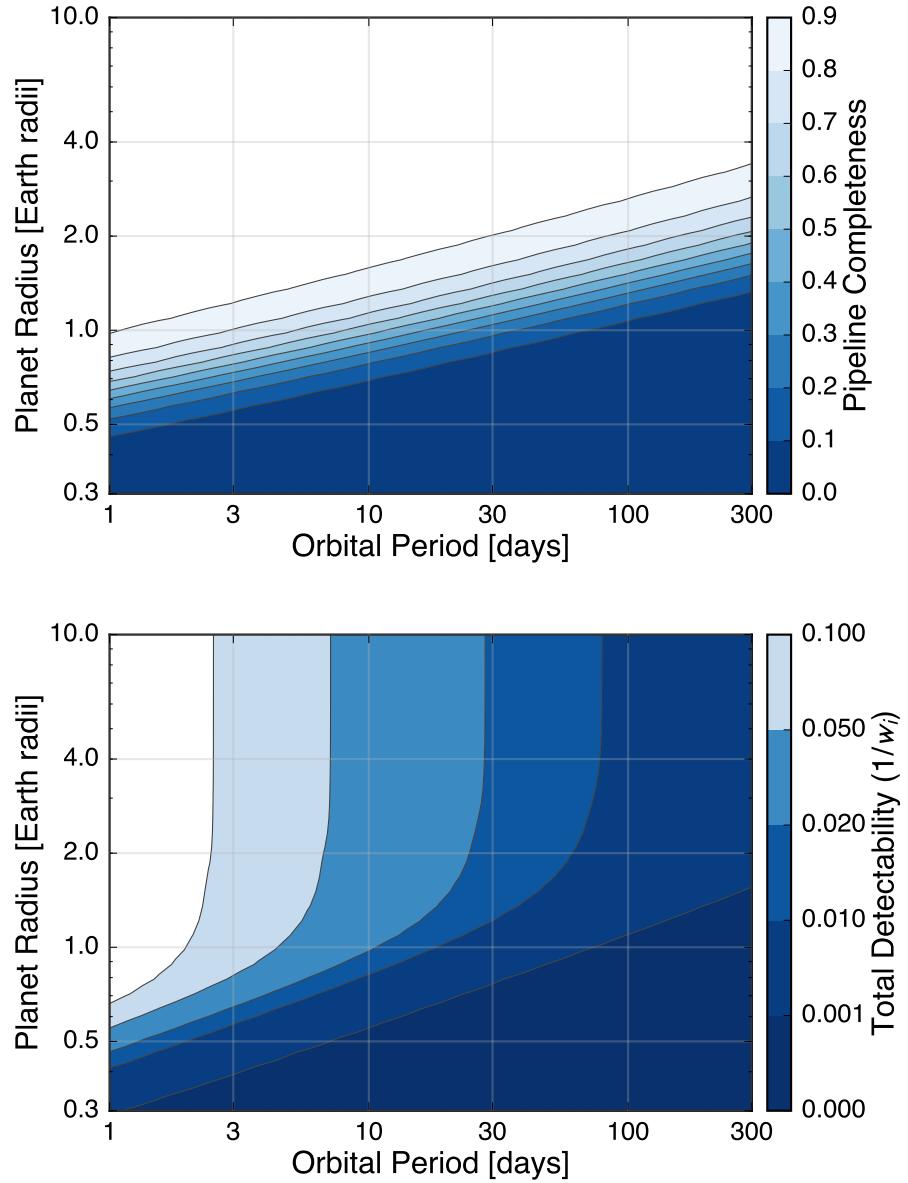


Figure 8.6 *Top:* Mean survey completeness for transiting planets orbiting the stars in our sample (p_{det}). *Bottom:* Mean survey completeness for all planets orbiting stars in our sample ($p_{\text{det}} \cdot p_{\text{tr}}$).

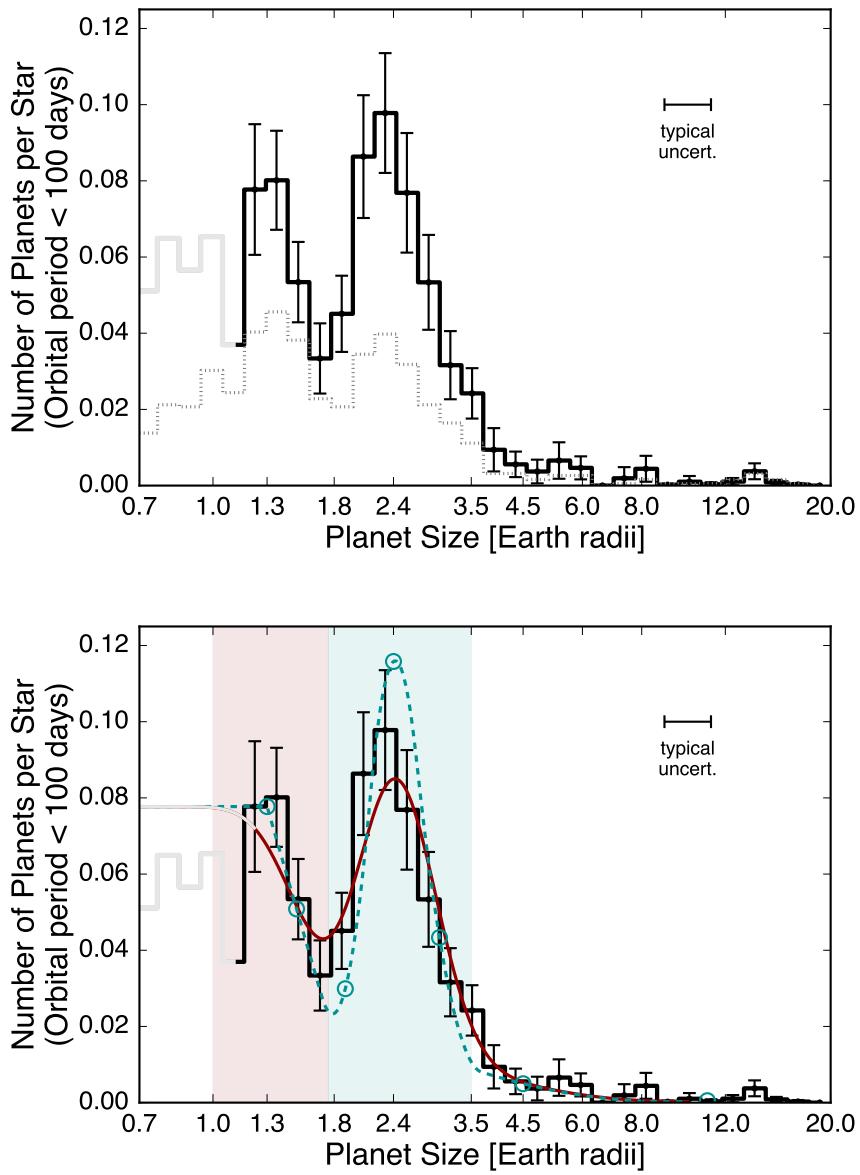


Figure 8.7 *Top*: Completeness-corrected histogram of planet radii for planets with orbital periods shorter than 100 days. Uncertainties in the bin amplitudes are calculated using the suite of simulated surveys described in Section A.3. The light gray region of the histogram for radii smaller than $1.14 R_{\oplus}$ suffers from low completeness. The histogram plotted in the dotted grey line is the same distribution of planet radii uncorrected for completeness. The median radius uncertainty is plotted in the upper right portion of the plot. *Bottom*: Same as top panel with the best-fit spline model over-plotted in the solid dark red line. The region of the histogram plotted in light grey is not included in the fit due to low completeness. Lightly shaded regions encompass our definitions of “super-Earths” (light red) and “sub-Neptunes” (light cyan). The dashed cyan line is a plausible model for the underlying occurrence distribution after removing the smearing caused by uncertainties on the planet radii measurements. The cyan circles on the dashed cyan line mark the node positions and values from the spline fit described in §8.4.3.

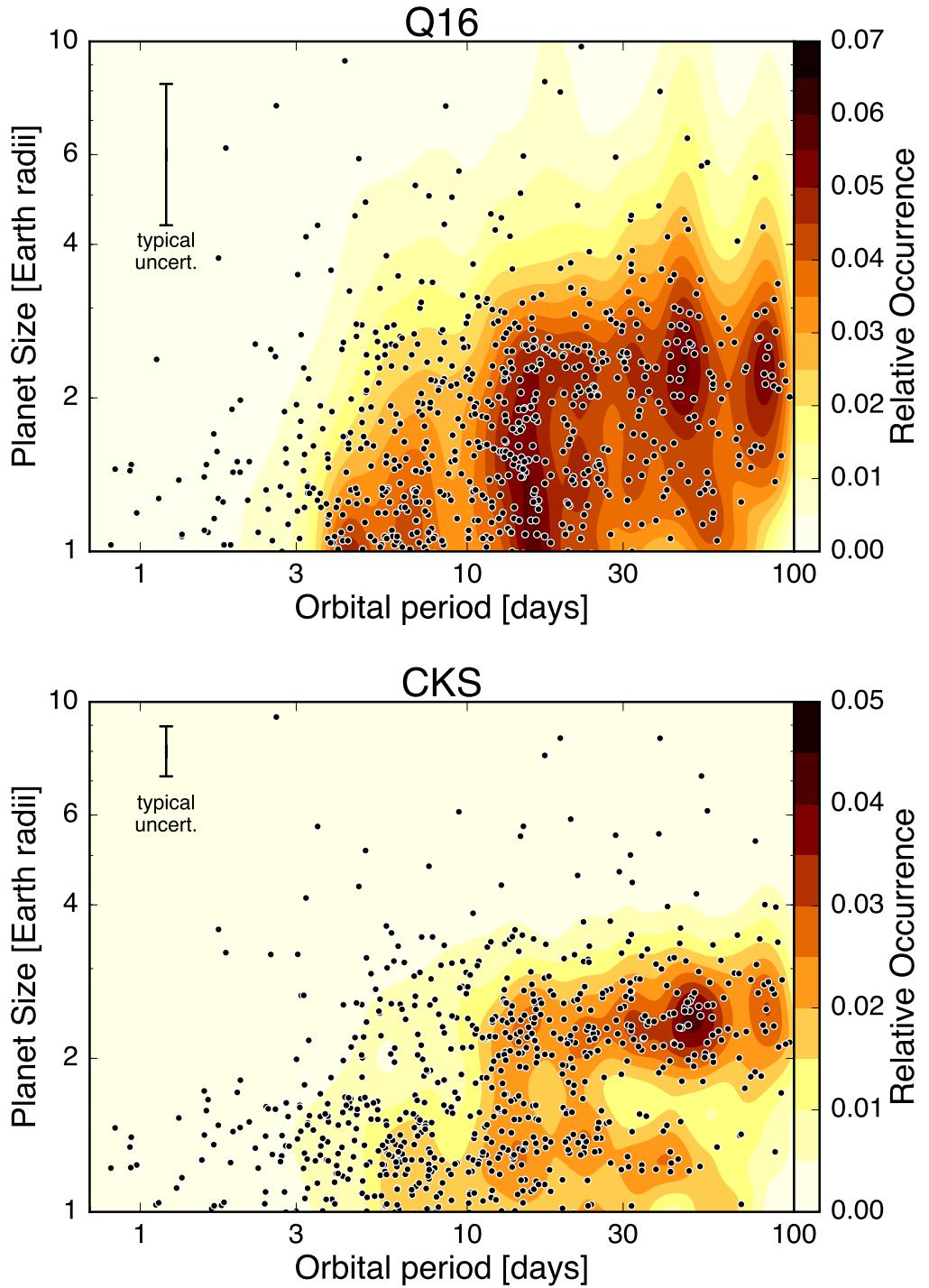


Figure 8.8 *Top:* Two-dimensional planet radius distribution as a function of orbital period using stellar parameters from the Q16 catalog. *Bottom:* Two-dimensional planet radius distribution as a function of orbital period using updated planet parameters from Paper II. In both cases the median uncertainty is plotted in the upper left. Individual planet detections are plotted as black points. The contours are corrected for completeness using the wKDE technique.

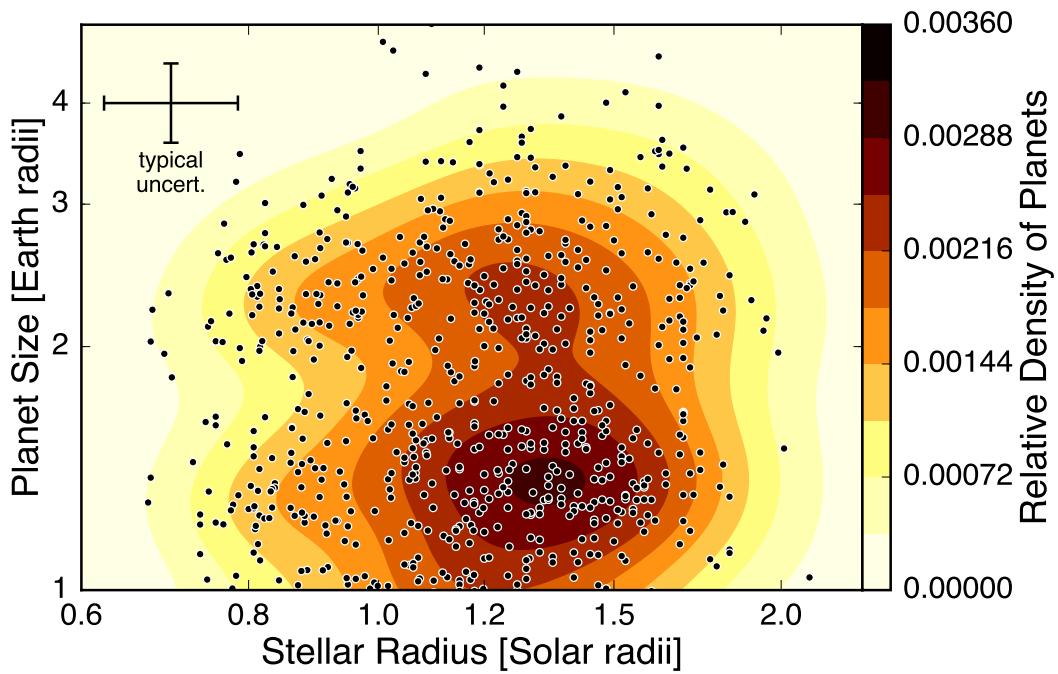


Figure 8.9 Two-dimensional planet radius distribution as a function of stellar radius using updated planet parameters from Paper II. The median uncertainty is plotted in the upper left. Individual planet detections are plotted as black points. The underlying contours are not corrected for completeness. The bifurcation of planet radii is independent of the size of the host star.

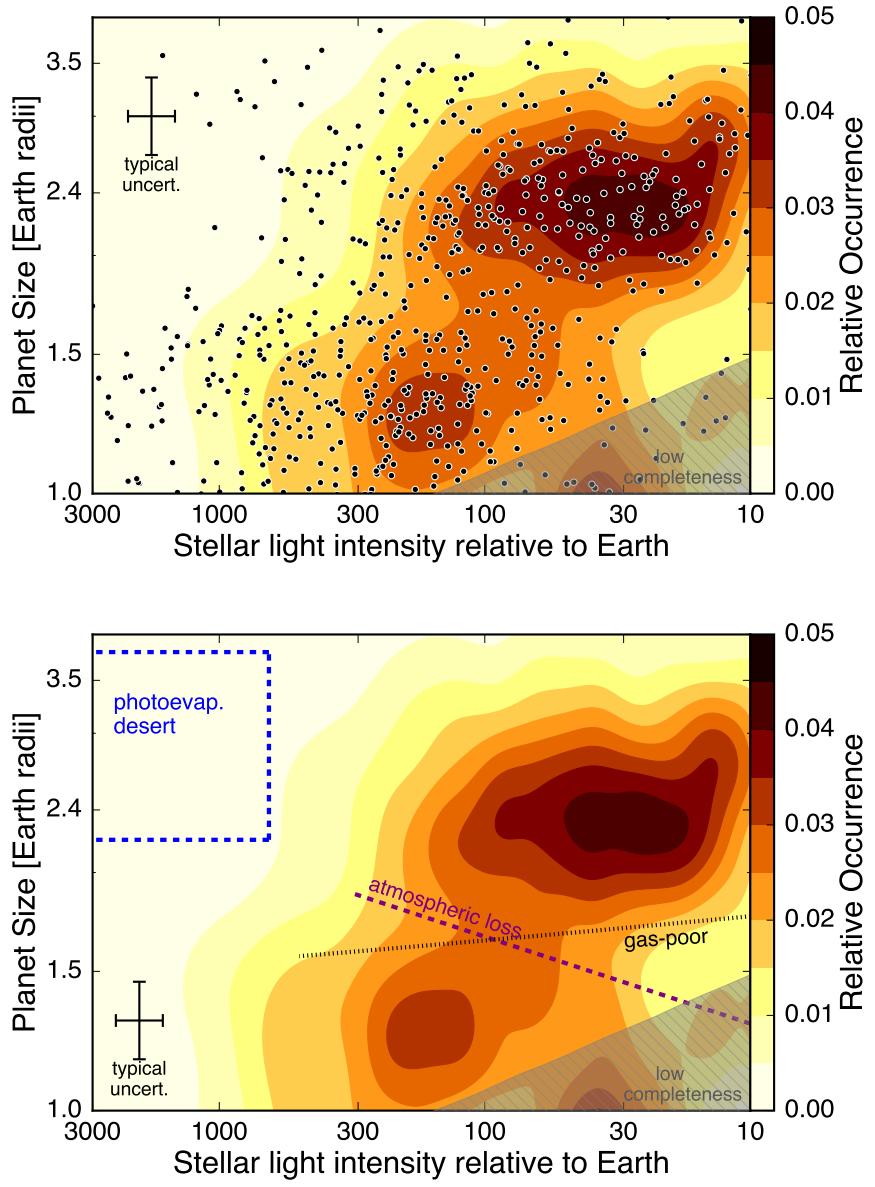


Figure 8.10 *Top*: Two-dimensional distribution of planet size and incident stellar flux. The median uncertainty is plotted in the upper left. There are at least two peaks in the distribution. One class of planets has typical radii of $\sim 1.3 R_{\oplus}$ and generally orbit in environments with $S_{\text{inc}} > 100 S_{\oplus}$, while another class of slightly larger planets with typical radii of $\sim 2.4 R_{\oplus}$ orbit in less irradiated environments with $S_{\text{inc}} < 200 S_{\oplus}$. *Bottom*: Same as top panel with individual planet detection points removed, annotations added, and vertical axis scaling changed. The region enclosed by the dashed blue lines marks the photoevaporation desert, or hot-Super Earth desert as defined by Lundkvist et al. (2016). The shaded region in the lower right indicates low completeness. Pipeline completeness in this region is less than 25%. The purple and black lines show the scaling relations for the photoevaporation valley predicted by Lopez & Rice (2016) for scenarios where these planets are the remnant cores of photoevaporated Neptune size planets (dashed purple line) or that these planets are formed at late times in a gas-poor disk (dotted black line).

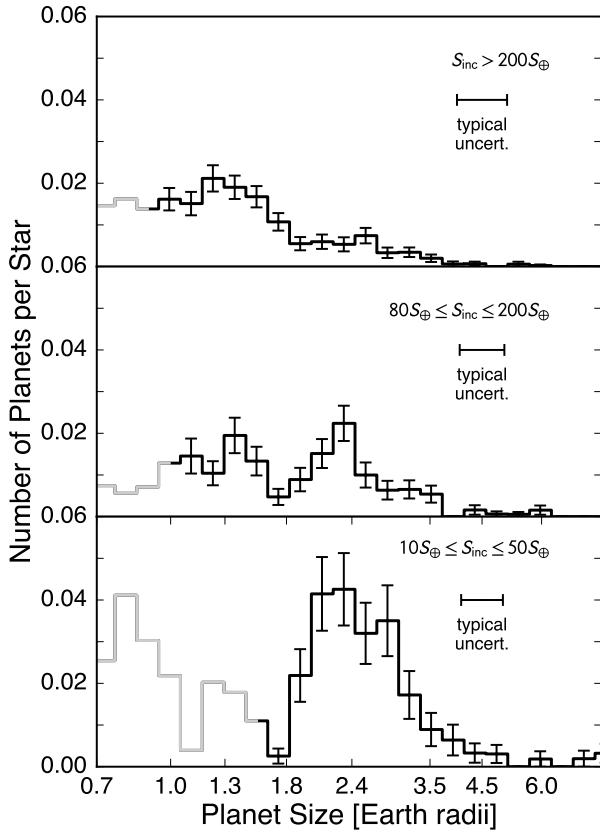


Figure 8.11 Histograms of planet radii broken up into the ranges of incident flux (S_{inc}) annotated in the upper right region of each panel. Planets orbiting in environments of higher S_{inc} tend to be smaller than those in low S_{inc} environments. Regions of the histograms plotted in light grey are highly uncertain due to pipeline completeness (<25%).

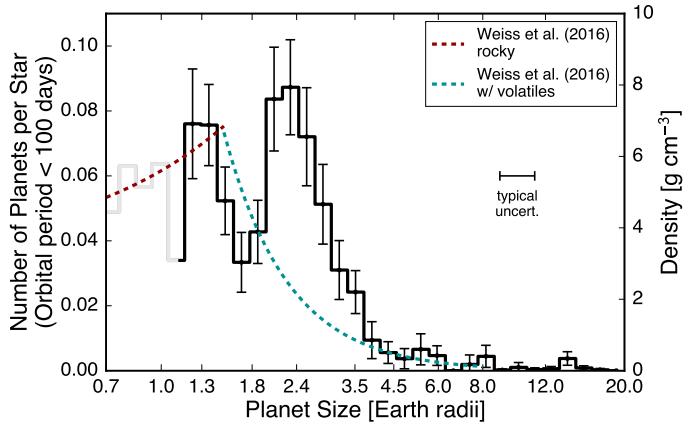


Figure 8.12 An empirical fit to planet radius and mass measurements from (Weiss et al., 2016) over-plotted on the completeness-corrected planet radius distribution derived in this work. The maximum in the planet density fit peaks near the gap in the planet radius distribution.

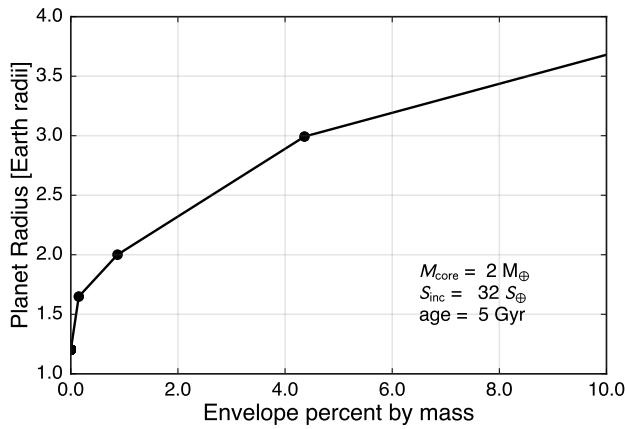


Figure 8.13 Model for planet radius as a function of envelope size from Lopez & Fortney (2014). The final planet radius is plotted for a simulated planet with a $2 M_{\oplus}$ core mass that has been irradiated by 32 times the incident flux received by Earth for a period of 5 Gyr. A bare $2 M_{\oplus}$ core has a radius of $1.2 R_{\oplus}$. Adding an envelope of H/He which is less than 0.2% of the planet's mass inflates the planet to over $1.6 R_{\oplus}$. An additional 0.7% envelope by mass inflates the planet to $2 R_{\oplus}$.

CHAPTER 9

CONCLUSION

This thesis presents a compilation of work dedicated to the discovery and characterization of exoplanets in all of their various forms. These discoveries and the tools developed to characterize the planets accurately and precisely enable the study of exoplanet demographics in unprecedented scope and detail.

In Chapters 2 and 3 we developed a set of tools that allowed us to precisely determine the mass of a planet, the orientation of it's orbit relative to the spin of the host star. We also developed a technique that allows for the extraction of precise iodine-based RV measurements from spectra that were previously thought to be too low signal-to-noise to make these measurements. This opens the door to the followup and characterization of many more planets orbiting much fainter host stars, and can save significant telescope time in the observation of moderate brightness targets.

These tools, along with the development of the automation software for the APF allowed for the discovery of several low-mass planets orbiting bright, nearby stars. HD 7924 is a system of three super-Earth mass planets orbiting a star only 17 pc distant. This compact, multi-planet system mirrors the many multi-planet systems discovered by *Kepler*, but gives some insight into the typical masses of planets in these systems. The three planets discovered in Chapter 5 all have interesting individual characteristics, but more importantly they add to the growing number of sub-Neptune mass planets with measured masses and serve to show that these planets may be the high-mass tail of the population of sub-Neptune size planets that the *Kepler* mission told us are so common in our galaxy.

We studied an interesting niche of the planet population by searching for transiting planets orbiting white dwarfs in Chapter 6. These stars are the remnants of stars slightly more massive than the sun. Similar white dwarfs had been known to show signs of silicates, and other minerals and elements on their surfaces. The degenerate nature of white dwarfs, and their incredibly high surface gravities should cause this material to sink below the photosphere relatively quickly. We did not detect any transiting planets orbiting the white dwarfs in our sample which allowed us to put strong constraints on the population of close-in planets orbiting white dwarfs. Shortly after publication of this work a disintegrating minor planet orbiting a white dwarf was discovered (Vanderburg et al., 2015). Combined with our statistical study it is clear that the population of planets that do orbit white dwarfs is much different than the population orbiting sun-like main-sequence stars.

Chapters 7 and 8 encompass the primary results of the APF-50 survey. We were successful in discovering \sim 10 new sub-Neptune mass planets using data autonomously collected from the APF. We demonstrated that observing these relatively inactive sun-like stars at high cadence (nearly once per night) provides huge gains in the detectability of small planets, and the ability to disentangle period aliasing and stellar activity from true planetary signals. We strove to obtain a uniform

dataset for all stars in our sample in order to minimize biases in the statistical study of the planets orbiting those stars. We successfully obtained a minimum of 82 observations per star and a mean of 125 observations per target. No target was observed more than 200 times which ensures relatively uniform sensitivity to small planets across our entire sample. The mass function of small planets shows a peak at around $10 M_{\oplus}$, which suggests that this might be a typical core mass. However, these statistics rely heavily on just a few planet detections and these will need to be more carefully vetted using other types of data (e.g. photometry).

Thanks to the increased precision on the stellar parameters of 1305 stars known to host *Kepler* planets from the California-Kepler Survey (Petigura et al., 2017; Johnson et al., 2017) we were able to view the radius distribution of small planets in great detail. Perhaps the highlight of this thesis is the discovery that planets smaller than Neptune are bifurcated into two distinct size classes. A population of planets that are probably mostly rocky with little to no atmospheres typically have sizes of $1.2 R_{\oplus}$, and a population of planets which must have a small, but non-negligible hydrogen/helium envelope have typical sizes of $2.4 R_{\oplus}$. We can extract an incredible amount of information about the formation and evolution of these planets from this feature in the radius distribution. First, it suggests that photo-evaporation is the primary mechanism that dictates the demographics of these small planets orbiting with periods shorter than 100 days. Second, it provides a way to estimate the core compositions of these planets. Owen & Wu (2017) and Jin & Mordasini (2017) both find that a large fraction of the cores of these planets must have Earth-like compositions and that there is very little spread in the compositions of those cores in order to explain the location and depth of the gap in the radius distribution. Third, the compositions of the cores suggest that they almost certainly formed inside the iceline. They did not form outside the iceline and migrate to their current close-in orbits as was previously invoked to explain the formation of these sub-Neptune planets. There is surely a multitude of information still to be extracted from this feature in the radius distribution.

Looking to the future there is still work to be done. Theorists have much work to do in order to incorporate the radius distribution into their planet formation models in order to determine the constraints on planet formation that can be extracted from that dataset. We will need to reconcile the mass distribution of small planets with the radius distribution. We now see a potential peak in the mass distribution of small planets, but a larger sample of stars and planets is needed to measure the mass function for planets smaller than $\sim 5 M_{\oplus}$. The majority of *Kepler* planets that were used to measure the radius distribution are most likely less massive than $5 M_{\oplus}$ (Owen & Wu, 2017). The Transiting Exoplanet Survey Satellite (TESS, Ricker et al., 2014) will provide a large sample of transiting exoplanets that are bright enough to be accessible to high-precision RVs which should allow us to make robust conclusions about the relationship between the radius distribution and the mass distribution of small planets.

One of the most consequential insights we have gained from the rise of the field of exoplanet

demographics is that small planets are nearly ubiquitous. This field is only just getting off the ground and there is certainly much work to be done in order to understand the context of our planet and solar system among the incredibly numerous and diverse population of planets orbiting stars throughout the galaxy.

APPENDIX A

SUPPLEMENTARY INFORMATION FOR CHAPTER 8

A.1 Non-cummulative Filters

We investigate the impact of each individual filter on the planet catalog by producing a figure similar to Figure 8.2. However, instead of plotting the distribution after all successive filters are applied to the original sample we plot the distributions after applying *only* the filter specified in the annotations and the figure caption (Figure A.1). The magnitude and impact parameter cuts have the greatest impact on the final sample since they subtract the greatest number of planets. However, no filter preferentially removes planets in the gap or preferentially preserves planets just outside the gap.

A.2 Weighted Kernel Density Estimation

The weights calculated in Section 8.3 can be used to estimate the occurrence rate distribution of any planet property using weighted kernel density estimation as an alternative to binned histograms (wKDE, Morton & Swift, 2014). We calculate the kernel density estimate as:

$$\phi(x) = \frac{1}{N_\star} \sum_{i=1}^{n_{\text{pl}}} w_i \cdot K(x - x_i, \sigma_{x,i}). \quad (\text{A.1})$$

K is the “kernel” and, in general, it can be any non-negative function that integrates to one and has a centroid of zero. x_i are the individual measurements for a given planet property and $\sigma_{x,i}$ are the uncertainties on those measurements. We treat double-sided uncertainties as symmetric Gaussian uncertainties by taking the mean of the reported upper and lower 1-sigma uncertainties. We adopt a standard Gaussian kernel to calculate the one-dimensional distributions of planet properties, and a bivariate Gaussian for two-dimensional distributions. In order to ensure smooth distributions and contours we limit fractional measurement uncertainty to $\geq 5\%$ in the calculation of the 2D wKDEs. Orbital period is the only parameter that is subject to this limit.

To investigate the possibility that the gap in the planet radius distribution is an artifact of binning we calculate the planet radius distribution using wKDE (Figure A.2). We choose a Gaussian kernel and a variable bandwidth that matches the radius uncertainty for each individual measurement. Again, there are two peaks in the radius distribution separated by a gap. The wKDE demonstrates that the presence or location of the gap does not depend on the particular choice of bin size. The contrast between the bottom of the gap and the top of the peaks is reduced in the wKDE-derived planet radius distribution. However, as shown in the simulations described in

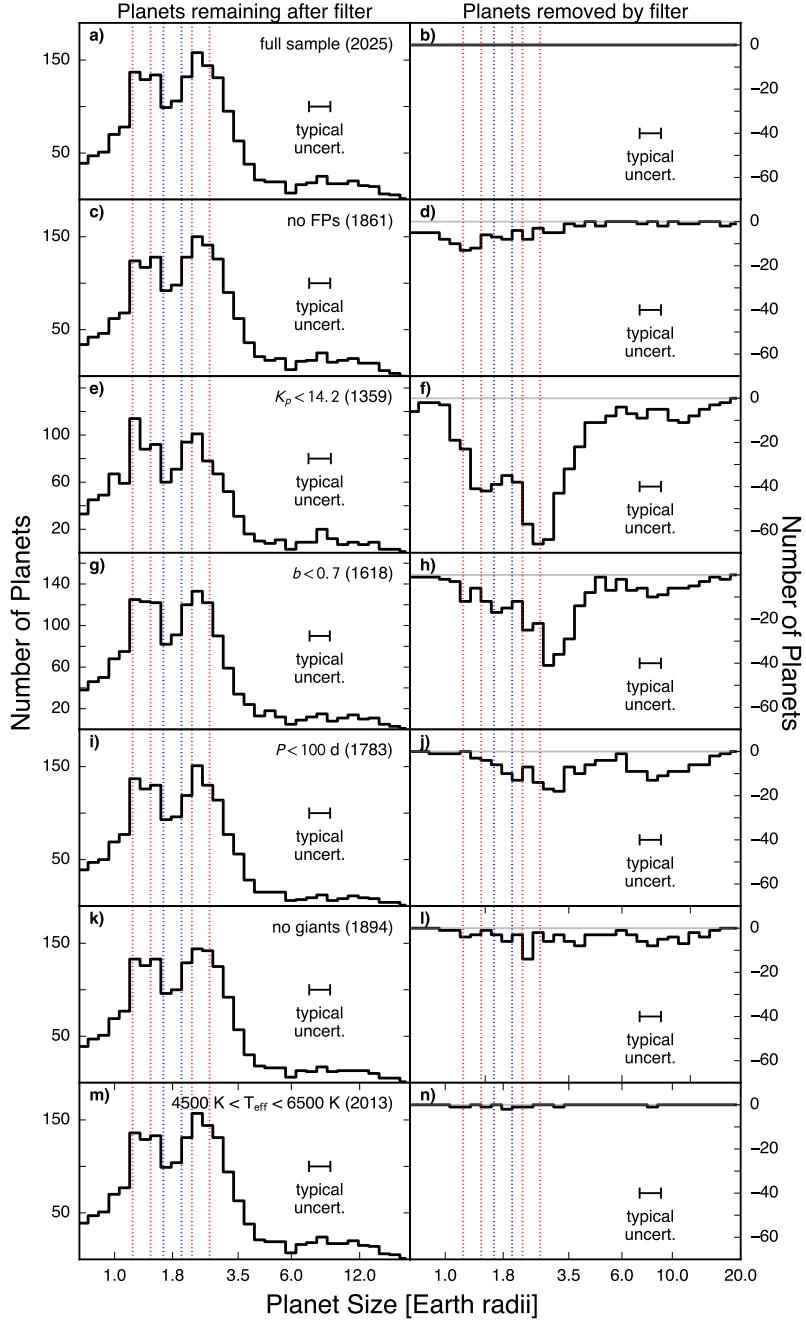


Figure A.1 (a) size distribution of planet candidates from the CKS sample. (b) planets removed by the specified filter. Panels (c)–(n) show the radius distribution and planets removed from the full sample after applying only a single cut removing known false positives (c), planets orbiting faint host stars (e), planets with grazing transits (g), planets with orbital periods longer than 100 days (i), planets orbiting giant host stars (k), and planets orbiting host stars cooler than 4700 K or hotter than 6500 K (m). No completeness corrections have been applied. The $b < 0.7$ cut and the $K_p < 14.2$ cut remove the most planet candidates, but no filter preferentially removes planets in the gap (between blue dotted lines).

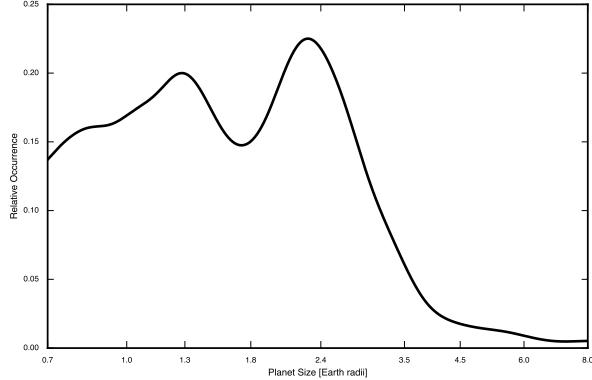


Figure A.2 Bin-free view of the planet radius distribution calculated using wKDE (Equation A.1). The 1-sigma uncertainty region is shaded in red and calculated using a suite of simulated transit surveys as described in Appendix A.2.

Appendix A.3, this is an artifact of the wKDE technique and probably not a good representation of the underlying radius distribution. The planet radius uncertainties are effectively being counted twice in both the scatter of the median values and the width of the Gaussians summed to create the wKDE. The simulations described in Appendix A.3 show the same dilution of the gap depth when using the wKDE to recover known distributions of simulated planets. Quantifying the valley depth from the wKDE radius distribution may require a careful exploration and justification of the kernel bandwidth selection. Our simulations show that the histograms better reproduce the known input distributions, so we choose leave this bandwith tuning for future studies and conclude that the histogram gives a more accurate picture of the planet radius distribution over this particular application and implementation of the wKDE.

A.3 Validation of the Completeness Corrections

We validate our occurrence calculations and estimate uncertainties by constructing a suite of 100 simulated transit surveys. For each simulation, we draw a distribution of 45000 planet radii and orbital periods from two lognormal distributions then sum those distributions together to create a bimodal distribution similar to the distribution observed in our real planet detections (Figure A.3). We assign each simulated planet to a star in our filtered sample of KOI hosts and calculate detection probabilities and weights as described in §8.3.3. These detection probabilities are used to decide which planets would have been detected in our real survey. The number of simulated planets (45000) was chosen such that the mean number of planets in the 100 simulated planet detection catalogues is equal to the total number of planets in our filtered KOI catalogue (900).

The stellar radii for the stars in the Stellar17 sample, which are used in the completeness

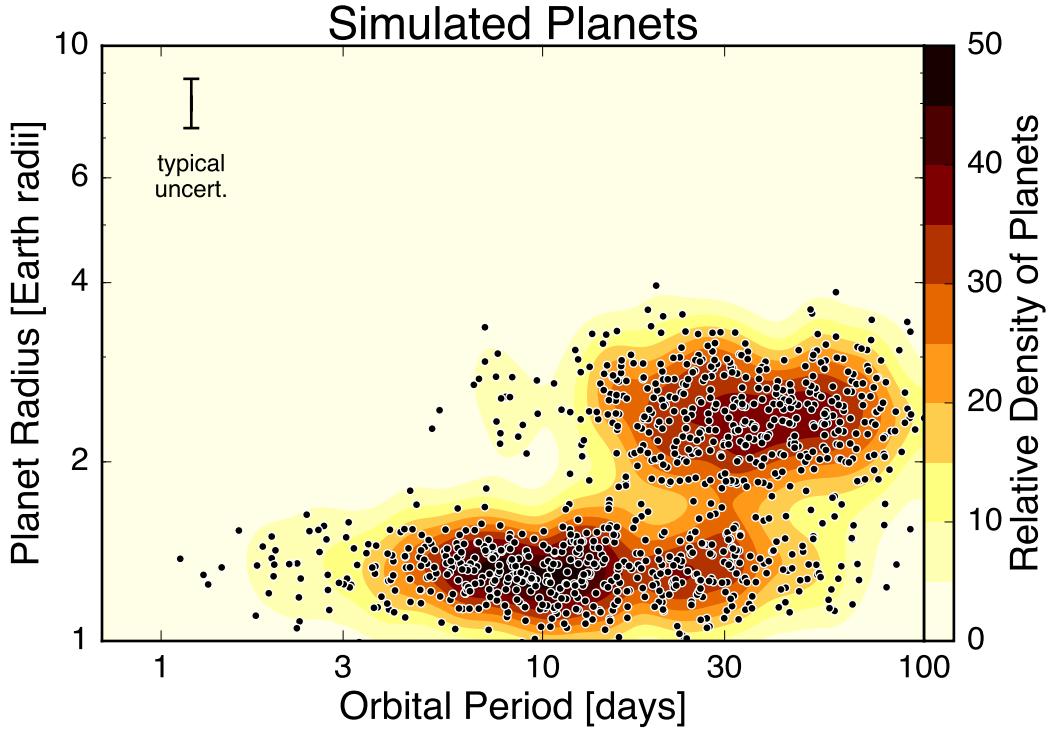


Figure A.3 Radius vs. period distribution for simulated sample of planets. For plotting clarity and speed we plot only 1,000 randomly chosen simulated planets out of the 45,000 simulated planets.

corrections, are perturbed in two different ways in each simulation. We multiply all of the Stellar17 stellar radii by a common constant drawn from a normal distribution centered at 1.0 with a width of 0.25 to simulate potential systematic offsets between the stellar radii in the Stellar17 catalogue and the stellar radii in the CKS catalog. We also add Gaussian noise to the stellar radii for all stars with distribution widths determined from their individual measurement uncertainties. The uncertainties in our final bin heights and occurrence ratios estimated from these simulations account for both systematic and Gaussian random errors in the stellar parameters in the Stellar17 catalog.

We produce histograms for each simulation and correct them for completeness as described in §8.3.3. The standard deviation of the values in each histogram bin become the uncertainty on the bin values. When compared with uncertainties calculated using Poisson statistics on the number of simulated detections in each bin we find that the Poisson uncertainties are underestimated by a factor of 1.5–2.9 depending on the radius bin. In order to avoid small number statistics for the histogram bins where the simulated distribution approaches zero we repeat the simulations with an input distribution of planets that is log-uniform in radius from 0.5–20.0 R_{\oplus} and log-uniform in period from 1–200 days solely for the purpose of calculating the uncertainty scaling factors for each radius bin. We adopt the scaling factors listed in Table A.1 in the calculation of all completeness-

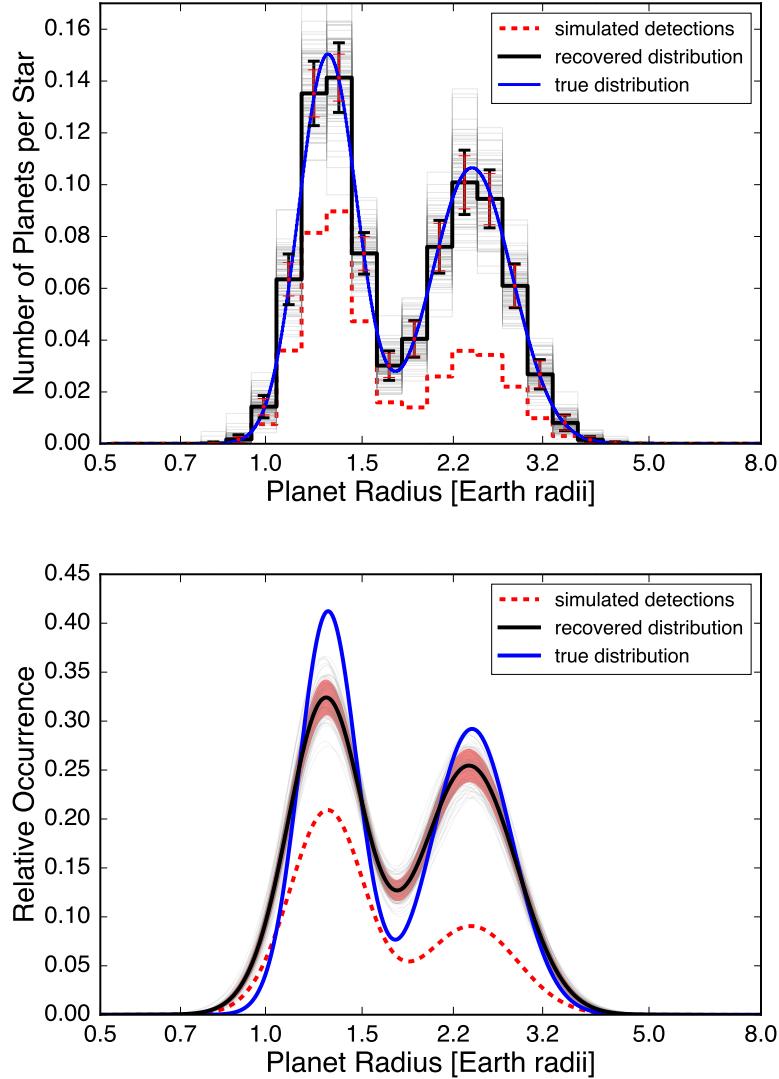


Figure A.4 *Top*: Results from simulating 100 transit surveys with a known input distribution of planets. The input distribution of simulated planets is plotted in blue, and the simulated detections are plotted in a red dashed line. The completeness-corrected distributions measured from each of the simulations are plotted as thin grey lines and the median of those recovered distributions is plotted in a thick black line. The thick black error bars are the standard deviation of all of the simulations in each bin and the thin red error bars are poisson uncertainties on the number of detections in each bin scaled by the completeness correction for that bin. *Bottom*: Same as *top* panel but calculated using the wKDE technique described in §8.3.3. The shaded red area encompasses the standard deviation of the resulting wKDEs over all 100 simulations. We adopt this fractional uncertainty for the one-dimensional KDE plotted in Figure A.2.

Table A.1. Bin
Uncertainty Scaling
Factors

Radius bin R_{\oplus}	Scaling Factor
0.50–0.56	2.82
0.56–0.62	2.50
0.62–0.69	2.30
0.69–0.76	2.54
0.76–0.85	2.35
0.85–0.94	2.09
0.94–1.05	1.92
1.05–1.16	1.95
1.16–1.29	1.89
1.29–1.43	1.46
1.43–1.59	1.65
1.59–1.77	1.81
1.77–1.97	1.38
1.97–2.19	1.50
2.19–2.43	1.39
2.43–2.70	1.58
2.70–3.00	1.48
3.00–3.33	1.58
3.33–3.70	1.25
3.70–4.12	1.48
4.12–4.57	1.47
4.57–5.08	1.46
5.08–5.65	1.63
5.65–6.27	1.45
6.27–6.97	1.50
6.97–7.75	1.52
7.75–8.61	1.34
8.61–9.56	1.44
9.56–10.63	1.46
10.63–11.81	1.52
11.81–13.12	1.57
13.12–14.58	1.36
14.58–16.20	1.35
16.20–18.00	1.45
18.00–20.00	1.44

corrected planet radius histograms and for fitting the distribution described in §8.4.

We calculate the occurrence ratio of super-Earths to sub-Neptunes in the same way as we do for the real planet catalogue in §8.4.4. The mean occurrence ratio is consistent with the same ratio for the input distribution of simulated planets and the standard deviation as a fraction of the ratio is 33%. We adopt this fractional uncertainty for the occurrence ratio calculation on the real planet catalogue.

We also calculate the radius distribution for each simulation using the wKDE technique described in Appendix A.2. We find that the wKDE slightly underestimates the contrast between the peaks of the radius distribution and the bottom of the gap. This is likely due to the fact that there is scatter in the radii measurements due to uncertainties. Those uncertainties are also being included as the widths of the Gaussians used to calculate the wKDE, in effect counting the uncertainty twice. Since we do not perform any quantitative analysis on the wKDE we choose not to “de-bias” the wKDE as described in Morton & Swift (2014), but instead limit our quantitative analysis to the histograms that seem to be a more accurate representation of the underlying distributions in our simulations. We use the resulting wKDEs from the simulated surveys in order to

estimate the fractional uncertainty as a function of planet radius for the wKDE calculated from the real planet catalog (Figure A.2).

BIBLIOGRAPHY

- Adams, E. R., Ciardi, D. R., Dupree, A. K., et al. 2012, AJ, 144, 42
- Adams, E. R., Dupree, A. K., Kulesa, C., & McCarthy, D. 2013, AJ, 146, 9
- Adams, E. R., Seager, S., & Elkins-Tanton, L. 2008, ApJ, 673, 1160
- Agol, E. 2011, ApJ, 731, L31
- Agol, E., Steffen, J., Sari, R., & Clarkson, W. 2005, MNRAS, 359, 567
- Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147
- Alard, C. 2000, A&AS, 144, 363
- Alard, C., & Lupton, R. H. 1998, ApJ, 503, 325
- Albrecht, S., Winn, J. N., Butler, R. P., et al. 2012a, ApJ, 744, 189
- Albrecht, S., Winn, J. N., Johnson, J. A., et al. 2012b, ApJ, 757, 18
- Alibert, Y., Carron, F., Fortier, A., et al. 2013, A&A, 558, A109
- Anderson, T. W., & Darling, D. A. 1952, Ann. Math. Statist., 23, 193
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
- Assef, R. J., Gaudi, B. S., & Stanek, K. Z. 2009, ApJ, 701, 1616
- Babu, G. J., & Feigelson, E. D. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 127
- Baglin, A., Auvergne, M., Barge, P., et al. 2009, in IAU Symposium, Vol. 253, IAU Symposium, ed. F. Pont, D. Sasselov, & M. J. Holman, 71–81
- Baglin, A., Michel, E., & CoRoT Team. 2012, in Astronomical Society of the Pacific Conference Series, Vol. 462, Progress in Solar/Stellar Physics with Helio- and Asteroseismology, ed. H. Shibashi, M. Takata, & A. E. Lynas-Gray, 492
- Bailey, III, J. I., White, R. J., Blake, C. H., et al. 2012, ApJ, 749, 16
- Bakos, G., Noyes, R. W., Kovács, G., et al. 2004, PASP, 116, 266
- Bakos, G. Á., Kovács, G., Torres, G., et al. 2007, ApJ, 670, 826

- Bakos, G. Å., Howard, A. W., Noyes, R. W., et al. 2009, *ApJ*, 707, 446
- Ballard, S., Fabrycky, D., Fressin, F., et al. 2011, *ApJ*, 743, 200
- Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 2002, *A&A*, 382, 563
- Baranec, C., Ziegler, C., Law, N. M., et al. 2016, *AJ*, 152, 18
- Baranec, C., Riddle, R., Law, N. M., et al. 2013, *Journal of Visualized Experiments*, 72, e50021
- . 2014, *ApJ*, 790, L8
- Baranne, A., Queloz, D., Mayor, M., et al. 1996, *A&AS*, 119, 373
- Barban, C., Deheuvels, S., Goupil, M. J., et al. 2013, *Journal of Physics Conference Series*, 440, 012030
- Batalha, N. M., Borucki, W. J., Koch, D. G., et al. 2010, *ApJ*, 713, L109
- Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013, *ApJS*, 204, 24
- Batygin, K., Bodenheimer, P., & Laughlin, G. 2009, *ApJ*, 704, L49
- Batygin, K., Stevenson, D. J., & Bodenheimer, P. H. 2011, *ApJ*, 738, 1
- Bayliss, D. D. R., & Sackett, P. D. 2011, *ApJ*, 743, 103
- Beatty, T. G., Pepper, J., Siverd, R. J., et al. 2012, *ApJ*, 756, L39
- Bensby, T., Feltzing, S., & Lundström, I. 2003, *A&A*, 410, 527
- Bento, J., Wheatley, P. J., Copperwheat, C. M., et al. 2014, *MNRAS*, 437, 1511
- Bianchi, L., Herald, J., Efremova, B., et al. 2011, *Ap&SS*, 335, 161
- Bieryla, A., Collins, K., Beatty, T. G., et al. 2015, *AJ*, 150, 12
- Bonneau, D., Delfosse, X., Mourard, D., et al. 2011, *A&A*, 535, A53
- Bonneau, D., Clausse, J.-M., Delfosse, X., et al. 2006, *A&A*, 456, 789
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, 327, 977
- Borucki, W. J., Koch, D. G., Basri, G., et al. 2011, *ApJ*, 736, 19
- Boyajian, T. S., McAlister, H. A., Baines, E. K., et al. 2008, *ApJ*, 683, 424
- Boyajian, T. S., von Braun, K., van Belle, G., et al. 2012, *ApJ*, 757, 112

- Burke, C. J., Christiansen, J. L., Mullally, F., et al. 2015, ApJ, 809, 8
- Burleigh, M. R., Hogan, E., Dobbie, P. D., Napiwotzki, R., & Maxted, P. F. L. 2006, MNRAS, 373, L55
- Burt, J., Vogt, S. S., Butler, R. P., et al. 2014, ApJ, 789, 114
- Butler, R. P., & Marcy, G. W. 1996, ApJ, 464, L153+
- Butler, R. P., Marcy, G. W., Vogt, S. S., et al. 2003, ApJ, 582, 455
- Butler, R. P., Marcy, G. W., Williams, E., et al. 1996a, PASP, 108, 500
- . 1996b, PASP, 108, 500
- Butler, R. P., Wright, J. T., Marcy, G. W., et al. 2006, ApJ, 646, 505
- Canto Martins, B. L., das Chagas, M. L., Alves, S., et al. 2011, Astronomy and Astrophysics, 530, A73
- Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, Science, 337, 556
- Cartier, K. M. S., Gilliland, R. L., Wright, J. T., & Ciardi, D. R. 2015, ApJ, 804, 97
- Catanzarite, J., & Shao, M. 2011, ApJ, 738, 151
- Chambers, J. E. 1999, MNRAS, 304, 793
- Chaplin, W. J., & Miglio, A. 2013, ARA&A, 51, 353
- Chatterjee, S., & Tan, J. C. 2014, ApJ, 780, 53
- Chen, H., & Rogers, L. A. 2016, ApJ, 831, 180
- Chiang, E., & Laughlin, G. 2013, MNRAS, 431, 3444
- Christensen-Dalsgaard, J., & Frandsen, S. 1983, Sol. Phys., 82, 469
- Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2015, ApJ, 810, 95
- . 2016, ApJ, 828, 99
- Chubak, C., Marcy, G., Fischer, D. A., et al. 2012, ArXiv e-prints, 1207.6212C, arXiv:1207.6212
- Claret, A., & Bloemen, S. 2011, A&A, 529, A75
- Cochran, W. D., Hatzes, A. P., Butler, R. P., & Marcy, G. W. 1997, ApJ, 483, 457
- Coelho, P. R. T. 2014, MNRAS, 440, 1027

- Coleman, G. A. L., & Nelson, R. P. 2014, MNRAS, 445, 479
- Collier Cameron, A., Bouchy, F., Hébrard, G., et al. 2007, MNRAS, 375, 951
- Collins, K. A., Eastman, J. D., Beatty, T. G., et al. 2014, AJ, 147, 39
- Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, PASP, 120, 531
- Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, VizieR Online Data Catalog, 2246, 0
- Cutri, R. M., Wright, E. L., Conrow, T., et al. 2014, VizieR Online Data Catalog, 2328, 0
- Dawson, R. I. 2014, ApJ, 790, L31
- Dawson, R. I., & Fabrycky, D. C. 2010, The Astrophysical Journal, 722, 937
- Dawson, R. I., & Murray-Clay, R. A. 2013, ApJ, 767, L24
- de Bruijne, J. H. J. 2012, Ap&SS, 341, 31
- Debes, J. H., Hoard, D. W., Kilic, M., et al. 2011, ApJ, 729, 4
- Debes, J. H., Sigurdsson, S., & Woodgate, B. E. 2005, AJ, 130, 1221
- Demory, B.-O. 2014, The Astrophysical Journal, 789, L20
- Demory, B.-O., & Seager, S. 2011, ApJS, 197, 12
- Devor, J. 2005, ApJ, 628, 411
- Díaz, R. F., Ségransan, D., Udry, S., et al. 2016, A&A, 585, A134
- Dong, S., & Zhu, Z. 2013, ApJ, 778, 53
- Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, ApJS, 178, 89
- Drake, A. J., Beshore, E., Catelan, M., et al. 2010, ArXiv e-prints, arXiv:1009.3048
- Dressing, C. D., Adams, E. R., Dupree, A. K., Kulesa, C., & McCarthy, D. 2014, AJ, 148, 78
- Dressing, C. D., & Charbonneau, D. 2015, ApJ, 807, 45
- Dumusque, X., Udry, S., Lovis, C., Santos, N. C., & Monteiro, M. J. P. F. G. 2011, A&A, 525, A140
- Eastman, J., Gaudi, B. S., & Agol, E. 2013, PASP, 125, 83
- Eastman, J., Siverd, R., & Gaudi, B. S. 2010, PASP, 122, 935

- Eaton, J. A., Henry, G. W., & Fekel, F. C. 2003, in *Astrophysics and Space Science Library*, Vol. 288, *Astrophysics and Space Science Library*, ed. T. D. Oswalt, 189
- Enoch, B., Collier Cameron, A., & Horne, K. 2012, *A&A*, 540, A99
- Everett, M. E., Barclay, T., Ciardi, D. R., et al. 2015, *AJ*, 149, 55
- Fabrycky, D., & Tremaine, S. 2007, *ApJ*, 669, 1298
- Fabrycky, D. C., Ford, E. B., Steffen, J. H., et al. 2012, *ApJ*, 750, 114
- Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, *ApJ*, 790, 146
- Faedi, F., West, R. G., Burleigh, M. R., Goad, M. R., & Hebb, L. 2011, *MNRAS*, 410, 899
- Fang, J., & Margot, J.-L. 2012, *ApJ*, 761, 92
- Farihi, J., Becklin, E. E., & Zuckerman, B. 2008, *ApJ*, 681, 1470
- Farihi, J., Zuckerman, B., & Becklin, E. E. 2005, *AJ*, 130, 2237
- Farrington, C. D., ten Brummelaar, T. A., Mason, B. D., et al. 2010, *AJ*, 139, 2308
- Fischer, D. A., Marcy, G. W., Butler, R. P., Vogt, S. S., & Apps, K. 1999, *PASP*, 111, 50
- Fischer, D. A., Valenti, J. A., & Marcy, G. W. 2003, in *ASP Conf. Ser.: IAU S219: Stars as Suns: Activity, Evolution and Planets*, Ed. A. Dupree
- Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, *PASP*, 128, 066001
- Ford, E. B. 2006, *ApJ*, 642, 505
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
- Fressin, F., Torres, G., Charbonneau, D., et al. 2013, *ApJ*, 766, 81
- Fuhrmann, K., Pfeiffer, M. J., & Bernkopf, J. 1998, *A&A*, 336, 942
- Fulton, B. J., Shporer, A., Winn, J. N., et al. 2011, *The Astronomical Journal*, 142, 84
- Fulton, B. J., Howard, A. W., Winn, J. N., et al. 2013, *ApJ*, 772, 80
- Fulton, B. J., Tonry, J. L., Flewelling, H., et al. 2014, *ApJ*, 796, 114
- Fulton, B. J., Collins, K. A., Gaudi, B. S., et al. 2015a, *ApJ*, 810, 30
- Fulton, B. J., Weiss, L. M., Sinukoff, E., et al. 2015b, *ApJ*, 805, 175
- Fulton, B. J., Howard, A. W., Weiss, L. M., et al. 2016, *ApJ*, 830, 46

- Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, ArXiv e-prints, arXiv:1703.10375
- Furlan, E., Ciardi, D. R., Everett, M. E., et al. 2017, The Astronomical Journal, 153, 71
- Gardner, J., Mather, J., Clampin, M., et al. 2006, Space Science Reviews, 123, 485
- Gatewood, G., Han, I., & Black, D. C. 2001, ApJ, 548, L61
- Gaudi, B. S. 2012, ARA&A, 50, 411
- Gautier, III, T. N., Charbonneau, D., Rowe, J. F., et al. 2012, ApJ, 749, 15
- Gellera, D. 1984, Double Star Section Circulars, 3, 3
- . 1990, Astronomische Nachrichten, 311, 117
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2003, Bayesian Data Analysis, 2nd edn. (Chapman and Hall)
- Ghezzi, L., Cunha, K., Smith, V. V., et al. 2010, ApJ, 720, 1290
- Gilliland, R. L., Cartier, K. M. S., Adams, E. R., et al. 2015, AJ, 149, 24
- Ginzburg, S., & Sari, R. 2015, ArXiv e-prints, arXiv:1501.02087
- Gladman, B. 1993, Icarus, 106, 247
- González Hernández, J. I., Israelian, G., Santos, N. C., et al. 2010, ApJ, 720, 1592
- Gould, A., Dorsher, S., Gaudi, B. S., & Udalski, A. 2006, Acta Astron., 56, 1
- Gray, R. O., Corbally, C. J., Garrison, R. F., McFadden, M. T., & Robinson, P. E. 2003, AJ, 126, 2048
- Hadden, S., & Lithwick, Y. 2014, ApJ, 787, 80
- . 2016, ArXiv e-prints, arXiv:1611.03516
- Hall, J. C., Henry, G. W., Lockwood, G. W., Skiff, B. A., & Saar, S. H. 2009, AJ, 138, 312
- Han, E., Wang, S. X., Wright, J. T., et al. 2014, PASP, 126, 827
- Hansen, B. M. S., & Murray, N. 2012, ApJ, 751, 158
- Hartigan, J. A., & Hartigan, P. M. 1985, Ann. Statist., 13, 70
- Hartkopf, W. I., Mason, B. D., Finch, C. T., et al. 2013a, AJ, 146, 76
- . 2013b, AJ, 146, 76

- Hartman, J. D., Gaudi, B. S., Holman, M. J., et al. 2008, *ApJ*, 675, 1233
- Hartman, J. D., Bakos, G. Á., Béky, B., et al. 2012, *AJ*, 144, 139
- Hébrard, G., Lecavelier Des Étangs, A., Vidal-Madjar, A., Désert, J.-M., & Ferlet, R. 2004, in Astronomical Society of the Pacific Conference Series, Vol. 321, Extrasolar Planets: Today and Tomorrow, ed. J. Beaulieu, A. Lecavelier Des Etangs, & C. Terquem, 203
- Henry, G. W. 1999, *PASP*, 111, 845
- Henry, G. W., Kane, S. R., Wang, S. X., et al. 2013, *ApJ*, 768, 155
- Hirano, T., Suto, Y., Winn, J. N., et al. 2011, *ApJ*, 742, 69
- Hirano, T., Narita, N., Sato, B., et al. 2012, *ApJ*, 759, L36
- Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, *A&A*, 355, L27
- Hogan, E., Burleigh, M. R., & Clarke, F. J. 2009, *MNRAS*, 396, 2074
- Holman, M. J., & Murray, N. W. 2005, *Science*, 307, 1288
- Holman, M. J., Winn, J. N., Latham, D. W., et al. 2006, *ApJ*, 652, 1715
- Holman, M. J., Fabrycky, D. C., Ragozzine, D., et al. 2010, *Science*, 330, 51
- Horch, E. P., Gomez, S. C., Sherry, W. H., et al. 2011, *AJ*, 141, 45
- Horch, E. P., Howell, S. B., Everett, M. E., & Ciardi, D. R. 2012, *AJ*, 144, 165
- . 2014, *ApJ*, 795, 60
- Horch, E. P., Veillette, D. R., Baena Gallé, R., et al. 2009, *AJ*, 137, 5057
- Horne, J. H., & Baliunas, S. L. 1986, *ApJ*, 302, 757
- Howard, A. W., & Fulton, B. J. 2016, ArXiv e-prints, arXiv:1606.03134
- Howard, A. W., Johnson, J. A., Marcy, G. W., et al. 2009, *ApJ*, 696, 75
- . 2010, *ApJ*, 721, 1467
- Howard, A. W., Marcy, G. W., Johnson, J. A., et al. 2010a, *Science*, 330, 653
- . 2010b, *Science*, 330, 653
- Howard, A. W., Johnson, J. A., Marcy, G. W., et al. 2011a, *ApJ*, 726, 73
- . 2011b, *ApJ*, 726, 73

- . 2011c, *ApJ*, 730, 10
- Howard, A. W., Bakos, G. Á., Hartman, J., et al. 2012a, *ApJ*, 749, 134
- Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012b, *ApJS*, 201, 15
- Howard, A. W., Sanchis-Ojeda, R., Marcy, G. W., et al. 2013, *Nature*, 503, 381
- Howard, A. W., Marcy, G. W., Fischer, D. A., et al. 2014, *ApJ*, 794, 51
- Howell, S. B., Everett, M. E., Sherry, W., Horch, E., & Ciardi, D. R. 2011, *AJ*, 142, 19
- Huber, D., Bedding, T. R., Stello, D., et al. 2011, *ApJ*, 743, 143
- Huber, D., Ireland, M. J., Bedding, T. R., et al. 2012, *ApJ*, 760, 32
- Huber, D., Silva Aguirre, V., Matthews, J. M., et al. 2014, *ApJS*, 211, 2
- Ida, S., & Lin, D. N. C. 2004, *ApJ*, 604, 388
- . 2008, *ApJ*, 673, 487
- . 2010, *ApJ*, 719, 810
- Inamdar, N. K., & Schlichting, H. E. 2016, *ApJ*, 817, L13
- Ireland, M. J., Mérand, A., ten Brummelaar, T. A., et al. 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7013, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 24
- Irwin, J., & Bouvier, J. 2009, in IAU Symposium, Vol. 258, IAU Symposium, ed. E. E. Mamajek, D. R. Soderblom, & R. F. G. Wyse, 363–374
- Isaacson, H., & Fischer, D. 2010, *The Astrophysical Journal*, 725, 875
- Ivezić, Ž., Beers, T. C., & Jurić, M. 2012, *ARA&A*, 50, 251
- Ivezic, Z., Tyson, J. A., Acosta, E., et al. 2008, ArXiv e-prints, arXiv:0805.2366
- Jenkins, J. M., Chandrasekaran, H., McCauliff, S. D., et al. 2010, in Proc. SPIE, Vol. 7740, Software and Cyberinfrastructure for Astronomy, 77400D
- Jensen, E. 2013, Tapir: A web interface for transit/eclipse observability, *Astrophysics Source Code Library*, , , ascl:1306.007
- Jin, S., & Mordasini, C. 2017, ArXiv e-prints, arXiv:1706.00251
- Jin, S., Mordasini, C., Parmentier, V., et al. 2014, *ApJ*, 795, 65

- Johnson, J. A., Marcy, G. W., Fischer, D. A., et al. 2006, ApJ, 647, 600
- Johnson, J. A., Fischer, D. A., Marcy, G. W., et al. 2007, ApJ, 665, 785
- Johnson, J. A., Clanton, C., Howard, A. W., et al. 2011, ApJS, 197, 26
- Johnson, J. A., Petigura, E. A., Fulton, B. J., et al. 2017, ArXiv e-prints, arXiv:1703.10402
- Joshi, Y. C., Pollacco, D., Collier Cameron, A., et al. 2009, MNRAS, 392, 1532
- Jura, M., Farihi, J., & Zuckerman, B. 2009, AJ, 137, 3191
- Kaiser, N., Burgett, W., Chambers, K., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7733, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
- Kallinger, T., Weiss, W. W., Barban, C., et al. 2010, A&A, 509, A77
- Kammer, J. A., Knutson, H. A., Howard, A. W., et al. 2014, ApJ, 781, 103
- Kane, S. R., Howell, S. B., Horch, E. P., et al. 2014, ApJ, 785, 93
- Kharchenko, N. V. 2001, Kinematika i Fizika Nebesnykh Tel, 17, 409
- Kibrick, R. I., Clarke, D. A., Deich, W. T. S., & Tucker, D. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6274, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1
- Kilic, M., Gould, A., & Koester, D. 2009, ApJ, 705, 1219
- Kipping, D. M., Nesvorný, D., Buchhave, L. A., et al. 2014, ApJ, 784, 28
- Kipping, D. M., Hartman, J., Bakos, G. Á., et al. 2011, ApJ, 142, 95
- Kjeldsen, H., & Bedding, T. R. 1995, A&A, 293, 87
- Knutson, H. A., Fulton, B. J., Montet, B. T., et al. 2014, ApJ, 785, 126
- Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, ApJ, 713, L79
- Koen, C., Kilkenny, D., van Wyk, F., & Marang, F. 2010, MNRAS, 403, 1949
- Kolbl, R., Marcy, G. W., Isaacson, H., & Howard, A. W. 2015, AJ, 149, 18
- Kolmogorov, A. 1933, Giornale dell' Istituto Italiano degli Attuari, 4, 83
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, The Astrophysical Journal, 765, 131

- Kovács, G., Bakos, G., & Noyes, R. W. 2005, MNRAS, 356, 557
- Kovács, G., Zucker, S., & Mazeh, T. 2002, A&A, 391, 369
- Kurucz, R. L. 1979, ApJS, 40, 1
- Law, N. M., Morton, T., Baranec, C., et al. 2014a, The Astrophysical Journal, 791, 35
- . 2014b, ApJ, 791, 35
- Lee, E. J., & Chiang, E. 2015, ArXiv e-prints, arXiv:1510.08855
- . 2016, ApJ, 817, 90
- Lee, E. J., Chiang, E., & Ormel, C. W. 2014, ApJ, 797, 95
- Léger, A., Rouan, D., Schneider, J., et al. 2009, A&A, 506, 287
- Liddle, A. R. 2007, MNRAS, 377, L74
- Lillo-Box, J., Barrado, D., & Bouy, H. 2012, A&A, 546, A10
- . 2014, A&A, 566, A103
- Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. 1996, Nature, 380, 606
- Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011, ApJS, 197, 8
- Liu, S.-F., Hori, Y., Lin, D. N. C., & Asphaug, E. 2015, ApJ, 812, 164
- Lockwood, A. C., Johnson, J. A., Bender, C. F., et al. 2014, ApJ, 783, L29
- Lockwood, G. W., Henry, G. W., Hall, J. C., & Radick, R. R. 2013, in Astronomical Society of the Pacific Conference Series, Vol. 472, New Quests in Stellar Astrophysics III: A Panchromatic View of Solar-Like Stars, With and Without Planets, ed. M. Chavez, E. Bertone, O. Vega, & V. De la Luz, 203
- Lomb, N. R. 1976, Ap&SS, 39, 447
- Lopez, E. D. 2014, PhD thesis, University of California, Santa Cruz
- Lopez, E. D., & Fortney, J. J. 2013, The Astrophysical Journal, 776, 2
- Lopez, E. D., & Fortney, J. J. 2014, ApJ, 792, 1
- Lopez, E. D., & Rice, K. 2016, ArXiv e-prints, arXiv:1610.09390
- Lovis, C., Mayor, M., Pepe, F., et al. 2006, Nature, 441, 305

- Lovis, C., Ségransan, D., Mayor, M., et al. 2011, A&A, 528, A112
- Lucy, L. B., & Sweeney, M. A. 1971, AJ, 76, 544
- Lundkvist, M. S., Kjeldsen, H., Albrecht, S., et al. 2016, Nature Communications, 7, 11201
- Lupton, W. F., & Conrad, A. R. 1993, in Astronomical Society of the Pacific Conference Series, Vol. 52, Astronomical Data Analysis Software and Systems II, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes, 315
- Magnier, E. A., Schlafly, E., Finkbeiner, D., et al. 2013, ApJS, 205, 20
- Mandel, K., & Agol, E. 2002, ApJ, 580, L171
- Mandushev, G., Torres, G., Latham, D. W., et al. 2005, ApJ, 621, 1061
- Mann, A. W., & von Braun, K. 2015, PASP, 127, 102
- Marcy, G., Butler, R. P., Fischer, D., et al. 2005, Progress of Theoretical Physics Supplement, 158, 24
- Marcy, G. W., & Butler, R. P. 1992, PASP, 104, 270
- . 1996, ApJ, 464, L147
- Marcy, G. W., Butler, R. P., Vogt, S. S., Fischer, D., & Liu, M. C. 1999, ApJ, 520, 239
- Marcy, G. W., Isaacson, H., Howard, A. W., et al. 2014, ApJS, 210, 20
- Mardling, R. A. 2010, MNRAS, 407, 1048
- Mason, B. D., Hartkopf, W. I., Raghavan, D., et al. 2011, AJ, 142, 176
- Mathur, S., Huber, D., Batalha, N. M., et al. 2016, ArXiv e-prints, arXiv:1609.04128
- Mayor, M., & Queloz, D. 1995, Nature, 378, 355
- Mayor, M., Udry, S., Lovis, C., et al. 2009a, A&A, 493, 639
- Mayor, M., Bonfils, X., Forveille, T., et al. 2009b, A&A, 507, 487
- Mayor, M., Marmier, M., Lovis, C., et al. 2011, arXiv:1109.2497, arXiv:1109.2497
- Mazeh, T., Holczer, T., & Faigler, S. 2016, A&A, 589, A75
- Medhi, B. J., Messina, S., Parihar, P. S., et al. 2007, A&A, 469, 713
- Meléndez, J., Asplund, M., Gustafsson, B., & Yong, D. 2009, ApJ, 704, L66

- Michel, E., Baglin, A., Auvergne, M., et al. 2008, *Science*, 322, 558
- Mordasini, C., Alibert, Y., & Benz, W. 2009a, *A&A*, 501, 1139
- Mordasini, C., Alibert, Y., Benz, W., & Naef, D. 2009b, *A&A*, 501, 1161
- Mordasini, C., Alibert, Y., Georgy, C., et al. 2012, *A&A*, 547, A112
- Morel, T., Rainer, M., Poretti, E., Barban, C., & Boumier, P. 2013a, *A&A*, 552, A42
- . 2013b, *A&A*, 552, A42
- Morton, T. D. 2012, *ApJ*, 761, 6
- . 2015, isochrones: Stellar model grid package, Astrophysics Source Code Library, , , ascl:1503.010
- Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, *ApJ*, 822, 86
- Morton, T. D., & Johnson, J. A. 2011, *ApJ*, 738, 170
- Morton, T. D., & Swift, J. 2014, *ApJ*, 791, 10
- Motalebi, F., Udry, S., Gillon, M., et al. 2015, *A&A*, 584, A72
- Muirhead, P. S., Johnson, J. A., Apps, K., et al. 2012, *ApJ*, 747, 144
- Mulders, G. D., Pascucci, I., & Apai, D. 2015, *ApJ*, 798, 112
- Mullally, F., Coughlin, J. L., Thompson, S. E., et al. 2015, *ApJS*, 217, 31
- Mullally, F. R. 2007, PhD thesis, The University of Texas at Austin
- Naef, D., Mayor, M., Korzennik, S. G., et al. 2003, *A&A*, 410, 1051
- Nagasawa, M., Ida, S., & Bessho, T. 2008, *ApJ*, 678, 498
- Naoz, S., Farr, W. M., Lithwick, Y., Rasio, F. A., & Teyssandier, J. 2011, *Nature*, 473, 187
- Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. 2014, LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, , , doi:10.5281/zenodo.11813
- Ngo, H., Knutson, H. A., Hinkley, S., et al. 2015, *ApJ*, 800, 138
- Noyes, R. W., Jha, S., Korzennik, S. G., et al. 1997, *ApJ*, 483, L111
- Nutzman, P., & Charbonneau, D. 2008, *PASP*, 120, 317
- Ohta, Y., Taruya, A., & Suto, Y. 2005, *ApJ*, 622, 1118

- O'Toole, S. J., Jones, H. R. A., Tinney, C. G., et al. 2009, ApJ, 701, 1732
- Owen, J. E., & Wu, Y. 2013, ApJ, 775, 105
- . 2017, ArXiv e-prints, arXiv:1705.10810
- Penev, K., Zhang, M., & Jackson, B. 2014, PASP, 126, 553
- Pepe, F., Mayor, M., Queloz, D., et al. 2004, A&A, 423, 385
- Pepe, F., Lovis, C., Ségransan, D., et al. 2011, A&A, 534, A58
- Pepe, F., Cameron, A. C., Latham, D. W., et al. 2013, Nature, 503, 377
- Pepper, J., Pogge, R. W., DePoy, D. L., et al. 2007, PASP, 119, 923
- Pepper, J., Siverd, R. J., Beatty, T. G., et al. 2013, ApJ, 773, 64
- Perryman, M. A. C., & ESA. 1997, The HIPPARCOS and TYCHO catalogues. Astrometric and photometric star catalogues derived from the ESA HIPPARCOS Space Astrometry Mission (The Hipparcos and Tycho catalogues. Astrometric and photometric star catalogues derived from the ESA Hipparcos Space Astrometry Mission, Publisher: Noordwijk, Netherlands: ESA Publications Division, 1997, Series: ESA SP Series vol no: 1200, ISBN: 9290923997 (set))
- Petigura, E. A. 2015, arXiv:1510.03902, arXiv:1510.03902
- Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013a, Proceedings of the National Academy of Science, 110, 19273
- . 2013b, Proceedings of the National Academy of Science, 110, 19273
- Petigura, E. A., Marcy, G. W., & Howard, A. W. 2013c, ApJ, 770, 69
- Petigura, E. A., Howard, A. W., Marcy, G. W., et al. 2017, ArXiv e-prints, arXiv:1703.10400
- Pickles, A., & Depagne, É. 2010, PASP, 122, 1437
- Pickles, A. J. 1998, PASP, 110, 863
- Pollacco, D. L., Skillen, I., Collier Cameron, A., et al. 2006, PASP, 118, 1407
- Pourteau, M. A. 1933, Observatoire de Paris. Catalogue des étoiles doubles de la zone + 24° de la carte photographique du ciel (Orléans: impr. de H. Tessier)
- Press, W. H., & Rybicki, G. B. 1989, ApJ, 338, 277
- Queloz, D., Henry, G. W., Sivan, J. P., et al. 2001, A&A, 379, 279

- Radovan, M. V., Cabak, G. F., Laiterman, L. H., Lockwood, C. T., & Vogt, S. S. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 4
- Radovan, M. V., Lanclos, K., Holden, B. P., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9145, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2
- Raghavan, D., Farrington, C. D., ten Brummelaar, T. A., et al. 2012, *ApJ*, 745, 24
- Ramírez, I., Fish, J. R., Lambert, D. L., & Allende Prieto, C. 2012, *ApJ*, 756, 46
- Ramírez, I., Meléndez, J., Bean, J., et al. 2014, *A&A*, 572, A48
- Rauer, H., Catala, C., Aerts, C., et al. 2014, *Experimental Astronomy*, 38, 249
- Raymond, S. N., & Cossou, C. 2014, *MNRAS*, 440, L11
- Reffert, S., & Quirrenbach, A. 2011, *A&A*, 527, A140
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9143, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 20
- Rivera, E. J., Butler, R. P., Vogt, S. S., et al. 2010, *ApJ*, 708, 1492
- Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, *A&A*, 409, 523
- Rogers, L. A. 2014, ArXiv:1407.4457, arXiv:1407.4457
- . 2015, *ApJ*, 801, 41
- Rowe, J. F., Bryson, S. T., Marcy, G. W., et al. 2014, *ApJ*, 784, 45
- Sanchis-Ojeda, R., Rappaport, S., Winn, J. N., et al. 2014, *ApJ*, 787, 47
- Santos, N. C., Mayor, M., Naef, D., et al. 2002, *A&A*, 392, 215
- Santos, N. C., Sousa, S. G., Mortier, A., et al. 2013, *A&A*, 556, A150
- Scargle, J. D. 1982, *ApJ*, 263, 835
- Schlafly, E. F., Finkbeiner, D. P., Jurić, M., et al. 2012, *ApJ*, 756, 158
- Schlaufman, K. C. 2010, *ApJ*, 719, 602
- Schlichting, H. E. 2014, *ApJ*, 795, L15

- Schlichting, H. E., Sari, R., & Yalinewich, A. 2015, Icarus, 247, 81
- Scholz, F. W., & Stephens, M. A. 1987, Journal of the American Statistical Association, 82, 918
- Schwarzenberg-Czerny, A. 1989, MNRAS, 241, 153
- . 1998, MNRAS, 301, 831
- Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, ApJ, 669, 1279
- Shporer, A., & Brown, T. 2011, ApJ, 733, 30
- Silburt, A., Gaidos, E., & Wu, Y. 2015, ApJ, 799, 180
- Sing, D. K., Wakeford, H. R., Showman, A. P., et al. 2015, MNRAS, 446, 2428
- Sinukoff, E., Fulton, B., Scuderi, L., & Gaidos, E. 2013, Space Sci. Rev., 180, 71
- Siverd, R. J., Beatty, T. G., Pepper, J., et al. 2012, ApJ, 761, 123
- Smirnov, N. 1948, Ann. Math. Statist., 19, 279
- Snellen, I. A. G., Brandl, B. R., de Kok, R. J., et al. 2014, Nature, 509, 63
- Southworth, J., Hinse, T. C., Dominik, M., et al. 2012, MNRAS, 426, 1338
- Spearman, C. 1904, The American Journal of Psychology, 15, pp. 72
- Spiegel, D. S., & Madhusudhan, N. 2012, ApJ, 756, 132
- Steele, P. R., Burleigh, M. R., Farihi, J., et al. 2009, A&A, 500, 1207
- Steffen, J. H., Ragozzine, D., Fabrycky, D. C., et al. 2012, Proceedings of the National Academy of Science, 109, 7982
- Stetson, P. B. 1987, PASP, 99, 191
- . 1990, PASP, 102, 932
- Swift, J. J., Bottom, M., Johnson, J. A., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 027002
- Takeda, G., Ford, E. B., Sills, A., et al. 2007, ApJS, 168, 297
- Tanner, A., White, R., Bailey, J., et al. 2012, ApJS, 203, 10
- ten Brummelaar, T. A., McAlister, H. A., Ridgway, S. T., et al. 2005a, ApJ, 628, 453
- . 2005b, ApJ, 628, 453

- Ter Braak, C. 2006, Statistics and Computing
- Tonry, J. L. 2011, PASP, 123, 58
- Tonry, J. L., Stubbs, C. W., Kilic, M., et al. 2012, ApJ, 745, 42
- Torres, G., Andersen, J., & Giménez, A. 2010, A&A Rev., 18, 67
- Torres, G., Konacki, M., Sasselov, D. D., & Jha, S. 2005, ApJ, 619, 558
- Torres, G., Bakos, G. Á., Kovács, G., et al. 2007, ApJ, 666, L121
- Traub, W. A. 2012, ApJ, 745, 20
- Valenti, J. A., & Fischer, D. A. 2005, ApJS, 159, 141
- Valenti, J. A., & Piskunov, N. 1996, A&AS, 118, 595
- van Belle, G. T., & von Braun, K. 2009, ApJ, 694, 1085
- van Leeuwen, F. 2007, A&A, 474, 653
- Vanderburg, A., Johnson, J. A., Rappaport, S., et al. 2015, Nature, 526, 546
- Veras, D., & Ford, E. B. 2012, MNRAS, 420, L23
- Vogt, S. S., Butler, R. P., Marcy, G. W., et al. 2005, ApJ, 632, 638
- Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 2198, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. D. L. Crawford & E. R. Craine, 362
- Vogt, S. S., Wittenmyer, R. A., Butler, R. P., et al. 2010, ApJ, 708, 1366
- Vogt, S. S., Radovan, M., Kibrick, R., et al. 2014a, PASP, 126, 359
- . 2014b, PASP, 126, 359
- von Braun, K., Boyajian, T. S., van Belle, G. T., et al. 2014, MNRAS, 438, 2413
- Wang, J., Fischer, D. A., Horch, E. P., & Xie, J.-W. 2015a, ApJ, 806, 248
- Wang, J., Fischer, D. A., Xie, J.-W., & Ciardi, D. R. 2015b, ApJ, 813, 130
- Weiss, L. M., & Marcy, G. W. 2014, ApJ, 783, L6
- Weiss, L. M., Marcy, G. W., Rowe, J. F., et al. 2013, ApJ, 768, 14
- Weiss, L. M., Deck, K., Sinukoff, E., et al. 2016, ArXiv e-prints, arXiv:1612.04856

- White, T. R., Huber, D., Maestro, V., et al. 2013, MNRAS, 433, 1262
- Winn, J. N. 2010, ArXiv e-prints, arXiv:1001.2010
- Winn, J. N., Fabrycky, D., Albrecht, S., & Johnson, J. A. 2010a, ApJ, 718, L145
- Winn, J. N., Johnson, J. A., Albrecht, S., et al. 2009, ApJ, 703, L99
- Winn, J. N., Johnson, J. A., Howard, A. W., et al. 2010b, ApJ, 718, 575
- Wolfgang, A., & Lopez, E. 2015, ApJ, 806, 183
- Wright, J. T., & Eastman, J. D. 2014, PASP, 126, 838
- Wright, J. T., & Howard, A. W. 2009, ApJS, 182, 205
- Wright, J. T., Marcy, G. W., Howard, A. W., et al. 2012, ApJ, 753, 160
- Wright, J. T., Upadhyay, S., Marcy, G. W., et al. 2009, ApJ, 693, 1084
- Wright, J. T., Marcy, G. W., Fischer, D. A., et al. 2007, ApJ, 657, 533
- Wright, J. T., Fakhouri, O., Marcy, G. W., et al. 2011, PASP, 123, 412
- Wu, Y., & Lithwick, Y. 2013, ApJ, 772, 74
- Xu, S., Jura, M., Koester, D., Klein, B., & Zuckerman, B. 2014, ApJ, 783, 79
- Youdin, A. N. 2011, ApJ, 742, 38
- Zahn, J.-P. 2008, in EAS Publications Series, Vol. 29, EAS Publications Series, ed. M.-J. Goupil & J.-P. Zahn, 67–90
- Zhao, M., O'Rourke, J. G., Wright, J. T., et al. 2014, ApJ, 796, 115
- Zucker, S., & Mazeh, T. 2001, arXiv, astro-ph/0104098, astro-ph/0104098
- Zuckerman, B., & Becklin, E. E. 1992, ApJ, 386, 260
- Zuckerman, B., Melis, C., Klein, B., Koester, D., & Jura, M. 2010, ApJ, 722, 725