

UCLA

UCLA Electronic Theses and Dissertations

Title

The Orbital Eccentricities of Small Planets

Permalink

<https://escholarship.org/uc/item/4fg33458>

Author

MacDougall, Mason Gordon

Publication Date

2023

Supplemental Material

<https://escholarship.org/uc/item/4fg33458#supplemental>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

**The Orbital Eccentricities of Small Planets**

A dissertation submitted in partial satisfaction  
of the requirements for the degree  
Doctor of Philosophy in Astronomy and Astrophysics

by

Mason Gordon MacDougall

2023

© Copyright by  
Mason Gordon MacDougall  
2023

# ABSTRACT OF THE DISSERTATION

## **The Orbital Eccentricities of Small Planets**

by

Mason Gordon MacDougall

Doctor of Philosophy in Astronomy and Astrophysics

University of California, Los Angeles, 2023

Professor Erik A. Petigura, Chair

Planetary orbital eccentricities are key to probing the formation and evolution pathways of exoplanets. Eccentricity measurements can be made through a variety of observing techniques that currently tend to be limited by signal-to-noise constraints, data quality limitations, or model biases, often leading to poorly constrained eccentricities for sub-Jovian-size planets. To better understand the typical and extreme values of planetary eccentricities for planets of all sizes, I investigated the photometric eccentricities of TESS planets and accurately constrained the eccentricity distributions of both individual planets and planet sub-populations. I used existing methods to identify several high-eccentricity sub-Jovian-size planets from TESS transit photometry and confirmed their eccentricities via follow-up radial velocity measurements (Chapters 2 and 3). Through these discoveries, I identified an unaccounted-for bias in a common transit modeling method which unfairly skewed photometric eccentricity constraints towards higher values. I worked with a small team of collaborators to mitigate this bias and proposed an alternative transit model parameterization that yields accurate transit fits and unbiased eccentricity constraints (Chapters 4 and 5).

I used our proposed modeling approach to homogeneously measure the transit properties

of 108 planets from the TESS-Keck Survey (TKS), including constraints on orbital periods, transit-timing variations, and planet-to-star radius ratios (Chapter 6). In addition to measuring these transit properties, I also performed precise stellar characterization that allowed me to measure planet radii and constrain orbital eccentricities via importance sampling with stellar density. These homogeneously-constrained posterior distributions of eccentricity from my photometric modeling served as the foundation for my hierarchical Bayesian analysis of the population-level eccentricity distribution of the TKS planet sample – a first among TESS planets (Chapter 7). Through this analysis, I found that sub-Jovian-size planets display a lower underlying eccentricity distribution than Jovian-size planets – the first confirmation of this trend shown via hierarchical Bayesian modeling. I found a similar distinction between the eccentricities of planets in single-planet systems versus multi-planet systems, with the latter displaying lower eccentricities. I also demonstrated that the results of such hierarchical analyses can be used to improve the individual measurement precisions of planet radii, impact parameters, and eccentricities. The parameter constraints and dynamical trends revealed through this hierarchical analysis of exoplanet eccentricities will help improve our understanding of the dynamical processes that drive the evolution of observed planetary systems.

The dissertation of Mason Gordon MacDougall is approved.

Tuan H. Do

Jean-Luc Margot

Smadar Naoz

Erik A. Petigura, Committee Chair

University of California, Los Angeles

2023

*For my family.*

# TABLE OF CONTENTS

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background and Motivation	2
1.1.1	Exoplanet Dynamics	2
1.1.2	Constraining Eccentricity	2
1.2	Structure of the dissertation	4
<b>2</b>	<b>The TESS-Keck Survey. VI. Two Eccentric sub-Neptunes Orbiting HIP-97166</b>	<b>5</b>
2.1	Abstract	5
2.2	Introduction	6
2.3	HIP-97166b: A High-Eccentricity Candidate	8
2.3.1	TESS Photometry	8
2.3.2	Photometric Transit Model	9
2.3.3	Eccentricity Constraints from Photometry	13
2.4	Spectroscopic Follow-up	14
2.4.1	HIRES RVs	14
2.4.2	APF RVs	14
2.5	Stellar Characterization	16
2.6	Keplerian Modeling	17
2.6.1	RV Detection of Planets b and c	17
2.6.2	RV-only Constraints	20
2.6.3	RV-Photometry Joint Model	21
2.7	System Dynamics	26
2.7.1	Eccentricity Constraints from Stability Requirements	26
2.7.2	Secular Eccentricity Variability	27

2.7.3	TTV Analysis . . . . .	29
2.8	System in Context . . . . .	29
2.8.1	Bulk Density and Core-Envelope Fraction . . . . .	29
2.8.2	Weak Eccentricity Damping . . . . .	30
2.8.3	Sub-Neptune Eccentricity . . . . .	31
2.9	Formation Scenarios . . . . .	32
2.10	Summary and Conclusions . . . . .	34
<b>3</b>	<b>The TESS-Keck Survey. XIII. An Eccentric Hot Neptune with a Similar-</b>	
	<b>Mass Outer Companion around TOI-1272 . . . . .</b>	<b>36</b>
3.1	Abstract . . . . .	36
3.2	Introduction . . . . .	37
3.3	TOI-1272b: A High-Eccentricity Candidate . . . . .	39
3.3.1	Photometry . . . . .	39
3.3.2	Photometric Transit Model . . . . .	40
3.3.3	Eccentricity Constraints from Photometry . . . . .	42
3.4	Spectroscopic Follow-up . . . . .	44
3.5	Stellar Characterization . . . . .	46
3.5.1	Bulk Properties . . . . .	46
3.5.2	Variability and Rotation . . . . .	46
3.5.3	Age . . . . .	48
3.6	Keplerian Modeling . . . . .	49
3.6.1	RV Detection of Planets b and c . . . . .	49
3.6.2	Additional RV Signals . . . . .	50
3.6.3	RV-only Model . . . . .	52
3.6.4	RV-Photometry Joint Model . . . . .	52
3.7	System Dynamics . . . . .	57
3.7.1	Eccentricity Constraints from Stability Requirements . . . . .	57

3.7.2	No Evidence for TTVs . . . . .	59
3.7.3	Strong Tidal Eccentricity Decay . . . . .	60
3.8	Context in the Hot Neptune Desert . . . . .	62
3.8.1	Bulk Density and Core-Envelope Fraction . . . . .	62
3.8.2	Eccentricities . . . . .	64
3.8.3	A Unique Formation and Evolution Pathway . . . . .	67
3.9	Summary and Conclusions . . . . .	68
<b>4</b>	<b>Implicit biases in transit models using stellar pseudo-density . . . . .</b>	<b>69</b>
4.1	Abstract . . . . .	69
4.2	Introduction . . . . .	69
4.3	Understanding parameter biases with fits to synthetic photometry . . . . .	73
4.4	Numerical sampling experiment . . . . .	76
4.4.1	Experimental setup . . . . .	77
4.4.2	Bias on impact parameter . . . . .	79
4.5	Mathematical origin of the bias . . . . .	81
4.6	Biased Kepler planet properties . . . . .	82
4.7	Summary and conclusions . . . . .	87
4.8	Appendix 4A: Derivation of Jacobian for $T_{14} \rightarrow \tilde{\rho}$ . . . . .	87
4.9	Appendix 4B: Derivation of Jacobian for $T_{14} \rightarrow \ln \tilde{\rho}$ . . . . .	89
4.10	Appendix 4C: Derivation of Jacobian for $T_{14} \rightarrow a/R_{\star}$ . . . . .	89
4.11	Appendix 4D: Derivation of Jacobian for $T_{14} \rightarrow \ln a/R_{\star}$ . . . . .	90
<b>5</b>	<b>Accurate and efficient photo-eccentric transit modeling . . . . .</b>	<b>91</b>
5.1	Abstract . . . . .	91
5.2	Introduction . . . . .	92
5.3	Synthetic Lightcurve Construction . . . . .	94
5.4	Method #1: Direct sampling in $e-\omega-\rho$ . . . . .	99

5.5	Method #2: Direct sampling in $T_{14}$ , then importance sampling in $e-\omega-\rho$ . . .	100
5.5.1	Transit fitting . . . . .	101
5.5.2	Importance sampling . . . . .	102
5.6	Results . . . . .	104
5.6.1	Both methods return equivalent eccentricity constraints . . . . .	104
5.6.2	Both methods return accurate results . . . . .	106
5.6.3	Our $T_{14} + umb$ method is more efficient than the $e-\omega-\rho$ method . . .	108
5.7	Conclusions . . . . .	111
5.8	Appendix 5A: Sampler comparison: NUTS vs. Nested Sampling . . . . .	112
5.8.1	Posterior comparison . . . . .	113
5.8.2	Accuracy . . . . .	113
5.8.3	Efficiency . . . . .	114

**6 The TESS-Keck Survey. XV. Precise Properties of 108 TESS Planets and Their Host Stars . . . . . 117**

6.1	Abstract . . . . .	117
6.2	Introduction . . . . .	118
6.3	Stellar Properties . . . . .	120
6.3.1	Spectroscopic Constraints . . . . .	120
6.3.2	Isochrone Modeling . . . . .	120
6.3.3	Limb Darkening . . . . .	126
6.4	Light Curve Modeling . . . . .	128
6.4.1	Input Data . . . . .	128
6.4.2	False-Positive Vetting . . . . .	128
6.4.3	Lightcurve Pre-Processing . . . . .	129
6.4.4	Estimating TTVs . . . . .	131
6.4.5	Full Lightcurve Model . . . . .	133
6.5	Planet Properties . . . . .	135

6.5.1	Importance-Weighted Posterior Distributions . . . . .	135
6.5.2	Planet-to-star Radius Ratio . . . . .	137
6.5.3	Derived Planet Properties . . . . .	139
6.6	Discussion . . . . .	141
6.6.1	Radius Gap . . . . .	141
6.6.2	Hot Neptune Desert . . . . .	143
6.6.3	Multis and TTVs . . . . .	143
6.6.4	Eccentricity . . . . .	146
6.7	Conclusions . . . . .	147
<b>7</b>	<b>Distribution of Orbital Eccentricities for 93 Small Planets and 15 Giant Planets from the TESS-Keck Survey . . . . .</b>	<b>158</b>
7.1	Abstract . . . . .	158
7.2	Introduction . . . . .	159
7.3	Framework for Hierarchical Bayesian Analysis . . . . .	162
7.4	A Hierarchical Analysis of the Full TKS Sample . . . . .	166
7.4.1	Measuring the true eccentricity distribution of TKS planets . . . . .	166
7.4.2	Improving $R_p/R_\star$ measurements via HBM-informed importance sampling . . . . .	171
7.5	A Hierarchical Analysis of Sub-populations of the TKS sample . . . . .	175
7.5.1	Planet Multiplicity-based Sub-populations . . . . .	175
7.5.2	Stellar Metallicity-based Sub-populations . . . . .	178
7.5.3	Planet Radius-based Sub-populations . . . . .	181
7.6	Conclusions . . . . .	183
<b>8</b>	<b>Conclusions . . . . .</b>	<b>185</b>

## LIST OF FIGURES

2.1	TESS photometry of HIP-97166. The data has been processed following the procedures discussed in §2.3.1. Transits of HIP-97166b are indicated by blue triangles and are 0.8 ppt. . . . .	9
2.2	Transit duration as a function of $e$ and $\omega$ . (a) Observer (blue triangle) viewing a planet transiting on a circular orbit (blue light curve). (b) Observer (orange triangle) viewing a planet transiting at periastron (orange light curve). (c) Observer (green triangle) viewing a planet transiting at apastron (green light curve). (d) Simulated light curves associated with scenarios a-c for a Jupiter-size planet orbiting edge-on around a Sun-like star with a period of 5 days. . . . .	10
2.3	Top: Transit models (orange; 50 samples) drawn from parameter posterior distributions fit from phase-folded TESS photometry of HIP-97166b. A simulated transit curve (blue) is shown for a theoretical circular orbit of HIP-97166b, modeled using median posterior distribution values for all other parameters. Details regarding fitting procedure are discussed in §2.3.2. Bottom: Residuals to our maximum <i>a posteriori</i> model. . . . .	12
2.4	2D joint posterior distribution of $e$ and $\omega$ for HIP-97166b, showing $1\text{-}\sigma$ and $2\text{-}\sigma$ credibility intervals. Best-fit values for both parameters are given: $e = 0.68^{+0.18}_{-0.16}$ and $\omega = 71^{\circ+81^{\circ}-48^{\circ}}$ . . . . .	13
2.5	Iterative Keplerian periodogram search of HIP-97166 RVs using <i>RVSearch</i> , confirming the 10.3-day transiting planet and identifying a significant non-transiting companion with a 16.8-day period. $\Delta\text{BIC}$ is used to discriminate between models with Keplerians at varying periods (Bayesian Information Criterion; Schwarz 1978), corresponding to a significance threshold of $\text{FAP} = 0.001$ at the yellow dashed line. Monthly and yearly aliases are shown with in green and red dashed lines, respectively. . . . .	18

2.6	Spearman rank-order linear correlation fit between 1-planet RV residuals and $S_{\text{HK}}$ activity metric shows a tenuous trend with a low correlation score $\rho_{\text{spear}}$ , suggesting that no significant relationship exists between the observed RV signals and stellar activity. . . . .	20
2.7	a) Best-fit model (blue) of radial velocity measurements from Keck/HIRES (black) and APF/Levy (green) using <code>Radvel</code> (Fulton et al. 2018) with no binning; b) Residuals to best-fit RV model; c) Phase-folded view of best-fit model and RV data for HIP-97166b, with binned points shown in red; d) Phase-folded view of best-fit model and RV data for HIP-97166c . . . . .	22
2.8	Posterior distributions of parameters from two-planet RV-only model with <code>RadVel</code> .	23
2.9	Distribution of initialized $e_b-e_c$ values for all <code>rebound</code> simulations (grey) of the HIP-97166 system. Blue-green contours show regions with the highest density of stable eccentricity configurations (simulations lasting $10^4$ years). . . . .	27
2.10	Three different dynamical scenarios of HIP-97166 b and c eccentricities and TTVs simulated using <code>rebound</code> and sampled from our previous results as described in §2.7.2. Rows: Stable eccentric scenario with negligible TTVs (Top). Stable eccentric scenario with periods oscillating about 5:3 MMR, displaying non-negligible TTVs (Middle). Unstable eccentric scenario lasting only $\sim 800$ years with moderate TTVs (Bottom). Columns: $e$ as a function of time over final 500 years of integration (Left). $e$ as a function of $\Delta\omega$ libration (Middle). TTV amplitude as a function of transit number over the same time baseline as TESS photometry (Right). . . . .	28

- 2.11 Left: Mass-radius distribution of known sub-Neptunes (Akeson et al. 2013a) with 20% measurement precision or better in both mass and radius (grey), shown with HIP-97166b (green). Planetary interior composition curves for various 2-component models (no atmosphere) from Zeng et al. (2016) are shown as dashed colored lines. Composition curves which include  $H_2$  envelopes of varying size on top of an Earth-like core are given by faded blue dot-dash lines (Zeng et al. 2019), assuming an equilibrium temperature of 700 K. Weiss & Marcy (2014) model shown as red dotted line (note the log-log axes). Right: Eccentricity distribution of planets with  $\sigma_e < 0.1$  (sub-Neptunes in red, other known planets in gray) as a function of orbital separation, showing HIP-97166 b and c in green and blue, respectively. While the eccentricity constraints of both planets place their upper bounds towards the eccentric tail of the sub-Neptune distribution, they are consistent with overall suppressed dynamical temperatures of sub-Neptunes relative to other planet populations. Periastron distance of 0.03 AU is also shown. 33
- 3.1 Top: Transit models drawn from parameter posterior distributions (orange; 50 samples) for phase-folded TESS photometry of TOI-1272b. Expected transit shape for a circular orbit shown in blue, modeled using median posterior distribution values for all other model parameters. Details regarding fitting procedure are discussed in §3.3.2. Bottom: Residuals to our maximum *a posteriori* model. 42
- 3.2 Transit photometry for two independent single transits measured by the MuSCAT2 and MuSCAT instruments, plotted with 10 minute binning and a photometric fits by the corresponding instrument pipeline. Horizontal dashed lines indicate the expected transit depth. Top: Combined photometry from MuSCAT2  $g$ ,  $r$ ,  $i$ , and  $z_s$  bands, observed on UT 2020 February 28. Bottom: Photometry from MuSCAT  $g$ ,  $r$ , and  $z$  bands, observed on UT 2021 April 8. . . . . 43

3.3	<p>Corner plot of <code>exoplanet</code> model posteriors for TOI-1272b, showing the effects of <math>e</math>-<math>\omega</math>-<math>b</math> degeneracy on our transit fit. Best-fit value for eccentricity given by <math>e = 0.40_{-0.22}^{+0.20}</math>. Argument of periastron <math>\omega</math> remained loosely constrained about <math>90^\circ</math>, suggesting a transit near periastron. . . . .</p>	44
3.4	<p>Top: TESS light curve of TOI-1272 from sectors 15, 16, and 22. The gray points show the original TESS SAP light curve and the black points show the TESS-SIP corrected light curve that is used to extract the variability period. The vertical dashed lines mark the gaps between TESS sectors. Bottom Left: Lomb-Scargle periodogram of the TESS-SIP corrected light curve. The orbital period and transit depth from the TOI catalog are listed at the top of the panel. Bottom Right: Phase-folded light curve based on the most significant period detected from the periodogram. The gray points show the TESS-SIP corrected photometry, the black points show the binned data, and the red curve is a sinusoidal fit to the phase-folded light curve. The period, amplitude, and their associated uncertainties from the best-fit sinusoidal function are listed at the top of the panel. . . . .</p>	47
3.5	<p>Lomb-Scargle periodogram search of TOI-1272 <math>S_{\text{HK}}</math> data using, identifying a significant 28.5 day variability signal consistent with the suspected stellar rotation period. . . . .</p>	49
3.6	<p>Iterative Keplerian periodogram search of TOI-1272 RV data using <code>RVSearch</code>. We confirmed the 3.3-day transiting planet (panel a) and identified a significant 8.7-day period with no corresponding transits (panel b). <math>\Delta\text{BIC}</math> was used to discriminate between models with additional Keplerians over a grid of periods (Bayesian Information Criterion; Schwarz 1978), corresponding to a significance threshold of <math>\text{FAP} = 0.001</math> at the yellow dashed horizontal line. Panel c shows a sub-significant signal at a 14.1 day period, likely corresponding to <math>P_{\text{rot}}/2</math>. . . . .</p>	51

3.7	(a) Best-fit radial velocity model (blue) for Keck/HIRES RV measurements (black) using <code>Radvel</code> (Fulton et al. 2018) in a joint RV-photometry model via <code>juliet</code> (Espinoza et al. 2018), with corresponding residuals shown below. (b)-(c) Phase-folded views of best-fit RV model for TOI-1272 b and c, with binned points shown in red. . . . .	54
3.8	Distribution of initialized $e_b$ - $e_c$ values for all <code>rebound</code> simulations of the TOI-1272 system. Green-blue contours show regions with highest density of stable configurations (simulations lasting $10^6$ orbits of the inner planet). . . . .	59
3.9	Hot Neptune Desert in $R_p$ - $P$ and $M_p$ - $P$ parameter space, where Neptune-like is defined as $M_p < 100 M_\oplus$ and $R_p = 2.0 - 6.0 R_\oplus$ . Data drawn from the NASA Exoplanet Archive (Akeson et al. 2013a; NASA Exoplanet Archive 2022a). TOI-1272 b and c depicted as green point (both panels) and blue point (left panel), respectively. Relations dictating the upper and lower boundaries of the Hot Neptune Desert from Mazeh et al. (2016) shown as black dashed lines. Left: Mass-period distribution of all planets, with Neptune-like planets shown in red. Right: Radius-period distribution of all planets, with Neptune-like planets shown in red and $R_p = 2 R_\oplus$ limit shown as black dotted line. . . . .	63
3.10	Eccentricity distribution of planets with eccentricity uncertainties of $\sigma_e < 0.1$ , comparing shorter period planets (blue) against longer period planets (orange). The lines show kernel density estimation fits to the data of the corresponding color. Top: Jupiter-like planets ( $> 6 R_\oplus$ ). Bottom: Neptune-like planets ( $2-6 R_\oplus$ ), including TOI-1272b (dashed gray line). . . . .	65
3.11	Eccentricity distribution of planets with $\sigma_e < 0.1$ (Neptune-like in red, other known planets in gray) as a function of orbital separation, showing TOI-1272 b and c in green and blue, respectively. Periastron distance of 0.03 AU is shown as an empirical threshold for rapid tidal eccentricity decay. Three eccentric Neptunes are labeled for reference (Kepler-1656b; TOI-269b, Cointepas et al. 2021; GJ 436b). 66	66

4.1	Simulated photometry for a mini-Neptune on a circular 15 day orbit around a Sun-like star, transiting at $b = 0.5$ . The orange line indicates the ground truth transit model. Grey points show simulated observations at a one minute observing cadence; black circles are binned to 30 minutes. The white noise level was set to $S/N = 16$ , close to the <i>Kepler</i> median. See Table 4.1 for ground-truth simulation parameters. . . . .	75
4.2	Posterior corner plots when sampling in the $\tilde{\rho}$ basis (orange) vs the $\log T_{14}$ basis (purple). The bias on impact parameter, $b$ , is apparent when sampling with $\tilde{\rho}$ but is resolved when sampling in $T_{14}$ . . . . .	77
4.3	Results of the numerical sampling experiment described in §4.4. Each row corresponds to the prior, likelihood, and posterior for a given model parameterization. For visual clarity, the height of the $T_{14}$ posterior has been reduced by a factor of 3 on all plots. The difference in the prior distribution on $b$ for rows 1 & 3 compared to rows 2 & 4 stems from the use of $r$ vs $\log r$ , respectively. Sampling with a uniform prior on $\tilde{\rho}$ (top row) produces a nonuniform prior on $T_{14}$ and a biased posterior for $b$ . In contrast, sampling in any of the other parameter bases produces a posterior estimate of $b$ which matches the prior, except in cases where constraints on $r$ would produce a non-transiting orbit. . . . .	80
4.4	Posterior samples of $b$ from the simulated transit fit (orange histogram, §4.3) and the numerical experiment (purple histogram, §4.4) are nearly perfectly matched by the expected bias from the analytically derived Jacobian (grey shaded region, §4.5). . . . .	82

4.5	Posterior distributions of impact parameter for a random selection of KOIs, organized in logarithmic bins on a $P - r$ grid. Data shown are the posterior MCMC chains from <i>Kepler</i> Data Release 25 (Thompson et al. 2018; Rowe et al. 2014a; Akeson et al. 2013b). Each posterior distribution is plotted with 20% opacity so that dark regions indicate where many distributions overlap; colors correspond to the median $b$ value for a given KOI, with a maximum of twelve KOIs plotted per panel. The horizontal axis of each panel ranges over $b = (0, 1.2)$ ; the vertical range of each row is different, but the dashed line indicates the same distribution height. The median S/N in each grid square is printed in the upper right-hand corners. Most of the objects (excluding the highest S/N objects) show qualitatively similar posterior distributions of $b$ . The similarity is particularly striking for small objects. . . . .	84
4.6	Posterior samples of $b$ from a representative selection of DR25 targets before and after reweighting by the Jacobian to correct for biases induced by sampling in $\tilde{\rho}$ . All targets have $0.02 < r < 0.04$ and $10 < P < 30$ days. <i>Left panel:</i> raw DR25 posteriors chains show a clear (biased) preference for low values of $b$ . <i>Right panel:</i> after reweighting, the (unbiased) distribution is nearly flat. To minimize spurious peaks and sampling noise in low probability regions, the lowest density 1% of samples have been excluded from our reweighting scheme. The slight increase in probability density near $b \approx 1$ in the reweighted posteriors reflects the presence of residual importance sampling noise rather than a real feature of the data. Because there is significant sampling noise (due to the large implied posterior mass in regions with few samples), our preferred method for ameliorating the pseudo-density bias is to refit the photometry. . . . .	85

4.7	Fractional change in median planet-to-star radius ratio for all planet candidates after correcting posterior chains from DR25 (Thompson et al., 2018) using the Jacobian reweighting scheme described in §4.6. $5\sigma$ outliers have been iteratively clipped in order to eliminate spurious values that are expected to arise due to insufficient sampling of low probability regions. There is a spike at $\delta_r = 0$ , indicating that some subset of targets were accurately measured, but the majority of targets are distributed around $\delta_r/r = 1.6\%$ . . . . .	86
5.1	Diagram showing (bottom middle) our grid of injected $e$ and $b$ values, along with (top) a gallery of phase-folded lightcurves for all combination of SNR and injected $\omega$ values (at $e = 0.2$ and $b = 0.7$ , as an example). The four panels to the left and right of the bottom grid show demonstrative examples of phase-folded lightcurves at the different extremes of our injected parameter grid (e.g. $b = 0.1$ or $0.9$ and $e = 0.05$ or $0.8$ ), all shown with $\omega = 132^\circ$ and SNR = 40. Each lightcurve is shown with the median final transit fits from both the $e-\omega-\rho$ (orange) and $T_{14} + umb$ (blue) modeling methods, which overlap almost entirely. . . . .	96
5.2	Diagram demonstrating an overview of our modeling procedure, from an input synthetic lightcurve to output $e$ and $\omega$ constraints via both the $e-\omega-\rho$ (orange) and $T_{14} + umb$ (blue) modeling methods. . . . .	99
5.3	Comparison of $e$ values measured from the $T_{14}+umb$ and $e-\omega-\rho$ modeling methods at the 15 <sup>th</sup> , 50 <sup>th</sup> , and 85 <sup>th</sup> percentiles of their distributions. We show $\Delta e_k = \{0.05, 0.1, 0.15\}$ in grey, as well as the ideal 1-to-1 line shown in black. These comparisons generally lie close to the 1-to-1 line, implying that the results of the two models are approximately equivalent. . . . .	105

5.4	Qualitative diagram showing the relative accuracy of measuring $e$ and $b$ from transit photometry in four distinct quadrants of $e - b$ parameter space, at three different $\omega$ values. The four scenarios shown are (Top Left) low- $e$ and low- $b$ , (Bottom Left) low- $e$ and high- $b$ , (Top Right) high- $e$ and low- $b$ , and (Bottom Right) high- $e$ and high- $b$ . Across all areas of $e - b$ parameter space, the $T_{14} + umb$ (blue) and $e - \omega - \rho$ (orange) modeling methods perform equivalently. . . . .	107
5.5	Ratio of sampling efficiencies $\eta_{T_{14}+umb}/\eta_{e\omega\rho}$ as a function of duration ratio $T_{14}/T_{14,ref}$ and SNR. We bin the data across every 10 <sup>th</sup> percentile of the duration ratio distribution, showing a single point per bin per SNR (bins are separated by vertical grey lines). Each point shows the 15 <sup>th</sup> , 50 <sup>th</sup> , and 85 <sup>th</sup> percentiles of a given bin. The increased efficiency of the $T_{14} + umb$ method relative to the $e - \omega - \rho$ method depends on the SNR of the modeled lightcurve. At higher SNR the $T_{14} + umb$ method is significantly more efficient, but at moderate-to-low SNR the two methods have more similar efficiencies. At shorter transit durations, the $T_{14} + umb$ method is always more efficient, but this behavior changes around $T_{14}/T_{14,ref} \approx 0.8$ . The large spread in some uncertainties reflects the heterogeneity of our injected lightcurve parameters. . . . .	110
5.6	Comparison of $e$ values measured from the $T_{14} + umb$ and $T_{14} + dyn$ modeling methods at the 15 <sup>th</sup> , 50 <sup>th</sup> , and 85 <sup>th</sup> percentiles of their distributions. We show the $\Delta e_k = \{0.05, 0.1, 0.15\}$ in grey, as well as the ideal 1-to-1 line shown in black. These comparisons generally lie close to the 1-to-1 line, implying that the results of the two models are approximately equivalent. However, we do observe 42 tests where the $T_{14} + umb$ and $T_{14} + dyn$ model results are discrepant by more than $ \Delta e_k  \lesssim 0.05$ . . . . .	114

5.7	Ratio of sampling efficiencies $\eta_{T_{14}+umb}/\eta_{T_{14}+dyn}$ as a function of duration ratio $T_{14}/T_{14,ref}$ and SNR. We bin the data across every 10 <sup>th</sup> percentile of the duration ratio distribution, showing a single point per bin per SNR (bins are separated by vertical grey lines). Each point shows the 15 <sup>th</sup> , 50 <sup>th</sup> , and 85 <sup>th</sup> percentiles of a given bin. The efficiency of the $T_{14} + umb$ method relative to the $T_{14} + dyn$ method depends partially on the SNR of the modeled lightcurve. At shorter transit durations, the $T_{14} + umb$ method is typically more efficient, but the opposite is true at longer transit durations. The large spread in some uncertainties reflects the heterogeneity of our injected lightcurve parameters. . . . .	115
6.1	Stellar radius as a function of effective temperature for the 85 host stars in the TKS sample, with metallicity shown as color gradient. The California- <i>Kepler</i> Survey sample (Petigura et al. 2022) and the <i>Gaia-Kepler</i> sample (Berger et al. 2020b) are shown for reference in dark and light grey, respectively. . . . .	122
6.2	Stellar radius (left) and effective temperature (right) as measured by our stellar characterization procedure (TKS) on the y-axis compared to the values reported by the <i>Gaia</i> DR2 pipeline (top) and the TOI Catalog (bottom) on the x-axis. The TKS results produce reduced uncertainties and are generally consistent with one or both external sources. The residuals of our comparison between $R_{\star,TKS}$ and $R_{\star,TOI}$ can be seen in Figure 6.6a. We also note that the <i>Gaia</i> DR2 pipeline does not report stellar properties when $R_{\star,Gaia} \lesssim 0.5R_{\odot}$ , leaving out two TKS targets. . . . .	125
6.3	Diagram demonstrating the flow of data throughout the TKS light curve modeling pipeline. Yellow arrows are data inputs, grey boxes are processes, and blue ovals are data products. . . . .	129

6.4	Before (top) and after (bottom) applying our de-trending and initial TTV fitting procedure to the photometry of TOI-1136, a multi-planet system with high stellar variability. We only show the first sector (Sector 14) of this target’s photometry here, but all sectors display similar variability. We model our de-trended transit photometry and measure $R_p/R_*$ values for all 4 planets which agree with the values reported by Dai et al. 2022 within $\sim 1.5\sigma$ , as opposed to the TOI Catalog values which are all more than $1.5\sigma$ discrepant. Black points highlight transits ( $\pm 0.75$ times $T_{14}$ ), colored vertical dashed lines show modeled transit mid-points, and dark grey vertical dashed lines show calculated transit mid-points assuming a linear ephemeris. In the sector shown, we find the most significant TTVs for TOI-1136.03 (green) as can be seen by the separation between the modeled and calculated mid-points. . . . .	134
6.5	Planet-to-star radius ratio $R_p/R_*$ as measured by our planet characterization pipeline (TKS; y-axis) compared to the radius ratio values reported by the TOI Catalog (x-axis). The TKS results have significantly reduced uncertainties and are generally consistent with the TOI Catalog results (see §6.5.2). . . . .	138
6.6	Residuals of (a) stellar radius $R_*$ , (b) planet-to-star radius ratio $R_p/R_*$ , and (c) planet radius $R_p$ between our results and the TOI Catalog measurements (i.e. residual = $(R_{*,\text{TKS}} - R_{*,\text{TOI}})/\sigma(R_{*,\text{TOI}})$ ). Dashed vertical lines show the 15 <sup>th</sup> , 50 <sup>th</sup> , and 85 <sup>th</sup> percentiles of each distribution of residuals. . . . .	139
6.7	Distribution of median planet radius measurements for the TKS sample from this work, compared against the median radii taken from the TOI Catalog (green). Shaded regions show specific planet categories, classified by planetary radius. We observe a sparser radius gap region using TKS results than is implied by radii from the TOI Catalog, likely owing to our higher precision measurements. . . .	140

6.8	Panel (a): radius and orbital period for all TKS planets, with stellar mass shown as a color scale. Grey background points show the distribution of planets from the CKS sample (Petigura et al. 2022). The radius gap, as defined by Petigura et al. (2022), is shown as the orange shaded region. The Hot Neptune desert, as defined by Mazeh et al. (2016), is shown as the region marked off by black dashed lines. Panel (b): same as (a), except x-axis is $S_{\text{inc}}$ . Panel (c): same as (a), except x-axis is $M_{\star}$ and color scale is $P$ . Panel (d): same as (c), except x-axis is $[\text{Fe}/\text{H}]$ .	142
6.9	Four examples of significant transit-timing variations detected among TKS systems, including three from multi-planet systems (panels a, b, and d) and one from a single-planet system (panel c). (Panel a) TOI-1136.01: TTV semi-amplitude $\approx 24$ minutes, TTV signal period $\approx 790$ days; (b) TOI-1246.03: TTV semi-amplitude $\approx 61$ minutes, TTV signal period $\approx 990$ days; (c) TOI-1279.01: TTV semi-amplitude $\approx 34$ minutes, TTV signal period $\approx 775$ days; (d) TOI-1339.03: TTV semi-amplitude $\approx 22$ minutes, TTV signal period $\approx 960$ days, with significant scatter likely attributed to TTV chopping. These TTVs are measured from the observed transit midpoints (fit via <code>exoplanet</code> ) minus the transit midpoints calculated assuming a linear ephemeris. An estimated sinusoidal signal is shown fit to each set of TTVs, using a Lomb-Scargle periodogram and regression fit from <code>astropy</code> . These significant TTV signals likely arise from dynamical interactions between planetary companions.	144
7.1	Illustrative example of the power of hierarchical modeling for some parameter(s) $\theta$ , individual distributions $p_i(\theta_i)$ for $i$ targets, population level distribution $P(\theta)$ , prior $\pi$ , and hyperparameters $\alpha$ . See §7.3 for full description.	163

7.2	Various distribution types fit to the population-level eccentricity distribution of the TKS sample using hierarchical Bayesian modeling. A commonly cited Beta distribution (Kipping 2013) is shown for reference (black/grey). Our distribution fits suggest that the TKS sample has a lower distribution of eccentricities than the sample used by Kipping 2013, possibly attributed to differences in observing techniques or sample selection biases. . . . .	168
7.3	Comparison between fitting the TKS population-level eccentricity distribution using hierarchical analysis (orange) versus estimating the underlying distribution from summary statistics (blue). The histogram of median $e$ values for each TKS planet is shown with a Gaussian kernel density estimate fit to this distribution with a bandwidth measured via Scott’s Rule (Scott 1992). . . . .	169
7.4	Distribution of orbital periods (from NASA Exoplanet Archive on 22 May 2023; Akeson et al. 2013a) for planets discovered via transit detection ( <i>Kepler</i> , blue; TESS, orange) or radial velocity detection (green). Generally, planets discovered via RVs have significantly longer orbital periods than transiting planets. Also, <i>Kepler</i> planets tend to have longer periods than TESS planets, with 15 <sup>th</sup> –to–85 <sup>th</sup> percentile period ranges of $P_{Kepler} \approx 3.5 - 45$ days versus $P_{TESS} \approx 2 - 20$ days. .	170
7.5	Distribution of the fractional uncertainty measurements $\sigma(R_p/R_\star)$ for all TKS planets using an uninformative $e$ prior (blue) versus an $e$ prior informed by our hierarchical analysis of the population-level $e$ distribution (orange). The HBM-informed $e$ prior leads to more precise radius ratios with narrower distribution of $\sigma(R_p/R_\star)$ uncertainties. Gaussian kernel density estimates are shown fit to each distribution. . . . .	172

7.6	Period-radius distribution of TKS planets, with markers shown at the radius values calculated from our newly re-weighted $R_p/R_\star$ distributions. Planets with radii that shifted by $ \Delta R_p  > 1\%$ are shown as red triangles, and all other planets in the TKS sample are shown as black triangles. The triangles are pointed up or down depending on the direction of the relative shift $\Delta R_p$ . The red triangles form part of an arrow that shows the full extent of the shift $\Delta R_p$ , with the most substantial shifts occurring among smaller planets. We also show noteworthy areas of the period-radius parameter space for reference, including the Radius Gap (Fulton et al. 2017a; Petigura et al. 2022) and the Hot Neptune Desert (Mazeh et al. 2016). . . . .	173
7.7	Best-fit hierarchical model fit to the shape of the underlying distribution of $b$ values among TKS planets both before (blue) and after (orange) applying our HBM-informed $e$ prior. We measure both curves with a hierarchical Bayesian model fit to the underlying distribution histogram via a Gaussian process model with a Matérn-3/2 kernel. The $b$ distribution achieved via our updated prior is closer to what is expected from orbital geometry and observational bias (grey; Kipping & Sandford 2016). We show only the best-fit solutions here for simplicity, noting that the $1\sigma$ uncertainties on these probability distributions are significant and thus this figure is only intended to convey qualitative results. . . . .	174

7.8	Rayleigh distribution (left) and half-Gaussian distribution (right) fit to multiplicity-based sub-populations of the TKS sample. We show the 15 <sup>th</sup> –to–85 <sup>th</sup> percentile ranges for our distribution fits (top), with the 50 <sup>th</sup> percentile highlighted by the solid line. We also show the distributions of the hyperparameters that we measure for each sub-population (bottom), highlighting their 15 <sup>th</sup> , 50 <sup>th</sup> , and 85 <sup>th</sup> percentiles as vertical lines. The mean eccentricity $\bar{e}$ for each distribution is shown in the legend. We find that planets with observed transiting companions (blue) have a significantly higher distribution of eccentricities than those without observed transiting companions (green), consistent with the results of past studies.	176
7.9	Mean eccentricity $\bar{e}$ from various Rayleigh distribution fits to exoplanet eccentricities, drawn from both this work and the literature (Moorhead et al. 2011; Hadden & Lithwick 2014; Xie et al. 2016; Mills et al. 2019; Van Eylen et al. 2019). Open circles are from multi-planet fits, filled circles are from single-planet fits, and open square points are from fits to a general planet population. Faint vertical lines are used as a visual reference to compare literature results to the maximum-likelihood fits from this work. The mean eccentricity is related to hyperparameter $\sigma_{\text{rayleigh}}$ via $\bar{e} = \sigma_{\text{rayleigh}} \sqrt{\pi/2}$ (Jurić & Tremaine 2008).	178
7.10	Rayleigh and half-Gaussian distribution fits to metallicity-based sub-populations of the TKS sample, with the metal-rich population ( $[\text{Fe}/\text{H}] > 0$ ) shown in blue and the metal-poor population ( $[\text{Fe}/\text{H}] < 0$ ) shown in green. We find that planets orbiting metal-poor stars tend to have a higher distribution of eccentricities than those orbiting metal-rich stars, counter to the results of past studies.	180
7.11	Rayleigh and half-Gaussian distribution fits to radius-based sub-populations of the TKS sample, with the sub-Jovian-size population ( $R_p \lesssim 8R_{\oplus}$ ) shown in blue and the Jovian-size population ( $R_p > 8R_{\oplus}$ ) shown in green. We find that Jovian-size planets tend to have a higher distribution of eccentricities than sub-Jovian-size planets, consistent with theoretical expectations.	182

## LIST OF TABLES

2.1	HIP-97166 Radial Velocity Measurements . . . . .	15
2.2	HIP-97166 System Properties . . . . .	24
2.2	HIP-97166 System Properties . . . . .	25
3.1	TOI-1272 Radial Velocity Measurements . . . . .	45
3.2	TOI-1272 System Properties . . . . .	55
3.2	TOI-1272 System Properties . . . . .	56
4.1	Ground-truth simulation parameters; simulated photometry is shown in Figure 4.1.	74
4.2	Priors on model parameters for simulated lightcurve. . . . .	76
5.1	Transit Model Parameters and Priors . . . . .	98
6.1	TKS System Parameters: Median Fractional Uncertainties . . . . .	127
6.2	Transit Model Priors . . . . .	132
6.3	TKS Stellar Properties . . . . .	149
6.3	TKS Stellar Properties . . . . .	150
6.3	TKS Stellar Properties . . . . .	151
6.3	TKS Stellar Properties . . . . .	152
6.4	TKS Planet Properties . . . . .	153
6.4	TKS Planet Properties . . . . .	154
6.4	TKS Planet Properties . . . . .	155
6.4	TKS Planet Properties . . . . .	156
6.4	TKS Planet Properties . . . . .	157
7.1	Select eccentricity distribution fits from the literature . . . . .	165
7.2	Eccentricity distribution fits from this work . . . . .	167

## ACKNOWLEDGMENTS

I acknowledge that these findings are based on manuscripts with additional coauthors: [MacDougall et al. 2021](#) (Chapter 2), [MacDougall et al. 2022](#) (Chapter 3), [Gilbert et al. 2022](#) (Chapter 4), [MacDougall, Gilbert, & Petigura 2023](#) (submitted; Chapter 5), [MacDougall et al. 2023a](#) (accepted; Chapter 6), and [MacDougall et al. 2023b](#) (in prep; Chapter 7).

I would like to thank Professor Erik Petigura for his guidance throughout the research presented in my thesis and for fostering my growth as a researcher and communicator. I would not be the scientist that I am today had it not been for his constant support throughout the highs and lows of my graduate studies. I also thank Dr. Greg Gilbert for his indispensable mentorship and close collaboration which encouraged the development my thesis work to its current state. I thank my collaborators within the TESS-Keck Survey team for allowing me to join them in their endeavor to improve our understanding of exoplanet demographics. I also thank my fellow students, staff, and faculty within the UCLA Physics & Astronomy department for their supportive community that allows aspiring scientists to thrive.

I would also like to thank my family for their endless love and support throughout this long, winding journey to become the first "doctor" in my family. I would not be where I am today without my parents, Brian and Ruth, who have always been there for me through thick and thin. Regardless of whether I am calling to vent about my frustrations, brag about my accomplishments, or just catch up, they are always there to listen and offer words of encouragement that only a parent could. Their constant support has encouraged me to work hard, pursue my passions, and always do my best – and my thesis is the ultimate display of these traits which they have instilled in me. I thank my sister, Alexis, for always being my biggest advocate in my pursuit of becoming a scientist. I also thank my bean, Tanmyaa, for joining me on this wild journey. Your love and beaniness have always kept me grounded, and I would not be the person I am today without you by my side over these past five years.

To the loved ones who I have lost over the years: although I cannot officially name planets after you, these discoveries are dedicated to you – wherever you are out there in the universe.

I also acknowledge support from the UCLA Cota-Robles Graduate Fellowship which helped to fund my graduate studies. This work was also supported by a NASA Keck PI Data Award, administered by the NASA Exoplanet Science Institute. Data presented herein were obtained at the W. M. Keck Observatory from telescope time allocated to the National Aeronautics and Space Administration through the agency’s scientific partnership with the California Institute of Technology and the University of California. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation.

I thank the time assignment committees of the University of California, the California Institute of Technology, NASA, and the University of Hawaii for supporting the TESS-Keck Survey with observing time at the Keck Observatory. I gratefully acknowledge the efforts and dedication of the Keck Observatory staff for support of HIRES and remote observing. I recognize and acknowledge the cultural role and reverence that the summit of Maunakea has within the indigenous Hawaiian community and I am deeply grateful to have the opportunity to conduct observations from this mountain.

This thesis is based on data collected by the TESS mission. Funding for the TESS mission is provided by the NASA Explorer Program. I also acknowledge the use of public TESS data from pipelines at the TESS Science Office and at the TESS Science Processing Operations Center (SPOC). This thesis also includes data that are publicly available from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. The specific observations analyzed can be accessed via [MAST 2021](#) and [STScI 2018](#). Resources supporting this work were provided by the NASA High-End Computing Program through the NASA Advanced Supercomputing Division at Ames Research Center for the production of the SPOC data products.

## VITA

- 2023– **Data Scientist**, Google, Mountain View, CA
- 2022 **Data Scientist Intern**, Google, Mountain View, CA
- 2020–present **Graduate Student Researcher**, Department of Physics and Astronomy,  
UCLA, Los Angeles, CA
- 2020 **M.S. – Astronomy and Astrophysics**, UCLA, Los Angeles, CA
- 2019–2020 **Teaching Assistant**, Department of Physics and Astronomy, UCLA, Los  
Angeles, CA
- 2018 **B.S. – Astrophysics (Minor in Planetary Science)**, California Insti-  
tute of Technology, Pasadena, CA
- 2018 **Research Assistant**, Division of Planetary Science, Jet Propulsion Lab-  
oratory, Pasadena, CA
- 2015–2017 **Undergraduate Researcher**, Department of Astronomy, California In-  
stitute of Technology, Pasadena, CA

## PUBLICATIONS

**MacDougall, M. G.**, Petigura, E. A., et al. 2023, "The TESS-Keck Survey. XVIII. A Hierarchical Analysis of Photo-eccentricities from the TESS-Keck Survey Sample", in prep.

**MacDougall, M. G.**, Gilbert, G. J., and Petigura, E. A. 2023, "Accurate and efficient

photo-eccentric transit modeling", *The Astronomical Journal*, Submitted.

**MacDougall, M. G.**, Petigura, E. A., et al. 2023, "The TESS-Keck Survey. XV. Precise Properties of 108 TESS Planets and Their Host Stars", *The Astronomical Journal*, Accepted.

**MacDougall, M. G.**, Petigura, E. A., et al. 2022, "The TESS-Keck Survey. XIII. An Eccentric Hot Neptune with a Similar-Mass Outer Companion around TOI-1272", *The Astronomical Journal*, 164, 97, doi: 10.3847/1538-3881/ac7ce1.

Gilbert, G. J., **MacDougall, M. G.**, and Petigura, E. A. 2022, "Implicit Biases in Transit Models using Stellar Pseudo-density", *The Astronomical Journal*, 164, 92, doi: 10.3847/1538-3881/ac7f2f.

Chontos, A., Murphy, J. M. A., **MacDougall, M. G.**, et al. 2022, "The TESS-Keck Survey: Science Goals and Target Selection", *The Astronomical Journal*, Accepted, 163, 297, doi: 10.3847/1538-3881/ac6266.

**MacDougall, M. G.**, Petigura, E. A., et al. 2021, "The TESS-Keck Survey. VI. Two Eccentric sub-Neptunes Orbiting HIP-97166", *The Astronomical Journal*, 162, 265, doi: 10.3847/1538-3881/ac295e.

Dalba, P. A., Kane, S. R., Li, Z., **MacDougall, M. G.**, et al. 2021, "Giant Outer Transiting Exoplanet Mass (GOT 'EM) Survey. II. Discovery of a Failed Hot Jupiter on a 2.7 Year, Highly Eccentric Orbit", *The Astronomical Journal*, 162, 154, doi: 10.3847/1538-3881/ac134b.

# CHAPTER 1

## Introduction

For most of human history, the existence of other worlds beyond our own Solar System was the stuff of fantasy or theory – until it wasn't. Since the first extrasolar planet was discovered about 30 years ago by [Wolszczan & Frail \(1992\)](#), the collective curiosity of both the general public and the scientific community has driven exoplanet researchers to develop a variety of techniques for discovering and characterizing new planets, with the most prolific of these being the transit method. The *Kepler* spacecraft ([Borucki et al. 2010a](#)) pioneered the use of the transit method on large scales by simultaneously observing thousands of stars in search of the periodic dips in brightness that are characteristic of transiting planets. The *Kepler* mission went on to contribute  $\sim 51\%$  of the 5,338 exoplanet discoveries as of 1 May 2023 ([Akeson et al. 2013a](#)), along with another  $\sim 10\%$  from its extended *K2* mission ([Howell et al. 2014](#)). *Kepler's* success paved the way for both current (TESS, [Ricker et al. 2015a](#)) and future (PLATO, [Rauer et al. 2014a](#)) space-based transit surveys which share the common goal of bolstering our sample size of well-characterized planetary systems.

In recent years, countless studies have sifted through the growing sample of known planets to piece together the bigger picture of exoplanet demographics, providing us with invaluable information about the distributions of planetary radii (e.g. [Fulton et al. 2017a](#)), masses (e.g. [Weiss & Marcy 2014](#)), and system architectures (e.g. [Fabrycky et al. 2014](#)). Such investigations have revealed a variety of planetary system outcomes which motivate the question: how do we probe the past evolution of planets in order to better explain the demographics

that we observe today. Orbital eccentricities are a crucial piece of this puzzle, offering a glimpse into the past dynamical interactions of bodies around distant stars. Through studying exoplanets and their eccentricities, we gain a better understand our own place in the evolutionary story told by the demographics of these distant worlds.

## 1.1 Background and Motivation

### 1.1.1 Exoplanet Dynamics

Planet formation occurs in flattened disks of gas and dust revolving around stars that recently formed from the collapse of a molecular cloud (Lissauer 1993; Armitage 2018; Raymond & Morbidelli 2022). Within these circumstellar disks, planets are believed to form on near-circular, co-planar orbits within  $\sim 100$  million years after the formation of the disk. A variety of physical processes can proceed to excite newly-formed planets onto more eccentric elliptical orbits, including dynamical processes like disk interactions (e.g. Goldreich & Sari 2003), resonance crossing (e.g. Chiang 2003), planet-planet scattering (e.g. Chatterjee et al. 2008; Jurić & Tremaine 2008), and more complex N-body interactions such as the Eccentric Kozai-Lidov Mechanism (e.g. Naoz 2016). Other processes can also damp eccentricities over time, such as Kozai cycles and tidal friction (e.g. Fabrycky & Tremaine 2007). Thus, a planet’s orbital eccentricity is a relic of past dynamical perturbations as well as a reflection of ongoing interactions. Eccentricity measurements can provide important clues about diverging planet formation and evolution pathways that cannot be directly drawn from other planet characteristics.

### 1.1.2 Constraining Eccentricity

The eccentricities of transiting planets have historically been constrained via follow-up radial velocity (RV) observations which detect the reflex motion of a planet’s host star. Instruments for measuring precise RVs such as the Keck Observatory’s High-Resolution Spectrograph

(HIRES, Vogt et al. 1994) are capable of constraining eccentricities for large and/or close-in planets, but obtaining enough RVs for a precise eccentricity measurement can be costly and time-consuming. In many cases, small planet eccentricities will remain entirely unconstrained from this method. On the other hand, transit photometry can be gathered in bulk via space-based transit surveys for many stars hosting planets of all sizes, but recovering precise eccentricity constraints directly from transit modeling is not so straight-forward. Therefore, in order to better characterize both the typical and extreme values of planet eccentricities, especially small planets, improved photometric eccentricity constraints may be the best approach.

While many difficulties in transit modeling have long been solved (see, e.g., Mandel & Agol 2002a; Barnes 2007a), sampling the full parameter space of a transit model efficiently and accurately is no easy task. Additionally, the complex relationship between impact parameter  $b$ , eccentricity  $e$ , and argument of periastron  $\omega$  (orbit orientation angle) makes sampling directly in these all three of these variables computationally intensive. Fortunately, simplifications that were used to efficiently model circular orbits (Seager & Mallén-Ornelas 2003a) helped to pave the way for efficient photo-eccentric transit modeling that avoids direct sampling in  $\{e, \omega\}$  (see Ford et al. 2008; Kipping et al. 2012). This particular parameterization has been used to model transit photometry for individual planets (see, e.g., Dawson & Johnson 2012), small planet populations (e.g. Kane et al. 2012; Xie et al. 2016), and most notably, the entire *Kepler* planet catalog (Rowe et al. 2014b). However, this transit modeling approach can introduce a bias on impact parameter which then propagates to other modeled parameters, including eccentricity (Gilbert et al. 2022). One can instead use an eccentricity-explicit transit model parameterization that includes direct sampling in  $\{e, \omega, \rho_\star\}$ , which achieves unbiased results but with great inefficiency and a lack of flexibility for incorporating future updated priors on these parameters. These shortcomings leave room for additional improvements to achieve a photo-eccentric transit model that is efficient, accurate, and flexible (or "future-proof").

## 1.2 Structure of the dissertation

My thesis focuses on revealing both the typical and extreme values of small planet eccentricities via improved transit fitting methods and robust statistical modeling techniques. In Chapters 2 and 3, I use existing photo-eccentric transit modeling methods to discover two eccentric sub-Jovian-size transiting planets, each with a non-transiting outer companion, in collaboration with the TESS-Keck Survey team (MacDougall et al. 2021; MacDougall et al. 2022). Through these works, I identify a previously unaccounted-for bias in one of the most common transit modeling methods which unfairly disfavors high impact parameters and low eccentricities. In Chapter 4, I work with a small team of collaborators to explain and correct this bias, proposing a duration-based transit model parameterization that allows for efficient and unbiased estimates of exoplanet impact parameters (Gilbert et al. 2022). I expand upon this work in Chapter 5 to measure accurate eccentricity constraints from our duration-based transit model and validate these constraints through a suite of injection-and-recovery tests (MacDougall, Gilbert, & Petigura 2023, submitted). In Chapter 6, I perform high-precision stellar characterization for 85 TESS-Keck Survey targets and then apply our duration-based transit model to characterize 108 transiting planets orbiting these host stars (MacDougall et al 2023a, accepted). The complete photometric detrending and modeling pipeline that I develop for this project measures planet sizes, orbital ephemerides, and transit timing variations along with unbiased measurements of impact parameter and eccentricity. In Chapter 7, I use these photometric eccentricity constraints to determine the underlying eccentricity distribution of the TESS-Keck Survey planet sample via hierarchical Bayesian modeling (MacDougall et al. 2023b, in prep). I also compare the derived eccentricity distributions for various sub-populations of this sample to identify trends in eccentricity as a function of planet multiplicity, planet radius, and stellar metallicity. I provide concluding remarks in Chapter 8 along with and a brief discussion of future extensions of this work.

## CHAPTER 2

# The TESS-Keck Survey. VI. Two Eccentric sub-Neptunes Orbiting HIP-97166

### 2.1 Abstract

We report the discovery of HIP-97166b (TOI-1255b), a transiting sub-Neptune on a 10.3-day orbit around a K0 dwarf 68 pc from Earth. This planet was identified in a systematic search of TESS Objects of Interest for planets with eccentric orbits, based on a mismatch between the observed transit duration and the expected duration for a circular orbit. We confirmed the planetary nature of HIP-97166b with ground-based radial velocity measurements and measured a mass of  $M_b = 20 \pm 2 M_\oplus$  along with a radius of  $R_b = 2.7 \pm 0.1 R_\oplus$  from photometry. We detected an additional non-transiting planetary companion with  $M_c \sin i = 10 \pm 2 M_\oplus$  on a 16.8-day orbit. While the short transit duration of the inner planet initially suggested a high eccentricity, a joint RV-photometry analysis revealed a high impact parameter  $b = 0.84 \pm 0.03$  and a moderate eccentricity. Modeling the dynamics with the condition that the system remain stable over  $>10^5$  orbits yielded eccentricity constraints  $e_b = 0.16 \pm 0.03$  and  $e_c < 0.25$ . The eccentricity we find for planet b is above average for the small population of sub-Neptunes with well-measured eccentricities. We explored the plausible formation pathways of this system, proposing an early instability and merger event to explain the high density of the inner planet at  $5.3 \pm 0.9$  g/cc as well as its moderate eccentricity and proximity to a 5:3 mean-motion resonance.

## 2.2 Introduction

One of the key features of the Solar System is its low dynamical temperature. The eight planets are arranged with wide orbital spacing, low eccentricities, and no significant mean-motion resonances. Even so, the low mean eccentricity of  $\sim 0.06$  within the Solar System had to arise from somewhere. One explanation involves dynamical excitation and subsequent eccentricity damping following the divergent migration of Jupiter and Saturn (e.g. [Tsiganis et al. 2005](#)). Given a different set of initial conditions, however, it is possible for this and other excitation processes to achieve even higher dynamical temperatures in other planetary systems.

One of the major surprises of early exoplanet observations was the prevalence of high-eccentricity Jovians, in direct contrast to the Solar System planets. For reference,  $\sim 25\%$  of known planets with  $M_p > 100 M_\oplus$  and  $a > 1$  AU have  $e > 0.4$  (NASA Exoplanet Archive; [Akeson et al. 2013a](#)). Various mechanisms have been proposed to explain the highly excited states of these Jovian orbits, including planet migration, resonances, and close approaches (see, [Fabrycky & Tremaine 2007](#); [Ford & Rasio 2008](#); [Winn & Fabrycky 2015](#)).

On the other hand, characterizing the orbits and eccentricities of sub-Jovians is more challenging. The standard method of measuring a planet's eccentricity through radial velocity (RV) observations relies on the detection of a significant, non-sinusoidal motion from the host star. While many sub-Jovian semi-amplitudes are detectable with current facilities, the departures from sinusoidal are often less clear. Of the other available methods for planet detection, the transit method is the most prolific to date, but precise eccentricities are typically achieved through transit-timing variations (TTVs) when planets are near resonance, which is not representative of all systems (see, e.g., [Hadden & Lithwick 2014](#)). This limitation has led to a low fraction of sub-Neptune discoveries with well-constrained eccentricities.

Fortunately, a connection between transit duration and eccentricity exists (see, e.g., [Ford](#)

& Rasio 2008), assuming one has a well-constrained estimate of impact parameter. This relationship has given rise to a variety of studies that derive dynamical insights from transit photometry alone (Kipping 2010a; Kipping 2014b; Xie et al. 2016). Such work has been successfully carried out in recent years to determine both the eccentricities of individual planets (Dawson & Johnson 2012; Van Eylen et al. 2014) and the eccentricity distributions of larger samples of planetary systems (Kane et al. 2012; Van Eylen & Albrecht 2015; Xie et al. 2016). In our ongoing study, we use this relationship as a pre-filter to identify planet candidates at the extremes of the eccentricity distribution, observed with the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015a).

One of such candidates is HIP-97166b (TOI-1255b, TIC 237222864b), a sub-Neptune around an early K dwarf located high in the Northern hemisphere ( $\text{dec} = +74^\circ$ ) at a distance of 68 pc (Gaia Collaboration; Lindgren et al. 2018). Our investigation is part of a larger effort by the TESS-Keck Survey (see Chontos et al. 2022 for more information on TKS and its goals), which will build upon the legacies of Kepler and K2 to address major outstanding questions about exoplanet compositions, atmospheres, and system architectures.

In this paper, we describe the HIP-97166 system and the transit profile modeling that we used to identify eccentric planet candidates from photometry (§2.3) as well as our follow-up radial velocity observations (§2.4). We analyze our spectroscopic measurements to characterize both the stellar (§2.5) and planetary properties. From our rich RV data set, we also detect the presence of a non-transiting outer planet (§2.6). Finally, we explore the dynamics of this system through N-body simulations, which we used to further constrain our eccentricity measurements (§2.7). We also place this system in context with past exoplanet discoveries (§2.8) and consider possible formation pathways that could have led to the observed system architecture (§2.9).

## 2.3 HIP-97166b: A High-Eccentricity Candidate

### 2.3.1 TESS Photometry

HIP-97166 was observed by TESS with 2-min-cadence photometry in 12 sectors between UT 2019 July 18 and 2020 June 9 (Sectors 14–17, 19–26). The time-series photometry was processed by the TESS Science Processing Operations Center pipeline (SPOC; [Jenkins et al. 2016](#)), which first detected the periodic transit signal of HIP-97166b in 2019 September with a wavelet-based, noise-compensating matched filter ([Jenkins 2002](#); [Jenkins et al. 2010](#)). An initial limb-darkened transit model fit was performed ([Li et al. 2019](#)) and the signature passed a suite of diagnostic tests ([Twicken et al. 2018](#)) but for the difference imaging centroid test, which located the source of the transit-like signal  $10.2'' \pm 3.1''$  from the target. As the data accumulated for this target, the difference imaging centroid test results improved, shrinking the maximum deviation from the known target star’s location to  $0.9'' \pm 3.5''$  in the analysis of the data from sectors 14 through 26. The TESS Science Office reviewed the vetting results and issued an alert to the community on UT 2019 October 17 ([Guerrero et al. 2021](#)).

We accessed the Pre-search Data Conditioning Simple Aperture Photometry (PDC-SAP; [Stumpe et al. 2012](#), [Stumpe et al. 2014](#); [Smith et al. 2012](#)) through the Mikulski Archive for Space Telescopes (MAST), stitching together the light curves from individual TESS sectors into a single time-series using `Lightkurve` ([Lightkurve Collaboration et al. 2018](#)). We cleaned the TESS photometry by keeping only points with `quality_flag = 0` and excluding outliers beyond  $10\text{-}\sigma$  from the baseline flux. We divided out the median background flux of the time-series data to normalize the light curve then searched for transits using a box least squares (BLS; [Kovács et al. 2002a](#)) transit search to recover the same planetary signal detected by SPOC with  $\text{SNR} = 29.6$ . We subtracted the known transits and applied the BLS search again but identified no additional periodic transit events.

With the measured transit mid-point  $t_0$  and period  $P$  of the transiting planet, we masked

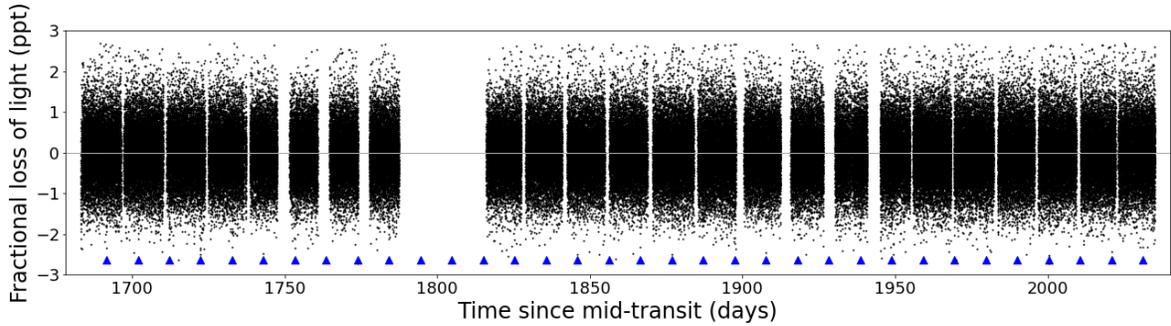


Figure 2.1: TESS photometry of HIP-97166. The data has been processed following the procedures discussed in §2.3.1. Transits of HIP-97166b are indicated by blue triangles and are 0.8 ppt.

out all transit events to detrend the light curve without obscuring the signal. We do note some photometric variability in the Simple Aperture Photometry (SAP) light curve, which we discuss in brief in §2.6.1. Interpolating over the masked transits, we fit a smoothed curve to systematics in the photometry with a Savitzky-Golay filter then subtracted out such additional structure to produce the flattened and normalized final light curve seen in Figure 2.1. Before unmasking the transit events, we clipped any individual outliers whose residuals to the smoothed fit were greater than  $5\text{-}\sigma$  discrepant.

### 2.3.2 Photometric Transit Model

The standard Mandel-Agol transit model can be specified by 5 transit parameters  $\{P, t_0, R_p/R_*, b, a/R_*\}$  in addition to stellar limb-darkening (Mandel & Agol 2002a). Eccentricity  $e$  and longitude of periastron  $\omega$  can also be directly sampled by such models, but including these dynamical parameters can significantly increase model run-time. Before devoting extensive computation time to every TOI system, a faster calculation involving transit duration can serve as a pre-filter for planets with potentially eccentric orbits.

While a planet on a circular orbit has a constant velocity, an eccentric planet will move faster when closer to its host star according to Kepler’s 2nd Law. As a result, an observer viewing

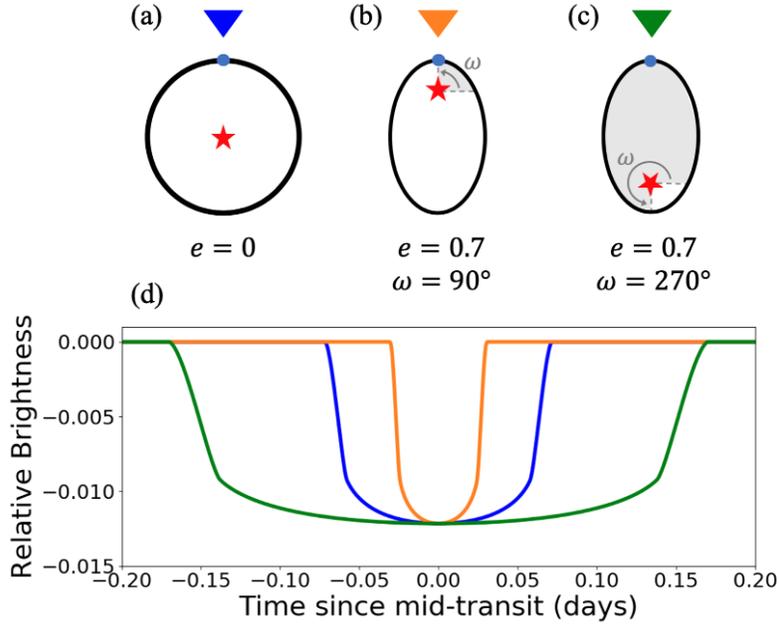


Figure 2.2: Transit duration as a function of  $e$  and  $\omega$ . (a) Observer (blue triangle) viewing a planet transiting on a circular orbit (blue light curve). (b) Observer (orange triangle) viewing a planet transiting at periastron (orange light curve). (c) Observer (green triangle) viewing a planet transiting at apastron (green light curve). (d) Simulated light curves associated with scenarios a-c for a Jupiter-size planet orbiting edge-on around a Sun-like star with a period of 5 days.

an eccentric planet transiting at periastron ( $\omega_* = 90^\circ$ , see Figure 2.2b) would see a shorter duration while another observer viewing this transit at apastron ( $\omega_* = 270^\circ$ , see Figure 2.2c) would see a longer duration. Thus, the ratio between a planet’s observed transit duration and its duration if it were on a circular orbit depends on the eccentricity and orbital orientation. The true duration (mid-ingress to mid-egress) can be computed using the geometric relation given in Winn 2010b:

$$T = \left( \frac{R_* P}{\pi a} \sqrt{1 - b^2} \right) \frac{\sqrt{1 - e^2}}{1 + e \sin \omega}. \quad (2.1)$$

We modeled this effect by re-parametrizing the standard transit model to sample duration rather than  $a/R_*$ . Our fitting basis included  $\{P, t_0, R_p/R_*, b, T, \mu, u, v\}$ , where  $\mu$  was mean out-of-transit stellar flux and  $u, v$  were quadratic limb darkening parameters. We fit the transit photometry of HIP-97166 using the statistical transit modeling package `exoplanet` (Foreman-Mackey et al. 2021a), which generates samples from the posterior probability density for these parameters conditioned on the observed TESS light curve. To generate these samples, `exoplanet` uses a gradient-based MCMC algorithm that is a generalization of the No U-Turn Sampling method (Hoffman & Gelman 2011a; Betancourt 2016). The versatility of `exoplanet` allowed us to build a model that best suited our goal here of obtaining a rapid transit duration fit. The sampled model produced a set of parameter posterior distributions from which we identified a median observed transit duration of  $T_{\text{obs}} = 0.06$  days.

To calculate the theoretical "circular" transit duration  $T_{\text{circ}}$  that the planet would have had given  $e = 0$ , we first point out that  $a/R_*$  in Eq. 2.1 maps to stellar density  $\rho_*$  through Kepler's 3rd Law (assuming  $m \ll M$ ):

$$\frac{a}{R_*} = \left( \frac{P^2 G \rho_*}{3\pi} \right)^{1/3}. \quad (2.2)$$

Since  $\rho_*$  can be directly measured through independent observations (see, e.g. §2.5), a more useful parametrization of Eq. 2.1 is obtained by substituting in Eq. 2.2, yielding

$$T \propto P^{1/3} (1 - b^2)^{1/2} \rho_*^{-1/3} \frac{\sqrt{1 - e^2}}{1 + e \sin \omega}. \quad (2.3)$$

Assuming one has an independently measured value for  $\rho_*$ , reliable estimates for  $b$  and  $P$ , and  $e = 0$ , Eq. 2.3 can be used to compute the expected transit duration of a planet if it were on a circular orbit (all else equal). For HIP-97166b, we found  $T_{\text{circ}} = 0.12 \pm 0.02$  days, for which a representative simulated transit curve can be seen in Figure 2.3. This calculation demonstrated that  $T_{\text{obs}}$  is significantly shorter than  $T_{\text{circ}}$ , with  $T_{\text{obs}}/T_{\text{circ}} \approx 0.5$ , implying a potential highly eccentric orbit transiting near periastron.

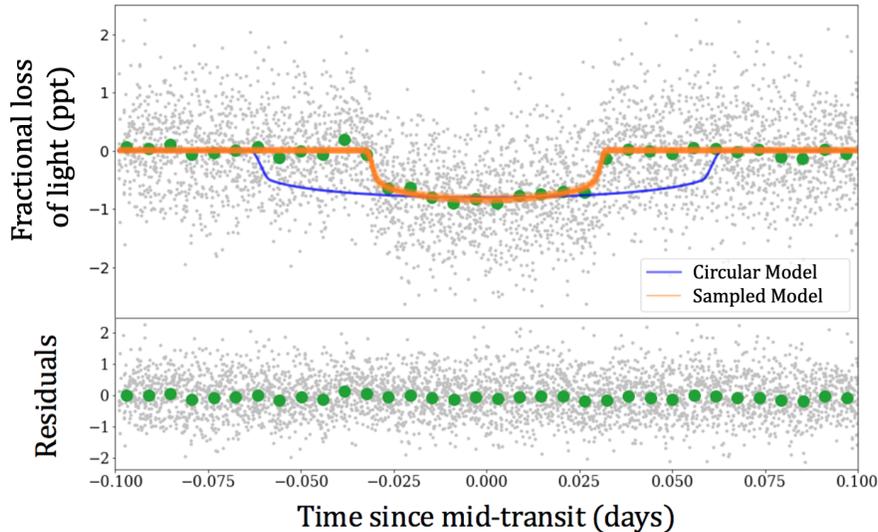


Figure 2.3: Top: Transit models (orange; 50 samples) drawn from parameter posterior distributions fit from phase-folded TESS photometry of HIP-97166b. A simulated transit curve (blue) is shown for a theoretical circular orbit of HIP-97166b, modeled using median posterior distribution values for all other parameters. Details regarding fitting procedure are discussed in §2.3.2. Bottom: Residuals to our maximum *a posteriori* model.

Although each of the variables in our transit duration model represented a unique transit property, they were not completely independent. In particular, there is significant  $e$ - $\omega$ - $b$  degeneracy for moderate-to-low SNR transits (Petigura 2020). For example, observed transit durations that are shorter than the expected duration for an edge-on, circular orbit can be caused by higher  $b$  and/or higher  $e$  (Moorhead et al. 2011; Dawson & Johnson 2012). On the other hand, this degeneracy is not a concern when modeling transits that are longer than expected since only eccentricity can have such an effect.

Given the short transit duration and these degeneracy concerns, we followed up our initial transit fit with a more detailed `exoplanet` model sampling 10 variables,  $\{P, t_0, R_p/R_*, b, \rho_*, \sqrt{e} \sin \omega, \sqrt{e} \cos \omega, \mu, u, v\}$ , each with weakly informative priors similar to those used by Sandford & Kipping (2017). We re-fit the TESS transit photometry of HIP-97166 with this

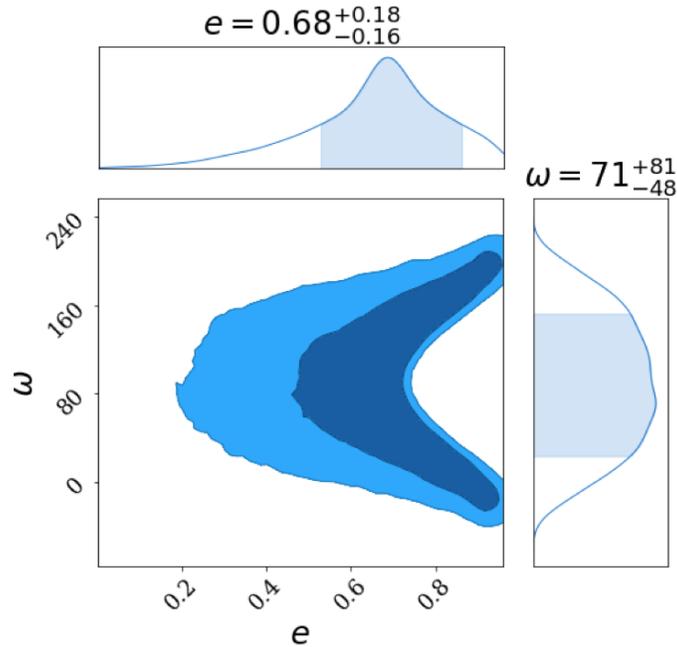


Figure 2.4: 2D joint posterior distribution of  $e$  and  $\omega$  for HIP-97166b, showing 1- $\sigma$  and 2- $\sigma$  credibility intervals. Best-fit values for both parameters are given:  $e = 0.68^{+0.18}_{-0.16}$  and  $\omega = 71^{+81}_{-48}^\circ$ .

model using 8,000 tuning steps and 6,000 sampling steps over 4 parallel chains, measuring an autocorrelation length of  $\sim 8$ . We find that the largest Gelman-Rubin statistic amongst the sampled parameters is  $R = 1.0004$ , suggesting convergence of the posterior chains according to [Gelman & Rubin \(1992\)](#). Figure 2.3 shows the final transit model sampled from the posteriors.

### 2.3.3 Eccentricity Constraints from Photometry

The photometrically-constrained eccentricity posterior distribution that we modeled for HIP-97166b was consistent with our initial high-eccentricity hypothesis, yielding a 1- $\sigma$  range of  $e = 0.52\text{--}0.86$ . While  $\omega$  is loosely constrained from this method, this analysis favors that the transit occurs closer to periastron. The joint 2D posterior for  $e$  and  $\omega$  is shown in Fig. 2.4, with the 1- $\sigma$  and 2- $\sigma$  credible regions shown in shades of blue.

We note, however, that our impact parameter distribution remains loosely constrained as well, with a  $1\text{-}\sigma$  range of  $b = 0.28\text{--}0.79$ , peaking in density towards the upper end of this range. Due to this high variance in the posterior distributions of both  $b$  and  $e$ , we required additional observations to break the  $e\text{-}\omega\text{-}b$  degeneracy of our transit model.

## 2.4 Spectroscopic Follow-up

### 2.4.1 HIRES RVs

We collected 44 spectra of HIP-97166 with the HIRES instrument at the Keck Observatory (Vogt et al. 1994) between UT 2020 May 30 and UT 2021 April 9 (Table 2.1). On average, the observations have a spectral resolution of  $R = 50,000$ , using a median exposure time of 900 s at  $5500 \text{ \AA}$ . We also obtained a high-SNR template spectrum on UT 2020 June 27 with 400 SNR  $\text{pixel}^{-1}$  at  $5000 \text{ \AA}$ .

For such observations, a heated cell of gaseous iodine was included along the light path just behind the entrance slit of the spectrometer, imprinting a dense forest of molecular absorption lines onto the observed stellar spectrum (Marcy & Butler 1992). These lines served as a wavelength reference for measuring the relative Doppler shift of each spectrum and tracking variations in the instrument profile using the standard forward-modeling procedures of the California Planet Search (Howard et al. 2010). Along with the measured RVs and corresponding uncertainties, the stellar activity S-index was computed for all 43 Keck/HIRES observations using the observed strengths of the Ca II H and K lines in our template spectrum, following the methods of Isaacson & Fischer (2010).

### 2.4.2 APF RVs

In addition to the Keck/HIRES follow-up, we also collected 124 iodine-in spectra of HIP-97166 with the Levy spectrograph on the Automated Planet Finder (APF) telescope (Vogt et al. 2014) between UT 2020 May 9 and UT 2021 April 11 at a spectral resolution of  $R =$

Table 2.1: HIP-97166 Radial Velocity Measurements

Time (BJD)	Tele.	RV (m s <sup>-1</sup> )	RV Unc. (m s <sup>-1</sup> )	$S_{\text{HK}}$
2458978.987	APF	0.549	3.407	0.202
2458979.905	APF	3.990	3.534	0.201
2458980.841	APF	-1.571	4.051	0.212
2458980.863	APF	-5.655	3.644	0.236
2458996.786	APF	5.818	3.092	0.278
2458996.807	APF	10.543	3.434	0.246
2459000.034	HIRES	-2.029	1.013	0.185
2459004.053	HIRES	-2.293	1.544	0.179
2459007.975	HIRES	-0.449	1.378	0.173
2459008.015	HIRES	-0.165	1.477	0.172

Note. — Only the first 10 RVs are displayed in this table. A complete list has been made available online.  $S_{\text{HK}}$  values were measured using procedures from [Isaacson & Fischer \(2010\)](#) with standard uncertainties of 0.002 for APF/Levy measurements and 0.001 for Keck/HIRES measurements.

100,000. For a majority of these, two observations were taken roughly 30 minutes apart over 70 separate nights. We also obtained 7 iodine-free template spectra using APF/Levy on UT 2020 May 26 with an average SNR  $\text{pixel}^{-1}$  of 102 at 5100 Å.

The APF/Levy Doppler code was developed based on the Keck/HIRES Doppler code and therefore follows a similar process for reducing spectra to RVs. However, this target is close to the APF/Levy magnitude limit with  $V = 9.85$ , contributing to a high cosmic ray rate that ultimately rendered the APF/Levy template unusable. Fortunately, we were able to substitute this with the the Keck/HIRES template to extract APF RVs, which has been successfully done in previous studies (see, e.g. [Dai et al. 2020](#)).

## 2.5 Stellar Characterization

We searched for nearby stars to rule out any contamination scenarios which would dilute the transit depth and, therefore, underestimate the planet size. We note that the star has a single neighbor listed in *Gaia* Data Release 2 (DR2) within 30'' which, at  $\Delta G = 8.75$ , contributes negligible dilution ([Gaia Collaboration et al. 2018](#)). Although *Gaia* is capable of spatially resolving sources down to  $\sim 0.5''$  separation in some instances, this still leaves a small region around our target star in which additional contaminating sources could exist.

We used Gemini/NIRI adaptive optics imager ([Hodapp et al. 2003](#)) to collect high resolution images of HIP-97166 on UT 2019 December 4. We collected nine images in the Br $\gamma$  filter, each with exposure time 1.75 s, in a grid dither pattern. We also collected flat frames and used the median-combined, dithered science frames as a sky background frame. For each frame we removed bad pixels, flat-fielded, and subtracted the sky background. We then aligned each image to the position of the frame and co-added the stack of images. We searched for companions visually, and did not detect companions anywhere in the field of view ( $26''.7 \times 26''.7$ , centered on the target). To assess the sensitivity of these observations, we injected fake PSFs at a number of position angles and separations from the host star

and scaled the brightness of these such that they could be detected at  $5\sigma$ . We reached a contrast of 5-mag relative to the host star beyond 270 mas and of 7.3-mag in the background limited regime beyond  $\sim 1''$ . We were thus able to rule out close-in diluting sources with high certainty.

We sought to further characterize HIP-97166 by inferring  $T_{\text{eff}}$  and  $[\text{Fe}/\text{H}]$  from our Keck/HIRES template spectrum using `SpecMatch-Synth`, as described in [Petigura et al. \(2017b\)](#). Following the methodology of [Fulton & Petigura \(2018\)](#), we then used these values as priors for stellar isochrone modeling with `isoclassify` ([Berger et al. 2020a](#); [Huber et al. 2017](#)). Our model also incorporated *2MASS* K-band magnitude and *Gaia* parallax to identify the best-fit stellar properties according to the MESA Isochrones and Stellar Tracks models (MIST; [Dotter 2016](#); [Choi et al. 2016](#)). We characterized  $\rho_*$  and several other stellar parameters using this method, accounting for model grid uncertainties according to [Tayar et al. \(2022\)](#), and we present these values in table 2.2.

Based on the solar-like values measured for  $\log(g)$ ,  $S_{\text{HK}}$ , and  $\log R'_{\text{HK}}$ , HIP-97166 places among the bottom quartile of expected activity-induced RV jitter, according to [Luhn et al. \(2020\)](#). The minimum jitter of similar stars in this study is  $\lesssim 2.5$  m/s, which is consistent with a jitter measurement of  $\sim 2.5$  m/s from our RVs.

## 2.6 Keplerian Modeling

### 2.6.1 RV Detection of Planets b and c

A preliminary search of our RV data for periodic signals using `RVSearch` ([Rosenthal et al. 2021](#)) revealed a Keplerian signal that matched the period of the transiting planet, with a false-alarm probability  $\text{FAP} \ll 0.001$  and a Doppler semi-amplitude of  $K \approx 6.6$  m/s. After subtracting, we identified an additional signal at 16.8 days, with  $K \approx 2.8$  m/s and  $\text{FAP} \approx 10^{-3}$  (Figure 2.5).

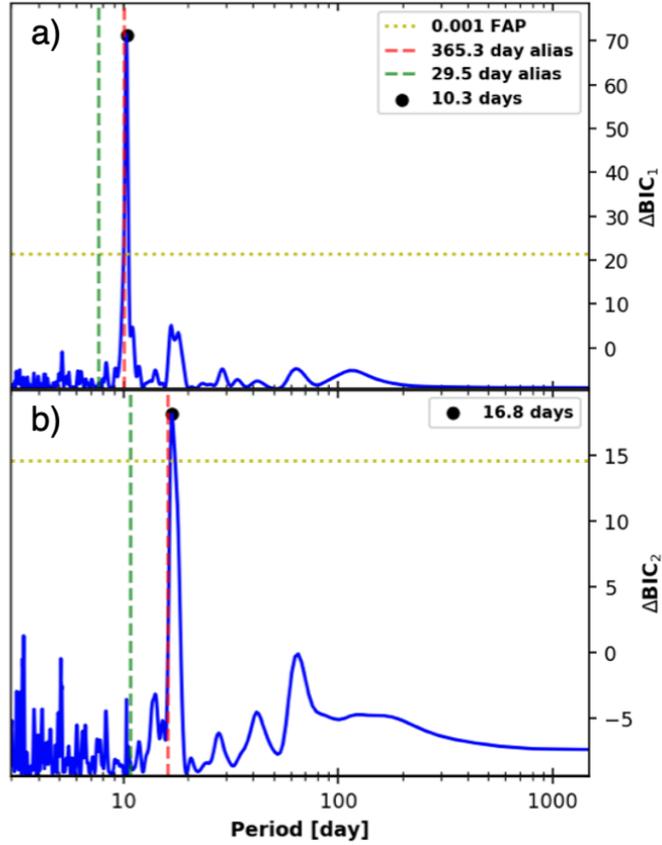


Figure 2.5: Iterative Keplerian periodogram search of HIP-97166 RVs using `RVSearch`, confirming the 10.3-day transiting planet and identifying a significant non-transiting companion with a 16.8-day period.  $\Delta\text{BIC}$  is used to discriminate between models with Keplerians at varying periods (Bayesian Information Criterion; Schwarz 1978), corresponding to a significance threshold of  $\text{FAP} = 0.001$  at the yellow dashed line. Monthly and yearly aliases are shown with in green and red dashed lines, respectively.

We also searched for planets using the  $l_1$  periodogram described in Hara et al. 2017, designed to reduce periodogram noise for unevenly sampled data as compared to the Lomb-Scargle method by solving the Basis Pursuit minimization problem (Chen et al. 2001). In our implementation, we used jitter term  $\sigma = 2.5$  m/s, correlation time  $\tau = 0$ , and maximum frequency of 1.5 cycles  $\text{d}^{-1}$ . Within the period range of 1.1 to 1000 days, the only clear detections in the  $l_1$ -periodogram occurred near the known period of 10.3 days and within  $1\text{-}\sigma$  of the suspected period of 16.8 days, with FAP values of  $\sim 10^{-9}$  and  $\sim 10^{-3}$ , respectively.

To confirm the significance of the 16.8 day period relative to other plausible signals that could be achieved by random fluctuations in noise, we re-sampled  $10^3$  synthetic data sets from the original RV data (Howard et al. 2010). We found that the 16.8 day signal was consistently the next strongest signal found in  $l_1$ -periodogram searches of the synthetic data, with the 10.3 day period always being the most significant.

While the 16.8-day RV signal was statistically significant, we did not identify a corresponding transit in §2.3.1. We confirmed this by phase-folding the detrended TESS photometry according to the RV-constrained period  $P$  and time of conjunction  $T_c$  for the outer RV signal, detecting no evidence for a transit event at this period. We therefore considered the possibility that this signal was stellar activity induced. We searched for trends in the  $S_{\text{HK}}$  activity time-series described in §2.5 and shown in Table 2.1. Similar to the RV data, we processed the nightly  $S_{\text{HK}}$  measurements using an  $l_1$ -periodogram, showing no indication of stellar activity with a 16.8-day period nor any other statistically significant periodicity (FAP  $\leq 10^{-3}$ ). Applying a similar approach to the SAP photometry of HIP-97166, we identified a significant  $l_1$ -periodogram signal at a period of 27.2 days, which was inconsistent with the 16.8 day RV signal. This periodicity in the SAP photometry is likely a systematic effect associated with the orbital period of the TESS spacecraft which has been seen in other TESS light curves as well.

To further test if the Keplerian signal was driven by stellar variability, we looked for a corre-

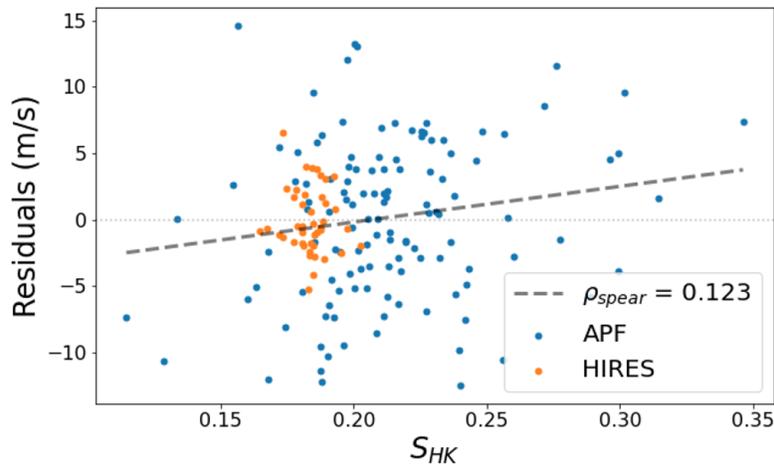


Figure 2.6: Spearman rank-order linear correlation fit between 1-planet RV residuals and  $S_{HK}$  activity metric shows a tenuous trend with a low correlation score  $\rho_{spear}$ , suggesting that no significant relationship exists between the observed RV signals and stellar activity.

lation between the  $S_{HK}$  index time-series and the RV residuals after removing the 10.3-day planet signal. Using the Spearman rank-order correlation test (Press et al. 1992), however, we found only a tenuous correlation with coefficient  $\rho_{spear} = 0.12$  (Figure 2.6). We, therefore, conclude that the 16.8-day periodicity is planetary in origin.

## 2.6.2 RV-only Constraints

We fit our complete RV data set using a two-planet model with `RadVel`, a Python package that applies maximum *a posteriori* model fitting and parameter estimation via MCMC to characterize planets from Keplerian RV signals (Fulton et al. 2018). The model that we selected consisted of the following free parameters for both planets:  $P$ ,  $T_c$ ,  $K$ ,  $\sqrt{e} \cos \omega$ , and  $\sqrt{e} \sin \omega$ . We also included RV offsets  $\gamma$  and RV jitter terms  $\sigma$  for each instrument, accounting for other astrophysical and instrumental uncertainty that are not already included in the model. Figure 2.7 shows the best-fit two-planet model as determined by the posterior distributions for each parameter shown in Figure 2.8.

Our best-fit model confirms the existence of two eccentric sub-Jovians orbiting HIP-97166, with a summary of planet properties provided in Table 2.2. Notably, the mass constraints for both planets b and c are significant, at the  $\sim 10\text{-}\sigma$  and  $\sim 5\text{-}\sigma$  levels, respectively. We also constrained the eccentricity of the transiting planet,  $e_b = 0.26 \pm 0.07$ , which was  $\sim 2.5\text{-}\sigma$  below the median  $e_b$  value of our transit model posterior distribution.

### 2.6.3 RV-Photometry Joint Model

While the RV-measured eccentricity was lower than expected, our RV and photometric posterior distributions on  $e_b$  remained consistent at the  $2\text{-}\sigma$  level. The degeneracy between  $e$  and  $b$  in our transit model was likely the source of this high spread, so we sought to build a complete model that more accurately accounted for this degeneracy. We performed a joint RV-transit fit of the data, using the posteriors on  $e$  and  $\omega$  from our RV fit as priors in a transit model to determine the combination of  $e$  and  $b$  needed to account for both the anomalously short transit duration and the RV signal.

In addition to  $e$  and  $\omega$  priors, we also placed a Gaussian prior on  $\rho_*$  in our RV-informed transit model, informed by our `isoclassify` stellar characterization. All other aspects of the model remained the same from the earlier implementation of `exoplanet` in §2.3.2. These RV-derived priors allowed us to constrain impact parameter to  $b = 0.84 \pm 0.03$ , significantly more constrained than our initial model fit and more consistent with what we observed in our RV-derived eccentricity posterior. Due to the covariance between  $b$  and  $R_p/R_*$ , a higher  $b$  also meant a higher  $R_p/R_*$  and subsequently resulted in our final planet radius measurement of  $R_p = 2.74 \pm 0.13 R_\oplus$ .

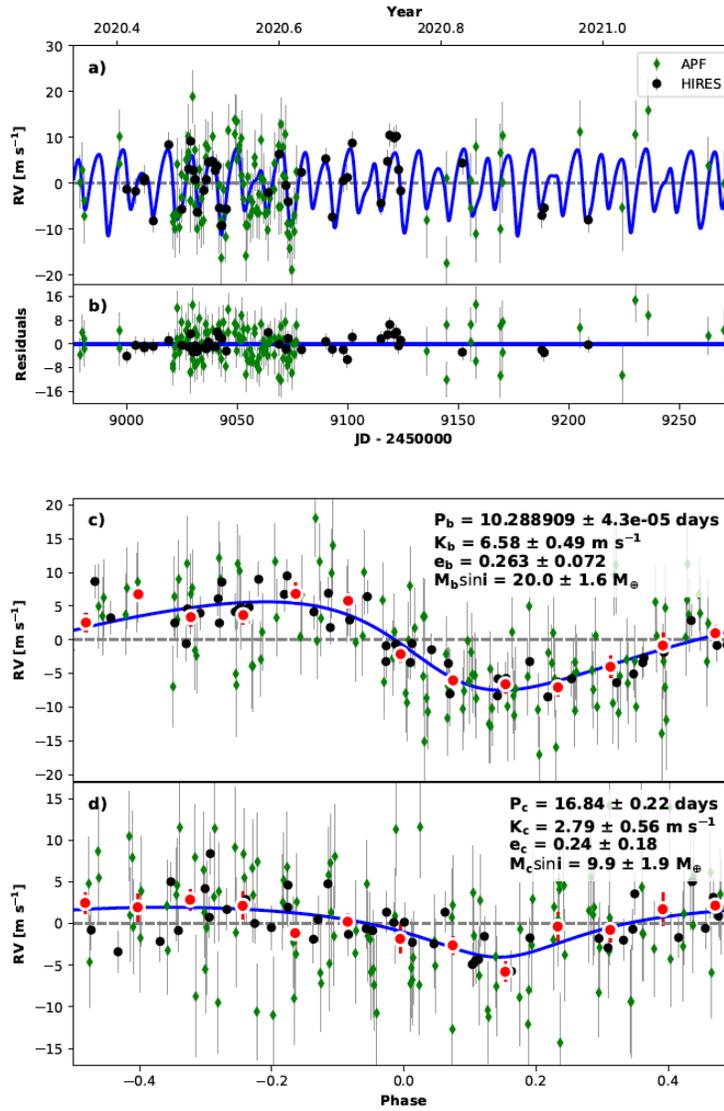


Figure 2.7: a) Best-fit model (blue) of radial velocity measurements from Keck/HIRES (black) and APF/Levy (green) using Radvel (Fulton et al. 2018) with no binning; b) Residuals to best-fit RV model; c) Phase-folded view of best-fit model and RV data for HIP-97166b, with binned points shown in red; d) Phase-folded view of best-fit model and RV data for HIP-97166c

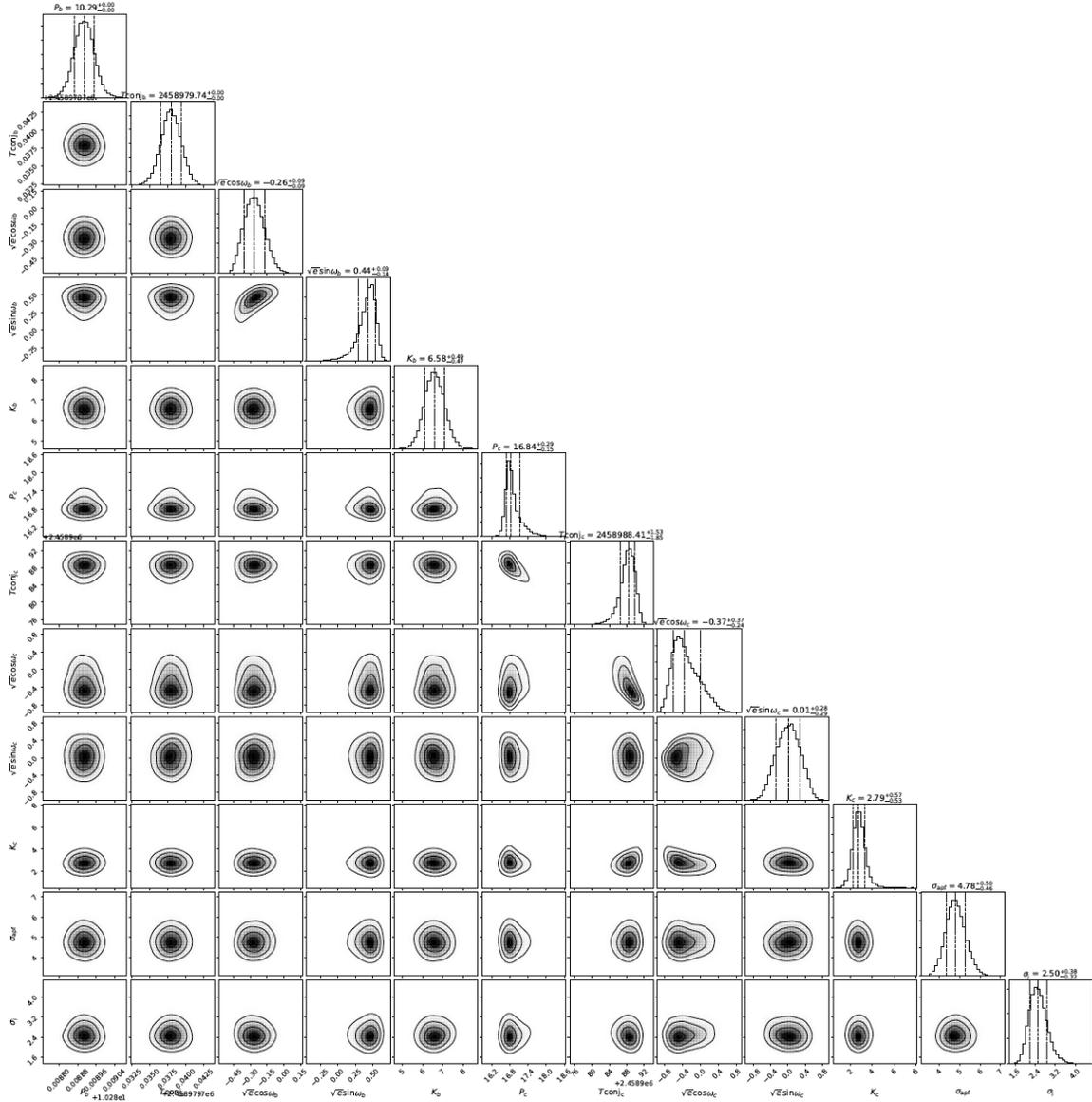


Figure 2.8: Posterior distributions of parameters from two-planet RV-only model with RadVel.

Table 2.2: HIP-97166 System Properties

Parameter	Value	Notes
<i>Stellar</i>		
RA ( $^{\circ}$ )	296.24462	A
Dec ( $^{\circ}$ )	74.06286	A
$\pi$ (mas)	$15.134 \pm 0.023$	A
$m_K$	$7.92 \pm 0.02$	B
$m_V$	$9.92 \pm 0.03$	C
$T_{\text{eff}}$ (K)	$5198 \pm 100$	D
[Fe/H] (dex)	$0.27 \pm 0.09$	D
$\log(g)$	$4.41 \pm 0.10$	D
age (Gyr)	$3.33 \pm 3.28$	E
$M_*$ ( $M_{\odot}$ )	$0.898 \pm 0.054$	E
$R_*$ ( $R_{\odot}$ )	$0.836 \pm 0.036$	E
$\rho_*$ (g/cc)	$2.154 \pm 0.312$	E
$S_{\text{HK}}$	0.182	F
$\log R'_{\text{HK}}$	-5.01	F
$u, v$	$0.47 \pm 0.05, 0.10 \pm 0.05$	G
<i>Planet b</i>		
$P$ (days)	$10.28891 \pm 0.00004$	H
$T_c$ (BJD-2457000)	$1691.6486 \pm 0.0007$	H
$b$	$0.836 \pm 0.027$	H
$R_p$ ( $R_{\oplus}$ )	$2.74 \pm 0.13$	H,E
$M_p$ ( $M_{\oplus}$ )	$20.0 \pm 1.5$	I
$\rho_p$ (g/cc)	$5.3 \pm 0.9$	I
$a$ (AU)	$0.089 \pm 0.001$	I
$\omega$ ( $^{\circ}$ )	$120.9 \pm 15.5$	I

Table 2.2 (cont'd): HIP-97166 System Properties

Parameter	Value	Notes
$e$	$0.16 \pm 0.03$	J
$T_{\text{eq}}$ (K)	$757 \pm 25$	K
<i>Planet c</i>		
$P$ (days)	$16.84 \pm 0.22$	I
$T_c$ (BJD-2457000)	$1988.4 \pm 1.6$	I
$M_p \sin i$ ( $M_{\oplus}$ )	$9.9 \pm 1.8$	I
$a$ (AU)	$0.124 \pm 0.002$	I
$\omega$ ( $^{\circ}$ )	$178.2 \pm 53.0$	I
$e$	$< 0.25$	J
$T_{\text{eq}}$ (K)	$642 \pm 22$	K

Note. — A: *Gaia* DR2 (Gaia Collaboration et al. 2018); B: *2MASS* (Skrutskie et al. 2006); C: TESS Input Catalog (TIC; Stassun et al. 2019); D: Derived with `SpecMatch-Synth`; E: Derived with `isoclassify`; F: Measured from Keck/HIRES template; G: Derived with LDTK (Parviainen & Aigrain 2015); H: Constrained from `exoplanet` transit model; I: Best-fit RV model with `RadVel`; J: Dynamically constrained with `rebound`; K: Derived.

## 2.7 System Dynamics

### 2.7.1 Eccentricity Constraints from Stability Requirements

The best-fit RV model suggested a significant ( $>3\text{-}\sigma$ ) non-zero eccentricity for the inner transiting planet but had only limited constraints for the eccentricity of the outer planet. Given this dynamical assessment and the relative proximity of these orbits ( $\Delta a \approx 0.035 \text{ AU} \approx 30 \Delta_{\text{Hill}}$ ), orbit crossing constraints revealed that the best-fit  $e_b$  and  $e_c$  values existed in an unstable region of parameter space. We subsequently evaluated the long-term stability and effects of planet-planet interactions within this system. We applied dynamical constraints on the orbital properties of both planets using N-body code `rebound` (Rein & Liu 2012), initializing 10,000 orbital simulations of the HIP-97166 system with properties drawn from our joint model posterior distributions.

For each simulation, we randomly drew the various system parameters ( $M_*$ ,  $M_b$ ,  $M_c$ ,  $P_b$ ,  $P_c$ ,  $\omega_b$ ,  $\omega_c$ ) from the posterior distributions of our joint model. For  $e_b$  and  $e_c$ , we performed a similar random sampling but with conditions that prevented orbit crossing at the initial state of the system based on derived values of  $a_b$  and  $a_c$ . We ran the simulations for up to  $10^4$  years ( $\sim 3 \times 10^5$  orbits of planet b) and found that only  $\sim 33\%$  of credible models were stable.

There is a region of  $e_b$ - $e_c$  parameter space where orbital stability is preferred that is consistent within  $\sim 1.5\text{-}\sigma$  of the values derived from the best-fit RV model (Figure 2.9). Simulations that successfully completed the  $10^4$  year run had overall eccentricity distributions given by  $e_b = 0.16 \pm 0.03$  and  $e_c < 0.25$ . Although the best-fit  $e$  values from our RV model are within the upper tail of the dynamically constrained distributions, it is clear that our `RadVel` model posteriors, which do not account for orbit crossing, skew towards higher values than allowed by stability criteria. Nonetheless, we confirmed that HIP-97166b can maintain a moderate eccentricity over long timescales even with a mildly eccentric, nearby companion.

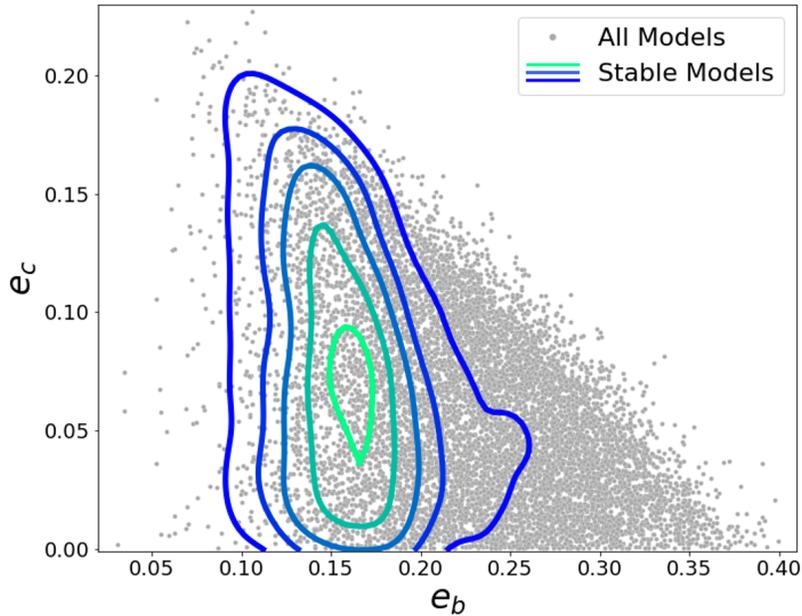


Figure 2.9: Distribution of initialized  $e_b$ - $e_c$  values for all **rebound** simulations (grey) of the HIP-97166 system. Blue-green contours show regions with the highest density of stable eccentricity configurations (simulations lasting  $10^4$  years).

### 2.7.2 Secular Eccentricity Variability

For simulations that experienced orbit crossing prior to  $10^4$  years, significant exchanges between  $e_b$  and  $e_c$  drove the system towards the instability region of  $e_b$ - $e_c$  parameter space. Stable systems, on the other hand, exhibited Laplace-Lagrange oscillations that remained stable long-term with a secular timescale on the order of  $\sim 10^2$  years (Figure 2.10). While neither planet’s eccentricity is expected to reach exceptionally high values during such oscillations, it is interesting to consider which formation scenarios could have led to this compact, excited orbital architecture. We note that these orbits are near 5:3 second-order mean-motion resonance (MMR),  $P_c/P_b \approx 1.64 \pm 0.02$ , but only in rare cases did we observe an impact of this resonance on the long-term stability of our simulations (see Figure 2.10, middle row). Although we cannot confirm the existence of MMR in this system, we cannot rule this scenario out either given the uncertainty on  $P_b$ .

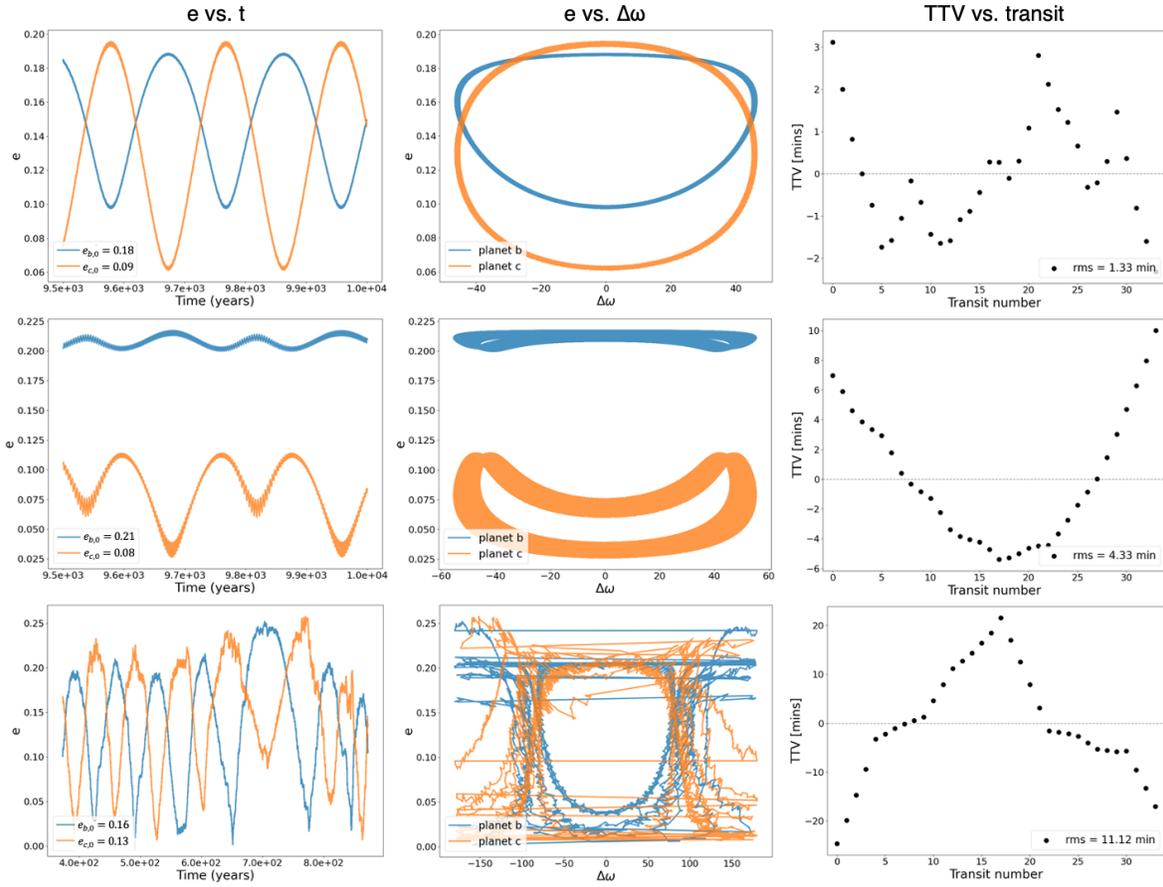


Figure 2.10: Three different dynamical scenarios of HIP-97166 b and c eccentricities and TTVs simulated using `rebound` and sampled from our previous results as described in §2.7.2. Rows: Stable eccentric scenario with negligible TTVs (Top). Stable eccentric scenario with periods oscillating about 5:3 MMR, displaying non-negligible TTVs (Middle). Unstable eccentric scenario lasting only  $\sim 800$  years with moderate TTVs (Bottom). Columns:  $e$  as a function of time over final 500 years of integration (Left).  $e$  as a function of  $\Delta\omega$  libration (Middle). TTV amplitude as a function of transit number over the same time baseline as TESS photometry (Right).

### 2.7.3 TTV Analysis

Through our dynamical simulations, we also computed the expected magnitude of the TTVs experienced by the inner planet due to interactions from the outer planet in the HIP-97166 system (Figure 2.10, 3rd column). Evaluating the RMS of these simulated TTVs over the same baseline as our TESS photometry, we found that stable simulations had a TTV RMS distribution of  $1.8 \pm 1.8$  mins. We also measured the magnitude of any observed TTVs from the transit photometry using `exoplanet`, finding an O-C RMS of  $\sim 2.9$  mins. This agreement between the simulated and empirical TTV RMS values demonstrated a consistency between the photometry and our proposed dynamics in which neither detected a significant TTV signal.

## 2.8 System in Context

### 2.8.1 Bulk Density and Core-Envelope Fraction

HIP-97166b is a sub-Neptune, a class of planets with radii  $\sim 2\text{--}4 R_{\oplus}$  that have been the subject of numerous studies in recent years (Marcy et al. 2014; Weiss & Marcy 2014; Lopez & Fortney 2014; Wolfgang et al. 2016). Sub-Neptunes display a wide range of characteristics and compositions but tend to have densities lower than that of Earth, suggesting H/He envelopes of a few percent (see, e.g., Rogers & Showman 2014; Owen & Wu 2017). With respect to the Weiss & Marcy (2014) relation, HIP-97166b is an outlier in the mass-radius parameter space, with bulk properties that place it among the denser and more massive planets in this class like K2-110b (Osborn et al. 2017).

We investigated the composition of HIP-97166b using a 2-component model, following the procedures of Petigura et al. (2017a) to quantify the core and envelope fractions of this planet. Assuming an Earth-like core composition and a solar-composition H/He envelope, we interpolated over a 4D grid of planetary and stellar parameters to quantitatively derive an estimate of the envelope mass fraction using Lopez & Fortney (2014) planet structure

models. We identified an envelope fraction of  $1.4 \pm 0.4\%$ , giving a core mass of  $\sim 19.6 M_{\oplus}$  (see Figure 2.11).

As for the outer planet, it is difficult to comment on its composition without a detectable transit. In the event that the outer planet is indeed transiting but below the detection threshold ( $\text{SNR} < 8$ ), we could place an upper limit on the radius of planet c. Assuming a transit duration of 0.1 days and no inclination, we find that  $R_{p,c} \leq 1.5 R_{\oplus}$ . A high-density planet of this size with expected mass  $\sim 10 M_{\oplus}$  would be an outlier in  $M_p - R_p$  space, but it cannot necessarily be ruled out. A sufficiently inclined orbit, however, could allow nearly any planet size, so we are left with only a minimum mass measurement at this time.

### 2.8.2 Weak Eccentricity Damping

While we found no indication of eccentricity decay in this system during our 10,000 year simulations, we considered the extent of tidal circularization on longer timescales following the procedure of Petigura et al. (2017a). Goldreich & Soter (1966) give the timescale for tidal eccentricity damping:

$$\tau_e = \frac{4}{63} \left( \frac{Q'}{n} \right) \left( \frac{M_p}{M_*} \right) \left( \frac{a}{R_p} \right)^5 \quad (2.4)$$

The mean motion and the reduced tidal quality factor are given by  $n = \sqrt{GM_*/a^3}$  and  $Q' = 3Q/2k_2$ , respectively, where  $Q$  is the specific dissipation function and  $k_2$  is the tidal Love number (Goldreich & Soter 1966; Murray & Dermott 1999; Mardling & Lin 2004). While  $Q'$  is highly uncertain even for Solar System planets, we estimated its value for the sub-Neptune HIP-97166b based on the range of values associated with Earth ( $\sim 10^3 - 2 \times 10^3$ ) and Uranus ( $\sim 10^5 - 6 \times 10^5$ ), drawn from Lainey (2016) and Petigura et al. (2017a). Assuming an Earth-like  $Q' = 10^3$  from the lower end of this range, we calculated  $\tau_e \approx 5$  Gyr. This was a similar timescale to the expected age of the system ( $3.3_{-2.3}^{+4.1}$  Gyr) based on

our `isoclassify` model, implying that significant eccentricity decay was unlikely to have occurred so far in this scenario. Since  $\tau_e$  scales linearly with  $Q'$ , larger values of  $Q'$  would only continue to inflate  $\tau_e$ . Thus, we conclude that the observed eccentricity of HIP-97166b is, at most, susceptible only to weak tidal damping over large timescales.

### 2.8.3 Sub-Neptune Eccentricity

To date, sub-Neptunes make up 29% of all planet discoveries, but only 6.4% of planets with eccentricity constraints of  $\sigma_e < 0.1$  fall within this radius range (Akeson et al. 2013a). In Figure 2.11, we show HIP-97166 b and c in context with other sub-Neptunes with well-constrained eccentricities. The small sample size makes it difficult to characterize the underlying eccentricity distribution of sub-Neptunes. Nevertheless, current observations suggest suppressed dynamical temperatures among this population, including the system presented in this paper. While our RV fit showed a 1- $\sigma$  eccentricity range of  $e_b = 0.19 - 0.33$ , our dynamically constrained N-body model lowered this 1- $\sigma$  range to  $e_b = 0.13 - 0.19$ . We display the latter result in Figure 2.11, adopting this as our final characterization. This finding is consistent with that of Correia et al. (2020), which suggests that warm Neptune-mass planets tend to present moderate, non-zero eccentricities.

Van Eylen & Albrecht (2015) also carried out an investigation into the eccentricity distribution of small planets using a sample of *Kepler* multi-planet systems. The eccentricity measurements in this study were performed using similar methods as we employed in our work in §2.3. The authors found that the overall eccentricity distribution was consistent with a Rayleigh function with dispersion  $\sigma_e = 0.049 \pm 0.013$ , which indicated that smaller planets ( $R_p \sim 2.6 R_\oplus$ ) generally had lower eccentricities when in the presence of planetary companions. A follow-up study investigating both multi-planet and single-planet transiting systems implemented a different distribution parametrization but still found that small planets in multi systems had lower eccentricities (Van Eylen et al. 2019). The modeled distributions for multi and single planet systems in this work were consistent with half-Gaussians with

dispersions  $\sigma_e = 0.083 \pm 0.018$  and  $\sigma_e = 0.32 \pm 0.06$ .

Given these past findings, we assert that HIP-97166b has a typical eccentricity relative to other single transiting systems of small planets. Similarly, the eccentricity of HIP-97166b is also consistent within  $\sim 1\text{-}\sigma$  of the typical well-characterized sub-Neptune. We do note, however, that in comparison to well-characterized Jovian planets, sub-Neptunes display a trend of suppressed eccentricities verified by a Kolmogorov-Smirnov test. The discovery of HIP-97166b is consistent with the observed trend.

## 2.9 Formation Scenarios

Although past studies have shown that *Kepler* multi-planet systems display no preference for being near mean-motion resonances (Fabrycky et al. 2014), resonances may still play a role in shaping some system architectures. One possible formation pathway for the HIP-97166 system is convergent migration, a process of interactions with smaller bodies in the circumstellar disk that lowers the period ratio and often drives planets into resonant configurations (Ford & Rasio 2008). While it has been noted that 5:3 MMR is a poor configuration for long-term stability (Lee & Thommes 2009), it is possible that the presently observed 5:3 near-MMR is a result of the past crossing of a stronger resonance. The dynamical instability and subsequent scattering from this compact resonance would be sufficient to excite the eccentricities that we observed for these two planets (Chiang et al. 2002; Izidoro et al. 2017).

While the HIP-97166 system is  $< 2\%$  away from the 5:3 second-order resonance, it is also only  $\sim 18\%$  away from the stronger 2:1 resonance, making this a plausible source of a past dynamical instability as well. A complication with this hypothesis, however, is that convergent migration requires dissipation and is expected to naturally damp eccentricities over time. While it is possible that the free eccentricities of HIP-97166 b and c may have been larger in the past if the orbits were spaced further apart, we also consider other possible explanations.

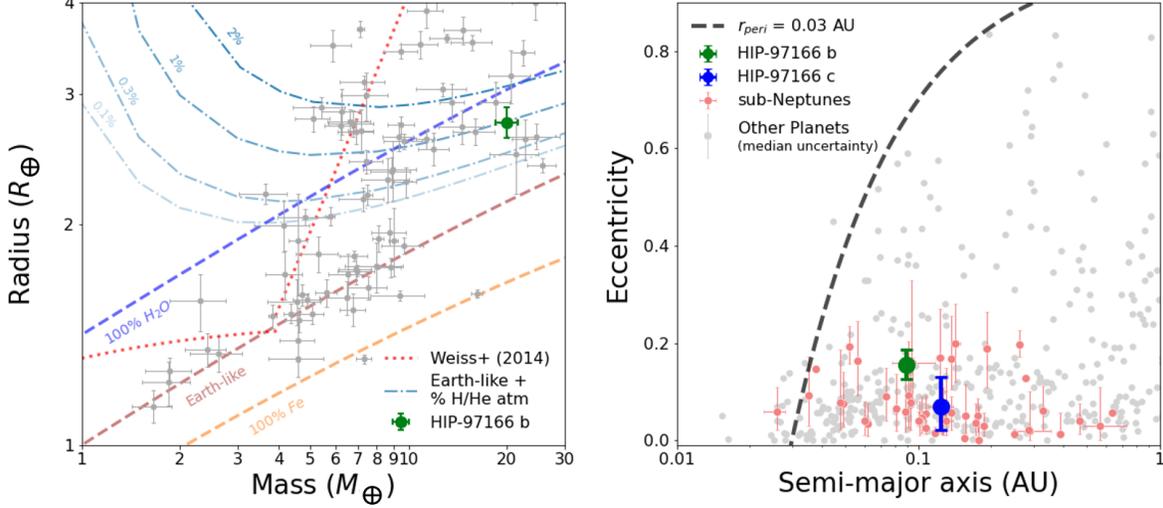


Figure 2.11: Left: Mass-radius distribution of known sub-Neptunes (Akeson et al. 2013a) with 20% measurement precision or better in both mass and radius (grey), shown with HIP-97166b (green). Planetary interior composition curves for various 2-component models (no atmosphere) from Zeng et al. (2016) are shown as dashed colored lines. Composition curves which include  $H_2$  envelopes of varying size on top of an Earth-like core are given by faded blue dot-dash lines (Zeng et al. 2019), assuming an equilibrium temperature of 700 K. Weiss & Marcy (2014) model shown as red dotted line (note the log-log axes). Right: Eccentricity distribution of planets with  $\sigma_e < 0.1$  (sub-Neptunes in red, other known planets in gray) as a function of orbital separation, showing HIP-97166 b and c in green and blue, respectively. While the eccentricity constraints of both planets place their upper bounds towards the eccentric tail of the sub-Neptune distribution, they are consistent with overall suppressed dynamical temperatures of sub-Neptunes relative to other planet populations. Periastron distance of 0.03 AU is also shown.

Another formation pathway for the HIP-97166 system is that the eccentricities were excited as the result of dynamical instabilities from planet-planet scattering and/or merger events (see, e.g., Jurić & Tremaine 2008). Chiang & Laughlin (2013a) demonstrated that in-situ formation of close-in super-Earths and sub-Neptunes can result in planet-planet mergers that lead to more massive planets on compact orbits. This led us to suggest that both the high density and moderate eccentricity of HIP-97166b may be a consequence of a past merger event between two  $\sim 10 M_{\oplus}$  planets. The occurrence of three or more planets of similar size on evenly spaced, compact orbits like this is not unheard of and has previously been referred to as the "peas in a pod" effect (Weiss et al. 2018). An instability in the dynamics of either inner planet in this proposed origin scenario could have easily resulted in a merged  $\sim 20 M_{\oplus}$  planet with  $e \sim 0.2$ , exciting the orbit of the outer planet in the process.

Finally, we considered the possibility that the observed planetary orbits may have been excited by a distant giant planetary companion. A giant perturber would have the capacity to increase inner planet eccentricities through dynamical interactions (Hansen 2017; Becker & Adams 2017; Pu & Lai 2018) or complex resonance effects such as the Eccentric Kozai-Lidov Mechanism (Naoz 2016; Denham et al. 2019; Barnes et al. 2020). Our RV observations presented only a marginal detection of acceleration, which allowed us to place loose constraints on a possible distant giant companion. With an observational baseline of  $\sim 12$ -months, we computed that an unobserved Jupiter-mass companion or larger could have a separation as low as a few AU but larger separations were more likely. These results indicated that the distant giant excitation scenario could not entirely be ruled out.

## 2.10 Summary and Conclusions

In this work, we identified HIP-97166b as an eccentric sub-Neptune candidate. This is the first in a series of investigations into TOIs with transit durations that significantly differ from expectations of a circular orbit. Combining our duration pre-filter and transit model with an

analysis of follow-up RV observations from Keck/HIRES and APF/Levy, we measured the mass, radius, and moderate eccentricity of HIP-97166b. This sub-Neptune is both denser ( $\rho_{p,b} = 5.3 \pm 0.9$  g/cc) and more eccentric ( $e_b = 0.16 \pm 0.03$ ) than is typical for planets of similar size, making it an interesting find among the TESS candidates. We also discovered a moderately eccentric outer companion ( $e_c < 0.25$ ) from RV observations, with a minimum mass of  $10 M_{\oplus}$  and a 16.8 day orbit.

N-body simulations of these orbits over time revealed a narrow region of dynamical stability that allowed us to measure the eccentricity of the inner planet with high precision, excluding a circular orbit to  $\sim 5\text{-}\sigma$ . Our leading hypothesis is that this system originally formed with a "peas in a pod" architecture, where the inner two of three original planets merged after a dynamical instability placed them on crossing orbits. The eccentricities that we observed in this system could have resulted from such an event and persisted on timescales of the age of the system. HIP-97166b is now among a small group of sub-Neptunes with high-precision eccentricity measurements.

## CHAPTER 3

# The TESS-Keck Survey. XIII. An Eccentric Hot Neptune with a Similar-Mass Outer Companion around TOI-1272

### 3.1 Abstract

We report the discovery of an eccentric hot Neptune and a non-transiting outer planet around TOI-1272. We identified the eccentricity of the inner planet, with an orbital period of 3.3 d and  $R_{p,b} = 4.1 \pm 0.2 R_{\oplus}$ , based on a mismatch between the observed transit duration and the expected duration for a circular orbit. Using ground-based radial velocity measurements from the HIRES instrument at the Keck Observatory, we measured the mass of TOI-1272b to be  $M_{p,b} = 25 \pm 2 M_{\oplus}$ . We also confirmed a high eccentricity of  $e_b = 0.34 \pm 0.06$ , placing TOI-1272b among the most eccentric well-characterized sub-Jovians. We used these RV measurements to also identify a non-transiting outer companion on an 8.7-d orbit with a similar mass of  $M_{p,c} \sin i = 27 \pm 3 M_{\oplus}$  and  $e_c \lesssim 0.35$ . Dynamically stable planet-planet interactions have likely allowed TOI-1272b to avoid tidal eccentricity decay despite the short circularization timescale expected for a close-in eccentric Neptune. TOI-1272b also maintains an envelope mass fraction of  $f_{\text{env}} \approx 11\%$  despite its high equilibrium temperature, implying that it may currently be undergoing photoevaporation. This planet joins a small population of short-period Neptune-like planets within the "Hot Neptune Desert" with a poorly understood formation pathway.

## 3.2 Introduction

The Solar System consists of eight planets from three size categories: terrestrial ( $0.4\text{--}1.0 R_{\oplus}$ ), ice giant ( $3.9\text{--}4.0 R_{\oplus}$ ), and gas giant ( $9.5\text{--}11.2 R_{\oplus}$ ). They are spread out across a 30 AU radial expanse and, with the exception of Mercury, their orbits are nearly circular. All three of these Solar System based patterns (distributions of planet sizes, orbit spacing, and orbital eccentricity) are contradicted by known exoplanetary systems. The prime *Kepler* mission (Borucki et al. 2010b) revealed that the most common exoplanet sizes are between Earth and Neptune ( $1.0\text{--}3.9 R_{\oplus}$ ; super-Earths / mini-Neptunes) in addition to a significant population with sizes between Uranus and Saturn ( $4.0\text{--}9.4 R_{\oplus}$ ; super-Neptunes or sub-Saturns). Moreover, planetary orbits interior to Mercury’s orbit are common, as well as orbits with eccentricities of  $e > 0.1$  ( $> 60\%$  of known planets, NASA Exoplanet Archive 2022a). Such broad demographics demonstrate a variety of possible outcomes to planet formation.

However, there are certain planet characteristics that are less common, even with observational biases accounted for. These include the dearth of extremely high eccentricity planets at short orbital periods, the "gap" in planet radius between super-Earths and mini-Neptunes at  $P < 100$  days (Fulton et al. 2017a), and the paucity of short-period Neptunian planets (Mazeh et al. 2016). The latter was first proposed as a natural consequence of photo-evaporation by planet characterization studies (Lopez & Fortney 2013; Owen & Wu 2013). The resulting “Hot Neptune desert” implies that for intermediate size planets between  $\sim 10\text{--}100 M_{\oplus}$  and  $\sim 2\text{--}6 R_{\oplus}$ , an inefficient formation pathway or an efficient mass loss mechanism sharply differentiates Neptunes from Jupiters at periods of  $\lesssim 5$  days. Owen & Lai (2018) proposed that the upper and lower boundaries of this sparse region of  $M_p\text{-}P$  and  $R_p\text{-}P$  parameter space can be mostly explained by limitations from either eccentricity decay of larger planets or photo-evaporation of smaller planets. The handful of observations of atmosphere-stripped Neptunian cores in the desert further supports this hypothesis (e.g. TOI-849b;

Armstrong et al. 2020), but many questions remain surrounding the formation pathways of such planets.

Eccentricity further complicates the long-term evolutionary history of sub-Jovians on compact orbits. To date, only 8 planets with sizes between 2.0–6.0  $R_{\oplus}$  and  $M_p < 100 M_{\oplus}$  have been found to have well-constrained eccentricities of  $e > 0.2$ , the greatest outlier being Kepler-1656b at  $e \approx 0.84 \pm 0.01$  (Brady et al. 2018). Only 2 of these planets, however, have orbital periods of  $P < 5$  days. Hot Jupiter-size planets tend to have longer tidal circularization time-scales and more massive cores that can retain their H/He envelopes during close-in periastron passage, but hot sub-Jovians are more susceptible to eccentricity decay and atmospheric loss. Consequently, the population of hot, eccentric Neptunes with  $\gtrsim 10\%$  H/He envelope mass fraction is small, consisting only of a handful of planets including HAT-P-11b (Yee et al. 2018) and GJ 436b (Lanotte et al. 2014).

In this paper, we discuss TOI-1272b, the latest Neptune to join the sparse population of hot, eccentric sub-Jovians. Leveraging the "photo-eccentric" methodology outlined by Dawson & Johnson (2012) and Kipping et al. (2012), we identified TOI-1272b as a candidate for high eccentricity based on a mismatch between the observed transit duration and the expected duration for a circular orbit. We used this technique as a pre-filter to vet for high-eccentricity candidates based on photometry alone, motivating follow-up radial velocity observations. Similar photometric modeling methods have been applied to *Kepler* target samples (Kane et al. 2012; Van Eylen & Albrecht 2015; Xie et al. 2016; Van Eylen et al. 2019), but those studies were not followed up by radial velocity campaigns. We present TOI-1272 as the second system from our photo-eccentric pre-filter study of TOIs, in association with the planetary demographics work being carried out by the TESS-Keck Survey collaboration (TKS; Chontos et al. 2022).

We introduce the TOI-1272 system and discuss the transit profile modeling that we used to identify TOI-1272b as an eccentric planet candidate from photometry (§3.3). We also

describe our follow-up radial velocity observations (§3.4) and analyze our spectroscopic measurements to characterize the properties of the host star, including stellar variability and age (§3.5). From our dense RV data set, we confirm the high eccentricity of TOI-1272b and detect the presence of a non-transiting outer planet (§3.6). Finally, we explore the long-term stability of this system through various dynamical criteria which we use to further constrain our eccentricity measurements (§3.7). We also place this system in context within the Hot Neptune Desert (§3.8) and consider possible formation and evolution pathways for TOI-1272b and other hot Neptunes.

### 3.3 TOI-1272b: A High-Eccentricity Candidate

#### 3.3.1 Photometry

TOI-1272 was observed by TESS with 2-min-cadence photometry in sectors 15, 16, and 22 between UT 2019 October 10 and 2020 May 11. The time-series photometry was processed by the TESS Science Processing Operations Center pipeline (SPOC; Jenkins et al. 2016), which first detected the periodic transit signal of TOI-1272b with a wavelet-based, noise-compensating matched filter (Jenkins 2002; Jenkins et al. 2010). An initial limb-darkened transit model fit was performed (Li et al. 2019) and the signature passed a suite of diagnostic tests described by Twicken et al. 2018, leading this target to be selected as a TOI.

We accessed the Pre-search Data Conditioning Simple Aperture Photometry (PDC-SAP; Stumpe et al. 2012; Stumpe et al. 2014; Smith et al. 2012) through the Mikulski Archive for Space Telescopes (MAST), stitching together the light curves from individual TESS sectors into a single time-series using `Lightkurve` (Lightkurve Collaboration et al. 2018). We performed outlier rejection, normalization, and de-trending of the full light curve following the procedures outlined in MacDougall et al. (2021). We then searched for transits using a box least squares (BLS; Kovács et al. 2002a) transit search to recover the same planetary signal detected by SPOC with  $\text{SNR} = 23.3$ . We subtracted the known transits and applied

the BLS search again but identified no additional periodic transit events.

To confirm that the observed transit events were on target and not the result of a background source, we referenced additional ground-based time-series photometry taken for TOI-1272. Independent observations were collected with MuSCAT2 (Narita et al. 2019) at a pixel scale of 0.44" in  $g$ ,  $r$ ,  $i$ , and  $z_s$  filters on UT 2020 February 28 and again  $\sim 1$  year later with MuSCAT (Narita et al. 2015) at a pixel scale of 0.36" in  $g$ ,  $r$ , and  $z_s$  filters on UT 2021 May 8. These detections confirmed that the expected transit was on target and presented no evidence of nearby eclipsing binaries. This target has one neighbor listed in *Gaia* Data Release 2 (DR2) within 30". At a separation of 8.45" and  $\Delta G = 5.93$ , the neighbor contributes  $< 1\%$  dilution to the light curve, which was already corrected for in the photometric data products that we used.

### 3.3.2 Photometric Transit Model

We characterized the planetary transit signal using a photometric light curve model to determine if TOI-1272b was a candidate for high-eccentricity. We made this determination by comparing the planet's observed transit duration ( $T$ ; mid-ingress to mid-egress) to the expected duration for a circular orbit  $T_{\text{circ}}$ . The ratio of these two values can be used to assess the orbital geometry of a transiting planet through the geometric relation for transit duration given by Winn 2010a:

$$T = \left( \frac{R_* P}{\pi a} \sqrt{1 - b^2} \right) \frac{\sqrt{1 - e^2}}{1 + e \sin \omega}, \quad (3.1)$$

Given the known period  $P = 3.316$  days, fixed  $b = 0$ , and the stellar characterization from §3.5.1, TOI-1272b would have a transit duration of  $T_{\text{circ}} = 0.094 \pm 0.004$  days if it were on a circular orbit. The observed transit, however, had a duration that was nearly 40% shorter than this at  $T_{\text{obs}} \approx 0.06$  days. The short transit duration suggests either a high eccentricity orbit transiting near periastron or an orbit with a high impact parameter, motivating our

follow-up analysis to constrain the true eccentricity.

To characterize the transit properties of TOI-1272b more precisely, we fit the available TESS photometry with the `exoplanet` package (Foreman-Mackey et al. 2021a). The `exoplanet` package uses a Hamiltonian Monte Carlo algorithm that is a generalization of the No U-Turn Sampling method (Hoffman & Gelman 2011a; Betancourt 2016). We used this model to generate samples from the posterior probability density for the parameters  $\{P, t_0, R_p/R_*, b, \rho_*, \sqrt{e} \sin \omega, \sqrt{e} \cos \omega, \mu, u, v\}$ , conditioned on the observed TESS light curve. Here,  $\mu$  is the mean out-of-transit stellar flux and  $\{u, v\}$  are quadratic limb darkening parameters. The model used here follows that of MacDougall et al. 2021.

We applied weakly informative priors to each of the 10 model parameters, similar to those used by Sandford & Kipping (2017). In particular, the prior used on our parameterization of eccentricity and argument of periastron  $\{\sqrt{e} \sin \omega, \sqrt{e} \cos \omega\}$  was uniform on both parameters, not accounting for transit probability or other astrophysically motivated considerations. Also, our prior on  $\rho_*$  was based on the stellar characterization discussed in §3.5.1. We fit the photometry of TOI-1272 with this model using 6,000 tuning steps and 4,000 sampling steps over 4 parallel chains. Figure 3.1 shows the final transit model sampled from the posteriors.

An independent fit to the MuSCAT2 transit photometry of TOI-1272b was performed and used to verify the results of our transit fit to the full TESS photometry (Figure 3.2). The raw MuSCAT2 data was reduced by the MuSCAT2 pipeline (Parviainen et al. 2019) which performed standard image calibration, aperture photometry, and modeled the instrumental systematics present in the data while simultaneously fitting a transit model to the light curve. We also applied our own transit model to the detrended MuSCAT2 photometry, achieving consistent posterior constraints on all transit parameters. The same process was repeated for independent transit photometry from MuSCAT, producing similar results.

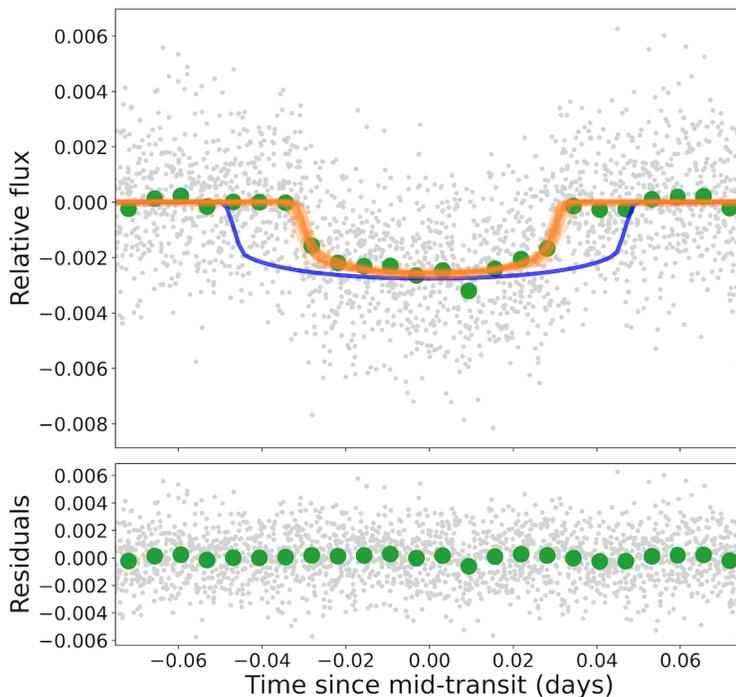


Figure 3.1: Top: Transit models drawn from parameter posterior distributions (orange; 50 samples) for phase-folded TESS photometry of TOI-1272b. Expected transit shape for a circular orbit shown in blue, modeled using median posterior distribution values for all other model parameters. Details regarding fitting procedure are discussed in §3.3.2. Bottom: Residuals to our maximum *a posteriori* model.

### 3.3.3 Eccentricity Constraints from Photometry

The photometrically-constrained eccentricity posterior distribution that we measured for TOI-1272b from  $\{\sqrt{e} \sin \omega, \sqrt{e} \cos \omega\}$  was consistent with our high-eccentricity hypothesis, yielding a  $1\sigma$  range of  $e = 0.18\text{--}0.60$  and an  $\omega$  suggestive of a transit near periastron. The individual posterior distributions of  $e$  and  $\omega$  are shown in Figure 3.3, along with their joint 2D posterior.

We do note, however, that our impact parameter distribution remains loosely constrained, with a  $1\sigma$  range of  $b = 0.19\text{--}0.73$ , peaking in density towards the upper end of this range

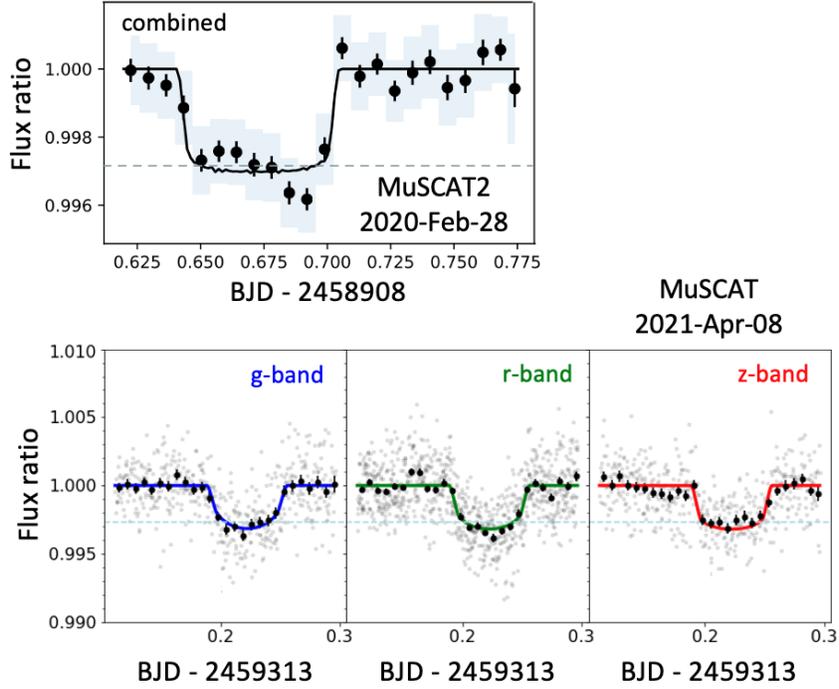


Figure 3.2: Transit photometry for two independent single transits measured by the MuSCAT2 and MuSCAT instruments, plotted with 10 minute binning and a photometric fits by the corresponding instrument pipeline. Horizontal dashed lines indicate the expected transit depth. Top: Combined photometry from MuSCAT2  $g$ ,  $r$ ,  $i$ , and  $z_s$  bands, observed on UT 2020 February 28. Bottom: Photometry from MuSCAT  $g$ ,  $r$ , and  $z$  bands, observed on UT 2021 April 8.

(Figure 3.3). The similarly loose constraints on both  $e$  and  $\omega$  implied that our photometric characterization of the orbital geometry was complicated by  $e$ - $\omega$ - $b$  degeneracy, as can be seen in the 2D joint posterior distributions in Figure 3.3. Nevertheless, the potential for a high eccentricity combined with the expected Neptune-like size of the planet and its short orbital period made TOI-1272b a prime target for follow-up radial velocity observations.

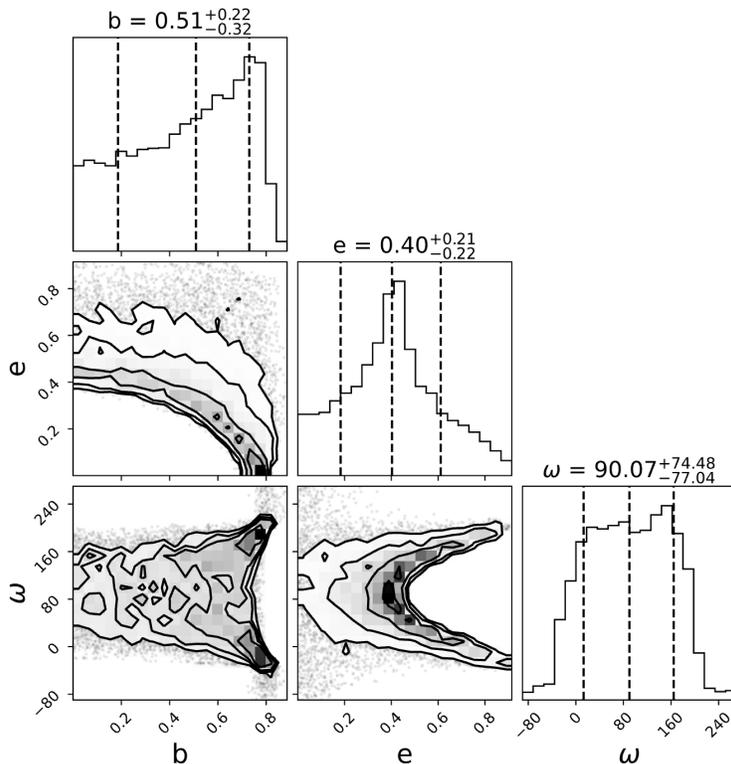


Figure 3.3: Corner plot of exoplanet model posteriors for TOI-1272b, showing the effects of  $e$ - $\omega$ - $b$  degeneracy on our transit fit. Best-fit value for eccentricity given by  $e = 0.40^{+0.20}_{-0.22}$ . Argument of periastron  $\omega$  remained loosely constrained about  $90^\circ$ , suggesting a transit near periastron.

### 3.4 Spectroscopic Follow-up

We obtained a high-SNR template spectrum of TOI-1272 with the HIRES instrument at the Keck Observatory (Vogt et al. 1994) on UT 2020 June 11 with  $282 \text{ SNR pixel}^{-1}$  at  $5000 \text{ \AA}$ . We also collected 62 spectra of TOI-1272 between UT 2020 Feb 5 and UT 2021 November 27 (Table 3.1). On average, the observations had a spectral resolution of  $R = 50,000$ , using a median exposure time of 900 s at  $5500 \text{ \AA}$ . Along with the RVs, we also measured the stellar activity S-index  $S_{\text{HK}}$  for all 62 Keck/HIRES observations using the observed strengths of the Ca II H and K lines in our spectra, following the methods of Isaacson & Fischer (2010).

Table 3.1: TOI-1272 Radial Velocity Measurements

Time (BJD)	RV (m s <sup>-1</sup> )	RV Unc. (m s <sup>-1</sup> )	$S_{\text{HK}}$
2458885.021035	11.00	1.68	0.315
2458904.925436	-0.66	1.55	0.277
2458911.896259	4.05	1.88	0.290
2458999.786695	4.78	1.51	0.304
2459003.840909	20.53	1.42	0.309
2459006.852128	-1.18	1.60	0.300
2459010.867973	4.47	1.55	0.294
2459024.775924	-3.86	1.58	0.300
2459024.829972	-3.96	1.36	0.295
2459024.883105	-6.08	1.70	0.302

Note. — Only the first 10 Keck/HIRES RVs are displayed in this table. A complete list has been made available online.  $S_{\text{HK}}$  values were measured using procedures from [Isaacson & Fischer \(2010\)](#) with standard uncertainties of 0.001.

For the RV observations, a heated cell of gaseous iodine was included along the light path just behind the entrance slit of the spectrometer, imprinting a dense forest of molecular absorption lines onto the observed stellar spectrum ([Marcy & Butler 1992](#)). These lines served as a wavelength reference for measuring the relative Doppler shift of each spectrum and tracking variations in the instrument profile using the standard forward-modeling procedures of the California Planet Search ([Howard et al. 2010](#)).

## 3.5 Stellar Characterization

### 3.5.1 Bulk Properties

Following the procedures outlined by [MacDougall et al. \(2021\)](#), we characterized the bulk properties of TOI-1272 by first inferring  $T_{\text{eff}}$  and  $[\text{Fe}/\text{H}]$  from our Keck/HIRES template spectrum using `SpecMatch-Synth` ([Petigura et al. 2017b](#)). We then modeled the stellar mass, radius, surface gravity, density, and age via stellar isochrone fitting with `isoclassify` ([Berger et al. 2020a](#); [Huber et al. 2017](#)). We report these values and their associated uncertainties in Table 3.2, accounting for small corrections due to model grid uncertainties discussed by [Tayar et al. \(2022\)](#). We note that the properties derived with `isoclassify` rely on *2MASS* K-band magnitude and *Gaia* parallax, also reported in Table 3.2. The properties that we measured were consistent with those reported to ExoFOP-TESS from two spectra obtained with the TRES instrument at the Whipple Observatory, analyzed using the Stellar Parameter Classification (SPC) tool ([Buchhave et al. 2012](#); [Buchhave et al. 2014](#)).

### 3.5.2 Variability and Rotation

To properly detrend our RV data and interpret any planetary signals, we first needed to characterize the intrinsic stellar variability for TOI-1272. We measured stellar variability from the TESS 2-min cadence SAP photometry where TOI-1272 was observed in 3 sectors, one of which partially overlapped with our RV observation baseline. Upon removing data that were flagged as being poor quality,  $\geq 5\sigma$  outliers, or part of the TOI-1272b transit events, we measured the stellar variability period from the trimmed SAP light curve using the TESS-SIP algorithm ([Hedges et al. 2020](#)). This systematics-insensitive Lomb-Scargle periodogram ([Angus et al. 2016](#)) yielded a clear variability signal at  $28.3 \pm 0.6$  days, likely associated with stellar rotation. The corrected TESS-SIP light curve, Lomb-Scargle periodogram, and phase-folded light curve with a sinusoidal fit are shown in Figure 3.4.

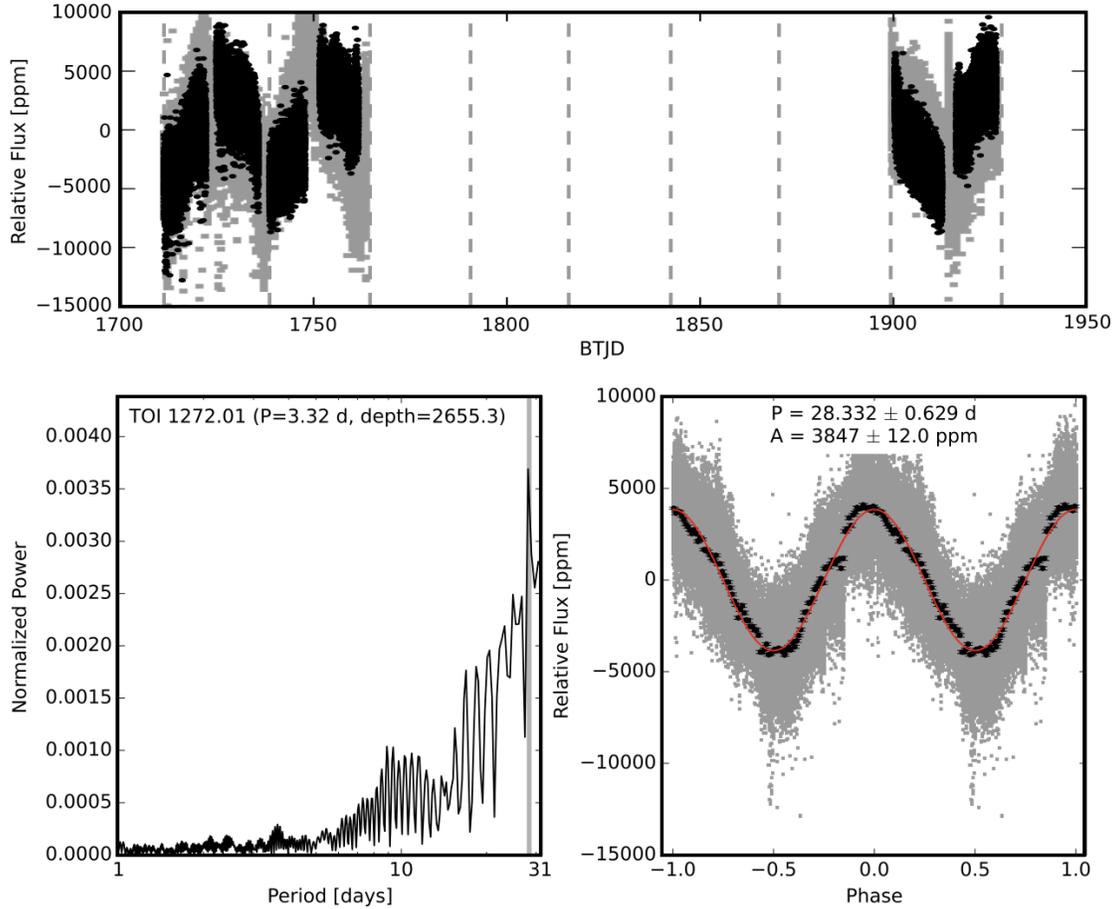


Figure 3.4: Top: TESS light curve of TOI-1272 from sectors 15, 16, and 22. The gray points show the original TESS SAP light curve and the black points show the TESS-SIP corrected light curve that is used to extract the variability period. The vertical dashed lines mark the gaps between TESS sectors. Bottom Left: Lomb-Scargle periodogram of the TESS-SIP corrected light curve. The orbital period and transit depth from the TOI catalog are listed at the top of the panel. Bottom Right: Phase-folded light curve based on the most significant period detected from the periodogram. The gray points show the TESS-SIP corrected photometry, the black points show the binned data, and the red curve is a sinusoidal fit to the phase-folded light curve. The period, amplitude, and their associated uncertainties from the best-fit sinusoidal function are listed at the top of the panel.

The stellar variability observed in the TESS photometry of TOI-1272 also allowed us to derive the expected stellar activity-driven variability in our RV measurements using the  $FF'$  method (Aigrain et al. 2012). This method uses the light curve flux ( $F$ ), its derivative ( $F'$ ), an estimate of relative spot coverage ( $f \sim 0.005$  in this work), and a simple spot model to simulate activity-induced RV variability. We estimated that this stellar activity would produce an RV variability signal with semi-amplitude  $K \approx 5.0 \text{ m s}^{-1}$ , assuming a sinusoidal signal. Given the partial overlap of our RV baseline with that of the TESS photometry, we used this RV variability estimate as the foundation for our consideration of activity-driven RV signals in §3.6.2. By doing so, we implicitly assumed that the variability signal in the region of overlap could be extrapolated out to the entire data set. Based on the measured value of  $\log(g)$  and the observed values of activity metrics  $S_{\text{HK}}$  and  $\log R'_{\text{HK}}$ , moderate stellar activity-driven RV jitter was also expected for TOI-1272 based on the classifications presented in Luhm et al. 2020,  $\sigma_{\text{jit}} \gtrsim 2.5 \text{ m s}^{-1}$ , consistent with our photometry-only estimate.

As a final consideration of stellar variability, we searched for periodic, activity-driven signals in the  $S_{\text{HK}}$  data series for TOI-1272 using a Lomb-Scargle periodogram. We identified a significant 28.5 day signal, consistent with our stellar variability measurement from TESS-SIP (Figure 3.5). We also detected additional sub-significant  $S_{\text{HK}}$  variability signals that did not correspond to any known sources. We consider the impact that activity may have on our RV measurements when constructing our RV-only model in §3.6.

### 3.5.3 Age

Given the short orbital period of TOI-1272b and the possibility of a high-eccentricity orbit, the age constraints for this system were valuable for interpreting the tidal circularization timescale of the transiting planet. Our isochrone fit using `isoclassify` yielded a poorly constrained age estimate of  $\sim 1\text{--}7$  Gyr. This was consistent with a first-order analytical estimate of the age of TOI-1272, 3.1 Gyr, based on  $G_{\text{Bp}} - G_{\text{Rp}}$  color and stellar rotation period via gyrochronology (Angus et al. 2019b).

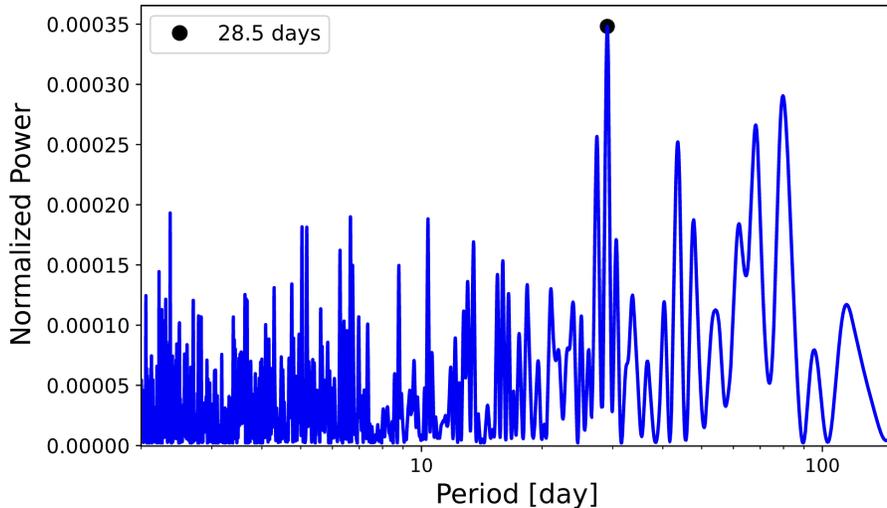


Figure 3.5: Lomb-Scargle periodogram search of TOI-1272  $S_{\text{HK}}$  data using, identifying a significant 28.5 day variability signal consistent with the suspected stellar rotation period.

We took this analysis a step further by using `stardate` (Angus et al. 2019a) to combine stellar isochrone modeling with gyrochronology to precisely measure the stellar age. Running the `stardate` MCMC sampler for  $10^5$  draws, we measured an age of  $3.65^{+4.17}_{-0.98}$  Gyr. While this age range remained broad and consistent with our `isoclassify` measurement, the increased median value and reduced lower uncertainty from `stardate` provided us with better constraints on the lower bound of the age of this system.

## 3.6 Keplerian Modeling

### 3.6.1 RV Detection of Planets b and c

We searched for periodic signals in our RV data using the `RVSearch` pipeline (Rosenthal et al. 2021). We set Gaussian priors on the period  $P_b$  and time of conjunction  $T_{c,b}$  of the 3.3-day planetary signal known from photometry. We then used `RVSearch` to iteratively search the RV data for additional Keplerian signals across the period range from 2 to 4000 days. This search yielded an eccentric Keplerian fit with  $K \approx 12.6 \text{ m s}^{-1}$  at the known period and an

outer 8.7-day Keplerian fit with  $K \approx 9.4 \text{ m s}^{-1}$  (Figure 3.6). Both signals surpassed our significance threshold, with false-alarm probabilities (FAP) measured by `RVSearch` FAP  $\approx 10^{-4}$  and  $10^{-5}$ , respectively.

We corroborated the significance of the 8.7-day signal by performing an independent search of the RV data set using an  $l_1$  periodogram (Hara et al. 2017), which minimizes the aliasing seen in a general Lomb-Scargle periodogram by evaluating all frequencies simultaneously rather than iteratively. We implemented our  $l_1$  periodogram with jitter  $\sigma = 5.0 \text{ m s}^{-1}$ , correlation time  $\tau = 0$ , and maximum frequency  $1.5 \text{ cycles d}^{-1}$  across the period range from 1.1 to 1000 days. Both the 3.3 and 8.7 day signals were clearly detected by from this  $l_1$  periodogram search, with consistent FAP values of  $\sim 10^{-4}$  and  $\sim 10^{-5}$ , respectively. Given the significance of the 8.7-day period and the lack of a corresponding signal in either the  $S_{\text{HK}}$  activity data or photometric time series (see §3.5.2), we concluded that this Keplerian signal was of planetary origin. A close inspection of the phase-folded and detrended TESS photometry at the RV-constrained period and time of conjunction for the outer RV signal showed no evidence for a corresponding transit event.

### 3.6.2 Additional RV Signals

While we did not identify any additional signals in our RV data that met our significance criteria, we did detect a sub-significant Keplerian signal at a 14.1 day period using both `RVSearch` and an  $l_1$  periodogram search. This signal persisted throughout our entire observing baseline and was detectable in the residuals to a preliminary two-planet RV fit with Keplerian modeling code `RadVel` (Fulton et al. 2018). We concluded that this signal was the first harmonic ( $P_{\text{rot}}/2$ ) of the 28.3 day stellar rotation as measured from the  $S_{\text{HK}}$  time series and TESS photometry. The  $P_{\text{rot}}/2$  harmonic of a star’s rotation period is known to induce strong periodic activity signatures such as this in RV time-series data (Boisse et al. 2011).

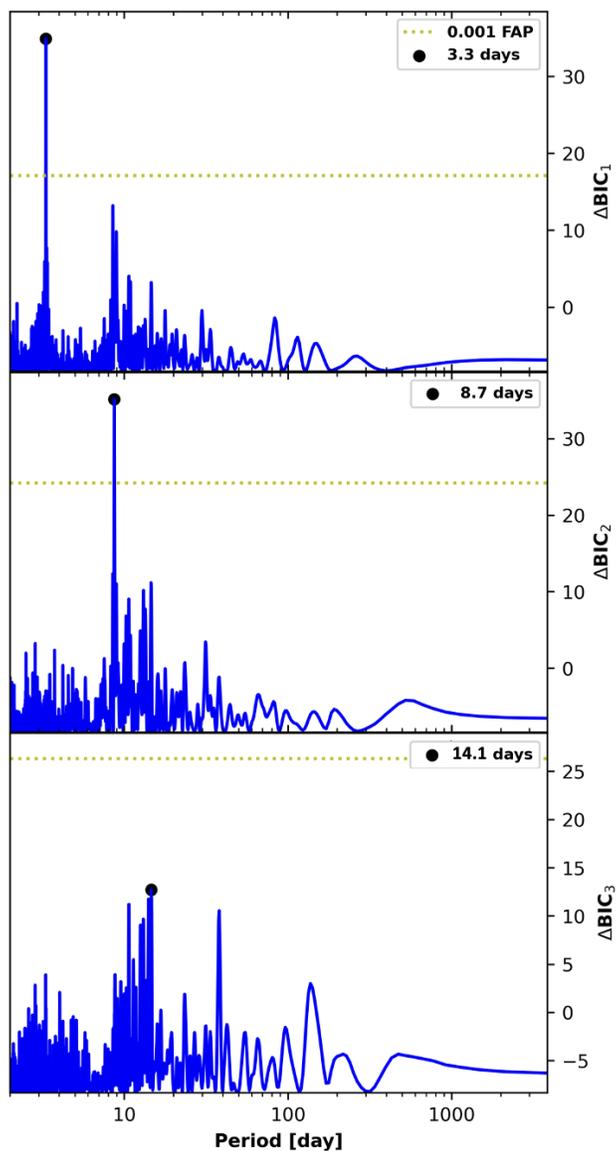


Figure 3.6: Iterative Keplerian periodogram search of TOI-1272 RV data using RVSearch. We confirmed the 3.3-day transiting planet (panel a) and identified a significant 8.7-day period with no corresponding transits (panel b).  $\Delta\text{BIC}$  was used to discriminate between models with additional Keplerians over a grid of periods (Bayesian Information Criterion; Schwarz 1978), corresponding to a significance threshold of  $\text{FAP} = 0.001$  at the yellow dashed horizontal line. Panel c shows a sub-significant signal at a 14.1 day period, likely corresponding to  $P_{\text{rot}}/2$ .

Preliminary RV modeling revealed no other RV signals and insignificant detections of a trend and curvature in our RV time-series, providing no evidence of further companions. We also found an estimated RV jitter of  $\sigma \approx 5.5 \text{ m s}^{-1}$ . This jitter measurement was consistent with both our  $FF'$  estimate of RV variability and the RV semi-amplitude of the marginal 14.1 day signal ( $K = 4 \pm 1 \text{ m s}^{-1}$ ). Given the low significance of this additional signal and its sub-jitter amplitude, we chose to only consider the two planetary signals in our final RV models. We therefore interpreted the spectroscopic data for TOI-1272 to reveal 2 planetary signals (3.3 days and 8.7 days), with a sub-significant activity signal driven by stellar rotation ( $P_{\text{rot}}/2 \approx 14.1$  days).

### 3.6.3 RV-only Model

We performed a two-planet fit to the RV time series for TOI-1272 using `RadVel`, a Python package used to characterize planets from Keplerian RV signals by applying maximum *a posteriori* model fitting and parameter estimation via MCMC (Fulton et al. 2018). Our model consisted of two planetary Keplerian signals with periods 3.3 and 8.7 days. We modeled the data by fitting the following free parameters for both planets:  $P$ ,  $T_c$ ,  $K$ ,  $\sqrt{e} \cos \omega$ , and  $\sqrt{e} \sin \omega$ . Our model also included RV offset  $\gamma$  and RV jitter term  $\sigma$  to account for astrophysical white noise and instrumental uncertainty. The best-fit RV-only `RadVel` model confirmed the existence of two eccentric sub-Jovian mass planets orbiting TOI-1272, and we used these results to inform the priors for a joint RV-photometry model.

### 3.6.4 RV-Photometry Joint Model

We obtained the most precise planet parameters for the TOI-1272 system by performing global RV-photometry modeling using `juliet` (Espinoza et al. 2018), a robust tool for modelling both transiting and non-transiting exoplanets. We used `juliet` to jointly fit the radial velocities through `RadVel` and the transit photometry through `batman` (Kreidberg 2015), with proper handling of limb-darkening coefficients (Kipping 2013). Estimation and comparison

of Bayesian evidences and posteriors was performed directly by the dynamic nested sampling package `dynesty` (Speagle 2020), one of several such tools offered through the `juliet` interface. Unlike the Monte Carlo algorithm used in our initial transit-only analysis, nested sampling algorithms break up complex posterior distributions into simpler nested slices, sampling from each slice individually then recombining the weighted results to reconstruct the complete posterior. This method becomes more efficient in the higher-dimensional posterior spaces of joint models.

We directly fit for each transit and Keplerian property with priors informed from our previous photometry-only and RV-only models. For the final global model, we fit for photometry-only properties  $\{R_p/R_*, b, \rho_*, \mu, u, v\}$ , joint properties  $\{P_b, t_{0,b}, \sqrt{e_b} \sin \omega_b, \sqrt{e_b} \cos \omega_b\}$ , and RV-only properties  $\{P_c, t_{0,c}, \sqrt{e_c} \sin \omega_c, \sqrt{e_c} \cos \omega_c, K_b, K_c, \gamma, \sigma\}$ . Our final measurements are included in Table 3.2 and the corresponding maximum *a posteriori* RV model is shown in Figure 3.7.

In summary, we measured mass constraints for TOI-1272 b and c at significance levels  $\sim 11\sigma$  and  $\sim 9\sigma$ , respectively, reflecting the strengths of the two periodogram signals discussed in §3.6.1. We also measured a high eccentricity of  $e_b = 0.34 \pm 0.06$  for TOI-1272b, consistent within  $1\sigma$  of our photometry-only eccentricity constraint from §3.3.3. The eccentricity of the outer planet was loosely constrained to  $e_c = 0.12^{+0.1}_{-0.08}$ . We note, however, that a model fit with  $e_c = 0$  performed nearly identically to the eccentric model, suggesting that the eccentricity of TOI-1272c is only marginally significant. We discuss these constraints on eccentricity further in §3.7.1. Our global model also served to minimize degeneracies between  $e$ - $\omega$ - $b$  and allowed us to obtain more precise  $b$  and  $R_{p,b}$  values than with our photometry-only model. Our loose posterior constraint on impact parameter from Figure 3.3 was improved to  $b = 0.45^{+0.15}_{-0.21}$ , subsequently yielding our final radius measurement of  $R_{p,b} = 4.14 \pm 0.21 R_\oplus$ .

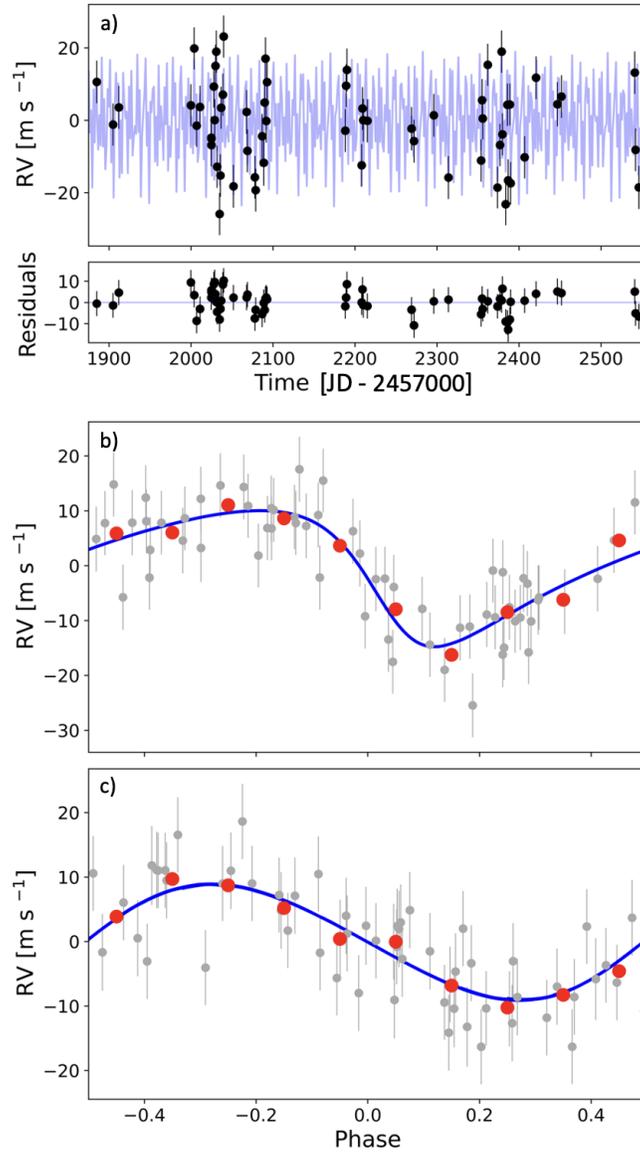


Figure 3.7: (a) Best-fit radial velocity model (blue) for Keck/HIRES RV measurements (black) using `Radvel` (Fulton et al. 2018) in a joint RV-photometry model via `juliet` (Espinoza et al. 2018), with corresponding residuals shown below. (b)-(c) Phase-folded views of best-fit RV model for TOI-1272 b and c, with binned points shown in red.

Table 3.2: TOI-1272 System Properties

Parameter	Value	Notes
<i>Stellar</i>		
RA ( $^{\circ}$ )	199.1966	A
Dec ( $^{\circ}$ )	49.86104	A
$\pi$ (mas)	$7.24 \pm 0.021$	A
$m_K$	$9.70 \pm 0.02$	B
$T_{\text{eff}}$ (K)	$4985 \pm 121$	C
[Fe/H] (dex)	$0.17 \pm 0.06$	C
$\log(g)$	$4.55 \pm 0.10$	C
$M_*$ ( $M_{\odot}$ )	$0.851 \pm 0.049$	C
$R_*$ ( $R_{\odot}$ )	$0.788 \pm 0.033$	C
$\rho_*$ ( $\text{g cm}^{-3}$ )	$2.453 \pm 0.343$	C
age (Gyr)	$3.65^{+4.17}_{-0.98}$	D
$P_{\text{rot}}$ (days)	$28.3 \pm 0.6$	E
$S_{\text{HK}}$	0.331	F
$\log R'_{\text{HK}}$	-4.705	F
$u, v$	$0.39 \pm 0.05, 0.09 \pm 0.05$	G
$\gamma$ ( $\text{m s}^{-1}$ )	$0.7 \pm 0.7$	H
$\sigma_{\text{jit}}$ ( $\text{m s}^{-1}$ )	$5.6 \pm 0.6$	H
<i>Planet b</i>		
$P$ (days)	$3.31599 \pm 0.00002$	H
$T_c$ (BJD-2457000)	$1713.0253 \pm 0.0006$	H
$b$	$0.45^{+0.15}_{-0.21}$	H
$R_p$ ( $R_{\oplus}$ )	$4.14 \pm 0.21$	H
$M_p$ ( $M_{\oplus}$ )	$24.6 \pm 2.3$	H
$\rho_p$ ( $\text{g cm}^{-3}$ )	$1.9 \pm 0.3$	H
$K$ ( $\text{m s}^{-1}$ )	$12.6 \pm 1.1$	H

Table 3.2 (cont'd): TOI-1272 System Properties

Parameter	Value	Notes
$a$ (AU)	$0.0412 \pm 0.0008$	H
$e$	$0.338^{+0.056}_{-0.062}$	H
$\omega$ ( $^\circ$ )	$123.6 \pm 11.5$	H
$T_{\text{eq}}$ (K)	$961 \pm 32$	I
<i>Planet c</i>		
$P$ (days)	$8.689 \pm 0.008$	H
$T_c$ (BJD-2457000)	$1885.34 \pm 0.48$	H
$M_p \sin i$ ( $M_\oplus$ )	$26.7 \pm 3.1$	H
$K$ ( $\text{m s}^{-1}$ )	$9.4 \pm 1.0$	H
$a$ (AU)	$0.0783 \pm 0.0014$	H
$e$	$\lesssim 0.35$	J
$\omega$ ( $^\circ$ )	$-80.8^{+97.4}_{-57.3}$	H
$T_{\text{eq}}$ (K)	$697 \pm 23$	I

Note. — A: *Gaia* DR2, epoch J2015.5 (Gaia Collaboration et al. 2018); B: *2MASS* (Skrutskie et al. 2006); C: Derived with `isoclassify`; D: Derived with `stardate` (Angus et al. 2019a); E: Derived with `TESS-SIP` (Hedges et al. 2020); F: Measured from Keck/HIRES template; G: Derived with `LDTK` (Parviainen & Aigrain 2015); H: Constrained from joint RV-photometry model with `juliet` (Espinoza et al. 2018; Kreidberg 2015, Fulton et al. 2018, Speagle 2020); I: Calculated from other parameters assuming albedo  $\alpha = 0.3$ ; J: Dynamically constrained with `rebound` (Rein & Liu 2012).

## 3.7 System Dynamics

### 3.7.1 Eccentricity Constraints from Stability Requirements

Despite the compact architecture of the TOI-1272 system, both planets had moderate RV-constrained eccentricities that were inconsistent with zero to  $\sim 5\sigma$  and  $\sim 1\sigma$  significance, respectively. Such excited dynamics put TOI-1272 b and c at risk of dynamical instability if orbit crossing were to occur:

$$\frac{a_c(1 - e_c)}{a_b(1 + e_b)} > 1. \quad (3.2)$$

Given our RV-constrained measurements of eccentricity and orbital separation in this system, we found the left-hand side of the above equation to be  $1.22 \pm 0.15$ , or  $< 2\sigma$  from the orbit crossing threshold. We note, however, that our confidence in long-term orbital stability based on this value is highly sensitive to our eccentricity uncertainties. Assuming a fixed true value of  $e_c = 0.05$ , this system would be firmly out of reach of geometric orbit crossing given its current configuration.

With the ambiguity in our orbit-crossing stability result, we also calculated the dynamical stability of the TOI-1272 system according to the stricter criterion from [Petrovich \(2015\)](#):

$$\frac{a_c(1 - e_c)}{a_b(1 + e_b)} - 2.4 \times \max(\mu_b, \mu_c)^{1/3} \left(\frac{a_c}{a_b}\right)^{1/2} > 1.15, \quad (3.3)$$

where  $\mu$  is  $M_p/M_*$ , drawn from our joint model results. This threshold marks an estimated empirical boundary in two-planet system stability, determined by applying a Support Vector Machine algorithm to a large number of numerical integrations. Planet-planet interactions resulting in ejecting a planet into the star or out of the system were considered by [Petrovich \(2015\)](#) in developing this criterion.

Systems that satisfy the condition in Eq. 3.3 are expected to maintain dynamical stability for integrations out to at least  $10^8$  orbits of the inner planet. When computed for this system, we measured the left-hand side of Eq. 3.3 to be  $1.1 \pm 0.15$ , or  $<1\sigma$  below the stated stability threshold of 1.15. Similar to the orbit crossing criterion, TOI-1272 straddles the stability boundary for the Petrovich (2015) empirical threshold. We again note that a fixed outer planet eccentricity of  $e_c = 0.05$  would promote the long-term stability of the TOI-1272 system according to this stability criterion.

We followed up these inconclusive analytical predictions of long-term stability with a full N-body treatment of the stability of the TOI-1272 system. Drawing initial conditions from our RV-photometry model posteriors, we ran  $10^4$  N-body simulations with `rebound` (Rein & Liu 2012) for  $\sim 10^6$  orbits of the inner planet. We restricted the initialized eccentricities of our simulations to avoid starting on crossing orbits, and we considered a simulation to be "unstable" after an orbit-crossing event or close dynamical encounter. Overall,  $\sim 81\%$  of simulations remained "stable" for the entirety of our integration time, suggesting that the eccentricities and masses measured from our RV model were largely consistent with a stable architecture on moderate time-scales (Figure 3.8). Our `rebound` simulations also showed that stable configurations of this system exhibit Laplace-Lagrange oscillations in eccentricity with a secular timescale on the order of  $\sim 10^2$  years.

While our  $e_b$  posterior remained mostly unchanged by this N-body model, our stability constraints on  $e_c$  allowed us to determine an upper bound of  $e_c \lesssim 0.35$ . Upon this redefinition of  $e_c$ , we inferred that the true eccentricity of TOI-1272c was likely in the lower tail of the acceptable range. Given that a `RadVel` model with  $e_c = 0$  performed nearly equivalently to the non-zero eccentricity model (§3.6.4), this interpretation is consistent with our RV-only analysis.

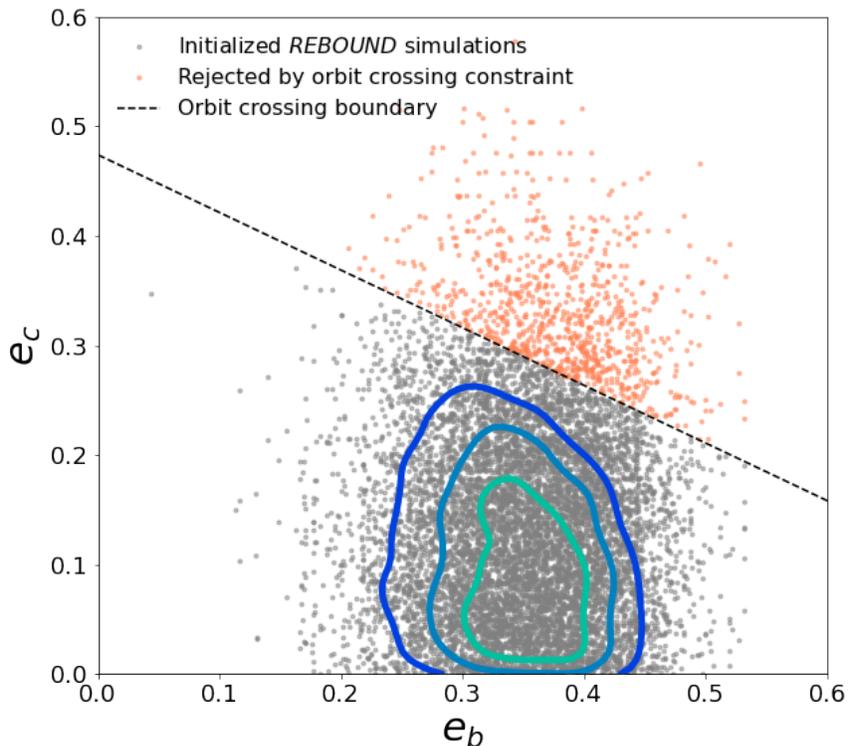


Figure 3.8: Distribution of initialized  $e_b$ - $e_c$  values for all `rebound` simulations of the TOI-1272 system. Green-blue contours show regions with highest density of stable configurations (simulations lasting  $10^6$  orbits of the inner planet).

### 3.7.2 No Evidence for TTVs

Along with our N-body integration, we also used `rebound` to model the transit-timing variations (TTVs) expected to be observed in this system over a similar baseline as our TESS photometry ( $\sim 215$  days). We estimated a TTV O-C RMS of  $0.3 \pm 0.3$  minutes, below our threshold of sensitivity for individual transits. We verified this empirically by modeling the TESS photometry with `exoplanet`, similar to §3.3.3 but this time including TTVs as an additional model parameter. From this fit, we measured a TTV O-C RMS of  $2 \pm 2$  minutes, consistent with the estimate from `rebound`.

With the additional photometric observations from MuSCAT, we extended our TTV search

to a total photometric baseline of  $\sim 600$  days. A fit to the single transit measured by MuSCAT yielded a transit mid-point of  $\text{BJD}-2457000 = 2313.223$ , only  $\sim 3$  minutes off from predicted mid-point of  $2313.221 \pm 0.003$  and within the  $\sim 5$  minute uncertainty of this prediction. The extended photometric baseline demonstrated again that the TTVs in this system are negligible, within  $\sim 1\sigma$  of showing no evidence for TTVs.

The lack of TTVs was also consistent with the non-resonant orbital period ratio between TOI-1272 b and c:  $P_c/P_b \approx 2.62$ . This period ratio is outside of the resonant width of any strong resonances, lying most closely to the 3:1 second-order mean-motion resonance (MMR), with a  $\sim 14\%$  difference in ratio. However, we cannot rule out the possibility of past planet migration leading to resonance-crossing, which could have played a role in the planet-planet excitation of  $e_b$  discussed briefly in §3.8.3.

### 3.7.3 Strong Tidal Eccentricity Decay

The age measurement of TOI-1272 from §3.5.3 is valuable when considering the potential tidal eccentricity decay of TOI-1272b. According to Millholland et al. (2020), which draws from Leconte et al. (2010), the timescale of orbital circularization due to tidal eccentricity damping for an eccentric orbit is given by

$$\tau_e = \frac{4}{99} \left( \frac{Q'}{n} \right) \left( \frac{M_p}{M_*} \right) \left( \frac{a}{R_p} \right)^5 \times \left( \Omega_e(e) \cos \epsilon \left( \frac{\omega_{\text{eq}}}{n} \right) - \frac{18}{11} N_e(e) \right)^{-1}. \quad (3.4)$$

Here, the mean motion is given by  $n = \sqrt{GM_*/a^3}$  and the reduced tidal quality factor  $Q'$  can be rewritten as  $Q' = 3Q/2k_2$ , with specific dissipation function  $Q$  and tidal Love number  $k_2$  (Murray & Dermott 1999; Mardling & Lin 2004). We defined  $\omega_{\text{eq}}$  as the spin rotation frequency of TOI-1272b at equilibrium, which we found to be  $3 \pm 0.5 \text{ day}^{-1}$  following the procedure outlined in Millholland et al. (2020). We assumed the obliquity  $\epsilon$  to be  $0^\circ$ . We have also introduced functions of eccentricity  $\Omega_e(e)$  and  $N_e(e)$  given by

$$\Omega_e(e) = \frac{1 + \frac{3}{2}e^2 + \frac{1}{8}e^4}{(1 - e^2)^5} \quad (3.5)$$

$$N_e(e) = \frac{1 + \frac{15}{4}e^2 + \frac{15}{8}e^4 + \frac{5}{64}e^6}{(1 - e^2)^{\frac{13}{2}}}. \quad (3.6)$$

A typical Neptune-like planet is generally assumed to have a tidal quality factor of  $Q' \approx 10^5$ , but the true value is highly uncertain. Assuming this fixed value for  $Q'$  and drawing the other parameters in Eq. 3.4 from our previous analysis, we estimated a circularization timescale of  $\tau_e \approx 0.21 \pm 0.09$  Gyr. This nominal value of  $\tau_e$  is  $>3\sigma$  below our age measurement of  $3.65^{+4.17}_{-0.98}$  Gyr, suggesting that TOI-1272b has experienced significant eccentricity decay due to tides. This is not reflected in the anomalously high eccentricity that we measured, suggesting that another mechanism must be driving the excited state of this system. We note, however, that  $Q'$  is highly uncertain and  $\tau_e \propto Q'$ , so a tidal quality factor of  $2 \times 10^6$  would make  $\tau_e$  consistent with the age of the system.

Continuing with the assumption of  $Q' \approx 10^5$ , we estimated the initial eccentricity that would have been needed for TOI-1272b to reach its currently observed  $e_b$  after 3.65 Gyr of tidal eccentricity decay. Assuming constant  $Q'$  and  $\tau_e$ , we followed the procedures of [Correia et al. \(2020\)](#) to derive the required post-formation eccentricity of  $e_b \approx 0.8$ . Without a significant restructuring of the TOI-1272 system architecture, however, such a high eccentricity would not have allowed for a stable companion at the orbital separation of TOI-1272c. Ruling out this "hot-start" scenario, we are left to consider whether the anomalously high eccentricity of the inner planet is due to an underestimated  $Q'$  or excitation by some other dynamical mechanism. We discuss such formation and evolution scenarios further in §3.8.3.

## 3.8 Context in the Hot Neptune Desert

### 3.8.1 Bulk Density and Core-Envelope Fraction

TOI-1272b is a Neptune-like planet for which we measured a mass of  $24.6 \pm 2.3 M_{\oplus}$  and a radius of  $4.1 \pm 0.2 R_{\oplus}$ , yielding a density of  $1.9 \pm 0.3 \text{ g cm}^{-3}$ . This system also contains a similar-mass outer companion ( $M_p \sin i = 26.7 \pm 3.1 M_{\oplus}$ ) that is not transiting. Planets in this size and mass range have been reported frequently in the literature, only a few of which also fall within the Hot Neptune Desert (Mazeh et al. 2016; Owen & Lai 2018). At moderate planet sizes ( $\sim 2\text{--}6 R_{\oplus}$ ;  $M_p \lesssim 100 M_{\oplus}$ ) and low orbital separations ( $P \lesssim 5$  days), a relative paucity of planets has been observed.

The triangular regions of  $R_p$ - $P$  and  $M_p$ - $P$  parameter space shown in Figure 3.9 highlight this phenomenon, as defined by Mazeh et al. (2016). TOI-1272b can be seen here among the small subset of Neptunes that fall within this otherwise sparse parameter space. Some notable inhabitants of the Hot Neptune Desert include GJ 436b (Lanotte et al. 2014) and HAT-P-11b (Yee et al. 2018) along with more recent finds from TESS photometry including LP 714-47b (Dreizler et al. 2020) and TOI-132b (Díaz et al. 2020). While the the exact mechanism responsible for clearing out this  $R_p$ - $P$  and  $M_p$ - $P$  region remains unknown, some models support a combination of photoevaporation and tidal disruption following high-eccentricity migration (Mazeh et al. 2016; Lundkvist et al. 2016; Owen & Lai 2018).

Planets within the Hot Neptune Desert range from dense, atmosphere-poor mini-Neptunes to atmosphere-rich, "puffy" super-Neptunes. TOI-1272b lies in the middle of this spectrum, with an elevated density relative to the upper end of the Weiss & Marcy (2014) relation. We used a 2-component composition model to determine the relative abundances of solid core and gaseous envelope for this dense Neptune, following the procedure of MacDougall et al. (2021). We interpolated over a 4D grid of stellar and planetary properties to derive the expected envelope mass fraction for TOI-1272b using the Lopez & Fortney (2014) planet structure

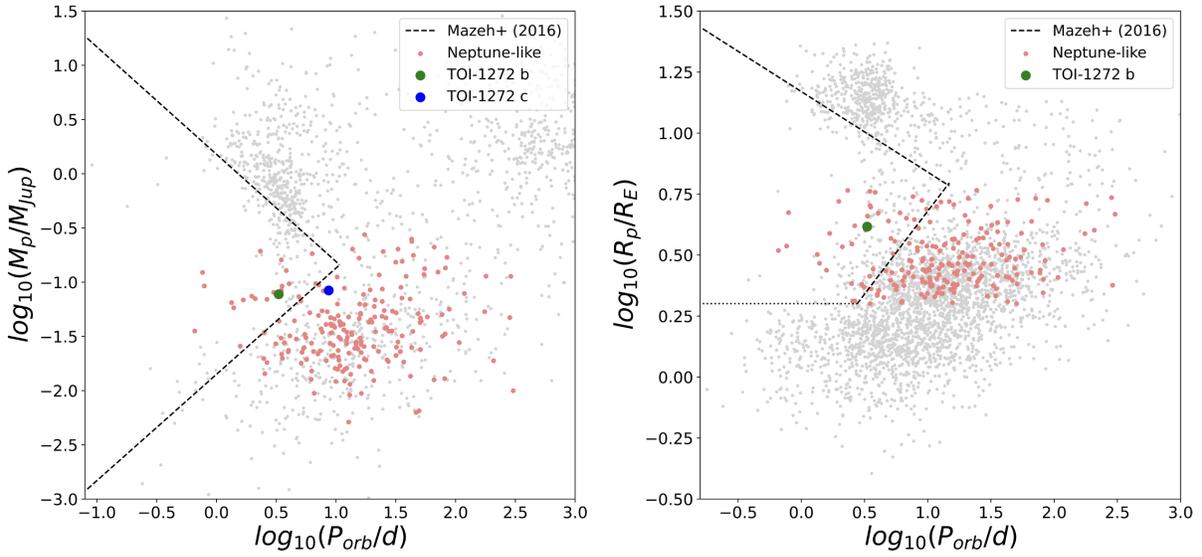


Figure 3.9: Hot Neptune Desert in  $R_p$ - $P$  and  $M_p$ - $P$  parameter space, where Neptune-like is defined as  $M_p < 100 M_\oplus$  and  $R_p = 2.0 - 6.0 R_\oplus$ . Data drawn from the NASA Exoplanet Archive (Akeson et al. 2013a; NASA Exoplanet Archive 2022a). TOI-1272 b and c depicted as green point (both panels) and blue point (left panel), respectively. Relations dictating the upper and lower boundaries of the Hot Neptune Desert from Mazeh et al. (2016) shown as black dashed lines. Left: Mass-period distribution of all planets, with Neptune-like planets shown in red. Right: Radius-period distribution of all planets, with Neptune-like planets shown in red and  $R_p = 2 R_\oplus$  limit shown as black dotted line.

models. Assuming an earth-like core composition and a solar-composition H/He envelope, we estimated  $f_{\text{env}} = 10.9 \pm 2.0\%$  and a core mass of  $21.9 \pm 2.0 M_\oplus$ . Given the strong stellar irradiance experienced by this planet, with  $T_{\text{eq}} \approx 960$  K (assuming albedo  $\alpha = 0.3$ ), TOI-1272b could have begun as a more atmosphere-rich Neptune similar to GJ 3470b (Kosiarek et al. 2019) and experienced subsequent atmosphere loss. TOI-1272b may then serve as a strong candidate for follow-up atmospheric observations, following the treatment of similar targets like those discussed by Crossfield & Kreidberg (2017).

The outer companion in this system, TOI-1272c, likely falls into the same size category as

TOI-1272b, with  $M_{p,c} \sin i = 27.4 \pm 3.2 M_{\oplus}$ . However, since no transit was detected in TESS photometry, we were unable to make any claims regarding its density or composition. One might suppose that a sufficiently low-radius planet on an 8.7-day orbit could produce a transit signal below the detection threshold of  $S/N \approx 7.1$ . Assuming a transit duration of  $T_{14} \approx 0.15$  days and the same noise properties as the TOI-1272b transit, this would require  $R_{p,c} \lesssim 2.3 R_{\oplus}$  and  $\rho_{p,c} \gtrsim 12.0 \text{ g cm}^{-3}$ . While this density is not entirely unreasonable (see, e.g., Kepler-411b; Sun et al. 2019), it is unlikely given the known sample of similar planets.

### 3.8.2 Eccentricities

The eccentricity distribution of hot Neptunes was discussed in depth by Correia et al. (2020) who noted that such planets exhibit elevated eccentricities despite being on compact orbits. We reconsidered this claim using a more recent set of confirmed planet data from the NASA Exoplanet Archive (Akeson et al. 2013a; NASA Exoplanet Archive 2022a), including TOI-1272b in our sample. We considered Neptunes to have radii  $\sim 2\text{--}6 R_{\oplus}$  and "hot" planets to have  $P < 5$  days. Constraining our sample to only planets with eccentricity uncertainties less than 0.1 (Figure 3.10), we found that hot Neptunes ( $N = 17$ ) displayed a broader eccentricity distribution than their longer period counterparts ( $N = 75$ ). We verified the distinction between the two distributions using a Kolmogorov-Smirnov (KS) test, finding  $p \approx 0.008$ . TOI-1272b contributed to this significant trend among hot Neptunes.

On the other hand, planets with radii  $> 6 R_{\oplus}$  showed the opposite trend as can also be seen in Figure 3.10, verified by a KS test with  $p \ll 0.01$ . In this radius range, 66 planets had  $P < 5$  d and 88 had  $P \geq 5$  d. We did not consider planets with radii  $< 2 R_{\oplus}$  in this analysis due to the low sample size of such planets with eccentricity uncertainties  $< 0.1$ . However, a cursory examination of this small subset suggested similar eccentricity distributions between shorter and longer period planets in this size range.

While these findings are statistically significant based on KS tests given the current data,

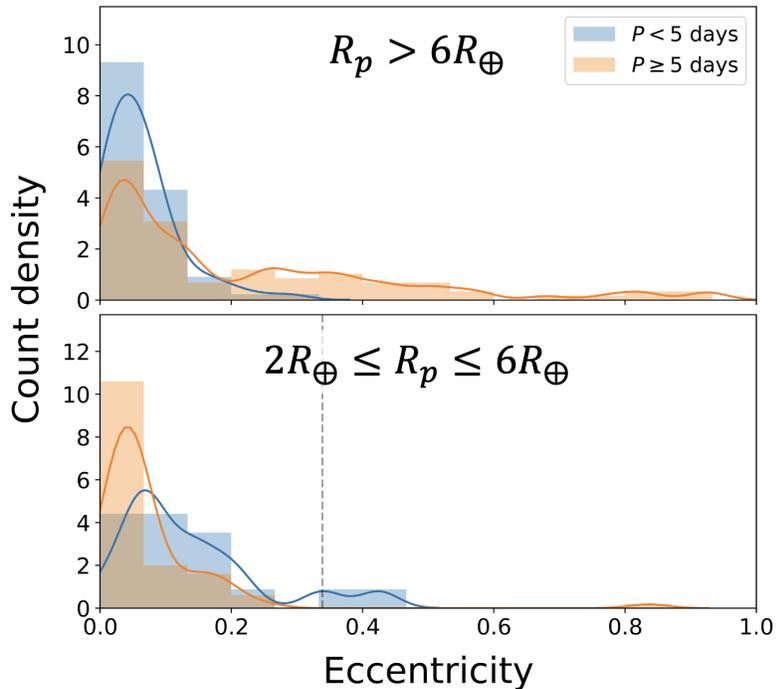


Figure 3.10: Eccentricity distribution of planets with eccentricity uncertainties of  $\sigma_e < 0.1$ , comparing shorter period planets (blue) against longer period planets (orange). The lines show kernel density estimation fits to the data of the corresponding color. Top: Jupiter-like planets ( $> 6 R_\oplus$ ). Bottom: Neptune-like planets ( $2\text{--}6 R_\oplus$ ), including TOI-1272b (dashed gray line).

several confounding factors led us to determine that these eccentricity trends are suggestive rather than definitive at this time. These factors include the small sample size of hot Neptunes, the reliability of the data reported by the NASA Exoplanet Archive, and potential observational biases. Nonetheless, the possible disagreement between the eccentricity trend for hot Neptunes versus hot Jupiters is an active area of research. The low eccentricities of hot Jupiters were consistent with the rapid tidal circularization timescales expected at lower semi-major axes. Conversely, hot Neptunes seem more likely to violate the rule.

A nominal empirically-derived periastron distance of  $r_{\text{peri}} \approx 0.03$  au is often used to approximate the boundary of rapid tidal eccentricity decay, shown in Figure 3.11. Here, we see

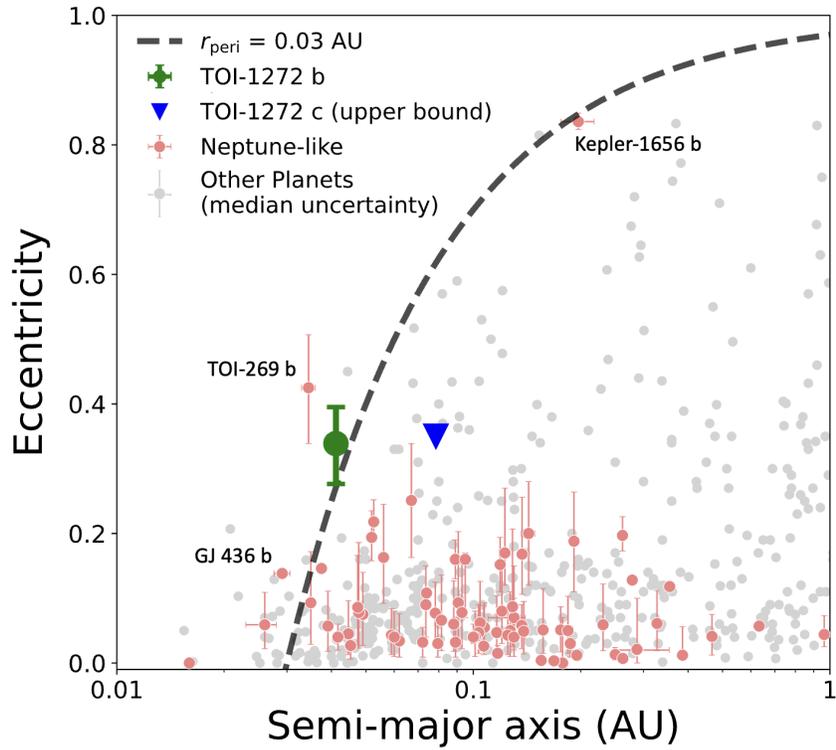


Figure 3.11: Eccentricity distribution of planets with  $\sigma_e < 0.1$  (Neptune-like in red, other known planets in gray) as a function of orbital separation, showing TOI-1272 b and c in green and blue, respectively. Periastron distance of 0.03 AU is shown as an empirical threshold for rapid tidal eccentricity decay. Three eccentric Neptunes are labeled for reference (Kepler-1656b; TOI-269b, [Cointepas et al. 2021](#); GJ 436b).

that only a small subset of well-characterized planets inhabit the high-eccentricity area of parameter space beyond this boundary, including TOI-1272b and a few other eccentric hot Neptunes. The well-studied planet GJ 436b is among these Neptunes with orbits that disagree with tidal circularization, making it a near-twin to TOI-1272b based on mass, radius, eccentricity, and period.

### 3.8.3 A Unique Formation and Evolution Pathway

The sparsity of the Hot Neptune Desert along with the counter-intuitive trend in hot Neptune eccentricities suggests a unique evolutionary pathway for hot planets within the Neptune size regime. Several studies have sought to explain the dearth of planets within the "desert" region of  $M_p$ - $P$  and  $R_p$ - $P$  parameter space. The leading hypothesis suggests a combination of photoevaporation (Owen & Wu 2013; Owen & Lai 2018) and high-eccentricity migration (Mazeh et al. 2016; Matsakos & Königl 2016). Interestingly, photoevaporation is also cited as a possible mechanism for maintaining non-zero eccentricities among hot Neptunes (Correia et al. 2020), along with planet-planet excitation (see, e.g. Jurić & Tremaine 2008; Chiang & Laughlin 2013b) or an Eccentric Kozai-Lidov (EKL) effect from a distant giant companion (Naoz 2016). The persisting eccentricities of some hot Neptunes could also simply be a result of  $Q'$  values underestimated by an order of magnitude or more, which would make them inconsistent with the  $Q'$  values measured for Neptune and Uranus through interior modeling.

An underestimated  $Q'$  could certainly be the case for TOI-1272b, contributing to a longer  $\tau_e$  and slower rate of eccentricity decay. TOI-1272b may also be undergoing significant photoevaporation given its  $f_{\text{env}}$  and close-in  $r_{\text{peri}}$ , contributing to both its location in the middle of the Hot Neptune Desert and its high eccentricity. However, TOI-1272b differs from the plausible formation and evolution pathways of other hot Neptunes due to the presence of a stable, nearby outer companion. Both high-eccentricity migration and perturbations from a distant companion through EKL effects are complicated by the presence of the mildly eccentric companion on an 8.7-day orbit. Such excitation mechanisms would have likely caused an orbit-crossing event and subsequent ejection of one or both planets.

Instead, we propose that, along with photoevaporation, TOI-1272b has experienced minor planet-planet excitation events with TOI-1272c, possibly involving close-approaches or resonance-crossing events during migration (Ford & Rasio 2008). These events could have

contributed to both the high eccentricity and possible inward migration of TOI-1272b into the Hot Neptune Desert region of parameter space, similar to the proposed evolution of mini-Neptune HIP-97166b (MacDougall et al. 2021). The elevated eccentricity of the inner planet may then persist in spite of strong tidal forces through Laplace-Lagrange oscillations that continue to force the eccentricity, as seen in our dynamical simulations mentioned in §3.7.1. However, additional considerations such as the relative inclination of the two planets may be necessary for a more detailed description of the dynamical evolution of this system.

### 3.9 Summary and Conclusions

In this work, we introduced a newly discovered planet within the Hot Neptune Desert around TOI-1272 with a 3.3-day orbital period. We predicted that this planet might have a high eccentricity based on a mismatch between the observed transit duration and the expected duration for a circular orbit upon modeling 3 sectors of transit photometry from TESS. We confirmed this high eccentricity with follow-up RV measurements and verified its stability through dynamical constraints, yielding  $e_b = 0.34 \pm 0.06$ . We also identified a non-transiting outer companion on an 8.7-day orbit, placing a limit on its eccentricity of  $e_c \lesssim 0.35$ . TOI-1272b is now one of only a handful of close-in Neptunes with a well-constrained high eccentricity. The high eccentricity of this inner planet persists in spite of strong tidal forces, likely as a result of either underestimated tidal quality factors for close-in exo-Neptunes or stable dynamical interactions with the outer planet that continue to pump the eccentricity. Nonetheless, the discovery of TOI-1272 b and c has boosted the sample size of a small and poorly understood class of planets, contributing to ongoing studies of hot Neptunes and eccentric short-period planets.

## CHAPTER 4

# Implicit biases in transit models using stellar pseudo-density

### 4.1 Abstract

The transit technique is responsible for the majority of exoplanet discoveries to date. Characterizing these planets involves careful modeling of their transit profiles. A common technique involves expressing the transit duration using a density-like parameter,  $\tilde{\rho}$ , often called the "circular density." Most notably, the *Kepler* project – the largest analysis of transit lightcurves to date – adopted a linear prior on  $\tilde{\rho}$ . Here, we show that such a prior biases measurements of impact parameter,  $b$ , due to the non-linear relationship between  $\tilde{\rho}$  and transit duration. This bias slightly favors low values ( $b \lesssim 0.3$ ) and strongly disfavors high values ( $b \gtrsim 0.7$ ) unless transit signal-to-noise ratio is sufficient to provide an independent constraint on  $b$ , a criterion that is not satisfied for the majority of *Kepler* planets. Planet-to-star radius ratio,  $r$ , is also biased due to  $r$ – $b$  covariance. Consequently, the median *Kepler* DR25 target suffers a 1.6% systematic underestimate of  $r$ . We present techniques for correcting these biases and for avoiding them in the first place.

### 4.2 Introduction

In the two decades since the discovery of the first transiting hot Jupiter (Charbonneau et al., 2000; Henry et al., 2000), the transit technique has grown to be the most prolific exoplanet

detection method to date, accounting for 77% of the current census. Contemporary work continues to rely heavily on the transit technique. To wit, several transit-focused NASA and ESA missions are either already on-sky (TESS, [Ricker et al. 2015b](#)) or slated for launch in the near future (PLATO, [Rauer et al. 2014b](#)), and next-generation radial velocity spectrographs have been designed for follow-up characterization of known transiting planets (e.g. KPF, [Gibson et al. 2016](#); MAROON-X, [Seifahrt et al. 2018](#)). The transit technique will remain indispensable for exoplanet astronomy for decades to come.

Accurate modeling of the transit lightcurve is a critical step for characterizing transiting planets. At the most basic level, transit modeling involves computing the time-dependent flux  $F(t)$  of a star obscured by a transiting planet relative to the unobscured flux  $F_0$ . If one assumes a spherical planet and star, this computation depends strictly on the planet-to-star size ratio  $r$ , the (time-dependent) center-to-center sky-projected planet-to-star separation  $z$  (measured in units of  $R_\star$ ), and the radial dependence of the stellar limb-darkening profile  $\{u\}$ . Early analyses computed  $F(z; r, \{u\})$  via numerical integration, but today the most widely used method is the [Mandel & Agol \(2002b\)](#) model, which expresses the transit lightcurve via an analytic solution to  $F(z; r, \{u\})$  for several limb darkening profiles which can be described by a small set of limb-darkening parameters.

In order to model time-series photometry, one must convert  $F(z; r, \{u\})$  into  $F(t; r, \{u\})$ . While  $z$  is the only parameter that varies with time, one may choose how to specify the function that maps  $t \rightarrow z$ . If one assumes strict periodicity of transits and a constant projected velocity during transit., then in the limit  $r \rightarrow 0$ ,  $z(t)$  may be specified completely by an orbital period,  $P$ , a transit mid-point,  $t_0$ , an impact parameter,  $b$ , and 1st-to-4th contact transit duration,  $T_{14}$ .<sup>1</sup> This parameterization —  $F(t; P, t_0, r, b, T_{14})$  — is convenient

---

<sup>1</sup>Several alternative transit durations may be substituted for  $T_{14}$ : (1) the 2nd-to-3rd contact duration,  $T_{23}$ , (2) the center-to-center contact duration,  $T_{cc}$ , also called the 1.5-to-3.5 contact duration, or (3) the full-width-half-max duration,  $T_{FWHM}$ , which may be defined in relation to the transit depth. While each has its merits (see [Kipping 2010b](#) for discussion), we adopt  $T_{14}$  throughout this work because it is the transit duration which is most readily defined for all grazing and non-grazing transit geometries.

and is closely linked to the transit geometry.

An alternative approach is to specify  $T_{14}$  from a combination of scaled separation  $a/R_\star$ , orbital eccentricity  $e$ , argument of periastron  $\omega$ , and projected inclination  $\cos i$ , following (Winn, 2010a) as

$$T_{14} \simeq \frac{P}{\pi} \sin^{-1} \left( \frac{R_\star}{a} \frac{\sqrt{(1+r)^2 - b^2}}{\sin i} \right) \left( \frac{\sqrt{1-e^2}}{1+e \sin \omega} \right) \quad (4.1)$$

$$b = \frac{a \cos i}{R_\star} \left( \frac{1-e^2}{1+e \sin \omega} \right) \quad (4.2)$$

Now the lightcurve is specified by the function  $F(t; P, t_0, r, a/R_\star, b, e, \omega)$ , which is similar to the parameterization used by the EXOFAST suite (Eastman et al., 2013a; Eastman, 2017).<sup>2</sup> A related approach is to replace  $a/R_\star$  with stellar density by employing Kepler’s third law. Thus, the light curve may also be parameterized by  $F(t; P, t_0, r, \rho_\star, \cos i, e, \omega)$ .

These two eccentricity-explicit parameterizations have the advantage that the lightcurve has been specified completely by properties of the star, planet, and planetary orbit; the disadvantage is that five parameters have been replaced by seven, and thus significant degeneracies between  $\{a/R_\star, e, \omega\}$  or  $\{\rho_\star, e, \omega\}$  are inevitable. These degeneracies lead to inefficiencies with light curve fitting and posterior sampling.

A common shortcut is to fit the lightcurve assuming that  $e = 0$  even though the orbit may, in fact, be eccentric. This assumption reduces the number of free parameters back to five, but  $\rho_\star$  can no longer be thought of as a stellar density. Rather, it is a stand-in for duration

---

<sup>2</sup>In practice, EXOFAST uses  $\log(a/R_\star)$  and expresses  $b$  as  $\cos i$ ;  $\{e, \omega\}$  is usually specified as  $\{\sqrt{e} \sin \omega, \sqrt{e} \cos \omega\}$  in order to establish uniform priors on  $e$  and  $\omega$  and to avoid a boundary issue at  $e = 0$ .

which merely has *units* of density, defined by Seager & Mallén-Ornelas (2003b) as

$$\tilde{\rho} \equiv \left( \frac{4\pi^2}{P^2 G} \right) \left( \frac{(1+r)^2 - b^2(1 - \sin^2[\pi T_{14}/P])}{\sin^2[\pi T_{14}/P]} \right)^{3/2} \quad (4.3)$$

where  $G$  is Newton’s gravitational constant. This quantity  $\tilde{\rho}$  is sometimes referred to as the “mean stellar density,” the “circular density,” or the “observed density,” but we prefer to call it the “pseudo-density” because (1) the other names are confusing, and (2)  $\tilde{\rho}$  matches the true stellar density only when numerous assumptions are met (see Kipping, 2014a).

Because the prior expectation for  $\tilde{\rho}$  is a complicated function of  $\rho_*$ ,  $b$ ,  $e$ , and  $\omega$ , naïvely placing a flat prior on  $\tilde{\rho}$  and adopting it as a fitting parameter induces undesired biases on  $T_{14}$  and  $b$ .

To date,  $\tilde{\rho}$  has enjoyed widespread use in the exoplanet literature. For example, the *Kepler* project (Borucki et al. 2010c; the largest analysis of transit lightcurves to date) fit their lightcurves with the  $F(t; P, t_0, r, \tilde{\rho}, b)$  parameterization (Rowe et al., 2014a, 2015; Mullally et al., 2015; Coughlin et al., 2016; Thompson et al., 2018). We discuss the effects of that choice in §4.6. More broadly, this paper investigates the implicit biases on impact parameter and other light curve parameters that result from the use of  $\tilde{\rho}$ .

Throughout this work, we assume that all transit signals under investigation have been thoroughly vetted such that the detected signal is known to be a real transit at high confidence. The methods employed in this work are thus appropriate for parameter estimation but not for transit detection or vetting.

This paper is organized as follows. In §4.3 we empirically demonstrate the origin of the  $\tilde{\rho}$  bias by fitting a transit lightcurve model to simulated photometry using the *Kepler* project parameterization; we then demonstrate that our preferred parameterization does not suffer from this bias. In §4.4 we present a numerical experiment which isolates the effects of various model assumptions on posterior inferences. In §4.5 we analytically derive the Jacobian of

the coordinate transformation  $T_{14} \rightarrow \tilde{\rho}$  which explains the origin of the empirical bias. In §4.6 we show that the  $\tilde{\rho}$  bias has affected most posterior inferences of  $b$  and  $r$  derived from *Kepler* data. In §4.7 we summarize our conclusions and discuss other biases which arise from using related parameterizations such as  $a/R_*$ .

### 4.3 Understanding parameter biases with fits to synthetic photometry

To illustrate the  $\tilde{\rho}$  bias, we simulated photometric observations of a warm mini-Neptune ( $P = 15$  days,  $r_p = 3.3 R_\oplus$ ) on a circular orbit around a Sun-like star, transiting at impact parameter  $b = 0.5$ . We simulated data with a 30 minute observing cadence (matching *Kepler*'s long cadence observing mode) within  $\pm T$  from the transit center. All photometric data were oversampled by a factor of 7 and integrated using Simpson's rule to account for the effects of finite integration time (Kipping, 2010c). The white noise level was tuned to produce  $S/N = 16$ , which is slightly lower than the median *Kepler* value and results in a posterior model with  $\sigma_r/r \approx 0.10$  and  $\sigma_T/T \approx 0.05$ , where  $\sigma_r/r$  corresponds to the fractional posterior measurement, and similar for  $T$ . We chose these values in order to produce a transit which is similar to those found by *Kepler*. Ground-truth simulation parameters are listed in Table 4.1, and simulated photometry is shown in Figure 4.1.

The transit model was specified using a standard pseudo-density parameterization:  $\{P, t_0, r, b, \tilde{\rho}\}$ . In order to minimize confounding factors, we held  $P$  and  $t_0$  fixed at their injected values; we also held the mean out-of-transit flux,  $F_0$ , and photometric white noise level,  $\sigma_{\text{phot}}^2$ , fixed to their true values, which is equivalent to assuming the raw photometry has been accurately pre-whitened. For the remaining transit parameters, we adopted broad weakly informative priors with permissive bounds (see Table 4.2 for details), for a total of three free parameters per model:  $\{r, b, \tilde{\rho}\}$  (the “ $\tilde{\rho}$  basis”), or  $\{\log r, b, \log T_{14}\}$  (the “log  $T$  basis”). We chose the later basis because  $T_{14}$  is typically well constrained by the data and furthermore

Table 4.1: Ground-truth simulation parameters; simulated photometry is shown in Figure 4.1.

Parameter	Value
$M_{\star}[M_{\odot}]$	1.0
$R_{\star}[R_{\odot}]$	1.0
$u_1, u_2$	0.40, 0.25
$P$ [days]	15.0
$r$	0.03
$b$	0.5
$T_{14}$ [hrs]	3.29
$S/N$	16

may be assigned priors in a sensible fashion; sampling in  $\log r$  and  $\log T_{14}$  is equivalent to placing log-uniform priors on  $r$  and  $T_{14}$  which facilitates exploration of posterior values over different orders of magnitude. We modeled a circular transit in all cases and held stellar mass, radius, and limb darkening to their true values during the fit; there is no loss of generality in this approach, because as long as we ignore minuscule ingress/egress asymmetry that exists for eccentric transits (Barnes, 2007b), there is no difference between a circular and eccentric transit. In order to avoid complications which arise when modeling grazing transits, we restricted impact parameters to  $b < 1 - r$ .<sup>3</sup> To confirm that this restriction is permissible, we explored the parameter space near the limb of the star following the methodology of Gilbert (2022) and verified that the simulated transit is inconsistent with a grazing geometry.

We drew samples from the posterior using Hamiltonian Monte Carlo (HMC; Neal, 2011) and

---

<sup>3</sup>A common approach (which we did not adopt) is to draw samples uniformly from the  $r - b$  plane using triangular sampling (Espinoza, 2018). However, naive application of this method induces a marginal prior on  $r$ , so caution must be taken to ensure that priors are established as intended.

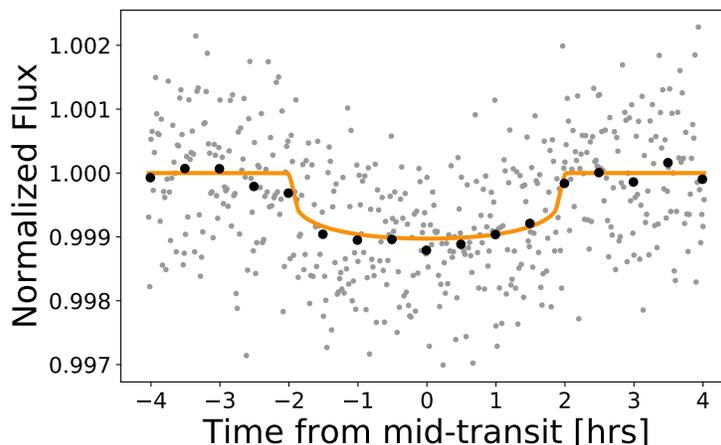


Figure 4.1: Simulated photometry for a mini-Neptune on a circular 15 day orbit around a Sun-like star, transiting at  $b = 0.5$ . The orange line indicates the ground truth transit model. Grey points show simulated observations at a one minute observing cadence; black circles are binned to 30 minutes. The white noise level was set to  $S/N = 16$ , close to the *Kepler* median. See Table 4.1 for ground-truth simulation parameters.

the No U-Turn Sampler (NUTS; Hoffman & Gelman, 2011b). Each model iteration consisted of two chains run for 5000 tuning steps and 20,000 draws, producing an effective number of samples greater than 11,000 for all parameters for each of the the two parameterizations.

Posterior corner plots for the quantities of interest are shown in Figure 4.2. The most notable difference is in the 1D marginalized distribution of impact parameter. When sampling using the  $\tilde{\rho}$  basis, the posterior is biased toward low  $b$ ; as a point of reference, 74% of the probability mass is below  $b = 0.5$ , the injected value. When sampling using the  $\log T$  basis, however, the distribution of  $b$  is nearly uniform over the allowed range, reflecting the fact that for a low signal-to-noise transit the impact parameter is largely unconstrained. Our results did not substantially change when simulating a one minute observing cadence (matching *Kepler*'s short cadence mode), indicating that the  $\tilde{\rho}$  bias arises from the model parameterization and is not an artifact of data binning. We also repeated the analysis using  $r = 0.1$  and  $r = 0.01$  and found that the results did not change.

Table 4.2: Priors on model parameters for simulated lightcurve.

Parameter	Value	Parameter	Value
$r$	$\mathcal{U}(0.01, 0.1)$	$\log r$	$\mathcal{U}(-2, -1)$
$b$	$\mathcal{U}(0, 1 - r)$	$b$	$\mathcal{U}(0, 1 - r)$
$\tilde{\rho}/\rho_{\odot}$	$\mathcal{U}(0.1, 10)$	$\log[T_{14}/\text{hr}]$	$\mathcal{U}(1, 10)$

Clearly, the results are inconsistent between models – which contain identical underlying physics and differ only in their parameter bases – so at least one of the two models has produced biased inference. In the sections that follow, we present both a numerical argument (§4.4) and an analytic argument (§4.5) which demonstrate that the  $\log T$  basis has produced the desired result.

#### 4.4 Numerical sampling experiment

We will now demonstrate that the bias on  $b$  seen in the previous section arises solely from the model parameterization and not from vagaries of the MCMC sampling algorithm or peculiarities of the noise realization in the photometry.

To do so, we performed a numerical experiment which approximated the lightcurve modeling procedure from §4.3 by drawing samples directly from the prior distributions and then applying an *a posteriori* importance weighting designed to mimic the constraints imposed by the photometry. When determining these importance weights, we employed a Gaussian likelihood function and approximated the (covariant) parameter constraints from §4.3 as independent univariate Gaussians. The key advantage of this method is that we no longer needed to directly fit the photometry, thereby eliminating potential confounding factors introduced by the photometry and the sampler.

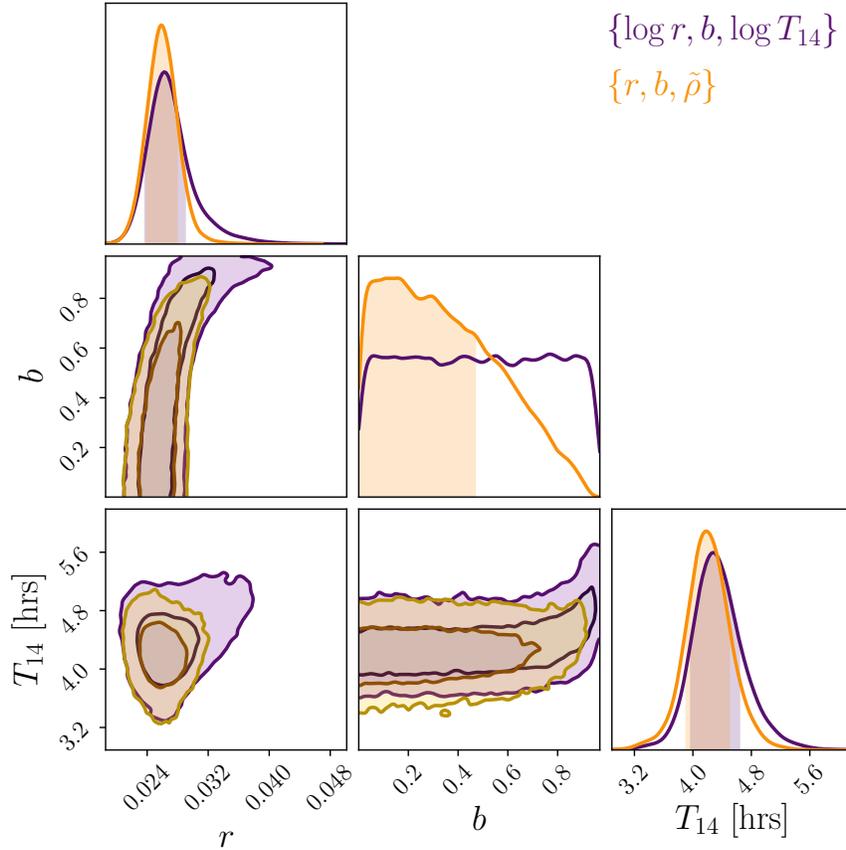


Figure 4.2: Posterior corner plots when sampling in the  $\tilde{\rho}$  basis (orange) vs the  $\log T_{14}$  basis (purple). The bias on impact parameter,  $b$ , is apparent when sampling with  $\tilde{\rho}$  but is resolved when sampling in  $T_{14}$ .

#### 4.4.1 Experimental setup

We adopted the same fiducial star-planet system as §4.3, placing a  $3.3R_{\oplus}$  mini-Neptune on a circular 15 day orbit around a solar twin. We fixed the ephemeris  $\{P, t_0\}$  throughout and placed uniform interval priors on all other parameters  $\{r, b, \tilde{\rho}, \log T_{14}\}$  as before (see Table 4.2), with the small modification that we now allow  $b$  to range over all detectable values, i.e.  $b \sim \mathcal{U}(0, 1 + r)$ ; this modification is acceptable because our sampling procedure (see below) avoids the usual issues which arise when fitting grazing transits (see Gilbert, 2022).

For the first iteration of the experiment we adopted the  $\tilde{\rho}$  basis  $\{r, b, \tilde{\rho}\}$  and drew random samples directly from the prior distributions. We next calculated transit duration using

$$T_{14} = \frac{PR_{\star}}{\pi a} \left( (1+r)^2 - b^2 \right)^{1/2} \quad (4.4)$$

for each sample. Here, we have approximated Equation 4.1 by using the small angle approximation  $\sin^{-1} \phi \approx \phi$  and  $i \approx \pi/2 \rightarrow \sin i \approx 1$ . The scaled separation can be calculated from Kepler’s Third Law as  $a/R_{\star} = [(GP^2\rho)/(3\pi)]^{1/3}$ .

For subsequent iterations of the experiment, we modified the procedure to use three alternative parameter bases: (1)  $\{\log r, b, \log \tilde{\rho}\}$ , (2)  $\{r, b, T_{14}\}$ , and (3)  $\{\log r, b, \log T_{14}\}$ . We chose these parameterizations in order to explore the effects of uniform vs log-uniform priors in addition to the effect of substituting  $\tilde{\rho} \rightarrow T_{14}$ . We followed the same sampling procedure as before, except when drawing samples of  $T_{14}$  or  $\log T_{14}$  we calculated  $\tilde{\rho}$  following Equation 4.3.

Mimicking the simulated light curve in §2, we assumed that we could constrain  $r$  to 10% accuracy and  $T_{14}$  to 5% accuracy, with independent Gaussian precision from the photometry (i.e.  $\sigma_r/r = 0.1$ ,  $\sigma_T/T = 0.05$ ). We further assumed that the impact parameter would be entirely unconstrained by the data. These uncertainties are representative of typical values, but we have removed the covariance and forced them to be Gaussian (or unconstrained), which eases interpretation.

We imposed our assumed measurement uncertainties on  $r$  and  $T_{14}$  by calculating the log-likelihood of each  $i^{\text{th}}$  sample

$$\log \mathcal{L}_i = -\frac{1}{2} \left( \frac{T_i - T_{\text{true}}}{\sigma_T} \right)^2 - \frac{1}{2} \left( \frac{r_i - r_{\text{true}}}{\sigma_r} \right)^2 \quad (4.5)$$

which assumes a Gaussian likelihood function. We then weighted each sample by

$$w_i = \frac{\mathcal{L}_i}{\sum_i \mathcal{L}_i} \quad (4.6)$$

to produce our synthetic posterior distributions.

#### 4.4.2 Bias on impact parameter

The results of our numerical experiment are summarized in Figure 4.3. As expected, when parameterizing the model as  $\{r, b, \tilde{\rho}\}$  with uniform priors, we obtain biased results that are qualitatively similar to those produced in §4.3 (i.e. by fitting the photometry directly). Notably, sampling in  $\tilde{\rho}$  produces a strong prior on  $T_{14}$  (purple) which is not physically motivated. Because  $T_{14}$  is constrained to 5%, the data overwhelm the prior and the  $T_{14}$  posterior distribution (orange) is only slightly biased. The posterior on impact parameter, however, is clearly different from the prior even though our model included no information about impact parameter. Because we have (by construction) placed no measurement constraint on  $b$ , the posterior distribution should match the prior. In reality however, the posterior is tilted toward  $b = 0$ , giving the illusion of a (modestly) constrained posterior.

The  $\tilde{\rho}$ - $b$  bias is resolved by using any of the alternative parameterizations which substitute  $\log \tilde{\rho}$ ,  $T_{14}$ , or  $\log T_{14}$  for  $\tilde{\rho}$ . Although using the substitution  $\tilde{\rho} \rightarrow \log \tilde{\rho}$  may seem at first glance to be the simplest choice (requiring little change from existing practices), we argue that using either of the duration-based parameterizations is preferable for two reasons. First, the results are insensitive to the exact choice of (reasonable) prior placed on  $T_{14}$ , whereas they are highly sensitive to the prior placed on  $\tilde{\rho}$ ; insensitivity to priors is in general a desirable feature of robust inference. Second, setting prior interval bounds on  $\tilde{\rho}$  is a non-intuitive task, requiring careful consideration of the true stellar density and orbital elements. In contrast, principled priors may be placed on the transit duration quite simply following inspection of the transit lightcurve. In fact, setting bounds on  $T_{14}$  is so straightforward that it could even be done algorithmically following the output of a box-least squares transit search (Kovács et al., 2002b). The bottom line is that given the choice between options which produce equivalent results, we prefer the simpler of the two.

In summary, because we have decoupled the posteriors from complicating factors (e.g. parameter covariances, sampler inefficiencies, etc.), we conclude that the differences between

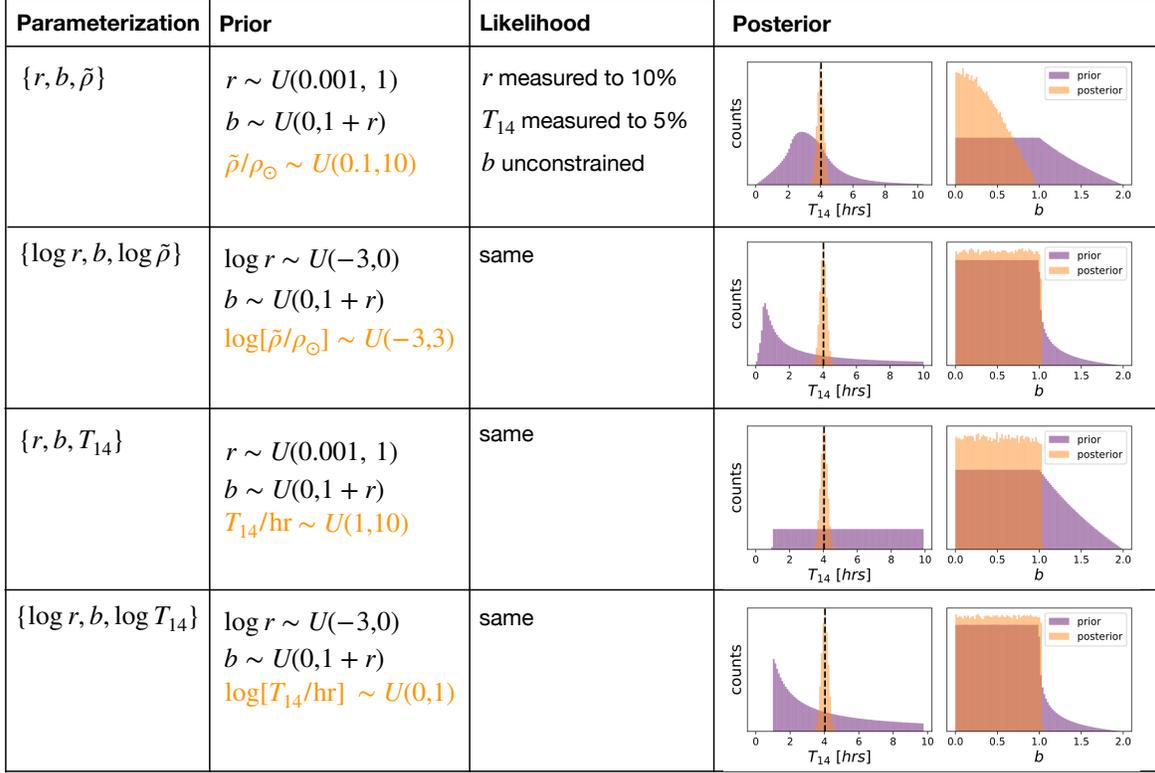


Figure 4.3: Results of the numerical sampling experiment described in §4.4. Each row corresponds to the prior, likelihood, and posterior for a given model parameterization. For visual clarity, the height of the  $T_{14}$  posterior has been reduced by a factor of 3 on all plots. The difference in the prior distribution on  $b$  for rows 1 & 3 compared to rows 2 & 4 stems from the use of  $r$  vs  $\log r$ , respectively. Sampling with a uniform prior on  $\tilde{\rho}$  (top row) produces a nonuniform prior on  $T_{14}$  and a biased posterior for  $b$ . In contrast, sampling in any of the other parameter bases produces a posterior estimate of  $b$  which matches the prior, except in cases where constraints on  $r$  would produce a non-transiting orbit.

posterior distributions obtained under the  $\tilde{\rho}$  basis versus the  $\log T_{14}$  basis arise solely due the parameterization. Furthermore, we conclude that the  $\tilde{\rho}$  basis (with a uniform prior) induces a bias on  $b$ , whereas the other options we have presented produce unbiased estimates.

## 4.5 Mathematical origin of the bias

In the previous sections, we illustrated the biases on  $b$  that result from uniform and log-uniform priors on  $\tilde{\rho}$  by exploring synthetic photometry fits and simple numerical experiments. In this section, we investigate the mathematical origins of this bias.

The transit parameter covariance matrix was previously derived by [Carter et al. \(2008\)](#), but where their treatment prioritized analytic interpretability (with a small sacrifice to accuracy), our treatment prioritizes accuracy (with a small sacrifice to interpretability). Most importantly, the covariance matrix derived by [Carter et al. \(2008\)](#) are least accurate as  $b \rightarrow 1$  and in the presence of non-negligible limb darkening, which are precisely the conditions under which the  $\tilde{\rho}$  bias we are investigating become most important. Thus, our work complements rather than supplants [Carter et al. \(2008\)](#).

When modeling light curves, our main goal is to derive the posterior probability density function,  $p(\vec{x})$ , i.e. the probability that a set of planet properties  $\vec{x}$  resides in an infinitesimal volume element spanning  $\vec{x}$  to  $\vec{x}+d\vec{x}$ . However, this probability is not invariant under changes in parameterization. Specifically, for our problem,  $p(T_{14})/dT_{14} \neq p(\tilde{\rho})/d\tilde{\rho}$ . To convert  $p(T_{14})$  to  $p(\tilde{\rho})$ , one must account for the change in infinitesimal volume element resulting from the  $T_{14} \rightarrow \tilde{\rho}$  transformation, i.e. the Jacobian

$$J = \frac{d\tilde{\rho}}{dT_{14}} = -\frac{12\pi^3}{P^3G} \left( (1+r)^2 - b^2 \right)^{3/2} \left( \frac{\pi T_{14}}{P} \right)^{-4} \quad (4.7)$$

which we derive in the Appendix. The Jacobian of the transformation  $T_{14} \rightarrow \log \tilde{\rho}$  is simply

$$J' = \frac{d \log \tilde{\rho}}{dT_{14}} = -\frac{3}{T_{14}} \quad (4.8)$$

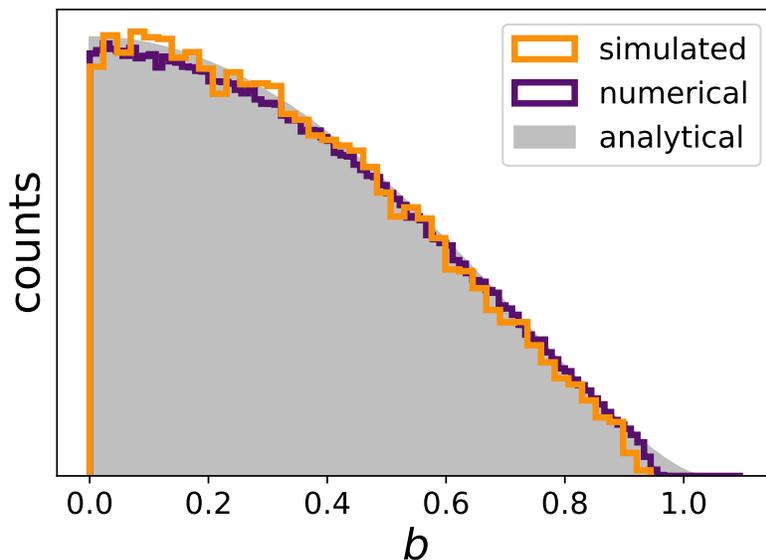


Figure 4.4: Posterior samples of  $b$  from the simulated transit fit (orange histogram, §4.3) and the numerical experiment (purple histogram, §4.4) are nearly perfectly matched by the expected bias from the analytically derived Jacobian (grey shaded region, §4.5).

which is independent of  $b$ , explaining why using  $\log \tilde{\rho}$  in place of  $\tilde{\rho}$  produces unbiased posteriors.

In Figure 4.4, we show the analytic Jacobian in Equation 4.7 alongside the simulated posterior samples of  $b$  obtained in §4.3 and the numerical results obtained in §4.4. It is evident from inspection that the distributions are in close agreement. We conclude that the non-uniform distribution of  $b$  arises from the combination of parameterization and (incorrect) prior, rather than from any real constraint imposed by the data.

## 4.6 Biased Kepler planet properties

We have shown that adopting a linear  $\tilde{\rho}$  prior results in a biased impact parameter. The *Kepler* project (Borucki et al., 2010c; Rowe et al., 2014a, 2015; Mullally et al., 2015; Coughlin et al., 2016; Thompson et al., 2018) used such a parameterization (Jason Rowe, private

communication). Therefore, we expect biased  $b$  in all cases except those where  $b$  is strongly constrained by the light curve itself. Because most *Kepler* planet candidates exhibit modest transit signal-to-noise (median  $S/N = 22.4$ ), the characteristic “hill” shape we have seen for biased posterior  $b$  distributions in the previous three sections is also present in the posterior distributions of nearly every *Kepler* planet candidate from DR25 (Thompson et al., 2018). Figure 4.5 illustrates the presence of the  $b$  bias over a grid of orbital periods and radii. Only the largest (and therefore highest  $S/N$ ) planets consistently exhibit meaningful constraints on  $b$ .

Due to signal-to-noise bias which disfavors the detection of high- $b$  transits (Kipping & Sandford, 2016), the prior expectation on impact parameter is not exactly flat, and so the posteriors exhibited in *Kepler* data will not exactly match the idealized distribution we derived in §4.3-4.5. However, most *Kepler* detections have  $S/N > 10$  and fall in the flat part of the detection completeness curve (Christiansen et al., 2020). Thus, the appropriate prior for the vast majority of *Kepler* planets should be nearly flat in  $b$ , with a fall off at the value of  $b$  that reduces  $S/N$  to  $\sim 10$ .

Detection biases notwithstanding, the  $\tilde{\rho}$  bias is easily understood and corrected. Because the relationship between  $\tilde{\rho}$ ,  $b$ ,  $r$ , and  $T_{14}$  is known analytically (Seager & Mallén-Ornelas, 2003b), one needs only to apply the appropriate Jacobian weighting in order to transform an unintended prior on  $\tilde{\rho}$  into the desired prior on  $b$  or  $T_{14}$  (or any other basis parameter derivable from these quantities). Unbiased parameter estimates can then be recovered from existing (biased) posterior chains by implementing an importance sampling scheme which accounts for this coordinate transformation, provided the chains are not too sparsely sampled in their low probability regions. Specifically, one can sample from a distribution  $p_1(\vec{x})$  by reweighting samples from a different distribution  $p_2(\vec{x})$ . An example of this reweighting scheme as applied to a selection of DR25 targets is shown in Figure 4.6. A caveat is there is increased sampling error since  $p_2(\vec{x})$  is a different distribution and the samples are not

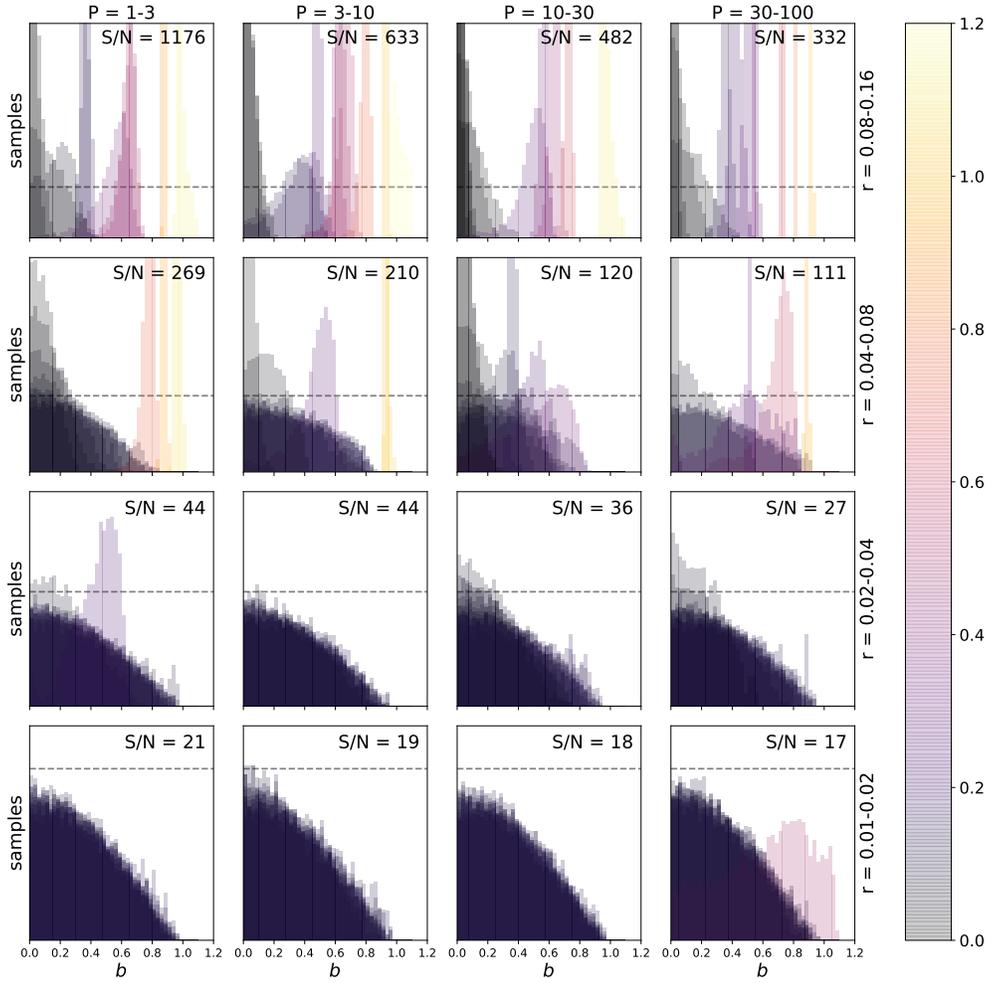


Figure 4.5: Posterior distributions of impact parameter for a random selection of KOIs, organized in logarithmic bins on a  $P - r$  grid. Data shown are the posterior MCMC chains from *Kepler* Data Release 25 (Thompson et al. 2018; Rowe et al. 2014a; Akeson et al. 2013b). Each posterior distribution is plotted with 20% opacity so that dark regions indicate where many distributions overlap; colors correspond to the median  $b$  value for a given KOI, with a maximum of twelve KOIs plotted per panel. The horizontal axis of each panel ranges over  $b = (0, 1.2)$ ; the vertical range of each row is different, but the dashed line indicates the same distribution height. The median S/N in each grid square is printed in the upper right-hand corners. Most of the objects (excluding the highest S/N objects) show qualitatively similar posterior distributions of  $b$ . The similarity is particularly striking for small objects.

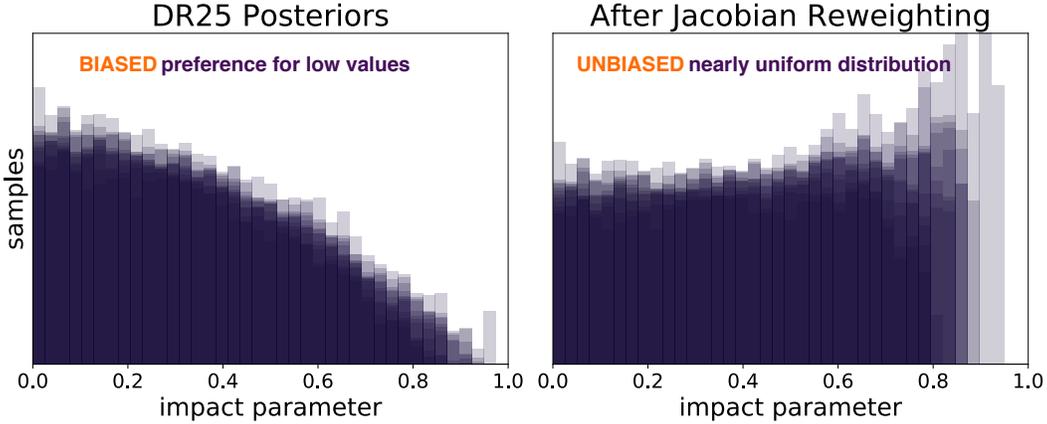


Figure 4.6: Posterior samples of  $b$  from a representative selection of DR25 targets before and after reweighting by the Jacobian to correct for biases induced by sampling in  $\tilde{\rho}$ . All targets have  $0.02 < r < 0.04$  and  $10 < P < 30$  days. *Left panel*: raw DR25 posteriors chains show a clear (biased) preference for low values of  $b$ . *Right panel*: after reweighting, the (unbiased) distribution is nearly flat. To minimize spurious peaks and sampling noise in low probability regions, the lowest density 1% of samples have been excluded from our reweighting scheme. The slight increase in probability density near  $b \approx 1$  in the reweighted posteriors reflects the presence of residual importance sampling noise rather than a real feature of the data. Because there is significant sampling noise (due to the large implied posterior mass in regions with few samples), our preferred method for ameliorating the pseudo-density bias is to refit the photometry.

optimally distributed in  $p_1(\vec{x})$ . In essence there are smaller number of “effective samples” after reweighting. Care must therefore be taken to ensure that Jacobian-corrected posteriors are reliable, and the reweighting scheme we have outlined here should not be applied blindly.

Because  $b$  is covariant with  $r$  (interacting via non-zero limb darkening), any bias on  $b$  translates to a bias on  $r$ . For measurements in the final *Kepler* data release, DR25, we find this covariance has produced a 1.6% median systematic underestimate of  $r$  (Figure 4.7), extending as high as  $\sim 6\%$  for some targets. This offset is comparable to the fractional un-

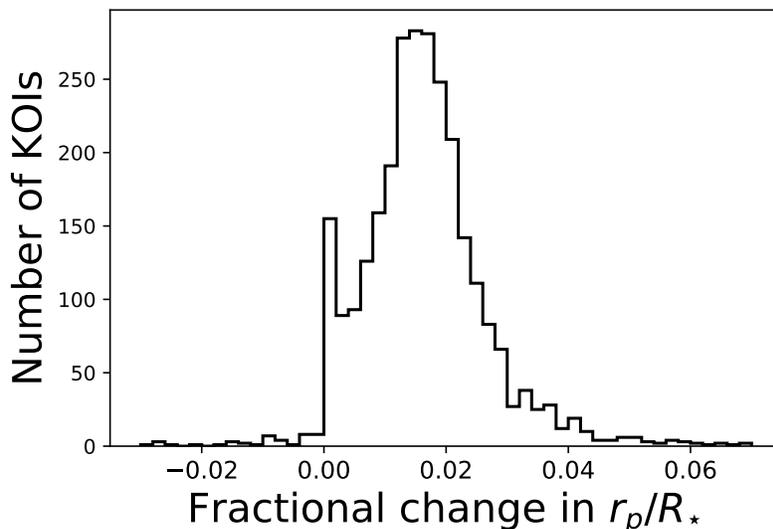


Figure 4.7: Fractional change in median planet-to-star radius ratio for all planet candidates after correcting posterior chains from DR25 (Thompson et al., 2018) using the Jacobian reweighting scheme described in §4.6.  $5\sigma$  outliers have been iteratively clipped in order to eliminate spurious values that are expected to arise due to insufficient sampling of low probability regions. There is a spike at  $\delta_r = 0$ , indicating that some subset of targets were accurately measured, but the majority of targets are distributed around  $\delta_r/r = 1.6\%$ .

certainty on  $R_*$  (Gaia Collaboration et al., 2018; Berger et al., 2018a) and so makes up a sizeable portion of the error budget for *Kepler* planetary radii. While a few percent difference in planetary radius for a *single* planet may be sub-significant, a systematic bias of a few percent on *all* planetary radii will significantly impact our interpretation of population demographics – for example, the precise characteristics of the radius valley (Fulton et al., 2017b) – thereby altering our understanding of the processes by which planets form and evolve.

## 4.7 Summary and conclusions

In this work, we explored the biases that result from using the popular stellar pseudo-density,  $\tilde{\rho}$ , as a parameter in light curve fits. Adopting a linear prior on this parameter results in a biased distribution on impact parameter due to the Jacobian that arises from the non-linear relationship between  $\tilde{\rho}$  and transit duration,  $T_{14}$ . Biased inferences on  $b$  lead to biased inferences on  $r$  due to covariances between the two parameters. We confirmed that these biases are present in *Kepler* modeling that used  $\tilde{\rho}$  as a fitting parameter, and we presented a method for de-biasing the distributions.

Although the  $\tilde{\rho}$  bias may be resolved by using  $\log \tilde{\rho}$  in place of  $\tilde{\rho}$  (or, equivalently, placing log-uniform priors on  $\tilde{\rho}$ ), we prefer sampling in duration over  $\tilde{\rho}$  for aesthetic and conceptual reasons. To avoid inducing biases, we recommend sampling directly in duration  $T_{14}$  or replacing  $T_{14}$  with the true stellar density and orbital eccentricity vector, i.e.  $\{\rho_\star, \sqrt{e} \sin \omega, \sqrt{e} \cos \omega\}$ .

This work focused on the biases induced from using  $\tilde{\rho}$  directly as a fitting parameter; similar biases may arise when using any related parameterization, for example  $a/R_\star$ , which is a popular choice (e.g. Crossfield et al., 2015; David et al., 2016; Stassun et al., 2017). As with  $\tilde{\rho}$ , adopting a log-uniform prior rather than a linear prior on  $a/R_\star$  avoids the unwanted bias. A log-uniform prior is a common choice, so most analyses which have used  $a/R_\star$  as a fitting parameter are probably unaffected by the bias. However, one should always verify what priors were adopted when interpreting the results of any transit model.

## 4.8 Appendix 4A: Derivation of Jacobian for $T_{14} \rightarrow \tilde{\rho}$

In this section, we derive the Jacobian of the coordinate transformation  $T_{14} \rightarrow \tilde{\rho}$ . The pseudo-density derived by Seager & Mallén-Ornelas (2003b) is

$$\tilde{\rho} \equiv \left( \frac{4\pi^2}{P^2 G} \right) \left( \frac{(1+r)^2 - b^2(1 - \sin^2[\pi T/P])}{\sin^2[\pi T/P]} \right)^{3/2} \quad (4.9)$$

where all variables are defined as in previous sections. For notational clarity, we also define  $T \equiv T_{14}$  and make the simplifying assumption  $r \approx \sqrt{\Delta F}$ , where  $\Delta F$  is the fractional change in flux. Substituting terms

$$\begin{aligned} x &= 4\pi^2/(P^2G) \\ y &= (1+r)^2 \\ z &= \sin^2[\pi T/P] \end{aligned} \quad (4.10)$$

yields

$$\tilde{\rho} = x \left( \frac{y - b^2(1-z)}{z} \right)^{3/2}. \quad (4.11)$$

By the chain rule,

$$\frac{d\tilde{\rho}}{dT} = \frac{d\tilde{\rho}}{dz} \frac{dz}{dT}. \quad (4.12)$$

The first term is

$$\frac{d\tilde{\rho}}{dz} = -\frac{3x}{2} \left( \frac{y - b^2}{z^2} \right) \left( \frac{y - b^2(1-z)}{z} \right)^{1/2} \quad (4.13)$$

and the second term is

$$\frac{dz}{dT} = \frac{\pi}{P} \sin \left[ \frac{2\pi T}{P} \right] \quad (4.14)$$

Combining equations 4.10, 4.12, 4.13, and 4.14 yields the exact Jacobian

$$J = \frac{d\tilde{\rho}}{dT} = -\frac{6\pi^3}{P^3G} \left( \frac{(1+r)^2 - b^2}{\sin^4[\pi T/P]} \right) \left( \frac{(1+r)^2 - b^2(1 - \sin^2[\pi T/P])}{\sin^2[\pi T/P]} \right)^{1/2} \sin \left[ \frac{2\pi T}{P} \right] \quad (4.15)$$

Making the small angle approximation  $\sin \phi \approx \phi$  (assuming  $\pi T \ll P$ ) and collecting terms yields

$$J = -\frac{12\pi^3}{P^3G} \left( (1+r)^2 - b^2 \right) \left( (1+r)^2 - b^2(1 - [\pi T/P]^2) \right)^{1/2} \left( \frac{\pi T}{P} \right)^{-4}. \quad (4.16)$$

Once again taking advantage of  $\pi T \ll P$  simplifies the expression further to

$$J = -\frac{12\pi^3}{P^3G} \left( (1+r)^2 - b^2 \right)^{3/2} \left( \frac{\pi T}{P} \right)^{-4}. \quad (4.17)$$

## 4.9 Appendix 4B: Derivation of Jacobian for $T_{14} \rightarrow \ln \tilde{\rho}$

To derive the Jacobian of the transformation  $T \rightarrow \ln \tilde{\rho}$ , we note that

$$\frac{d \ln \tilde{\rho}}{dT} = \frac{1}{\tilde{\rho}} \frac{d\tilde{\rho}}{dT}. \quad (4.18)$$

Adopting our usual approximations  $\sin \phi \approx \phi$ ,  $\pi T \ll P$ , we may rewrite Equation 4.9 in the simplified form

$$\tilde{\rho} \equiv \left( \frac{4\pi^2}{P^2 G} \right) \left( (1+r)^2 - b^2 \right)^{3/2} \left( \frac{\pi T}{P} \right)^{-3}. \quad (4.19)$$

Combining Equations 4.17, 4.18, and 4.19 and cancelling terms yields

$$\frac{d \ln \tilde{\rho}}{dT} = -\frac{3}{T}. \quad (4.20)$$

We see that  $d \ln \tilde{\rho}/dT$  is independent of  $b$ .

## 4.10 Appendix 4C: Derivation of Jacobian for $T_{14} \rightarrow a/R_\star$

To derive the Jacobian of the transformation  $T \rightarrow a/R_\star$ , we define  $\alpha \equiv a/R_\star$  and recognize that from Seager & Mallén-Ornelas (2003b) (their Equations 8 & 9),

$$\alpha = \left( \frac{4\pi^2}{P^2 G} \right)^{-1/3} \tilde{\rho}^{1/3}. \quad (4.21)$$

By the chain rule,

$$\frac{d\alpha}{dT} = \frac{d\alpha}{d\tilde{\rho}} \frac{d\tilde{\rho}}{dT}. \quad (4.22)$$

The first term is

$$\frac{d\alpha}{d\tilde{\rho}} = \frac{1}{3} \left( \frac{4\pi^2}{P^2 G} \right)^{-1/3} \tilde{\rho}^{-2/3}. \quad (4.23)$$

and the second term we derived previously. Adopting our usual approximations  $\sin \phi \approx \phi$ ,  $\pi T \ll P$  and combining Equations 4.17, 4.19, 4.22, and 4.23 yields

$$\frac{d\alpha}{dT} = -\frac{\pi}{P} \left( (1+r)^2 - b^2 \right)^{1/2} \left( \frac{\pi T}{P} \right)^{-2}. \quad (4.24)$$

#### 4.11 Appendix 4D: Derivation of Jacobian for $T_{14} \rightarrow \ln a/R_\star$

To derive the Jacobian of the transformation  $T \rightarrow \ln a/R_\star$ , we note that

$$\frac{d \ln \alpha}{dT} = \frac{1}{\alpha} \frac{d\alpha}{dT} \quad (4.25)$$

where as before  $\alpha \equiv a/R_\star$ . Following our usual strategies and combining Equations 4.21, 4.24, and 4.25, we arrive at

$$\frac{d \ln \alpha}{dT} = -\frac{1}{T} \quad (4.26)$$

which is independent of  $b$ .

## CHAPTER 5

### Accurate and efficient photo-eccentric transit modeling

#### 5.1 Abstract

Kepler- and TESS-quality transit light curves typically allow robust first-order determination of planet sizes (i.e. planet-to-star radius ratio  $r$ ) and orbital ephemerides (i.e. period  $P$  and epoch  $t_0$ ). Other orbital parameters (i.e. impact parameter  $b$ , eccentricity  $e$ , and argument of periastron  $\omega$ ) are more challenging to measure because they imprint themselves on the transit lightcurve through second-order effects on the transit duration  $T_{14}$ . In [Gilbert et al. 2022](#), we investigated using a five parameter model basis  $\{P, t_0, r, b, T_{14}\}$  to recover unbiased measurements of  $r$  and  $b$ . Here, we expand upon our previous work and use post-hoc importance sampling to recover unbiased measurements of  $e$  and  $\omega$ . We perform a suite of injection-and-recovery test to demonstrate that eccentricities derived using our method are consistent with those derived from a standard model basis  $\{P, t_0, r, b, \rho_*, e, \omega\}$  which explicitly fits eccentricity vectors. The advantage of our five-parameter photo-eccentric transit model is that it is both simpler than the seven-parameter model and is “future-proof”: our posterior measurements can be quickly resampled to accommodate updated priors as more precise stellar characterization becomes available without requiring an expensive re-run of the transit fitting routine. We conclude by offering suggestions for accurate measurement of  $b$ ,  $e$ , and  $\omega$  from transit photometry.

## 5.2 Introduction

Out of more than 5,300 confirmed planets to date,  $\sim 75\%$  were discovered via the transit method. These discoveries have paved the way for keystone scientific advancements in our understanding of planetary formation, evolution, and demographics. To ensure the reliability of these discoveries, we must also ensure that the individual transiting planet characterizations are accurately and consistently derived. Previously, uncertainties on stellar parameters significantly limited the achievable precision of planet properties (i.e.  $\sigma(R_*) \sim 27\%$  and  $\sigma(\rho_*) \sim 51\%$ ; Thompson et al. 2018). Now, in the era of *Gaia* (Gaia Collaboration et al. 2018) and high-precision stellar characterizations (i.e.  $\sigma(R_*) \lesssim 2\%$  and  $\sigma(\rho_*) \lesssim 10\%$ ), the determination of key planet properties is limited by light curve modeling.

A variety of methods exist for fitting transit signals, including various parameterizations (e.g. Seager & Mallén-Ornelas 2003b; Dawson & Johnson 2012; Eastman et al. 2013a; Thompson et al. 2018; Gilbert et al. 2022) and sampling techniques (e.g. Feroz & Hobson 2008; Foreman-Mackey et al. 2013; Foreman-Mackey et al. 2021b; Speagle 2020; Gilbert 2022). Substantial effort has been put into vetting such methods for transit signal detection (see, e.g., Christiansen et al. 2015), but less effort has been devoted to validating and comparing the final characterizations achieved by these models. Differences between sampling techniques are typically assumed to be insignificant relative to parameter uncertainties, but differences in model parameterizations can have a substantial impact on the subsequent planet characterization (see, e.g., Gilbert et al. 2022; *G22* hereafter). In particular, the implementation of a given transit model must be done with careful consideration of the covariances that exist between transit parameters, including planet-to-star radius ratio versus impact parameter ( $r-b$ ), impact parameter versus eccentricity ( $b-e$ ), and eccentricity versus argument of periastron ( $e-\omega$ ). Here, we show that our implementation of a duration-based transit model parameterization can model these covariances with equivalent accuracy as an eccentricity-explicit model.

Typically, transit model parameterizations will include orbital period  $P$ , transit mid-point  $t_0$ , planet-to-star radius ratio  $r$ , impact parameter  $b$ , and some additional parameter(s) to describe the transit duration. Such models also tend to include quadratic limb darkening coefficients  $\{u_1, u_2\}$  and additional parameters relating to the zero-point photometric flux  $\mu_{\text{flux}}$  and photometric noise  $\sigma_{\text{flux}}$ .

A popular transit modeling basis characterizes a transit using 7 primary parameters, with transit duration described via three parameters: eccentricity  $e$ , argument of periastron  $\omega$ , and stellar density  $\rho_\star$  (or, alternatively, scaled orbital separation  $a/R_\star$  via Kepler’s third law). This eccentricity-explicit parameter basis ( $e\text{-}\omega\text{-}\rho$  hereafter) benefits from being fully characterized by properties of the star, planet, and planetary orbit. However, transit photometry typically only has enough information from observable properties to constrain  $\sim 4\text{--}5$  parameters well. Additionally, the  $e\text{-}\omega\text{-}\rho$  basis is computationally expensive to run because the sampler must trace a thin, curving degeneracy in  $\{e, \omega\}$  space and also solve Kepler’s equation at each sampler step.

This basis also relies on an input constraint on  $\rho_\star$ , preventing model outputs from being “future-proof” and necessitating an expensive re-run of the model when the stellar characterization is inevitably updated at a later time (e.g. from a new *Gaia* data release). An alternative approach that improves upon these limitations requires fitting the lightcurve assuming a circular orbit,  $e = 0$ , regardless of what the true underlying eccentricity might be (see, e.g., Seager & Mallén-Ornelas 2003b; Dawson & Johnson 2012). This shortcut approach reduces the total number of model parameters by two, relative to the  $e\text{-}\omega\text{-}\rho$  parameterization, and allows for  $e$  and  $\omega$  to instead be derived post-modeling via a non-physical pseudo-density parameter  $\tilde{\rho}$ . As *G22* points out, however, this alternative parameterization can introduce a bias on  $b$  that artificially disfavors  $b \gtrsim 0.7$ . This bias propagates through to other model parameters, shifting  $e$  towards higher values and underestimating  $r$ .

*G22* offers a more intuitive alternate parameterization that achieves similar improvements

over the  $e\text{-}\omega\text{-}\rho$  basis but without imposing an unintended bias on  $b$ . This proposed parameterization is motivated by observable transit parameters, which includes the transit duration  $T_{14}$  – defined from 1<sup>st</sup>-to-4<sup>th</sup> contact:  $\{P, t_0, r, b, T_{14}\}$ . *G22* demonstrated that the duration-based parameterization achieves improved  $b$  and  $r$  constraints relative to the biased pseudo-density parameter basis. Here, we provide a recipe for extracting  $\{e, \omega\}$  from this alternative basis quickly and accurately, relative to the  $e\text{-}\omega\text{-}\rho$  basis. We combine the duration-based parameterization with umbrella sampling (Gilbert, 2022) to ensure full exploration of the  $b$  posterior space. We then perform post-hoc importance sampling to derive the posterior distributions of  $e$  and  $\omega$ , without needing to include them as explicit model parameters. These improvements resolve various limitations of both the  $e\text{-}\omega\text{-}\rho$   $\{P, t_0, r, b, e, \omega, \rho_\star\}$  and pseudo-density  $\{P, t_0, r, b, \tilde{\rho}\}$  parameterizations.

In this work, we use simulated photometry to perform injection-and-recovery tests over a range of injected transit properties to compare the performances of our proposed modeling approach and the  $e\text{-}\omega\text{-}\rho$  modeling basis for modeling  $e$  and  $\omega$ . We lay out our methodology for lightcurve synthesis and transit injection-and-recovery in §5.3. We then highlight the procedural differences between the  $e\text{-}\omega\text{-}\rho$  method (§5.4) and our  $T_{14}+umb$  modeling approach (§5.5). In §5.6, we analyze the results of our injection-and-recovery tests and compare the performances of the two parameterizations. We provide a summary of our conclusions in §5.7.

### 5.3 Synthetic Lightcurve Construction

Our objective is to compare the performance of the physical  $e\text{-}\omega\text{-}\rho$  parameter basis to the simpler  $T_{14}$  basis. We aim to demonstrate whether or not these methods return equivalent and accurate posterior results and determine their relative efficiencies. To achieve these objectives, we perform a suite of injection-and-recovery tests over a grid of parameters which spans a wide range of values of eccentricity  $e$ , argument of periastron  $\omega$ , inclination (param-

eterized as impact parameter  $b$ ), and signal-to-noise (see Figure 5.1). Injection-and-recovery is a standard tool used to evaluate transit signal detection methods (see, e.g., Christiansen et al. 2015), but it has not been applied to transit model validation on nearly the same scale. Here, we construct a set of synthetic lightcurves, then we proceed to use two distinct transit modeling methods to recover the injected transit properties and compare the relative model performances.

For all injection-recovery tests in this work, we inject the transit signal of a sub-Neptune-size planet orbiting a Sun-like star with an orbital period close to the average among *Kepler* planets. We synthesize 10 transits per lightcurve with a photometric zero-point flux of  $\mu_{\text{flux}} = 0$  and a fixed photometric noise  $\sigma_{\text{flux}}$ , consistent with raw photometry that has been accurately prewhitened. We calculate the duration of each injected transit signal  $T_{14}$  according to the following equation from Winn (2010a):

$$T_{14} = \frac{P}{\pi} \sin^{-1} \left( \frac{\sqrt{\frac{(1+r)^2 - b^2}{\left(\frac{GP^2\rho_\star}{3\pi}\right)^{2/3} - b^2}}}{1 + e \sin \omega} \right) \frac{\sqrt{1 - e^2}}{1 + e \sin \omega}. \quad (5.1)$$

We construct our synthetic lightcurves at three different signal-to-noise ratio (SNR) levels:  $\text{SNR} \sim [20, 40, 80]$ . We show example lightcurves for each SNR level in Figure 5.1. At an SNR of 20, the injected signal has a slightly lower significance compared to the median *Kepler* planet signal. At the higher SNR levels of 40 and 80, we seek to identify any differences that emerge between our two models as the transit ingress and egress become more distinct from the photometric noise, making  $b$  measurements more precise. From the selected SNR and other injected lightcurve properties, we generate Gaussian white noise per lightcurve centered on  $\sigma_{\text{flux}}$ , which we calculate according to:

$$\sigma_{\text{flux}} = \sqrt{\frac{T_{14,\text{true}} N_{\text{transits}}}{t_{\text{exp}}} \frac{r_{\text{true}}^2}{\text{SNR}}}, \quad (5.2)$$

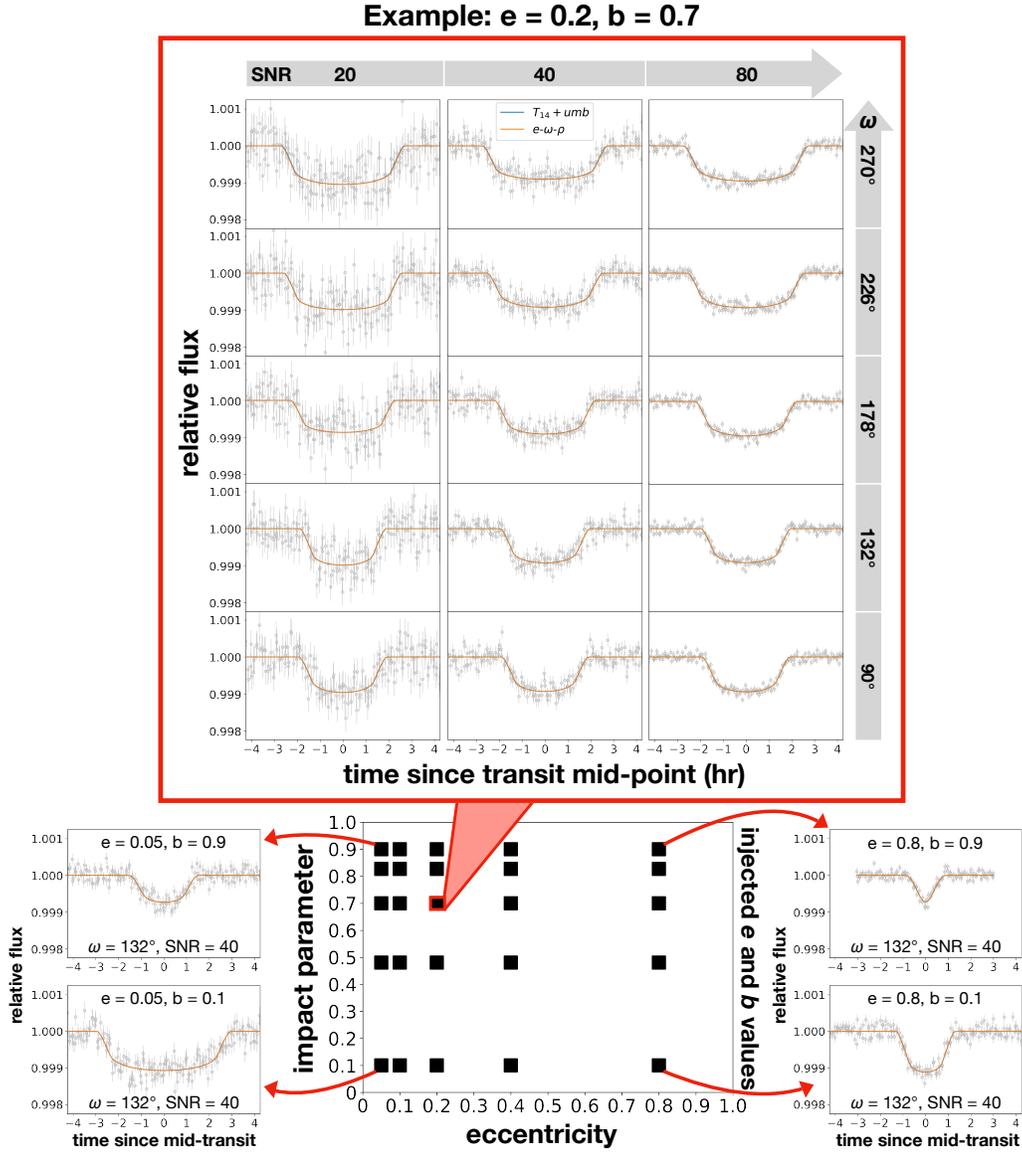


Figure 5.1: Diagram showing (bottom middle) our grid of injected  $e$  and  $b$  values, along with (top) a gallery of phase-folded lightcurves for all combination of SNR and injected  $\omega$  values (at  $e = 0.2$  and  $b = 0.7$ , as an example). The four panels to the left and right of the bottom grid show demonstrative examples of phase-folded lightcurves at the different extremes of our injected flux parameter grid (e.g.  $b = 0.1$  or  $0.9$  and  $e = 0.05$  or  $0.8$ ), all shown with  $\omega = 132^\circ$  and SNR = 40. Each lightcurve is shown with the median final transit fits from both the  $e-\omega-\rho$  (orange) and  $T_{14} + umb$  (blue) modeling methods, which overlap almost entirely.

where  $N_{\text{transits}}$  is the number of injected transits and  $t_{\text{exp}}$  is the simulated exposure time of our synthetic lightcurve. The random seed used to generate the synthetic white noise is unique to each injection-recovery test.

We also assign a unique set of transit parameter values  $\{b, e, \omega\}$  for each injection-recovery test, where each of these inputs is drawn from a grid of discrete parameter values (see Figure 5.1). We specifically choose a parameter grid that emphasizes the region of parameter space where the  $e - \omega - b$  degeneracy is strongest (see, e.g., [Van Eylen & Albrecht 2015](#)) since this is where the two parameterizations are more likely to yield differing results. As a result, our injected planet signals do not exactly mirror the distribution of *Kepler* planets, but they do include a broad range of realistic planet characteristics.

Since the transit shape is more sensitive to small changes in  $b$  at high values, we select injected values of  $b$  with tighter spacing towards higher values, spanning the non-grazing parameter space. We construct an array of  $b$  values that are evenly spaced on a reversed log scale:  $b \sim [0.1, 0.48, 0.7, 0.83, 0.9]$ . We also prefer to use  $e$  values that span the range of eccentricities with tighter spacing towards low-to-moderate values, since these are more common. We select an array of possible  $e$  values which are evenly spaced on a log scale:  $e \sim [0.05, 0.1, 0.2, 0.4, 0.8]$ . Additionally, the  $\omega$  values that we draw upon for our grid of injected parameters are intentionally selected to include the inflection points of periastron ( $\pi/2$  or  $90^\circ$ ) and apastron ( $3\pi/2$  or  $270^\circ$ ) along with three roughly evenly spaced values in between:  $\omega \sim [90^\circ, 132^\circ, 178^\circ, 226^\circ, 270^\circ]$ .

We construct a set of 375 unique transit lightcurves from all combinations of  $\{b, e, \omega, \text{SNR}\}$  using the `batman` transit modeling package ([Kreidberg 2015](#)). We synthesize these injected lightcurve models with an oversampling rate of 11 and  $t_{\text{exp}} = 30$  minutes, similar to real *Kepler* photometry. These lightcurves serve as inputs to the two modeling methods that we are comparing, described below, in order to demonstrate similarities and differences in model performance across a range of potential transit signals (see Figure 5.2 for an overview).

Table 5.1: Transit Model Parameters and Priors

Parameter	Input Value(s)	Prior
$P$ (d)	26.1	fixed
$t_0$ (d)	1.0	$t_0 \sim \text{N}(1.0, 0.1)$
$r$	0.03	$\log r \sim \text{U}(-9, 0)$
$b$	[0.1, 0.48, 0.7, 0.83, 0.9]	$b \sim \text{U}(0, 1 + r)$
$u_1, u_2$	{0.4, 0.25}	fixed
$\mu_{\text{flux}}$	0	fixed
$\sigma_{\text{flux}}$	derived	fixed per lightcurve
$T_{14}$ (d)	derived	$\log T_{14} \sim \text{U}(-9, 0)$
$e$	[0.05, 0.1, 0.2, 0.4, 0.8]	$e \sim \text{U}(0, 0.92)$
$\omega$ ( $^\circ$ )	[90, 132, 178, 226, 270]	$\omega \sim \text{U}(-90, 270)$
$\rho_*$ ( $g/cm^3$ )	1.41 (e.g. $\rho_\odot$ )	$\rho_* \sim \text{N}(1.41, 0.141)$

Note. — All parameters used in the models discussed throughout this analysis, along with their units, input values, and associated priors (if applicable). Priors include normal (N) and uniform (U) distributions.  $b$ ,  $e$ , and  $\omega$  each have five input value options that form a grid of possible injected transit signal properties. We note that the priors on  $e$  and  $\omega$  can also be represented via the transform  $\{e, \omega\} \rightarrow \{\sqrt{e} \sin \omega, \sqrt{e} \cos \omega\} \sim \text{UniformDisk}(\sqrt{0.92})$ .

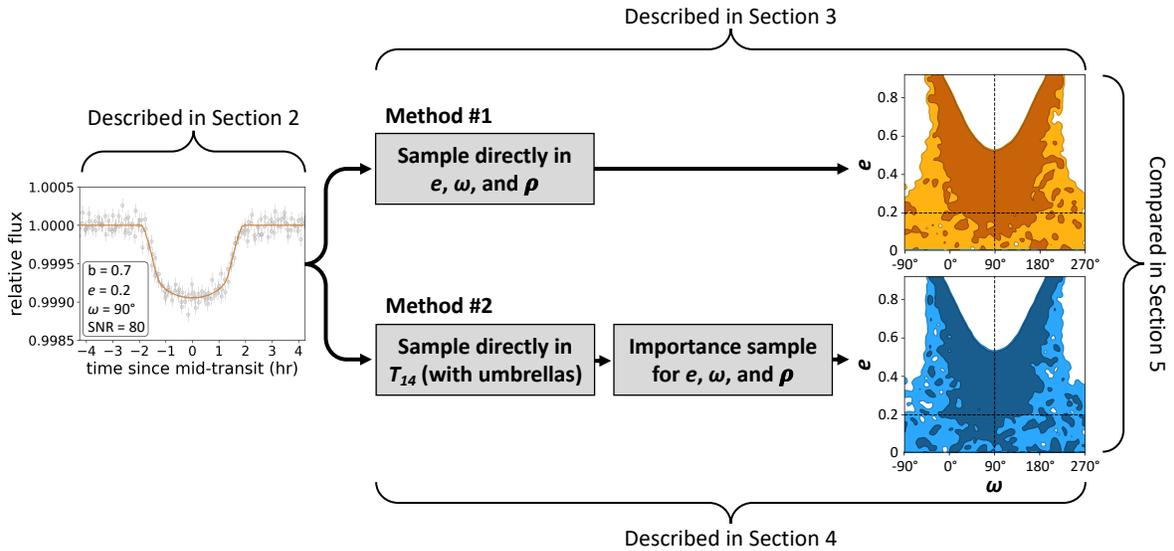


Figure 5.2: Diagram demonstrating an overview of our modeling procedure, from an input synthetic lightcurve to output  $e$  and  $\omega$  constraints via both the  $e$ - $\omega$ - $\rho$  (orange) and  $T_{14} + umb$  (blue) modeling methods.

## 5.4 Method #1: Direct sampling in $e$ - $\omega$ - $\rho$

We first model our synthetic transit lightcurves using the  $e$ - $\omega$ - $\rho$  model, which serves as our baseline model and standard reference when evaluating the performance of our proposed  $T_{14} + umb$  model. This physically-motivated transit model is parameterized by  $\{P, t_0, r, b, \rho_\star, e, \omega\}$ , along with quadratic limb darkening parameters  $\{u_1, u_2\}$ . Since we simulate lightcurves with white noise, we fix  $\mu_{\text{flux}}$  and  $\sigma_{\text{flux}}$  which would otherwise be directly sampled parameters when modeling real transit photometry.

We construct the  $e$ - $\omega$ - $\rho$  model using uninformative priors that are of standard use in transit fitting literature or drawn directly from *G22*, summarized in Table 5.1. We apply a normal prior on  $\rho_\star$  which assumes that the stellar density is known with 10% uncertainty through independent measurements. To mitigate boundary issues that can occur when sampling  $e$  and  $\omega$  directly, we use a common redefinition of these parameters  $\{\sqrt{e} \sin \omega, \sqrt{e} \cos \omega\}$  (see,

e.g. [Eastman et al. 2013a](#)), with implicit uniform priors on both  $e$  and  $\omega$ . These priors do not account for transit probability or other astrophysically motivated considerations (see [Barnes 2007a](#)).

We implement this model using `exoplanet` ([Foreman-Mackey et al., 2021b](#)), with sampling performed by the NUTS algorithm via `PyMC3` ([Salvatier et al., 2016b](#)). We use 30,000 tuning steps with an additional 40,000 sampler draws to ensure that the sampler converges with an effective sample size  $N_{\text{eff}} \approx 10^3$ . We also set a high target acceptance fraction of 0.99 to encourage the sampler to adequately explore complex topologies in the posterior parameter space, such as the  $b - r$  and  $e - \omega$  degeneracies. This sampling process is performed via two sampler chains across two CPU cores.

From initial experimentation, we found that sampler limitations exist which restrict the valid parameter space of eccentricity modeling when applying the  $e - \omega - \rho$  parameterization via NUTS sampling with `exoplanet`. When sampling  $e \gtrsim 0.92$ , this implementation of the  $e - \omega - \rho$  model can have convergence issues due to the high curvature of the posterior parameter space being traversed. This also roughly corresponds with the upper eccentricity limit where we expect transit duration approximations to begin breaking down (see, e.g. [Kipping 2014a](#)). Given that only 5 known planets have  $e > 0.9$  and only one of these was discovered via transit modeling, we choose to restrict our eccentricity sampling to  $e < 0.92$  for all modeling approaches considered in this work. By doing so, we avoid conflating our primary interest – differences in modeling methods – with rare edge cases that are beyond the scope of this work.

## 5.5 Method #2: Direct sampling in $T_{14}$ , then importance sampling in $e - \omega - \rho$

Our alternative transit modeling approach, the  $T_{14} + umb$  model, has a parameter basis that includes the observable transit duration  $T_{14}$  as an explicit parameter. This parameterization

avoids explicitly sampling the complex degeneracies introduced by  $e$  and  $\omega$ , allowing us to instead measure these parameters post-hoc via importance sampling (see 5.5.2). We couple this duration-based parameterization with umbrella sampling (see Gilbert 2022) to ensure that our model accurately samples the complicated topology of the high- $b$  “grazing” parameter space. Based on the arguments made in both *G22* and Gilbert (2022), we expect that our  $T_{14} + umb$  approach should achieve results that are consistent with those from the  $e-\omega-\rho$  model with a potential boost in efficiency.

### 5.5.1 Transit fitting

Similar to our implementation of the baseline  $e-\omega-\rho$  model, we also construct our  $T_{14} + umb$  model via `exoplanet` with NUTS sampling and use it to model our synthetic transit signals. This parameterization is motivated by observable transit properties and characterized by the basis  $\{P, t_0, r, b, T_{14}\}$ . Like the  $e-\omega-\rho$  model, the  $T_{14} + umb$  model also includes quadratic limb darkening parameters  $\{u_1, u_2\}$  as well as fixed values of  $\mu_{\text{flux}}$  and  $\sigma_{\text{flux}}$ . The priors used here are identical to those used in our  $e-\omega-\rho$  model, summarized in Table 5.1. Neither  $e$  nor  $\omega$  is explicitly constrained during the sampling process here, and their values are instead estimated from post-model importance sampling. This parameterization is thus agnostic to orbital eccentricity, except for the implicit assumption of a symmetric transit. This is a reasonable approximation since the acceleration of an eccentric planet during its transit is unlikely to introduce detectable asymmetry given modern photometry (Barnes, 2007a).

To improve both the sampling convergence and the exploration of complex posterior topologies, we follow Gilbert (2022) to implement umbrella sampling. We separate our NUTS sampler into three windows (i.e. “umbrellas”) defined within the joint  $\{r, b\}$  parameter space, which allows us to sample the full posterior parameter space in smaller pieces that are easier to explore. The resulting posteriors can later be stitched together by applying the appropriate umbrella weights. The three umbrella windows that we use correspond to non-grazing and grazing orbits separated by a region that we refer to as the transition umbrella, which

partially overlaps with the other two (see [Gilbert 2022](#) for full description). In our implementation, we apply the three umbrella models in series but emphasize that this task can easily be parallelized to reduce the apparent wall-clock run-time. In the Appendix, we also discuss a potential alternative to umbrella sampling, known as dynamic nested sampling (see, e.g., [Skilling 2004](#); [Skilling 2006](#)), which achieves roughly comparable results.

### 5.5.2 Importance sampling

To recover  $\{e, \omega\}$  samples from the  $T_{14} + umb$  modeling approach, we apply post-hoc importance sampling to the combined umbrella model posterior distributions. Importance sampling (see, e.g., [Oh & Berger 1993](#); [Gilks et al. 1995](#); [Madras & Piccioni 1999](#)) allows one to measure the properties of a given parameter’s probability distribution based on samples generated from a different (typically easier to sample) parameter’s distribution. This method was first incorporated into exoplanet characterization models by [Ford \(2005\)](#) and [Ford \(2006\)](#), used in combination with MCMC sampling to improve radial velocity model efficiency. Such methods can be useful to correct for observational biases post-hoc or derive the distributions of more complicated distributions outside of the MCMC sampling routine. Importance sampling is closely related to umbrella sampling, and the former can be thought of as a single-window special case of the latter. In our implementation, importance sampling only marginally increase the total run-time of the  $T_{14} + umb$  approach by a few seconds.

We first compute the relative weights of the three umbrella models following [Gilbert \(2022\)](#) and combine our posterior chains into a single set of weighted posterior distributions. Since the umbrella weights effectively reduce the total number of samples, we up-sample the merged posterior distributions via random resampling to generate a total of  $10^5$  samples per parameter for convenience. We then perform importance sampling to weigh how well the measured values of  $\{P, r, b, T_{14}\}$  at each sampler step can be described by an independently measured density of the host star. We will refer to this independent stellar density as  $\rho_{\star, \text{true}}$ , with some uncertainty  $\sigma_{\rho_{\star, \text{true}}}$ . To determine the appropriate importance weights, we first cal-

culate the sampler-derived stellar density,  $\rho_{\star,\text{samp}}$ , at each point in the umbrella-weighted posterior. This calculation directly follows from the transit duration equation described by Winn (2010a):

$$\rho_{\star,\text{samp}} = \frac{3\pi}{GP^2} \left( \frac{(1+r)^2 - b^2}{\sin^2 \left( \frac{T_{14}\pi}{P} \frac{1+e\sin\omega}{\sqrt{1-e^2}} \right)} + b^2 \right)^{3/2}. \quad (5.3)$$

We note that Equation 5.3 explicitly includes  $e$  and  $\omega$ , for which we do not yet have any information. We substitute these parameters with random draws of  $\{e, \omega\}$  from uniform priors  $e \sim U(0, 0.92)$  and  $\omega \sim U(-\frac{\pi}{2}, \frac{3\pi}{2})$ . By deriving  $\rho_{\star,\text{samp}}$  from measured values of  $\{P, r, b, T_{14}\}$  and random uniform values of  $\{e, \omega\}$ , we ensure that  $\rho_{\star,\text{samp}}$  reflects a true stellar density as opposed to the pseudo-stellar density parameterization which assumes  $e = 0$  and was deemed unreliable by *G22*.

We compare the samples of  $\rho_{\star,\text{samp}}$  against the independently measured  $\rho_{\star,\text{true}}$  by computing the log-likelihood of each  $i^{\text{th}}$  sample,

$$\log \mathcal{L}_i = -\frac{1}{2} \left( \frac{\rho_{\star,\text{samp},i} - \rho_{\star,\text{true}}}{\sigma_{\rho_{\star,\text{true}}}} \right)^2, \quad (5.4)$$

assuming a Gaussian likelihood function. We then weight each sample from our umbrella-weighted posterior distributions by

$$w_i = \frac{\mathcal{L}_i}{\sum_i \mathcal{L}_i} \quad (5.5)$$

to produce the final, importance-weighted posterior distributions for each parameter. We apply these same weights to the random uniform  $\{e, \omega\}$  samples to derive the final posterior distributions of these two parameters. All analysis in this work regarding the  $T_{14} + \text{umb}$  model is based on these posterior distributions that have been umbrella-weighted, up-sampled, and importance-weighted. The final posterior distribution of  $e$  that we measure using our  $T_{14} + \text{umb}$  modeling approach can thus be directly compared to the  $e$  posterior from the  $e-\omega-\rho$  model.

We therefore can use the  $T_{14}$  basis  $\{P, t_0, b, r, T_{14}\}$  along with an independently constrained  $\rho_{\text{true}}$  to derive posterior distributions for all parameters represented by the  $e-\omega-\rho$  basis  $\{P, t_0, b, r, e, \omega, \rho_\star\}$ . With the  $T_{14}$  basis, we have the advantage of avoiding introducing significant stellar constraints (i.e.  $\rho_\star$ ) until *after* the transit has already been fully modeled. Thus, our  $T_{14}+umb$  model only needs to be run once while the  $e-\omega-\rho$  model would have to be re-run for each updated measurement of stellar density. The post-hoc importance sampling step can easily be re-run for an updated  $\rho_{\star, \text{true}}$  value (or different priors on  $e$  or  $\omega$ ) within only a few seconds, making our  $T_{14} + umb$  modeling approach essentially future-proof. In the era of *Gaia* and high-precision stellar characterization, such future-proofing will become increasingly valuable.

## 5.6 Results

### 5.6.1 Both methods return equivalent eccentricity constraints

We fit 375 injected transit signals from our grid of injection-recovery tests using both the  $e-\omega-\rho$  baseline model and our  $T_{14}+umb$  modeling approach. We measure all transit parameters using both modeling approaches, including  $e$  and  $\omega$ . The posterior distributions of  $e$ ,  $\omega$ , and  $b$  serve as our primary points of comparison between the baseline model and our alternative modeling approach. Here, we specifically focus our analysis on  $e$ , since  $b$  (and its relationship with  $r$ ) was already covered in *G22* and  $\omega$  is often a nuisance parameter in photometric modeling. We use posterior comparisons of  $\omega$  and  $b$  for secondary analysis when necessary.

We perform a quantile-quantile comparison of the posterior values  $e_k$  at the  $k = 15^{\text{th}}$ ,  $50^{\text{th}}$ , and  $85^{\text{th}}$  percentiles of the  $e_{e\omega\rho}$  and  $e_{T_{14}+umb}$  eccentricity distributions. In Figure 5.3, we present a comparison of  $e_k$  from both modeling methods at each of the key percentiles for all injection-recovery tests. We see that all tests at each percentile are close to the 1-to-1 line (black), demonstrating that the two modeling methods produce nearly equivalent posterior results for  $e$ .

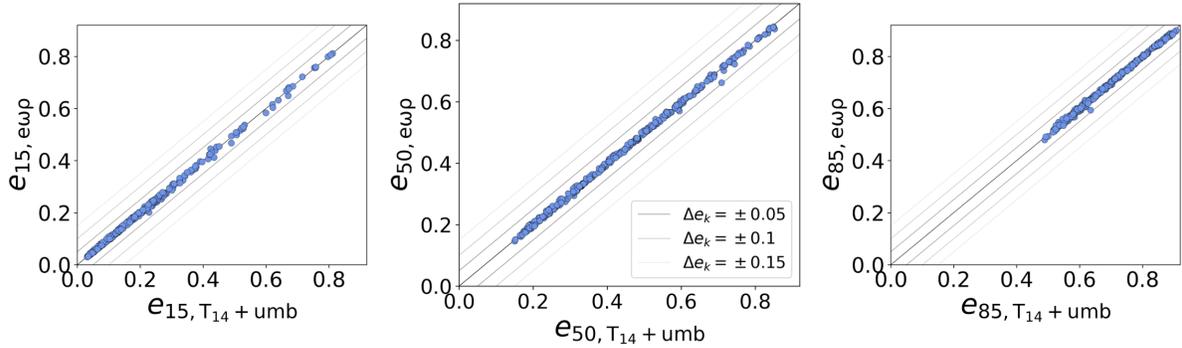


Figure 5.3: Comparison of  $e$  values measured from the  $T_{14} + umb$  and  $e-\omega-\rho$  modeling methods at the 15<sup>th</sup>, 50<sup>th</sup>, and 85<sup>th</sup> percentiles of their distributions. We show  $\Delta e_k = \{0.05, 0.1, 0.15\}$  in grey, as well as the ideal 1-to-1 line shown in black. These comparisons generally lie close to the 1-to-1 line, implying that the results of the two models are approximately equivalent.

We compute the difference  $\Delta e_k$  (e.g.  $\Delta e_{50} = e_{50,umb} - e_{50,e\omega\rho}$ ) and use this as a measure of similarity between the two model results. To estimate the significance of  $\Delta e_k$  for each posterior comparison, we assume a standard eccentricity uncertainty of  $\sigma_e = 0.05$ , informed by the typical uncertainty on  $e$  measured among all known planets ( $\sigma_{\text{median}}(e) \approx 0.05$ ; [NASA Exoplanet Science Institute 2020<sup>1</sup>](#)). For injection-recovery tests where  $|\Delta e_k| \lesssim 0.05$  at the 15<sup>th</sup>, 50<sup>th</sup>, and 85<sup>th</sup> percentiles of eccentricity, we assert that the  $e-\omega-\rho$  and  $T_{14}+umb$  methods produce equivalent results. Among multiple iterations of our suite of injection-recovery tests, we did not identify any tests which consistently produced posterior measurements for  $e$  that differed by  $|\Delta e_k| \lesssim 0.05$  (see Figure 5.3). This suggests that our approach is an excellent alternative to the  $e-\omega-\rho$  method, since the two methods should converge on identical results (as opposed to  $\sim 68\%$  identical).

We also consider how  $\Delta e_k$  differs as a function of both the lightcurve SNR and the injected transit duration  $T_{14}$ . Specifically, we consider the ratio between  $T_{14}$  and the expected du-

---

<sup>1</sup>NASA Exoplanet Archive data retrieved on 2023 February 23

ration of the same planet on a circular, centrally transiting orbit (the reference duration,  $T_{14,\text{ref}}$ ):  $T_{14}/T_{14,\text{ref}}$ . This duration ratio is a more concise metric to interpret the effects of  $e$ ,  $\omega$ , and  $b$  on the duration of a transit. While we observe no trend in  $\Delta e_k$  with respect to SNR, we do note a marginal trend in  $\Delta e_k$  as a function of  $T_{14}/T_{14,\text{ref}}$  across our sample. We find that the  $T_{14} + umb$  model estimates slightly higher  $e$  values than the  $e-\omega-\rho$  model at short transit durations and vice-versa at long transit durations, but the deviations that contribute to this trend are sub-significant. We ultimately conclude that the two modeling methods produce equivalent eccentricity measurements (within a reasonable uncertainty) for virtually all tenable combinations of  $\{e, \omega, b, \text{SNR}\}$ .

### 5.6.2 Both methods return accurate results

We have demonstrated that our alternative transit modeling approach produces equivalently accurate results relative to our baseline model, but we have not yet considered if these models yield the correct results (relative to the injected parameters). It is known in the field of exoplanet characterization that photometric eccentricity constraints (and  $\omega$  constraints) tend to have large uncertainties for individual planets (see, e.g. [Van Eylen et al. 2019](#)). Here, we qualitatively assess these uncertainties across our set of injection-recovery tests.

Since our sample is not representative of the observed planet population, we describe the observed trends among our  $e$  measurements according to different quadrants of  $e - b$  parameter space. We split up our tests into four broad scenarios based on their injected transit properties: (1) low  $e$  and low  $b$ , (2) low  $e$  and high  $b$ , (3) high  $e$  and low  $b$ , (4) high  $e$  and high  $b$ . We show demonstrative examples of these four scenarios in [Figure 5.4](#) with several  $\omega$  values, all at  $\text{SNR} = 20$ . In all four quadrants, the posterior distributions of  $e$  and  $b$  are broad, non-Gaussian, and display a range of outcomes, but we describe the general trends that we observe below. We also offer some additional discussion regarding how  $\omega$  can affect these posterior constraints. We limit our discussion to only the posterior distributions of the  $T_{14} + umb$  modeling approach since the two approaches produce nearly equivalent results.

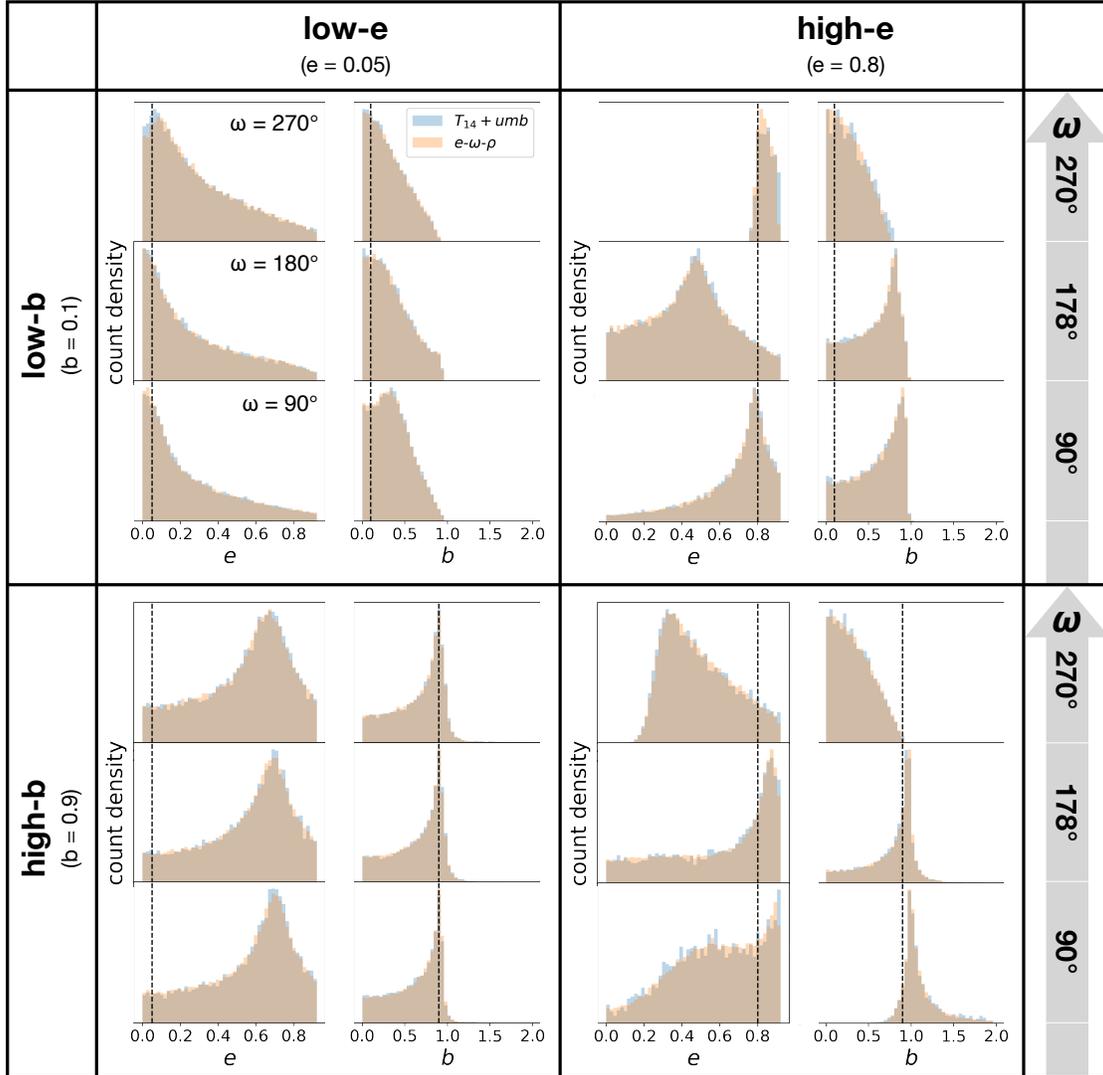


Figure 5.4: Qualitative diagram showing the relative accuracy of measuring  $e$  and  $b$  from transit photometry in four distinct quadrants of  $e - b$  parameter space, at three different  $\omega$  values. The four scenarios shown are (Top Left) low- $e$  and low- $b$ , (Bottom Left) low- $e$  and high- $b$ , (Top Right) high- $e$  and low- $b$ , and (Bottom Right) high- $e$  and high- $b$ . Across all areas of  $e - b$  parameter space, the  $T_{14} + umb$  (blue) and  $e - \omega - \rho$  (orange) modeling methods perform equivalently.

In scenario 1 (low  $e$  and low  $b$ ), transit models accurately measure low values for both  $e$  and  $b$  with little posterior mass at higher values (Figure 5.4, top left), regardless of  $\omega$ . In scenario 2 (low  $e$  and high  $b$ ), models tends to significantly overestimate  $e$  but produce more accurate measurements of  $b$  (Figure 5.4, bottom left), regardless of  $\omega$ . The opposite is true in scenario 3 (high  $e$  and low  $b$ ) where  $b$  tends to be overestimated while  $e$  is measured more accurately (Figure 5.4, top right), except near apastron where both are measured fairly accurately. In scenario 4 (high  $e$  and high  $b$ ), transit models tend to accurately measure high values for both parameters with little posterior mass at lower values (Figure 5.4, bottom right), except near apastron where neither is measured well. We avoid providing a quantitative description of these observed trends because the non-Gaussian posterior distributions are not well-represented by simple summary statistics.

When  $e$  is high (e.g. scenarios 3 and 4), the value of  $\omega_{\text{true}}$  can significantly impact the posterior constraints on  $e$  and  $b$  due to the degenerate influence that these parameters can have on the observed transit duration, particularly near apastron. On the other hand, we do not observe any noteworthy trends in model accuracy as a function of SNR. For a typical *Kepler* planet which has low  $e$ , non-grazing  $b$ , and  $\omega$  closer to periastron, we would generally expect to measure  $e$  and  $b$  posterior distributions that are somewhat consistent with the true underlying orbital geometry of the planet based on the trends that we observe in Figure 5.4. In Appendix 5.8, we briefly explore whether using a different sampler (i.e. dynamic nested sampling via *dynesty*; Speagle 2020) might yield even more accurate posterior constraints, but our findings there are inconclusive.

### 5.6.3 Our $T_{14} + umb$ method is more efficient than the $e-\omega-\rho$ method

We have shown that the  $T_{14} + umb$  basis can be used as an alternative to the  $e-\omega-\rho$  basis, achieving equivalent results while also reducing the number of parameters by two. This parameter reduction should increase the efficiency of the  $T_{14} + umb$  model, but this approach also requires three separate sampling runs – one for each of the three umbrellas. To evaluate

the overall model efficiencies, we compared the number of effective samples per second ( $\eta$ ) achieved by each method for all injection-recovery tests.

For the  $e\text{-}\omega\text{-}\rho$  method, we measure the number of effective samples from the  $r$  posterior distribution for each test using Geyer’s initial monotone sequence criterion via `arviz` (Geyer 1992; Gelman et al. 2013; Kumar et al. 2019). We select  $r$  because it is a common output between our models and is less affected by complicated parameter degeneracies. We then divide  $N_{\text{eff}}$  by the total run-time for this model to achieve the  $e\text{-}\omega\text{-}\rho$  sampling efficiency:  $\eta_{e\omega\rho}$ . For the  $T_{14} + umb$  method, we average  $N_{\text{eff}}$  of the  $r$  posteriors from each umbrella model, weighted by their respective umbrella weights. We divide this weighted average by the sum of the run-times for the three umbrella models (e.g. the CPU run-time) to achieve the overall  $T_{14} + umb$  sampling efficiency:  $\eta_{T_{14}+umb}$ .

We calculate the ratio of these two efficiencies for all injection-recovery tests and find that  $\eta_{T_{14}+umb}/\eta_{e\omega\rho} > 1$  for  $\sim 73\%$  of tests, suggesting that the  $T_{14} + umb$  approach is generally more efficient across our set of injected planet parameters. The median value of  $\eta_{T_{14}+umb}/\eta_{e\omega\rho}$  across our sample is 2.0, implying that the  $T_{14} + umb$  approach is typically  $2\times$  more efficient than the  $e\text{-}\omega\text{-}\rho$  method, although the range of this efficiency ratio is broad. When we consider  $\eta_{T_{14}+umb}/\eta_{e\omega\rho}$  as a function of SNR, however, we measure a median efficiency increase of  $5.7\times$  at SNR = 80,  $1.2\times$  at SNR = 40, and  $1.1\times$  at SNR = 20 (see Figure 5.5). We also find that the  $T_{14} + umb$  method is only more efficient than the  $e\text{-}\omega\text{-}\rho$  method in  $\sim 52\%$  of low-SNR tests. These findings suggest that the  $T_{14} + umb$  method tends to be less efficient when the transit signal is weaker.

From Figure 5.5, we also see that the efficiency ratio changes with respect to the duration ratio  $T_{14}/T_{14,\text{ref}}$ . For tests with SNR = 20, the median efficiency ratio  $\eta_{T_{14}+umb}/\eta_{e\omega\rho}$  decreases significantly as the duration ratio increases, dropping from  $2.1\times$  at  $T_{14}/T_{14,\text{ref}} \leq 0.8$  to  $0.6\times$  at  $T_{14}/T_{14,\text{ref}} > 0.8$ . This trend is likely due to differences in how the two methods explore the high- $b$  grazing regime. As the duration ratio approaches unity or higher, high  $b$  values

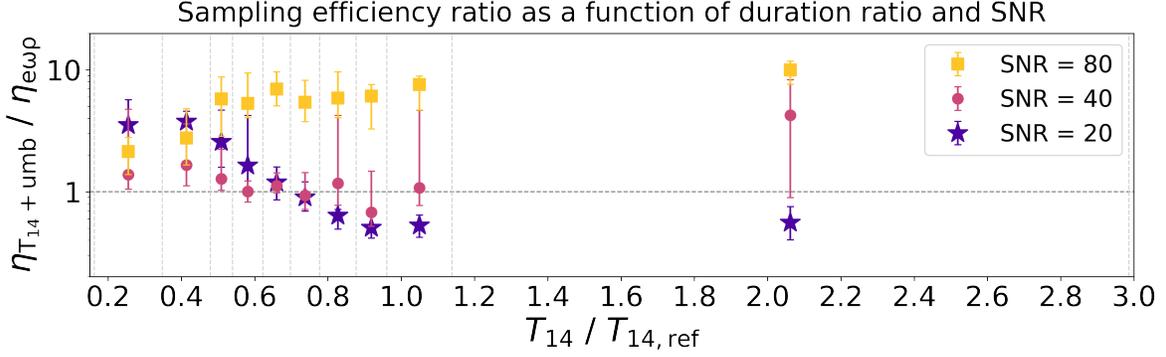


Figure 5.5: Ratio of sampling efficiencies  $\eta_{T_{14}+umb}/\eta_{e\omega\rho}$  as a function of duration ratio  $T_{14}/T_{14,ref}$  and SNR. We bin the data across every 10<sup>th</sup> percentile of the duration ratio distribution, showing a single point per bin per SNR (bins are separated by vertical grey lines). Each point shows the 15<sup>th</sup>, 50<sup>th</sup>, and 85<sup>th</sup> percentiles of a given bin. The increased efficiency of the  $T_{14} + umb$  method relative to the  $e-\omega-\rho$  method depends on the SNR of the modeled lightcurve. At higher SNR the  $T_{14} + umb$  method is significantly more efficient, but at moderate-to-low SNR the two methods have more similar efficiencies. At shorter transit durations, the  $T_{14} + umb$  method is always more efficient, but this behavior changes around  $T_{14}/T_{14,ref} \approx 0.8$ . The large spread in some uncertainties reflects the heterogeneity of our injected lightcurve parameters.

are significantly less likely, but the  $T_{14} + umb$  approach continues to carefully explore the high- $b$  regime via three umbrella models even when it is not necessary. On the other hand, injection-recovery tests with higher  $b$  values (and generally shorter transit durations) are more efficiently sampled by the  $T_{14} + umb$  approach. This behavior is consistent with what we would expect, given that umbrella sampling is specifically intended to ensure accurate measurements of the high- $b$  parameter space.

Our set of injected transit properties, however, is not completely representative of observed planet demographics. To make a more representative comparison, we estimate the efficiency ratio for a typical *Kepler* planet based on both SNR and duration ratio  $T_{14}/T_{14,ref}$ . We use

the latter metric because it reflects the combined effects of  $e$ ,  $b$ , and  $\omega$  in a single variable. For a typical confirmed *Kepler* planet with  $\text{SNR} \approx 20 - 40$  and  $T_{14}/T_{14,\text{ref}} \approx 0.6 - 1.1$ , we estimate an efficiency ratio of  $\sim 0.9\times$ . Based on these findings, we assert that the two methods generally have similar sampling efficiencies for real planetary transit signals, with the  $T_{14} + umb$  approach excelling for signals with higher SNR or lower duration ratio.

The efficiency increase from the  $T_{14} + umb$  approach is more significant when we consider wall-clock time rather than CPU time. Since the three umbrella models can be run in parallel, we can reduce the apparent run-time of the  $T_{14} + umb$  approach by up to a factor of a few. In this parallelized case, the apparent sampling efficiency of the  $T_{14} + umb$  method is  $\sim 1.2\times$  faster than the  $e-\omega-\rho$  method for a typical *Kepler* planet. As another added benefit, the posteriors of the  $T_{14} + umb$  approach can be importance sampled for updated values of  $\rho_*$  (as they become available) *without* re-running the NUTS sampling process (see §5.5.2), which is a major advantage in the long-term efficiency of the  $T_{14} + umb$  parameterization.

## 5.7 Conclusions

In this work, we presented an updated photo-eccentric transit modeling method using a duration-based parameterization  $\{P, t_0, r, b, T_{14}\}$  (with umbrella sampling) and post-hoc importance sampling which efficiently achieves accurate constraints on  $e$ ,  $\omega$ , and  $b$ . Through a suite of synthetic injection-and-recovery tests, we demonstrated that our approach produces equivalent eccentricity constraints relative to the more common eccentricity-explicit transit model parameterization  $\{P, t_0, r, b, e, \omega, \rho\}$ . We find that our modeling method generally has a higher sampling efficiency than the  $e-\omega-\rho$  method when the true  $e$  or  $b$  value is high or a similar efficiency otherwise. Our approach can also be parallelized to increase its relative sampling efficiency several-fold more.

A key advantage of our modeling method is that post-hoc importance sampling allows us to successfully derive accurate  $e$  and  $\omega$  posterior distributions (relative to the  $e-\omega-\rho$  method)

without including  $e$ ,  $\omega$ , or  $\rho$  as explicit model parameters. Our importance sampling routine is fast and flexible enough to easily incorporate an updated prior on  $e$  and/or  $\omega$ , which is critical for hierarchical modeling approaches at the population level. Our method also allows us to update parameter posterior distributions according to updated values of  $\rho_\star$  (e.g. from new *Gaia* data releases) without any loss of generality. In the modern era of high-precision stellar characterization, this sort of “future-proofing” will be invaluable as the number of transit candidates around well-characterized stars continues to grow.

## 5.8 Appendix 5A: Sampler comparison: NUTS vs. Nested Sampling

In the previous sections, we demonstrated that the baseline  $e$ - $\omega$ - $\rho$  model and our alternative  $T_{14} + umb$  approach yield equivalent results when posterior samples are obtained using MCMC methods. Unfortunately, we also saw that posterior inferences of eccentricity can be significantly over- or under-estimated relative to their true values. Here, we explore whether using a different sampling technique – dynamic nested sampling (Skilling 2004; Skilling 2006) – can yield more accurate results and/or serve as a potential alternative to NUTS sampling with umbrella sampling.

We implement the duration-based parameterization using the `dynesty` framework for dynamic nested sampling (Speagle 2020), which does not necessitate the use of umbrella sampling because it already accomplishes the same goal of thoroughly exploring complicated posterior topologies. To model the transit shape and measure the log-likelihood at each sampler step, we use a modified version of `batman` which takes  $\{P, t_0, r, b, T_{14}\}$  as explicit transit parameters (in contrast to the default set  $\{P, t_0, r, b, e, \omega, \rho_\star\}$ ). As before, we perform post-hoc importance sampling to obtain  $\{e, \omega, \rho_\star\}$  samples. We apply this alternative modeling method,  $T_{14} + dyn$ , to all 375 injection-recovery tests in an identical manner as the previous models.

### 5.8.1 Posterior comparison

For each injection-recovery test, we measure the values  $e_{k,T_{14}+dyn}$  from the  $T_{14} + dyn$  eccentricity posterior at the  $k = 15^{\text{th}}$ ,  $50^{\text{th}}$ , and  $85^{\text{th}}$  percentiles of the distribution and compare to the results of the  $T_{14} + umb$  method like in §5.6.1 (Figure 5.6). We find that the  $T_{14} + umb$  and  $T_{14} + dyn$  methods yield eccentricity results that are broadly in agreement. However, there appears to be more differences between samplers ( $T_{14} + umb$  versus  $T_{14} + dyn$ ) than between parameterizations ( $T_{14} + umb$  versus  $e-\omega-\rho$ ). The comparison between parameterizations yielded no test results that consistently differed by  $|\Delta e_k| \geq 0.05$ , but the comparison between samplers yields 42 of such discrepancies. Among these, there are three tests that differ by  $|\Delta e_k| \geq 0.15$  and yield entirely different posterior topologies for  $e$ .

The discrepant measurements of  $\Delta e_k$  are most common at the  $k = 15^{\text{th}}$  percentile, implying that the two sampling methods differ most at sampling the low- $e$  tail of the eccentricity distribution. We observe that the  $T_{14} + dyn$  method produces  $e$  posterior distributions with much less posterior weight in the low- $e$  tail as compared to the results of the  $T_{14} + umb$  approach. We also see a similar divergence of the two methods in the upper tail of the  $b$  posterior distributions. This is consistent with our additional observation that the majority of the discrepancies occur in tests with shorter duration ratios ( $T_{14}/T_{14,\text{ref}} \lesssim 0.5$ ). Most discrepancies also occur at higher SNR levels, counter to expectations. Together, these criteria for discrepant results only match with  $\sim 1\%$  of observed *Kepler* transit signals, implying that real systems are highly unlikely to fall into this subset.

### 5.8.2 Accuracy

We compare the true underlying eccentricity of each injection-recovery test with the measured posterior distribution of  $e$  from the  $T_{14} + dyn$  modeling method. Overall, we find that the qualitative trends in  $e$  and  $b$  measured via the  $T_{14} + dyn$  method are roughly equivalent to those measured from the  $T_{14} + umb$  method (see §5.6.2). We do, however, find a signifi-

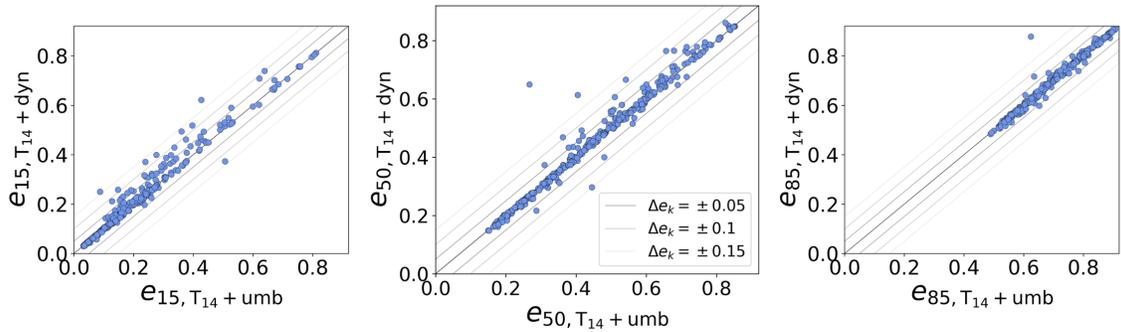


Figure 5.6: Comparison of  $e$  values measured from the  $T_{14} + umb$  and  $T_{14} + dyn$  modeling methods at the 15<sup>th</sup>, 50<sup>th</sup>, and 85<sup>th</sup> percentiles of their distributions. We show the  $\Delta e_k = \{0.05, 0.1, 0.15\}$  in grey, as well as the ideal 1-to-1 line shown in black. These comparisons generally lie close to the 1-to-1 line, implying that the results of the two models are approximately equivalent. However, we do observe 42 tests where the  $T_{14} + umb$  and  $T_{14} + dyn$  model results are discrepant by more than  $|\Delta e_k| \lesssim 0.05$ .

cant difference between the accuracies of the two modeling methods among the three most discrepant injection-recovery tests, where  $|\Delta e_{50}| \geq 0.15$ . For these discrepant tests, seen as outliers in Figure 5.6, the  $T_{14} + dyn$  method achieves more accurate posterior constraints on both  $e$  and  $b$ . This may suggest that differences between samplers can, in some cases, lead to significant differences in the accuracy of modeled parameters. While all three of these tests have  $e_{\text{true}} = 0.8$ , we unfortunately do not find any discernible rules by which to distinguish when sampler differences will lead to substantial differences in the accuracy of posterior results.

### 5.8.3 Efficiency

We also compare these two modeling approaches according their sampling efficiencies. We calculate the efficiency  $\eta_{T_{14}+dyn}$  of the  $T_{14} + dyn$  approach for each injection-recovery test, based on the number of effective samples measured via the Kish 1965 approach using `dynesty`. Similar to §5.6.3, we compute the efficiency ratio between the  $T_{14} + dyn$  model and

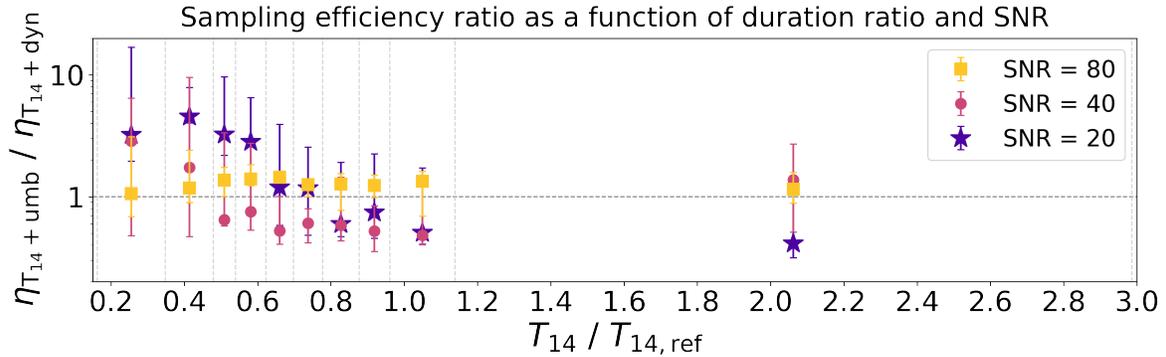


Figure 5.7: Ratio of sampling efficiencies  $\eta_{T_{14}+umb}/\eta_{T_{14}+dyn}$  as a function of duration ratio  $T_{14}/T_{14,ref}$  and SNR. We bin the data across every 10<sup>th</sup> percentile of the duration ratio distribution, showing a single point per bin per SNR (bins are separated by vertical grey lines). Each point shows the 15<sup>th</sup>, 50<sup>th</sup>, and 85<sup>th</sup> percentiles of a given bin. The efficiency of the  $T_{14} + umb$  method relative to the  $T_{14} + dyn$  method depends partially on the SNR of the modeled lightcurve. At shorter transit durations, the  $T_{14} + umb$  method is typically more efficient, but the opposite is true at longer transit durations. The large spread in some uncertainties reflects the heterogeneity of our injected lightcurve parameters.

our  $T_{14} + umb$  approach ( $\eta_{T_{14}+umb}/\eta_{T_{14}+dyn}$ ) and show these results in Figure 5.7. The distribution of efficiency ratios among our sample is broad but suggests that the two methods generally have similar sampling efficiencies, with a median efficiency ratio of  $\eta_{T_{14}+umb}/\eta_{T_{14}+dyn} \approx 1.1$ . At lower duration ratios, the  $T_{14} + umb$  approach is  $\sim 1.4\times$  more efficient, which is to be expected since this part of parameter space includes higher  $b$  values – the specialty of umbrella sampling as implemented by Gilbert (2022).

For a typical *Kepler* planet, however, we estimate that the  $T_{14} + dyn$  method is  $\sim 1.6\times$  faster than the  $T_{14} + umb$  approach. This observation, along with an occasional improvement in accuracy, leans in favor of dynamic nested sampling compared to NUTS sampling + umbrella sampling for our tests, but there are many other compounding factors that are beyond the scope of our experiment. Overall, both sampling methods offer their own benefits with

neither winning out 100% of the time, but it is clear that the duration-based parameterization performs well regardless of the underlying sampling method.

## CHAPTER 6

# The TESS-Keck Survey. XV. Precise Properties of 108 TESS Planets and Their Host Stars

### 6.1 Abstract

We present the stellar and planetary properties for 85 TESS Objects of Interest (TOIs) hosting 108 planet candidates which comprise the TESS-Keck Survey (TKS) sample. We combine photometry, high-resolution spectroscopy, and *Gaia* parallaxes to measure precise and accurate stellar properties. We then use these parameters as inputs to a lightcurve processing pipeline to recover planetary signals and homogeneously fit their transit properties. Among these transit fits, we detect significant transit-timing variations among at least three multi-planet systems (TOI-1136, TOI-1246, TOI-1339) and at least one single-planet system (TOI-1279). We also reduce the uncertainties on planet-to-star radius ratios  $R_p/R_\star$  across our sample, from a median fractional uncertainty of 8.8% among the original TOI Catalog values to 3.0% among our updated results. With this improvement, we are able to recover the Radius Gap among small TKS planets and find that the topology of the Radius Gap among our sample is broadly consistent with that measured among *Kepler* planets. The stellar and planetary properties presented here will facilitate follow-up investigations of both individual TOIs and broader trends in planet properties, system dynamics, and the evolution of planetary systems.

## 6.2 Introduction

The NASA Transiting Exoplanet Survey Satellite (TESS; [Ricker et al. 2015a](#)) is currently in its 5<sup>th</sup> year of carrying out an all-sky survey in search of transiting planets orbiting nearby bright stars. So far, TESS has revealed over 6,000 planet candidates, building upon the legacies of its predecessors, NASA’s *Kepler* ([Borucki et al. 2010a](#)) and *K2* ([Howell et al. 2014](#)). These TESS Objects of Interest (TOIs) are now the subjects of numerous follow-up studies to verify their planetary nature and measure their properties (e.g. [Kane et al. 2021](#); [Teske et al. 2021](#); [Chontos et al. 2022](#); [Yee et al. 2022](#)). The precise characterization of planet properties, however, is greatly dependent on the robustness of the input stellar parameters and lightcurve modeling procedure.

The TESS Input Catalog (TIC; [Stassun et al. 2018](#), [Stassun et al. 2019](#); [ExoFOP 2019](#)) contains stellar parameters for all TESS targets which were measured from broadband photometry and parallaxes. These stellar properties were used as inputs to the TESS data processing pipeline developed by the Science Processing Operations Center (SPOC; [Jenkins et al. 2016](#)), which flags planet candidates as TOIs and performs an initial characterization ([Guerrero et al. 2021](#); [NASA Exoplanet Archive 2022b](#)). The propagation of large uncertainties throughout this process can lead to miscalculated or poorly constrained planet radii for TOIs, with a median fractional radius uncertainty of  $\sigma(R_p)/R_p \approx 7.4\%$  and a mean fractional uncertainty of  $\sim 20.2\%$ . This large discrepancy between the median and mean uncertainty is due to a significant high- $\sigma(R_p)$  tail, with  $\sim 1$  in 5 TOIs having  $\sigma(R_p)/R_p > 20\%$ .

The TESS-Keck Survey (TKS) has set out to confirm and characterize a sample of 108 TOIs orbiting 85 TESS host stars ([Chontos et al. 2022](#)). The TKS collaboration aims to precisely measure the stellar properties of this sample using spectra taken with the High-Resolution Spectrograph (HIRES; [Vogt et al. 1994](#)) on the Keck I telescope at the W. M. Keck Observatory. These spectra also allow us to dynamically confirm the planetary nature of the TKS TOIs through the measurement of precise radial velocities (RVs), leading to

numerous planet discoveries among the TKS team (e.g. Dalba et al. 2020; Dai et al. 2020; Weiss et al. 2021; Rubenzahl et al. 2021; Scarsdale et al. 2021; MacDougall et al. 2021; Dalba et al. 2022; Lubin et al. 2022; Dai et al. 2022; Turtelboom et al. 2022; MacDougall et al. 2022; Van Zandt et al. 2022).

To ensure consistency among the planetary and stellar parameters reported for TKS targets, we use the stellar parameters measured with our spectroscopic constraints as inputs to a lightcurve detrending and transit fitting pipeline applied homogeneously to all TKS targets. The planet properties derived by our pipeline allow us to begin addressing the major TKS Science Cases (see Chontos et al. 2022), such as improving our understanding of planet bulk compositions, system architectures and dynamics, planetary atmospheres, and the role of stellar evolution in shaping planetary systems.

In this work, we describe our homogeneous characterization of both planetary and stellar properties of TKS targets, derived from transit photometry and single stellar spectra. In §6.3, we update the stellar properties reported by Chontos et al. 2022 with additional photometric constraints and compare these results against those reported by the TOI Catalog and *Gaia* Data Release 2. We also derive estimates of the quadratic limb darkening parameters for each TOI host to set priors on our transit model. In §6.4, we describe the data processing pipeline used to retrieve, detrend, and fit the lightcurve photometry for our target sample. We account for both stellar variability and transit-timing variations in order to properly model the transit signals. We use the results from our transit fits and stellar characterization to derive various planet parameters in §6.5, including planet radius which we measure to a median uncertainty of 3.8%. In §6.6, we offer some preliminary insight gained through this analysis to begin addressing the science goals of TKS which will be evaluated in more depth in future works by the TKS team. Planet masses, precise eccentricities, and non-transiting companions are not addressed in this transit-focused study, and we leave that analysis to a full radial velocity analysis by TKS collaborators (Polanski et al. 2023, in prep).

## 6.3 Stellar Properties

### 6.3.1 Spectroscopic Constraints

The TKS collaboration has collected spectra for the 85 TOIs from [Chontos et al. \(2022\)](#) at a spectral resolution of  $R = 50,000$  using the HIRES instrument at the W.M. Keck Observatory ([Vogt et al. 1994](#)) from Summer 2019 to Fall 2022. We used iodine-free reconnaissance spectra, with  $S/N \approx 40/\text{pixel}$  across  $3600 - 9000 \text{ \AA}$ , to check for rapid stellar rotation and rule out spectroscopic false-positives. We also extracted stellar effective temperature  $T_{\text{eff}}$ , metallicity  $[\text{Fe}/\text{H}]$ , and surface gravity  $\log g$  from these spectra using two different methods according to the estimated effective temperature per target. The two methods used were `SpecMatch-Syn` (synthetic; [Petigura 2015](#)) and `SpecMatch-Emp` (empirical; [Yee et al. 2017](#)). `SpecMatch-Syn` compares a given stellar spectrum to synthetic spectra generated by interpolating within a grid of modeled spectra from a library described by [Coelho et al. \(2005\)](#). `SpecMatch-Emp` fits a stellar spectrum via direct comparison to observed optical spectra from a dense library of stars with well-measured properties.

We initially processed the spectra of each target using both techniques and used their respective valid ranges of effective temperature to determine which set of derived properties is more reliable for a given star. Specifically, we use `SpecMatch-Syn` measurements for stars with estimated effective temperature between  $4800 - 6500 \text{ K}$ , as determined by `SpecMatch-Emp`. For stars with estimated  $T_{\text{eff}}$  beyond this range (i.e.  $T_{\text{eff}} < 4800 \text{ K}$  and  $T_{\text{eff}} > 6500 \text{ K}$ ), we use `SpecMatch-Emp` measurements. In cases where `SpecMatch-Emp` is used, the `SpecMatch` model does not output an estimate of  $\log g$ .

### 6.3.2 Isochrone Modeling

#### 6.3.2.1 Input values

We derive additional stellar properties such as mass  $M_{\star}$ , radius  $R_{\star}$ , density  $\rho_{\star}$ , luminosity  $L_{\star}$ , age, and extinction  $A_v$  using `isoclassify` ([Berger et al. 2020a](#); [Huber et al. 2017](#)).

This software interpolates between the MESA Isochrones and Stellar Tracks models (MIST; Dotter 2016; Choi et al. 2016) to measure the best-fitting solution of stellar properties for a given set of input spectroscopic and photometric constraints. For M-dwarfs, `isoclassify` uses empirical relations from Mann et al. (2019) to derive fundamental stellar properties (see Berger et al. 2020a for details). `isoclassify` has been shown to produce robust results and was the foundation for deriving the *Gaia-Kepler* stellar properties catalog (Berger et al. 2020a).

We used the outputs of the preferred `SpecMatch` model per target as priors to constrain the isochrone grid space explored by our `isoclassify` models. This includes priors on  $T_{\text{eff}}$  and  $[\text{Fe}/\text{H}]$  from either `SpecMatch` method, along with a prior on  $\log g$  when `SpecMatch-Syn` is preferred. We also include priors on stellar parallax drawn from *Gaia* DR2, with a median uncertainty of 0.03 mas. We incorporate photometric inputs for several all-sky photometric bands from the *2MASS* (Skrutskie et al. 2006) and *Gaia* missions to further constrain our `isoclassify` results. These bands are *2MASS*  $J$ ,  $H$ ,  $K$  and *Gaia* DR2  $G$ ,  $B_p$ ,  $R_p$ , from which various photometric colors are used to calibrate the underlying isochrone model grid. We also select *2MASS*  $K$  as the photometric band to be used for absolute magnitude calculations during `isoclassify` modeling. All input parameter priors are assumed to be normally distributed.

The inclusion of multiple photometric constraints allows us to fit for extinction directly using `isoclassify` without needing to specify a generalized dust map (see, e.g. Lallement et al. 2019), making our extinction and metallicity estimates more accurate on a target-by-target basis. This improves the accuracy of our stellar characterizations as compared to Chontos et al. (2022) where only a single photometric band was used with a generalized dust map. The median uncertainty for the *2MASS* photometry that we use is 0.02 mag. To account for systematic zero-point differences in photometric systems, we use a standardized uncertainty floor of 0.01 mag for *Gaia* photometry.

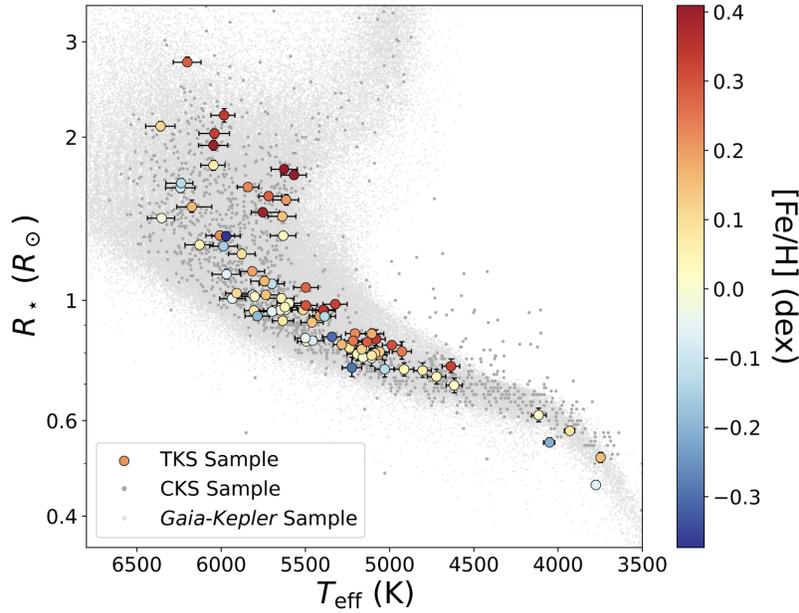


Figure 6.1: Stellar radius as a function of effective temperature for the 85 host stars in the TKS sample, with metallicity shown as color gradient. The California-*Kepler* Survey sample (Petigura et al. 2022) and the *Gaia-Kepler* sample (Berger et al. 2020b) are shown for reference in dark and light grey, respectively.

With these inputs to `isoclassify`, we perform stellar characterization via isochrone grid modeling to measure  $M_*$ ,  $R_*$ ,  $\rho_*$ ,  $L_*$ , age, and  $A_v$  for all 85 TKS targets, in addition to spectroscopically constrained  $T_{\text{eff}}$ ,  $[\text{Fe}/\text{H}]$ , and  $\log g$ . This catalog of precise homogeneously derived stellar properties will serve as a reference for all future TKS studies and as a resource for the broader astronomy community. We show the relationship between  $R_*-T_{\text{eff}}-[\text{Fe}/\text{H}]$  in Figure 6.1 for this sample, as compared to the distributions of both the California-*Kepler* Survey sample (Petigura et al. 2017a) and the *Gaia-Kepler* sample (Berger et al. 2020b). We note that the TKS sample includes a broad range of stellar types, similar to the distributions of stars studied of similar past works.

### 6.3.2.2 Uncertainties

The median uncertainties that we measure for  $R_\star$ ,  $M_\star$ ,  $T_{\text{eff}}$ , and  $\log g$  are reported in Table 6.1, as compared against the original TOI Catalog source uncertainties (Guerrero et al. 2021; NASA Exoplanet Archive 2022b). We do not currently account for model-dependent uncertainties associated with the MIST models to maintain consistency with the reported uncertainties by the TOI Catalog and *Gaia* DR2 catalog. The appropriate corrections (described by Tayar et al. 2022) can be added in quadrature to our error values to encapsulate this systematic error. When we perform this correction, we find updated median fractional uncertainties of 4.5%, 4.0%, 1.3%, and 0.6% for  $R_\star$ ,  $M_\star$ ,  $T_{\text{eff}}$ , and  $\log g$ , respectively. Most notable among these is the change in the stellar radius uncertainty, which increased from 1.7% to 4.5% when the systematic error was introduced. This ultimately propagates to an increase in the median  $R_p$  uncertainty as well, from 3.8% to 5.8%. Even with this additional systematic uncertainty on  $R_p$ , our results remain better constrained compared to the TOI Catalog.

We also note that two targets from the TKS sample have missing or erroneous *2MASS* catalog uncertainties. TOI-1807 has an anomalously large  $K$ -mag uncertainty ( $\sigma_K = 10$  mag) and TOI-1246 is missing an uncertainty value for  $J$ -mag. We replaced both of these with a standardized uncertainty value of 0.02 mag, roughly consistent with the median errors for these bands among our sample.

### 6.3.2.3 Comparison with TOI Catalog and *Gaia* results

To validate our stellar parameters, we compared our final values of  $T_{\text{eff}}$  and  $R_\star$  from *isoclassify* with the TOI Catalog and *Gaia* DR2. We show these comparisons in Figure 6.2. The stellar parameters measured from these external sources were calculated using either photometry with parallaxes (TOI Catalog; Stassun et al. 2018) or a combination of photometry with parallaxes and low-resolution spectra (*Gaia* DR2; Bailer-Jones et al. 2013).

For each comparison, we calculated the ratio of the given parameter between our TKS results and the external source data (i.e.  $R_{\star,\text{TKS}}/R_{\star,\text{TOI}}$ ) to demonstrate the lack of bias in these results. We measured the mean and standard deviation of these ratio distributions for each comparison and found: (Fig. 6.2a)  $R_{\star,\text{TKS}}/R_{\star,\text{Gaia}} = 0.99 \pm 0.04$ , (Fig. 6.2b)  $T_{\text{eff,TKS}}/T_{\text{eff,Gaia}} = 1.00 \pm 0.01$ , (Fig. 6.2c)  $R_{\star,\text{TKS}}/R_{\star,\text{TOI}} = 0.99 \pm 0.03$ , and (Fig. 6.2d)  $T_{\text{eff,TKS}}/T_{\text{eff,TOI}} = 1.00 \pm 0.02$ . The scatter between our TKS results and the external data source for each comparison was consistent with the reported parameter uncertainty. This implies a general agreement between our TKS stellar parameters and those reported by external data sources. We do find that the *Gaia* DR2 pipeline produces overestimated  $T_{\text{eff}}$  for cooler TKS stars, but this is consistent with a known degradation in the precision of the *Gaia* DR2 pipeline around  $T_{\text{eff}} \approx 4000$  K (Andrae, René et al. 2018).

From these comparisons, we flag two targets with discrepancies of  $>3\sigma$  between our results and the values reported from the external data sources. TOI-2145 had an estimated effective temperature whose uncertainty spanned across the threshold of reliability between *SpecMatch-Emp* and *SpecMatch-Syn*. Selecting between these two methods following the criteria in §6.3.1 yielded a radius measurement for this target that was inconsistent with both the TOI Catalog and *Gaia* DR2 estimates. By manually swapping methods and allowing for an exception to these rules, we found good alignment between our final radius measurement and the values reported from the other two sources. The remaining flagged target (TOI-2114), which can be seen in Figure 6.2d, has a discrepant  $T_{\text{eff}}$  when we compare our result to the TOI Catalog. However, our result remains highly consistent with the  $T_{\text{eff}}$  reported by *Gaia* DR2 and other external sources. Given this observation and the relatively large uncertainties reported by the TOI Catalog, we attribute the disagreement to slight differences in modeling methods and proceed without making any adjustments for this particular target.

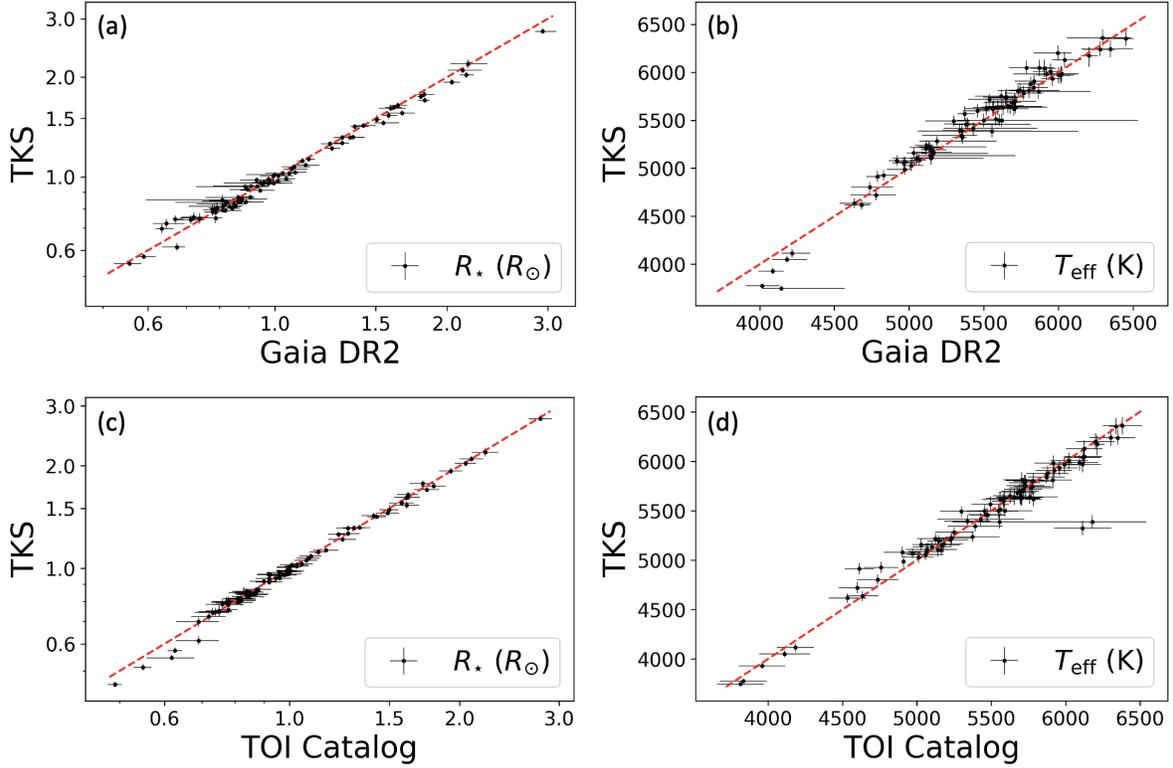


Figure 6.2: Stellar radius (left) and effective temperature (right) as measured by our stellar characterization procedure (TKS) on the y-axis compared to the values reported by the *Gaia* DR2 pipeline (top) and the TOI Catalog (bottom) on the x-axis. The TKS results produce reduced uncertainties and are generally consistent with one or both external sources. The residuals of our comparison between  $R_{*,\text{TKS}}$  and  $R_{*,\text{TOI}}$  can be seen in Figure 6.6a. We also note that the *Gaia* DR2 pipeline does not report stellar properties when  $R_{*,\text{Gaia}} \lesssim 0.5R_\odot$ , leaving out two TKS targets.

### 6.3.3 Limb Darkening

Stellar limb darkening is the decrease in brightness of the disk of star from its center towards its edge or limb. A set of quadratic coefficients is often used to define this limb darkening function, which describes the ingress and egress of a transit profile (Mandel & Agol 2002a). An accurate estimate of limb darkening coefficients is crucial for measuring accurate planet properties, especially for planets that transit closer to the limb of a star (e.g. higher impact parameter or inclination).

There are two leading stellar atmosphere models that are widely used for calculating quadratic limb darkening coefficients: the PHOENIX model (Hauschildt & Baron 1999) as implemented by Husser et al. (2013) and the ATLAS models (Castelli & Kurucz 2004) as implemented by Claret & Bloemen (2011). Investigations into the accuracy of these two methods (see, e.g., Patel & Espinoza 2022) have demonstrated that such models produce quadratic limb darkening coefficients with an average discrepancy of  $\sim 0.1$ . For reference, a typical quadratic limb darkening value for a solar-like star is in the range 0.1 – 0.5, so an uncertainty of 0.1 is typically  $>20\%$  of the nominal measurement. Additionally, this formal uncertainty between methods only encompasses part of the true uncertainty that arises when measuring these coefficients from atmospheric models. With this in mind, we use the `limb-darkening` code (Espinoza & Jordán 2015) to easily compute both the PHOENIX-derived and ATLAS-derived limb darkening coefficients for all targets in our sample. This derivation requires several stellar inputs, which we draw from our `isoclassify` results:  $T_{\text{eff}}$ ,  $[\text{Fe}/\text{H}]$ , and  $\log g$ . We then use these values to adopt nominal estimates of coefficients  $\{u_1, u_2\}$  for each target with a systematic noise floor of 0.1.

Given the large uncertainty in these measurements, we establish a simple set of rules to determine which set of limb darkening coefficients ( $u_i$ ) to use in deriving our final catalog of system properties. (1) When  $|u_{i,\text{ATLAS}} - u_{i,\text{PHOENIX}}| < 0.1$ , we select  $\text{AVG}(u_{i,\text{ATLAS}}, u_{i,\text{PHOENIX}})$  as the final coefficient values. (2) When  $|u_{i,\text{ATLAS}} - u_{i,\text{PHOENIX}}| \geq 0.1$ , we select

Table 6.1: TKS System Parameters: Median Fractional Uncertainties

Parameter		TKS Result Unc.	TOI Catalog Unc.
		%	%
<i>Star</i>	$R_\star$	1.7	5.1
	$M_\star$	3.0	12.6
	$T_{\text{eff}}$	1.2	2.3
	$\log g$	0.52	1.9
<i>Planet</i>	$R_p$	3.8	6.7
	$R_p/R_\star$	3.0	8.8
	$T_{14}$	1.8	9.7
	$P$	$2.2 \times 10^{-4}$	$3.8 \times 10^{-4}$
	$t_0$	$7.1 \times 10^{-5}$	$7.3 \times 10^{-5}$

Note. — Reported fractional uncertainties for several stellar and planetary parameters compared between our TKS results and those reported in the TOI Catalog, demonstrating our improved precision across the TKS sample. The values presented here do not account for systematic uncertainties that may arise from differences in underlying stellar model grids (see §6.3.2.2).

$u_{i,\text{ATLAS}}$  as the final coefficient values. These rules were established upon consideration of the systematic comparison of these two methods carried out by [Patel & Espinoza 2022](#). We also assume a standardized uncertainty of 0.1 for all limb darkening coefficients among the TKS sample, which is later propagated into a model prior during our transit fitting procedure.

## 6.4 Light Curve Modeling

### 6.4.1 Input Data

We perform our light curve detrending and modeling using transit photometry from Sectors 1–60 of the TESS mission. We accessed the Pre-search Data Conditioning Simple Aperture Photometry (PDC-SAP; [Stumpe et al. 2012](#); [Stumpe et al. 2014](#); [Smith et al. 2012](#)) through the Mikulski Archive for Space Telescopes (MAST). For the majority of TKS targets, we used 2-min-cadence time-series photometry processed by the primary TESS Science Processing Operations Center pipeline (SPOC; [Jenkins et al. 2016](#)). For TOI-1386, 1601, and 2045, photometry from the primary SPOC pipeline was unavailable, so we instead used PDC-SAP photometry from the TESS-SPOC updated pipeline which produces lightcurves for targets selected from full-frame image observations ([Caldwell et al. 2020](#)).

### 6.4.2 False-Positive Vetting

While false-positive vetting is not a primary focus of this work, we acknowledge that the resolved false-positive rate amongst *all* TOIs is on the order of  $\sim 20\text{--}30\%$  ([NASA Exoplanet Archive 2022b](#); [Cacciapuoti et al. 2022](#); [Magliano et al. 2023](#)). However, this is an extreme upper bound for the TKS sample. We specifically selected our sample to include *only* TOIs with unambiguous planetary dispositions confirmed by both the initial TOI vetting pipeline ([Guerrero et al. 2021](#)) and the TESS Follow-up Program Working Group.

Additionally, all TKS targets were individually vetted by members of the TKS collaboration, as detailed in [Chontos et al. \(2022\)](#). This vetting process allowed us to establish a set of

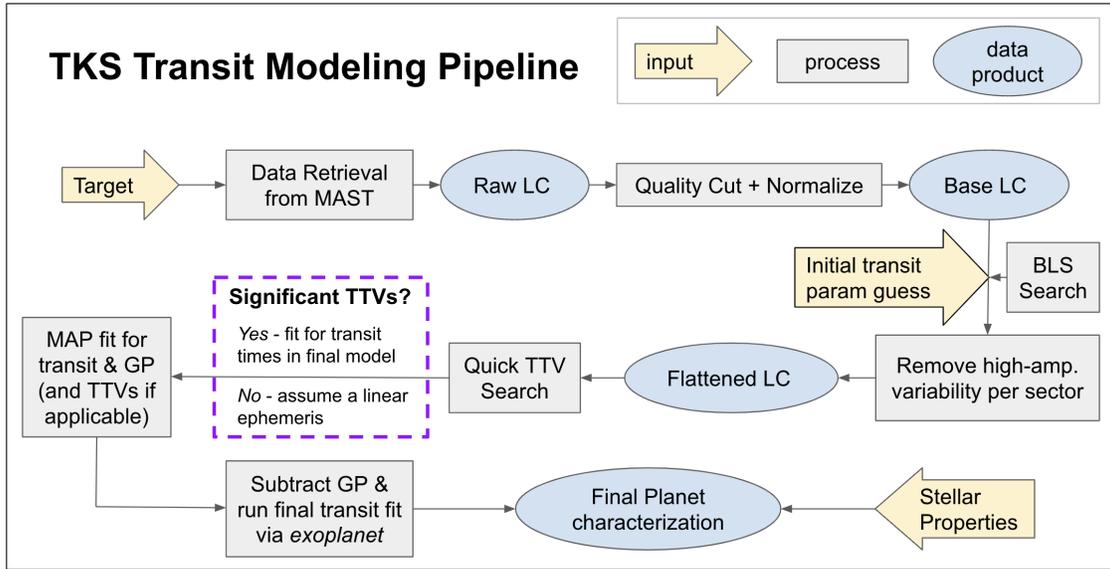


Figure 6.3: Diagram demonstrating the flow of data throughout the TKS light curve modeling pipeline. Yellow arrows are data inputs, grey boxes are processes, and blue ovals are data products.

selection rules to minimize the risk of including unresolved false-positives in our final target sample by making cuts based on transit S/N, out-of-transit centroid offsets, and the proximity of stellar companions. Therefore, we suspect that the false-positive rate in the TKS sample is much lower than the quoted upper bound, but we leave a quantitative assessment of this value for later work. Should any of the TKS TOIs be deemed false-positives, the planetary properties presented here would need to be revisited.

### 6.4.3 Lightcurve Pre-Processing

We apply the following pre-processing routine homogeneously across all lightcurves for the 85 TKS targets. A complete outline of this pre-processing routine and our subsequent modeling pipeline can be found in a flowchart in Figure 6.3. We first mask all points that were flagged for poor quality (quality flag  $> 0$ ), with the exception of TOI-1456 for which a valid transit signal is mistakenly flagged as scattered light contamination by the SPOC pipeline (see Dalba

et al. 2020). We then normalize each lightcurve sector iteratively and subtract 1 from the normalized flux values such that  $\mu_{\text{flux}} = 0$  for all sectors. We also flag four lightcurves as having high stellar variability, which we determine based on whether or not the standard deviation of a target’s out-of-transit flux is  $>3\sigma_{\text{flux}}$ , where  $\sigma_{\text{flux}}$  is the median uncertainty of the individual photometric data points. These four targets are TOI-1136, 1726, 1807, and 2076. We also apply an upper sigma cut of  $5\sigma$  to remove high photometric outliers among targets that do not meet this high-variability criteria.

For light curves with low stellar variability, we perform a box least squares (BLS; Kovács et al. 2002a) search to determine initial estimates of the orbital period and first transit mid-point of each planet. We compare the BLS search results to estimates taken from the primary TOI Catalog (NASA Exoplanet Archive 2022b), flagging any results with  $>3\sigma$  discrepancy when compared to this source. We then follow up flagged targets with a manual inspection, determining that all discrepancies are likely the result of new photometry assisting in pinning down the orbital period with greater precision. In cases of high stellar variability, we do not perform this initial BLS search and instead proceed with the values drawn directly from the TOI Catalog or other literature as our initial guesses.

We use these period and mid-point values, along with transit durations from the TOI catalog, to mask out all transits with a mask width of  $\pm 0.75$  times the transit duration from the transit mid-point. While this mask does not yet account for transit-timing variations (TTVs), we use it here for an initial detrending step which we improve upon in our subsequent fit in §6.4.5. With the masked photometry, we modeled stellar variability in each lightcurve sector using a Gaussian Process (GP) via `celerite2` (Foreman-Mackey 2018) through the `exoplanet` interface (Foreman-Mackey et al. 2021a). The GP fit includes a term to model the variability via a stochastically-driven damped harmonic oscillator with quality factor  $Q = 1/\sqrt{2}$ , with log-normal priors on both the undamped period ( $\rho_{\text{GP}}$ ) and the standard deviation of the process ( $\sigma_{\text{GP}}$ ) with means  $\mu_{\rho_{\text{GP}}} = 0$  and  $\mu_{\sigma_{\text{GP}}} = \ln(\sigma_{\text{flux}})$ , respectively.

We interpolate over the masked transits to create a smooth fit to the variability for each lightcurve sector and subtract this trend from the original unmasked data to produce an initial flattened lightcurve. We stitch together the flattened photometry for each sector and save the maximum *a posteriori* (MAP) parameters for the GP fit which we use later as informed initial guesses in a more complete lightcurve model.

#### 6.4.4 Estimating TTVs

Along with determining an initial fit to the stellar variability trend, we also use the `exoplanet` lightcurve modeling package to find an initial MAP fit to the transit times for each planet. We use our previous  $P$  and  $t_0$  estimates to calculate a set of initial guesses for the transit times assuming a linear ephemeris. We model the true transit times for each planet using an initial Gaussian prior for each transit time, centered on their linear ephemeris values with a standard deviation of 0.1 day. Along with the transit times, we also fit the planet-to-star radius ratio ( $R_p/R_*$ ), impact parameter ( $b$ ), transit duration from 1st-to-4th contact ( $T_{14}$ ), quadratic limb darkening parameters ( $u_1, u_2$ ), mean out-of-transit flux ( $\mu$ ), and lightcurve jitter ( $\sigma_{LC}$ ). The orbital period ( $P$ ) and initial transit mid-point ( $t_0$ ) are also measured as derived values. The priors used on all modeled parameters are given in Table 6.2 and described in more detail in §6.4.5.

This MAP fit to the transit shape and transit times allows us to get a quick estimate of any TTVs detected for each planet. Based on these estimates, we determine for each planet whether to proceed with a final model that assumes a linear ephemeris or includes transit-time fits with a Gaussian prior. We use the following criteria to determine if the estimated TTVs warrant further consideration in our full lightcurve model: (1) more than 3 high-quality transits, (2) a TTV standard deviation greater than the photometric exposure time (i.e. 2 minutes), and (3) a single-transit signal-to-noise ratio of 2 or greater. Since we expect a linear ephemeris for most planets in our sample, we do not offer an in-depth analysis of these variations here and instead leave that analysis to later work (see §6.6.3 for more details).

Table 6.2: Transit Model Priors

Parameter	Prior
$P$ (days)	$\log P \sim \text{N}(\log P_{\text{init}}, 1)$
$t_0$ (days)	$t_0 \sim \text{N}(t_{0, \text{init}}, 1)$
$R_p/R_\star$	$\log(R_p/R_\star) \sim \text{U}(-9, 0)$
$b$	$b \sim \text{U}(0, 1 + R_p/R_\star)$
$T_{14}$ (days)	$\log T_{14} \sim \text{U}(-9, 0)$
$u_i$	$u_i \sim \text{N}(u_{i, \text{init}}, 0.1)$
$\mu$	$\mu \sim \text{N}(0, 1)$
$\sigma_{\text{LC}}$	$\log \sigma_{\text{LC}} \sim \text{N}(\log \sigma_{\text{flux}}, 1.0)$

Note. — Priors on final transit model parameters.  $P$ ,  $T_{14}$ , and  $t_0$  are all given in units of days, with  $t_0$  using the reference frame BJD - 2457000. Priors include normal (N) and uniform (U) distributions.

### 6.4.5 Full Lightcurve Model

We use the MAP values of all model parameters from both our trend model and TTV model as well-informed initial guesses for a complete MAP fit to the normalized lightcurve photometry. We simultaneously fit for the stellar variability trend, transit times, and transit shape using `exoplanet` with the same model priors as were used in our previous MAP fits. This full model allows us to measure photometric variability via Gaussian Processes (as was done in §6.4.3) while also accounting for the true transit shapes and times. We subtract this MAP fit of the photometric variability to produce the final de-trended lightcurve for our analysis (see Figure 6.4 for an example). We create a transit mask using the MAP values from this full model fit to exclude out-of-transit photometry beyond  $\pm 2.5 T_{14, \text{MAP}}$  from each transit mid-point time, removing data that is uninformative of the transit fit in order to improve posterior sampling efficiency.

With our fully de-trended photometry and TTV estimates from this full MAP fit, we build our final model intended for posterior sampling and robustly fitting the transit shape. For planets with TTV estimates below the thresholds described in §6.4.4, we assume a linear ephemeris and do not directly fit transit times in the final model. Altogether, we fit for  $\{P, t_0, R_p/R_\star, b, T_{14}, u_1, u_2, \mu, \sigma_{\text{LC}}\}$  and sometimes transit times  $\{TT_i\}$ , conditioned on our fully de-trended TESS lightcurve.

We apply priors to each of the lightcurve model parameters, briefly outlined below and in Table 6.2. We apply Gaussian priors on  $t_0$ ,  $\mu$ ,  $u_1$ , and  $u_2$ , each centered on estimates that were reasonably well constrained from earlier analysis. We use a log-normal prior on  $\sigma_{\text{LC}}$  to ensure positive values, with  $\mu(\sigma_{\text{LC}}) = \log(\sigma_{\text{flux}})$  and  $\sigma(\sigma_{\text{LC}}) = 1$ . For  $R_p/R_\star$  and  $T_{14}$ , we use log-uniform priors with upper and lower bounds described in Table 6.2, designed to minimize the effect that these boundaries have on the completeness of the posterior sampling. We also model  $b$  with a uniform prior from 0 to  $1 + R_p/R_\star$  to encompass the regime of grazing transits in our posterior space.

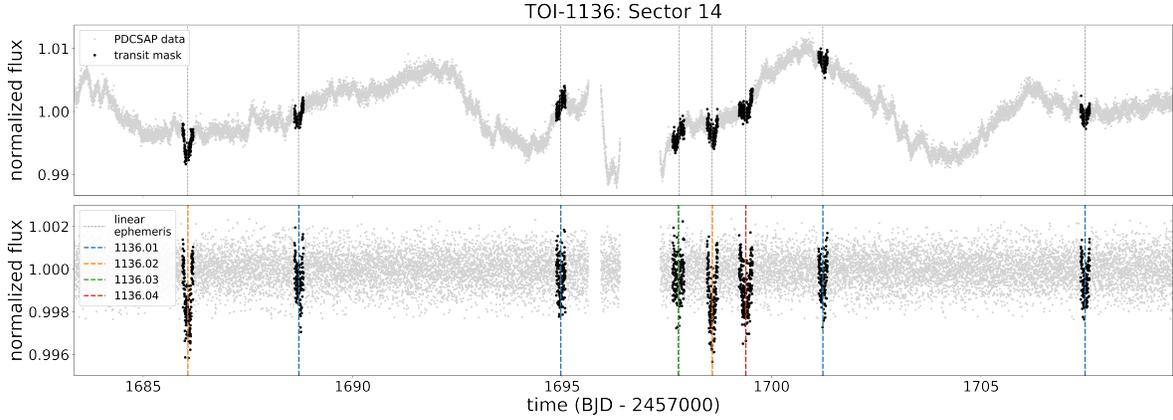


Figure 6.4: Before (top) and after (bottom) applying our de-trending and initial TTV fitting procedure to the photometry of TOI-1136, a multi-planet system with high stellar variability. We only show the first sector (Sector 14) of this target’s photometry here, but all sectors display similar variability. We model our de-trended transit photometry and measure  $R_p/R_\star$  values for all 4 planets which agree with the values reported by Dai et al. 2022 within  $\sim 1.5\sigma$ , as opposed to the TOI Catalog values which are all more than  $1.5\sigma$  discrepant. Black points highlight transits ( $\pm 0.75$  times  $T_{14}$ ), colored vertical dashed lines show modeled transit mid-points, and dark grey vertical dashed lines show calculated transit mid-points assuming a linear ephemeris. In the sector shown, we find the most significant TTVs for TOI-1136.03 (green) as can be seen by the separation between the modeled and calculated mid-points.

Here, we use  $T_{14}$  to directly fit the transit duration rather than measuring it through indirect means such as sampling in "circular" stellar density (see, e.g. Dawson & Johnson 2012) or simultaneously fitting the true stellar density ( $\rho_*$ ), eccentricity ( $e$ ), and argument of periastron ( $\omega$ ). While the former method has been found to produce biased results (see Gilbert et al. 2022), the latter is unbiased but requires three parameters to measure that which can be described by one ( $T_{14}$ ) more efficiently. Our duration-sampling method also allows us to calculate posterior distributions for  $e$  and  $\omega$  post-modeling through importance sampling, described further in §6.5. Regardless of the method used, we emphasize that most  $\{e, \omega\}$  constraints measured from photometry alone are imprecise for individual planets and are better applied towards population-level studies (§6.6.4).

We sample the posterior probability density of all model parameters with `exoplanet` using the gradient-based No U-Turn Sampling method (Hoffman & Gelman 2011a; Betancourt 2016) as implemented by PyMC3 (Salvatier et al. 2016a). For all posterior sampling performed in this work, we use 5,000 tuning steps with an additional 3,000 sampler draws and a target acceptance fraction of 0.95. This sampling process is performed via two sampler chains across two CPU cores, and we consider the process to be converged when the Gelman-Rubin statistic  $\hat{R}$  for these chains is  $\hat{R} < 1.01$  (Gelman & Rubin 1992).

## 6.5 Planet Properties

### 6.5.1 Importance-Weighted Posterior Distributions

Our final planet property measurements are based on the posterior distributions measured with `exoplanet`. These properties are modeled via a fit to the transit shape, conditioned on the input lightcurve data. However, our selected modeling basis does not inherently assume that any information is known about  $\rho_*$ ,  $e$ , or  $\omega$  and, therefore, the posterior chains produced directly by our sampler can be considered *unweighted*. Given our robust measurements of  $\rho_*$  from our stellar characterization, we can apply an "importance weight" (Oh & Berger 1993;

Gilks et al. 1995) to these posterior samples according to how well the *unweighted* results can be described by the  $\rho_\star$  values measured with `isoclassify`. We will refer to these precise stellar density measurements as  $\rho_{\star,\text{iso}}$ . The importance weights also allow us to produce properly weighted posterior distributions for  $e$  and  $\omega$  without needing to directly sample them during the lightcurve modeling – a step that takes seconds rather than hours. This method has been used extensively in past literature for similar applications of improving modeling efficiency (see, e.g., Ford 2005; Ford 2006) or measuring eccentricity (see, e.g., Dawson & Johnson 2012; Van Eylen et al. 2019).

To determine the appropriate importance weights, we first calculate the estimated stellar density at each sampler step based on the sampled quantities  $P$ ,  $R_p/R_\star$ ,  $b$ , and  $T_{14}$ . We refer to this derived density as  $\rho_{\star,\text{derived}}$ . We calculate these values according to Equation 6.1, which is a rearrangement of the transit duration equation described by Winn (2010b):

$$\rho_{\star,\text{derived}} = \frac{3\pi}{GP^2} \left( \frac{(1 + R_p/R_\star)^2 - b^2}{\sin^2 \left( \frac{T_{14}\pi}{P} \frac{1+e\sin\omega}{\sqrt{1-e^2}} \right)} + b^2 \right)^{3/2}. \quad (6.1)$$

This equation, however, also includes  $e$  and  $\omega$ , for which we do not yet have any information. Here, we draw random samples of  $\{e, \omega\}$  from uniform prior distributions  $e \sim U(0, 1)$  and  $\omega \sim U(-\frac{\pi}{2}, \frac{3\pi}{2})$  to be used in the calculation. We prefer to use uninformative uniform priors on these parameters to remain agnostic to eccentricity here but will explore the parameterization of this eccentricity prior more in depth in MacDougall et al. 2023 (in prep). The stellar density values that we calculate here produce a derived posterior distribution based on our *unweighted* posterior samples and uniform  $\{e, \omega\}$  samples.

We then compare the samples of  $\rho_{\star,\text{derived}}$  against the independently measured  $\rho_{\star,\text{iso}}$  for a given target by computing the log-likelihood of each  $i^{\text{th}}$  sample,

$$\log \mathcal{L}_i = -\frac{1}{2} \left( \frac{\rho_{\star,\text{derived},i} - \rho_{\star,\text{iso}}}{\sigma(\rho_{\star,\text{iso}})} \right)^2, \quad (6.2)$$

assuming a Gaussian likelihood function. We then weight each sample from our original *unweighted* posterior distributions by

$$w_i = \frac{\mathcal{L}_i}{\sum_i \mathcal{L}_i} \quad (6.3)$$

to produce the final importance-weighted posterior distributions for each parameter. We also apply these same weights to the random, uniform  $\{e, \omega\}$  samples to derive their weighted posterior distributions. All summary statistics that are reported or used throughout the remaining analysis are based on the 15<sup>th</sup>, 50<sup>th</sup>, and 85<sup>th</sup> percentiles of these importance-weighted posteriors.

### 6.5.2 Planet-to-star Radius Ratio

Before deriving final planet radii, we first confirm that our final modeled  $R_p/R_\star$  values are unbiased and generally consistent with those reported by the TOI Catalog (Figure 6.5). We find no evidence of systematic bias in our results, measuring the average ratio between the TKS results and TOI Catalog values  $(R_p/R_\star)_{\text{TKS}}/(R_p/R_\star)_{\text{TOI}} = 1.02 \pm 0.11$ . This is consistent within the average fractional uncertainty of  $R_p/R_\star$  from the TOI Catalog. We show the residuals of this comparison,  $((R_p/R_\star)_{\text{TKS}} - (R_p/R_\star)_{\text{TOI}})/\sigma(R_p/R_\star)_{\text{TOI}}$ , in Figure 6.6b. Here we see that 4 out of 108 TKS planets have  $(R_p/R_\star)_{\text{TKS}}$  measurements that are  $\gtrsim 3\sigma_{\text{TOI}}$  discrepant from  $(R_p/R_\star)_{\text{TOI}}$ .

Upon closer inspection, we find that these discrepancies occur in instances where our careful de-trending and TTV modeling led to more reliable  $R_p/R_\star$  measurements. An example of this can be seen in Figure 6.4, where the initial PDC-SAP photometry of TOI-1136 displays significant variability which we remove before performing our final transit fits that include TTV modeling. Our measured  $R_p/R_\star$  values are consistent within  $\sim 1.5\sigma$  of the published values reported by Dai et al. 2022 while the  $R_p/R_\star$  values reported by the TOI Catalog are far more discrepant.

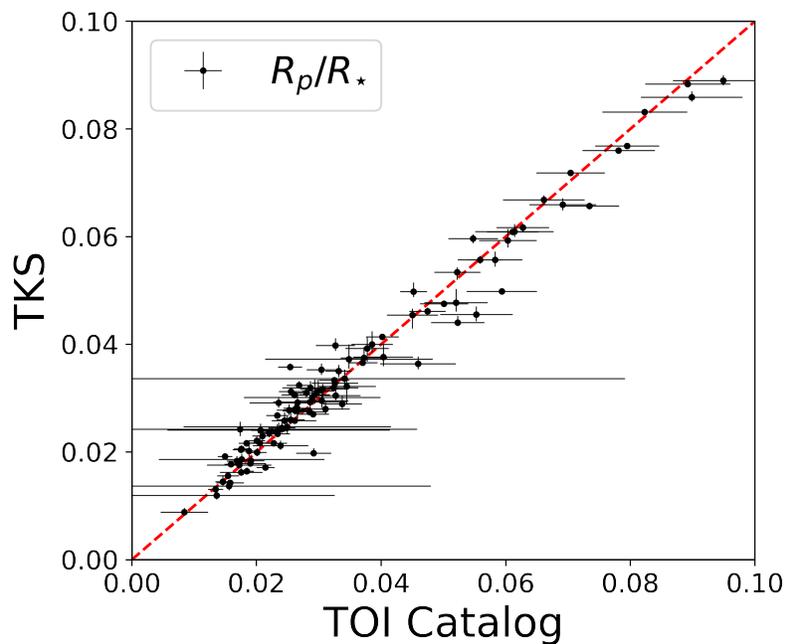


Figure 6.5: Planet-to-star radius ratio  $R_p/R_*$  as measured by our planet characterization pipeline (TKS; y-axis) compared to the radius ratio values reported by the TOI Catalog (x-axis). The TKS results have significantly reduced uncertainties and are generally consistent with the TOI Catalog results (see §6.5.2).

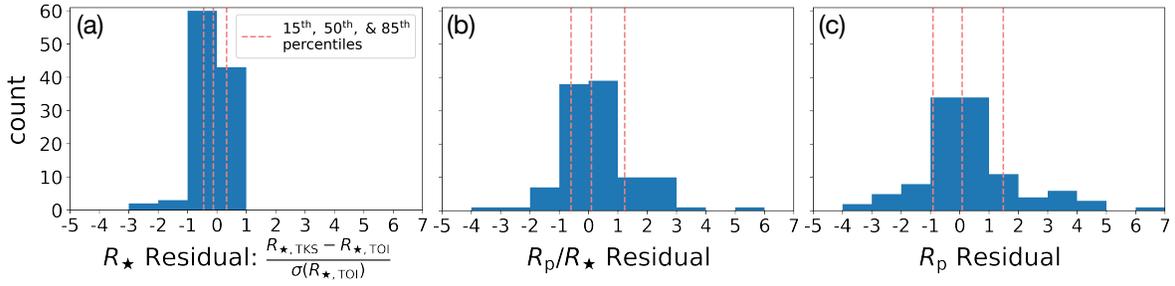


Figure 6.6: Residuals of (a) stellar radius  $R_*$ , (b) planet-to-star radius ratio  $R_p/R_*$ , and (c) planet radius  $R_p$  between our results and the TOI Catalog measurements (i.e. residual =  $(R_{*,\text{TKS}} - R_{*,\text{TOI}})/\sigma(R_{*,\text{TOI}})$ ). Dashed vertical lines show the 15<sup>th</sup>, 50<sup>th</sup>, and 85<sup>th</sup> percentiles of each distribution of residuals.

### 6.5.3 Derived Planet Properties

We calculate the radius of each planet in our sample ( $R_p$ ) from our final weighted posterior distributions of  $R_p/R_*$  and our isochrone-modeled  $R_*$  values. When compared against the TOI Catalog measurements, we find that our planet radii have a lower median uncertainty, reduced from  $\sigma(R_p)/R_p \approx 6.7\%$  for TOI Catalog values to 3.8% for TKS results. When looking at the mean uncertainty rather than the median, the difference is even more significant due to large outliers among the SPOC pipeline outputs, with mean uncertainties reduced from  $\sigma(R_p)/R_p \approx 16.1\%$  to 3.9%.

The planet radius distribution that we measure for the TKS sample is shown in Figure 6.7, displayed against the distribution of TKS planet radii reported by the TOI Catalog. We show the residuals of this comparison,  $(R_{p,\text{TKS}} - R_{p,\text{TOI}})/\sigma(R_{p,\text{TOI}})$ , in Figure 6.6c. We observe similar residual distributions for both  $R_p/R_*$  and  $R_p$  but far less significant scatter in the  $R_*$  residuals (Figure 6.6a), suggesting that the uncertainties on  $R_p$  are dominated by differences in lightcurve modeling rather than stellar characterization (Petigura 2020).

Along with planet radii, we also compute semi-major axis  $a$ , incident stellar flux  $S_{\text{inc}}$ , and equilibrium temperature  $T_{\text{eq}}$  for each planet. We calculate these parameters assuming con-

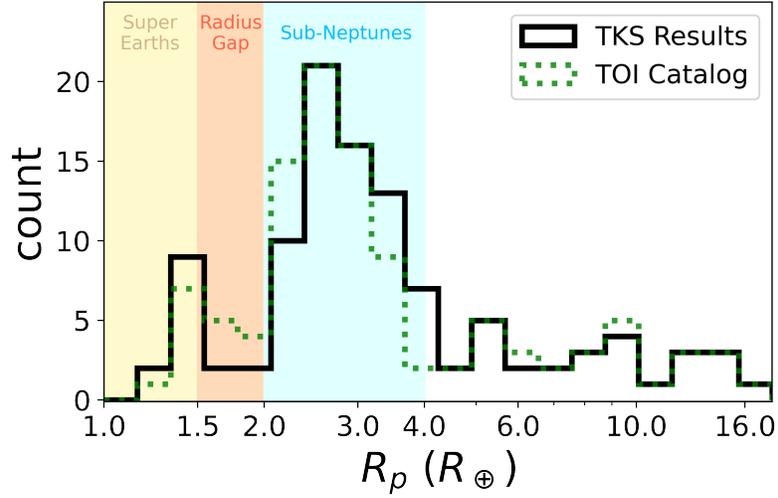


Figure 6.7: Distribution of median planet radius measurements for the TKS sample from this work, compared against the median radii taken from the TOI Catalog (green). Shaded regions show specific planet categories, classified by planetary radius. We observe a sparser radius gap region using TKS results than is implied by radii from the TOI Catalog, likely owing to our higher precision measurements.

stant planet-to-star separation, which is not a valid approximation for eccentric orbits so these values should only be considered as estimates in such cases. We compute semi-major axes from Kepler’s Third Law, using precise values of  $P$  and our well-constrained measurements of  $M_{\star}$ . The median uncertainty among these measurements is  $\sigma(a)/a \approx 1\%$ . We also calculate  $S_{\text{inc}}$  and  $T_{\text{eq}}$  using the standard equations described in [Johnson et al. \(2017\)](#), assuming a Bond albedo of  $\alpha = 0.3$  which is typical of a super-Earth-size *Kepler* planet ([Demory 2014](#)). This assumption is not uniformly applicable across our sample, but it is a common simplification among similar studies such as the California-*Kepler* Survey ([Johnson et al. 2017](#)). The median uncertainties that we measure for  $S_{\text{inc}}$  and  $T_{\text{eq}}$  are 6.7% and 1.7%, respectively, without accounting for the uncertainty in  $\alpha$ .

## 6.6 Discussion

### 6.6.1 Radius Gap

We note from [Chontos et al. \(2022\)](#) that a primary goal of TKS Science Case 1A is to better understand the bimodal distribution of small planet radii first described by [Fulton et al. \(2017a\)](#). This radius valley has been studied extensively from both *Kepler* and *K2* planets (see, e.g., [Berger et al. 2018b](#); [Cloutier & Menou 2020](#)) which have improved our understanding of the physical mechanisms leading to this phenomenon ([Van Eylen et al. 2018](#); [Gupta & Schlichting 2019](#)). For the first time among TESS planets, we observe a valley-like structure between 1.5–2.0  $R_{\oplus}$  among the TKS sample. In [Figure 6.7](#), we show the distribution of TKS planet radii and highlight the radius valley, which we find is insensitive to differences in binning. We also find this structure to be more significant among our precisely measured TKS radii from this work as compared to the TOI Catalog radii. We attribute this difference primarily to our improved constraints on  $R_p/R_{\star}$  from our lightcurve detrending and modeling procedure (see, e.g., [Petigura 2020](#)).

The radius gap has also been found to have some reliance on various other system properties such as orbital period, incident flux ([Cloutier & Menou 2020](#)), and stellar parameters as well ([Fulton & Petigura 2018](#); [Berger et al. 2020c](#)). This was recently explored by [Petigura et al. 2022](#) for the sample of 1,246 planets in the California-*Kepler* Survey (CKS; [Petigura et al. 2017a](#)), which was the *Kepler*-focused predecessor to the TESS-Keck Survey. Although the TKS planet sample is only 8.4% the size of the CKS sample, our results do suggest the existence of a radius valley in various 2D representations of radius as a function of other parameters, including  $P$ ,  $S_{\text{inc}}$ ,  $M_{\star}$ , and  $[\text{Fe}/\text{H}]$  ([Figure 6.8](#)). The orange shaded regions in these figures show the radius valley as measured by [Petigura et al. \(2022\)](#) in the various parameter spaces, which generally agree with our results despite differences between the CKS and TKS samples. Future investigations of the radius valley may combine these samples to gain further insight to the mechanisms driving the bi-modality in the planet radius distribution.

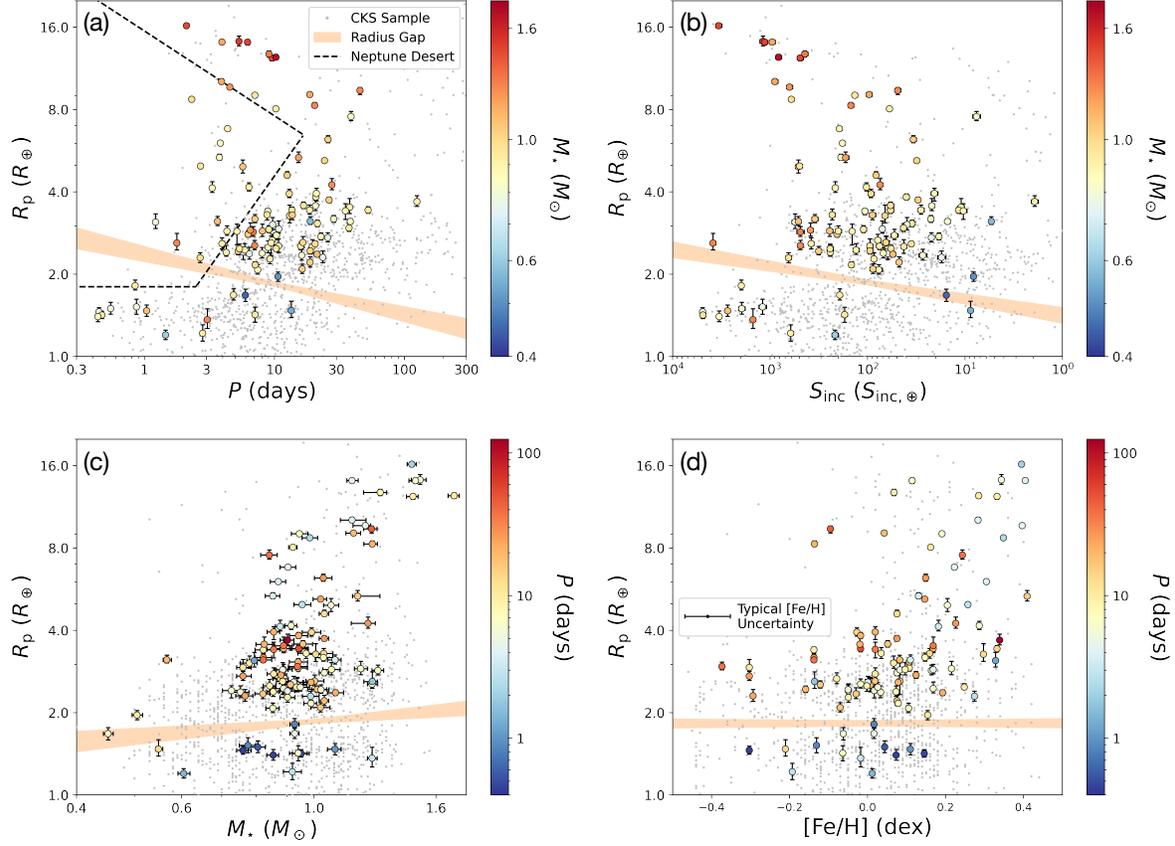


Figure 6.8: Panel (a): radius and orbital period for all TKS planets, with stellar mass shown as a color scale. Grey background points show the distribution of planets from the CKS sample (Petigura et al. 2022). The radius gap, as defined by Petigura et al. (2022), is shown as the orange shaded region. The Hot Neptune desert, as defined by Mazeh et al. (2016), is shown as the region marked off by black dashed lines. Panel (b): same as (a), except x-axis is  $S_{\text{inc}}$ . Panel (c): same as (a), except x-axis is  $M_*$  and color scale is  $P$ . Panel (d): same as (c), except x-axis is  $[\text{Fe}/\text{H}]$ .

### 6.6.2 Hot Neptune Desert

Similar to the radius gap, the Hot Neptune Desert has been proposed as a region of  $R_p - P$  parameter space that is sparsely populated, with few Neptune-size planets at  $P \lesssim 5$  d. The mechanism that would drive such an effect is currently unknown, but this sparsity has been proposed as a natural consequence of photo-evaporation (Lopez & Fortney 2013; Owen & Wu 2013). The Hot Neptune Desert region can be seen in Figure 6.8a, outlined by boundaries defined by Mazeh et al. (2016). We find that the TKS sample includes at least 12 planets which fall firmly within this region (see, e.g., MacDougall et al. 2022; Van Zandt et al. 2022), supplying future investigations with an increased sample size to better understand the nature of the Hot Neptune Desert. Further studies will be carried out by the TKS team to determine the compositions (TKS Science Case 1B) and atmospheric properties (TKS Science Case 3) of such TKS planets to probe this topic further with planet masses retrieved from ongoing radial velocity measurements.

### 6.6.3 Multis and TTVs

The TKS sample was selected to contain several systems with multiple transiting planet signals, fulfilling TKS Science Case 2C. Among our 85 target systems, we report 69 with a single transiting planet signal, 11 with two transit signals, 3 with three transit signals (TOI-561, Weiss et al. 2021; TOI-1339, Lubin et al. 2022; TOI-2076, Osborn et al. 2022), 1 with four transit signals (TOI-1246, Turtelboom et al. 2022), and 1 with more than four transit signals (TOI-1136). For TOI-1136, we only report on 4 transiting planets in this work, but ongoing investigations by TKS collaborators are working towards revealing the true number of transiting planets in this system (see Dai et al. 2022). We also measure orbital periods in some multi-planet systems that differ from those reported by the SPOC pipeline. These include TOI-561.03 (see Lacedelli et al. 2022), TOI-266.02 (Akana Murphy et al. 2023, in

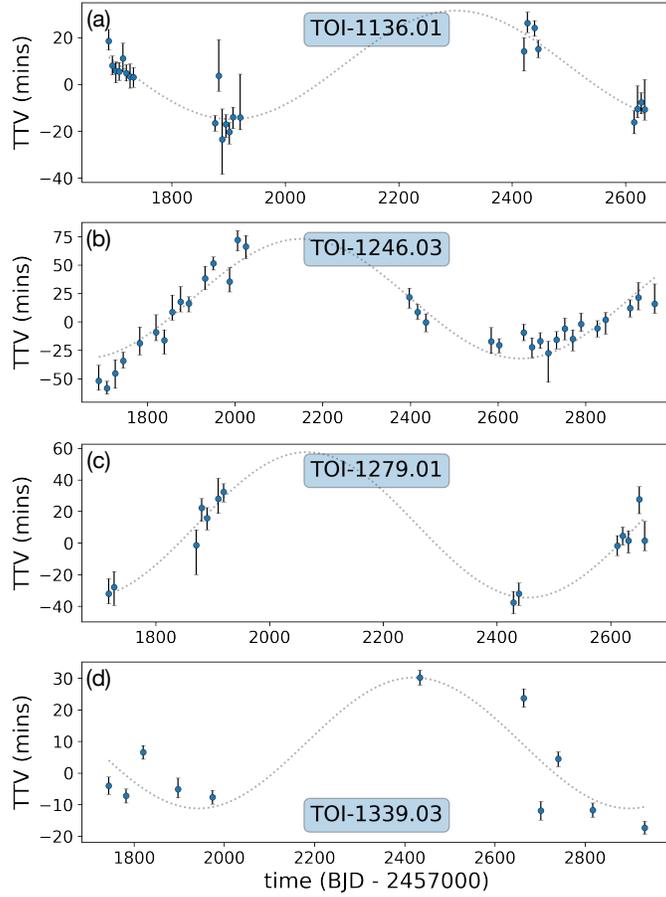


Figure 6.9: Four examples of significant transit-timing variations detected among TKS systems, including three from multi-planet systems (panels a, b, and d) and one from a single-planet system (panel c). (Panel a) TOI-1136.01: TTV semi-amplitude  $\approx 24$  minutes, TTV signal period  $\approx 790$  days; (b) TOI-1246.03: TTV semi-amplitude  $\approx 61$  minutes, TTV signal period  $\approx 990$  days; (c) TOI-1279.01: TTV semi-amplitude  $\approx 34$  minutes, TTV signal period  $\approx 775$  days; (d) TOI-1339.03: TTV semi-amplitude  $\approx 22$  minutes, TTV signal period  $\approx 960$  days, with significant scatter likely attributed to TTV chopping. These TTVs are measured from the observed transit midpoints (fit via `exoplanet`) minus the transit midpoints calculated assuming a linear ephemeris. An estimated sinusoidal signal is shown fit to each set of TTVs, using a Lomb-Scargle periodogram and regression fit from `astropy`. These significant TTV signals likely arise from dynamical interactions between planetary companions.

prep<sup>1</sup>), and TOI-1471.02 (Osborn et al. 2023, in prep). These findings help to better resolve the architectures of several multi-planet systems and are discussed in existing or upcoming literature which verify by our analysis.

We also note that several of the multi-planet systems in our sample display resonant or near-resonant orbital configurations, such as the 1:2 orbital period ratios among the inner planets in the TOI-1136 system (see Dai et al. 2022) as well as the outer planets in the TOI-1246 system (Turtelboom et al. 2022). Such orbital configurations can give rise to significant transit-timing variations, which we observe for TOI-1136.01 and TOI-1246.03 with TTV semi-amplitudes of  $\sim 24$  and  $\sim 61$  minutes, respectively (Figure 6.9a–b). The resonant counterparts to these two planets, TOI-1136.02 and TOI-1246.04, also display significant TTVs with similar semi-amplitudes as their companions but with possible anti-correlation and less distinguishable periodicity. We assert that our measured TTVs for a given planet are significant when  $>32\%$  of the measured transit times differ from a linear ephemeris by  $>1\sigma$ .

The near-resonant planets in the TOI-2076 system also display possibly significant TTVs, as first noted by Osborn et al. (2022), but we do not identify a clear structure to this signal. Additionally, we observe possible TTV chopping for TOI-1339.03 (TTV semi-amplitude  $\approx 22$  minutes; Figure 6.9d) which orbits near the 3:4 mean-motion resonance with TOI-1339.02 (TTV semi-amplitude  $\approx 10$  minutes), as suggested by Badenas-Agusti et al. (2020). The semi-amplitudes that we report here are only an approximation, estimated as half the difference between the maximum and minimum TTV value for a given planet.

It is also possible for transit-timing variations to arise in systems without a known transiting companion, suggesting the presence of an additional non-transiting planet. Although most single-planet orbits among the TKS sample can be described by a linear ephemeris (see

---

<sup>1</sup>Period ambiguity resolved via private communications with the CHEOPS team (point of contact: Hugh Osborn)

§6.4.4), we identify at least one planet that displays significant TTVs without a known companion: TOI-1279.01 with a TTV semi-amplitude of  $\sim 34$  minutes (Figure 6.9c). At least two other single-planet systems among the TKS sample also display possible TTVs, TOI-1611.01 and TOI-1742.01, but these signals are less significant relative to the other TTV systems described above. The dynamics and multiplicities of the TKS systems discussed in this section will be examined further in future TKS studies of multi-planet systems (TKS Science Case 2C) and distant giant companions (TKS Science Case 2A).

#### 6.6.4 Eccentricity

Through importance sampling (§6.5.1), we measure posterior distributions for  $e$  and  $\omega$  for all planets in the TKS sample using stellar density and the results of our photometric modeling. The goal of TKS Science Case 2Bi is to characterize the eccentricities of sub-Jovian planets and better measure their underlying eccentricity distribution. Generally, however, photometric modeling alone is known to produce imprecise measurements of  $\{e, \omega\}$  on a planet-by-planet basis. Since  $\omega$  is typically unconstrained from transits and  $b$  is often loosely constrained for small planets due to ambiguous ingress / egress signals, strong covariances between  $b$ ,  $e$ , and  $\omega$  drive large uncertainties in photometrically derived eccentricities. We observe this with sub-Neptune TOI-1255.01 from the TKS sample (MacDougall et al. 2021), for which the photometric eccentricity  $e_{\text{phot}}$  was highly degenerate with  $b$  until additional eccentricity constraints were introduced from radial velocity measurements to determine  $e \approx 0.16$ . On the other hand, we also identified the Neptune-size TKS target TOI-1272.01 (MacDougall et al. 2022) as moderately eccentric based on  $e_{\text{phot}}$  and later confirmed a self-consistent eccentricity of  $e \approx 0.34$  via joint modeling with radial velocities.

Rather than using our photometrically-derived eccentricities to draw conclusions about planets on an individual level, we are conducting an ongoing study to determine the population-level eccentricity distribution underlying the TKS sample (MacDougall et al. 2023, in prep). In this way, we are less sensitive to parameter covariances like  $b$ - $e$ - $\omega$  and other effects like

the Lucy-Sweeney bias which leads to preferentially non-circular orbit fits (Lucy & Sweeney 1971; Eastman et al. 2013b). An investigation is also underway by members of TKS to measure more precise eccentricities for planets in our sample using radial velocity measurements (Polanski et al. 2023, in prep), which will aid in breaking the degeneracy between  $e$  and  $b$  for lower signal-to-noise planets.

## 6.7 Conclusions

In this work, we characterized 85 TESS target stars presented by Chontos et al. (2022) as the TESS-Keck Survey sample. Through a combination of photometry, high-resolution spectroscopy, and *Gaia* parallaxes, we measured stellar properties for these targets with improved precision relative to past works. We used these stellar characterizations to facilitate in homogeneously modeling the lightcurve photometry of each TKS target, including fitting for photometric variability, estimating transit times, and measuring the transit properties of planet candidates recovered from the reduced lightcurve photometry. We characterized a total of 108 transiting planet candidates orbiting the 85 TKS target stars - the largest sample of homogeneously characterized planets from the TESS mission to date.

In measuring the radii of these planets, we found a substantial increase in precision amongst our radii relative to the TOI Catalog ( $\sigma(R_{p, \text{TOI}})/R_{p, \text{TOI}} \approx 6.7\%$  versus  $\sigma(R_{p, \text{TKS}})/R_{p, \text{TKS}} \approx 3.8\%$ ), largely attributed to our improved accuracy in measuring  $R_p/R_\star$ . From our updated radii, we successfully recovered the radius gap in both one-dimensional ( $R_p$ ) and two-dimensional ( $R_p - P$ ,  $R_p - S_{\text{inc}}$ ,  $R_p - M_\star$ , and  $R_p - [\text{Fe}/\text{H}]$ ) parameter spaces. While we do not provide new models to fit the topology of the radius gap in these parameter spaces, we find that the distribution of radii in the radius gap region is broadly consistent with that of *Kepler* planets. We also identify significant transit-timing variations among several TKS systems, including multi-planet systems TOI-1136, TOI-1246, and TOI-1339 as well as the single-planet system TOI-1279. These TTVs will inform future investigations into the

dynamics and multiplicity of TKS systems.

The TKS Collaboration will continue use the planetary and stellar characterizations from this work to answer key questions that remain in the field of exoplanet science, covering the topics of bulk planet compositions, system architectures and dynamics, planet atmospheres, and the effects of stellar evolution. Our photometric analysis will soon be accompanied by a radial velocity analysis of all TKS targets as well (Polanski et al. 2023, in prep), which will provide us with new planetary system characteristics such as planet masses, precise eccentricities, and discoveries of non-transiting planets. The combination of insights from the photometry described here and upcoming radial velocity measurements will allow us to more fully answer the questions laid out by the TKS Science Cases.

The stellar and planetary properties measured from this work, along with additional supporting materials, will be made available in machine-readable format on GitHub<sup>2</sup>. The TESS-Keck Survey sample adds to the legacy of in-depth surveys of host stars and their transiting exoplanets, such as the *California-Kepler* Survey, which will continue to serve the community as a source of precise, homogeneously derived planetary system properties.

---

<sup>2</sup>[https://github.com/mason-macdougall/tks\\_system\\_properties.git](https://github.com/mason-macdougall/tks_system_properties.git)

Table 6.3: TKS Stellar Properties

TOI	$m_K$	$\pi$	$T_{\text{eff}}$	[Fe/H]	$R_*$	$M_*$	$\rho_*$	age	SpecMatch
	mag	mas	K	dex	$R_\odot$	$M_\odot$	$g\text{ cm}^{-3}$	Gyr	
260	6.55	49.51	4050 <sup>+35</sup> <sub>-30</sub>	-0.21 <sup>+0.06</sup> <sub>-0.09</sub>	0.55 <sup>+0.01</sup> <sub>-0.01</sub>	0.55 <sup>+0.01</sup> <sub>-0.01</sub>	4.7 <sup>+0.14</sup> <sub>-0.2</sub>	15.9 <sup>+2.8</sup> <sub>-5.5</sub>	emp
266	8.45	9.8	5617 <sup>+74</sup> <sub>-73</sub>	0.01 <sup>+0.06</sup> <sub>-0.06</sub>	0.96 <sup>+0.02</sup> <sub>-0.02</sub>	0.94 <sup>+0.03</sup> <sub>-0.03</sub>	1.48 <sup>+0.11</sup> <sub>-0.1</sub>	6.0 <sup>+2.4</sup> <sub>-2.3</sub>	syn
329	9.68	3.49	5613 <sup>+78</sup> <sub>-73</sub>	0.2 <sup>+0.06</sup> <sub>-0.06</sub>	1.53 <sup>+0.03</sup> <sub>-0.03</sub>	1.07 <sup>+0.03</sup> <sub>-0.03</sub>	0.41 <sup>+0.03</sup> <sub>-0.03</sub>	8.7 <sup>+1.0</sup> <sub>-1.0</sub>	syn
465	9.34	8.16	4986 <sup>+57</sup> <sub>-56</sub>	0.3 <sup>+0.06</sup> <sub>-0.06</sub>	0.83 <sup>+0.01</sup> <sub>-0.01</sub>	0.87 <sup>+0.02</sup> <sub>-0.03</sub>	2.15 <sup>+0.15</sup> <sub>-0.14</sub>	5.6 <sup>+3.9</sup> <sub>-3.2</sub>	syn
469	7.59	14.64	5322 <sup>+72</sup> <sub>-70</sub>	0.33 <sup>+0.06</sup> <sub>-0.06</sub>	0.98 <sup>+0.02</sup> <sub>-0.02</sub>	0.95 <sup>+0.03</sup> <sub>-0.03</sub>	1.39 <sup>+0.1</sup> <sub>-0.1</sub>	8.6 <sup>+2.8</sup> <sub>-2.7</sub>	syn
480	6.01	18.31	6173 <sup>+89</sup> <sub>-118</sub>	0.16 <sup>+0.08</sup> <sub>-0.07</sub>	1.49 <sup>+0.04</sup> <sub>-0.03</sub>	1.28 <sup>+0.03</sup> <sub>-0.03</sub>	0.54 <sup>+0.04</sup> <sub>-0.05</sub>	2.7 <sup>+0.6</sup> <sub>-0.5</sub>	syn
509	6.88	20.39	5512 <sup>+71</sup> <sub>-72</sub>	0.11 <sup>+0.05</sup> <sub>-0.06</sub>	0.96 <sup>+0.02</sup> <sub>-0.02</sub>	0.94 <sup>+0.03</sup> <sub>-0.03</sub>	1.49 <sup>+0.11</sup> <sub>-0.1</sub>	6.6 <sup>+2.5</sup> <sub>-2.4</sub>	syn
554	5.71	21.89	6352 <sup>+84</sup> <sub>-76</sub>	-0.02 <sup>+0.05</sup> <sub>-0.06</sub>	1.42 <sup>+0.02</sup> <sub>-0.02</sub>	1.25 <sup>+0.02</sup> <sub>-0.02</sub>	0.61 <sup>+0.04</sup> <sub>-0.03</sub>	2.3 <sup>+0.4</sup> <sub>-0.4</sub>	syn
561	8.39	11.63	5342 <sup>+57</sup> <sub>-54</sub>	-0.3 <sup>+0.06</sup> <sub>-0.06</sub>	0.86 <sup>+0.01</sup> <sub>-0.01</sub>	0.76 <sup>+0.02</sup> <sub>-0.02</sub>	1.69 <sup>+0.11</sup> <sub>-0.09</sub>	16.1 <sup>+2.4</sup> <sub>-3.1</sub>	syn
669	9.13	6.97	5597 <sup>+71</sup> <sub>-72</sub>	0.0 <sup>+0.06</sup> <sub>-0.06</sub>	0.99 <sup>+0.02</sup> <sub>-0.02</sub>	0.92 <sup>+0.03</sup> <sub>-0.03</sub>	1.34 <sup>+0.1</sup> <sub>-0.09</sub>	7.9 <sup>+2.3</sup> <sub>-2.2</sub>	syn
1136	8.03	11.8	5805 <sup>+57</sup> <sub>-57</sub>	0.08 <sup>+0.05</sup> <sub>-0.06</sub>	0.96 <sup>+0.01</sup> <sub>-0.01</sub>	1.04 <sup>+0.02</sup> <sub>-0.02</sub>	1.66 <sup>+0.06</sup> <sub>-0.07</sub>	0.7 <sup>+0.9</sup> <sub>-0.4</sub>	syn
1173	9.13	7.53	5414 <sup>+62</sup> <sub>-65</sub>	0.19 <sup>+0.06</sup> <sub>-0.06</sub>	0.94 <sup>+0.02</sup> <sub>-0.01</sub>	0.95 <sup>+0.03</sup> <sub>-0.03</sub>	1.61 <sup>+0.11</sup> <sub>-0.1</sub>	5.9 <sup>+2.4</sup> <sub>-2.4</sub>	syn
1174	8.96	10.54	5157 <sup>+59</sup> <sub>-57</sub>	-0.0 <sup>+0.06</sup> <sub>-0.06</sub>	0.78 <sup>+0.03</sup> <sub>-0.02</sub>	0.84 <sup>+0.02</sup> <sub>-0.02</sub>	2.44 <sup>+0.18</sup> <sub>-0.25</sub>	2.7 <sup>+3.6</sup> <sub>-1.9</sub>	syn
1180	8.59	13.86	4803 <sup>+53</sup> <sub>-52</sub>	0.07 <sup>+0.08</sup> <sub>-0.09</sub>	0.74 <sup>+0.02</sup> <sub>-0.02</sub>	0.78 <sup>+0.02</sup> <sub>-0.03</sub>	2.66 <sup>+0.22</sup> <sub>-0.28</sub>	5.0 <sup>+6.3</sup> <sub>-3.6</sub>	emp
1181	9.22	3.27	6045 <sup>+88</sup> <sub>-84</sub>	0.39 <sup>+0.05</sup> <sub>-0.06</sub>	1.93 <sup>+0.04</sup> <sub>-0.04</sub>	1.46 <sup>+0.02</sup> <sub>-0.03</sub>	0.28 <sup>+0.02</sup> <sub>-0.02</sub>	2.4 <sup>+0.3</sup> <sub>-0.3</sub>	syn
1184	8.32	17.04	4616 <sup>+47</sup> <sub>-48</sub>	0.02 <sup>+0.08</sup> <sub>-0.09</sub>	0.7 <sup>+0.02</sup> <sub>-0.02</sub>	0.73 <sup>+0.02</sup> <sub>-0.03</sub>	3.0 <sup>+0.24</sup> <sub>-0.27</sub>	5.3 <sup>+6.6</sup> <sub>-3.9</sub>	emp
1194	9.34	6.65	5393 <sup>+64</sup> <sub>-62</sub>	0.35 <sup>+0.06</sup> <sub>-0.06</sub>	0.96 <sup>+0.02</sup> <sub>-0.01</sub>	0.98 <sup>+0.03</sup> <sub>-0.03</sub>	1.55 <sup>+0.1</sup> <sub>-0.1</sub>	4.9 <sup>+2.3</sup> <sub>-2.1</sub>	syn
1244	9.42	9.72	4721 <sup>+55</sup> <sub>-54</sub>	0.03 <sup>+0.09</sup> <sub>-0.09</sub>	0.72 <sup>+0.02</sup> <sub>-0.02</sub>	0.75 <sup>+0.02</sup> <sub>-0.03</sub>	2.78 <sup>+0.24</sup> <sub>-0.27</sub>	5.9 <sup>+6.8</sup> <sub>-4.2</sub>	emp
1246	9.91	5.87	5213 <sup>+70</sup> <sub>-67</sub>	0.17 <sup>+0.06</sup> <sub>-0.06</sub>	0.86 <sup>+0.02</sup> <sub>-0.02</sub>	0.89 <sup>+0.03</sup> <sub>-0.03</sub>	1.92 <sup>+0.16</sup> <sub>-0.15</sub>	6.3 <sup>+3.6</sup> <sub>-3.2</sub>	syn
1247	7.5	13.51	5697 <sup>+75</sup> <sub>-72</sub>	-0.12 <sup>+0.06</sup> <sub>-0.06</sub>	1.07 <sup>+0.02</sup> <sub>-0.02</sub>	0.91 <sup>+0.03</sup> <sub>-0.03</sub>	1.03 <sup>+0.07</sup> <sub>-0.07</sub>	9.8 <sup>+2.0</sup> <sub>-2.0</sub>	syn
1248	9.87	5.9	5205 <sup>+56</sup> <sub>-61</sub>	0.22 <sup>+0.06</sup> <sub>-0.06</sub>	0.87 <sup>+0.01</sup> <sub>-0.01</sub>	0.9 <sup>+0.03</sup> <sub>-0.03</sub>	1.93 <sup>+0.13</sup> <sub>-0.12</sub>	5.5 <sup>+3.0</sup> <sub>-2.7</sub>	syn
1249	9.13	7.14	5496 <sup>+68</sup> <sub>-66</sub>	0.3 <sup>+0.06</sup> <sub>-0.06</sub>	0.98 <sup>+0.02</sup> <sub>-0.02</sub>	1.01 <sup>+0.03</sup> <sub>-0.03</sub>	1.5 <sup>+0.11</sup> <sub>-0.1</sub>	4.0 <sup>+2.3</sup> <sub>-2.0</sub>	syn
1255	7.92	15.13	5216 <sup>+52</sup> <sub>-52</sub>	0.25 <sup>+0.08</sup> <sub>-0.05</sub>	0.84 <sup>+0.01</sup> <sub>-0.01</sub>	0.93 <sup>+0.02</sup> <sub>-0.02</sub>	2.19 <sup>+0.08</sup> <sub>-0.09</sub>	1.2 <sup>+1.5</sup> <sub>-0.8</sub>	syn
1269	9.89	5.78	5499 <sup>+63</sup> <sub>-60</sub>	-0.05 <sup>+0.06</sup> <sub>-0.06</sub>	0.85 <sup>+0.01</sup> <sub>-0.01</sub>	0.9 <sup>+0.02</sup> <sub>-0.03</sub>	2.04 <sup>+0.12</sup> <sub>-0.12</sub>	3.1 <sup>+2.3</sup> <sub>-1.8</sub>	syn
1272	9.7	7.24	5065 <sup>+52</sup> <sub>-50</sub>	0.18 <sup>+0.05</sup> <sub>-0.06</sub>	0.79 <sup>+0.01</sup> <sub>-0.01</sub>	0.88 <sup>+0.01</sup> <sub>-0.02</sub>	2.46 <sup>+0.08</sup> <sub>-0.09</sub>	1.1 <sup>+1.6</sup> <sub>-0.8</sub>	syn

Table 6.3 (cont'd): TKS Stellar Properties

TOI	$m_K$ mag	$\pi$ mas	$T_{\text{eff}}$ K	[Fe/H] dex	$R_*$ $R_\odot$	$M_*$ $M_\odot$	$\rho_*$ $g\text{ cm}^{-3}$	age Gyr	SpecMatch
1279	8.89	9.31	$5457^{+66}_{-67}$	$-0.08^{+0.06}_{-0.06}$	$0.84^{+0.01}_{-0.01}$	$0.88^{+0.03}_{-0.03}$	$2.05^{+0.13}_{-0.13}$	$4.2^{+2.5}_{-2.2}$	syn
1288	8.78	8.68	$5388^{+68}_{-66}$	$0.26^{+0.06}_{-0.06}$	$0.96^{+0.02}_{-0.02}$	$0.96^{+0.03}_{-0.03}$	$1.52^{+0.11}_{-0.1}$	$6.5^{+2.6}_{-2.5}$	syn
1294	9.96	2.98	$5718^{+58}_{-77}$	$0.28^{+0.06}_{-0.06}$	$1.56^{+0.03}_{-0.03}$	$1.16^{+0.06}_{-0.05}$	$0.43^{+0.03}_{-0.03}$	$6.2^{+1.4}_{-1.4}$	syn
1296	9.74	3.06	$5567^{+77}_{-72}$	$0.4^{+0.06}_{-0.02}$	$1.7^{+0.03}_{-0.03}$	$1.16^{+0.03}_{-0.02}$	$0.33^{+0.02}_{-0.02}$	$7.1^{+0.6}_{-0.7}$	syn
1298	10.01	3.11	$5752^{+68}_{-98}$	$0.4^{+0.06}_{-0.06}$	$1.45^{+0.03}_{-0.02}$	$1.22^{+0.02}_{-0.06}$	$0.55^{+0.03}_{-0.04}$	$4.3^{+1.7}_{-0.6}$	syn
1339	7.18	18.62	$5385^{+69}_{-64}$	$-0.14^{+0.06}_{-0.06}$	$0.93^{+0.02}_{-0.02}$	$0.82^{+0.03}_{-0.03}$	$1.41^{+0.1}_{-0.09}$	$13.9^{+2.8}_{-2.7}$	syn
1347	9.62	6.75	$5493^{+53}_{-55}$	$0.02^{+0.06}_{-0.05}$	$0.84^{+0.01}_{-0.01}$	$0.93^{+0.02}_{-0.02}$	$2.19^{+0.08}_{-0.08}$	$0.8^{+1.1}_{-0.6}$	syn
1386	9.09	6.78	$5734^{+75}_{-76}$	$0.15^{+0.06}_{-0.06}$	$1.02^{+0.02}_{-0.02}$	$1.04^{+0.03}_{-0.04}$	$1.35^{+0.09}_{-0.09}$	$3.2^{+1.9}_{-1.6}$	syn
1410	8.58	13.72	$4635^{+52}_{-51}$	$0.33^{+0.08}_{-0.09}$	$0.76^{+0.02}_{-0.02}$	$0.79^{+0.02}_{-0.02}$	$2.57^{+0.22}_{-0.28}$	$6.6^{+6.6}_{-4.6}$	emp
1411	7.25	30.76	$4115^{+41}_{-45}$	$0.01^{+0.07}_{-0.07}$	$0.61^{+0.02}_{-0.02}$	$0.6^{+0.01}_{-0.01}$	$3.63^{+0.23}_{-0.23}$	$9.6^{+6.6}_{-6.3}$	emp
1422	9.19	6.42	$5810^{+81}_{-80}$	$-0.03^{+0.06}_{-0.06}$	$1.02^{+0.02}_{-0.02}$	$0.99^{+0.04}_{-0.04}$	$1.29^{+0.09}_{-0.09}$	$4.6^{+2.0}_{-1.8}$	syn
1430	7.08	24.26	$5086^{+62}_{-59}$	$0.11^{+0.06}_{-0.06}$	$0.8^{+0.03}_{-0.02}$	$0.85^{+0.02}_{-0.02}$	$2.34^{+0.19}_{-0.26}$	$3.4^{+4.3}_{-2.4}$	syn
1436	9.72	7.57	$5029^{+58}_{-57}$	$-0.13^{+0.06}_{-0.06}$	$0.75^{+0.03}_{-0.02}$	$0.77^{+0.02}_{-0.03}$	$2.58^{+0.28}_{-0.32}$	$5.9^{+5.9}_{-4.0}$	syn
1437	7.82	9.65	$5985^{+82}_{-81}$	$-0.16^{+0.03}_{-0.05}$	$1.26^{+0.02}_{-0.02}$	$1.01^{+0.04}_{-0.04}$	$0.71^{+0.05}_{-0.05}$	$6.7^{+1.4}_{-1.3}$	syn
1438	9.09	9.02	$5237^{+55}_{-55}$	$0.08^{+0.05}_{-0.05}$	$0.82^{+0.01}_{-0.01}$	$0.89^{+0.02}_{-0.02}$	$2.29^{+0.1}_{-0.13}$	$2.2^{+2.2}_{-1.5}$	syn
1439	9.05	4.31	$5840^{+63}_{-64}$	$0.23^{+0.06}_{-0.07}$	$1.62^{+0.03}_{-0.03}$	$1.23^{+0.04}_{-0.08}$	$0.4^{+0.02}_{-0.03}$	$4.6^{+1.6}_{-0.7}$	syn
1443	8.67	11.63	$5223^{+58}_{-57}$	$-0.29^{+0.06}_{-0.06}$	$0.75^{+0.03}_{-0.03}$	$0.77^{+0.02}_{-0.03}$	$2.49^{+0.33}_{-0.35}$	$7.4^{+5.9}_{-4.8}$	syn
1444	9.06	7.94	$5460^{+66}_{-66}$	$0.15^{+0.06}_{-0.06}$	$0.91^{+0.02}_{-0.01}$	$0.95^{+0.03}_{-0.03}$	$1.76^{+0.12}_{-0.11}$	$3.8^{+2.5}_{-2.1}$	syn
1451	8.07	10.85	$5800^{+77}_{-78}$	$0.02^{+0.06}_{-0.06}$	$1.02^{+0.02}_{-0.02}$	$1.0^{+0.04}_{-0.03}$	$1.34^{+0.09}_{-0.09}$	$4.0^{+1.9}_{-1.7}$	syn
1456	7.24	12.37	$6127^{+87}_{-79}$	$0.04^{+0.06}_{-0.06}$	$1.27^{+0.02}_{-0.02}$	$1.16^{+0.03}_{-0.03}$	$0.8^{+0.05}_{-0.04}$	$2.9^{+0.9}_{-0.8}$	syn
1467	8.57	26.68	$3775^{+22}_{-20}$	$-0.06^{+0.04}_{-0.05}$	$0.46^{+0.01}_{-0.01}$	$0.45^{+0.01}_{-0.01}$	$6.65^{+0.22}_{-0.24}$	$13.4^{+4.4}_{-6.4}$	emp
1471	7.56	14.78	$5648^{+74}_{-70}$	$-0.02^{+0.06}_{-0.06}$	$0.96^{+0.02}_{-0.02}$	$0.94^{+0.03}_{-0.03}$	$1.49^{+0.11}_{-0.1}$	$5.6^{+2.3}_{-2.3}$	syn
1472	9.28	8.18	$5132^{+54}_{-53}$	$0.29^{+0.06}_{-0.06}$	$0.84^{+0.01}_{-0.01}$	$0.92^{+0.02}_{-0.02}$	$2.18^{+0.1}_{-0.13}$	$2.3^{+2.5}_{-1.5}$	syn
1473	7.39	14.76	$5934^{+74}_{-78}$	$-0.05^{+0.06}_{-0.05}$	$1.01^{+0.02}_{-0.02}$	$1.03^{+0.03}_{-0.04}$	$1.41^{+0.09}_{-0.09}$	$2.4^{+1.6}_{-1.3}$	syn
1601	9.19	2.94	$5982^{+78}_{-64}$	$0.34^{+0.05}_{-0.06}$	$2.2^{+0.06}_{-0.06}$	$1.51^{+0.03}_{-0.05}$	$0.2^{+0.02}_{-0.02}$	$2.4^{+0.4}_{-0.2}$	syn

Table 6.3 (cont'd): TKS Stellar Properties

TOI	$m_K$ mag	$\pi$ mas	$T_{\text{eff}}$ K	[Fe/H] dex	$R_*$ $R_\odot$	$M_*$ $M_\odot$	$\rho_*$ $g\text{ cm}^{-3}$	age Gyr	SpecMatch
1611	6.31	35.31	$5106^{+66}_{-63}$	$-0.03^{+0.06}_{-0.06}$	$0.78^{+0.03}_{-0.03}$	$0.82^{+0.02}_{-0.03}$	$2.38^{+0.25}_{-0.31}$	$4.7^{+5.1}_{-3.2}$	syn
1669	8.46	8.96	$5497^{+73}_{-72}$	$0.27^{+0.06}_{-0.06}$	$1.06^{+0.02}_{-0.02}$	$0.99^{+0.03}_{-0.03}$	$1.17^{+0.08}_{-0.08}$	$7.7^{+2.2}_{-2.1}$	syn
1691	8.54	8.94	$5641^{+80}_{-74}$	$0.05^{+0.06}_{-0.06}$	$1.01^{+0.02}_{-0.02}$	$0.96^{+0.04}_{-0.03}$	$1.3^{+0.1}_{-0.09}$	$6.6^{+2.3}_{-2.2}$	syn
1694	9.43	7.99	$5058^{+60}_{-55}$	$0.13^{+0.05}_{-0.06}$	$0.8^{+0.01}_{-0.01}$	$0.85^{+0.02}_{-0.02}$	$2.3^{+0.14}_{-0.15}$	$4.3^{+3.5}_{-2.6}$	syn
1710	7.96	12.28	$5684^{+63}_{-67}$	$0.15^{+0.04}_{-0.05}$	$0.96^{+0.02}_{-0.02}$	$1.03^{+0.02}_{-0.03}$	$1.64^{+0.08}_{-0.09}$	$1.5^{+1.6}_{-1.0}$	syn
1716	7.93	9.57	$5877^{+78}_{-78}$	$0.1^{+0.06}_{-0.06}$	$1.22^{+0.02}_{-0.02}$	$1.07^{+0.04}_{-0.04}$	$0.83^{+0.06}_{-0.06}$	$5.4^{+1.4}_{-1.3}$	syn
1723	8.13	9.9	$5742^{+82}_{-80}$	$0.16^{+0.05}_{-0.06}$	$1.09^{+0.02}_{-0.02}$	$1.04^{+0.04}_{-0.04}$	$1.13^{+0.09}_{-0.08}$	$4.8^{+1.9}_{-1.8}$	syn
1726	5.26	44.61	$5634^{+62}_{-64}$	$0.08^{+0.05}_{-0.05}$	$0.92^{+0.01}_{-0.01}$	$0.99^{+0.02}_{-0.03}$	$1.8^{+0.08}_{-0.09}$	$1.5^{+1.6}_{-1.0}$	syn
1736	7.28	11.21	$5636^{+86}_{-79}$	$0.15^{+0.06}_{-0.06}$	$1.43^{+0.03}_{-0.02}$	$1.04^{+0.04}_{-0.03}$	$0.5^{+0.04}_{-0.03}$	$9.2^{+1.3}_{-1.4}$	syn
1742	7.34	13.69	$5814^{+75}_{-72}$	$0.19^{+0.06}_{-0.06}$	$1.13^{+0.02}_{-0.02}$	$1.09^{+0.03}_{-0.03}$	$1.05^{+0.07}_{-0.06}$	$3.9^{+1.5}_{-1.3}$	syn
1751	7.93	8.79	$5970^{+85}_{-83}$	$-0.37^{+0.06}_{-0.06}$	$1.31^{+0.03}_{-0.02}$	$0.94^{+0.04}_{-0.03}$	$0.58^{+0.04}_{-0.04}$	$9.0^{+1.4}_{-1.4}$	syn
1753	10.18	4.38	$5620^{+75}_{-71}$	$0.03^{+0.06}_{-0.06}$	$0.97^{+0.02}_{-0.02}$	$0.95^{+0.03}_{-0.03}$	$1.44^{+0.12}_{-0.11}$	$5.9^{+2.4}_{-2.3}$	syn
1758	8.81	10.32	$5150^{+64}_{-59}$	$-0.02^{+0.06}_{-0.06}$	$0.82^{+0.01}_{-0.01}$	$0.83^{+0.03}_{-0.03}$	$2.13^{+0.15}_{-0.15}$	$7.5^{+3.7}_{-3.2}$	syn
1759	7.93	24.93	$3930^{+32}_{-29}$	$0.08^{+0.04}_{-0.04}$	$0.57^{+0.01}_{-0.01}$	$0.57^{+0.01}_{-0.01}$	$4.18^{+0.15}_{-0.17}$	$15.1^{+3.4}_{-5.4}$	emp
1775	9.72	6.67	$5283^{+50}_{-51}$	$0.16^{+0.03}_{-0.05}$	$0.83^{+0.01}_{-0.01}$	$0.92^{+0.01}_{-0.02}$	$2.27^{+0.07}_{-0.08}$	$0.8^{+1.1}_{-0.5}$	syn
1776	6.69	22.37	$5784^{+74}_{-71}$	$-0.19^{+0.06}_{-0.06}$	$0.94^{+0.02}_{-0.02}$	$0.92^{+0.03}_{-0.03}$	$1.57^{+0.11}_{-0.1}$	$5.1^{+2.2}_{-2.0}$	syn
1778	7.62	10.03	$6007^{+83}_{-87}$	$0.22^{+0.06}_{-0.06}$	$1.32^{+0.02}_{-0.01}$	$1.2^{+0.03}_{-0.03}$	$0.73^{+0.03}_{-0.05}$	$3.1^{+0.9}_{-0.7}$	syn
1794	8.69	6.39	$5631^{+77}_{-73}$	$0.03^{+0.05}_{-0.06}$	$1.32^{+0.02}_{-0.02}$	$0.97^{+0.03}_{-0.03}$	$0.59^{+0.04}_{-0.04}$	$10.8^{+1.5}_{-1.5}$	syn
1797	7.78	12.12	$5907^{+65}_{-71}$	$0.11^{+0.06}_{-0.06}$	$1.03^{+0.02}_{-0.02}$	$1.08^{+0.03}_{-0.03}$	$1.38^{+0.07}_{-0.08}$	$1.3^{+1.2}_{-0.8}$	syn
1798	9.24	8.81	$5106^{+53}_{-55}$	$0.07^{+0.06}_{-0.05}$	$0.79^{+0.01}_{-0.01}$	$0.85^{+0.02}_{-0.02}$	$2.42^{+0.1}_{-0.13}$	$2.6^{+2.6}_{-1.7}$	syn
1799	7.38	16.07	$5697^{+74}_{-68}$	$-0.06^{+0.06}_{-0.05}$	$0.95^{+0.02}_{-0.02}$	$0.94^{+0.03}_{-0.03}$	$1.52^{+0.1}_{-0.1}$	$5.1^{+2.2}_{-2.0}$	syn
1801	7.8	32.57	$3747^{+24}_{-23}$	$0.15^{+0.05}_{-0.04}$	$0.51^{+0.01}_{-0.01}$	$0.51^{+0.01}_{-0.01}$	$5.23^{+0.23}_{-0.23}$	$13.3^{+4.6}_{-6.6}$	emp
1807	7.57	23.46	$4914^{+60}_{-57}$	$0.04^{+0.09}_{-0.1}$	$0.75^{+0.02}_{-0.02}$	$0.8^{+0.02}_{-0.03}$	$2.72^{+0.18}_{-0.2}$	$2.1^{+3.4}_{-1.5}$	emp
1823	8.33	13.93	$4926^{+57}_{-56}$	$0.24^{+0.09}_{-0.09}$	$0.8^{+0.03}_{-0.03}$	$0.84^{+0.03}_{-0.03}$	$2.24^{+0.25}_{-0.31}$	$6.3^{+6.3}_{-4.4}$	emp
1824	7.76	16.8	$5165^{+56}_{-56}$	$0.12^{+0.06}_{-0.06}$	$0.81^{+0.01}_{-0.01}$	$0.88^{+0.02}_{-0.02}$	$2.3^{+0.1}_{-0.13}$	$2.5^{+2.5}_{-1.6}$	syn

Table 6.3 (cont'd): TKS Stellar Properties

TOI	$m_K$	$\pi$	$T_{\text{eff}}$	[Fe/H]	$R_*$	$M_*$	$\rho_*$	age	SpecMatch
	mag	mas	K	dex	$R_{\odot}$	$M_{\odot}$	$g \text{ cm}^{-3}$	Gyr	
1836	8.53	5.19	$6237^{+91}_{-68}$	$-0.14^{+0.07}_{-0.03}$	$1.65^{+0.03}_{-0.03}$	$1.25^{+0.02}_{-0.04}$	$0.39^{+0.02}_{-0.02}$	$3.2^{+0.7}_{-0.3}$	syn
1842	8.45	4.45	$6039^{+93}_{-88}$	$0.33^{+0.06}_{-0.06}$	$2.03^{+0.04}_{-0.04}$	$1.46^{+0.03}_{-0.04}$	$0.24^{+0.02}_{-0.02}$	$2.5^{+0.4}_{-0.3}$	syn
1898	6.66	12.52	$6241^{+88}_{-84}$	$-0.1^{+0.06}_{-0.07}$	$1.61^{+0.03}_{-0.03}$	$1.25^{+0.03}_{-0.04}$	$0.41^{+0.03}_{-0.02}$	$3.1^{+0.6}_{-0.3}$	syn
2019	8.6	5.02	$5625^{+77}_{-77}$	$0.41^{+0.06}_{-0.06}$	$1.75^{+0.03}_{-0.03}$	$1.18^{+0.11}_{-0.02}$	$0.31^{+0.03}_{-0.02}$	$6.4^{+0.6}_{-1.9}$	syn
2045	9.85	2.7	$6045^{+75}_{-68}$	$0.07^{+0.06}_{-0.04}$	$1.78^{+0.04}_{-0.04}$	$1.29^{+0.03}_{-0.08}$	$0.32^{+0.02}_{-0.02}$	$3.4^{+1.2}_{-0.4}$	syn
2076	7.12	23.83	$5191^{+61}_{-57}$	$0.02^{+0.06}_{-0.06}$	$0.8^{+0.03}_{-0.02}$	$0.86^{+0.02}_{-0.02}$	$2.36^{+0.19}_{-0.25}$	$2.7^{+3.6}_{-1.9}$	syn
2088	9.52	7.88	$5080^{+55}_{-54}$	$0.34^{+0.06}_{-0.06}$	$0.85^{+0.03}_{-0.02}$	$0.9^{+0.02}_{-0.03}$	$2.05^{+0.2}_{-0.26}$	$4.5^{+4.6}_{-3.1}$	syn
2114	9.09	3.12	$6359^{+88}_{-85}$	$0.11^{+0.05}_{-0.06}$	$2.1^{+0.04}_{-0.04}$	$1.48^{+0.03}_{-0.04}$	$0.23^{+0.01}_{-0.01}$	$2.0^{+0.2}_{-0.2}$	syn
2128	5.82	27.31	$5967^{+85}_{-84}$	$-0.07^{+0.05}_{-0.06}$	$1.12^{+0.02}_{-0.02}$	$1.03^{+0.04}_{-0.04}$	$1.02^{+0.08}_{-0.08}$	$4.7^{+1.7}_{-1.5}$	syn
2145	7.76	4.42	$6200^{+83}_{-81}$	$0.29^{+0.06}_{-0.05}$	$2.75^{+0.06}_{-0.05}$	$1.72^{+0.03}_{-0.04}$	$0.11^{+0.01}_{-0.01}$	$1.6^{+0.2}_{-0.1}$	syn

Note. — Properties of 85 planet hosting stars from the TKS sample (Chontos et al. 2022).  $m_K$  is 2MASS  $K$ -band apparent magnitude and  $\pi$  is *Gaia* DR2 parallax.  $T_{\text{eff}}$ , [Fe/H],  $R_*$ ,  $M_*$ ,  $\rho_*$  and age were derived using `isoclassify` isochrone modeling (Huber et al. 2017) as described in §6.3.2. The given values reflect median measurements with upper and lower uncertainties. The last column denotes which `SpecMatch` method was used to derive the initial stellar property inputs to `isoclassify`: "syn" – `SpecMatch-Synthetic` or "emp" – `SpecMatch-Empirical` (see §6.3.1).

Table 6.4: TKS Planet Properties

Planet	$P$ d	$t_0$ BTJD	$R_p/R_\star$ %	$R_p$ $R_\oplus$	$T_{14}$ hr	$T_{\text{circ}}$ hr	TTVs? y/n
260.01	$13.475832^{+4.5e-05}_{-5.6e-05}$	$1392.2944^{+0.0021}_{-0.0023}$	$2.47^{+0.2}_{-0.14}$	$1.47^{+0.12}_{-0.08}$	$2.0^{+0.14}_{-0.13}$	$2.96^{+0.03}_{-0.04}$	n
266.01	$10.751013^{+7.2e-05}_{-7e-05}$	$1393.0862^{+0.0035}_{-0.0032}$	$2.42^{+0.14}_{-0.12}$	$2.54^{+0.16}_{-0.14}$	$3.0^{+0.16}_{-0.13}$	$4.04^{+0.1}_{-0.09}$	n
266.02	$19.605464^{+0.000228}_{-0.000237}$	$1398.2927^{+0.0087}_{-0.0059}$	$2.41^{+0.14}_{-0.13}$	$2.52^{+0.15}_{-0.14}$	$4.38^{+0.27}_{-0.21}$	$4.94^{+0.12}_{-0.11}$	n
329.01	$5.70442^{+0.000116}_{-0.000101}$	$2090.7934^{+0.0043}_{-0.0057}$	$2.96^{+0.16}_{-0.14}$	$4.94^{+0.29}_{-0.26}$	$4.21^{+0.21}_{-0.18}$	$5.04^{+0.13}_{-0.11}$	n
465.01	$3.836162^{+2e-06}_{-2e-06}$	$1414.1361^{+0.0004}_{-0.0003}$	$6.68^{+0.09}_{-0.07}$	$6.03^{+0.13}_{-0.12}$	$2.4^{+0.02}_{-0.02}$	$2.64^{+0.06}_{-0.06}$	n
469.01	$13.630829^{+2.6e-05}_{-2.6e-05}$	$1474.5691^{+0.0011}_{-0.0011}$	$3.19^{+0.07}_{-0.06}$	$3.43^{+0.1}_{-0.09}$	$4.33^{+0.05}_{-0.04}$	$4.5^{+0.11}_{-0.11}$	n
480.01	$6.865906^{+1.7e-05}_{-1.9e-05}$	$1469.5655^{+0.0016}_{-0.0015}$	$1.76^{+0.06}_{-0.05}$	$2.85^{+0.13}_{-0.1}$	$3.57^{+0.06}_{-0.06}$	$4.84^{+0.13}_{-0.15}$	n
509.01	$9.058805^{+1.8e-05}_{-1.6e-05}$	$1494.4467^{+0.001}_{-0.001}$	$2.79^{+0.07}_{-0.06}$	$2.92^{+0.09}_{-0.08}$	$3.7^{+0.04}_{-0.04}$	$3.82^{+0.09}_{-0.09}$	n
509.02	$21.402464^{+0.001794}_{-0.001919}$	$1504.1438^{+0.067}_{-0.0622}$	$2.93^{+0.07}_{-0.06}$	$3.06^{+0.09}_{-0.08}$	$5.42^{+0.08}_{-0.06}$	$5.1^{+0.12}_{-0.12}$	n
554.01	$3.04405^{+9e-06}_{-8e-06}$	$1438.4733^{+0.002}_{-0.0025}$	$0.88^{+0.08}_{-0.06}$	$1.36^{+0.13}_{-0.1}$	$1.35^{+0.06}_{-0.06}$	$3.52^{+0.08}_{-0.05}$	n
554.02	$7.049157^{+1.3e-05}_{-1.4e-05}$	$1442.6179^{+0.0016}_{-0.0016}$	$1.65^{+0.07}_{-0.05}$	$2.55^{+0.11}_{-0.08}$	$3.35^{+0.06}_{-0.04}$	$4.68^{+0.1}_{-0.07}$	n
561.01	$0.4465691^{+4e-07}_{-4e-07}$	$1517.9451^{+0.0007}_{-0.0008}$	$1.56^{+0.05}_{-0.04}$	$1.46^{+0.05}_{-0.05}$	$1.35^{+0.02}_{-0.02}$	$1.36^{+0.03}_{-0.03}$	n
561.02	$10.778858^{+3.3e-05}_{-3.5e-05}$	$1527.0608^{+0.0025}_{-0.0023}$	$3.13^{+0.21}_{-0.09}$	$2.93^{+0.2}_{-0.1}$	$3.83^{+0.16}_{-0.11}$	$3.89^{+0.08}_{-0.07}$	n
561.03	$25.712443^{+0.000111}_{-0.000122}$	$1521.8835^{+0.0038}_{-0.0036}$	$2.91^{+0.11}_{-0.09}$	$2.72^{+0.11}_{-0.1}$	$5.19^{+0.17}_{-0.12}$	$5.19^{+0.11}_{-0.09}$	n
669.01	$3.945154^{+1.3e-05}_{-1.1e-05}$	$1546.1417^{+0.0021}_{-0.003}$	$2.4^{+0.11}_{-0.09}$	$2.59^{+0.13}_{-0.11}$	$2.87^{+0.1}_{-0.09}$	$2.99^{+0.07}_{-0.07}$	n
1136.01	$6.258767^{+1.5e-05}_{-1.4e-05}$	$1688.7067^{+0.0013}_{-0.0012}$	$2.65^{+0.06}_{-0.05}$	$2.77^{+0.08}_{-0.07}$	$3.49^{+0.05}_{-0.04}$	$3.25^{+0.04}_{-0.04}$	y
1136.02	$12.518595^{+1.5e-05}_{-1.5e-05}$	$1686.0634^{+0.0005}_{-0.0006}$	$4.43^{+0.13}_{-0.08}$	$4.62^{+0.15}_{-0.1}$	$4.13^{+0.07}_{-0.04}$	$4.17^{+0.05}_{-0.06}$	y
1136.03	$18.797505^{+0.000141}_{-0.000121}$	$1697.799^{+0.004}_{-0.0045}$	$2.44^{+0.09}_{-0.08}$	$2.55^{+0.1}_{-0.09}$	$4.46^{+0.11}_{-0.1}$	$4.68^{+0.06}_{-0.06}$	y
1136.04	$26.317767^{+4.7e-05}_{-4.7e-05}$	$1699.3798^{+0.0011}_{-0.0011}$	$3.58^{+0.06}_{-0.06}$	$3.74^{+0.09}_{-0.08}$	$5.18^{+0.05}_{-0.05}$	$5.3^{+0.07}_{-0.07}$	y
1173.01	$7.064397^{+3e-06}_{-3e-06}$	$1688.7154^{+0.0002}_{-0.0002}$	$8.83^{+0.06}_{-0.06}$	$9.02^{+0.16}_{-0.15}$	$2.56^{+0.02}_{-0.02}$	$3.63^{+0.08}_{-0.08}$	n
1174.01	$8.9534^{+4.1e-05}_{-3.9e-05}$	$1690.0626^{+0.0022}_{-0.0026}$	$2.93^{+0.13}_{-0.1}$	$2.51^{+0.14}_{-0.11}$	$3.0^{+0.12}_{-0.1}$	$3.23^{+0.08}_{-0.11}$	n
1180.01	$9.686755^{+1.2e-05}_{-1.2e-05}$	$1691.0488^{+0.0009}_{-0.0009}$	$3.75^{+0.2}_{-0.12}$	$3.04^{+0.19}_{-0.13}$	$2.76^{+0.08}_{-0.05}$	$3.25^{+0.09}_{-0.11}$	n
1181.01	$2.1031937^{+3e-07}_{-3e-07}$	$1957.8213^{+0.0001}_{-0.0001}$	$7.68^{+0.02}_{-0.02}$	$16.2^{+0.33}_{-0.34}$	$4.12^{+0.01}_{-0.01}$	$4.33^{+0.09}_{-0.08}$	n
1184.01	$5.748431^{+3e-06}_{-3e-06}$	$1684.3594^{+0.0005}_{-0.0005}$	$3.17^{+0.09}_{-0.16}$	$2.41^{+0.1}_{-0.14}$	$1.88^{+0.03}_{-0.03}$	$2.61^{+0.07}_{-0.08}$	n
1194.01	$2.310645^{+1e-06}_{-1e-06}$	$1684.9235^{+0.0002}_{-0.0002}$	$8.32^{+0.06}_{-0.06}$	$8.72^{+0.16}_{-0.15}$	$1.63^{+0.01}_{-0.01}$	$2.53^{+0.06}_{-0.05}$	n

Table 6.4 (cont'd): TKS Planet Properties

Planet	$P$ d	$t_0$ BTJD	$R_p/R_\star$ %	$R_p$ $R_\oplus$	$T_{14}$ hr	$T_{\text{circ}}$ hr	TTVs? y/n
1244.01	$6.400317^{+8e-06}_{-8e-06}$	$1684.9474^{+0.0008}_{-0.0008}$	$3.01^{+0.14}_{-0.1}$	$2.38^{+0.13}_{-0.11}$	$2.22^{+0.04}_{-0.04}$	$2.77^{+0.08}_{-0.09}$	n
1246.01	$4.30744^{+4e-06}_{-3e-06}$	$1686.5661^{+0.0007}_{-0.0008}$	$3.05^{+0.12}_{-0.08}$	$2.87^{+0.13}_{-0.1}$	$2.18^{+0.04}_{-0.03}$	$2.75^{+0.08}_{-0.07}$	n
1246.02	$5.904141^{+1.3e-05}_{-1.4e-05}$	$1683.4664^{+0.0016}_{-0.0016}$	$2.6^{+0.14}_{-0.08}$	$2.45^{+0.14}_{-0.09}$	$2.89^{+0.07}_{-0.07}$	$3.04^{+0.08}_{-0.08}$	n
1246.03	$18.654958^{+4.9e-05}_{-4.7e-05}$	$1688.9746^{+0.002}_{-0.002}$	$3.64^{+0.14}_{-0.1}$	$3.43^{+0.15}_{-0.12}$	$3.95^{+0.1}_{-0.08}$	$4.5^{+0.12}_{-0.12}$	y
1246.04	$37.925389^{+0.000149}_{-0.000156}$	$1700.6956^{+0.0033}_{-0.0034}$	$3.73^{+0.27}_{-0.18}$	$3.51^{+0.27}_{-0.18}$	$3.54^{+0.16}_{-0.12}$	$5.71^{+0.16}_{-0.15}$	y
1247.01	$15.923365^{+3.4e-05}_{-3.6e-05}$	$1687.6494^{+0.0014}_{-0.0014}$	$2.16^{+0.06}_{-0.05}$	$2.53^{+0.08}_{-0.07}$	$4.33^{+0.07}_{-0.07}$	$5.18^{+0.12}_{-0.11}$	n
1248.01	$4.360156^{+1e-06}_{-1e-06}$	$1687.1212^{+0.0002}_{-0.0002}$	$7.19^{+0.05}_{-0.06}$	$6.81^{+0.12}_{-0.12}$	$2.31^{+0.01}_{-0.01}$	$2.87^{+0.06}_{-0.06}$	n
1249.01	$13.079163^{+6.3e-05}_{-6.3e-05}$	$1694.3786^{+0.0027}_{-0.0029}$	$3.06^{+0.28}_{-0.17}$	$3.27^{+0.31}_{-0.19}$	$3.15^{+0.26}_{-0.15}$	$4.32^{+0.1}_{-0.1}$	n
1255.01	$10.288889^{+4e-06}_{-4e-06}$	$1691.6544^{+0.0004}_{-0.0003}$	$2.7^{+0.07}_{-0.05}$	$2.48^{+0.07}_{-0.05}$	$1.53^{+0.02}_{-0.01}$	$3.5^{+0.04}_{-0.05}$	n
1269.01	$4.252989^{+6e-06}_{-6e-06}$	$1686.6058^{+0.0009}_{-0.0013}$	$2.58^{+0.16}_{-0.07}$	$2.4^{+0.15}_{-0.07}$	$2.62^{+0.07}_{-0.05}$	$2.67^{+0.05}_{-0.05}$	n
1269.02	$9.237875^{+2e-05}_{-2e-05}$	$1685.9772^{+0.0022}_{-0.0017}$	$2.45^{+0.12}_{-0.09}$	$2.27^{+0.12}_{-0.09}$	$2.72^{+0.07}_{-0.07}$	$3.45^{+0.07}_{-0.07}$	n
1272.01	$3.315979^{+6e-06}_{-5e-06}$	$1713.0315^{+0.0006}_{-0.0006}$	$4.77^{+0.25}_{-0.17}$	$4.13^{+0.23}_{-0.15}$	$1.57^{+0.06}_{-0.05}$	$2.36^{+0.03}_{-0.03}$	n
1279.01	$9.614216^{+3.7e-05}_{-3.5e-05}$	$1717.4777^{+0.0026}_{-0.0027}$	$2.9^{+0.14}_{-0.11}$	$2.66^{+0.14}_{-0.11}$	$2.76^{+0.11}_{-0.09}$	$3.5^{+0.07}_{-0.07}$	y
1288.01	$2.699831^{+1e-06}_{-1e-06}$	$1712.3587^{+0.0002}_{-0.0002}$	$4.75^{+0.06}_{-0.05}$	$4.97^{+0.11}_{-0.1}$	$2.38^{+0.01}_{-0.01}$	$2.59^{+0.06}_{-0.06}$	n
1294.01	$3.915289^{+1.6e-05}_{-1.6e-05}$	$2393.0074^{+0.0007}_{-0.0007}$	$5.96^{+0.08}_{-0.08}$	$10.12^{+0.24}_{-0.22}$	$3.03^{+0.05}_{-0.05}$	$4.51^{+0.1}_{-0.12}$	n
1296.01	$3.944373^{+1e-06}_{-1e-06}$	$1930.7553^{+0.0002}_{-0.0002}$	$7.6^{+0.04}_{-0.04}$	$14.12^{+0.28}_{-0.26}$	$4.86^{+0.01}_{-0.01}$	$5.04^{+0.12}_{-0.1}$	n
1298.01	$4.537142^{+2e-06}_{-2e-06}$	$1934.1225^{+0.0003}_{-0.0003}$	$6.09^{+0.05}_{-0.04}$	$9.66^{+0.2}_{-0.17}$	$4.03^{+0.02}_{-0.02}$	$4.38^{+0.09}_{-0.1}$	n
1339.01	$8.880322^{+3e-06}_{-3e-06}$	$1715.3558^{+0.0003}_{-0.0002}$	$3.33^{+0.05}_{-0.07}$	$3.4^{+0.08}_{-0.09}$	$3.07^{+0.02}_{-0.02}$	$3.89^{+0.09}_{-0.08}$	n
1339.02	$28.579984^{+2.4e-05}_{-2.4e-05}$	$1726.0547^{+0.0006}_{-0.0006}$	$3.12^{+0.08}_{-0.07}$	$3.18^{+0.1}_{-0.09}$	$4.47^{+0.05}_{-0.04}$	$5.73^{+0.14}_{-0.12}$	y
1339.03	$38.352218^{+4.2e-05}_{-4.1e-05}$	$1743.556^{+0.0008}_{-0.0007}$	$3.06^{+0.05}_{-0.04}$	$3.12^{+0.07}_{-0.07}$	$5.51^{+0.05}_{-0.04}$	$6.31^{+0.15}_{-0.13}$	y
1347.01	$0.8474247^{+4e-07}_{-4e-07}$	$1683.5587^{+0.0003}_{-0.0003}$	$1.98^{+0.09}_{-0.06}$	$1.81^{+0.09}_{-0.06}$	$0.88^{+0.02}_{-0.01}$	$1.53^{+0.02}_{-0.02}$	n
1347.02	$4.84195^{+1.3e-05}_{-1.2e-05}$	$1688.1903^{+0.0017}_{-0.002}$	$1.83^{+0.09}_{-0.07}$	$1.68^{+0.09}_{-0.07}$	$2.24^{+0.09}_{-0.08}$	$2.7^{+0.03}_{-0.03}$	n
1386.01	$25.840653^{+0.003119}_{-0.003192}$	$1752.3213^{+0.0022}_{-0.0022}$	$5.57^{+0.16}_{-0.12}$	$6.22^{+0.21}_{-0.17}$	$5.67^{+0.15}_{-0.12}$	$5.75^{+0.13}_{-0.13}$	n
1410.01	$1.216876^{+1e-06}_{-1e-06}$	$1739.7293^{+0.0006}_{-0.0006}$	$3.76^{+0.24}_{-0.17}$	$3.1^{+0.22}_{-0.16}$	$1.11^{+0.05}_{-0.04}$	$1.66^{+0.05}_{-0.06}$	n
1411.01	$1.452053^{+2e-06}_{-2e-06}$	$1739.474^{+0.0005}_{-0.0005}$	$1.79^{+0.05}_{-0.04}$	$1.2^{+0.05}_{-0.04}$	$1.51^{+0.02}_{-0.02}$	$1.53^{+0.03}_{-0.03}$	n

Table 6.4 (cont'd): TKS Planet Properties

Planet	$P$ d	$t_0$ BTJD	$R_p/R_\star$ %	$R_p$ $R_\oplus$	$T_{14}$ hr	$T_{\text{circ}}$ hr	TTVs? y/n
1422.01	13.001782 $^{+5e-05}_{-4.2e-05}$	1745.9141 $^{+0.0018}_{-0.0016}$	3.53 $^{+0.12}_{-0.09}$	3.94 $^{+0.15}_{-0.13}$	4.44 $^{+0.11}_{-0.09}$	4.55 $^{+0.11}_{-0.11}$	n
1430.01	7.434098 $^{+9e-06}_{-9e-06}$	1690.779 $^{+0.0009}_{-0.0009}$	2.38 $^{+0.07}_{-0.05}$	2.08 $^{+0.09}_{-0.07}$	2.67 $^{+0.04}_{-0.04}$	3.06 $^{+0.08}_{-0.11}$	n
1436.01	0.867617 $^{+2e-06}_{-3e-06}$	1711.8961 $^{+0.0012}_{-0.0011}$	1.87 $^{+0.1}_{-0.09}$	1.52 $^{+0.1}_{-0.09}$	1.32 $^{+0.05}_{-0.05}$	1.45 $^{+0.05}_{-0.06}$	n
1437.01	18.840952 $^{+7.7e-05}_{-7.7e-05}$	1700.7335 $^{+0.0029}_{-0.003}$	1.77 $^{+0.04}_{-0.03}$	2.43 $^{+0.07}_{-0.06}$	6.19 $^{+0.12}_{-0.11}$	6.19 $^{+0.15}_{-0.14}$	n
1438.01	5.139659 $^{+5e-06}_{-5e-06}$	1683.6266 $^{+0.0007}_{-0.0007}$	3.22 $^{+0.2}_{-0.34}$	2.87 $^{+0.18}_{-0.3}$	1.01 $^{+0.06}_{-0.1}$	2.75 $^{+0.04}_{-0.05}$	n
1438.02	9.428074 $^{+1.6e-05}_{-1.5e-05}$	1689.9152 $^{+0.0014}_{-0.0014}$	2.92 $^{+0.17}_{-0.3}$	2.6 $^{+0.15}_{-0.27}$	1.56 $^{+0.11}_{-0.1}$	3.36 $^{+0.05}_{-0.06}$	n
1439.01	27.643869 $^{+0.000133}_{-0.000138}$	1703.4782 $^{+0.004}_{-0.0039}$	2.4 $^{+0.12}_{-0.09}$	4.24 $^{+0.23}_{-0.17}$	5.63 $^{+0.21}_{-0.14}$	8.52 $^{+0.17}_{-0.2}$	n
1443.01	23.540714 $^{+6.6e-05}_{-6.2e-05}$	1693.2443 $^{+0.0023}_{-0.0025}$	2.81 $^{+0.09}_{-0.07}$	2.3 $^{+0.13}_{-0.1}$	4.26 $^{+0.08}_{-0.07}$	4.43 $^{+0.2}_{-0.21}$	n
1444.01	0.4702737 $^{+3e-07}_{-3e-07}$	1711.3666 $^{+0.0005}_{-0.0008}$	1.43 $^{+0.04}_{-0.03}$	1.42 $^{+0.05}_{-0.04}$	1.28 $^{+0.02}_{-0.02}$	1.36 $^{+0.03}_{-0.03}$	n
1451.01	16.537903 $^{+6.9e-05}_{-6.3e-05}$	1694.3138 $^{+0.0021}_{-0.0029}$	2.36 $^{+0.11}_{-0.08}$	2.61 $^{+0.13}_{-0.1}$	3.09 $^{+0.09}_{-0.07}$	4.82 $^{+0.11}_{-0.11}$	n
1456.01	18.711796 $^{+1.4e-05}_{-1.4e-05}$	1692.2609 $^{+0.0005}_{-0.0005}$	6.57 $^{+0.06}_{-0.05}$	9.07 $^{+0.19}_{-0.18}$	6.17 $^{+0.03}_{-0.03}$	6.21 $^{+0.14}_{-0.12}$	y
1467.01	5.971148 $^{+1.1e-05}_{-1e-05}$	1766.9883 $^{+0.0013}_{-0.0022}$	3.37 $^{+0.17}_{-0.16}$	1.68 $^{+0.09}_{-0.08}$	1.58 $^{+0.07}_{-0.05}$	2.03 $^{+0.02}_{-0.02}$	n
1471.01	20.772891 $^{+4.8e-05}_{-5e-05}$	1767.422 $^{+0.0014}_{-0.0014}$	3.66 $^{+0.07}_{-0.06}$	3.83 $^{+0.1}_{-0.09}$	4.94 $^{+0.06}_{-0.05}$	5.08 $^{+0.12}_{-0.11}$	n
1471.02	52.563553 $^{+0.000166}_{-0.000185}$	1779.1909 $^{+0.0016}_{-0.0014}$	3.28 $^{+0.07}_{-0.06}$	3.43 $^{+0.1}_{-0.09}$	6.57 $^{+0.08}_{-0.06}$	6.9 $^{+0.17}_{-0.15}$	n
1472.01	6.363386 $^{+9e-06}_{-8e-06}$	1765.6096 $^{+0.0012}_{-0.0013}$	4.55 $^{+0.15}_{-0.12}$	4.16 $^{+0.16}_{-0.13}$	2.61 $^{+0.06}_{-0.05}$	3.04 $^{+0.05}_{-0.06}$	n
1473.01	5.254479 $^{+1.2e-05}_{-1.3e-05}$	1769.7834 $^{+0.0026}_{-0.0019}$	2.21 $^{+0.08}_{-0.06}$	2.43 $^{+0.1}_{-0.07}$	3.24 $^{+0.1}_{-0.07}$	3.23 $^{+0.07}_{-0.07}$	n
1601.01	5.333298 $^{+0.000889}_{-0.000968}$	1793.2752 $^{+0.0017}_{-0.0016}$	5.93 $^{+0.21}_{-0.13}$	14.2 $^{+0.65}_{-0.51}$	6.49 $^{+0.13}_{-0.11}$	6.5 $^{+0.18}_{-0.18}$	n
1611.01	16.201708 $^{+1.3e-05}_{-1.2e-05}$	1796.495 $^{+0.0006}_{-0.0006}$	2.74 $^{+0.09}_{-0.05}$	2.34 $^{+0.12}_{-0.09}$	2.9 $^{+0.03}_{-0.03}$	3.96 $^{+0.14}_{-0.17}$	y
1669.01	2.680055 $^{+3e-06}_{-3e-06}$	1816.9447 $^{+0.0009}_{-0.0008}$	1.99 $^{+0.08}_{-0.07}$	2.3 $^{+0.1}_{-0.09}$	1.93 $^{+0.04}_{-0.03}$	2.75 $^{+0.07}_{-0.06}$	n
1691.01	16.7369 $^{+3.2e-05}_{-3e-05}$	1818.0907 $^{+0.0013}_{-0.0014}$	3.24 $^{+0.08}_{-0.05}$	3.57 $^{+0.11}_{-0.09}$	4.97 $^{+0.07}_{-0.05}$	4.92 $^{+0.12}_{-0.12}$	n
1694.01	3.770137 $^{+8.8e-05}_{-8.9e-05}$	1817.2664 $^{+0.0006}_{-0.0006}$	6.09 $^{+0.14}_{-0.1}$	5.34 $^{+0.15}_{-0.12}$	2.87 $^{+0.04}_{-0.03}$	2.55 $^{+0.05}_{-0.05}$	n
1710.01	24.283384 $^{+1.9e-05}_{-2e-05}$	1836.9629 $^{+0.0005}_{-0.0005}$	4.98 $^{+0.04}_{-0.04}$	5.2 $^{+0.1}_{-0.09}$	5.26 $^{+0.03}_{-0.02}$	5.25 $^{+0.09}_{-0.1}$	n
1716.01	8.082366 $^{+3.5e-05}_{-3.6e-05}$	1843.8553 $^{+0.0034}_{-0.0031}$	2.17 $^{+0.12}_{-0.09}$	2.88 $^{+0.17}_{-0.13}$	3.51 $^{+0.17}_{-0.16}$	4.45 $^{+0.1}_{-0.1}$	n
1723.01	13.726468 $^{+0.000386}_{-0.000398}$	1852.7027 $^{+0.0027}_{-0.0023}$	2.78 $^{+0.11}_{-0.09}$	3.29 $^{+0.14}_{-0.13}$	4.4 $^{+0.22}_{-0.22}$	4.81 $^{+0.12}_{-0.12}$	n
1726.01	7.107941 $^{+7e-06}_{-6e-06}$	1845.3735 $^{+0.0005}_{-0.0006}$	2.16 $^{+0.05}_{-0.04}$	2.16 $^{+0.06}_{-0.05}$	3.24 $^{+0.04}_{-0.03}$	3.29 $^{+0.05}_{-0.06}$	n

Table 6.4 (cont'd): TKS Planet Properties

Planet	$P$ d	$t_0$ BTJD	$R_p/R_*$ %	$R_p$ $R_\oplus$	$T_{14}$ hr	$T_{\text{circ}}$ hr	TTVs? y/n
1726.02	$20.543827^{+2.1e-05}_{-2.1e-05}$	$1844.0589^{+0.0006}_{-0.0006}$	$2.58^{+0.04}_{-0.04}$	$2.58^{+0.06}_{-0.06}$	$4.07^{+0.04}_{-0.04}$	$4.7^{+0.07}_{-0.08}$	n
1736.01	$7.073092^{+1.7e-05}_{-1.7e-05}$	$1792.7939^{+0.0018}_{-0.0019}$	$2.04^{+0.07}_{-0.05}$	$3.18^{+0.12}_{-0.1}$	$3.98^{+0.08}_{-0.08}$	$5.04^{+0.12}_{-0.11}$	n
1742.01	$21.269084^{+5.4e-05}_{-5.1e-05}$	$1725.352^{+0.0018}_{-0.0022}$	$1.92^{+0.03}_{-0.03}$	$2.37^{+0.06}_{-0.05}$	$6.29^{+0.07}_{-0.07}$	$5.66^{+0.12}_{-0.11}$	y
1751.01	$37.468555^{+0.000141}_{-0.000138}$	$1733.6315^{+0.0028}_{-0.0029}$	$2.06^{+0.07}_{-0.04}$	$2.95^{+0.11}_{-0.08}$	$7.65^{+0.16}_{-0.1}$	$8.35^{+0.21}_{-0.19}$	n
1753.01	$5.384623^{+9e-06}_{-1.1e-05}$	$1684.504^{+0.0016}_{-0.0013}$	$2.34^{+0.08}_{-0.06}$	$2.48^{+0.1}_{-0.08}$	$3.18^{+0.05}_{-0.05}$	$3.23^{+0.09}_{-0.08}$	n
1758.01	$20.705061^{+4.8e-05}_{-4.8e-05}$	$1806.6981^{+0.0015}_{-0.0016}$	$4.0^{+0.25}_{-0.15}$	$3.56^{+0.23}_{-0.14}$	$3.67^{+0.1}_{-0.08}$	$4.52^{+0.11}_{-0.1}$	n
1759.01	$18.850009^{+2.2e-05}_{-2.3e-05}$	$1745.4661^{+0.0009}_{-0.001}$	$4.98^{+0.17}_{-0.1}$	$3.12^{+0.12}_{-0.08}$	$3.59^{+0.09}_{-0.05}$	$3.53^{+0.04}_{-0.05}$	n
1775.01	$10.240554^{+1e-05}_{-1e-05}$	$1877.5645^{+0.0005}_{-0.0005}$	$8.9^{+0.1}_{-0.09}$	$8.05^{+0.14}_{-0.13}$	$3.66^{+0.04}_{-0.03}$	$3.66^{+0.04}_{-0.04}$	n
1776.01	$2.799865^{+2.5e-05}_{-3.2e-05}$	$1871.498^{+0.0038}_{-0.0037}$	$1.19^{+0.09}_{-0.07}$	$1.22^{+0.09}_{-0.08}$	$1.93^{+0.19}_{-0.18}$	$2.5^{+0.06}_{-0.05}$	n
1778.01	$6.527337^{+4.4e-05}_{-4e-05}$	$1876.0046^{+0.0025}_{-0.0026}$	$2.02^{+0.12}_{-0.1}$	$2.9^{+0.19}_{-0.14}$	$2.98^{+0.18}_{-0.15}$	$4.31^{+0.05}_{-0.09}$	n
1794.01	$8.765566^{+7.3e-05}_{-6.8e-05}$	$1715.309^{+0.004}_{-0.0033}$	$2.3^{+0.11}_{-0.09}$	$3.3^{+0.17}_{-0.14}$	$3.92^{+0.13}_{-0.13}$	$5.11^{+0.12}_{-0.1}$	n
1797.01	$1.039141^{+5e-06}_{-3e-06}$	$1900.238^{+0.0014}_{-0.0016}$	$1.31^{+0.06}_{-0.05}$	$1.47^{+0.07}_{-0.07}$	$1.82^{+0.06}_{-0.06}$	$1.89^{+0.03}_{-0.04}$	n
1797.02	$3.645143^{+8e-06}_{-8e-06}$	$1902.8746^{+0.0014}_{-0.0014}$	$2.77^{+0.11}_{-0.08}$	$3.12^{+0.14}_{-0.1}$	$2.36^{+0.06}_{-0.06}$	$2.89^{+0.05}_{-0.06}$	n
1798.01	$0.437815^{+1e-06}_{-2e-06}$	$1739.0717^{+0.001}_{-0.001}$	$1.62^{+0.08}_{-0.06}$	$1.4^{+0.07}_{-0.06}$	$1.28^{+0.04}_{-0.05}$	$1.19^{+0.02}_{-0.02}$	n
1798.02	$8.021543^{+2.9e-05}_{-3.3e-05}$	$1741.5941^{+0.0022}_{-0.0022}$	$2.76^{+0.13}_{-0.1}$	$2.39^{+0.12}_{-0.09}$	$3.14^{+0.11}_{-0.08}$	$3.12^{+0.04}_{-0.06}$	n
1799.01	$7.085754^{+8.4e-05}_{-9.2e-05}$	$1904.8305^{+0.0084}_{-0.008}$	$1.37^{+0.09}_{-0.08}$	$1.42^{+0.09}_{-0.09}$	$3.14^{+0.18}_{-0.18}$	$3.45^{+0.08}_{-0.08}$	n
1801.01	$10.643985^{+2.7e-05}_{-2.5e-05}$	$1903.5423^{+0.0013}_{-0.0014}$	$3.51^{+0.11}_{-0.11}$	$1.96^{+0.08}_{-0.08}$	$2.7^{+0.06}_{-0.05}$	$2.67^{+0.04}_{-0.04}$	n
1807.01	$0.549371^{+1e-06}_{-1e-06}$	$1900.4436^{+0.0004}_{-0.0003}$	$1.84^{+0.09}_{-0.07}$	$1.5^{+0.08}_{-0.07}$	$0.98^{+0.02}_{-0.02}$	$1.23^{+0.03}_{-0.03}$	n
1823.01	$38.81359^{+3.3e-05}_{-3.3e-05}$	$1715.179^{+0.0004}_{-0.0004}$	$8.59^{+0.11}_{-0.08}$	$7.54^{+0.33}_{-0.25}$	$5.92^{+0.04}_{-0.04}$	$5.72^{+0.21}_{-0.26}$	n
1824.01	$22.80853^{+6e-05}_{-6.2e-05}$	$1879.5468^{+0.0013}_{-0.0012}$	$3.1^{+0.09}_{-0.07}$	$2.74^{+0.09}_{-0.07}$	$4.11^{+0.06}_{-0.05}$	$4.51^{+0.07}_{-0.08}$	n
1836.01	$1.772745^{+6e-06}_{-5e-06}$	$1929.5241^{+0.0018}_{-0.0024}$	$1.45^{+0.12}_{-0.08}$	$2.6^{+0.22}_{-0.14}$	$1.79^{+0.06}_{-0.08}$	$3.46^{+0.07}_{-0.07}$	n
1836.02	$20.380831^{+2.5e-05}_{-2.7e-05}$	$1933.166^{+0.0008}_{-0.0008}$	$4.61^{+0.07}_{-0.06}$	$8.28^{+0.2}_{-0.17}$	$6.62^{+0.07}_{-0.05}$	$7.97^{+0.16}_{-0.16}$	n
1842.01	$9.573924^{+1.5e-05}_{-1.5e-05}$	$1933.3359^{+0.0009}_{-0.0009}$	$5.57^{+0.07}_{-0.07}$	$12.35^{+0.28}_{-0.29}$	$4.31^{+0.08}_{-0.07}$	$7.32^{+0.16}_{-0.15}$	n
1898.01	$45.522288^{+6.1e-05}_{-6.2e-05}$	$1894.252^{+0.0007}_{-0.0007}$	$5.34^{+0.09}_{-0.14}$	$9.4^{+0.23}_{-0.31}$	$4.27^{+0.06}_{-0.06}$	$10.28^{+0.22}_{-0.19}$	n
2019.01	$15.32222^{+0.000181}_{-0.000172}$	$1942.9799^{+0.0042}_{-0.0054}$	$2.8^{+0.13}_{-0.1}$	$5.33^{+0.26}_{-0.21}$	$7.57^{+0.24}_{-0.26}$	$7.67^{+0.26}_{-0.13}$	n

Table 6.4 (cont'd): TKS Planet Properties

Planet	$P$ d	$t_0$ BTJD	$R_p/R_\star$ %	$R_p$ $R_\oplus$	$T_{14}$ hr	$T_{\text{circ}}$ hr	TTVs? y/n
2045.01	$9.077535^{+8.5e-05}_{-8.4e-05}$	$1765.5949^{+0.0014}_{-0.0014}$	$6.59^{+0.12}_{-0.1}$	$12.77^{+0.36}_{-0.34}$	$5.9^{+0.09}_{-0.08}$	$6.62^{+0.16}_{-0.16}$	n
2076.01	$10.355307^{+1.4e-05}_{-1.5e-05}$	$1743.724^{+0.001}_{-0.0009}$	$3.19^{+0.12}_{-0.07}$	$2.77^{+0.14}_{-0.1}$	$3.24^{+0.05}_{-0.04}$	$3.44^{+0.09}_{-0.12}$	y
2076.02	$21.015324^{+2.3e-05}_{-2.4e-05}$	$1748.6895^{+0.0006}_{-0.0006}$	$4.26^{+0.06}_{-0.05}$	$3.69^{+0.14}_{-0.11}$	$4.19^{+0.04}_{-0.03}$	$4.4^{+0.12}_{-0.16}$	y
2076.03	$35.125611^{+9.6e-05}_{-8.5e-05}$	$1762.6658^{+0.0016}_{-0.0019}$	$3.95^{+0.2}_{-0.27}$	$3.43^{+0.21}_{-0.25}$	$2.97^{+0.12}_{-0.09}$	$5.2^{+0.14}_{-0.18}$	n
2088.01	$124.730182^{+0.00061}_{-0.000575}$	$1769.6077^{+0.0031}_{-0.0032}$	$3.98^{+0.13}_{-0.1}$	$3.68^{+0.19}_{-0.14}$	$7.36^{+0.15}_{-0.14}$	$8.32^{+0.27}_{-0.36}$	n
2114.01	$6.209837^{+0.000206}_{-0.000197}$	$2719.047^{+0.0011}_{-0.0011}$	$6.17^{+0.07}_{-0.08}$	$14.1^{+0.33}_{-0.32}$	$4.6^{+0.08}_{-0.08}$	$6.55^{+0.11}_{-0.13}$	n
2128.01	$16.341418^{+0.000145}_{-0.000128}$	$1987.2651^{+0.0022}_{-0.0022}$	$1.71^{+0.06}_{-0.05}$	$2.09^{+0.09}_{-0.08}$	$4.7^{+0.14}_{-0.11}$	$5.22^{+0.14}_{-0.13}$	n
2145.01	$10.261125^{+1.1e-05}_{-1.2e-05}$	$2013.2802^{+0.0006}_{-0.0006}$	$4.14^{+0.05}_{-0.04}$	$12.42^{+0.31}_{-0.25}$	$7.47^{+0.05}_{-0.04}$	$9.52^{+0.21}_{-0.19}$	n

Note. — Properties of 108 planets orbiting the 85 star TKS sample. Orbital period  $P$ , transit mid-point  $t_0$  (given in BTJD = BJD - 2457000), transit duration  $T_{14}$ , and planet-to-star radius ratio  $R_p/R_\star$  were measured from a final fit using `exoplanet` (Foreman-Mackey et al. 2021a) and reweighted via importance sampling. The given values reflect median measurements with upper and lower uncertainties.  $R_p$  follows from  $R_p/R_\star$  and  $R_\star$ . The expected duration of a centrally transiting object on a circular orbit  $T_{\text{circ}}$  and semi-major axis  $a$  are determined from Kepler’s Third Law (Winn 2010b). The last column shows a flag (yes or no) indicating whether or not observed TTVs are significant for a given planet.

## CHAPTER 7

# Distribution of Orbital Eccentricities for 93 Small Planets and 15 Giant Planets from the TESS-Keck Survey

### 7.1 Abstract

We determine the underlying eccentricity distribution for a sample of transiting planets from the TESS-Keck Survey. We use a hierarchical Bayesian model to fit the population-level eccentricity distribution of 108 planets from their individual posterior samples. The eccentricities are consistent with a Rayleigh distribution with a mean value of  $\bar{e} = 0.135^{+0.018}_{-0.016}$ . One feature of our hierarchical analysis is the ability to simultaneously fit the true eccentricity distribution while also improving measurements of impact parameter and planet size via importance sampling. We also apply hierarchical analysis to study sub-populations of our planet sample according to planet multiplicity, stellar metallicity, and planet radius. We find that planets in single-planet systems tend to have higher eccentricities than those in multi-planet systems, but our single-planet eccentricity distribution (with  $\bar{e} = 0.158^{+0.024}_{-0.020}$ ) is significantly lower than the findings of comparable studies of single *Kepler* planets (e.g.  $\bar{e} = 0.30^{+0.05}_{-0.05}$  from [Van Eylen et al. 2019](#)). In contrast to past findings, we also measure a broader eccentricity distribution for planets around metal-poor stars versus those around metal-rich stars. Additionally, we offer the first hierarchical analysis of eccentricities for a sample of transiting planets with heterogeneous sizes, allowing us to identify a broad eccentricity distribution for Jovian-size planets that is distinct from the lower distribution of

eccentricities among sub-Jovian-size planets in our sample. Hierarchical Bayesian modeling thus allows us to probe potential diverging pathways of exoplanet formation and evolution using large sets of noisy individual observations in a statistically robust manner.

## 7.2 Introduction

Orbital eccentricities are a crucial yet elusive clue to understanding the evolution of planets and planetary systems. Planets orbiting a given star are believed to form on near-circular, coplanar orbits, but the observed distribution of exoplanet eccentricities points towards a broad range of dynamical behaviors. A variety of dynamical processes have thus been proposed as potential means of exciting (see, e.g., [Goldreich & Sari 2003](#); [Chiang 2003](#); [Chatterjee et al. 2008](#); [Naoz 2016](#)) or damping (see, e.g., [Fabrycky & Tremaine 2007](#)) planetary orbital eccentricities. Several of such interactions may have even played out within the evolutionary history of our own Solar System (e.g. [Tsiganis et al. 2005](#); [Gomes et al. 2005](#)). However, further observational evidence is needed to better distinguish between the formation and evolution pathways that may have led to observed trends in the dynamical outcomes of planetary systems.

To better understand the breadth of orbital eccentricities, several studies have sought to estimate the underlying distribution of best-fit eccentricity  $e$  values for planets measured from both radial velocities (e.g. [Kipping 2013](#)) and transit photometry (e.g. [Van Eylen & Albrecht 2015](#); [Mills et al. 2019](#); [Xie et al. 2016](#)). These studies come to qualitatively similar conclusions regarding  $e$  and its relationship with other planetary system properties. Noteworthy trends include a general tendency towards near-circular orbits, lower  $e$  values in high-multiplicity systems, and a possible positive correlation between  $e$  and host star metallicity. However, deriving a general model to describe a population-level parameter is often not as simple as adding the best-fit results of all samples together. This issue is amplified for a variable like  $e$  that is generally loosely constrained or noisy on an individual

basis. In such cases, it is crucial to retain information regarding the complete structure of the posterior distribution of the modeled parameter when developing a generalized population-level model.

A tool that has recently gained popularity in exoplanet science for its use in addressing this issue is hierarchical Bayesian modeling (HBM). HBM enables simultaneous inference of the individual system parameters and the hyperparameters that govern the underlying distribution of the population by applying Bayes' Theorem at both levels. HBM was first introduced in the context of exoplanet eccentricities by [Hogg et al. \(2010\)](#) who used it to infer the eccentricity distribution of simulated planets from radial velocity (RV) model posteriors. The technique was later expanded upon by [Foreman-Mackey et al. \(2014\)](#) to be more robust and generalized. Other studies have since used HBM to measure the densities of small planets ([Rogers 2015](#)), the planetary mass–radius relationship ([Wolfgang et al. 2016](#)), and the rotation period distribution of stars ([Masuda et al. 2022](#)), demonstrating the power of this methodology for studying exoplanetary system properties.

For eccentricity modeling, [Kipping \(2013\)](#) points out that HBM is difficult to apply to RV-fit planets in practice due to the inhomogeneity of RV modeling and the lack of a centralized database for RV model posterior data. Instead, [Bowler et al. 2020](#) applied hierarchical modeling to direct imaging data to infer the underlying eccentricity distribution of 9 giant planets and 18 brown dwarfs. While their measurements of individual objects were noisy, the hierarchical analysis showed that the giant planets had a much smaller mean eccentricity than the brown-dwarfs – a clear sign of alternate formation pathways. However, the small sample size of homogeneously available data from either RV modeling or direct imaging can limit the generalizations that can be made from such hierarchical analyses.

On the other hand, transit photometry offers a significant number of available targets with a more complete description of the entire exoplanet population. Accordingly, [Van Eylen et al. \(2019\)](#) used HBM to estimate the underlying eccentricity distributions of single- and multi-

planet systems from *Kepler* photometry, identifying a clear distinction between these two populations. The multi-planet population from [Van Eylen et al. \(2019\)](#) could be described by a Rayleigh distribution with a mean eccentricity of  $\bar{e}_{\text{multi}} = 0.076_{-0.015}^{+0.013}$ , which is far less eccentric than their single-planet population with  $\bar{e}_{\text{single}} = 0.30_{-0.05}^{+0.05}$ . Their multi-planet population findings are consistent with past works (which also use *Kepler* planets) but their single-planet population findings differ from past works, with [Xie et al. \(2016\)](#) measuring  $\bar{e}_{\text{single}} = 0.32_{-0.02}^{+0.02}$  and [Mills et al. \(2019\)](#) measuring  $\bar{e}_{\text{single}} = 0.209_{-0.010}^{+0.016}$ . These previous studies, however, rely on summary statistics to describe their individual planet eccentricities whereas [Van Eylen et al. \(2019\)](#) incorporates the full posterior distributions via HBM analysis.

Here, we implement a hierarchical Bayesian analysis to infer the underlying distribution of eccentricity from transit photometry for a sample of 108 heterogeneous TESS planets. We draw these 69 single-planet systems and 16 multi-planet systems (hosting 39 planets) from the TESS-Keck Survey sample (TKS, [Chontos et al. 2022](#); [MacDougall et al. 2023a](#)), with individual  $e$  posterior distributions homogeneously measured via the methods described in [MacDougall, Gilbert, & Petigura 2023](#). We then estimate the underlying true eccentricity distribution of this sample following a framework similar to past studies: (1) implement the HBM statistical methodology from [Hogg et al. \(2010\)](#), (2) try describing the underlying eccentricity distribution using various distribution forms (see, e.g., [Van Eylen et al. 2019](#)), and (3) use Monte Carlo sampling to model the population-level hyperparameters.

We also re-analyze the parameter posterior distributions of individual TKS planets by re-weighting their samples via importance sampling, using the underlying  $e$  distribution measured from our HBM analysis as an updated prior. This leads to improved measurements of planet-to-star radius ratio  $R_p/R_\star$  and impact parameter  $b$ . Additionally, we investigate how the underlying  $e$  distribution changes as a function of other system properties, including planet multiplicity, stellar metallicity, and planet radius. We compare our results to

other works and consider how the details of sample selection and sample size can impact the outcomes of hierarchical eccentricity distribution modeling.

### 7.3 Framework for Hierarchical Bayesian Analysis

The objective of any Bayesian analysis is to infer the posterior probability density  $p(\theta)$  for some parameter(s)  $\theta$ . We typically wish to obtain an estimate of both the population-level distribution  $P(\theta)$  and individual  $p_i(\theta_i)$  for  $i$  independent targets. In practice, we are often presented with a large sample of noisy measurements leading to poor independent constraints on  $p_i(\theta_i)$ , which then makes inferring  $P(\theta)$  challenging. Instead, one can use hierarchical Bayesian modeling to leverage the large volume of individual measurements in order to improve estimates of both  $P(\theta)$  and  $p_i(\theta_i)$  (see Figure 7.1 for a graphical explanation), assuming that the  $i$  independent targets belong to the same underlying population (or a mixture of distinguishable sub-populations).

Using standard Bayesian inference (Figure 7.1, left), information flows in one direction: from the individual level  $p_i(\theta_i)$  to the population level  $P(\theta)$ . Under this paradigm, parameter posterior distributions for individual targets cannot simultaneously and self-consistently be constrained with the population, so noisy measurements lead to poorly constrained  $p_i(\theta_i)$  and the inferred posterior  $P(\theta)$  thus depends strongly on choice of prior  $\pi(\theta)$ . In reality, we do not infer  $P(\theta)$ , but rather  $P(\theta|\Lambda)$  for some choice of hyperparameters  $\Lambda$  used to specify the functional form of the prior  $\pi(\theta) = \pi(\theta|\Lambda)$ .

One might be tempted to circumvent these challenges by iteratively adopting the posterior  $P(\theta)$  as an updated prior, but this method, is incorrect as it double-conditions the data. The self-consistent solution for determining both  $p_i(\theta_i)$  and  $P(\theta)$  is hierarchical Bayesian modeling (Figure 7.1, right). The trick is to open up the prior  $\pi(\theta)$  itself to modeling by making the transformation  $\pi(\theta) \rightarrow P(\theta|\Lambda)$  and placing hyperpriors  $\pi(\Lambda)$  on  $\Lambda$ . This allows information to flow bidirectionally between individuals and populations. We then simultaneously and self

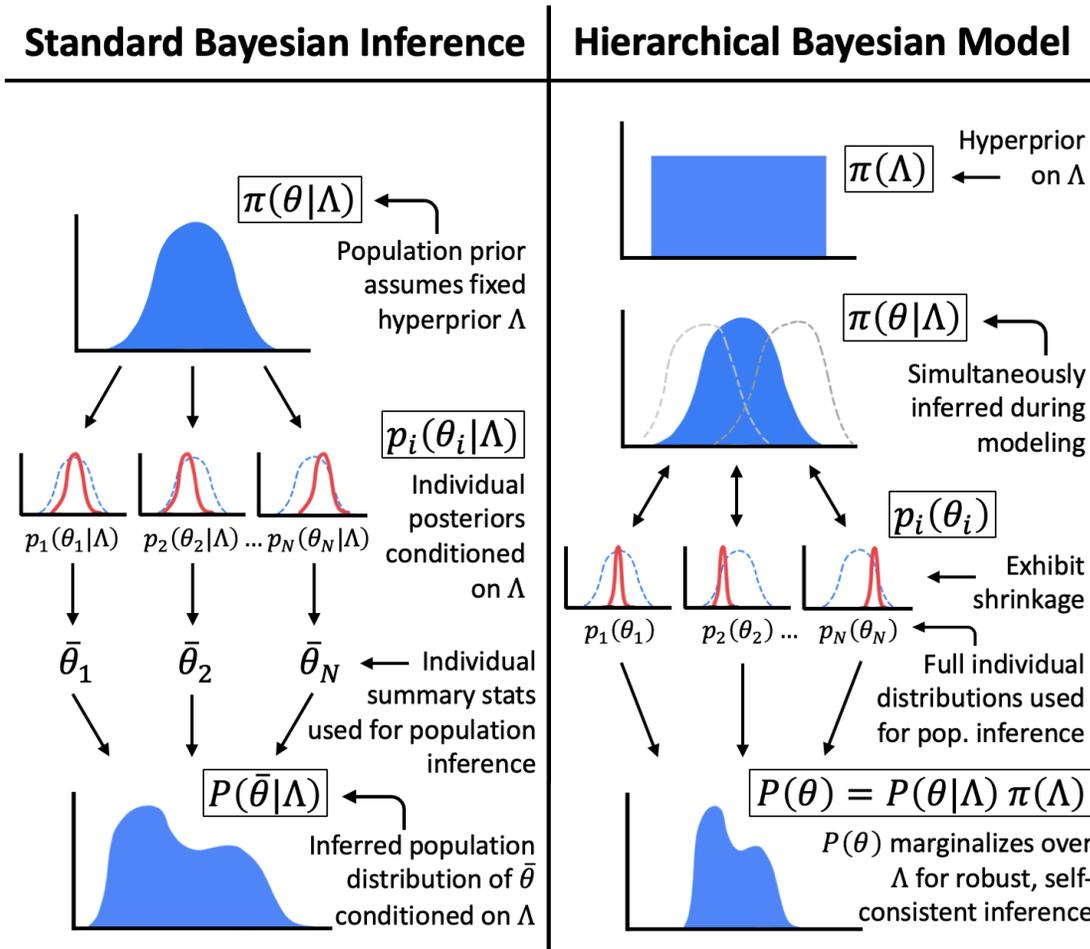


Figure 7.1: Illustrative example of the power of hierarchical modeling for some parameter(s)  $\theta$ , individual distributions  $p_i(\theta_i)$  for  $i$  targets, population level distribution  $P(\theta)$ , prior  $\pi$ , and hyperparameters  $\alpha$ . See §7.3 for full description.

consistently infer both the true population distribution  $P(\theta) = P(\theta|\Lambda)\pi(\Lambda)$  and more precise individual  $p_i(\theta_i)$ , which exhibit hierarchical shrinkage and are meaningfully constrained by learning from the population.

For our hierarchical Bayesian analysis of the TKS dataset, we adopt the HBM framework established by [Hogg et al. \(2010\)](#) and further developed by [Foreman-Mackey et al. \(2014\)](#) and [Van Eylen et al. \(2019\)](#). The likelihood function  $\mathcal{L}$  that describes the population-level distribution of the data  $d$  conditioned on the population-level hyperparameter(s)  $\Lambda$  is approximately given by

$$\mathcal{L}(d|\Lambda) \approx \prod_{n=1}^N \frac{1}{N} \sum_{k=1}^K \frac{P(e_{nk}|\Lambda)}{\pi(e_{nk})}, \quad (7.1)$$

where  $N$  is the number of planets being considered,  $K$  is the number of samples from the posterior eccentricity distribution,  $e_{nk}$  is the  $k^{\text{th}}$  random sample for each eccentricity distribution  $n$ ,  $P(e_{nk}|\Lambda)$  is the probability density of the population-level eccentricity distribution at  $e_{nk}$  and conditioned on the hyperparameter(s)  $\Lambda$ , and  $\pi(e_{nk})$  is the probability density of the prior probability distribution evaluated at  $e_{nk}$  ([Bowler et al. 2020](#)). Equation 7.1 is approximate for finite  $K$ , approaching an exact solution as  $K$  increases. The population-level probability density function  $P$  can take on any distribution form, including a Beta distribution with hyperparameters  $\Lambda = \{\alpha, \beta\}$  (e.g. [Kipping 2013](#); [Bowler et al. 2020](#)), Rayleigh distribution with  $\Lambda = \{\sigma_{\text{rayleigh}}\}$  (e.g. [Mills et al. 2019](#); [Xie et al. 2016](#)), or zero-mean positive-half-Gaussian distribution with  $\Lambda = \{\mu = 0, \sigma_{\text{gaussian}}\}$  (e.g. [Van Eylen et al. 2019](#)).

We implement this analytical framework using the PyMC statistical modeling software ([Salvatier et al. 2016b](#)), with sampling performed via the No-U-Turns Sampling method (NUTS; [Hoffman & Gelman 2011b](#)). For all HBM analysis performed in this work, we use 3,000 tuning steps with an additional 2,000 sampler draws and a target acceptance fraction of 0.95. This sampling process is performed via two sampler chains across two CPU cores, and we consider the process to be converged when the Gelman-Rubin statistic  $\hat{R}$  for these chains is  $\hat{R} < 1.01$  ([Gelman & Rubin 1992](#)).

Table 7.1: Select eccentricity distribution fits from the literature

Population(s)	Source	Distribution	Parameter(s)	Measurement(s)	Estimated $\bar{e}$
Heterogeneous Sample	A	Rayleigh	$\sigma_{\text{rayleigh, all}}$	0.1 – 0.2	0.125 – 0.25
	B	Beta	$\{\alpha_{\text{all}}, \beta_{\text{all}}\}$	$\{0.87^{+0.04}_{-0.04}, 3.03^{+0.17}_{-0.16}\}$	$0.222^{+0.013}_{-0.013}$
Planet Multiplicity	C	Rayleigh	$\sigma_{\text{rayleigh, multi}}$	$0.018^{+0.005}_{-0.004}$	$0.023^{+0.006}_{-0.005}$
	D	Rayleigh	$\sigma_{\text{rayleigh, single}}$	$0.26^{+0.02}_{-0.02}$	$0.32^{+0.02}_{-0.02}$
			$\sigma_{\text{rayleigh, multi}}$	$0.03^{+0.02}_{-0.03}$	$0.04^{+0.03}_{-0.04}$
	E	Rayleigh	$\sigma_{\text{rayleigh, single}}$	$0.167^{+0.013}_{-0.008}$	$0.209^{+0.016}_{-0.010}$
			$\sigma_{\text{rayleigh, multi}}$	$0.0355^{+0.012}_{-0.012}$	$0.044^{+0.015}_{-0.015}$
	F	Rayleigh	$\sigma_{\text{rayleigh, single}}$	$0.24^{+0.04}_{-0.04}$	$0.301^{+0.050}_{-0.050}$
			$\sigma_{\text{rayleigh, multi}}$	$0.061^{+0.010}_{-0.012}$	$0.076^{+0.013}_{-0.015}$
		Half-Gaussian	$\sigma_{\text{gaussian, single}}$	$0.32^{+0.06}_{-0.06}$	$0.255^{+0.048}_{-0.048}$
			$\sigma_{\text{gaussian, multi}}$	$0.083^{+0.015}_{-0.020}$	$0.066^{+0.012}_{-0.016}$
		Beta	$\{\alpha_{\text{single}}, \beta_{\text{single}}\}$	$\{1.58^{+0.59}_{-0.93}, 4.4^{+1.8}_{-2.2}\}$	$0.264^{+0.108}_{-0.150}$
$\{\alpha_{\text{multi}}, \beta_{\text{multi}}\}$			$\{1.52^{+0.50}_{-0.85}, 29^{+9}_{-17}\}$	$0.050^{+0.021}_{-0.038}$	

Note. — A: [Moorhead et al. \(2011\)](#); distribution fit to best-fit  $e$  values measured from transit photometry for 104 *Kepler* Objects of Interest with cool host stars ( $T_{\text{eff}} < 5100$  K). B: [Kipping \(2013\)](#); distribution fit to best-fit  $e$  values measured from RV data for 396 planets. C: [Hadden & Lithwick \(2014\)](#); distribution fit to best-fit  $e$  values measured from transit-timing variations for 139 *Kepler* planets. D: [Xie et al. \(2016\)](#); distribution fit via forward modeling of best-fit  $e$  measurements from transit photometry for 368 single planets and 330 planets from multi-planet systems. E: [Mills et al. \(2019\)](#); distribution fit to best-fit  $e$  values measured from *Kepler* transit data for a sample of 439 single planets and 870 planets from multi-planet systems. F: [Van Eylen et al. \(2019\)](#); distributions fit via hierarchical analysis of full posterior distributions measured from *Kepler* transit photometry for a sample of 51 single planets and 66 planets from multi-planet systems.

## 7.4 A Hierarchical Analysis of the Full TKS Sample

The TKS sample is composed of 16 multi-planet systems hosting 39 planets along with another 69 single-planet systems, displaying heterogeneous radii, periods, and orbital architectures (see [Chontos et al. 2022](#) for sample selection criteria). These planets are hosted mostly by inactive, solar-like, main-sequence stars with a wide range of metallicities and  $3700 \lesssim T_{\text{eff}} \lesssim 6500$ . Both stellar characterization and lightcurve modeling were performed by [MacDougall et al. 2023a](#) for all targets in the TKS sample, producing a set of homogeneously-measured posterior distributions for all transit parameters including eccentricity. As noted in [MacDougall et al. 2023a](#), these  $e$  posteriors are sampled post-modeling via importance sampling using an uninformative uniform prior on  $e$  from 0 to 1 and also using stellar densities measured with `isoclassify` ([Huber et al. 2017](#); [Berger et al. 2020b](#)) with typical uncertainties of  $\sim 4\text{--}8\%$ .

We can apply the hierarchical Bayesian modeling procedure described in §7.3 to measure the underlying eccentricity distribution of the 108 planets in the TESS-Keck Survey sample. We draw our  $e$  posterior distributions for individual planets in this sample from [MacDougall et al. 2023a](#). We take these posterior distributions in their entirety as inputs to our HBM analysis. Following similar studies in the literature, we use our hierarchical model to try fitting the population-level  $e$  distribution with various distribution shapes, including a Rayleigh distribution, zero-mean half-Gaussian distribution, and Beta distribution. Here, we use this method to measure the  $e$  distribution of the full TKS sample. Later, in §7.5, we investigate how  $e$  may vary as a function of other system parameters by analyzing various sub-populations of this sample.

### 7.4.1 Measuring the true eccentricity distribution of TKS planets

We analyze the population-level eccentricity distribution for the full TKS sample first without considering separate sub-populations. We apply the analytical framework described

Table 7.2: Eccentricity distribution fits from this work

Population(s)	Distribution	Parameter(s)	Measurement(s)	Estimated $\bar{e}$
Full TKS Sample	Rayleigh	$\sigma_{\text{rayleigh, all}}$	$0.106^{+0.014}_{-0.013}$	$0.135^{+0.018}_{-0.016}$
	Half-Gaussian	$\sigma_{\text{gaussian, all}}$	$0.149^{+0.026}_{-0.021}$	$0.119^{+0.021}_{-0.017}$
	Beta	$\alpha_{\text{all}}, \beta_{\text{all}}$	$\{0.360^{+0.097}_{-0.067}, 4.03^{+1.49}_{-1.09}\}$	$0.082^{+0.035}_{-0.025}$
Planet Multiplicity	Rayleigh	$\sigma_{\text{rayleigh, single}}$	$0.126^{+0.019}_{-0.016}$	$0.158^{+0.024}_{-0.020}$
		$\sigma_{\text{rayleigh, multi}}$	$0.038^{+0.018}_{-0.014}$	$0.048^{+0.023}_{-0.018}$
	Half-Gaussian	$\sigma_{\text{gaussian, single}}$	$0.176^{+0.030}_{-0.025}$	$0.140^{+0.024}_{-0.020}$
		$\sigma_{\text{gaussian, multi}}$	$0.058^{+0.030}_{-0.020}$	$0.046^{+0.024}_{-0.016}$
	Beta	$\alpha_{\text{single}}, \beta_{\text{single}}$	$\{0.355^{+0.128}_{-0.078}, 3.32^{+1.42}_{-1.01}\}$	$0.097^{+0.049}_{-0.033}$
		$\alpha_{\text{multi}}, \beta_{\text{multi}}$	$\{0.580^{+1.006}_{-0.264}, 12.38^{+20.08}_{-7.32}\}$	$0.045^{+0.102}_{-0.032}$
Stellar Metallicity	Rayleigh	$\sigma_{\text{rayleigh, met-poor}}$	$0.148^{+0.033}_{-0.027}$	$0.185^{+0.041}_{-0.034}$
		$\sigma_{\text{rayleigh, met-rich}}$	$0.086^{+0.015}_{-0.012}$	$0.108^{+0.019}_{-0.015}$
	Half-Gaussian	$\sigma_{\text{gaussian, met-poor}}$	$0.196^{+0.058}_{-0.042}$	$0.156^{+0.046}_{-0.034}$
		$\sigma_{\text{gaussian, met-rich}}$	$0.123^{+0.028}_{-0.022}$	$0.098^{+0.022}_{-0.018}$
	Beta	$\alpha_{\text{met-poor}}, \beta_{\text{met-poor}}$	$\{0.309^{+0.200}_{-0.092}, 2.61^{+2.19}_{-1.18}\}$	$0.106^{+0.100}_{-0.051}$
		$\alpha_{\text{met-rich}}, \beta_{\text{met-rich}}$	$\{0.392^{+0.149}_{-0.088}, 4.71^{+2.40}_{-1.49}\}$	$0.077^{+0.045}_{-0.028}$
Planet Radius	Rayleigh	$\sigma_{\text{rayleigh, sub-jov}}$	$0.085^{+0.016}_{-0.012}$	$0.107^{+0.020}_{-0.015}$
		$\sigma_{\text{rayleigh, jov}}$	$0.155^{+0.038}_{-0.030}$	$0.194^{+0.048}_{-0.038}$
	Half-Gaussian	$\sigma_{\text{gaussian, sub-jov}}$	$0.119^{+0.026}_{-0.021}$	$0.095^{+0.021}_{-0.017}$
		$\sigma_{\text{gaussian, jov}}$	$0.217^{+0.072}_{-0.048}$	$0.173^{+0.057}_{-0.038}$
	Beta	$\alpha_{\text{sub-jov}}, \beta_{\text{sub-jov}}$	$\{0.368^{+0.133}_{-0.079}, 4.93^{+2.43}_{-1.59}\}$	$0.069^{+0.040}_{-0.025}$
		$\alpha_{\text{jov}}, \beta_{\text{jov}}$	$\{0.318^{+0.217}_{-0.114}, 1.98^{+1.94}_{-1.02}\}$	$0.139^{+0.143}_{-0.075}$

Note. — Eccentricity distribution fits from this work for various populations drawn from the TKS planet sample. Metal-poor ( $[\text{Fe}/\text{H}] < 0$ ) versus metal-rich ( $[\text{Fe}/\text{H}] > 0$ ) and sub-Jovian-size ( $R_{\text{p}} < 8R_{\oplus}$ ) versus Jovian-size ( $R_{\text{p}} > 8R_{\oplus}$ ) sub-populations based on characterizations drawn from MacDougall et al. 2023a.

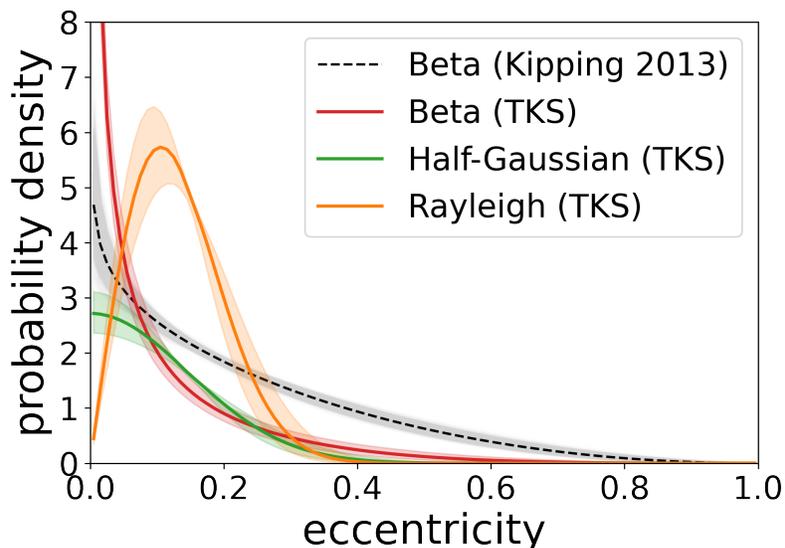


Figure 7.2: Various distribution types fit to the population-level eccentricity distribution of the TKS sample using hierarchical Bayesian modeling. A commonly cited Beta distribution (Kipping 2013) is shown for reference (black/grey). Our distribution fits suggest that the TKS sample has a lower distribution of eccentricities than the sample used by Kipping 2013, possibly attributed to differences in observing techniques or sample selection biases.

in §7.3 to fit our observed eccentricities using a simple Rayleigh distribution with a single hyperparameter  $\sigma_{\text{rayleigh}}$ , finding  $\sigma_{\text{rayleigh}} = 0.106^{+0.014}_{-0.013}$ . This distribution fit implies that planets in our sample tend to have low orbital eccentricities, with a mean eccentricity of  $\bar{e}_{\text{rayleigh}} = 0.135^{+0.018}_{-0.016}$  among the TKS population. Here, we prefer to describe the population-level eccentricities using a Rayleigh distribution because it is commonly used in the literature and also because it describes the magnitude of a two-dimensional random vector where each component is independent and normally distributed, approximately fitting the description of  $e$  in its relationship to argument of periastron  $\omega$ .

However, various other distribution forms have also been considered in the literature when modeling population-level eccentricity distributions since the true topology is loosely constrained. For completeness, we also try fitting both a zero-mean positive-half-Gaussian

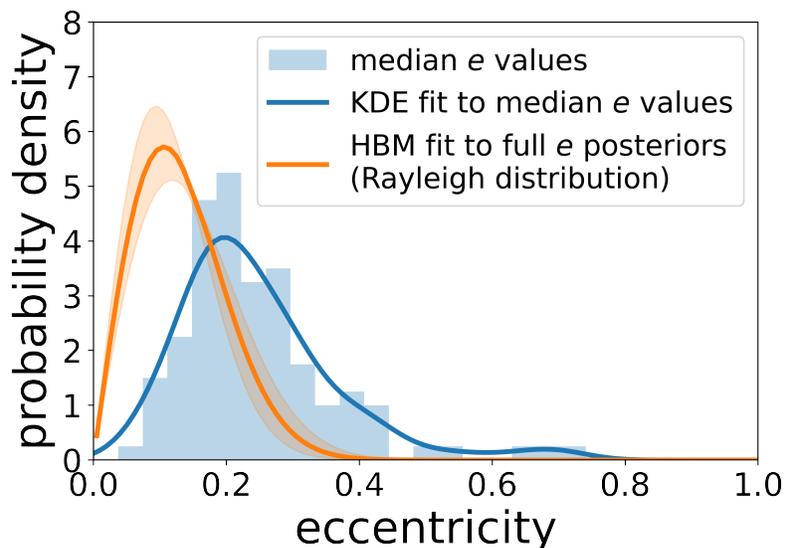


Figure 7.3: Comparison between fitting the TKS population-level eccentricity distribution using hierarchical analysis (orange) versus estimating the underlying distribution from summary statistics (blue). The histogram of median  $e$  values for each TKS planet is shown with a Gaussian kernel density estimate fit to this distribution with a bandwidth measured via Scott’s Rule (Scott 1992).

distribution (e.g. Van Eylen et al. 2019) and a Beta distribution (e.g., Kipping 2013; Van Eylen et al. 2019; Bowler et al. 2020) to our TKS eccentricities. While each fit uses a different distribution shape to describe the eccentricities, they all achieve similar  $\bar{e}$  estimates for the population within  $\sim 1.5\sigma$  of each other. These results are summarized in Table 7.2 and shown in Figure 7.2. We could also consider mixture models or non-parametric models, but such considerations are beyond the scope of this analysis. We proceed with our analysis focusing on the results from the physically-motivated Rayleigh distribution fit to our population-level eccentricities.

We compare the results of our hierarchical analysis of TKS eccentricities to a simple analysis of these same eccentricities based on individual planet summary statistics. We take the median  $e$  values measured for each TKS planet and calculate the mean eccentricity of the

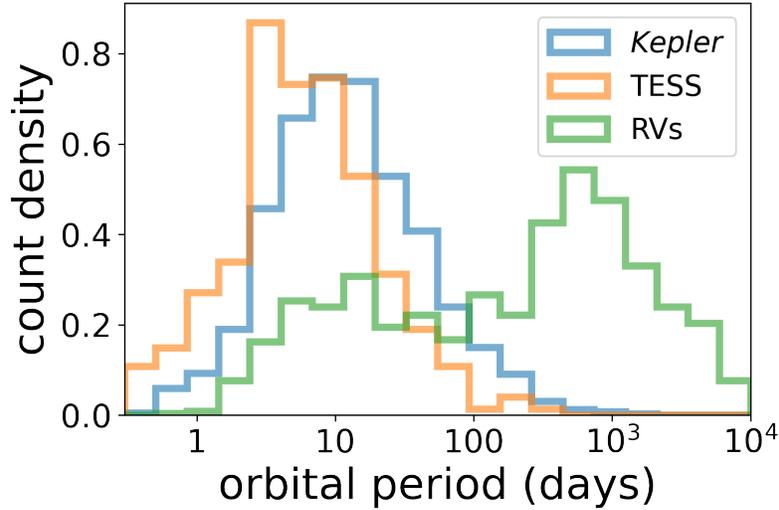


Figure 7.4: Distribution of orbital periods (from NASA Exoplanet Archive on 22 May 2023; [Akeson et al. 2013a](#)) for planets discovered via transit detection (*Kepler*, blue; TESS, orange) or radial velocity detection (green). Generally, planets discovered via RVs have significantly longer orbital periods than transiting planets. Also, *Kepler* planets tend to have longer periods than TESS planets, with 15<sup>th</sup>-to-85<sup>th</sup> percentile period ranges of  $P_{Kepler} \approx 3.5 - 45$  days versus  $P_{TESS} \approx 2 - 20$  days.

histogram of these summary statistics, measuring  $\bar{e}_{\text{histogram}} \approx 0.25$ . If we use the mean  $e$  values from individual planet posteriors rather than medians, this estimate increases to  $\bar{e}_{\text{histogram}} \approx 0.31$ . The underlying eccentricity distribution estimated from summary statistics is thus significantly broader than the distribution that we measure from our hierarchical analysis ( $\bar{e}_{\text{rayleigh}} \approx 0.14$ ). We show this comparison in [Figure 7.3](#) where the two distributions clearly differ. The hierarchical Bayesian modeling approach retains all posterior information of  $e$  from the individual targets which leads to a more reliable interpretation of the underlying dynamics.

Our population-level  $\bar{e}_{\text{rayleigh}}$  estimate is marginally consistent with that of [Moorhead et al. \(2011\)](#) measured from *Kepler* planet transits ( $\bar{e}_{\text{rayleigh}} \approx 0.125 - 0.25$ ), but our estimate is

significantly lower than the result of [Kipping \(2013\)](#) measured from RVs ( $\bar{e}_{\text{beta}} \approx 0.222^{+0.013}_{-0.013}$ ) – where both studies considered heterogeneous planet samples. A possible explanation for this disagreement is that TESS planets have an inherently lower period distribution than either *Kepler* planets or planets found via RVs (see [Figure 7.4](#)). The TESS mission generally detects fewer longer-period planets than other surveys due to its short observing baseline of  $<30$  days per observing sector. This may lead to lower observed eccentricities among the TESS planet population since short-period planets tend to have low eccentricities due to tidal circularization (e.g. [Matsumura et al. 2010](#)).

#### 7.4.2 Improving $R_p/R_*$ measurements via HBM-informed importance sampling

The eccentricities that we use as inputs to our HBM analysis were produced through importance sampling with an uninformative uniform prior on  $e$  (see [MacDougall et al. 2023a](#)). This post-model importance-weighting process also allows one to easily go back and re-weight parameter posterior distributions with an updated  $e$  prior. This is useful for post-hoc re-analysis of transit model results following a hierarchical evaluation of the population-level eccentricity distribution. We can thus use our HBM-informed  $e$  distribution from [§7.4.1](#) to update the individual target parameter posterior distributions from [MacDougall et al. 2023a](#) via importance sampling.

We use a Rayleigh distribution with  $\sigma_{\text{rayleigh}} = 0.106$  as our new prior to re-weight the TKS sample posteriors, drawing this prior from our hierarchical fit to the population-level eccentricity distribution. We re-weight the parameter posterior distributions of all TKS planets individually via importance sampling using this prior, then we calculate updated summary statistics for all planet parameters. Since the updated  $e$  prior is a Rayleigh distribution, it enforces zero posterior weight at  $e = 0$  and low weight at high  $e$  values. We prefer the Rayleigh distribution here over the half-Gaussian or Beta distributions because of its simplicity, physical motivation, and common usage in the literature.

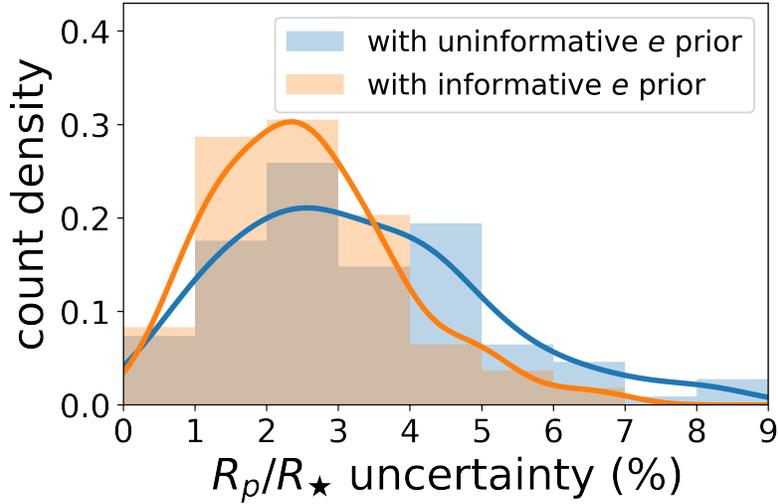


Figure 7.5: Distribution of the fractional uncertainty measurements  $\sigma(R_p/R_\star)$  for all TKS planets using an uninformative  $e$  prior (blue) versus an  $e$  prior informed by our hierarchical analysis of the population-level  $e$  distribution (orange). The HBM-informed  $e$  prior leads to more precise radius ratios with narrower distribution of  $\sigma(R_p/R_\star)$  uncertainties. Gaussian kernel density estimates are shown fit to each distribution.

Post re-weighting, we find that the newly-constrained (lower) eccentricities generally lead to better constrained (higher) impact parameters and, subsequently, more precise (higher) radius ratios  $R_p/R_\star$ . Following this re-weighting, we re-measure radius ratios for all TKS planets and obtain more precise radius measurements than those from MacDougall et al. 2023a. This can be seen in Figure 7.5 where we compare the before-and-after distributions of fractional radius ratio uncertainties  $\sigma(R_p/R_\star)$ . We measure a 15<sup>th</sup>-to-85<sup>th</sup> percentile uncertainty range of  $\sigma(R_p/R_\star) \sim 1.3 - 3.8\%$  for our re-weighted planet population. These uncertainties are thus more precise than our initial radius ratios measured with an uninformative  $e$  prior, where the corresponding 15<sup>th</sup>-to-85<sup>th</sup> percentile range of uncertainties was broader:  $\sigma(R_p/R_\star) \sim 1.4 - 4.9\%$ . The largest  $R_p/R_\star$  uncertainty amongst our sample subsequently shrank from 8.3% to 6.9% after our informative re-weighting.

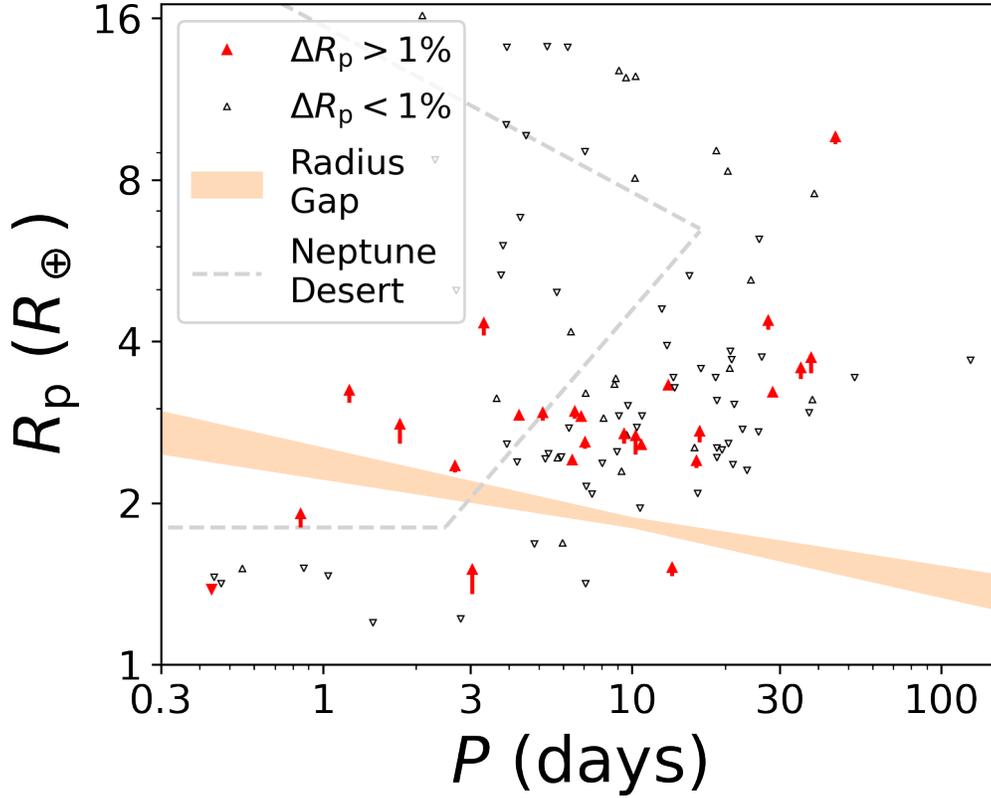


Figure 7.6: Period-radius distribution of TKS planets, with markers shown at the radius values calculated from our newly re-weighted  $R_p/R_\star$  distributions. Planets with radii that shifted by  $|\Delta R_p| > 1\%$  are shown as red triangles, and all other planets in the TKS sample are shown as black triangles. The triangles are pointed up or down depending on the direction of the relative shift  $\Delta R_p$ . The red triangles form part of an arrow that shows the full extent of the shift  $\Delta R_p$ , with the most substantial shifts occurring among smaller planets. We also show noteworthy areas of the period-radius parameter space for reference, including the Radius Gap (Fulton et al. 2017a; Petigura et al. 2022) and the Hot Neptune Desert (Mazeh et al. 2016).

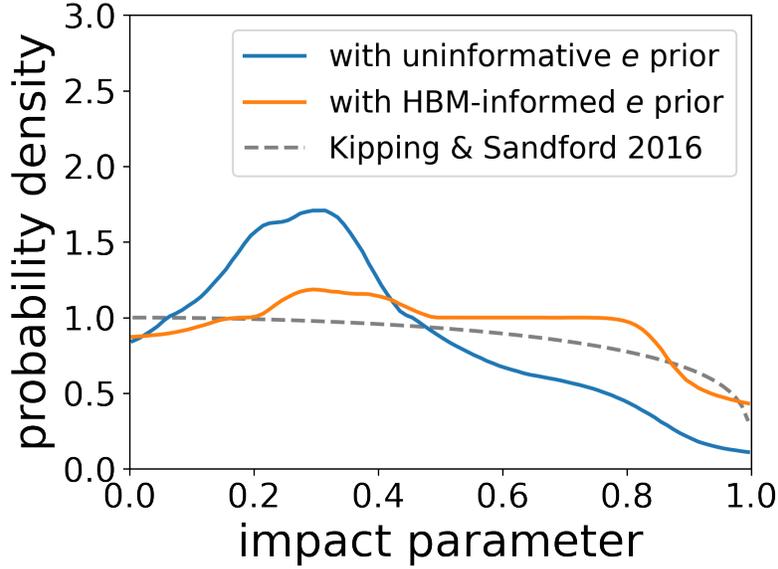


Figure 7.7: Best-fit hierarchical model fit to the shape of the underlying distribution of  $b$  values among TKS planets both before (blue) and after (orange) applying our HBM-informed  $e$  prior. We measure both curves with a hierarchical Bayesian model fit to the underlying distribution histogram via a Gaussian process model with a Matérn-3/2 kernel. The  $b$  distribution achieved via our updated prior is closer to what is expected from orbital geometry and observational bias (grey; [Kipping & Sandford 2016](#)). We show only the best-fit solutions here for simplicity, noting that the  $1\sigma$  uncertainties on these probability distributions are significant and thus this figure is only intended to convey qualitative results.

Overall,  $\sim 20\%$  of TKS planets experienced a shift of  $|\Delta R_p/R_\star| > 1\%$ , highlighted in [Figure 7.6](#) where we show the significant corresponding shift in  $\Delta R_p$  in some cases. The largest  $\Delta R_p/R_\star$  shifts mostly occurred for smaller planets, like a  $\sim 10\%$  increase in planet size for the super-Earth TOI-554.01. Such planets generally have lower signal-to-noise transits with initially overestimated  $e$  values and underestimated  $b$  values due to the inherent degeneracy between  $b$ ,  $e$ , and argument of periastron  $\omega$  (see [MacDougall et al. 2023b](#)). By including an informative eccentricity prior, we are able to minimize this degeneracy to obtain more accurate radius measurements for smaller planets.

In support of our updated  $R_p/R_\star$  constraints, we also find that our informative  $e$  prior yields more realistic impact parameter measurements. Before applying our updated  $e$  prior, the underlying distribution of impact parameter values among our sample was skewed in favor of lower impact parameters due to the  $b$ - $e$ - $\omega$  degeneracy (Figure 7.7). After implementing our HBM-informed  $e$  prior, we find that our impact parameter distribution is closer to the true distribution expected from orbital geometry and observational bias (see Kipping & Sandford 2016). We estimate the  $b$  distribution in both cases using a hierarchical Bayesian model fit to the binned underlying distribution via a Gaussian process model with a Matérn-3/2 kernel.

## 7.5 A Hierarchical Analysis of Sub-populations of the TKS sample

### 7.5.1 Planet Multiplicity-based Sub-populations

We apply our HBM fitting procedure to measure the underlying eccentricity distribution for various sub-populations of the TKS sample, starting with a comparison of systems with multiple transiting planets versus those with a single transiting planet. This is a common basis for separating a planet sample into sub-populations (see, e.g., Van Eylen et al. 2019; Mills et al. 2019) because it allows us to consider the dynamical differences that arise from the different formation and evolution pathways that planets undergo with or without additional observed companions. It has previously been observed that multi-planet systems tend to have smaller eccentricities than single planets (e.g. Wright et al. 2009; Limbach & Turner 2015; Xie et al. 2016), typically attributed to dynamical interactions with remnant protoplanetary disk material (e.g. Dawson et al. 2016) or between planetesimals (e.g. Mulders et al. 2020) which preferentially damp orbital eccentricities over time in high-multiplicity systems.

Given the differences identified in previous studies, we seek to perform a similar comparison of the population-level eccentricity distributions of the TKS single-planet ( $N = 69$ ) and multi-planet ( $N = 39$ ) populations through our hierarchical Bayesian framework. As we did in §7.4.1, we try fitting three different distribution shapes to each sub-population: Rayleigh,

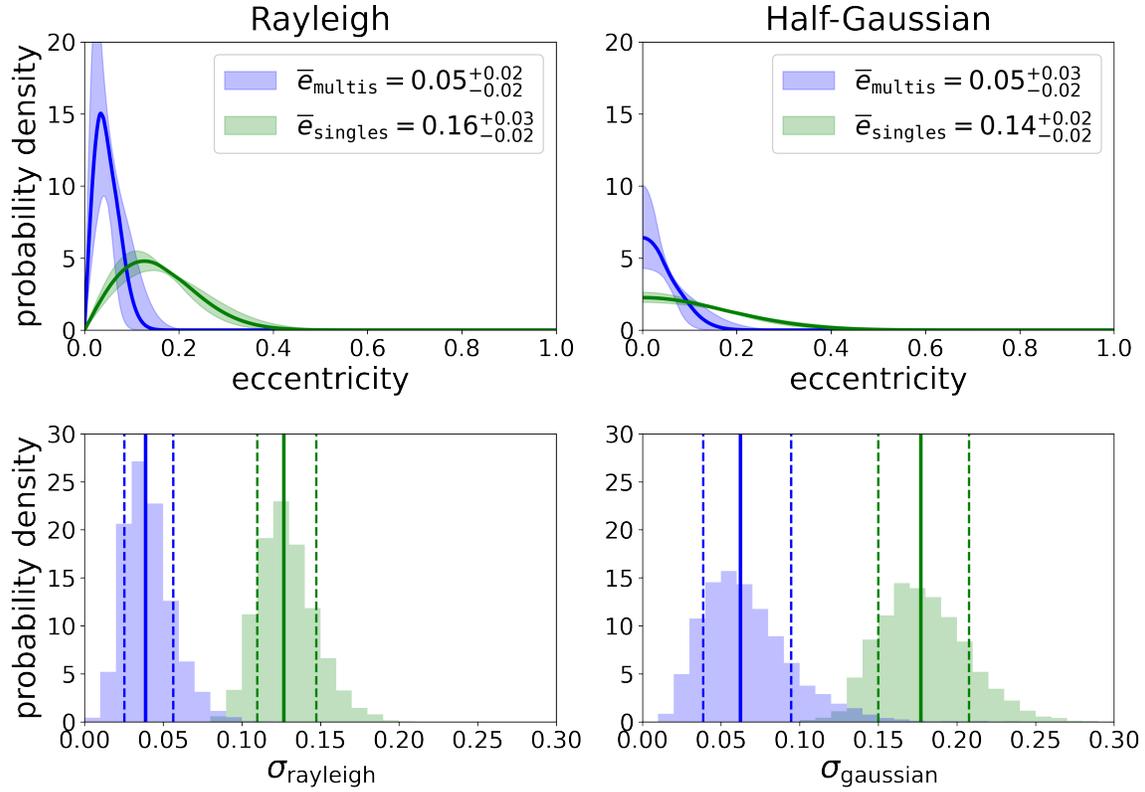


Figure 7.8: Rayleigh distribution (left) and half-Gaussian distribution (right) fit to multiplicity-based sub-populations of the TKS sample. We show the 15<sup>th</sup>–to–85<sup>th</sup> percentile ranges for our distribution fits (top), with the 50<sup>th</sup> percentile highlighted by the solid line. We also show the distributions of the hyperparameters that we measure for each sub-population (bottom), highlighting their 15<sup>th</sup>, 50<sup>th</sup>, and 85<sup>th</sup> percentiles as vertical lines. The mean eccentricity  $\bar{e}$  for each distribution is shown in the legend. We find that planets with observed transiting companions (blue) have a significantly higher distribution of eccentricities than those without observed transiting companions (green), consistent with the results of past studies.

half-Gaussian, and Beta distributions. We measure the hyperparameters for each distribution fit to our two sub-populations and summarize these results in Table 7.2. We also show the measured Rayleigh and half-Gaussian distributions along with their hyperparameter distributions in Figure 7.8, but we mainly focus on the Rayleigh fit throughout our analysis. Overall, our findings suggest that single-planet systems in the TKS sample display a higher eccentricity distribution than our multi-planet systems – consistent with past studies.

When we fit a Rayleigh distribution to our eccentricities, we measure  $\sigma_{\text{rayleigh, multi}}$  consistent with Mills et al. (2019) but we measure an inconsistent  $\sigma_{\text{rayleigh, single}}$ . Those authors used a large sample size and high-quality data, but their results are based on best-fit  $e$  values rather than complete posterior distributions. We also compare our Rayleigh distribution results with those from Van Eylen et al. (2019), again finding consistent  $\sigma_{\text{rayleigh, multi}}$  values but inconsistent  $\sigma_{\text{rayleigh, single}}$  values (see Figure 7.9). This inconsistency persists even when we apply cuts to our sample to mimic those used by Van Eylen et al. (2019) (i.e.  $R_p < 6R_{\oplus}, P > 5d$ ). We also consider a half-Gaussian fit to our eccentricity distributions, following Van Eylen et al. (2019), and again measure consistent  $\sigma_{\text{gaussian, multi}}$  but inconsistent  $\sigma_{\text{rayleigh, single}}$ .

These differences suggest that single-planet systems in the TKS sample display lower eccentricities than single-planet populations examined in similar studies. A possible explanation for this disagreement is that there are inherent difference between the TESS planet population (our study) and the *Kepler* population (all similar past studies). The TESS mission tends to detect fewer long-period planets than *Kepler* due to its significantly shorter observing baseline (see Figure 7.4), which may lead to lower observed eccentricities among TESS planets due to tidal effects. Alternatively, the short observing baseline of TESS may also limit our ability to detect additional companions that have longer orbital periods. This may lead to some true multi-planet systems being mistakenly considered as single-planet systems in our analysis, polluting our single-planet population with low- $e$  planets from true multi-planet systems.

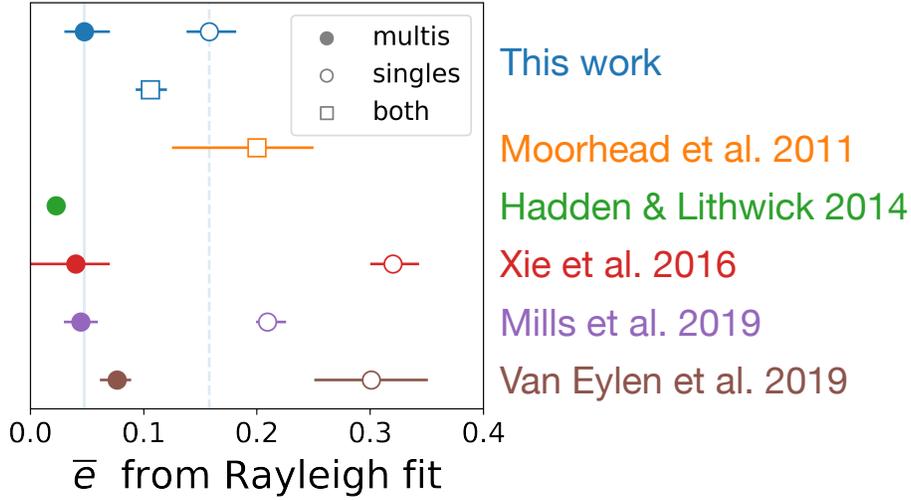


Figure 7.9: Mean eccentricity  $\bar{e}$  from various Rayleigh distribution fits to exoplanet eccentricities, drawn from both this work and the literature (Moorhead et al. 2011; Hadden & Lithwick 2014; Xie et al. 2016; Mills et al. 2019; Van Eylen et al. 2019). Open circles are from multi-planet fits, filled circles are from single-planet fits, and open square points are from fits to a general planet population. Faint vertical lines are used as a visual reference to compare literature results to the maximum-likelihood fits from this work. The mean eccentricity is related to hyperparameter  $\sigma_{\text{rayleigh}}$  via  $\bar{e} = \sigma_{\text{rayleigh}} \sqrt{\pi/2}$  (Jurić & Tremaine 2008).

### 7.5.2 Stellar Metallicity-based Sub-populations

We also consider metallicity-based sub-populations of the TKS sample in our hierarchical Bayesian analysis. We split the sample into two groups: planets orbiting metal-poor stars ( $[\text{Fe}/\text{H}] < 0$ ;  $N = 31$ ) versus planets orbiting metal-rich stars ( $[\text{Fe}/\text{H}] \gtrsim 0$ ;  $N = 77$ ), as done by Mills et al. (2019). Those authors modeled the distribution of best-fit  $e$  values from transit data and found that planets around metal-rich stars tend to have higher eccentricities than those around metal-poor stars, supported by An et al. (2023) using transit data. This relationship may be linked to the early evolution of planetary systems when metal-rich stars would likely have had more solid protoplanetary disk material to form more planetesimals which could then excite each other’s orbits (e.g. Dawson & Murray-Clay 2013).

We split our sample into these two sub-populations and fit the eccentricities of each group separately following the same hierarchical framework as before (see Table 7.2 for a summary of these results). We find a statistically significant difference between the  $e$  distributions of metal-poor versus metal-rich populations based on both  $\sigma_{\text{rayleigh}}$  and  $\sigma_{\text{gaussian}}$  (see Figure 7.10). However, contrary to past studies, our results show that planets around metal-rich stars in the TKS sample have *lower* eccentricities while those around metal-poor stars have *higher* eccentricities.

One possible explanation for this discrepancy between ours and past studies could stem from loosely constrained metallicity measurements that mistakenly fall on the wrong side of our binary sub-population grouping. If we instead require that a host star’s metallicity measurement from 15<sup>th</sup>-to-85<sup>th</sup> percentile falls entirely above or below  $[\text{Fe}/\text{H}] = 0$ , we end up with  $N = 20$  planets around metal-poor stars and  $N = 59$  planets around metal-rich stars, with another 29 planets orbiting stars with more moderate Sun-like metallicities. For this exaggerated set of metallicity-based sub-populations, we more confidently recover the same result as before. We also repeat this process with metallicities drawn from Berger et al. (2023) rather than MacDougall et al. 2023a, again producing similar results.

It is also possible that the specific selection criteria of the TKS sample (Chontos et al. 2022) may have led to a planet population that defies the broader metallicity-eccentricity relationship. On the other hand, Van Eylen et al. (2019) found no evidence for *any* metallicity-eccentricity relationship through their own hierarchical Bayesian analysis – the only other such analysis to-date in the literature. In summary, the relationship between metallicity and eccentricity remains loosely constrained, but our sample demonstrates that planetary orbital eccentricities may be anti-correlated to host star metallicities for TESS targets.

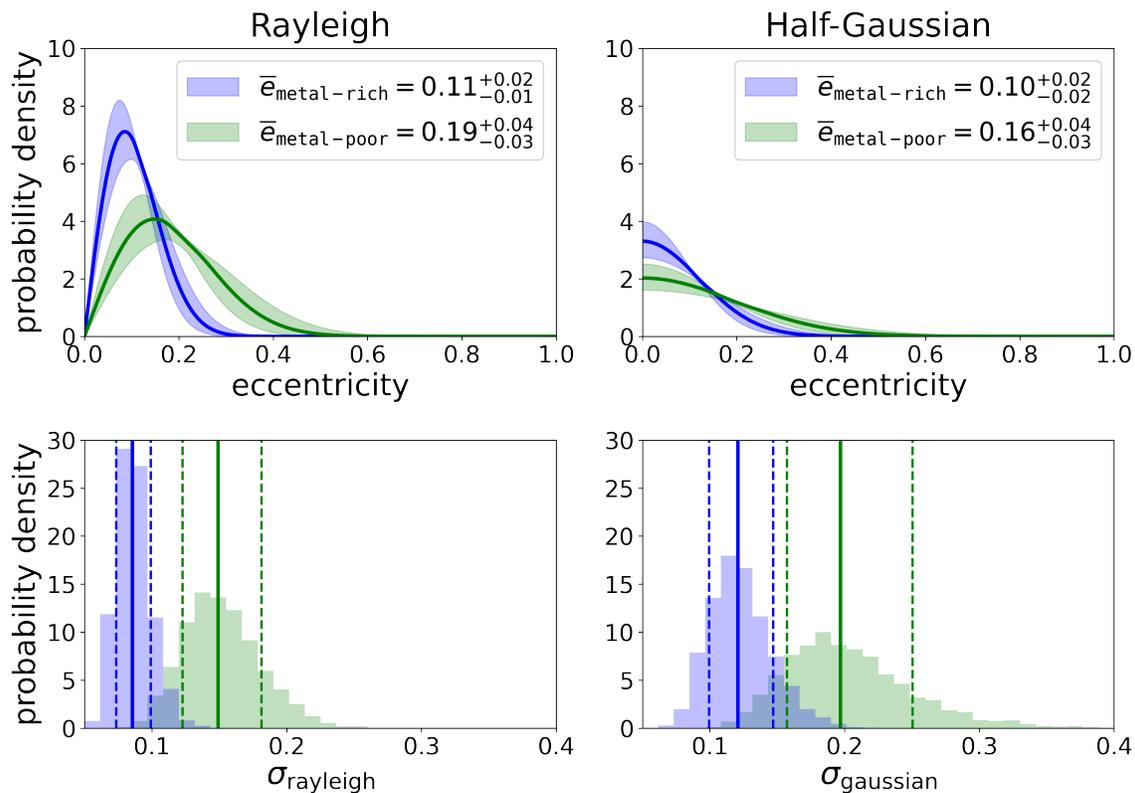


Figure 7.10: Rayleigh and half-Gaussian distribution fits to metallicity-based subpopulations of the TKS sample, with the metal-rich population ( $[\text{Fe}/\text{H}] > 0$ ) shown in blue and the metal-poor population ( $[\text{Fe}/\text{H}] < 0$ ) shown in green. We find that planets orbiting metal-poor stars tend to have a higher distribution of eccentricities than those orbiting metal-rich stars, counter to the results of past studies.

### 7.5.3 Planet Radius-based Sub-populations

We also use our hierarchical Bayesian analysis to measure the true eccentricity distribution of sub-Jovian-size TKS planets ( $R_p < 8R_\oplus$ ) as compared to that of Jovian-size TKS planets ( $R_p \gtrsim 8R_\oplus$ ). There is substantial evidence indicating that smaller planets tend to have lower eccentricities while giant planets have a much broader eccentricity distribution (see, e.g., [Wright et al. 2009](#); [Mayor et al. 2011](#); [Mills et al. 2019](#)). Several evolutionary pathways have been proposed to explain this  $e$  divergence based on planet size, including strong dynamical interactions between multiple planetary bodies that result in orbital excitation for giant planets while small eccentric planets are destroyed or ejected (e.g. [Bitsch et al. 2020](#)). There are no radius-based hierarchical analyses of exoplanet eccentricities in the literature, so we offer the first HBM-based investigation of this radius-eccentricity trend.

As before, we split our sample into two sub-populations: small planets ( $N = 93$ ) versus large planets ( $N = 15$ ), separated at  $R_p = 8R_\oplus$ . We try fitting several different distribution shapes to the eccentricities of both sub-populations and consistently find that giant planets in our sample have a broader underlying eccentricity distribution than sub-Jovian-size planets. Our Rayleigh and half-Gaussian distribution fits yield hyperparameter measurements that significantly differentiate between the two sub-populations (see [Figure 7.11](#) and [Table 7.2](#)).

We try repeating this process with different radius values separating the two sub-populations (instead of  $8R_\oplus$ ), including all integer values from  $2-7R_\oplus$ . We find that the measured eccentricity distributions of the two populations start to diverge significantly when the separation radius reaches  $\gtrsim 6R_\oplus$ . Including a third sub-population (i.e. small, medium, and large planets) helps to distinguish more strongly between small planets and large planets, but the middle population always has significant overlap with one or both of the outer populations. These different ways of splitting up our sample into sub-populations also lead to different sub-sample sizes which can impact the statistical significance these supplementary tests' results. Generally, our findings strongly support a differentiation between the dynamical outcomes of

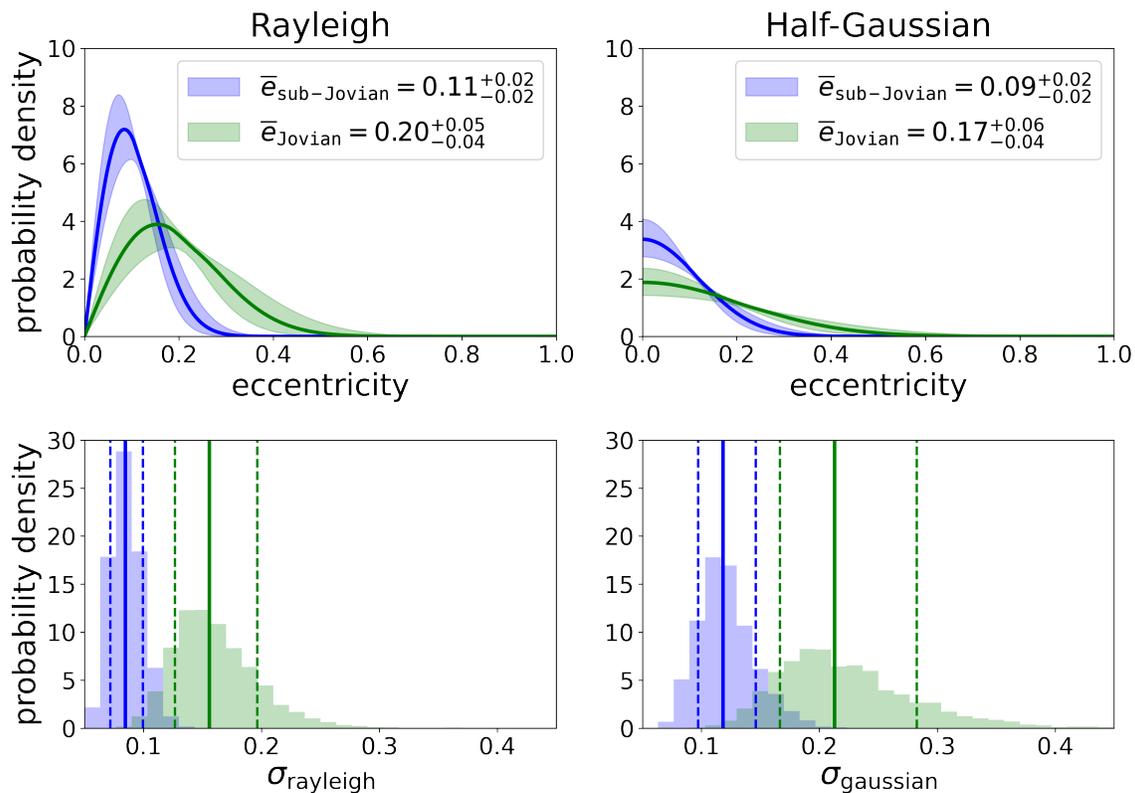


Figure 7.11: Rayleigh and half-Gaussian distribution fits to radius-based sub-populations of the TKS sample, with the sub-Jovian-size population ( $R_p \lesssim 8R_\oplus$ ) shown in blue and the Jovian-size population ( $R_p > 8R_\oplus$ ) shown in green. We find that Jovian-size planets tend to have a higher distribution of eccentricities than sub-Jovian-size planets, consistent with theoretical expectations.

small versus large planets, with the latter consistently displaying higher eccentricities with statistical significance.

## 7.6 Conclusions

In this work, we used hierarchical Bayesian modeling to measure the eccentricity distribution of a sample of TESS planets for the first time. Drawing our input data and planet population from MacDougall et al. 2023a, we recover an underlying eccentricity distribution for this sample that is somewhat less eccentric than past results from other planet populations found in the literature. We propose using the output eccentricity distribution fit from our hierarchical analysis as an input prior for re-weighting the parameter posterior distributions of our planet sample, including for  $R_p/R_\star$  and  $b$ . The updated planet-to-star radius ratios that we measure from this re-weighting process are more precise than those measured using an uninformative uniform prior on eccentricity. We also find preliminary evidence to suggest that our updated re-weighted impact parameter measurements have a more realistic underlying distribution than before implementing our updated  $e$  prior.

To better investigate trends between eccentricity and other system properties, we also apply this hierarchical analysis to various sub-populations of our sample based on planet multiplicity, stellar metallicity, and planet radius. We find that single-planet systems in our sample display a broader eccentricity distribution than multi-planet systems, but this distribution is less eccentric than the single-planet  $e$  distributions recovered by comparable works using *Kepler* planets. This likely points towards an underestimate in the multiplicity of TESS systems due to the shorter baseline of TESS observations. We also find evidence for an anti-correlation between planetary orbital eccentricity and stellar metallicity, counter to expectations.

We offer the first hierarchical Bayesian analysis to investigate the eccentricity distribution of sub-Jovian-size planets ( $R_p < 8R_\oplus$ ) versus Jovian-size planets. We identify a significant

difference between these two populations, with Jovian-size planets generally having higher eccentricities. This observation is supported by various proposed dynamical processes that preferentially excite the orbits of giant transiting planets. Thus, by measuring the true eccentricity distribution underlying all or sub-populations of exoplanets, we can both improve our constraints on exoplanet properties and better identify trends that highlight the dynamical processes governing planetary evolution.

## CHAPTER 8

### Conclusions

To understand the origin of Earth, we must look beyond our own Solar System and study the processes that govern the formation and evolution of distant planetary systems. Large-scale exoplanet surveys such as *Kepler* and TESS have revealed that the range of planetary system outcomes stretches far beyond what we observe among the Solar System planets. Planetary orbital eccentricities help us to make sense of observed planet characteristics and system architectures by providing a glimpse into the dynamical evolution that a planet must have undergone to reach its current state. Although individual exoplanet eccentricities can be difficult to constrain in many cases, trends between eccentricity and other system properties can be measured among planetary sub-populations as evidence of diverging evolutionary pathways.

I have presented a study of photometrically-constrained eccentricities of planets from the TESS-Keck Survey and used my results to identify population-level trends in eccentricity along with individual high-eccentricity sub-Jovian-size planets. I accomplished this by developing a new approach to measuring orbital eccentricities from transit photometry which I then applied towards modeling the transit properties of 108 planets from the TESS mission. I demonstrated that photometric eccentricity constraints can be used to identify potential high-eccentricity planets which can then be followed up through additional observations. I also demonstrated that a large sample of individual noisy eccentricity constraints from TESS photometry can be used to estimate the population-level eccentricity distribution of

transiting planets via hierarchical Bayesian modeling. I used such hierarchical analysis to show that Jovian-size planets have a much broader underlying eccentricity distribution than sub-Jovian-size planets. I also found that planets in single-planet systems display a higher eccentricity distribution than planets in multi-planet systems, consistent with the results of similar studies of *Kepler* planets found in the literature.

My collaborators and I propose using the techniques discussed in this thesis to homogeneously analyze *all* transiting planets in order to improve individual exoplanet properties and reveal new population-level trends in exoplanet demographics. A precise homogeneous re-analysis of exoplanet transit photometry on such a large scale has yet to be performed, but I have a good feeling that we will soon see new discoveries made from such an analysis in the near future. Although I am now stepping away from exoplanet research, I am excited to see the future progress made by my collaborators and fellow exoplaneteers.

## Bibliography

- Aigrain S., Pont F., Zucker S., 2012, *MNRAS*, 419, 3147
- Akeson R. L., et al., 2013a, *PASP*, 125, 989
- Akeson R. L., et al., 2013b, *PASP*, 125, 989
- An D.-S., Xie J.-W., Dai Y.-Z., Zhou J.-L., 2023, *AJ*, 165, 125
- Andrae, René et al., 2018, *A&A*, 616, A8
- Angus R., Foreman-Mackey D., Johnson J. A., 2016, *ApJ*, 818, 109
- Angus R., Morton T., Foreman-Mackey D., 2019a, *The Journal of Open Source Software*, 4, 1469
- Angus R., et al., 2019b, *AJ*, 158, 173
- Armitage P. J., 2018, in Deeg H. J., Belmonte J. A., eds, , *Handbook of Exoplanets*. p. 135, doi:10.1007/978-3-319-55333-7\_135
- Armstrong D. J., et al., 2020, *Nature*, 583, 39–42
- Badenas-Agusti M., et al., 2020, *AJ*, 160, 113
- Bailer-Jones C. A. L., et al., 2013, *A&A*, 559, A74
- Barnes J. W., 2007a, *PASP*, 119, 986
- Barnes J. W., 2007b, *PASP*, 119, 986
- Barnes J. R., et al., 2020, *Nature Astronomy*, 4, 419
- Becker J. C., Adams F. C., 2017, *MNRAS*, 468, 549

Berger T. A., Huber D., Gaidos E., van Saders J. L., 2018a, *ApJ*, 866, 99

Berger T. A., Huber D., Gaidos E., van Saders J. L., 2018b, *ApJ*, 866, 99

Berger T. A., Huber D., van Saders J. L., Gaidos E., Tayar J., Kraus A. L., 2020a, *AJ*, 159, 280

Berger T. A., Huber D., Gaidos E., van Saders J. L., Weiss L. M., 2020b, *AJ*, 160, 108

Berger T. A., Huber D., Gaidos E., van Saders J. L., Weiss L. M., 2020c, *AJ*, 160, 108

Berger T. A., Schlieder J. E., Huber D., 2023, *arXiv e-prints*, p. arXiv:2301.11338

Betancourt M., 2016, *arXiv e-prints*, p. arXiv:1601.00225

Bitsch B., Trifonov T., Izidoro A., 2020, *A&A*, 643, A66

Boisse I., Bouchy F., Hébrard G., Bonfils X., Santos N., Vauclair S., 2011, *A&A*, 528, A4

Borucki W. J., et al., 2010a, *Science*, 327, 977

Borucki W. J., et al., 2010b, *Science*, 327, 977

Borucki W. J., et al., 2010c, *Science*, 327, 977

Bowler B. P., Blunt S. C., Nielsen E. L., 2020, *AJ*, 159, 63

Brady M. T., et al., 2018, *AJ*, 156, 147

Buchhave L. A., et al., 2012, *Nature*, 486, 375

Buchhave L. A., et al., 2014, *Nature*, 509, 593

Cacciapuoti L., et al., 2022, *MNRAS*, 513, 102

Caldwell D. A., et al., 2020, *Research Notes of the American Astronomical Society*, 4, 201

Carter J. A., Yee J. C., Eastman J., Gaudi B. S., Winn J. N., 2008, *ApJ*, 689, 499

Castelli F., Kurucz R. L., 2004, New Grids of ATLAS9 Model Atmospheres ([arXiv:astro-ph/0405087](#))

Charbonneau D., Brown T. M., Latham D. W., Mayor M., 2000, *ApJ*, 529, L45

Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., 2008, *ApJ*, 686, 580

Chen S. S., Donoho D. L., Saunders M. A., 2001, *SIAM Review*, 43, 129

Chiang E. I., 2003, *ApJ*, 584, 465

Chiang E., Laughlin G., 2013a, *MNRAS*, 431, 3444

Chiang E., Laughlin G., 2013b, *MNRAS*, 431, 3444

Chiang E. I., Fischer D., Thommes E., 2002, *ApJ*, 564, L105

Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, *ApJ*, 823, 102

Chontos A., et al., 2022, *AJ*, 163, 297

Christiansen J. L., et al., 2015, *ApJ*, 810, 95

Christiansen J. L., et al., 2020, *AJ*, 160, 159

Claret A., Bloemen S., 2011, *A&A*, 529, A75

Cloutier R., Menou K., 2020, *AJ*, 159, 211

Coelho P., Barbuy B., Meléndez J., Schiavon R. P., Castilho B. V., 2005, *A&A*, 443, 735

Cointepas M., et al., 2021, *A&A*, 650, A145

Correia A. C. M., Bourrier V., Delisle J. B., 2020, *A&A*, 635, A37

Coughlin J. L., et al., 2016, [ApJS](#), 224, 12

Crossfield I. J. M., Kreidberg L., 2017, [AJ](#), 154, 261

Crossfield I. J. M., et al., 2015, [ApJ](#), 804, 10

Dai F., et al., 2020, [AJ](#), 160, 193

Dai F., et al., 2022, arXiv e-prints, p. [arXiv:2210.09283](#)

Dalba P. A., et al., 2020, [AJ](#), 159, 241

Dalba P. A., et al., 2022, [AJ](#), 163, 61

David T. J., et al., 2016, [Nature](#), 534, 658

Dawson R. I., Johnson J. A., 2012, [ApJ](#), 756, 122

Dawson R. I., Murray-Clay R. A., 2013, [ApJ](#), 767, L24

Dawson R. I., Lee E. J., Chiang E., 2016, [ApJ](#), 822, 54

Demory B.-O., 2014, [ApJ](#), 789, L20

Denham P., Naoz S., Hoang B.-M., Stephan A. P., Farr W. M., 2019, [MNRAS](#), 482, 4146

Díaz M. R., et al., 2020, [MNRAS](#), 493, 973

Dotter A., 2016, [ApJS](#), 222, 8

Dreizler S., et al., 2020, [A&A](#), 644, A127

Eastman J., 2017, EXOFASTv2: Generalized publication-quality exoplanet modeling code  
([ascl:1710.003](#))

Eastman J., Gaudi B. S., Agol E., 2013a, [PASP](#), 125, 83

Eastman J., Gaudi B. S., Agol E., 2013b, *PASP*, 125, 83

Espinoza N., 2018, *Research Notes of the American Astronomical Society*, 2, 209

Espinoza N., Jordán A., 2015, *MNRAS*, 450, 1879

Espinoza N., Kossakowski D., Brahm R., 2018, arXiv e-prints, p. arXiv:1812.08549

ExoFOP 2019, Exoplanet Follow-up Observing Program - TESS, doi:10.26134/EXOFOP3, <https://catcopy.ipac.caltech.edu/doi/doi.php?id=10.26134/EXOFOP3>

Fabrycky D., Tremaine S., 2007, *ApJ*, 669, 1298

Fabrycky D. C., et al., 2014, *ApJ*, 790, 146

Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449

Ford E. B., 2005, *AJ*, 129, 1706

Ford E. B., 2006, *ApJ*, 642, 505

Ford E. B., Rasio F. A., 2008, *ApJ*, 686, 621

Ford E. B., Quinn S. N., Veras D., 2008, *ApJ*, 678, 1407

Foreman-Mackey D., 2018, *Research Notes of the American Astronomical Society*, 2, 31

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306

Foreman-Mackey D., Hogg D. W., Morton T. D., 2014, *ApJ*, 795, 64

Foreman-Mackey D., et al., 2021b, *The Journal of Open Source Software*, 6, 3285

Foreman-Mackey D., et al., 2021a, *Journal of Open Source Software*, 6, 3285

Fulton B. J., Petigura E. A., 2018, *AJ*, 156, 264

- Fulton B. J., et al., 2017a, *AJ*, 154, 109
- Fulton B. J., et al., 2017b, *AJ*, 154, 109
- Fulton B. J., Petigura E. A., Blunt S., Sinukoff E., 2018, *PASP*, 130, 044504
- Gaia Collaboration et al., 2018, *A&A*, 616, A1
- Gelman A., Rubin D. B., 1992, *Statistical Science*, 7, 457
- Gelman A., Carlin J., Stern H., Dunson D., Vehtari A., Rubin D., 2013, *Bayesian Data Analysis, Third Edition*. Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis, <https://books.google.com/books?id=ZXL6AQAQBAJ>
- Geyer C. J., 1992, *Statistical Science*, 7, 473
- Gibson S. R., Howard A. W., Marcy G. W., Edelstein J., Wishnow E. H., Poppett C. L., 2016, in Evans C. J., Simard L., Takami H., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI*. p. 990870, [doi:10.1117/12.2233334](https://doi.org/10.1117/12.2233334)
- Gilbert G. J., 2022, arXiv e-prints, p. [arXiv:2201.08350](https://arxiv.org/abs/2201.08350)
- Gilbert G. J., MacDougall M. G., Petigura E. A., 2022, *AJ*, 164, 92
- Gilks W., Richardson S., Spiegelhalter D., 1995, *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC Interdisciplinary Statistics, Taylor & Francis, [http://books.google.com/books?id=TRXrMWY\\_i2IC](http://books.google.com/books?id=TRXrMWY_i2IC)
- Goldreich P., Sari R., 2003, *ApJ*, 585, 1024
- Goldreich P., Soter S., 1966, *Icarus*, 5, 375
- Gomes R., Levison H. F., Tsiganis K., Morbidelli A., 2005, *Nature*, 435, 466

Guerrero N. M., et al., 2021, [ApJS](#), 254, 39

Gupta A., Schlichting H. E., 2019, [MNRAS](#), 487, 24

Hadden S., Lithwick Y., 2014, [ApJ](#), 787, 80

Hansen B. M. S., 2017, [MNRAS](#), 467, 1531

Hara N. C., Boué G., Laskar J., Correia A. C. M., 2017, [MNRAS](#), 464, 1220

Hauschildt P. H., Baron E., 1999, [Journal of Computational and Applied Mathematics](#), 109, 41

Hedges C., Angus R., Barentsen G., Saunders N., Montet B. T., Gully-Santiago M., 2020, [Research Notes of the American Astronomical Society](#), 4, 220

Henry G. W., Marcy G. W., Butler R. P., Vogt S. S., 2000, [ApJ](#), 529, L41

Hodapp K. W., et al., 2003, [PASP](#), 115, 1388

Hoffman M. D., Gelman A., 2011a, arXiv e-prints, p. [arXiv:1111.4246](#)

Hoffman M. D., Gelman A., 2011b, arXiv e-prints, p. [arXiv:1111.4246](#)

Hogg D. W., Myers A. D., Bovy J., 2010, [ApJ](#), 725, 2166

Howard A. W., et al., 2010, [ApJ](#), 721, 1467

Howell S. B., et al., 2014, [PASP](#), 126, 398

Huber D., et al., 2017, [ApJ](#), 844, 102

Husser T. O., Wende-von Berg S., Dreizler S., Homeier D., Reinert A., Barman T., Hauschildt P. H., 2013, [A&A](#), 553, A6

Isaacson H., Fischer D., 2010, [ApJ](#), 725, 875

Izidoro A., Ogihara M., Raymond S. N., Morbidelli A., Pierens A., Bitsch B., Cossou C., Hersant F., 2017, [MNRAS](#), 470, 1750

Jenkins J. M., 2002, [ApJ](#), 575, 493

Jenkins J. M., et al., 2010, in Radziwill N. M., Bridger A., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7740, Software and Cyberinfrastructure for Astronomy. p. 77400D, [doi:10.1117/12.856764](#)

Jenkins J. M., et al., 2016, in Software and Cyberinfrastructure for Astronomy IV. p. 99133E, [doi:10.1117/12.2233418](#)

Johnson J. A., et al., 2017, [AJ](#), 154, 108

Jurić M., Tremaine S., 2008, [ApJ](#), 686, 603

Kane S. R., Ciardi D. R., Gelino D. M., von Braun K., 2012, [MNRAS](#), 425, 757

Kane S. R., et al., 2021, [PASP](#), 133, 014402

Kipping D. M., 2010a, [MNRAS](#), 407, 301

Kipping D. M., 2010b, [MNRAS](#), 407, 301

Kipping D. M., 2010c, [MNRAS](#), 408, 1758

Kipping D. M., 2013, [MNRAS](#), 434, L51

Kipping D. M., 2014a, [MNRAS](#), 440, 2164

Kipping D. M., 2014b, [MNRAS](#), 444, 2263

Kipping D. M., Sandford E., 2016, [MNRAS](#), 463, 1323

Kipping D. M., Dunn W. R., Jasinski J. M., Manthri V. P., 2012, [MNRAS](#), 421, 1166

Kish L., 1965, [American Political Science Review](#), 59, 1025–1025

Kosiarek M. R., et al., 2019, [AJ](#), 157, 97

Kovács G., Zucker S., Mazeh T., 2002a, [A&A](#), 391, 369

Kovács G., Zucker S., Mazeh T., 2002b, [A&A](#), 391, 369

Kreidberg L., 2015, [Publications of the Astronomical Society of the Pacific](#), 127, 1161

Kumar R., Carroll C., Hartikainen A., Martin O., 2019, [Journal of Open Source Software](#), 4, 1143

Lacedelli G., et al., 2022, [MNRAS](#), 511, 4551

Lainey V., 2016, [Celestial Mechanics and Dynamical Astronomy](#), 126, 145

Lallement R., Babusiaux C., Vergely J. L., Katz D., Arenou F., Valette B., Hottier C., Capitanio L., 2019, [A&A](#), 625, A135

Lanotte A. A., et al., 2014, [A&A](#), 572, A73

Leconte J., Chabrier G., Baraffe I., Levrard B., 2010, [A&A](#), 516, A64

Lee M. H., Thommes E. W., 2009, [ApJ](#), 702, 1662

Li J., Tenenbaum P., Twicken J. D., Burke C. J., Jenkins J. M., Quintana E. V., Rowe J. F., Seader S. E., 2019, [PASP](#), 131, 024506

Lightkurve Collaboration et al., 2018, Lightkurve: Kepler and TESS time series analysis in Python (ascl:1812.013)

Limbach M. A., Turner E. L., 2015, *Proceedings of the National Academy of Sciences*, 112, 20

Lindegren L., et al., 2018, *A&A*, 616, A2

Lissauer J. J., 1993, *ARA&A*, 31, 129

Lopez E. D., Fortney J. J., 2013, *ApJ*, 776, 2

Lopez E. D., Fortney J. J., 2014, *ApJ*, 792, 1

Lubin J., et al., 2022, *AJ*, 163, 101

Lucy L. B., Sweeney M. A., 1971, *AJ*, 76, 544

Luhn J. K., Wright J. T., Howard A. W., Isaacson H., 2020, *AJ*, 159, 235

Lundkvist M. S., et al., 2016, *Nature Communications*, 7, 11201

MAST 2021, TESS Light Curves - All Sectors, doi:10.17909/T9-NMC8-F686, <http://archive.stsci.edu/doi/resolve/resolve.html?doi=10.17909/t9-nmc8-f686>

MacDougall M. G., et al., 2021, *AJ*, 162, 265

MacDougall M. G., et al., 2022, *AJ*, 164, 97

Madras N., Piccioni M., 1999, *The Annals of Applied Probability*, 9, 1202

Magliano C., et al., 2023, *arXiv e-prints*, p. arXiv:2303.00624

Mandel K., Agol E., 2002a, *ApJ*, 580, L171

Mandel K., Agol E., 2002b, *ApJ*, 580, L171

Mann A. W., et al., 2019, *ApJ*, 871, 63

Marcy G. W., Butler R. P., 1992, *PASP*, 104, 270

Marcy G. W., et al., 2014, *ApJS*, 210, 20

Mardling R. A., Lin D. N. C., 2004, *ApJ*, 614, 955

Masuda K., Petigura E. A., Hall O. J., 2022, *MNRAS*, 510, 5623

Matsakos T., Königl A., 2016, *ApJ*, 820, L8

Matsumura S., Peale S. J., Rasio F. A., 2010, *ApJ*, 725, 1995

Mayor M., et al., 2011, arXiv e-prints, p. arXiv:1109.2497

Mazeh T., Holczer T., Faigler S., 2016, *A&A*, 589, A75

Millholland S., Petigura E., Batygin K., 2020, *ApJ*, 897, 7

Mills S. M., Howard A. W., Petigura E. A., Fulton B. J., Isaacson H., Weiss L. M., 2019, *AJ*, 157, 198

Moorhead A. V., et al., 2011, *ApJS*, 197, 1

Mulders G. D., O'Brien D., Ciesla F., Apai D., Pascucci I., 2020, in AAS/Division for Planetary Sciences Meeting Abstracts. p. 312.04

Mullally F., et al., 2015, *ApJS*, 217, 31

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

NASA Exoplanet Archive 2022a, Planetary Systems, doi:10.26133/NEA12, <https://catcopy.ipac.caltech.edu/does/doi.php?id=10.26133/NEA12>

NASA Exoplanet Archive 2022b, TESS Objects of Interest, doi:10.26134/ExoFOP3, <https://catcopy.ipac.caltech.edu/does/doi.php?id=10.26134/ExoFOP3>

NASA Exoplanet Science Institute 2020, Planetary Systems Composite Table,  
[doi:10.26133/NEA13](https://catcopy.ipac.caltech.edu/doi/doi.php?id=10.26133/NEA13), <https://catcopy.ipac.caltech.edu/doi/doi.php?id=10.26133/NEA13>

Naoz S., 2016, *ARA&A*, 54, 441

Narita N., et al., 2015, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1,  
045001

Narita N., et al., 2019, *Journal of Astronomical Telescopes, Instruments, and Systems*, 5,  
015001

Neal R., 2011, MCMC Using Hamiltonian Dynamics. pp 113–162, [doi:10.1201/b10905](https://doi.org/10.1201/b10905)

Oh M., Berger J., 1993, *Journal of the American Statistical Association*, 88, 450

Osborn H. P., et al., 2017, *A&A*, 604, A19

Osborn H. P., et al., 2022, *A&A*, 664, A156

Owen J. E., Lai D., 2018, *MNRAS*, 479, 5012

Owen J. E., Wu Y., 2013, *ApJ*, 775, 105

Owen J. E., Wu Y., 2017, *ApJ*, 847, 29

Parviainen H., Aigrain S., 2015, *MNRAS*, 453, 3821

Parviainen H., et al., 2019, *A&A*, 630, A89

Patel J. A., Espinoza N., 2022, *AJ*, 163, 228

Petigura E. A., 2015, PhD thesis, University of California, Berkeley

Petigura E. A., 2020, *AJ*, 160, 89

- Petigura E. A., et al., 2017a, *AJ*, 153, 142
- Petigura E. A., et al., 2017b, *AJ*, 154, 107
- Petigura E. A., et al., 2022, *AJ*, 163, 179
- Petrovich C., 2015, *ApJ*, 808, 120
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in  
C. The art of scientific computing
- Pu B., Lai D., 2018, *MNRAS*, 478, 197
- Rauer H., et al., 2014a, *Experimental Astronomy*, 38, 249
- Rauer H., et al., 2014b, *Experimental Astronomy*, 38, 249
- Raymond S. N., Morbidelli A., 2022, in Biazzo K., Bozza V., Mancini L., Sozzetti A.,  
eds, *Astrophysics and Space Science Library Vol. 466, Demographics of Exoplanetary  
Systems, Lecture Notes of the 3rd Advanced School on Exoplanetary Science*. pp 3–82  
(arXiv:2002.05756), doi:10.1007/978-3-030-88124-5\_1
- Rein H., Liu S. F., 2012, *A&A*, 537, A128
- Ricker G. R., et al., 2015a, *Journal of Astronomical Telescopes, Instruments, and Systems*,  
1, 014003
- Ricker G. R., et al., 2015b, *Journal of Astronomical Telescopes, Instruments, and Systems*,  
1, 014003
- Rogers L. A., 2015, *ApJ*, 801, 41
- Rogers T. M., Showman A. P., 2014, *ApJ*, 782, L4

Rosenthal L. J., et al., 2021, *ApJS*, 255, 8

Rowe J. F., et al., 2014a, *ApJ*, 784, 45

Rowe J. F., et al., 2014b, *ApJ*, 784, 45

Rowe J. F., et al., 2015, *ApJS*, 217, 16

Rubenzahl R. A., et al., 2021, *AJ*, 161, 119

STScI 2018, TESS Input Catalog and Candidate Target List, doi:10.17909/FWDT-2X66, <http://archive.stsci.edu/doi/resolve/resolve.html?doi=10.17909/fwdt-2x66>

Salvatier J., Wiecki T. V., Fonnesbeck C., 2016a, PyMC3: Python probabilistic programming framework (ascl:1610.016)

Salvatier J., Wiecki T. V., Fonnesbeck C., 2016b, *PeerJ Comput. Sci.*, 2, e55

Sandford E., Kipping D., 2017, *AJ*, 154, 228

Scarsdale N., et al., 2021, *AJ*, 162, 215

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

Scott D. W., 1992, *Multivariate Density Estimation*

Seager S., Mallén-Ornelas G., 2003a, *ApJ*, 585, 1038

Seager S., Mallén-Ornelas G., 2003b, *ApJ*, 585, 1038

Seifahrt A., Stürmer J., Bean J. L., Schwab C., 2018, in Evans C. J., Simard L., Takami H., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 10702, Ground-based and Airborne Instrumentation for Astronomy VII*. p. 107026D (arXiv:1805.09276), doi:10.1117/12.2312936

Skilling J., 2004, in Fischer R., Preuss R., Toussaint U. V., eds, American Institute of Physics Conference Series Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering. pp 395–405, doi:10.1063/1.1835238

Skilling J., 2006, [Bayesian Analysis](#), 1, 833

Skrutskie M. F., et al., 2006, [AJ](#), 131, 1163

Smith J. C., et al., 2012, [PASP](#), 124, 1000

Speagle J. S., 2020, [MNRAS](#), 493, 3132

Stassun K. G., Collins K. A., Gaudi B. S., 2017, [AJ](#), 153, 136

Stassun K. G., et al., 2018, [AJ](#), 156, 102

Stassun K. G., et al., 2019, [AJ](#), 158, 138

Stumpe M. C., et al., 2012, [PASP](#), 124, 985

Stumpe M. C., Smith J. C., Catanzarite J. H., Van Cleve J. E., Jenkins J. M., Twicken J. D., Girouard F. R., 2014, [PASP](#), 126, 100

Sun L., Ioannidis P., Gu S., Schmitt J. H. M. M., Wang X., Kouwenhoven M. B. N., 2019, [A&A](#), 624, A15

Tayar J., Claytor Z. R., Huber D., van Saders J., 2022, [ApJ](#), 927, 31

Teske J., et al., 2021, [ApJS](#), 256, 33

Thompson S. E., et al., 2018, [ApJS](#), 235, 38

Tsiganis K., Gomes R., Morbidelli A., Levison H. F., 2005, [Nature](#), 435, 459

Turtelboom E. V., et al., 2022, *AJ*, 163, 293

Twicken J. D., et al., 2018, *PASP*, 130, 064502

Van Eylen V., Albrecht S., 2015, *ApJ*, 808, 126

Van Eylen V., et al., 2014, *ApJ*, 782, 14

Van Eylen V., Agentoft C., Lundkvist M. S., Kjeldsen H., Owen J. E., Fulton B. J., Petigura E., Snellen I., 2018, *MNRAS*, 479, 4786

Van Eylen V., et al., 2019, *AJ*, 157, 61

Van Zandt J. E., et al., 2022, arXiv e-prints, p. [arXiv:2209.06958](https://arxiv.org/abs/2209.06958)

Vogt S. S., et al., 1994, in Crawford D. L., Craine E. R., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 2198, Instrumentation in Astronomy VIII. p. 362, [doi:10.1117/12.176725](https://doi.org/10.1117/12.176725)

Vogt S. S., et al., 2014, *PASP*, 126, 359

Weiss L. M., Marcy G. W., 2014, *ApJ*, 783, L6

Weiss L. M., et al., 2018, *AJ*, 155, 48

Weiss L. M., et al., 2021, *AJ*, 161, 56

Winn J. N., 2010a, arXiv: Earth and Planetary Astrophysics

Winn J. N., 2010b, arXiv e-prints, p. [arXiv:1001.2010](https://arxiv.org/abs/1001.2010)

Winn J. N., Fabrycky D. C., 2015, *ARA&A*, 53, 409

Wolfgang A., Rogers L. A., Ford E. B., 2016, *ApJ*, 825, 19

Wolszczan A., Frail D. A., 1992, [Nature](#), 355, 145

Wright J. T., Upadhyay S., Marcy G. W., Fischer D. A., Ford E. B., Johnson J. A., 2009, [ApJ](#), 693, 1084

Xie J.-W., et al., 2016, [Proceedings of the National Academy of Science](#), 113, 11431

Yee S. W., Petigura E. A., von Braun K., 2017, [ApJ](#), 836, 77

Yee S. W., et al., 2018, [AJ](#), 155, 255

Yee S. W., et al., 2022, [AJ](#), 164, 70

Zeng L., Sasselov D. D., Jacobsen S. B., 2016, [ApJ](#), 819, 127

Zeng L., et al., 2019, [Proceedings of the National Academy of Science](#), 116, 9723