



Edge-of-the-Multis: Evidence for a Transition in the Outer Architectures of Compact Multiplanet Systems

Sarah C. Millholland^{1,2,3} , Matthias Y. He^{4,5} , and Jon K. Zink⁶ ¹ Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; sarah.milholland@mit.edu² MIT Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA³ Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA⁴ Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA⁵ Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA⁶ Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

Received 2022 April 20; revised 2022 June 21; accepted 2022 June 26; published 2022 August 1

Abstract

Although the architectures of compact multiple-planet systems are well characterized, there has been little examination of their “outer edges,” or the locations of their outermost planets. Here we present evidence that the observed high-multiplicity Kepler systems truncate at smaller orbital periods than can be explained by geometric and detection biases alone. To show this, we considered the existence of hypothetical planets orbiting beyond the observed transiting planets with properties dictated by the “peas-in-a-pod” patterns of intrasystem radius and period ratio uniformity. We evaluated the detectability of these hypothetical planets using (1) a novel approach for estimating the mutual inclination dispersion of multitransiting systems based on transit chord length ratios, and (2) a model of transit probability and detection efficiency that accounts for the impacts of planet multiplicity on completeness. Under the assumption that the “peas-in-a-pod” patterns continue to larger orbital separations than observed, we find that $\gtrsim 35\%$ of Kepler compact multis should possess additional detected planets beyond the known planets, constituting a $\sim 7\sigma$ discrepancy with the lack of such detections. These results indicate that the outer ($\sim 100\text{--}300$ days) regions of compact multis experience a truncation (i.e., an “edge-of-the-multis”) or a significant breakdown of the “peas-in-a-pod” patterns, in the form of systematically smaller radii or larger period ratios. We outline future observations that can distinguish these possibilities, and we discuss implications for planet formation theories.

Unified Astronomy Thesaurus concepts: [Exoplanets \(498\)](#); [Super Earths \(1655\)](#); [Mini Neptunes \(1063\)](#); [Exoplanet systems \(484\)](#); [Exoplanet dynamics \(490\)](#); [Exoplanet formation \(492\)](#)

1. Introduction

Short-period, sub-Neptune-sized planets are one of the most prevalent types of exoplanets in the local galaxy. Discovered in abundance by NASA’s Kepler mission (Borucki et al. 2010), these planets are frequently found in multiple-planet systems with orbital periods ranging from days to months (Lissauer et al. 2011, 2014; Rowe et al. 2014; Fabrycky et al. 2014). Over the last decade, the architectures of these compact multiple-planet systems (“compact multis”) have been characterized in detail. The observed planets have low eccentricities (e.g., Van Eylen & Albrecht 2015; Xie et al. 2016), low inclinations (e.g., Fang & Margot 2012; Fabrycky et al. 2014), and tight orbital spacings that are generally not near resonances (Lissauer et al. 2011; Fabrycky et al. 2014; Winn & Fabrycky 2015). Within a given system, the period ratios, planet radii, and planet masses tend to be significantly more uniform than would be expected by random chance—a set of patterns known collectively as the “peas-in-a-pod” patterns or “intrasystem uniformity” (Weiss et al. 2018; Millholland et al. 2017; for a review, see Weiss et al. 2022).

The statistical properties of the observed compact multis have in turn revealed corresponding properties of the underlying distribution of planetary systems (i.e., including

undetected planets) through the help of forward models. For instance, the fraction of stars with multiple planets interior to ~ 1 au is estimated at $\sim 30\%\text{--}70\%$ (Mulders et al. 2018; Zhu et al. 2018; Zink et al. 2019; He et al. 2019), with the occurrence increasing for cooler stars (Howard et al. 2012; Mulders et al. 2015; Yang et al. 2020; He et al. 2021). There is evidence that the eccentricities and inclinations correlate with the intrinsic planet multiplicity (Zhu et al. 2018; He et al. 2020; Zhu & Dong 2021) and that the inclinations are both relatively small overall and not bifurcated into dynamically cool and dynamically hot subpopulations (Millholland et al. 2021). Finally, forward models have shown that the patterns of intrasystem uniformity in planetary sizes and orbital spacings must be present to a strong degree in the underlying population (He et al. 2019; Gilbert & Fabrycky 2020; He et al. 2020; Mishra et al. 2021), indicating that these patterns cannot be explained by detection biases and require astrophysical origins.

There are additional features of compact multis that, relative to the properties mentioned above, have received comparatively little attention: that is, their inner and outer edges, which can be summarized statistically as the distributions of the orbital periods of the innermost and outermost planets in the systems. The edges of compact multis are not well understood from either an observational or theoretical perspective. However, these features (specifically those of the underlying population of systems) are crucial for a complete characterization of short-period, tightly-packed systems.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

The inner edges are the simpler of the two from an observational perspective, since the geometric and detection biases are comparatively minimal. A simple calculation based on a sample of Kepler systems with four or more observed planets (to be defined later in this paper) shows that the distribution of observed innermost planet periods has a median of 3.9 days, with the 16th–84th percentile interval equal to [2.2, 7.0] days. After accounting for biases, Mulders et al. (2018) showed that the underlying distribution of inner edges likely peaks at slightly longer orbital periods of about ~ 10 days. This is consistent with being a signature from the protoplanetary disk inner edges (e.g., Millan-Gabet et al. 2007), which may limit where planets are able to form in situ (e.g., Lee & Chiang 2017) or may act as a planet trap that halts inward migration (e.g Terquem & Papaloizou 2007; Izidoro et al. 2017).

The outer edges are considerably more difficult to understand because they are heavily impacted by geometric and detection biases. Specifically, the transit probability and transit signal-to-noise ratio (S/N) decrease with increasing orbital period as $p_{\text{trans}} \sim P^{-2/3}$ and $S/N \sim P^{-1/3}$, respectively. Moreover, transit surveys require a minimum number of transits for detection (three for the Kepler pipeline), which, together with the observation baseline and duty cycle, limits the largest detectable orbital period. This upper limit is $P \lesssim 500$ days for the Kepler prime mission. For a sample of Kepler systems with four or more observed planets (the same sample as discussed in the previous paragraph), the distribution of outermost planet periods has a median of 40.6 days, with the 16th–84th percentile interval equal to [16.3, 76.9] days. Naively speaking, this distribution already seems inconsistent with expectations, since almost all of the outermost periods are more than an order of magnitude smaller than the upper limit of ~ 500 days. However, such comparisons are not very meaningful until the relevant geometric and detection biases have been thoroughly accounted for.

In this work, we examine the outer edges of compact multis discovered by the Kepler prime mission.⁷ Our main goal is to understand whether the outer edges of the observed systems are consistent with sculpting purely from geometric and detection biases acting upon an underlying population of systems that extend out to (and potentially beyond) the largest periods probed by the Kepler photometry. Alternatively, if we find that the observed systems truncate at smaller orbital periods than required by their detection, then this could indicate either that (1) the underlying planets in compact multis are present only within some restricted range of orbital separations, and/or (2) there are other significant changes in the planet properties (e.g., periods, radii) at larger orbital separations. To explore this, we will consider the existence of hypothetical planets orbiting beyond the outermost planets in Kepler multis, and we will estimate the number of these hypothetical planets that we would expect to be detectable.

It is important to note that this investigation of the outer edges of compact multis has consequences beyond demographics; specifically, the outer edges are signatures of planet formation and dynamical evolution. Just as the inner edges of compact multis are likely relics from the disk inner edges, a

possible truncation in the outer systems could be the result of disk migration traps (e.g., Zawadzki et al. 2022), planet formation in pebble rings (e.g Chatterjee & Tan 2014), or dynamical perturbations from exterior giant planets (e.g., Pu & Lai 2018), among other possibilities. We will review such theoretical considerations in Section 7. However, it is helpful to keep this motivation in mind from the outset.

This paper is organized as follows. We begin with a heuristic demonstration of the detectability of potential transiting outer planets, as a motivation for the more detailed calculations in subsequent sections (Section 2). We then investigate the geometric aspects of the problem and develop a new method for estimating a system’s mean inclination and mutual inclination dispersion (Section 3). Based on these estimates, we define a model for the transit probability, as well as a detection-efficiency model (Section 4). We then validate these approaches on populations of simulated planets (Section 5). Moving onto observed systems, we estimate the number of hypothetical additional outer planets that we would expect to be transiting and detectable in Kepler multiple-planet systems (Section 6). Finally, we discuss the implications of our results for planet formation theories (Section 7).

2. Heuristic Calculation

We begin with a heuristic exploration of the detectability of potential transiting outer planets in Kepler multiplanet systems. This offers a simplified preview of later detailed calculations. We consider a sample of 64 Kepler systems with four or more transiting planets. (The details of the sample selection will be provided in Section 6.1.) We imagine that each system has an additional transiting outer planet with properties consistent with the “peas-in-a-pod” patterns. Thus, we set the radii and periods of these “hypothetical planets” to $R_{p,\text{hypo}} = R_{p,\text{out}}$ and $P_{\text{hypo}}/P_{\text{out}} = P_{\text{out}}/P_{2\text{nd-out}}$, where P_{out} and $R_{p,\text{out}}$ are the period and radius of the outermost observed planet, and $P_{2\text{nd-out}}$ is the period of the second outermost observed planet.

Next, we compare $R_{p,\text{hypo}}$ to the minimum planet radius, $R_{p,\text{min}}$, that the hypothetical planets would require in order to be detected with signal-to-noise ratio, $S/N > S/N_{\text{min}} = 7$, which is approximately equal to the detection threshold of the Kepler pipeline. We define the S/N as (Christiansen et al. 2012)

$$S/N = \frac{(R_p/R_*)^2}{\text{CDPP}_{\text{eff}}} \sqrt{\frac{t_{\text{obs}} f_0}{P}}. \quad (1)$$

Here, t_{obs} is the time interval over which data were collected (~ 4 yr in most cases), f_0 is the duty cycle, and CDPP_{eff} is the effective combined differential photometric precision,

$$\text{CDPP}_{\text{eff}} = \text{CDPP}_{6\text{ hr}} \sqrt{\frac{6\text{ hr}}{T}}, \quad (2)$$

where $T \propto P^{-1/3}$ is the transit duration (see Equation (4) for the full expression). In Section 4.2, we will use a more advanced definition of the S/N, but this definition suffices for the purposes of this illustration. Given that $S/N \propto P^{-1/3} R_p^2$, we can calculate the minimum planet radius, $R_{p,\text{min}}$, for which $S/N > S/N_{\text{min}} = 7$ at $P = P_{\text{hypo}}$ using a scaling to S/N_{out} , the

⁷ Although K2 and TESS have also discovered some compact multis, we restrict our focus to the large sample of systems discovered by the Kepler prime mission, since its homogeneous coverage out to large orbital periods is required for this study.

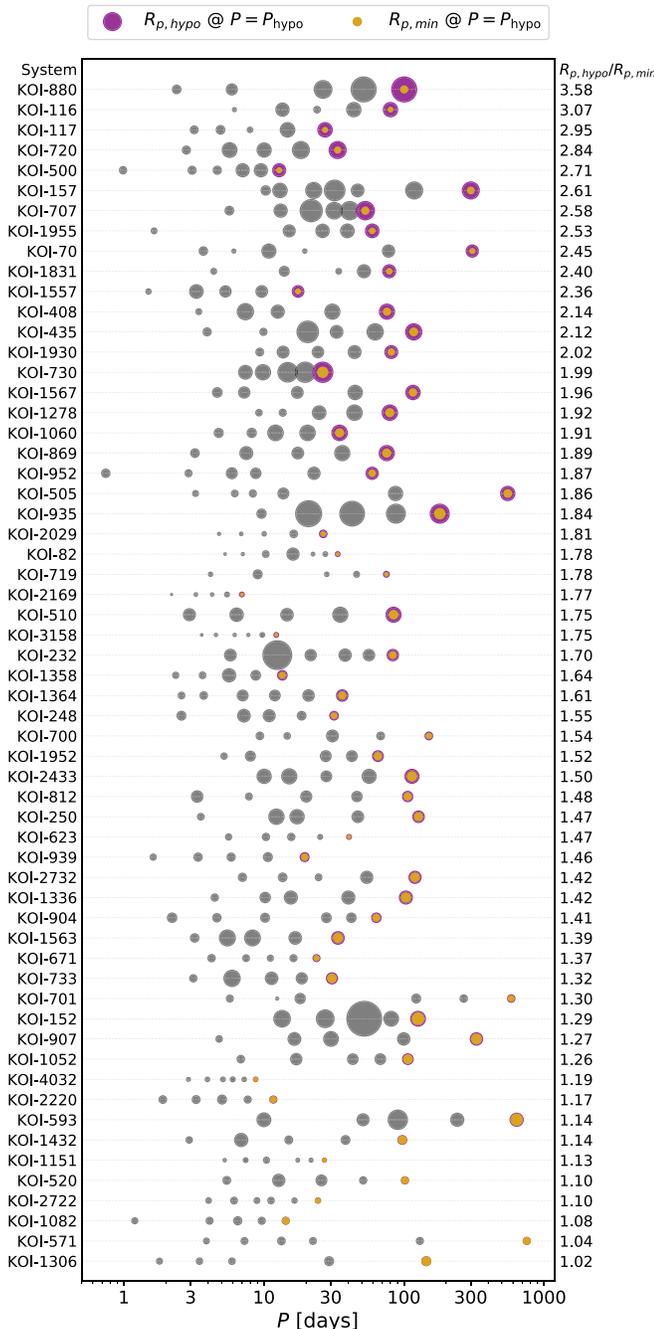


Figure 1. Architectures of Kepler systems with four or more transiting planets, where all planets in the system are smaller than $8 R_{\oplus}$. The dot size is proportional to the planet radius. Observed planets are indicated in gray, while the “hypothetical planets” with $R_{p,\text{hypo}} = R_{p,\text{out}}$ and $P_{\text{hypo}}/P_{\text{out}} = P_{\text{out}}/P_{2\text{nd-out}}$ are shown in purple. Yellow dots indicate the minimum planet radius, $R_{p,\text{min}}$, necessary for $S/N > 7$ at $P = P_{\text{hypo}}$. The systems are reverse-ordered by the ratio, $R_{p,\text{hypo}}/R_{p,\text{min}}$, as shown at right. In every system, $R_{p,\text{hypo}}$ is greater than $R_{p,\text{min}}$.

S/N of the outermost observed planet in the system,

$$R_{p,\text{min}} = R_{p,\text{out}} \left(\frac{S/N_{\text{min}}}{S/N_{\text{out}}} \right)^{1/2} \left(\frac{P_{\text{hypo}}}{P_{\text{out}}} \right)^{1/6}. \quad (3)$$

Figure 1 shows the architectures of Kepler systems with four or more planets, along with dots indicating the radii $R_{p,\text{hypo}}$ and $R_{p,\text{min}}$ at period P_{hypo} . For the purposes of the visualization, we

include only systems for which all planets in the system have $R_p < 8 R_{\oplus}$, leaving out five systems. Remarkably, in all 59 systems, $R_{p,\text{hypo}} > R_{p,\text{min}}$. In other words, the hypothetical planet’s nominal radius is larger than the minimum radius required for $S/N > 7$ in every single case. This result can be seen by the purple outlines around the yellow dots in Figure 1 and the $R_{p,\text{hypo}}/R_{p,\text{min}}$ ratios on the right-hand side. If we increase P_{hypo} , then $R_{p,\text{hypo}} > R_{p,\text{min}}$ remains generally true but not in every case. For instance, if we set $P_{\text{hypo}} = 300$ days in all systems, then $R_{p,\text{hypo}} > R_{p,\text{min}}$ in 45 out of 59 systems.

To summarize, this brief experiment has demonstrated the following: if additional transiting planets exist beyond the known transiting planets, and if they have properties in line with expectations from the “peas-in-a-pod” patterns, then we would generally expect them to be detectable. In the remainder of this paper, we explore this simple premise in greater detail. We consider similar experiments involving hypothetical outer planets, but we consider both transit probabilities and detection probabilities and formulate detailed expectations of the frequency of such planets we would expect to be transiting and detectable. In the next section, we begin with the key geometric aspect of the problem: mutual orbital inclinations.

3. Mutual Inclination Estimation Using Transit Chord Ratios

3.1. Overview of the Transit Chord Ratio Method

One of our subsequent objectives is to quantify the likelihood that planets orbiting beyond the known planets in a multitransiting system are also transiting. In order to do this, we must first obtain estimates of the mean inclination and inclination dispersion. Here we develop a new procedure, called the “transit chord ratio method,” which uses the ratios of the transit chord lengths of pairs of planets in the same system (Steffen et al. 2010) as a means of constraining the planets’ inclinations. This has previously been done at the population level (e.g., Fang & Margot 2012; Fabrycky et al. 2014) but not the level of individual systems. One benefit of working with ratios of transit chord lengths (rather than each planet’s individual transit duration) is that it significantly reduces the impact of the degeneracy between the stellar density and impact parameter, which are difficult to accurately measure with Kepler long-cadence photometry (Petigura 2020).

It is important to note that transit chord ratios are sensitive specifically to sky-plane inclinations, the inclinations between the planetary orbit planes and the sky plane. Accordingly, they cannot directly constrain the dispersion of the inclinations between the planetary orbits and the invariable plane (the plane perpendicular to the system’s total angular momentum vector). Moreover, since the transit chord lengths are symmetric on the two stellar hemispheres bisected by the $b=0$ chord, the method cannot determine whether two planets are transiting on the same hemisphere or opposite ones. Despite these limitations, the method is still capable of providing useful estimates of the mean inclination and inclination dispersion of multi-transiting systems, as we will soon demonstrate.

In the limit $R_p \ll R_* \ll a$, a planet’s transit duration is given by

$$T = \frac{2R_* \sqrt{1 - b^2}}{v_{\text{mid}}}, \quad (4)$$

where R_* is the stellar radius, b is the dimensionless impact parameter, and v_{mid} is the sky-projected orbital velocity at

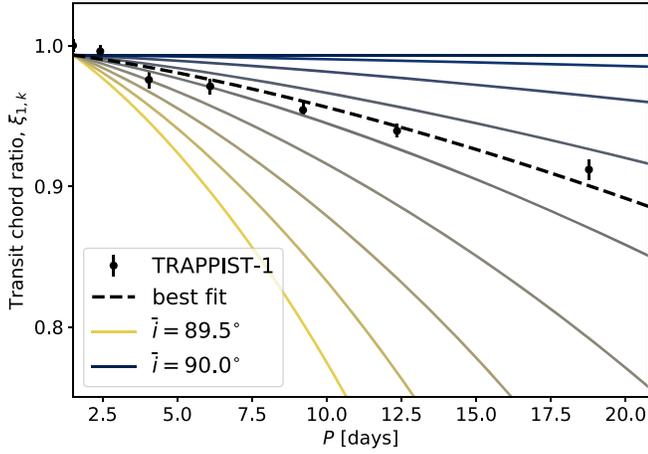


Figure 2. Transit chord ratios for planets in the TRAPPIST-1 system. The black dots with error bars indicate the observed ratios, $(\xi_{1,k})_{\text{obs}} = (T_k/T_1)(P_k/P_1)^{-1/3}$. The solid curves represent coplanar models, $(\xi_{1,k})_{\text{mod}}$, with different inclinations, ranging from $\bar{i} = 89.5^\circ$ to $\bar{i} = 90^\circ$ in steps of 0.0625° . The black dashed line indicates the model curve with best-fit parameters $\bar{i} = 89.781^\circ$ and $\gamma = 0.993$.

mid-transit. Here we will make the simplifying assumption of circular orbits,⁸ such that $v_{\text{mid}} = na$, where $n = 2\pi/P$ is the mean motion and a is the semimajor axis. The ratio of transit chord lengths (Tv_{mid}) of a pair of planets (j and k with $j < k$) in the same system is

$$\begin{aligned} \xi_{j,k} &\equiv \left(\frac{T_k}{T_j}\right) \left(\frac{P_k}{P_j}\right)^{-1/3} = \sqrt{\frac{1 - (a_k/R_\star)^2 \cos^2 i_k}{1 - (a_j/R_\star)^2 \cos^2 i_j}} \\ &= \sqrt{\frac{1 - \left(\frac{1}{3}G\pi^{-1}\rho_\star P_k^2\right)^{2/3} \cos^2 i_k}{1 - \left(\frac{1}{3}G\pi^{-1}\rho_\star P_j^2\right)^{2/3} \cos^2 i_j}}, \end{aligned} \quad (5)$$

where on the right-hand side of the first line, we have made the substitution $bR_\star = a \cos i$, with i the inclination between the planet’s orbit plane and the sky plane. In the final line, we expressed a/R_\star in terms of the orbital period and stellar density, ρ_\star . Since T and P are well measured from transit data, the observed ratios $(\xi_{j,k})_{\text{obs}} = (T_k/T_j)(P_k/P_j)^{-1/3}$ can provide constraints on the inclinations (Fabrycky et al. 2014). The uncertainties on $(\xi_{j,k})_{\text{obs}}$ can be calculated through standard error propagation using the reported uncertainties on T and P .

Consider the ratio of transit chord lengths of each planet in a multiplanet system with respect to that of the innermost planet, $\xi_{1,k}$, where $k = 1, 2, \dots, N$. For a perfectly coplanar system with constant sky-plane inclination \bar{i} , data points of $(P, \xi_{1,k})$ would trace out a smooth and monotonic curve. Examples of these curves are depicted in Figure 2 and are calculated by setting the inclinations to \bar{i} in Equation (5). For a set of observed $(\xi_{1,k})_{\text{obs}}$ values in a system, we can identify a model curve, $(\xi_{1,k})_{\text{mod}}$, that best fits the observations. The spread of the $(\xi_{1,k})_{\text{obs}}$ measurements around the best-fit curve indicates the degree of sky-plane inclination dispersion about the mean sky-plane inclination.

⁸ This assumption will slightly increase our estimates of the inclination dispersion because any eccentricity-driven spread of the transit durations from the circular orbit expectations will be interpreted as a spread in impact parameters. In practice, this is a small effect due to the nearly circular orbits of planets in compact multis.

When fitting the model curve, we allow $(\xi_{1,k})_{\text{mod}}$ to be modified by a vertical scaling factor, γ , which accounts for the fact that i_1 may not be near \bar{i} . In other words, $(\xi_{1,k})_{\text{mod}}$ need not pass through unity at $P = P_1$, since this can otherwise bias the fit. We use nonlinear least squares to find the best-fit values of \bar{i} and γ that minimize the sum of the squared residuals of $(\xi_{1,k})_{\text{mod}} - (\xi_{1,k})_{\text{obs}}$. We require $\bar{i} \leq 90^\circ$ in the fit, which is not a loss of generality due to the symmetry of the transit chord lengths on the two stellar hemispheres bisected by the $b = 0$ chord. The best-fit values of \bar{i} and γ uniquely yield i_1 , which we then use to solve for the other planets’ inclinations from Equation (5). Finally, we measure the sky-plane inclination dispersion as the rms of $\Delta i_k = i_k - \bar{i}$:

$$\sigma_i = \sqrt{\frac{1}{N} \sum_{k=1}^N (i_k - \bar{i})^2}. \quad (6)$$

We illustrate the transit chord ratio method using the TRAPPIST-1 system (Gillon et al. 2017) as an example. With well-determined parameters for all seven planets (Agol et al. 2021), the TRAPPIST-1 system allows us to compare the constraints on the mean inclination \bar{i} and dispersion σ_i from the transit chord ratio method to those obtained from independent techniques. We adopt the values of ρ_\star and each planet’s T and P from Agol et al. (2021). Figure 2 shows the observed ratios, $(\xi_{1,k})_{\text{obs}} = (T_k/T_1)(P_k/P_1)^{-1/3}$, along with the model curve with best-fit parameters, $\bar{i} = 89.781^\circ$ and $\gamma = 0.993$. The observed values are tightly distributed around the best-fit curve, yielding a dispersion equal to $\sigma_i = 0.105^\circ$. The calculated estimates of \bar{i} and σ_i agree well with those derived from the photodynamical model in Agol et al. (2021), where the mean and dispersion of the i values are $89.783^\circ \pm 0.032^\circ$ and $0.053^\circ \pm 0.034^\circ$.

The TRAPPIST-1 system is exceptionally close to coplanar. By experimenting with synthetic systems with larger inclination dispersion (such as some of the systems that will be described in the next section), we identified that when $\sigma_i \gtrsim 0.4^\circ$, the fit sometimes forces \bar{i} to nearly its maximum possible value, 90° . This occurs when several planets have $|b_k| < |b_1|$, which leads to the corresponding ratios $(\xi_{1,k})_{\text{obs}} > 1$. The model attempts to capture these points in the fit (along with those with $(\xi_{1,k})_{\text{obs}} < 1$) by forcing \bar{i} to 90° , which corresponds to a horizontal line of $(\xi_{1,k})_{\text{mod}}$ versus P . To remedy these cases, we utilize a second estimate of \bar{i} that is derived by first estimating the inclination for each planet directly from its measured transit duration (Equation (4)) and then taking the mean over all planets, \bar{i}_{circ} . For cases where the transit chord ratio method yields $\sigma_i > 0.4^\circ$ and $\bar{i} > 89.99^\circ$, we replace the original value of \bar{i} with $(\bar{i} + \bar{i}_{\text{circ}})/2$. Whenever we replace \bar{i} , we also recalculate σ_i using Equation (6).

3.2. Validation Using Simulated Planetary Systems

We can examine the accuracy of the transit chord ratio method using simulated planetary systems, for which we can compare the derived mean inclination and inclination dispersion to the system’s true values. We consider simulated planetary systems from the SysSim (short for “Planetary Systems Simulator”) forward-modeling framework⁹ (Hsu et al. 2018, 2019; He et al. 2019, 2020, 2021). SysSim is a tool used

⁹ The core SysSim code and the specific forward model explored in this study are available at <https://github.com/ExoJulia/ExoplanetsSysSim.jl> and <https://github.com/ExoJulia/SysSimExClusters>.

to generate simulated planetary systems according to flexible statistical models, the parameters of which are determined using a calibration to summary statistics of the observed Kepler planet population (including the observed distributions of multiplicities, orbital periods and period ratios, transit depths and depth ratios, etc.).

It is important to distinguish between two different planetary populations that are generated within a SysSim forward model. First, there is the “physical catalog,” the underlying population of planetary systems, which is drawn directly from the statistical model. Second, there is the “observed catalog,” which is obtained by passing the physical catalog through a simulated Kepler detection pipeline and determining which planets would be detected and labeled as planet candidates during the automated vetting process. Once the best-fit parameters of the forward model have been identified (see, e.g., Hsu et al. 2019 and He et al. 2019, for details), the observed catalog closely resembles the population of observed Kepler systems.

In this study, we will consider simulated planetary systems from the latest edition of SysSim: the “maximum AMD model” (He et al. 2020). This model is based on the argument that a system’s long-term orbital stability is closely related to its angular momentum deficit (AMD), the difference between the total orbital angular momentum of the system and what it would be if all orbits were circular and coplanar (e.g., Laskar & Petit 2017). The key assumption of the model is that all systems have a critical AMD for stability. Although we do not expect this to be true for all systems in reality, the model offers an intuitive framework for assigning system orbital properties and has been shown to reproduce many aspects of the Kepler data (He et al. 2020), including Kepler planet transit duration variations, which probe mutual orbital inclinations (Millholland et al. 2021). In this work, we will use pairs of physical and observed catalogs generated from the maximum AMD model. Each catalog pair corresponds to a different set of model parameters sampled from the posterior distributions derived by He et al. (2020).

We utilize 10 physical/observed catalog pairs and consider only the systems with four or more detected planets (“4+ systems”), equaling 374 systems in total. The observed catalogs contain estimates of T and P of the simulated observed planets with realistic measurement precision. They also contain the “true” values of the stellar density (i.e., assuming perfect measurement precision). In order to simulate realistic uncertainties of the stellar density measurements, we utilize the fractional uncertainties, $f_{\rho_*} = \sigma_{\rho_*}/\rho_*$, of Kepler host stars from the Gaia–Kepler Stellar Properties Catalog (Berger et al. 2020a), where we consider only the subset of host stars with four or more observed planets. The median f_{ρ_*} of this subsample is 9%. For each $\rho_{*,\text{true}}$ value in the SysSim catalog, we draw a random fractional uncertainty, $f_{\rho_*,\text{rand}}$. We then draw a new value of the density as $\rho_* \sim \mathcal{N}(\rho_{*,\text{true}}, \rho_{*,\text{true}} \times f_{\rho_*,\text{rand}})$. It is important to note that the transit chord ratio method is not very sensitive to ρ_* because the process of using ratios of transit chord lengths allows this dependency to partially cancel out.

We use our method to estimate \bar{i} and σ_i in each SysSim 4+ system using T , P , and ρ_* values as inputs. Figure 3 shows the comparison between the derived values and true values of \bar{i} and σ_i . There is good agreement overall, with relatively tight clustering of the calculated versus true values around the

one-to-one line. The median and 16th and 84th percentiles of the distribution of differences in σ_i are $-0.04^{+0.16}_{-0.24}$. In addition, the offset of the calculated σ_i values from the true values is weakly correlated with the true \bar{i} . This occurs because systems with \bar{i} far from 90° are more likely to have all planets transiting on one hemisphere of the star, which yields σ_i values that are more accurate but biased high due to the effects of the unmodeled eccentricities. On the contrary, systems with \bar{i} close to 90° are more likely to have some planets transiting on separate hemispheres, which is not captured within the transit chord ratios and leads to underestimated σ_i values. As for the \bar{i} estimates, the median and 16th and 84th percentiles of the distribution of differences in \bar{i} are $0.03^{+0.29}_{-0.28}$. We also compare (but do not plot) the true values of σ_i with $\sigma_{i,\text{inv}}$, the dispersion of the inclinations with respect to the invariable plane. We find that median $(\sigma_{i,\text{inv}}/\sigma_i) \approx 1.6$.

Before proceeding to utilize \bar{i} and σ_i to calculate transit probabilities, we introduce an additional scaling factor to σ_i . Recall that σ_i is the sky-plane inclination dispersion of the observed *transiting* planets. By definition, σ_i does not contain information about the sky-plane inclinations of the *nontransiting* planets. In order to approximately rectify this, we consider each SysSim system with four or more detected planets, and we calculate the sky-plane inclination dispersion, $\sigma_{i,\text{all}}$, of *all* planets in the system (i.e., including undetected planets) using the true values of the inclinations. We then compute the ratio of $\sigma_{i,\text{all}}$ and the calculated σ_i of the detected planets (as defined in Equation (6)). We find that median $(\sigma_{i,\text{all}}/\sigma_i) = 1.2$. We thus define a scaled sky-plane inclination dispersion, $\sigma'_i = 1.2\sigma_i$, that now approximately reflects the spread of inclinations of the full system. We will use σ_i whenever we are referring to the dispersion measured from the detected planets and σ'_i whenever we are working in the context of transit probabilities, as in the next section.

4. Model of Transit Probabilities and Detection Probabilities

4.1. Transit Probabilities

Once the mean inclination, \bar{i} , and scaled inclination dispersion, σ'_i , have been estimated for a given system, it is possible to calculate the probability that a hypothetical planet with an arbitrary period transits the star. This requires an assumption that the inclinations of individual planets in the system are well described by some distribution with mean and standard deviation equal to \bar{i} and σ'_i . While there are multiple distributions one could consider, we find that a normal distribution, $i \sim \mathcal{N}(\bar{i}, \sigma'_i)$, is a good approximation.

The calculation of the transit probability first requires specifying the bounded range of inclinations for which a planet is transiting. Again assuming circular orbits, a planet transits with $|b| < b_{\text{max}}$ if its inclination falls between $i_{\text{min}} < i < \pi - i_{\text{min}}$, where

$$i_{\text{min}} \equiv \arccos \left[b_{\text{max}} \left(\frac{3\pi}{G\rho_*P^2} \right)^{1/3} \right]. \quad (7)$$

The probability, p_{trans} , that a planet with period P transits with $|b| < b_{\text{max}}$ is thus

$$p_{\text{trans}} = \text{Prob}(i_{\text{min}} < i < \pi - i_{\text{min}}); \quad i \sim \mathcal{N}(\bar{i}, \sigma'_i). \quad (8)$$

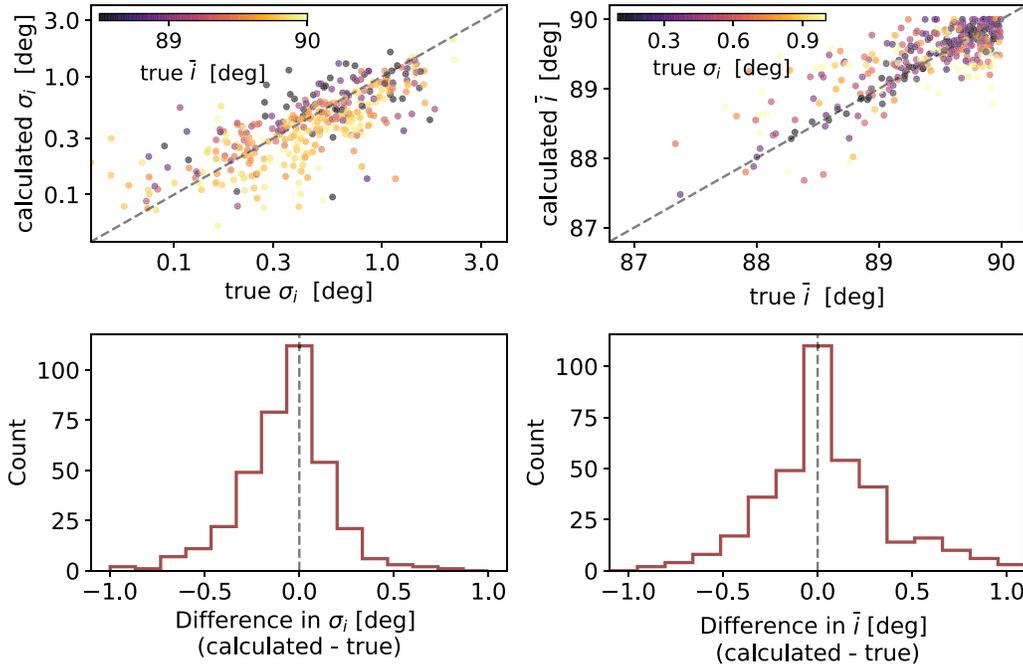


Figure 3. Calculated values vs. true values of the sky-plane inclination dispersion, σ_i , and mean inclination, \bar{i} , of 374 simulated SysSim systems with four or more detected planets. True values of \bar{i} greater than 90° are reported as $180^\circ - \bar{i}$, corresponding to a reflection across the $b = 0$ chord. The top-left panel shows the calculated values of σ_i from the transit chord ratio method vs. the true values, with the color bar corresponding to the true mean inclination. The bottom-left panel shows the distribution of differences in σ_i values. The top-right panel shows the calculated values of \bar{i} vs. the true values, with the color bar corresponding to the true inclination dispersion. The bottom-right panel shows the distribution of differences in \bar{i} .

We illustrate the transit probability calculations using two examples of SysSim systems with four detected planets. Figure 4 shows the inclination range within which a planet would transit with $|b| < b_{\max} = 1$, as well as the transit probability assuming that $i \sim \mathcal{N}(\bar{i}, \sigma'_i)$. The transit probability is unity out to a period large enough such that the transiting region crosses over the region of higher probability according to $i \sim \mathcal{N}(\bar{i}, \sigma'_i)$. There is close agreement between the two transit probability curves associated with the calculated (solid purple) and true (dashed purple) values of \bar{i} and σ'_i . As in these example SysSim systems, each system has its own unique curve of transit probability versus P given the set of parameters ρ_* , \bar{i} , and σ'_i . One can thus evaluate the transit probability of observed or hypothetical planets at arbitrary periods.

4.2. Detection Probabilities

Even if a planet is transiting, it is not guaranteed to be detected. Here we outline the calculation of the detection probability, which depends on the statistical significance of the transit signal. The Kepler transiting planet search (TPS) pipeline required a planet to transit at least three times and to have a statistical significance exceeding a threshold of 7.1σ in order to be detected (Christiansen et al. 2020). The statistical significance of the transit signal is computed with the multiple event statistic (MES), which cannot be specified exactly without applying the TPS pipeline to a given lightcurve but can be approximately computed using the one-sigma depth function (OSDF). The OSDF is a data product provided by Kepler Data Release 25 (DR25; Thompson et al. 2018) that quantifies the transit signal that would be expected to result in an MES equal to unity for a given target star, orbital period, and transit duration, after averaging over the epoch of transit

(Burke & Catanzarite 2017). Thus, the expected MES is simply

$$\text{MES} = \delta / \text{OSDF}, \quad (9)$$

where δ is the transit depth. Kepler DR25 provided large tables of OSDF values for each target star as a function of 14 transit durations and $\sim 10^4$ orbital periods; we use the downsampled versions of these tables provided by Hsu et al. (2019). Given a target star and a planet's P and T , we use bilinear interpolation to the star's OSDF table and extrapolate whenever P or T are beyond the range.

We can now define a detection-efficiency model. We build upon the model from Hsu et al. (2019), which combines the probability of a transiting planet being detected and passing vetting, such that it is labeled a planet candidate. The model is calibrated to the Kepler DR25 pixel-level transit injection tests (Christiansen 2017) and corresponding robovetter results (Coughlin 2017). The probability of the planet passing through both detection and vetting is (Hsu et al. 2019)¹⁰

$$p_{\text{det \& vet}}(\text{MES}, N_{\text{tr}}) = c_{N_{\text{tr}}} \frac{\gamma(\alpha_{N_{\text{tr}}}, \text{MES}/\beta_{N_{\text{tr}}})}{\Gamma(\alpha_{N_{\text{tr}}})}, \quad (10)$$

where N_{tr} is the number of valid transits observed by Kepler, and $\alpha_{N_{\text{tr}}}$, $\beta_{N_{\text{tr}}}$, and $c_{N_{\text{tr}}}$ are N_{tr} -dependent parameters, provided in Table 1. The number of transits can be approximated as $N_{\text{tr}} = \text{floor}(t_{\text{obs}} f_0 / P)$, where t_{obs} and f_0 are the target-specific data span and duty cycle.

We apply additional modifications to our model to account for biases that can occur specifically in multiplanet systems, as shown by Zink et al. (2019). For example, the Kepler TPS pipeline masks each existing transit detection before searching

¹⁰ Note that there is a typo in this equation in Hsu et al. (2019); there, the equation contains $\beta_{N_{\text{tr}}} \times \text{MES}$ in place of $\text{MES}/\beta_{N_{\text{tr}}}$.

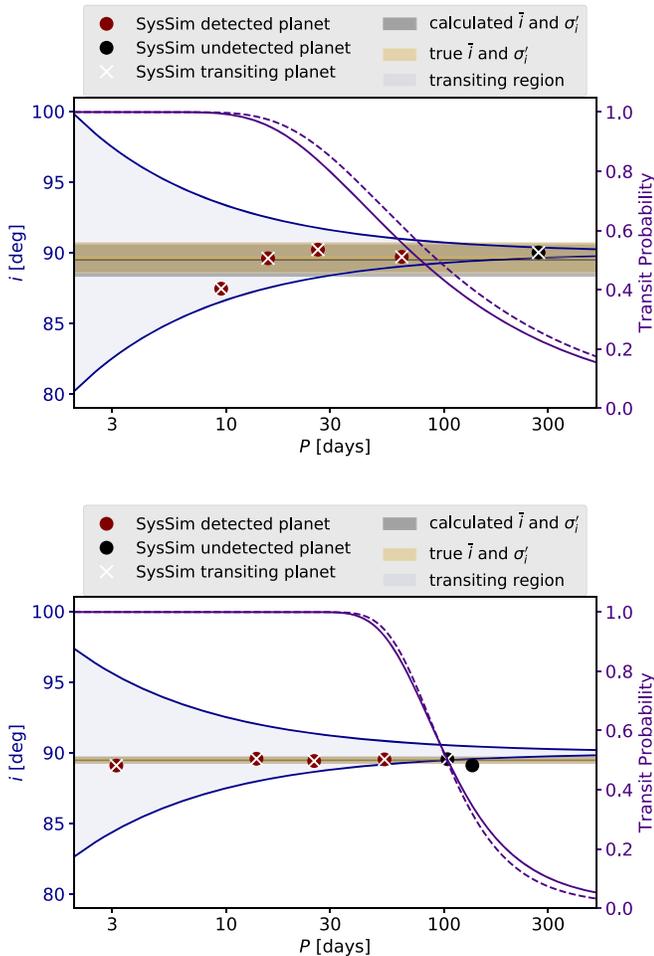


Figure 4. Transit probability as a function of orbital period for two example SysSim simulated systems. The blue curves and solid envelope regions (left y-axis) show the inclinations at which a planet would transit with $|b| < b_{\max} = 1$ as a function of P (Equation (7)). The gray (yellow) lines and banding indicate the calculated (true) values of \tilde{i} and σ'_i . The purple curves (right y-axis) show the transit probability, p_{trans} , as a function of period (Equation (8)), with the solid (dashed) curve corresponding to the calculated (true) values of \tilde{i} and σ'_i . The maroon (black) dots indicate the detected (undetected) planets in the synthetic SysSim system. The white “x” points indicate the transiting planets in the system.

for the next highest S/N candidate, reducing the available photometry for lower S/N candidate searches. These completeness issues are relevant since the theoretical planets considered in this study have the longest periods in their systems, with the fewest number of available transits. To ensure our estimates are conservative, we consider the case where these additional biases impact all the outermost planets.

The completeness functions provided by Zink et al. (2019) do not account for vetting as done by Hsu et al. (2019). However, the multiplicity completeness will not be improved through the vetting process. To renormalize the Zink et al. (2019) results for vetting, we take the difference of the Zink et al. (2019) multi- and single-planet completeness functions and subtract that from the Hsu et al. (2019) values. We then fit for the new parameters. The results of this manipulation are provided in Table 1. We acknowledge that a more accurate measure of multiplicity completeness would involve an injection/recovery test of multiplanet systems through the entirety of the Kepler TPS pipeline, a procedure which is beyond the scope of this paper. Our estimates provide a

Table 1
Parameters for $p_{\text{det\&vet}}$ (Equation (10))

N_{tr}	$\alpha_{N_{\text{tr}}}$	$\beta_{N_{\text{tr}}}$	$c_{N_{\text{tr}}}$
Hsu et al.’s (2019) values			
3	33.3884	0.264472	0.699093
4	32.8860	0.269577	0.768366
5	31.5196	0.282741	0.833673
6	30.9919	0.286979	0.859865
7 – 9	30.1906	0.294688	0.875042
10 – 18	31.6342	0.279425	0.886144
19 – 36	32.6448	0.268898	0.889724
≥ 37	27.8185	0.32432	0.945075
New values			
3	21.6423	0.493516	0.535406
4	30.4538	0.317441	0.713207
5	29.7524	0.324572	0.778662
6	28.8620	0.333193	0.804942
7 – 9	28.1473	0.341453	0.820238
10 – 18	28.4430	0.335102	0.831143
19 – 36	28.3006	0.334147	0.834582
≥ 37	27.4994	0.352761	0.890567

Note. We show the original values from Hsu et al. (2019) and our updated values, after accounting for the effects of planet multiplicity.

first-order approximation, and the results in the following sections exceed any expected changes from a more refined completeness assessment.

Before returning the final probability of detecting a planet, the probability in Equation (10) is multiplied by an additional factor, called the window function (WF), or the fraction of unique transit ephemeris epochs that permit three or more transits to be observed as a function of orbital period (Burke & Catanzarite 2017). This factor is necessary and complex because it provides the probability that a planet has $N_{\text{tr}} \geq 3$ given existing data gaps within the limited span of Kepler photometry, whereas Equation (10) assumes $N_{\text{tr}} \geq 3$ to be the case. Similar to the OSDFs, Kepler DR25 provided tabulations of WFs for each star as a function of T and P . We again use the downsampled versions provided by Hsu et al. (2019). We use bilinear interpolation to calculate the target star’s WF as a function of P and T , extrapolating whenever these values are beyond the range.

5. Validation Tests with Simulated Planetary Systems

The previous section outlined the ingredients necessary to calculate an expected distribution of detections of outer planets orbiting beyond observed planets in multiplanet systems. Before applying this to the Kepler compact multis (Section 6), we will first test our methods on subpopulations of simulated planets from the SysSim maximum AMD model (He et al. 2020). We will essentially “resimulate” the transit and detection observations using the methods developed in Section 4 as a means of validating that our approach results in accurate total numbers of transiting and detected planets. We will specifically examine outer planets in SysSim systems with four or more detected planets (“4+ systems”) and verify that our calculations are in line with the true properties of the synthetic systems.

We begin with the same 10 pairs of SysSim physical catalogs and observed catalogs as used in Section 3.2. Recall

that a single physical/observed catalog pair represents an entire Kepler population of planetary systems orbiting FGK dwarfs. We consider the 4+ systems only. In each system, we identify the innermost undetected planet beyond the detected planets, which we will call the “next outer undetected planet.” For instance, in each example system in Figure 4, the fifth planet from left is the next outer undetected planet. Not all systems have such a planet; in $\sim 25\%$ of systems, the outermost planet is a detected planet, so the next outer undetected planet is undefined. We focus on the remaining $\sim 75\%$ of systems for which this planet is defined. The true number of detections in this subsample is zero by definition, so when we calculate the expected number of detections, we should get a result that is consistent with zero.

For each physical/observed catalog pair, we calculate the distribution of the number of next outer undetected planets that we would expect to be transiting and detectable. We use the following steps:

1. For each next outer undetected planet, we estimate the transit probability, p_{trans} , according to Equation (8), where \bar{i} and σ'_i are either the true values or the calculated values from the transit chord ratio method.
2. We approximate each planet’s transit duration and transit depth (if it were to transit) using scalings to the duration and depth of the outermost observed planet:

$$\begin{aligned} T_{\text{next}} &= T_{\text{out}} (P_{\text{next}}/P_{\text{out}})^{1/3}, \\ \delta_{\text{next}} &= \delta_{\text{out}} (R_{p,\text{next}}/R_{p,\text{out}})^2. \end{aligned} \quad (11)$$

3. Using the period, P_{next} , transit duration, T_{next} , and transit depth, δ_{next} , we estimate the OSDF, WF, and MES for each planet using the target-specific data tables and the procedure described in Section 4.2.
4. We use the MES and WF to estimate the detection probability, $p_{\text{det \& vet}}$, using Equation (10).
5. Using the calculated probabilities, we draw two Bernoulli random variables,

$$\begin{aligned} X_{\text{trans}} &\sim \text{Bernoulli}(p_{\text{trans}}) \\ X_{\text{detect}} &\sim \text{Bernoulli}(p_{\text{trans}} \times p_{\text{det \& vet}}), \end{aligned} \quad (12)$$

with X_{trans} indicating whether the planet is transiting and X_{detect} indicating whether the planet is both transiting and detected.

6. The sum of Bernoulli random variables across all the systems in a given catalog pair gives a number of “successes” (transits or detections) in that catalog,

$$\begin{aligned} N_{\text{trans}} &= \sum_{N_{4+}} X_{\text{trans}} \\ N_{\text{detect}} &= \sum_{N_{4+}} X_{\text{detect}}. \end{aligned} \quad (13)$$

7. Finally, by repeating steps five and six 1000 times, we create distributions of N_{trans} and N_{detect} for each catalog pair (with 1000 values in each distribution).

Figure 5 shows the distributions of N_{trans} and N_{detect} of the next outer undetected planets across 10 SysSim catalog pairs. The first thing to note is that the distributions of N_{trans} and N_{detect} , which use calculated values of \bar{i} and σ'_i for the p_{trans} calculation (blue histograms), agree well with the corresponding distributions that use the true values of \bar{i} and σ'_i (yellow

histograms). Thus, we can trust that using calculated values of these key parameters will provide accurate distributions when we proceed to working with the Kepler multis.

A second observation from Figure 5 is that the distributions of N_{trans} and N_{detect} are in reasonable agreement with the true values of N_{trans} and N_{detect} . (Recall that the true value of N_{detect} is always zero because these experiments are conditioned upon predictions for the first outer *undetected* planet in each system.) The distributions are consistent with the true values to within $\sim 1\sigma$ – 3σ , with the average being $\sim 2.2\sigma$ for N_{trans} and $\sim 2.4\sigma$ for N_{detect} . However, the distributions are always biased high. This is by construction of our exploration. Because the next outer undetected planets were missed within the actual SysSim simulations, they are more likely to be nontransiting than other planets in the same system. Thus, our assumption of $i \sim \mathcal{N}(\bar{i}, \sigma'_i)$ in p_{trans} (Equation (8)) partially breaks down and overestimates the transit probability. Regardless, the fact that the distributions agree with the true N_{trans} and N_{detect} to within $\lesssim 3\sigma$ ensures that we can reliably use these methods on the observed Kepler multis, as long as any results we find are significant by much more than 3σ .

6. Detection Expectations for Hypothetical Outer Planets in Kepler Multiplanet Systems

In the previous section, we used populations of simulated planetary systems to validate our calculated expectations of the number of planets that would be transiting and detected. We now extend these methods to Kepler high-multiplicity systems. We aim to answer the questions: If the Kepler multis hosted additional outer planets with similar properties as the known planets, how many of them would we expect to be transiting and detectable? Moreover, are these expectations reconcilable with the lack of detections beyond the outermost known transiting planets?

6.1. Sample of Kepler Multiplanet Systems

We begin by defining our sample of Kepler multiplanet systems. We use the Kepler DR25 KOI catalog (Thompson et al. 2018; NASA Exoplanet Archive 2022) as our starting point, using all planets with “confirmed” and “candidate” dispositions. Where possible, we replace the stellar parameters and planet radii in the DR25 catalog with parameters from the Gaia–Kepler Stellar Properties Catalog (Berger et al. 2020a, 2020b). In addition, we apply a small set of quality cuts. We consider only planets smaller than $16 R_{\oplus}$ with fractional radius uncertainties less than 100%. To avoid stars with large systematic radius errors, we discard targets for which Furlan et al. (2017) found a companion star that contributed more than 5% of the light in the photometric aperture. After these cuts, we are left with 64 Kepler systems with four or more observed transiting planets (“4+ systems”).

We now apply the transit chord ratio method (Section 3) to the transiting planets in the Kepler 4+ systems to estimate the mean sky-plane inclination, \bar{i} , and the sky-plane inclination dispersion, σ_i (from which the scaled dispersion $\sigma'_i = 1.2\sigma_i$ is subsequently derived). Figure 6 shows the distributions of \bar{i} and σ_i for the Kepler 4+ systems and the SysSim 4+ systems examined in Sections 3.2 and 5. For the SysSim data, we show both the calculated values from the transit chord ratio method and the true values. Overall, the distributions for the Kepler

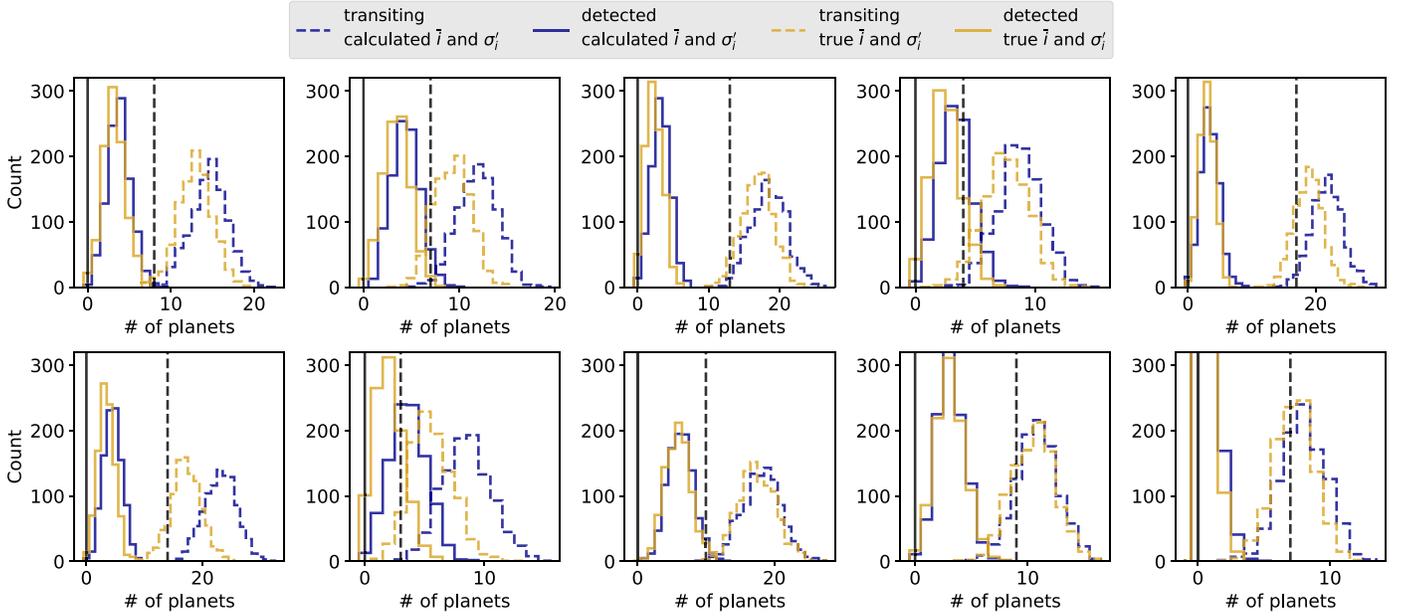


Figure 5. Distributions of the number of next outer undetected planets that we would expect to be transiting, N_{trans} , and detected, N_{detect} , in systems with four or more detected planets within 10 SysSim catalogs. Each panel represents a different catalog. The dashed histograms represent N_{trans} , while the solid histograms represent N_{detect} . The blue histograms use calculated values of \tilde{i} and σ'_i (from the transit chord ratio method) for the p_{trans} calculation, while the yellow histograms values use true values of \tilde{i} and σ'_i . The dashed (solid) vertical black lines correspond to the true numbers of next outer undetected planets that are transiting (detected) in the catalog.

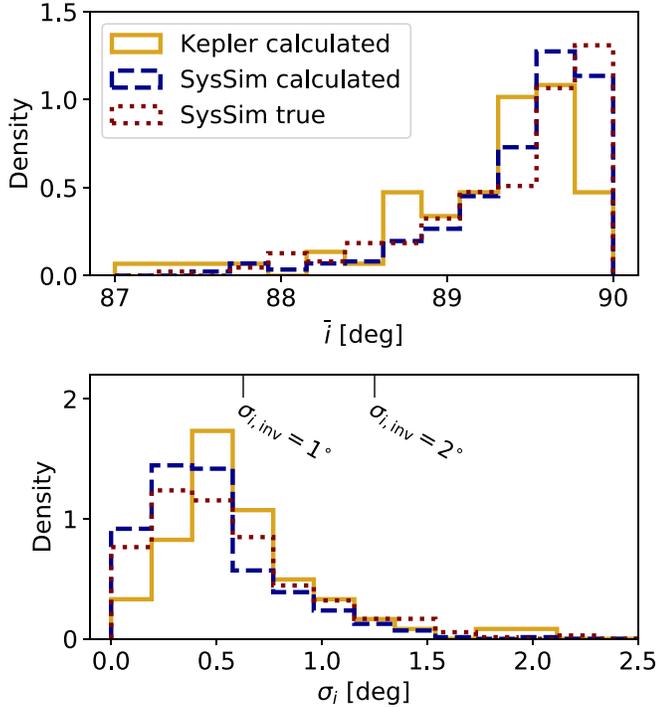


Figure 6. Comparison of \tilde{i} and σ_i distributions for Kepler 4+ systems and SysSim 4+ systems. The top (bottom) panels show the distributions of \tilde{i} (σ_i), with the calculated values for Kepler multis in solid yellow, the calculated values for SysSim multis in dashed blue, and the true values for SysSim multis in dotted maroon. True values of \tilde{i} greater than 90° are reported as $180^\circ - \tilde{i}$, corresponding to a reflection across the $b = 0$ chord. Benchmark values of the dispersion of the inclinations with respect to the invariable plane, $\sigma_{i,\text{inv}} \approx 1.6\sigma_i$ (Section 3.2), are also labeled.

systems bear close resemblance to those for the SysSim systems. This offers another validation of the \tilde{i} and σ_i estimates, which will be used in transit probability calculations in the next section.

6.2. Set-up of Hypothetical Planet Experiments

We now construct several experiments consisting of “hypothetical planets” orbiting beyond the observed planets in Kepler 4+ systems. In general, the hypothetical planets are designed to emulate the continuation of the “peas-in-a-pod” architectures (Weiss et al. 2018; Millholland et al. 2017) by positing the existence of one additional outer planet in each system. We first assign the periods, P_{hypo} , and radii, $R_{p,\text{hypo}}$, of the hypothetical planets. We consider three separate approaches, outlined below, and denoted “fixed,” “random,” and “targeted” sampling.

Fixed sampling. The most straightforward approach in assigning P_{hypo} and $R_{p,\text{hypo}}$ is to consider that they perfectly follow the “peas-in-a-pod” patterns (Weiss et al. 2018; Millholland et al. 2017), or the statistical tendency for systems to exhibit intrasystem uniformity in planetary radii and period ratios. Accordingly, we can assign

$$\begin{aligned} R_{p,\text{hypo}} &= R_{p,\text{out}} \\ P_{\text{hypo}}/P_{\text{out}} &= P_{\text{out}}/P_{2\text{nd-out}}, \end{aligned} \quad (14)$$

where P_{out} and $R_{p,\text{out}}$ are the period and radius of the outermost known planet, and $P_{2\text{nd-out}}$ is the period of the second outermost known planet. (This is identical to the initial experiments shown in Section 2.)

Random sampling. The patterns of intrasystem uniformity of radii and period ratios have an inherent degree of scatter (e.g., Millholland & Winn 2021). Thus, a second and more physical approach to assigning P_{hypo} and $R_{p,\text{hypo}}$ is to utilize empirical distributions of radius ratios and ratios of period ratios from our sample. We assign

$$\begin{aligned} R_{p,\text{hypo}}/R_{p,\text{out}} &= R_{p,i+1}/R_{p,i}, \\ P_{\text{hypo}}/P_{\text{out}} &= P_{\text{out}}/P_{2\text{nd-out}} \left(\frac{P_{i+1}/P_i}{P_i/P_{i-1}} \right), \end{aligned} \quad (15)$$

where $R_{p,i+1}/R_{p,i}$ is randomly drawn from the observed distribution of the radius ratios of adjacent planets, and $(p_{i+1}/p_i)/(p_i/p_{i-1})$ is drawn from the observed distribution of the ratios of period ratios of adjacent planet pairs. However, in an effort to preserve covariances of the distributions with respect to other variables (e.g., correlations between $R_{p,i+1}/R_{p,i}$ and p_{i+1}/p_i), our random draws are not taken from the distributions as a whole but rather from predefined subsets. We bin the distribution of $R_{p,i+1}/R_{p,i}$ in five bins of p_i and four bins of p_{i+1}/p_i with an average of about 40 points in each bin. The bins are chosen so as to capture the large-scale correlations within the distribution, and the overall results are not sensitive to them. When sampling from the distribution of $R_{p,i+1}/R_{p,i}$, we identify the appropriate bin and only sample from the corresponding subset of the distribution. Similarly, we bin the distribution of $(p_{i+1}/p_i)/(p_i/p_{i-1})$ in six bins of the inner period ratio, p_i/p_{i-1} , and we sample from the distribution by first identifying the appropriate bin and only drawing from the corresponding subset of the distribution.

Targeted sampling. Finally, a third approach of assigning P_{hypo} and $R_{p,\text{hypo}}$ is designed to allow us to systematically vary these parameters and observe the resulting changes to the distributions of N_{trans} and N_{detect} . Specifically, we sample $R_{p,\text{hypo}}/R_{p,\text{out}} \in [0.2, 1]$ and $(P_{\text{hypo}}/P_{\text{out}})/(P_{\text{out}}/P_{2\text{nd-out}}) \in [1, 4]$, thus probing configurations in which the hypothetical planet is either smaller or further away than the “peas-in-a-pod” expectation. We target P_{hypo} and $R_{p,\text{hypo}}$ one at a time using this sampling method, with the other parameter assigned using the “fixed sampling” approach.

In addition to the targeted sampling of P_{hypo} and $R_{p,\text{hypo}}$, we also extend our targeted sampling experiments to two other parameters. First, we consider assigning hypothetical planets to only a fraction of systems, rather than all systems. We sample this fraction in the range $f_{\text{hypo}} \in [0.1, 1]$. This allows us to investigate how many systems can have additional outer planets and still be consistent with the observations. Lastly, in order to explore the effects of potentially biased estimations of the sky-plane inclination dispersion, we consider a constant scaling factor, $\gamma_{\sigma'_i}$, such that $\sigma_{i,\text{new}}' = \gamma_{\sigma'_i} \sigma'_i$. We sample the scaling factor in the range $\gamma_{\sigma'_i} \in [1, 10]$. We reiterate that we use targeted sampling on only one of the four parameters (P_{hypo} , $R_{p,\text{hypo}}$, f_{hypo} , and $\gamma_{\sigma'_i}$) at a time, with all others assigned using the “fixed sampling” approach (which, in the case of the latter two parameters, means $f_{\text{hypo}} = 1$ and $\gamma_{\sigma'_i} = 1$).

With the sampling approach defined and P_{hypo} and $R_{p,\text{hypo}}$ assigned for each hypothetical planet, we proceed to calculate the distribution of the number of these planets that we would expect to be transiting and detectable. We use the same seven steps as outlined in Section 5, where the transit probabilities and detection probabilities are calculated for each hypothetical planet, and repeated trials of summed Bernoulli random variables yield distributions of N_{trans} and N_{detect} .

6.3. Results from Fixed Sampling and Random Sampling

Figure 7 shows the results of the fixed sampling and random sampling approaches. Ten iterations of the distributions are shown, indicating the degree of variability resulting from the sampling of P_{hypo} and $R_{p,\text{hypo}}$ combined with the seven steps outlined in Section 5. Immediately, we can see that the distributions of N_{detect} are inconsistent with zero detections in

both the fixed sampling and random sampling approaches. With the fixed sampling approach, the mean and standard deviation of the composite distribution of N_{detect} (consisting of the 10 separate distributions pooled together) is 27.55 ± 3.39 detected planets, or $\sim 8.1\sigma$ greater than zero detections. This indicates that $\sim 44\%$ of the 64 Kepler 4+ systems would be expected to show detections of the hypothetical outer planets, if they existed.

As for the distributions resulting from the random sampling approach, there is more variability between different iterations due to each iteration having a distinct set of P_{hypo} and $R_{p,\text{hypo}}$ values. The median and standard deviation of the composite distribution of N_{detect} is 22.11 ± 3.25 detected planets, or $\sim 6.8\sigma$ greater than zero detections. This indicates that the hypothetical planets would be detectable in $\sim 35\%$ of the 64 Kepler 4+ systems.

Both the fixed sampling and random sampling approaches indicate the same general result: if we posit the existence of additional planets orbiting beyond the known transiting planets with properties dictated by the expected “peas-in-a-pod” architectures, then these planets would be detectable in roughly $\sim 35\%–45\%$ of systems, yielding a $\sim 7\sigma–8\sigma$ discrepancy with the lack of detections of such planets in the observed systems. However, it is important to note that the distributions of N_{detect} are always positive by definition, so the more meaningful comparison is that with the corresponding results for the SysSim simulated systems (Section 5; Figure 5), where we found that the N_{detect} distributions were biased high but always consistent with zero to within 3σ .

A final point is that this result is a lower limit on the expected number of detections since we only assigned one hypothetical planet per system. If we considered more than one outer planet per system, this would potentially generate an even larger number of expected detections. For instance, one possible experiment would be to add multiple sequential hypothetical planets to each system, assigned as $P_{\text{hypo},i} = P_{\text{out}}(P_{\text{out}}/P_{2\text{nd-out}})^i$ and restricted to $P_{\text{hypo},i} < 500$ days. We find that this would involve a total of 327 hypothetical planets, compared to the 64 in this study.

6.4. Results from Targeted Sampling

The discrepancy between the expected number of detections and the lack of such detections in the real systems indicates that one or more of our assumptions are breaking down. These assumptions and their related implications are enumerated below:

1. We assumed that each system has a single additional planet orbiting beyond the known transiting planets. However, perhaps not all systems have a such a planet. This would imply that the architectures of compact multis are truncated at a detectable orbital period in at least some fraction of systems.
2. We assumed that $P_{\text{hypo}}/P_{\text{out}} \approx P_{\text{out}}/P_{2\text{nd-out}}$. However, perhaps the period ratios are on average larger at sufficiently large orbital period.¹¹

¹¹ We searched for a trend between orbital period and period ratio in the observed sample of Kepler 4+ systems and found none. However, it is still possible that the period ratios increase at orbital periods at or near the edge of detectability.

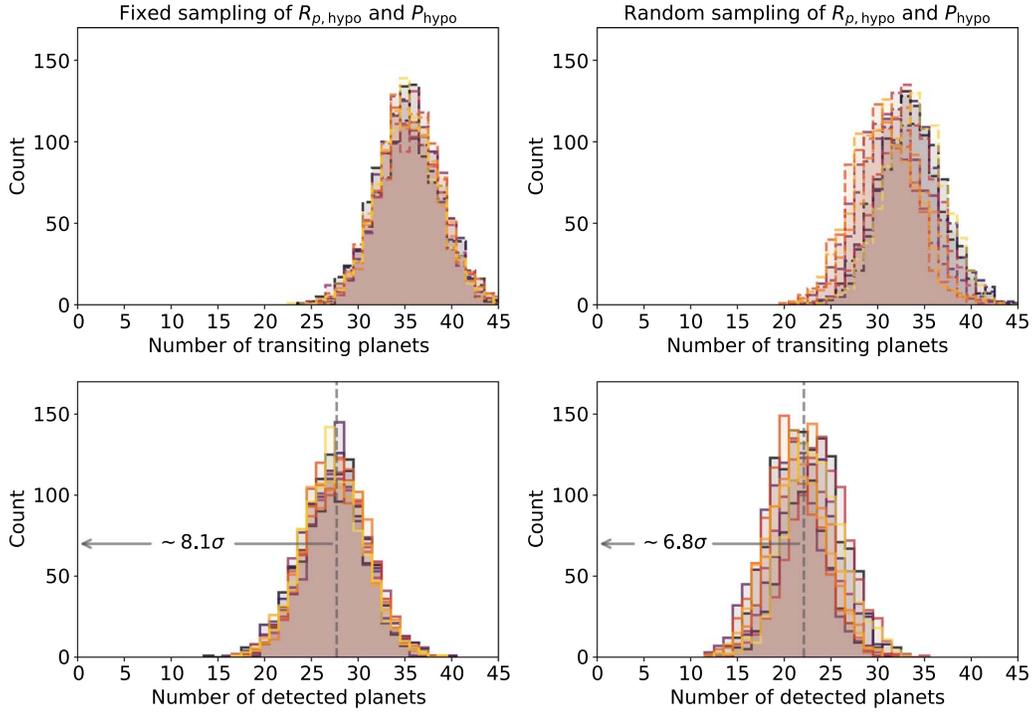


Figure 7. Distributions of the number of “hypothetical planets” that we would expect to be transiting, N_{trans} , and detected, N_{detect} , if they existed at the outer edges of the 64 Kepler systems with four or more known transiting planets. The left column corresponds to the fixed sampling approach with $R_{p,\text{hypo}} = R_{p,\text{out}}$ and $P_{\text{hypo}}/P_{\text{out}} = P_{\text{out}}/P_{2\text{nd-out}}$. The right column corresponds to the random sampling approach with $R_{p,\text{hypo}}/R_{p,\text{out}}$ and $(P_{\text{hypo}}/P_{\text{out}})/(P_{\text{out}}/P_{2\text{nd-out}})$ sampled from empirical distributions. The dashed histograms in the top panels represent N_{trans} , while the solid histograms in the bottom panels represent N_{detect} . The different colors represent 10 separate iterations and have no significance otherwise.

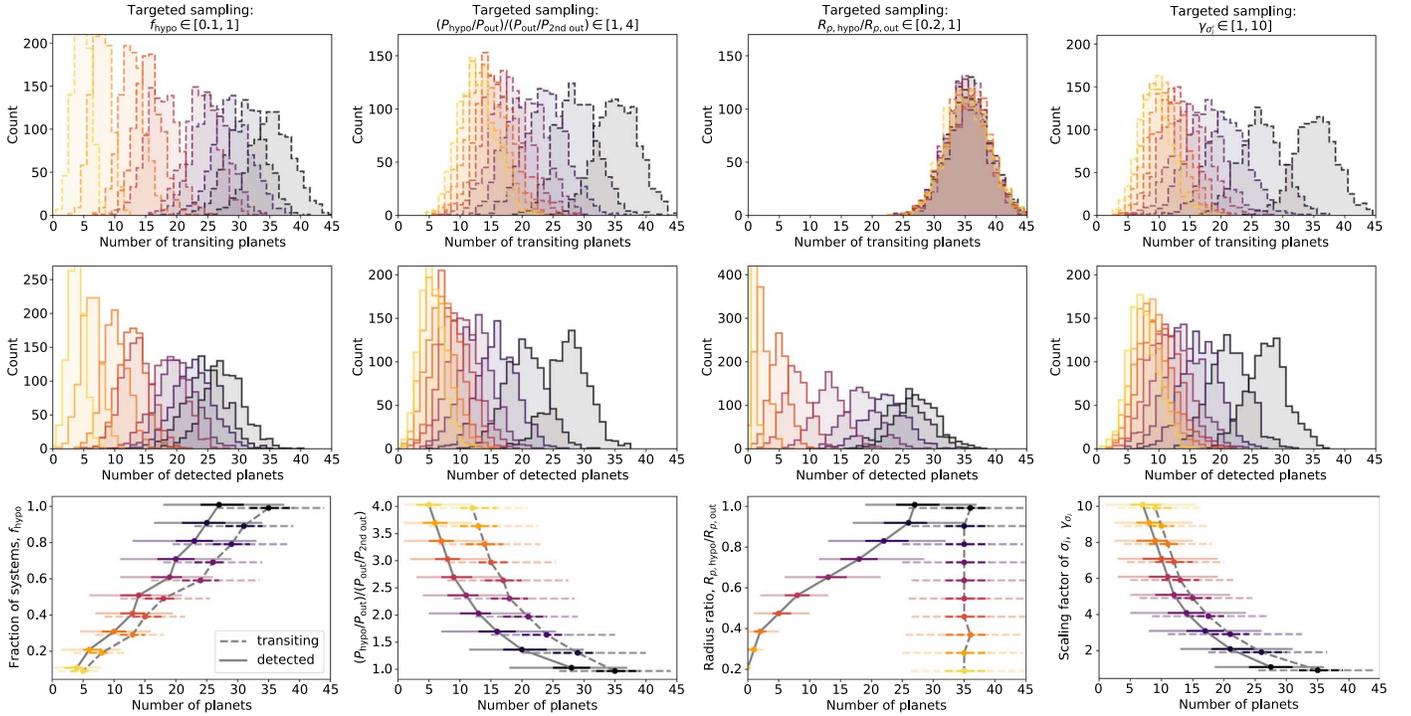


Figure 8. Distributions of the number of “hypothetical planets” that we would expect to be transiting, N_{trans} , and detected, N_{detect} , if they existed at the outer edges of the 64 Kepler systems with four or more known transiting planets. All columns indicate the results of the targeted sampling approach, with the columns corresponding to (from left to right) sampling of f_{hypo} (the fraction of systems with hypothetical planets), $(P_{\text{hypo}}/P_{\text{out}})/(P_{\text{out}}/P_{2\text{nd-out}})$, $R_{p,\text{hypo}}/R_{p,\text{out}}$ and $\gamma_{\sigma'_i}$ (the scaling factor of the inclination dispersion). These parameters are sampled from 10 evenly spaced values within the ranges indicated at the top of the columns. The dashed histograms in the top panels represent N_{trans} , and the solid histograms in the middle panels represent N_{detect} . The colors indicate the value of the targeted parameter, as shown in the bottom panels, where the 1σ and 3σ intervals around the medians of the distributions are indicated with the darker and lighter error bars. The two curves in the bottom panels are artificially offset vertically such that the error bars can be seen clearly.

3. We also assumed that $R_{p,\text{hypo}} \approx R_{p,\text{out}}$. However, perhaps the radii of planets are on average smaller at sufficiently large orbital period.
4. Finally, we also assumed a mutual inclination dispersion for each system based on our estimation from the transit chord ratio method. Although we validated the method in Section 3, it is possible that the σ'_i estimates are systematically low or that the true dispersion increases at large periods.

The targeted sampling approach described earlier in this section allows us to investigate these four possibilities by systematically varying f_{hypo} , P_{hypo} , $R_{p,\text{hypo}}$, and $\gamma_{\sigma'_i}$, which directly map to the four possibilities outlined above. All other parameters are held fixed while the targeted parameters are varied. Figure 8 shows the results of the targeted sampling. The bottom panels summarize the distributions by showing the medians and 1σ and 3σ intervals as a function of the targeted parameter. This allows us to visualize how much the targeted parameter must be varied in order to have the N_{detect} distribution be approximately consistent with zero detections to within 3σ .

As shown in the figure, the N_{detect} distribution is consistent (or, in some cases, nearly consistent) with zero to within 3σ when one or more of the following conditions are met: $f_{\text{hypo}} \lesssim 0.2$, $(P_{\text{hypo}}/P_{\text{out}})/(P_{\text{out}}/P_{2\text{nd-out}}) \gtrsim 3$, $R_{p,\text{hypo}}/R_{p,\text{out}} \lesssim 0.5$, or $\gamma_{\sigma'_i} \gtrsim 6$. These thresholds are both approximate and conservative. The conditions on P_{hypo} and $R_{p,\text{hypo}}$ indicate that the systems would require a significant deviation from the typical “peas-in-a-pod” architectures. Moreover, the condition on f_{hypo} indicates that only $\sim 20\%$ of systems can host an additional outer planet with the expected architecture if the population is to be consistent with the observations. However, we note that the four conditions mentioned above could be relaxed if multiple parameter variations were considered simultaneously (e.g., both larger period ratios and smaller radii).

The condition on the σ'_i scaling factor, $\gamma_{\sigma'_i}$, indicates that our inclination dispersion estimates would have to be very off (by more than a factor of ~ 6) before the distribution of N_{detect} is consistent with zero detections at 3σ . However, this raises questions about how large σ'_i can actually be before it becomes too unlikely that the observed planets are still cotransiting. In other words, perhaps we can rule out σ'_i values that are as large as those required to significantly shift the N_{detect} distribution. We investigate this using synthetic trials in which we systematically increase $\gamma_{\sigma'_i}$ and determine how frequently the number of transiting planets equals the observed number. For 15 equally spaced values of $\gamma_{\sigma'_i}$, we consider each Kepler 4+ system, and we reassign each observed planet in the system a random sky-plane inclination according to $i \sim \mathcal{N}(\bar{i}, \gamma_{\sigma'_i} \sigma'_i)$. We then determine how many planets are transiting, N_{sim} , and compare to the observed number of transiting planets, N_{obs} . We calculate the fraction of all 64 systems for which $N_{\text{sim}} = N_{\text{obs}}$. Finally, we repeat these random trials 100 times.

Figure 9 shows the results of these simulations. The fraction of Kepler 4+ systems for which $N_{\text{sim}} = N_{\text{obs}}$ rapidly decreases as a function of $\gamma_{\sigma'_i}$. Note that the fraction is not unity for $\gamma_{\sigma'_i} = 1$ because the process of resampling the inclinations makes it unlikely to recover all of the exact same planets in transit, given that the observed sample is conditioned upon the planets transiting. With $\gamma_{\sigma'_i}$ as large as ~ 6 (the value required earlier), less than 15% of systems would still have as many

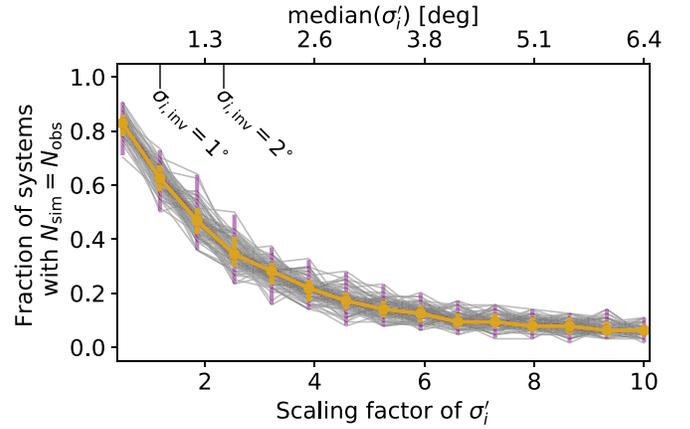


Figure 9. Results of simulations that examine the impact of varying $\gamma_{\sigma'_i}$, the scaling factor of σ'_i , on the transit multiplicity of the observed planets in the Kepler 4+ systems. The y-axis shows the fraction of the 64 Kepler 4+ systems for which the simulated number of transiting planets, N_{sim} , is equal to the observed number of planets in the system, N_{obs} . The gray lines show the individual results from 100 simulations for each value of $\gamma_{\sigma'_i}$. The yellow (purple) lines show the median and the 1σ (3σ) intervals. Benchmark values of the dispersion of the inclinations with respect to the invariable plane, $\sigma_{i,\text{inv}} \approx 1.6\sigma_i$ (Section 3.2), are also labeled.

transiting planets as the observed systems. We can thus safely rule out the possibility that we have underestimated σ'_i by a factor as large as required to make the N_{detect} distribution consistent with zero. However, this still leaves open the possibility that the edge-of-the-multis discrepancy is explained by a substantial increase in the mutual inclinations at long periods (e.g., Shahaf et al. 2021).

7. Discussion

7.1. Constraints on the Truncation Location

The hypothetical planet experiments from the previous section indicate that the outer edges of the observed Kepler 4+ systems likely cannot be sculpted by geometric and detection biases alone. Rather, these systems provide evidence either for an average truncation (i.e., occurrence rate drop-off) in the underlying architectures and/or some breakdown of the “peas-in-a-pod” patterns at larger orbital separations. We have not yet discussed where this transition occurs, since our hypothetical planet experiments were primarily designed to assess whether it exists rather than where it is. Even so, these experiments can still provide us with some insight.

Figure 10 shows the distributions of P_{out} , P_{hypo} assuming the default “peas-in-a-pod” expectation ($P_{\text{hypo}}/P_{\text{out}} = P_{\text{out}}/P_{2\text{nd-out}}$), and P_{hypo} assuming larger period ratios ($P_{\text{hypo}}/P_{\text{out}} = 3P_{\text{out}}/P_{2\text{nd-out}}$), which corresponds to the approximate condition for which the N_{detect} distribution would be consistent with observations according to our targeted sampling experiments (Section 6.4). The medians of the three distributions are 40.6 days, 78.9 days, and 236.6 days, respectively. The distribution of P_{out} shows a steep drop-off at around ~ 100 days. Meanwhile, the third distribution rises beyond ~ 100 days and peaks around ~ 300 days. Since the third distribution corresponds to a set of hypothetical planets that would be consistent with zero detections to within 3σ , this indicates that the average edge must occur somewhere within the range $P_{\text{edge}} \approx 100\text{--}300$ days, or $a_{\text{edge}} \approx 0.5\text{--}1$ au.

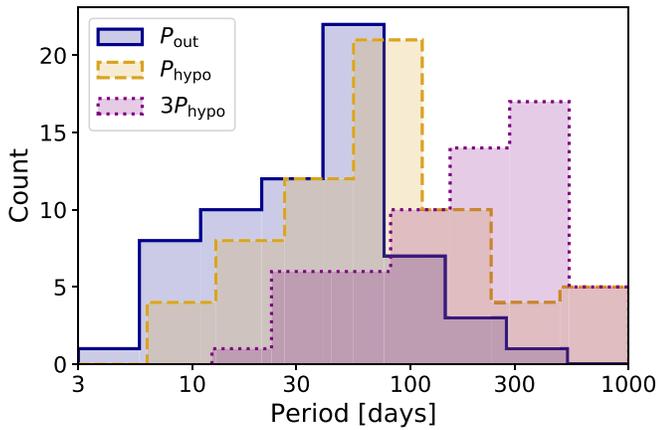


Figure 10. Distributions of orbital periods of outermost planets. Blue/solid: orbital periods of the outermost observed planets in the 64 Kepler 4+ systems. Yellow/dashed: orbital periods of the 64 hypothetical planets assuming $P_{\text{hypo}}/P_{\text{out}} = P_{\text{out}}/P_{2\text{nd-out}}$. Purple/dotted: orbital periods of the 64 hypothetical planets assuming $P_{\text{hypo}}/P_{\text{out}} = 3P_{\text{out}}/P_{2\text{nd-out}}$ (which corresponds to the yellow distribution scaled by a factor of 3). These distributions have not been corrected for completeness.

7.2. Theoretical Interpretations

Whether our results are consistent with an average truncation in the underlying architectures or some breakdown of the “peas-in-a-pod” patterns, these findings have important implications for theories of the formation and dynamical evolution of compact multis. The theoretical interpretations will differ depending on which effect is dominant, but the same physical mechanism could, in principle, lead to both an occurrence rate decrease and a change in the architectures at larger separations. Nevertheless, we discuss outer truncation and a “peas-in-a-pod” breakdown separately below.

7.2.1. Outer Truncation

With regard to outer truncation, one possibility is that, if compact multis form via orbital migration, they may often experience migration traps that prevent inward migration of some planets and cause large gaps in the systems at $\sim 100\text{--}300$ days (Zawadzki et al. 2022; see also, e.g., Paardekooper et al. 2010; Coleman & Nelson 2014; Bitsch et al. 2015; Izidoro et al. 2017; Carrera et al. 2019). Recently, Zawadzki et al. (2022) outlined that migration traps should occur in a time-evolving range of planet mass and semimajor axis (e.g., $M_p \sim 0.5\text{--}2 M_{\oplus}$, $a \lesssim 2$ au at ~ 5 Myr) where the corotation torque dominates over the Lindblad torque and planets thus migrate outwards instead of inwards. Planets in this mass and semimajor axis range are prevented from reaching the inner regions of the disk, whereas more massive planets avoid the traps and end up on close-in and dynamically cold orbits. Although Zawadzki et al. (2022) studied this theoretical framework in the context of possible explanations for the “Kepler dichotomy” (e.g., Lissauer et al. 2011; Johansen et al. 2012), it is relevant that their simulations naturally yield planetary systems that are bifurcated into two clusters with a large gap at $\sim 100\text{--}300$ days, since this is precisely consistent with our observational findings.

Another physical mechanism that may cause outer architecture truncation is a proposed channel of compact multi formation called inside-out planet formation (Chatterjee & Tan 2014, 2015). In this framework, planets are formed sequentially

from successive gravitationally unstable rings of centimeter-to-meter-sized pebbles that drift inwards via gas drag and buildup at the pressure maximum associated with the dead-zone boundary, which separates the inner and outer regions of the disk where the magnetorotational instability is and is not active.¹² Once a planet forms out of the pebble ring, it may migrate or isolate itself from the accretion flow, after which the dead-zone boundary retreats and the process repeats. The process is limited by the finite extent of the dead-zone boundary’s retreat and by the mass reservoir that remains available for planet formation after the initial onset (Hu et al. 2018). Accordingly, this model predicts that planet formation will halt at some orbital separation, creating a break in the outer regions of compact multis.

In contrast to disk effects at the formation epoch, the outer edges of compact multis may also be limited by exterior perturbing planets. Recent results indicate that distant giant planets (with $a \gtrsim 1$ au, $M_p \gtrsim 0.5 M_{\text{Jup}}$) are common in systems with inner super-Earths (Zhu & Wu 2018; Bryan et al. 2019). These distant giant planets are often eccentric and may be mutually inclined with respect to the inner planets (Dawson & Murray-Clay 2013; Masuda et al. 2020). Several authors have shown that distant giant planets can dynamically excite an inner system of super-Earths, provided that the distant perturber exerts a stronger gravitational influence on the inner planets than their mutual gravitational coupling (e.g., Lai & Pu 2017; Pu & Lai 2018; Denham et al. 2019; Spalding & Millholland 2020; Tamayo et al. 2021). The strength of the distant giant planet’s perturbation relative to the inner planets’ coupling depends on the separation between the outer perturber and the inner system (among other parameters). Accordingly, planets in compact multis that extend out to sufficiently large orbital separations may be dynamically excited or even destabilized by a distant giant. However, since distant giants are present in only a fraction of compact multi systems, their presence likely cannot be the sole explanation of our findings.

7.2.2. Breakdown of the “Peas-in-a-pod” Patterns

A different set of theories may be relevant in the case where our results are explained by smaller and/or more widely spaced planets at larger orbital separations. One possibility is that the dominant channel of compact multi formation—whether it is in situ accretion (e.g., Hansen & Murray 2012, 2013; Chiang & Laughlin 2013; Dawson et al. 2015), distant assembly followed by inward migration (e.g., Terquem & Papaloizou 2007; Cossou et al. 2014; Coleman & Nelson 2014; Izidoro et al. 2017; Carrera et al. 2019), or some variation thereof (e.g., Chatterjee & Tan 2014)—might naturally produce smaller planet masses at larger orbital separations ($\gtrsim 0.5\text{--}1$ au) due to limitations of solid material or growth timescales.

For instance, in the planet formation simulations by Izidoro et al. (2017), in which sub-Neptune systems are produced through inward migration and disruption of resonant chains, the resulting systems appear to have a gap at around $\sim 0.5\text{--}1$ au, with the planets beyond the gap being less massive ($\lesssim 1 M_{\oplus}$). This is also seen in other works such as Zawadzki et al. (2022) and is related to the slower migration timescales for the smaller-mass planets. These small planets can also encounter migration

¹² Other theories (e.g., Izidoro et al. 2021) have proposed that planet formation may occur within rings of planetesimals created by other types of pressure maxima in the disk, such as the bump associated with the silicate sublimation line.

traps where the corotation torque is dominant (as discussed in Section 7.2.1), whereas the larger-mass planets avoid the traps and migrate inwards. This process naturally leads to a bifurcation of the system, with larger sub-Neptunes inside of ~ 0.5 –1 au and smaller planets beyond that.

Stepping back from the specific details of planet formation physics, one can consider the general process of energy optimization at the assembly epoch, which also leads to the prediction that the “peas-in-a-pod” patterns should break down at larger orbital separations. Adams et al. (2020; see also Adams 2019) calculated the lowest-energy states available to forming planetary pairs subject to conservation of angular momentum, constant total mass, and fixed orbital spacing. They found that a configuration with approximately equal masses (i.e., “peas-in-a-pod”) is the most energetically favorable when the total mass in the planets is less than some critical threshold, $m_c \sim 40 M_\oplus$. Above this threshold, the optimum state is one in which most of the mass is in one planet. The critical threshold m_c decreases with increasing a , which indicates that mass uniformity is no longer energetically favorable at sufficiently large orbital distances ($\gtrsim 0.5$ –1 au). Thus, the breakdown of the “peas-in-a-pod” patterns might simply be a consequence of planets optimizing the available energy.

To summarize, a variety of physical processes may be responsible for an outer architecture truncation or transition. It is possible that the same physical mechanism (e.g., migration traps) can cause multiple effects (e.g., occurrence rate decrease and smaller planets beyond ~ 0.5 –1 au). We note that the list of theories above is by no means exhaustive, and it is beyond the scope of this paper to investigate all possibilities in detail. This topic is ripe for future theoretical investigations.

8. Conclusion

The inner and outer edges of compact multiplanet systems are fundamental signatures of their formation and evolution. The distribution of the innermost planet orbital periods of the underlying distribution of compact multis (i.e., after accounting for observational biases) peaks around ~ 10 days (Mulders et al. 2018), consistent with being a relic from the protoplanetary disk inner edges (e.g., Terquem & Papaloizou 2007; Lee & Chiang 2017). The outer edges are significantly more affected by observational biases and thus have received no prior in-depth investigations, as far as we are aware. However, these biases are well understood, meaning that a robust characterization is possible.

In this paper, we presented evidence that the outermost planets in the observed Kepler high-multiplicity systems truncate at smaller orbital periods than expected from geometric and detection biases alone. We showed this using experiments on the detectability of “hypothetical planets” orbiting beyond the outermost observed planets in Kepler high-multiplicity systems. These experiments were first demonstrated heuristically in Section 2 and then developed robustly in the remainder of the paper.

We considered one hypothetical outer planet per system for 64 Kepler systems with four or more observed planets, with each hypothetical planet’s properties dictated by expectations from the “peas-in-a-pod” patterns (Weiss et al. 2018; Millholland et al. 2017). Using models of the transit and detection probabilities, we estimated that 22.11 ± 3.25 (approximately $\sim 35\%$) of these 64 planets would be transiting and detectable, constituting a $\sim 7\sigma$

difference with zero detections. This is significantly different than analogous results using simulated planetary systems, which were consistent with zero to within ~ 2 – 3σ . We thus identified a strong discrepancy between expectations based on hypothetical planets and the lack of additional outer planet detections in the observed systems. Crucially, these results are a lower limit on the true extent of the discrepancy, since we would expect more detections if we assigned more than one hypothetical planet per system.

The discrepancy reveals that there is some truncation or transition in the underlying system architectures at roughly ~ 100 – 300 days. There are four distinct possibilities, enumerated below:

1. There is an average truncation in the underlying architectures. If the “peas-in-a-pod” patterns continue to larger separations than observed, consistency with the observations would require that fewer than $\sim 20\%$ of systems host additional planets at the next outermost expected periods.
2. The period ratios increase at larger orbital separations. This would require an average increase in the period ratios (relative to the “peas-in-a-pod” expectations) by a factor of $\gtrsim 3$.
3. The planet radii are smaller at larger orbital separations. This would require an average decrease in the radii (relative to the “peas-in-a-pod” expectations) by a factor of $\lesssim 0.5$.
4. The mutual inclinations increase at larger orbital separations. This would require an average increase in the inclination dispersions by a factor of $\gtrsim 6$.

Whether the outer architectures of the compact multis have some truncation or a breakdown of the “peas-in-a-pod” patterns, these results have important consequences for theories of planet formation and dynamical evolution. Several physical processes could potentially give rise to these features (Section 7.2). However, one particularly complementary prediction from the migration model of close-in sub-Neptune formation is that planets in a certain mass and semimajor axis range should experience migration traps that prevent inward migration (e.g., Izidoro et al. 2017), leading to a gap in the resulting systems at ~ 100 – 300 days and the existence of smaller-mass planets outside of this gap (Zawadzki et al. 2022). We encourage future work on this and other theoretical processes pertinent to the outer regions of compact multis.

With the Kepler data set alone, it is impossible to determine whether our results provide evidence for a truncation or some other transition at $\gtrsim 0.5$ –1 au. However, future observations may characterize the outer architectures enough to distinguish these possibilities. In particular, extreme-precision radial velocity observations (e.g., Pepe et al. 2021) may allow for a detailed understanding of the outer architectures of some compact multis orbiting particularly bright or quiet stars, with systems like TOI-178 (Leleu et al. 2021) being a prime example. Long-term monitoring of transit timing variations can also signal the presence of more distant, nontransiting planets (e.g., Nesvorný et al. 2012). Future high-precision photometric monitoring, especially with the PLATO mission (Rauer et al. 2014), will shed more light on the outer architectures of compact multis by revisiting the Kepler field and potentially targeting some new fields for multiyear baselines. Finally, continued efforts at characterization of gas giants at larger

(~few au) separations (e.g., Fulton et al. 2021; Rosenthal et al. 2021), especially in systems with inner sub-Neptunes (e.g., Chachan et al. 2022), will gradually provide a more holistic understanding of inner and outer system interactions and dynamics.

We thank the anonymous referee for their careful review and constructive comments. We are grateful to Fred Adams, Chris Burke, Eric Ford, Dan Tamayo, Jonathan Tan, and Josh Winn for helpful discussions. S.C.M. was supported by NASA through a NASA Hubble Fellowship grant #HST-HF2-51465 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

ORCID iDs

Sarah C. Millholland  <https://orcid.org/0000-0003-3130-2282>

Matthias Y. He  <https://orcid.org/0000-0002-5223-7945>

Jon K. Zink  <https://orcid.org/0000-0003-1848-2063>

References

- Adams, F. C. 2019, *MNRAS*, **488**, 1446
- Adams, F. C., Batygin, K., Bloch, A. M., & Laughlin, G. 2020, *MNRAS*, **493**, 5520
- Agol, E., Dorn, C., Grimm, S. L., et al. 2021, *PSJ*, **2**, 1
- Berger, T. A., Huber, D., Gaidos, E., van Saders, J. L., & Weiss, L. M. 2020a, *AJ*, **160**, 108
- Berger, T. A., Huber, D., van Saders, J. L., et al. 2020b, *AJ*, **159**, 280
- Bitsch, B., Johansen, A., Lambrechts, M., & Morbidelli, A. 2015, *A&A*, **575**, A28
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Sci*, **327**, 977
- Bryan, M. L., Knutson, H. A., Lee, E. J., et al. 2019, *AJ*, **157**, 52
- Burke, C. J., & Catanzarite, J. 2017, in Kepler Science Document KSCI-19101-002, ed. M. R. Haas & N. M. Batalha (Washington, DC: NASA), 14
- Carrera, D., Ford, E. B., & Izidoro, A. 2019, *MNRAS*, **486**, 3874
- Chachan, Y., Dalba, P. A., Knutson, H. A., et al. 2022, *ApJ*, **926**, 62
- Chatterjee, S., & Tan, J. C. 2014, *ApJ*, **780**, 53
- Chatterjee, S., & Tan, J. C. 2015, *ApJL*, **798**, L32
- Chiang, E., & Laughlin, G. 2013, *MNRAS*, **431**, 3444
- Christiansen, J. L. 2017, in Kepler Science Document KSCI-19110-001, ed. M. R. Haas & N. M. Batalha (Washington, DC: NASA), 18
- Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. 2012, *PASP*, **124**, 1279
- Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2020, *AJ*, **160**, 159
- Coleman, G. A. L., & Nelson, R. P. 2014, *MNRAS*, **445**, 479
- Cossou, C., Raymond, S. N., Hersant, F., & Pierens, A. 2014, *A&A*, **569**, A56
- Coughlin, J. L. 2017, in Kepler Science Document KSCI-19114-002, ed. N. Batalha & M. R. Haas (Washington, DC: NASA), 22
- Dawson, R. I., Chiang, E., & Lee, E. J. 2015, *MNRAS*, **453**, 1471
- Dawson, R. I., & Murray-Clay, R. A. 2013, *ApJL*, **767**, L24
- Denham, P., Naoz, S., Hoang, B.-M., Stephan, A. P., & Farr, W. M. 2019, *MNRAS*, **482**, 4146
- Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, *ApJ*, **790**, 146
- Fang, J., & Margot, J.-L. 2012, *ApJ*, **761**, 92
- Fulton, B. J., Rosenthal, L. J., Hirsch, L. A., et al. 2021, *ApJS*, **255**, 14
- Furlan, E., Ciardi, D. R., Everett, M. E., et al. 2017, *AJ*, **153**, 71
- Gilbert, G. J., & Fabrycky, D. C. 2020, *AJ*, **159**, 281
- Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, *Natur*, **542**, 456
- Hansen, B. M. S., & Murray, N. 2012, *ApJ*, **751**, 158
- Hansen, B. M. S., & Murray, N. 2013, *ApJ*, **775**, 53
- He, M. Y., Ford, E. B., & Ragozzine, D. 2019, *MNRAS*, **490**, 4575
- He, M. Y., Ford, E. B., & Ragozzine, D. 2021, *AJ*, **161**, 16
- He, M. Y., Ford, E. B., Ragozzine, D., & Carrera, D. 2020, *AJ*, **160**, 276
- Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, *ApJS*, **201**, 15
- Hsu, D. C., Ford, E. B., Ragozzine, D., & Ashby, K. 2019, *AJ*, **158**, 109
- Hsu, D. C., Ford, E. B., Ragozzine, D., & Morehead, R. C. 2018, *AJ*, **155**, 205
- Hu, X., Tan, J. C., Zhu, Z., et al. 2018, *ApJ*, **857**, 20
- Izidoro, A., Dasgupta, R., Raymond, S. N., et al. 2021, *NatAs*, **6**, 357
- Izidoro, A., Ogihara, M., Raymond, S. N., et al. 2017, *MNRAS*, **470**, 1750
- Johansen, A., Davies, M. B., Church, R. P., & Holmelin, V. 2012, *ApJ*, **758**, 39
- Lai, D., & Pu, B. 2017, *AJ*, **153**, 42
- Laskar, J., & Petit, A. C. 2017, *A&A*, **605**, A72
- Lee, E. J., & Chiang, E. 2017, *ApJ*, **842**, 40
- Leleu, A., Alibert, Y., Hara, N. C., et al. 2021, *A&A*, **649**, A26
- Lissauer, J. J., Marcy, G. W., Bryson, S. T., et al. 2014, *ApJ*, **784**, 44
- Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011, *ApJS*, **197**, 8
- Masuda, K., Winn, J. N., & Kawahara, H. 2020, *AJ*, **159**, 38
- Millan-Gabet, R., Malbet, F., Akeson, R., et al. 2007, in Protostars and Planets V, ed. B. Reipurth, D. Jewitt, & K. Keil (Tucson, AZ: Univ. of Arizona Press), 539
- Millholland, S., Wang, S., & Laughlin, G. 2017, *ApJL*, **849**, L33
- Millholland, S. C., He, M. Y., Ford, E. B., et al. 2021, *AJ*, **162**, 166
- Millholland, S. C., & Winn, J. N. 2021, *ApJL*, **920**, L34
- Mishra, L., Alibert, Y., Leleu, A., et al. 2021, *A&A*, **656**, A74
- Mulders, G. D., Pascucci, I., & Apai, D. 2015, *ApJ*, **814**, 130
- Mulders, G. D., Pascucci, I., Apai, D., & Ciesla, F. J. 2018, *AJ*, **156**, 24
- NASA Exoplanet Archive 2022, Kepler Objects of Interest DR25, Version: 2022-02-21 15:06, NExSci-Caltech/IPAC
- Nesvorný, D., Kipping, D. M., Buchhave, L. A., et al. 2012, *Sci*, **336**, 1133
- Paardekooper, S. J., Baruteau, C., Crida, A., & Kley, W. 2010, *MNRAS*, **401**, 1950
- Pepe, F., Cristiani, S., Rebolo, R., et al. 2021, *A&A*, **645**, A96
- Petigura, E. A. 2020, *AJ*, **160**, 89
- Pu, B., & Lai, D. 2018, *MNRAS*, **478**, 197
- Rauer, H., Catala, C., Aerts, C., et al. 2014, *ExA*, **38**, 249
- Rosenthal, L. J., Fulton, B. J., Hirsch, L. A., et al. 2021, *ApJS*, **255**, 8
- Rowe, J. F., Bryson, S. T., Marcy, G. W., et al. 2014, *ApJ*, **784**, 45
- Shahaf, S., Mazeh, T., Zucker, S., & Fabrycky, D. 2021, *MNRAS*, **505**, 1293
- Spalding, C., & Millholland, S. C. 2020, *AJ*, **160**, 105
- Steffen, J. H., Batalha, N. M., Borucki, W. J., et al. 2010, *ApJ*, **725**, 1226
- Tamayo, D., Murray, N., Tremaine, S., & Winn, J. 2021, *AJ*, **162**, 220
- Terquem, C., & Papaloizou, J. C. B. 2007, *ApJ*, **654**, 1110
- Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, *ApJS*, **235**, 38
- Van Eylen, V., & Albrecht, S. 2015, *ApJ*, **808**, 126
- Weiss, L. M., Marcy, G. W., Petigura, E. A., et al. 2018, *AJ*, **155**, 48
- Weiss, L. M., Millholland, S. C., Petigura, E. A., et al. 2022, arXiv:2203.10076
- Winn, J. N., & Fabrycky, D. C. 2015, *ARA&A*, **53**, 409
- Xie, J.-W., Dong, S., Zhu, Z., et al. 2016, *PNAS*, **113**, 11431
- Yang, J.-Y., Xie, J.-W., & Zhou, J.-L. 2020, *AJ*, **159**, 164
- Zawadzki, B., Carrera, D., & Ford, E. B. 2022, arXiv:2202.05342
- Zhu, W., & Dong, S. 2021, *ARA&A*, **59**, 291
- Zhu, W., Petrovich, C., Wu, Y., Dong, S., & Xie, J. 2018, *ApJ*, **860**, 101
- Zhu, W., & Wu, Y. 2018, *AJ*, **156**, 92
- Zink, J. K., Christiansen, J. L., & Hansen, B. M. S. 2019, *MNRAS*, **483**, 4479