A New Metric Revealing an Overabundance of Multiple Planet Systems Near Mean-Motion Resonance

by

Kyle McGregor
Class of 2024

A thesis submitted to the
faculty of Wesleyan University
in partial fulfillment of the requirements for the
Degree of Bachelor of Arts
with Departmental Honors in Astronomy

Abstract

With resonant configurations commonly observed on the scale of the solar system, the mutual gravitational interactions which lead to resonance are assumed to play a role in sculpting the architectures of planetary systems beyond our own. We present a metric, the commensurability offset index (COI), which measures the offset from a system containing a chain of orbital periods in pure integer commensurability ratios and allows for comparisons across the entire sample of transiting systems with $N > 2$ planets. This summary statistic is a measure of how close to commensurability a system needs to be for resonance to have impacted the orbital evolution of a system. We explore a range of simulated planetary populations and orbital architectures the evaluate the robustness and magnitude of this overabundance of systems near resonance. We compare the COI distributions of the observed transiting sample to various model populations designed to mimic the transiting multiplanet sample, dominated by systems discovered by the *Kepler* space telescope. In this analysis, we find an overabundance of observed systems near resonance peaking at COI $= 0.0211$ with a standard deviation of $\sigma = 0.0141$. We discuss how this summary statistic may be used in predictive modeling for searching for additional planets in $N = 2$ planet systems and compare the application of COI to other relevant populations, including angular substructures in planet-forming disks and various groupings of solar system objects.

"By a small sample, we may judge of the whole piece"

–Miguel de Cervantes, *Don Quixote*

Acknowledgments

First and foremost I want to thank my advisor, Seth Redfield, for his support and belief in me these past four years that led to this thesis. I am so grateful to have you as my advisor and role model, and will carry with me the kindness and diligence you have given into the years to come. Your passion for this science and for expanding our horizons in it will always be an inspiration to me. Thank you for all your work.

I also want to thank Roy Kilgard for being a great mentor as well, especially through our work with the 24-inch telescope work these past few years. Thank you for always having your door open and making me feel at home in this department. Thank you also to Meredith Hughes for introducing me to this field during that first pandemic fall, and for always having your door open as well. Thank you to Jonathan Jackson for being a great mentor and for having great discussions during group meetings. Thank you to Ed Moran for teaching me about cosmology and to Sarah Wellons for teaching me about stars and galaxies. Thank you also to Stef Dinneen for all the behind-the-scenes work you do for this department. I have learned so much from you all here in our home atop Foss. Thank you also to Duncan Lorimer for the wonderful work we embarked upon this summer. Words cannot express my gratitude at having worked with you.

I am incredibly grateful for all my peers in this department these past four

years. Thank you to all the graduate students for all your hard work and help though classes and research, particularly to Katie Bennett, Alaina Einsig, Rewa Bush, Lisseth Gonzales Quevedo, Katie Ciurleo, and Cat Sarosi. Thank you to my fellow undergraduates who I have worked with in research and classes, especially Cassidy Soloff, Chris (Qiushi) Tian, Uday Narayanan, Jamar Kittling, and Owen Gonzales. Thank you also to Anna Fehr, Seth Larner, Eric Rumsfeld, and Hannah Lewis for an unforgettable first summer of research and for many fun memories.

Thank you to all my friends and mentors outside of astronomy for all your support these past four years. Thank you to Wesleyan Swimming and Diving for being my rock and my happy place, with special thanks to Coach Peter and Coach Vanessa for teaching me to be a leader both in and out of the pool. Thank you to Taua Fitisemanu, Maya Milrod, Bella Kellner, and countless others for being wonderful friends and leaders. I can't believe it's the end of our time here at Wesleyan, but I'm really excited to see what's to come for us. Thank you for being there for me.

Last and most importantly, I want to thank my family. Thank you to my mom and dad for all their work in helping me pursue this dream. Thank you to my twin brother, Matt, for always lending a listening ear and being there for me. I love you all so much and I am so proud of how far we've all come these past four years. Thank you to Uncle VJ and Erin for opening your home up to me, and being there when I needed it. Thank you to Aunt Nancy, Camille, and Marissa for the same.

This thesis is dedicated to Marie Geronimo, Charlotte and Francesca Rich, and my goddaughter Gianna Santinelli. I can't wait to see you all grow to be curious and kindhearted adults.

Contents

Chapter 1

Introduction

Exoplanets are planetary bodies orbiting a star other than the Sun. For much of the past century, the existence of exoplanets was largely assumed but unsupported by any observational evidence (e.g. Struve 1952). From the first confirmed detections in the 1990s, the subsequent decade saw the refinement of an array of techniques now widely used to uncover a booming and diverse population of exoplanets and their host stars (e.g. Wolszczan & Frail 1992; Mayor & Queloz 1995). These include both photometric and spectroscopic methods—methods of measuring the light of a host star to infer the presence of exoplanets—which can be readily done with even modest ground-based telescopes and have also motivated several space-based planet-hunting surveys in subsequent decades (e.g. *Kepler*, *TESS*). From this ongoing observational effort now more than 5500 planets have been observed in more than 4100 systems, with hundreds of new exoplanets and systems confirmed each year.

Here we introduce the most fruitful of the detection methods, the transit method, which has evidenced more than 70% of all confirmed exoplanets, and the ways that this technique has been applied through large surveys to produce the sample we see today. We also introduce the biases inherent to the detection of transiting exoplanets, as well as some notable trends in the dynamics of exoplanets that particularly characterize the subset of systems that contain multiple

planets. This will help inform the nature of this large and growing transiting population, which in future sections we model with statistical simulations and dynamical analysis. To get there, though, we must first understand this population at the base observational level.

1.1   Transiting Exoplanets
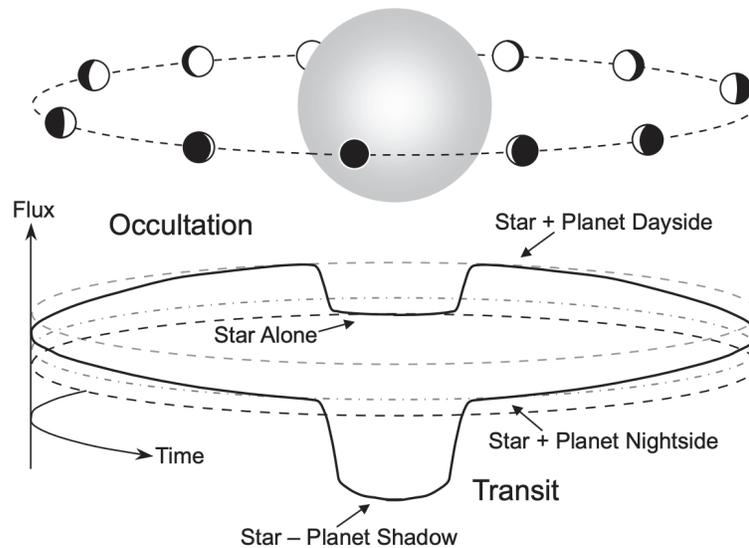
Figure 1.1: Exoplanet transit and occultation diagram from Winn (2010). Transit detections search for periodic dips in the total flux from the star caused by a planet passing in front of an observer's view of the star. These events are accompanied by occultations, where the total flux dips even more slightly due to the star blocking the reflected light by the planet as it passes behind the star.

A transit occurs when an exoplanet passes through the line of sight from Earth to its host star (see Fig. 1.1) and is characterized by a slight decrease in the light received from the star as the planet passes over it. These occur periodically with each orbit, and are accompanied by an even slighter occultation event when the planet passes behind the star (also seen in Fig. 1.1), though transits are more easily observed due to the much greater change in the stellar flux that they cause.

Due to their periodic nature of occurring with every orbit, and as the change in flux is caused by the planet blocking out a part of the star's light, the orbital period and radius of the planet are direct observables by the transit method. The orbital period P is the difference between the midpoints t_{mid} of two succesive transits, given by

$$P = t_{\mathrm{mid},2} - t_{\mathrm{mid},1}. \tag{1.1}$$

While two transits are required to constrain orbital period, at least three are required for a confirmed planet detection. Planets with an insufficient number of transits detected by a survey are referred to as *candidate* planets and are often cataloged for follow-up monitoring by other facilities (e.g. Stassun et al. 2018).

The planet radius R_p is given by

$$R_p = \sqrt{\delta} R_s, \tag{1.2}$$

where δ is the observed transit depth and R_s is the radius of the star. Note that transit depths are often measured in parts per thousand (ppt) or parts per million (ppm) normalized to the base out-of-transit flux of the star. Ground-based telescopes are generally limited by atmospheric interference and produce most significant detections for transit depths on the order of greater than 10 ppt, while space-based observatories can detect planets at the same significance with depths at ppm scale (e.g. Mallonn et al. 2022).

The primary observational limitation of the transit method is the precise viewing geometry required for a transit event to occur. The inclination angle (i) of a detected planet's orbit must be such that the planet passes over the face of the star, leaving only a small window of observability for all possible planet configu-
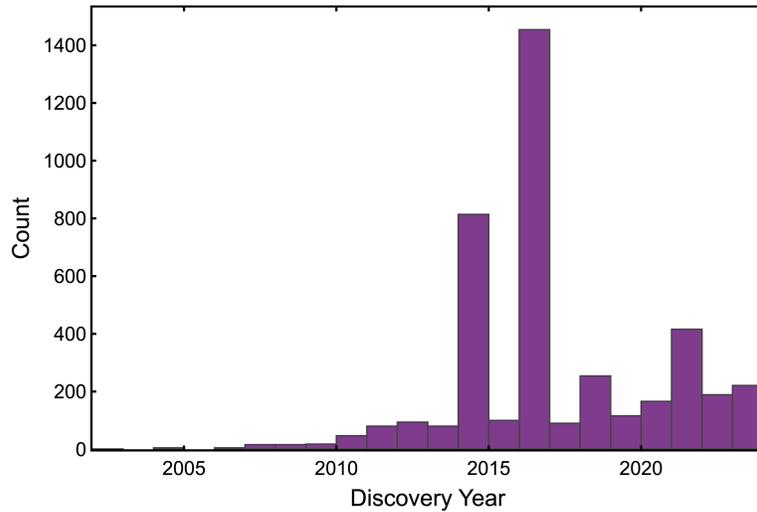
Figure 1.2: Detection count of new transiting planets by year. Source: NASA Exoplanet Archive

rations. This window is generally within $\pm 1°$ of coplanarity with the line of sight to the observer, leaving only about 1% of all planetary systems observable by the transit method.

Despite this limitation, transit surveys have been very effective at detecting a large population of planets, which we introduce further in Sec. 1.2. In Fig. 1.2 and Fig. 1.3 we provide the timeline for the increase in detected transiting exoplanets over the past two decades. Here note two large increases in 2014 and 2017. These correspond to large data releases by the *Kepler* space telescope, both part of its initial observations and its successor mission *K2* (see Howell et al. 2014; Fulton et al. 2017; Petigura et al. 2017; Johnson et al. 2017; Thompson et al. 2018). *Kepler* is the most prolific planet-hunting survey conducted to date. Over 10 years of observations, in a single well-populated observing field for the initial mission (2009–2013) and along the ecliptic for *K2* (2014–2018), the facility discovered more than 2700 exoplanets. These fields are particularly noted in Fig. 1.4, where we give the sky position (R.A./Dec.) of all confirmed transiting systems.
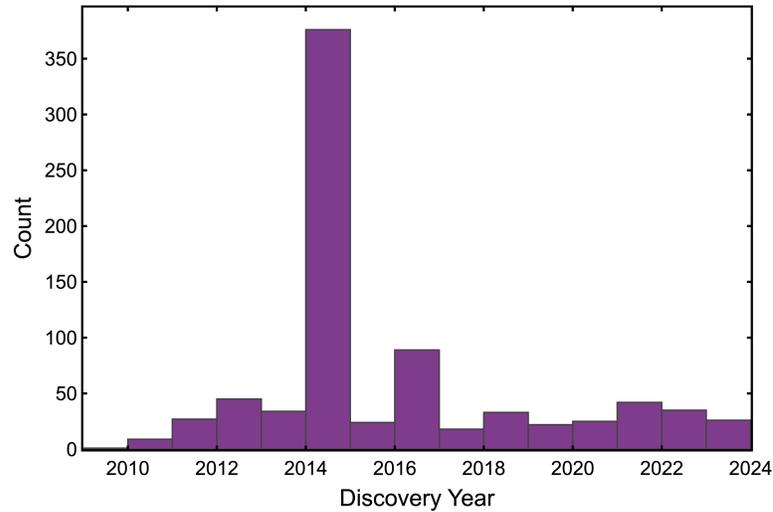
Figure 1.3: Same as Fig. 1.2, but for transiting planets in systems containing $N > 2$ planets. Source: NASA Exoplanet Archive

The *Kepler* space telescope took its final observations in 2018, but subsequent planet-hunting survey work is led by the *Transiting Exoplanet Survey Satellite (TESS)*. Rather than probing a single patch of sky with multiple years of observations, *TESS* observes 85% of the sky for a minimum of 27.4 days. This gives a much wider survey range but a lower likelihood of confirming planets with low transit probability (Ricker et al. 2015). About 400 exoplanets have been confirmed from *TESS* detections to date, with several thousand planet candidates awaiting follow-up observation for confirmation.

These two surveys have provided the most robust and standardized population of exoplanets to date. In the next section we discuss the nature of this population, introducing the observed sample to date and the detection biases of these transit surveys that shape what we see. Particularly, we introduce and focus on the population of systems containing multiple planets (see Fig. 1.3), which display resonant dynamical interactions that, in further sections, we seek to contextualize within the entire transiting population at large.
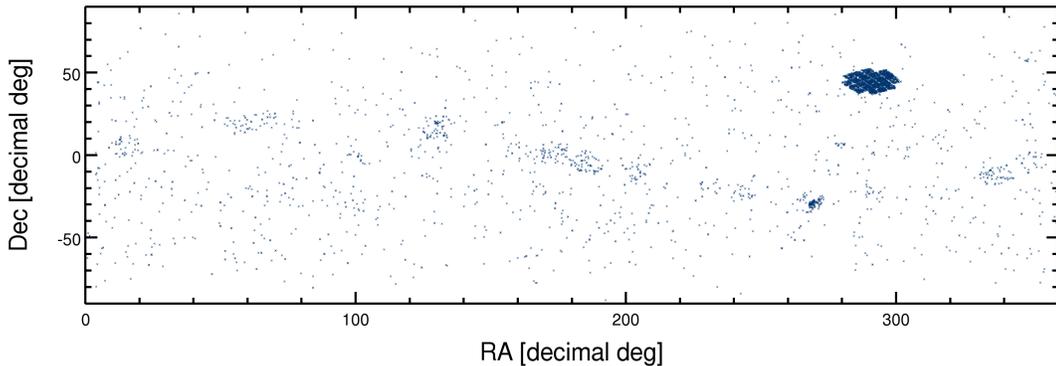
Figure 1.4: Sky position of all confirmed exoplanet detections across all detection techniques. Note the concentration of sources in the upper right and along the ecliptic, these are sources in the *Kepler* observing field and on the area of highest cadence for *K2*, the most productive planet-hunting surveys to date. Source: NASA Exoplanet Archive.

1.2 Exoplanets as a Population

As introduced in Sec. 1.1, the past two decades have seen a massive increase in the number of known exoplanets, particularly thanks to the planet-hunting space telescopes *Kepler* (launched in 2009, retired 2018) and *TESS* (launched in 2018, ongoing). More information on these missions and their specifications is given by Borucki (2016) and Ricker et al. (2015), respectfully.

Across the board, transit surveys are biased toward detecting short-period (close-in) planets and against detecting long-period (far-out) planets. This is because the probability of observing a transit of a planet scales against orbital period and planet size. Assuming a uniform distribution of inclination angles and that the orbital eccentricity $e = 0$ for all planets across a survey, the probability that a planet will transit its host star (excluding grazing passes) is given by

$$\mathrm{P_{tra}} = \frac{R_s - R_p}{a},\tag{1.3}$$

where R_s is the radius of the star, R_p is the radius of the planet, and a is
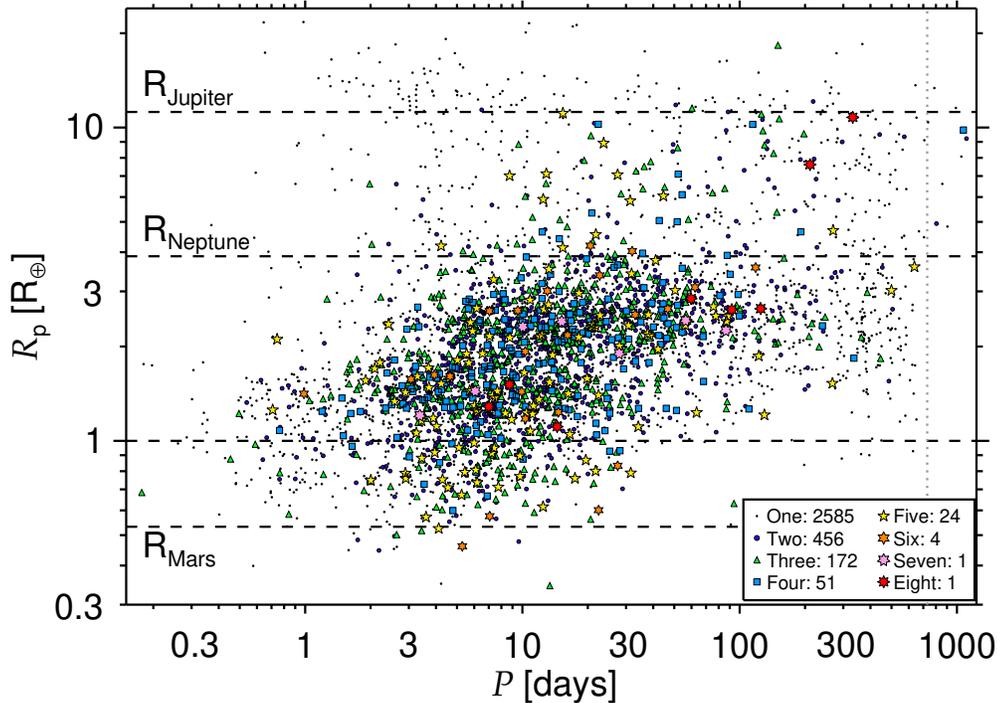
Figure 1.5: Fig. 1 from Lissauer et al. (2023). Period-radius relationship for the *Kepler* candidate systems studied therein, labeled by the number of transiting planet candidates observed in each system.

the semimajor axis (note that $a \propto P^{2/3}$ by Kepler's third law, where P is the orbital period). As there is a large variation in possible orbital periods, from less than a day to upwards of hundreds of days, compared to the spread in planet radius (visible in Fig. 1.5), the decrease in probability of detection is particularly dominated by orbital distance (Lissauer et al. 2023).

Combining these selection factors, consider the prospect of observing the Earth as an exoplanet (Robinson & Reinhard 2018). No transit survey designed to date would be able to detect an earth-sized planet in an Earth-sized orbit around a sun-like star using the transit method, due to the weak transit depth (rearranging Eqn. 1.2 gives $\delta_\oplus = 0.08$ ppm) and low transit probability ($P_{\text{tra}} = 0.005$ due to its long-period $\approx$365 day orbit). Note that $P_{\text{tra}} = 0.005$ implies that in an observed sample population not limited by observing cadence (i.e. constantly observed
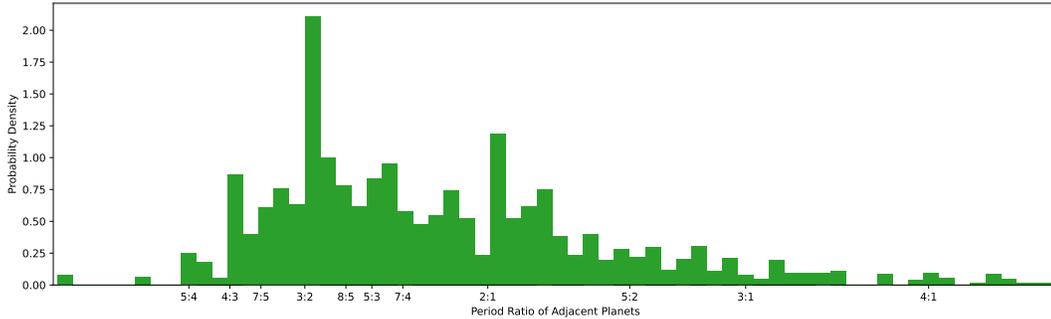
Figure 1.6: Period ratios of adjacent planet pairs for the sample of planetary systems with three or more transiting planets.

for a minimum of 365 days), only 1 in 200 stars containing earth-sized planets would have a transit detection. This also assumes that the survey could resolve a sufficiently small $\delta_\oplus$ and have a short enough exposure to capture the duration of the transit (for earth this would be ≈ 12 hr) with a sufficiently high signal-to-noise ratio. This is an incredibly unlikely proposition.

We show the *Kepler* transiting sample in Fig. 1.5, taken from Lissauer et al. (2023). Note the slight upward relationship between period and radius, which can be interpreted as a weak period-radius relation. This makes sense intuitively as, for one, planets within the solar system tend to have bigger radii further out from the Sun. This may also be an observational artifact, as small planets at high orbital separations have the lowest transit detection probability according to Eqn. 1.3. The cause is still debated, and we discuss the underlying observed phenomenon that may lead to this observed trend (the "peas in a pod" model) more in Chapter 5.

There is a tendency for exoplanets to converge around "resonant" configurations, where the orbits of successive planets occur in integer ratios. We show this in Fig. 1.6, where the orbital period ratios of adjacent planets in systems containing

$N > 2$ planets cluster around resonant configurations. This was notably identified in data from *Kepler/K2* (Fabrycky et al. 2014) and reflects a phenomenon that is observed widely within the solar system (Murray & Dermott 2000). There also appears to be spread off of the exact resonance values, particularly the strong 3:2 and 2:1 resonances, so that the abundance of systems caused by the resonance is slightly offset, leaving ambiguity as to the exact definition of whether a system is "in proximity" to the resonance or not. We discuss the significance of this, and how it frames the inquiry of this thesis, in Chapter 2.

1.3   A Note on Astrophysical Population Science

Each and every object studied by astronomers is, without exception, a constituent of some much larger and ultimately unobservable population. Astrophysics is a science of samples and surveys, limited by many considerations intrinsic both to our instruments and the objects we design them to observe. These limitations, which we call "incompleteness factors" and "observational biases", ultimately prevent us from gathering a complete catalog of every possible source, and cause certain sources and events to be observed more effectively and others to be suppressed in survey observations. Such is seen all across astrophysics, from surveys of stars (Gaia Collaboration et al. 2018), galaxies (Almeida et al. 2023), and compact objects (Stovall et al. 2014), but are especially pertinent in transient astrophysics (e.g. Pereira et al. 2023; McGregor & Lorimer 2024) and exoplanet science.
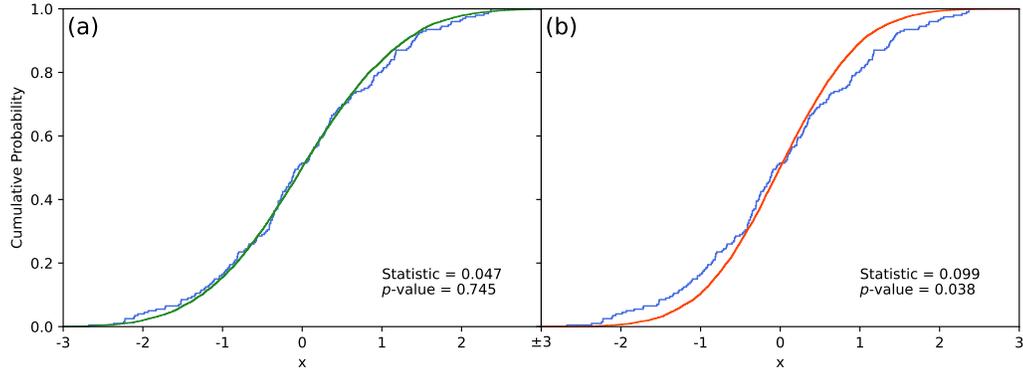
Figure 1.7: KS test demonstration plots described in Sec. 1.4.1. We include the KS statistic and p-value for a two-sided KS test between the model (red) and empirical (blue) cumulative distribution functions.

1.4 Statistical Methods

Fortunately, such a population-limited science can make good use out of statistical techniques to generate models and compare them to observations, which can be used to inform and seek out biases in observed sample populations. In this section we introduce two widely-used statistical methods, the Kolmogorov-Smirnov test and Markov Chain Monte Carlo, which will be applied in Chapters 3 and 4 to simulated exoplanet populations.

1.4.1 Kolmogorov-Smirnov (KS) Test

In general, the Kolmogorov-Smirnov test compares the cumulative distribution function (CDF) of two probability distributions to determine some shared property about the two distributions. There are three variants of the test (Babu & Feigelson 1996), with an "upper" variant that gives the likelihood a given distribution is systematically greater than another, and a "lower" variant that gives the opposite. However, the most used version for this work, and most astrophysical populations,

is the "two-sided" variant. This test gives the likelihood that the two distributions share some underlying source distribution or may have been drawn randomly from each other, which is useful for population science as it serves as a "goodness-of-fit" type measurement for comparing empirical data against a well-populated model.

An example of the KS test procedure is shown in Fig. 1.7, comparing a well-populated ($n_{\mathrm{model}} = 10^5$) Gaussian CDF against a smaller "empirical" ($n_{\mathrm{emp}} = 200$) CDF drawn from the same underlying distribution ($\sigma = 1$). The test essentially determines the probability that the largest difference between the two CDFs (the KS statistic) occurred randomly, given by the p-value. A p-value above 0.05 is considered statistically significant for the empirical distribution being drawn from the model. As the distributions share the same underlying probability density function the p-value from the test is very significant and the KS statistic is low. If we vary the parameters of the model CDF, say ($\sigma = 0.8, x_0 = 0$), this results in a higher KS statistic and a much lower p-value, as shown on the figure.

With a sufficiently physically-minded model distribution to compare to observed data, the KS test can serve as evidence of parity between a model and observations. Such procedure is, as mentioned, ubiquitous across astrophysical population science (see, e.g., James 2023; McGregor & Lorimer 2024).

1.4.2   Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is a suite of algorithms for multidimensional parameter estimation. Essentially, it allows for data to be fit to a model described by an arbitrary number of parameters through iteratively sampling distributions that converge on the parameters' most likely values. More specifically used in astrophysics, we use the Metropolis-Hastings algorithm, which is particu-

larly useful for sampling models with a large number of possibly degenerate (that is, dependent) parameters.

To give a high-level overview, this algorithm uses an iterative process for finding the distribution of possible values of a given parameter, known as a Bayesian framework. Using Bayes' theorem, the value of a parameter is slightly varied over the course of the algorithm for a certain number of steps. Using this Bayesian framework for many parameter variations along a random walk, MCMC allows us to construct a posterior distribution for each model parameter θ given an empirically-determined dataset d. Each iteration of the algorithm applies the theorem to evolve the posterior distribution given the fit of the parameters to the dataset. Bayes' theorem is given by

$$P(\theta|d) = \frac{P(\theta)P(d|\theta)}{P(d)}, \tag{1.4}$$

where $P(\theta|d)$ is the posterior distribution for each parameter, $P(\theta)$ is the prior distribution for each parameter, $P(d|\theta)$ is the likelihood function for the data being generated by a certain parameter value, and $P(d)$ is an arbitrary normalization factor unimportant for running MCMC. Each MCMC run requires some specified number of steps and walkers, where the number of steps dictates the length of the Markov process and the number of walkers gives the number of samples taken from the posterior distribution. This essentially sets the precision of the parameter estimation, which scales with computational expense.

For further reading on the applicability of MCMC to exoplanet science and astrophysical populations in general, see Ford (2006); Andreon & Weaver (2015); Babu & Feigelson (1996). To implement the algorithm practically, the most commonly-used python library is `emcee` (Foreman-Mackey et al. 2013).

Chapter 2

Mean Motion Resonance

Within and beyond our solar system, the gravity of orbiting satellites will influence the orbits of each companions within a system. Often this influence is small and assumed to be negligible, but certain systems show evidence of gravitational interactions influencing system architectures through tugging on bodies well beyond that predicted by Kepler's laws. The strongest observational evidence for such interactions is the phenomenon known as mean motion resonance, in which orbiting bodies move in whole number ratios of their orbits relative to other planets within a system. In other words, in the time it takes an outer body to complete a single orbit an inner planet will complete some integer $N > 1$ number of orbits. Chains of planets in these resonant configurations may also form and can dominate the orbital architecture of a system.

Observations of mean motion resonance have been made since the very beginning of modern Astronomy. Within our own solar system, the Galilean moons of Jupiter–the system of Io, Europa, Ganymede, and Callisto–show orbital motions near a tight 1:2:4 chain of orbital periods between the system's inner three bodies. This configuration is readily observable through timing with even a small backyard telescope, which the eponymous Galileo Galilei discovered in 1610 (*Sidereus Nuncius*, Galilei 1610). Galileo's discovery, of objects clearly revolving around a central planetary body, served among the first strong observational evidence

against the geocentric model, full stop. Now we know of entire solar systems orbiting stars other than our own, and we have discovered many other similarly interesting orbital configurations beyond our own host system.

2.1   A Tendency toward Commensurability
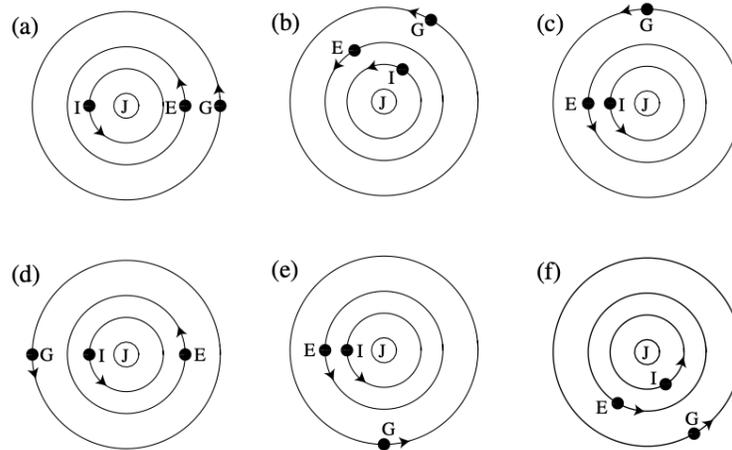
Figure 2.1: Fig. 8.30 from Murray & Dermott (2000) showing the orbital conjunctions of the Galilean moons of Jupiter, a tight first-order commensurability chain indicating a Laplace-type resonance.

As the observable products of planet formation and dynamical evolution, each confirmed exoplanet system provides a window into the results of these processes. For systems containing multiple bodies, the motion of each body in the system will deviate from pure Keplerian motion with the gravitational influence of other perturbing bodies. This ultimately sculpts the orbital architecture of such systems in their evolution, which have notably been observed in planet-hunting surveys to date (Fabrycky et al. 2014).

As discussed in the previous section, survey incompleteness factors result in a known sample of exoplanets observationally biased towards short-period planets, particularly for multiplanet systems. Due to the close proximity of bodies in

these systems, planet-planet gravitational interactions will strongly influence the formation and migration of these systems–these are elements of a system's *architecture*, which we can directly probe with precise timing of the orbital periods of the planets. Such interactions which may be observable in young (< 5 Gyr old) planetary systems (Dai et al. 2023).

The phenomenon of resonance occurring near integer ratios brings up an interesting ramification. Since all observable period ratios can be approximated by some ratio of natural numbers, are all pairs of planets in resonance? While the analytical derivations of orbital perturbations in such systems is outside the scope of this work (see Tremaine 2023), it is most notable when a single pair or multiplanet chain of adjacent planets are found in proximity to some commensurability ratio $(p + q)/q$ with $p \leq 3$. Specifically, pairs near a commensurability with $p = 1$ as first-order, $p = 2$ as second-order, and $p = 3$ as third-order resonant configurations. These low-order configurations are generally considered the most stable in planet-planet interactions.

Some tight resonant systems will display the phenomenon of Laplace resonance, in which a linear combination of the average orbital longitudes (λ_i) of each body equals a constant resonant angle Φ_L (i.e. it does not precess across successive orbits). For the three-body case, this is given by the Laplace relation

$$\Phi_L = a_1\lambda_1 - a_2\lambda_2 + a_3\lambda_3. \tag{2.1}$$

The Galilean moons system notably is situated within a Laplace resonance with $\Phi_L = \lambda_{\text{Io}} - 3\lambda_{\text{Europa}} + 2\lambda_{\text{Callisto}} = 180°$. As this is the time-averaged relation the Laplace relation is *exact*, but in reality Φ_L librates with an amplitude of $0.064°$ and a period of 2071 d (values from Murray & Dermott (2000), citing Liseke

1998). This libration occurs only in systems containing a Laplace-like resonant chain, and has the physical meaning that no three-body orbital conjunctures will take place and that two-body conjunctures will occur in integer fractions of Φ_L. Orbits displaying this phenomenon are illustrated in Fig. 2.1.

The seminal orbital mechanics textbook by Murray & Dermott (2000) identifies this "tendency toward commensurability" in dynamical systems within our solar system and beyond. Such examples include the aforementioned Galilean moons of Jupiter (shown in Fig. 2.1), radial structure in the asteroid belt and the rings of Saturn, as well as the co-orbital "trojan" satellites of Jupiter and other outer solar-system planets, among many others. The advent of large-scale exoplanet surveys has expanded the study of these dynamical phenomena to systems beyond our solar system, and a wide diversity of planetary system architectures has been revealed as a result of this observational effort.

2.2 The Commensurability Offset Index

Systems converging upon resonant configurations will never actually contain a so-called "pure" resonant chain. That is, the period ratio between two bodies will never be directly situated on the ratio of two integers–though many systems indeed come very close. For example, we return to the prototypical Laplace-type resonant system of the Galilean moons of Io, Europa, and Callisto. Each have orbital periods (in days) $P_{\mathrm{Io}} = 1.769138, P_{\mathrm{Ganymede}} = 3.551810, P_{\mathrm{Callisto}} = 7.154553$ (Murray & Dermott 2000). These periods are similar in order of magnitude to the close-in transiting exoplanet sample. By convention we use the pairwise ratio between each successive pair of bodies, giving $P_{\mathrm{Ganymede}}/P_{\mathrm{Io}} = 2.007650$ and $P_{\mathrm{Callisto}}/P_{\mathrm{Ganymede}} = 2.014340$. These ratios are very close to the pairwise 2/1

ratios between the three bodies in a 4:1 resonance, with offsets representing a deviation on the order of minutes. Such offset measurements are aided by the precise orbital periods measured for the Jovian moons through direct observation, which is similarly granted with the high timing resolution of transit photometry.

Here we introduce a test statistic to examine the proximity of multiple adjacent period ratios to a chain of resonances. This will allow us to probe the entire sample of transiting multiplanet systems to compare systems' absolute proximity to containing a pure resonant chain. Similar to the above calculation and the Δ value calculated for a given pair of exoplanets in Dai et al. (2023), we desire an empirical measurement for the distance of a whole system of exoplanets from containing a resonant chain. This system-wide summary statistic, which we have called the Commensurability Offset Index (COI), is defined as follows:

$$\text{COI}_n = \sum_{i=1}^{n} \left| \frac{P_i/P_{i-1}}{p/q} - 1 \right|. \tag{2.2}$$

Where P_i/P_{i-1} is the period ratio of adjacent planets (also written as $\mathcal{P}_i$ elsewhere in this thesis) and p/q is its closest low-order resonance. As the majority of multi-body resonant chains are among three planets and the $N = 3$ planet population is the most populated within this sample, we usually take the sum to $n = 2$ and will do so unless otherwise noted for the rest of the work. We give the calculated COI values for the 12 lowest-offset systems in Table 1. We have also plotted the period ratios for each system of COI < 1.0 in Fig. 2.3 and plot the two period ratios put into the calculation in Fig. 2.2 to show their proximity to pure resonant chains.

Fig. 2.3 is inspired by Fig. 1 and Fig. 2a in Fabrycky et al. (2014) and informs the underlying dynamical trends of the multiplanet transiting population. No-
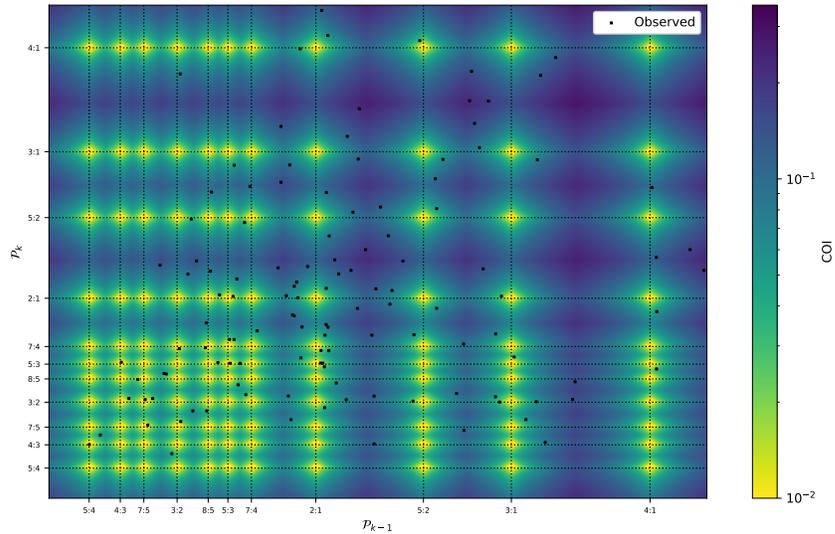
Figure 2.2: Period ratios used in calculation of COI for each system. We include a heatmap of the proximity to a "pure" resonant chain, shown as the intersection points between the vertical and horizontal lines representing the commensurability ratios.

tably, the right histogram indicates that the population tends to decrease around

$COI = 10^{-1}$ and 10^{-2}, showing several systems with significantly lower COI values.

These low-COI systems correspond to systems with tight and well-constrained

resonant chains and often contain many detected planets, such as TOI-1136 (Dai

et al. 2023), Kepler-223 (Mills et al. 2016), and TRAPPIST-1 (Gillon et al. 2017).

The period ratios within these systems tend to fall on first-order ratios, partic-

ularly the 3:2 and 4:3 ratios, as most low-COI systems tend to contain strong

resonant chains dominated by these first-order commensurabilities.

This observation is supported by the period ratio histogram, which shows a

grouping of period ratios around certain commensurabilities, particularly the first-

order 3:2 and 2:1 resonances (e.g. Ghosh & Chatterjee 2023). This phenomenon

was originally identified for the *Kepler/K2* sample and continues to hold for the

entire transiting sample to date.

Table 1: Calculated COI system values for lowest offset systems as of January 1st, 2024.

System	N	COI
TOI-1136	6	0.00041
Kepler-223	4	0.00041
Kepler-60	3	0.00076
TRAPPIST-1	7	0.00504
K2-384	5	0.00672
Kepler-138	3	0.0117
Kepler-398	3	0.00723
Kepler-444	5	0.00898
Kepler-85	4	0.0142
Kepler-1254	3	0.0149
Kepler-54	3	0.0176
V1298 Tau	4	0.0179

2.3   Case study systems

Within the entire known population of exoplanetary systems containing multiple transiting planets, we identify multiple "special case" systems that are unique within the population due to their orbital architectures. As is shown in Fig. 2.3, there are numerous systems with COI values on the order of 10^{-3} to 10^{-4}. These are incredibly tight resonant chains and are often the targets of a follow-up study to probe planet-planet interactions (e.g. Shallue & Vanderburg 2018), and as such these systems tend to contain many known planets. Outside of this regime, resonant interactions are still possible (and noted) within the population, though chains of near-resonant planets are less common. Here we explore the diversity of this population, focusing on two systems containing distinct resonant chains.
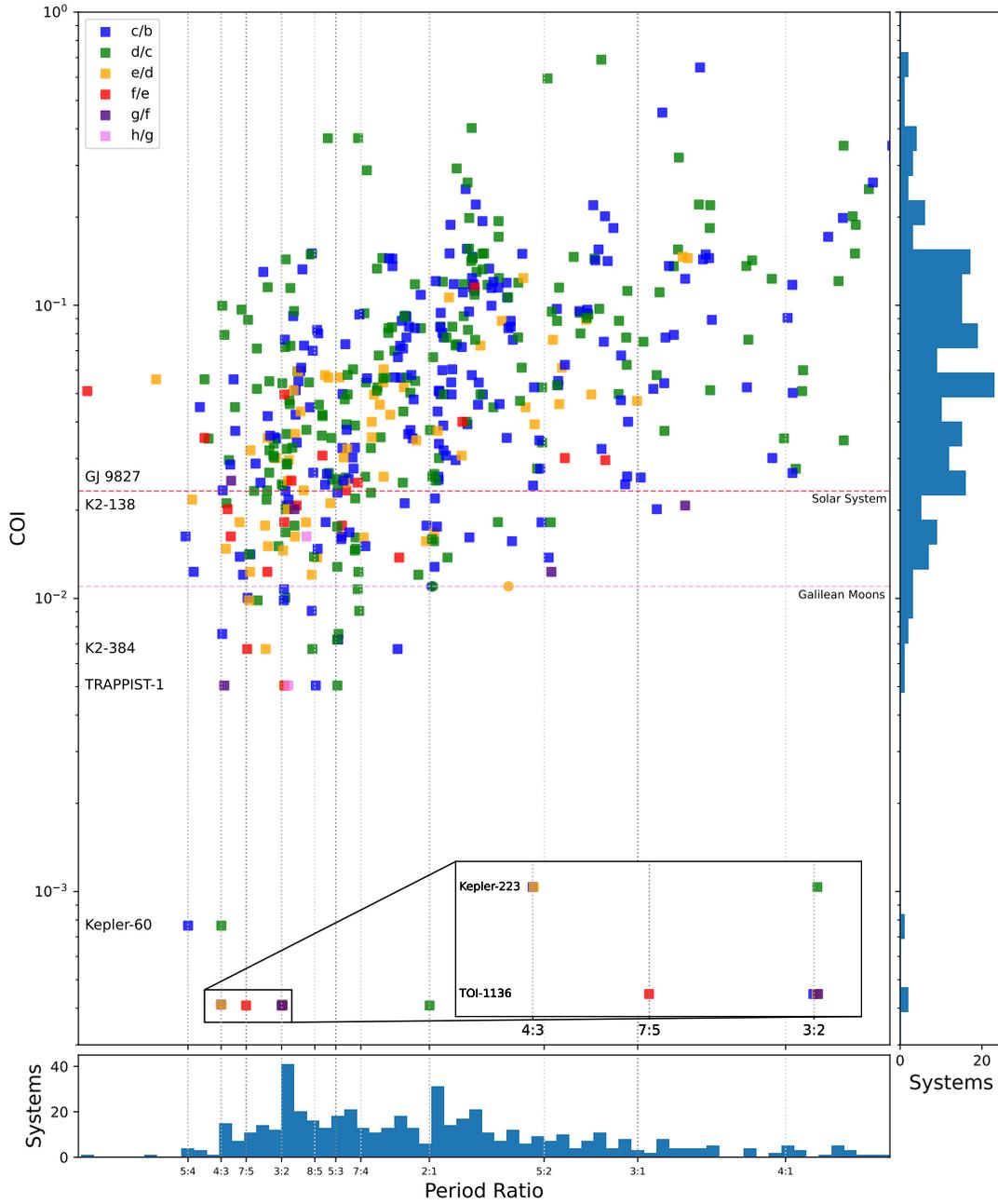
Figure 2.3: COI values for all multiple-planet systems plotted against period ratio for each successive planet pair therein. Low-order commensurabilities considered in this work are shown with dashed vertical lines distinguishing their order. We include the calculated COI values as horizontal lines for the Galilean moons and the solar system. Symbols are colored denoting the planet pair.
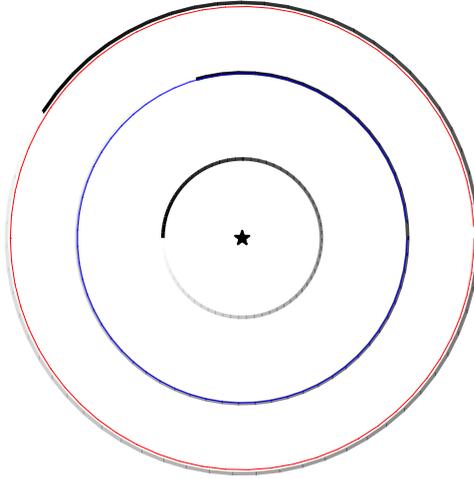
Figure 2.4: Orbit plot for GJ 9827, generated by the N-body code `REBOUND` (Rein & Liu 2012) We include the orbital trajectories of a pure resonant chain (COI = 0), assuming the same inner planet orbital period for both scenarios.

2.3.1 GJ 9827

GJ 9827 is a system of three known super-earths orbiting a K-type dwarf star in the constellation of Pisces. First observed during the *K2* mission and discovered by a team at Wesleyan (Niraula et al. 2017), this system is notable as it is in proximity to the second-order 1:3:5 resonant chain. As noted in Section 2.2, there is a notable dearth of systems converged on second-order resonances. GJ 9827 is the system of lowest COI in proximity to a chain of only second-order resonances, which is notable as second-order resonances tend to be observed less frequently within the solar system–with the location of the second-order 3:1 resonance with Jupiter forming the outer limit for inner main-belt asteroids, known as one of the Kirkwood gaps (Xu & Lai 2017). While the Kirkwood gaps are considered unstable, as the repeated gravitational perturbation by Jupiter excites bodies

away from the resonance, the mutual interactions in the second-order chain in GJ 9827 are generally considered long-term stable (Prieto-Arranz et al. 2018, see Sec. 6.3).

2.3.2   Kepler-223
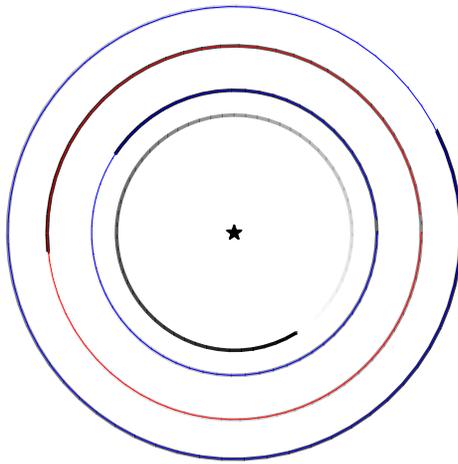
Figure 2.5: Same as Fig. 2.4, but for Kepler-223

As of January 1st, 2024, Kepler-223 is the system with the second-lowest COI value. In April 2023 the system was surpassed by the discovery of the 6-planet TOI-1136 (MacDougall et al. 2023). In this work, we have chosen to highlight Kepler-223 due to an abundance of literature probing its dynamics. This system, identified with the *Kepler/K2* mission, is unique in that it was the first identified 4-planet first-order resonant chain. Originally thought to contain a pair of co-orbital planets, Kepler-223 contains a tight 8:6:4:3 resonant chain between all planets within the system.

Chapter 3

Simulating Exoplanet Populations

With the definition of the commensurability offset index (COI) as presented in Eqn. 2.2, we seek to apply this dynamical summary statistic to a simulated population of multiplanet systems. Here we introduce two models for simulating the multiplanet sample—a simple best-fit model for independent draws of orbital period in $N = 3$ systems ("Model A") and an adapted code from the Exoplanet Population Observation Simulator (`EPOS`, Mulders et al. 2018) which accounts for observed trends in the orbital spacings between adjacent planets.

These methods differ in how they simulate orbital periods but are similar in how they generate planet radius and mass. For both methods we enforce physicality through N-body stability analysis, described in Sec. 3.4, in which we eliminate unstable systems from the samples. With these models, we run a Monte Carlo simulation to construct a mock-observed catalog for each case. We compare these to the observed sample in Chapter 4.

3.1 Model A

As demonstrated in Fig. 1.5, surveys of the observed multiplanet population have provided a large sample of $N = 3$ planet systems, particularly dominated by high-cadence observations by the *Kepler* space telescope (Borucki et al. 2010; Fabrycky et al. 2014). Examining the orbital periods of the $N > 2$ planet systems,

we fit a lognormal parameterization to the period distributions for the innermost three planets of each system, shown in Fig. 3.1. The mean of each distribution increases monotonically as one would expect for three populations of planets with systematically increasing orbital periods, with values of 6.24 days for the innermost planet (planet b), 11.09 days for the second (planet c), and 17.42 days for the third (planet d). This shape and trend arises from the observational bias against planets with longer periods inherent to transit surveys, implicitly accounting for this selection effect in constructing each model system. This model fits to the orbital period distributions of the inner three planets for all systems containing $N > 2$ planets as of January 1st, 2024. There are 218 systems confirmed as such.
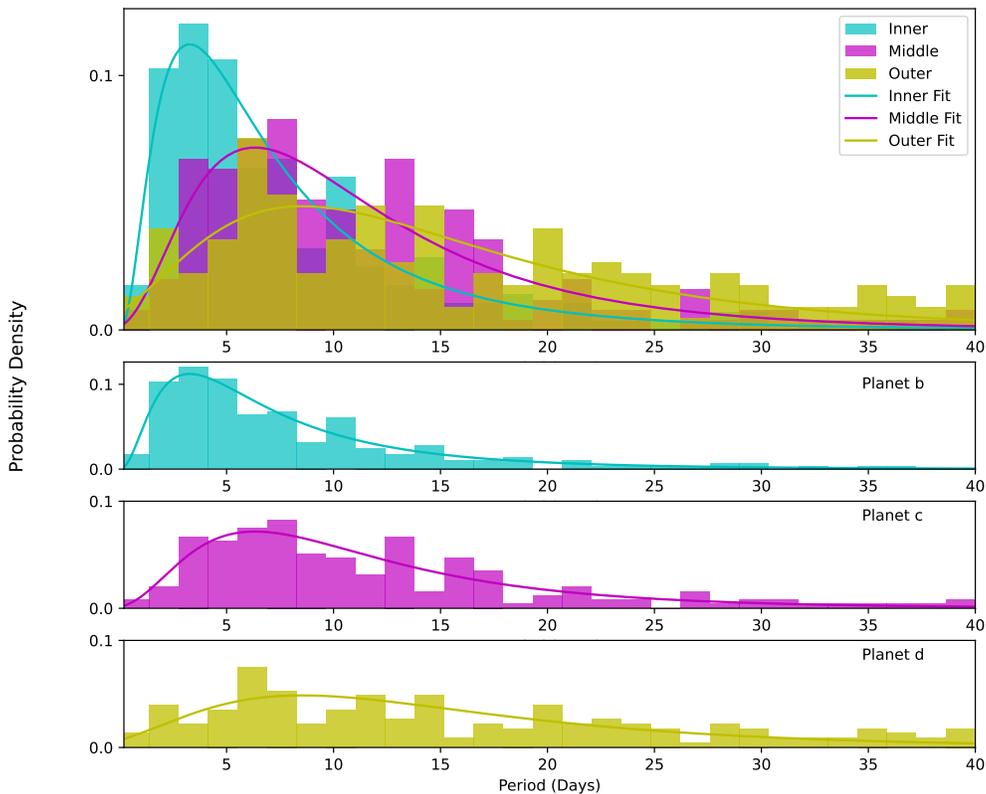
Figure 3.1: Log-normal distributions of all the $N = 3$ planetary systems, constituting the "square-one" model for orbital period.

Running a two-sided KS test (Sec. 1.4.1) between the observed period histograms and best-fit lognormal distributions (shown in Fig. 3.1) yields p-values of 0.83, 0.92, and 0.32 for the inner, middle, and outer cases, respectfully. This test allows us to determine the likelihood that a given empirical distribution (in this case, each observed split orbital period distribution) is randomly drawn from some given underlying distribution (here, the lognormal fits). Specifically, it determines the likelihood that the maximum difference between the two distributions occurred randomly, with a result of $p > 0.05$ considered significant for the empirical distribution being drawn from the underlying. As such, the fit distributions are all considered statistically significant, motivating the lognormal fits. We discuss the KS test as a summary statistic more in Sec. 1.4.1, and its utility in astrophysical population science cannot be overstated.

With this model, we construct a large number ($n = 10,000$) mock $N = 3$ planetary systems from independent draws from each distribution, enforcing that the innermost planet period must be drawn from the innermost distribution and that each successive planet is drawn from its respective distribution. This gives us a base catalog of orbital periods over which we can construct a radius and mass distribution, which we outline in Sec. 3.3. As these distributions are empirically determined from the observed population, for simplicity we do not implement any procedure for probing detectability and do not simulate characteristics for the host stars of these systems. Each system is simulated to be considered observable, with each planet at an inclination of 90° and eccentricity of 0.

With the only dependence between the three lognormal distributions being that each successive planet must have an orbital period larger than its inner companion(s), the main characteristic of this model is that it is completely agnostic to planet-planet interactions and other mechanisms that can gravitationally influ-
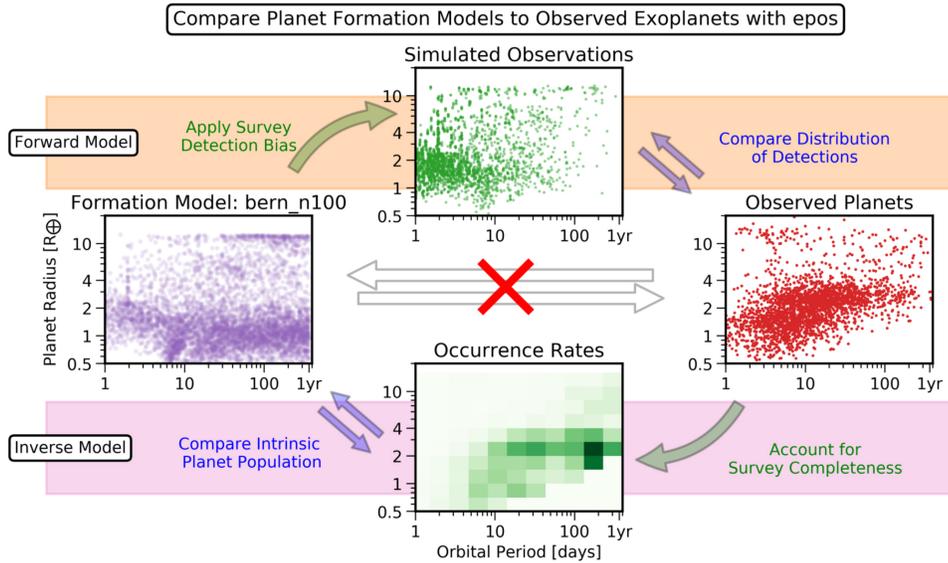
Figure 3.2: From Mulders et al. (2018). Population synthesis can both serve to infer future populations with forward modeling and explain existing populations through inverse modeling, which can account for the completeness of a given exoplanet survey. This allows us to parametrically compare models for planet formation to the observed sample of exoplanets, which will be affected by observational biases and survey incompleteness.

ence system architectures. With the removal of unstable systems through N-body analysis described in Sec. 3.4, this model primarily serves as a benchmark to compare to other populations, both modeled and observed, the strength and influence of such resonant interactions in constructing the known population.

3.2   EPOS

The Exoplanet Population Observation Simulator (`EPOS`) is an incredibly versatile exoplanet population synthesis code in its ability to apply selection and incompleteness effects and provide parameter estimations to a wide variety of exoplanet surveys, including the transit and radial velocity search methods. For this work, we primarily use its functionality to mimic the Kepler Data Release

25 (DR25) population, the final *Kepler* data release which to date dominates the transiting multiplanet population (see Fig. 1.1).

EPOS (Mulders et al. 2018, 2019) was initially developed to inform and forward-model the *Kepler/K2* sample. For our purposes, it is useful to leverage as a more sophisticated method for simulating populations as it can generate model exoplanet catalogs based on empirically determined planetary occurrence rates which take into account survey incompleteness factors and detection biases. Fig. 3.2 shows the versatility in population synthesis methods that EPOS affords. For this work, we employ fits to the *Kepler/K2* DR25 population as the base catalog from which we draw our underlying population.

In contrast to the independent draws for each consecutive orbital period as used in Model A, this model assumes a dependence between the orbital periods for outer planets and that of the innermost planet, with each planet's position drawn from a lognormal distribution scaled by orbital period relative to its inner adjacent neighbor. For outer planets, this is accomplished by the dimensionless spacing parameter D (Malhotra 2015). Where $\mathcal{P}_k = \frac{P_k}{P_{k-1}}$ is the pairwise period ratio for planet k relative to its inner companion, $D_k = 2\frac{\mathcal{P}_k^{2/3}-1}{\mathcal{P}_k^{2/3}+1}$. We provide the logarithm of the D_k distribution for our observed transiting sample in Fig. 3.3. Peaks near resonant ratios are clearly seen, particularly near the 2:1 and 3:2 resonances, which have D values of -0.34 and -0.57, respectively.

First, the orbital period of innermost "seed" ($k = 0$) planet is drawn from a broken power-law distribution given by

$$f_P(P_{k=1}) = \begin{cases} (P_{k=0}/P_{\text{break}})^{a_P}, & P_{k=0} < P_{\text{break}} \\ (P_{k=0}/P_{\text{break}})^{b_P}, \end{cases} \tag{3.1}$$

where $a_P \approx 1.5$, $b_P \approx 0$, and $P_{\text{break}} \approx 10$ days are assumed prior values, reflecting

the flat occurrence rates for subneptune-sized planets observed above P_{break} (see Mulders et al. 2018; Howard et al. 2012; Mulders et al. 2015). Exact values for a_P, b_P, and P_{break} are determined with a Markov Chain Monte Carlo parameter estimation implemented in EPOS (see Sec. 1.4.2) for each run of the Monte Carlo simulation. The MCMC run is used to minimize the distance between the observed and simulated distributions for varying values of these parameters and others discussed below. For these on-the-fly parameter estimations, we run 100 steps with 20 walkers. These are sufficient to produce a plausible model population with good computational efficiency (see Mulders et al. 2018, and EPOS documentation).

With the innermost orbital period generated with Eqn. 3.1, the period of each successive outer planet is drawn from the lognormal distribution

$$f(P_{k+1}|P_k) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(\log D_k - D)^2}{2\sigma^2}} \tag{3.2}$$

where the lognormal fit parameters σ and D are approximately 0.2 and -0.4 respectfully, given by Malhotra (2015). This allows for each planet's orbital spacing to be drawn recursively and interdependently, while building up a system's successive orbital periods through random draws from Eqn. 3.2.

To generate an observable population to mimic the *Kepler* DR25 catalog, a power-law dropoff in planets per star is fit to the observed occurrence rates using an MCMC parameter search (Sec. 1.4.2). The effect of this is shown in Fig. 3.4. As multi-planet mode for EPOS is designed to simulate the *Kepler* DR25 population, each simulation run generates approximately 250 observable $N > 2$ systems. For the purposes of this work, we desire a larger population than observed to be generated and to have this population sample a variety of results from the MCMC parameter estimations, while reducing computational expense. As such, we run
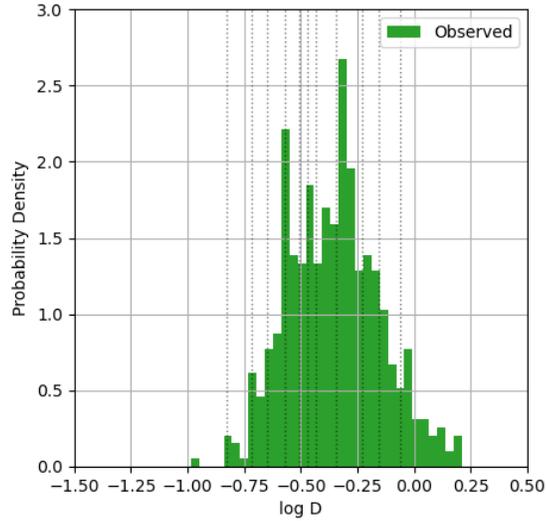
Figure 3.3: Logarithm of D histogram for the observed $N > 2$ population, with vertical lines at low-order commensurabilities. We note peaks at $D = -0.34$ and $D = -0.57$, the 2:1 and 3:2 resonances, respectfully.

the simulation several times to draw a population of approximately 1500 systems. This is a sufficient size to compare statistically against a much smaller observed population, which we show in Chapter 4.

3.3   Radius and Mass Determination

Both models adopt a broken power law parameterization for the planet radius distribution,

$$f_R(R_k) = \begin{cases} (R_k/R_{\text{break}})^{a_R}, & R_k < R_{\text{break}} \\ (R_k/R_{\text{break}})^{b_R}, \end{cases} \tag{3.3}$$

with values from Mulders et al. (2018) of $a_R \approx 0$, $b_R \approx -4$, and $R_{\text{break}} \approx 3M_\oplus$. For the Model A simulation we set the parameters at these fit values, but for the EPOS population, as for the orbital period distribution, these parameters are refined for each population using through an MCMC parameter estimation (see Sec. 1.4.2).
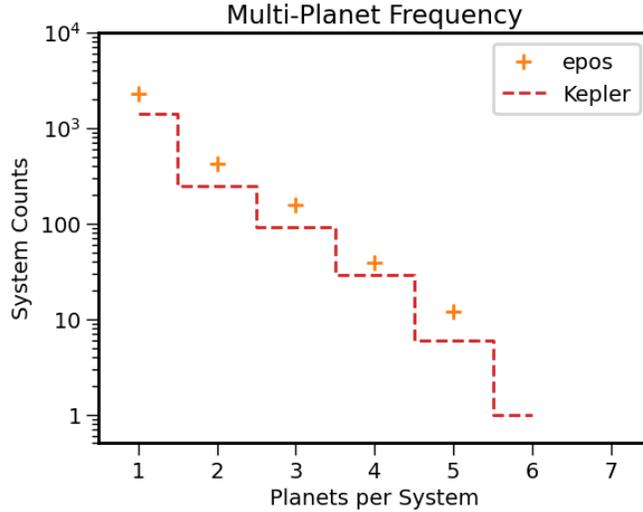
Figure 3.4: Planets per star between EPOS and *Kepler* sample.

We draw from Eqn. 3.3 to determine a planet radius value for each planet in both models. Model A draws independently from $f(R_k)$ for each planet in each system, while the EPOS model assumes a single value for planet radius for all planets in a system. Mulders et al. (2018) makes this assumption noting the tendency for *Kepler* multiplanet systems to have similar masses across their component planets. Such systems, as is visible in Fig. 1.5, do not tend to contain Jupiter mass planets in succession with relatively lower mass planets, so simulating a consistent mass across N planets both serves as a physically-motivated approximation to simplify the generation of the synthetic population.

With radius values generated for the Model A and EPOS synthetic catalogs, we use the `forecaster` python code from Chen & Kipping (2017) to generate probabilistic masses for each simulated planet. This code samples from a broken power law mass-radius relation shown in Fig. 3.5. To aid in computational expense and more precisely factor in the variance off of a direct mass-radius relationship, this code breaks up the broken power law into a configurable number of grid
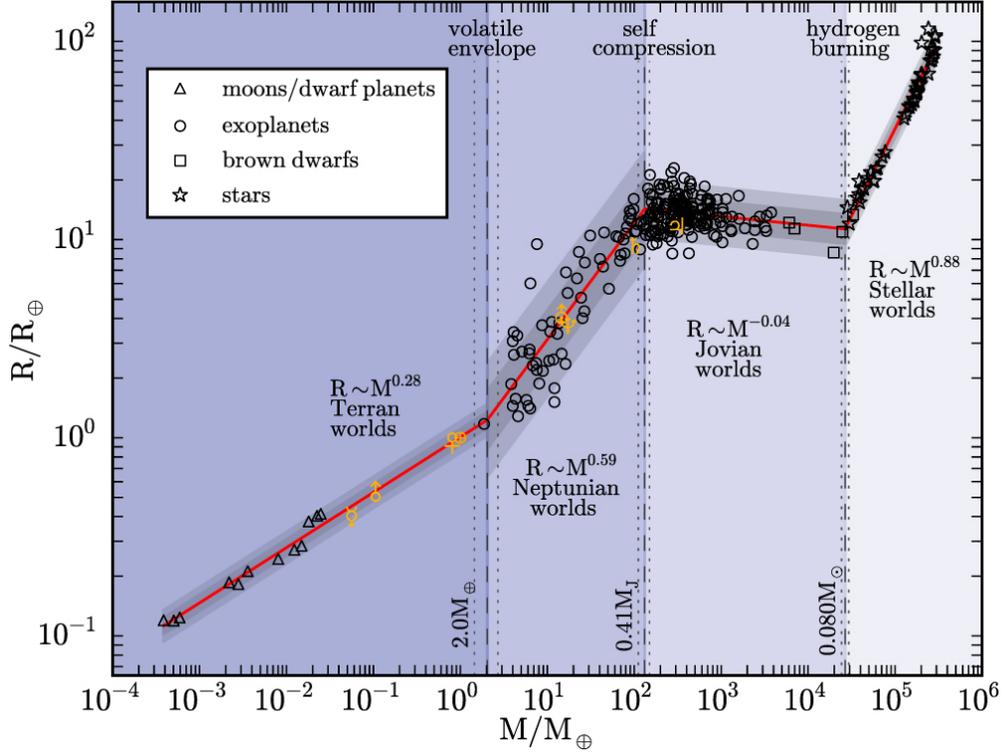
Figure 3.5: Continuous power law mass-radius relation plot from Chen & Kipping (2017), including 68% and 95% confidence intervals for dispersion off the direct relation.

spaces, in this case we use $N_{\text{spaces}} = 1000$. To "forecast" the mass value for a given radius, the code determines the possible mass values given in Fig. 3.5, including the associated spread of the broken power law, and samples accordingly. For a sufficiently large population and a large value for N_{spaces} this results in a well-populated mass-radius distribution for both model cases, shown in Fig. 3.6.

3.4 Stability Analysis and Dynamical Filtering

Unlike the orbital configurations of the observed population, there is no guarantee of short-term stability for our simulated populations. Unstable configurations may have planets undergo ejection or migration on rapid timescales that are not seen in the observed sample. This is a particular concern for Model A, where
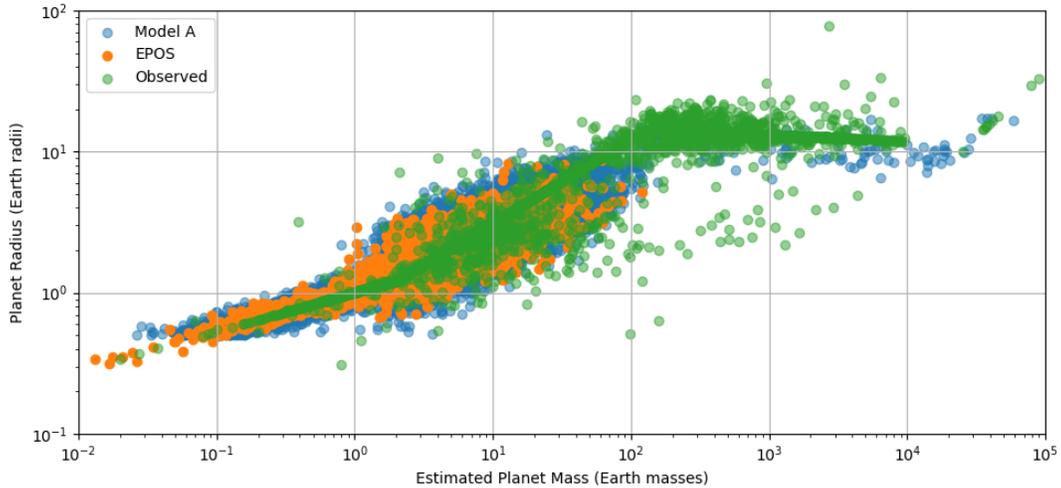
Figure 3.6: Forecasted mass-radius distribution using triple broken power law from Chen & Kipping (2017) for the two model cases, including the observed population.

orbital periods for the $N = 3$ planet case are drawn with no dependence on the location of other planets in the system. Fortunately, there are several numerical stability indicators that are easily implemented in N-body codes, which we can use to remove dynamically unstable systems from our simulated population.

Simulations in this paper made use of the REBOUND N-body code (Rein & Liu 2012), which includes a fast symplectic integration functionality for calculating a widely-used stability criterion in exoplanet science–the Mean Exponential Growth factor for Nearby Orbits (MEGNO). This metric, represented by $\langle Y \rangle$, is essentially a summary statistic gauging the influence of chaotic interactions in perturbing successive orbits over an integration. This requires an N-body integration for each system in our model populations, so to maintain computational expense we set the integration time at $10 \times P_{\max}$. This gives at least 10 orbits for any given planet in a system simulated by our models. Systems with inherently unstable orbital configurations can then be filtered out of the model populations. Simulations of systems with the ejection of a planet are halted and automatically removed
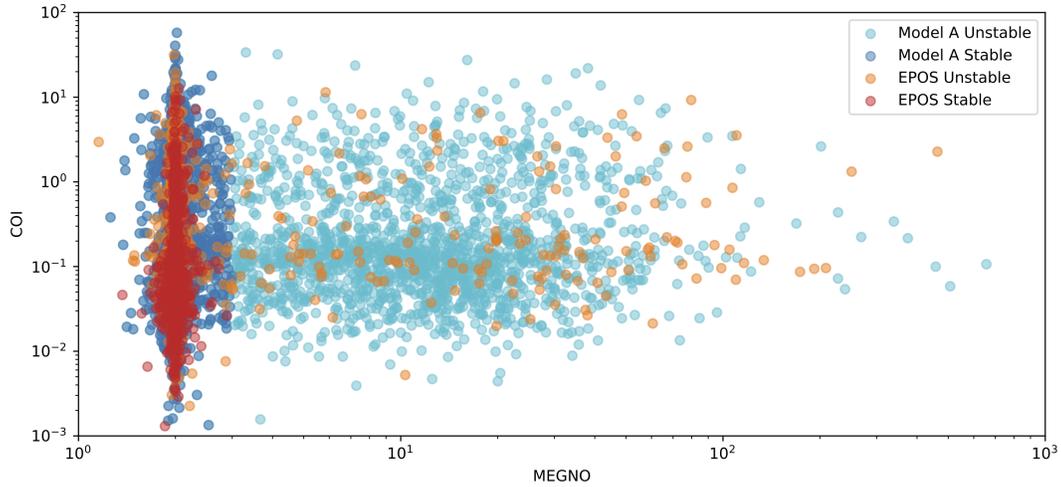
Figure 3.7: MEGNO versus COI for the Model A and EPOS populations, where the stable population is limited to systems with $\langle Y \rangle < 3$. Note that the majority of systems with low COI (< 0.05) are stable whereas at higher COI there is a small, but significant, fraction that are unstable.

from the sample. We discuss the effect of this on the simulated populations in Sec. 4.1. The simulations were integrated using `WHFast`, a symplectic Wisdom-Holman integrator (Rein & Tamayo 2015; Wisdom & Holman 1991). Variational equations were used to calculate trajectories of nearby orbits (Rein & Tamayo 2016). We run the N-body code for each system in the populations, rejecting systems with $\langle Y \rangle > 3$ as unstable. We provide the scatter plot for MEGNO and COI for each system in Fig. 3.7, which include all simulated stable and unstable systems.

Chapter 4

Simulation Results

Here we analyze the results of a Monte Carlo simulation for the populations described in Chapter 3 and compare these simulated populations to the observed sample of transiting multiplanet systems. For each model, we generate $n_{\text{Model A}} = 10000$ and $n_{\text{EPOS}} = 1458$ model systems to compare against $n_{\text{observed}} = 218$ known/candidate systems. We find these simulated samples sufficiently well-populated enough to make meaningful statistical comparison between them and the observed sample.

4.1 Orbital Spacings
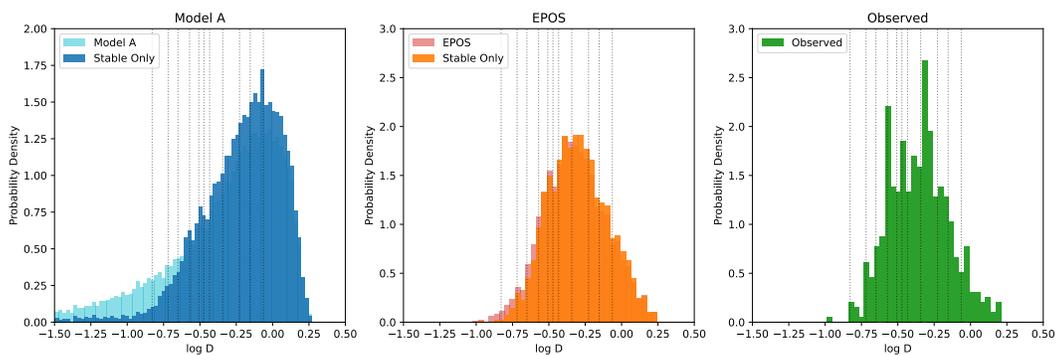
Figure 4.1: $\log D$ distributions for the Model A, EPOS, and observed populations. Here we show the distributions before and after removing unstable systems from the Model A and EPOS populations, and include the low-order commensurabilities used in calculating Eqn. 2.2 as vertical lines.

We begin the evaluation of the simulated samples with a discussion of the stability analysis described in Sec. 3.4, which is primarily affected by the orbital spacings between planets in a system. The effect of this procedure ultimately restricts, in principle, the models to physically feasible values. That is, it is unrealistic for two planets of comparable to be located very close to each other as their mutual gravitational interactions would render such a configuration as unstable. This is different from the gravitational processes leading to resonance, however, (see Chapter 2) as this orbital evolution tends to take place during the later stages of planet formation (see Armitage 2010). Such unstable configurations are very unlikely to make it past these stages, and as such we remove the systems demonstrating chaotic instability from our model populations.

Recalling the dimensionless parameter D from Eqn. 3.2, adjacent planet pairs that are closely situated in orbital period will have a low value of D while pairs situated further apart will have a higher value. In Fig. 4.1 and Fig. 4.2 we plot the distribution of the logarithm of D for the Model A and EPOS populations both before and after the removal of unstable $\langle Y \rangle > 3$ systems. In Fig. 4.3 we provide the corresponding cumulative distribution functions. We note that Model A shows a particularly large change in its $\log D$ distribution before and after enforcing stability, as many systems are generated with planet pairs in close orbital configurations that are generally unstable on short timescales. The EPOS population also shows some tendency for instability in the same regime, though Fig. 4.1 shows the distribution thinning out towards a lower limit of $\log D = -1$ more rapidly than Model A. This is consistent with the simulation procedure discussed in Sec. 3.2, where D is determined through successive lognormal draws in a system, and covers the same domain as the observed distribution. In general, the EPOS method produces a much more stable population than Model A.

We also note the tendency for resonant configurations to stand out in the $\log D$ distributions in the form of peaks along low-order resonant ratios, similar to Fig. 1.6. This is particularly seen in Fig. 4.1 for the observed systems, compared to the relative smoothness of the model distributions. While the dimensionless spacing parameter D gives information on the relative orbital spacings between planet pairs, it does not give any information on the proximity of these pairs to any given period ratio value, such as low-order commensurabilities. It also does not give a description of orbital spacings across an entire system, with planetary having independent orbital spacing values in line with the simulations in Sec. 3.2. In effect, to summarize the system-wide tendency of a system to converge near resonant ratios, COI "collapses down" the offsets in D-space for these values of interest as well. We discuss the COI distribution for the observed and model samples in the following section.
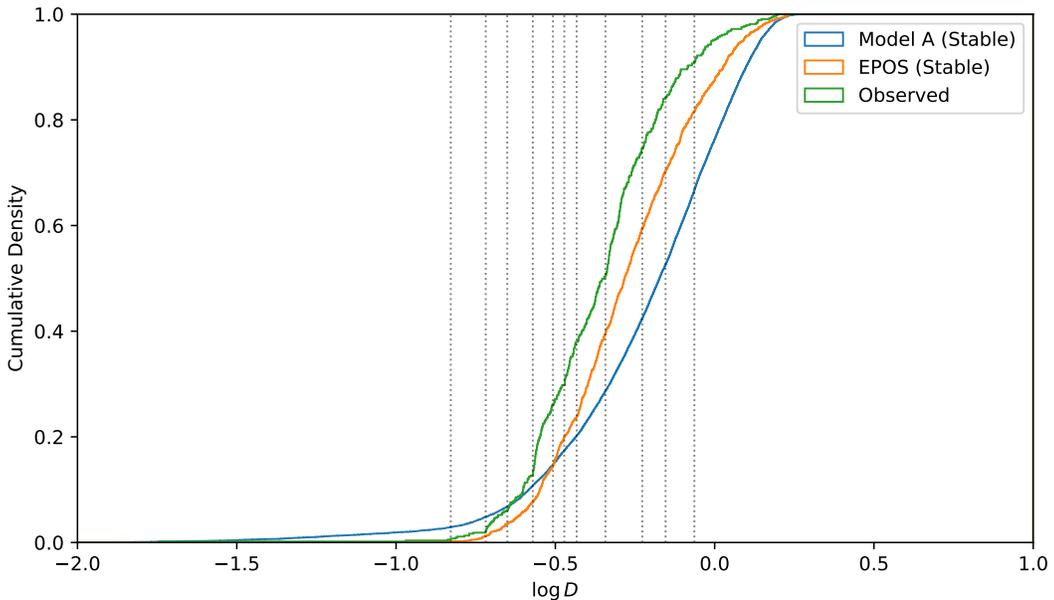
Figure 4.2: $\log D$ distribution for simulated stable and observed populations.
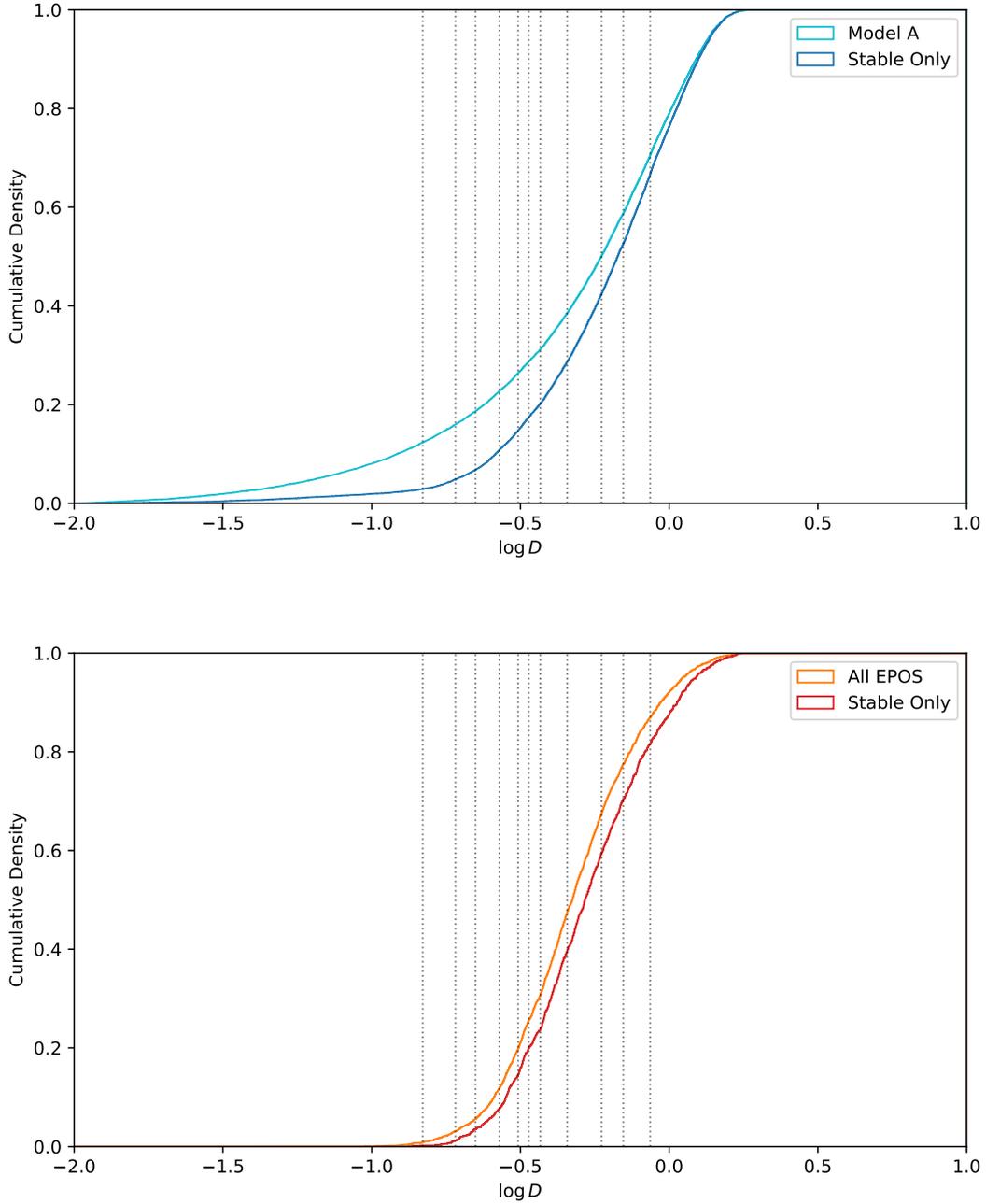
Figure 4.3: $\log D$ cumulative distribution for Model A and EPOS, both for the whole simulated populations and those limited by $\langle Y \rangle < 3$, which are considered stable (see Sec. 3.4).

4.2   System Abundance at Low-COI

In Fig. 4.4, we plot the probability density function and cumulative distribution function of the calculated COI values for the Model A, EPOS, and observed populations. Note that the distribution for the observed population is also shown in Fig. 2.3, though not as a normalized probability distribution. Here, we are particularly interested in comparing the nature of the distributions at low values of COI, between 10^{-3} and 10^{-1}.
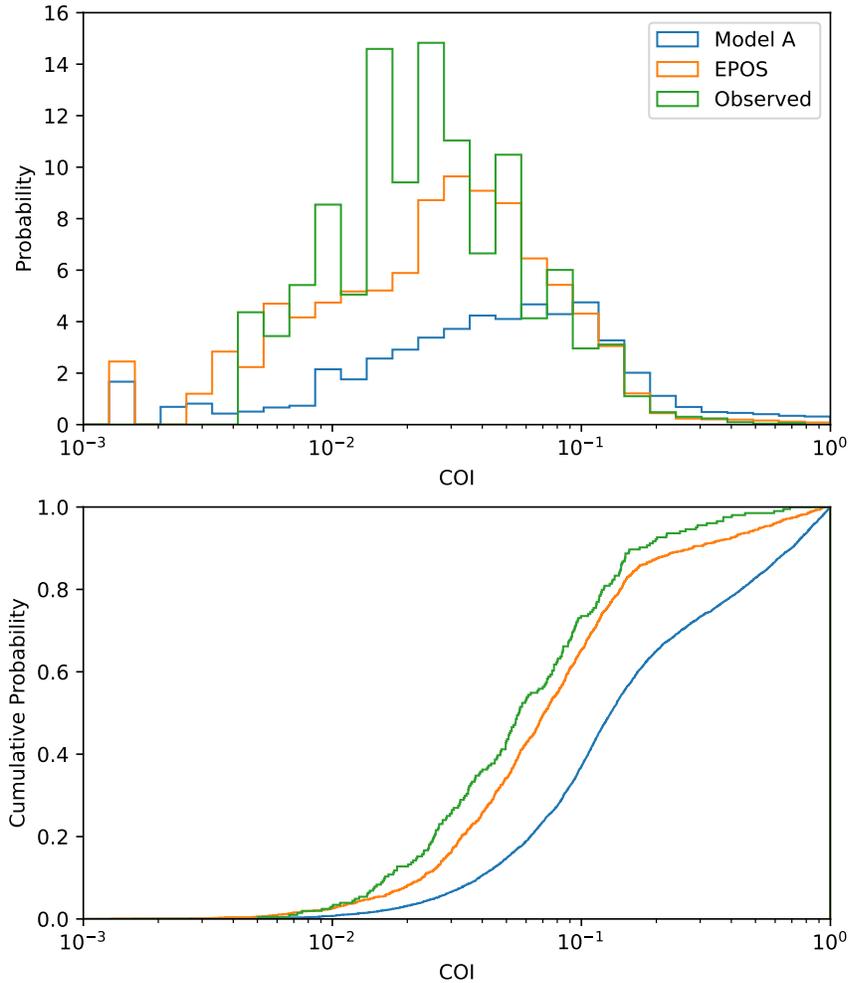
Figure 4.4: COI distributions for Model A, EPOS, and the observed population.

We note an overabundance of systems within this low-COI regime in the observed population compared to EPOS and Model A, shown in Fig. 4.4. This is particularly notable for the Model A population, which peaks around $COI = 10^{-1}$, compared to the peak around $COI = 2 \times 10^{-2}$ for the observed distribution. This overabundance does not appear to be an artifact of normalization on a logarithmic horizontal axis, and appears with linear binning as well.
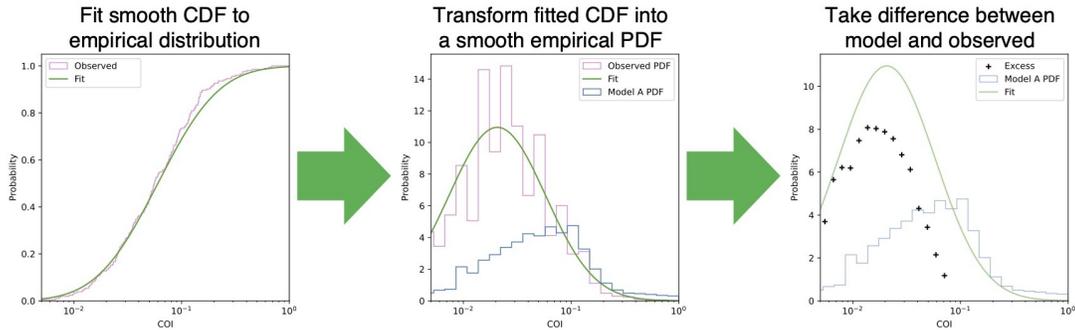
Figure 4.5: Probability density estimation procedure for probing the overabundance in the observed COI distribution compared to Model A (see text).

To quantify the difference between the observed and Model A distributions we run a density estimation from the cumulative distribution function of the observed sample. We desire a smooth analytic form for the observed distribution of COI in the form of a lognormal, which can be fit to the cumulative distribution and then transformed into a probability density function. This procedure is shown in Fig. 4.5. First, we fit a lognormal cumulative distribution function to the cumulative distribution of COI for the observed distribution. We fit this to the cumulative distribution as, unlike the probability density function, it will not change with arbitrarily large binning (here we use $N_{\text{bins}} = 1000$). We then transform the cumulative distribution into a probability density function to serve as a smooth parametric form for the observed COI distribution. As the Model A distribution is well-populated and does not appear to follow a lognormal trend, we choose not

to fit a smooth parametric form to its probability distributions. This enables us to take the difference between a smooth observed distribution and the Model A normalized histogram. We plot the positive values of the difference as the "excess" between the two distributions in Fig. 4.5, and note that this excess distribution has a population mean $\mu = 0.0211$ and a variance $\sigma^2 = 0.0002$, implying a standard deviation of $\sigma = 0.0141$.

The significance of this overabundance lies in the assumption made *a priori* that the Model A population is completely agnostic to the arising of resonant architectures within a system as an end-product of system formation. It only seeks to mimic the population in terms of individual orbital period values rather than in terms of relative orbital spacings. As such, it may serve as a benchmark for a physically-minded population independent of effects leading to convergence upon resonant configurations. This may indicate that the tendency toward commensurability noted in Sec. 2.1 observed on the scale of the solar system and some exoplanet systems plays a role in sculpting the architectures of multiplanet systems well beyond the regime of strictly resonant systems.

4.3   Occurrence Rates of Commensurabilities

We provide the relative occurrence rates of pairs in proximity to the low-order commensurabilities considered herein in Fig. 4.6 for the Model A, EPOS, and observed populations. We do not establish a cutoff for the proximity of a given period ratio to a certain commensurability in this visualization. Here the similarity between the EPOS and the observed *Kepler* DR25 population in terms of this metric is notable, with similar occurrence rates of the many resonances demonstrated between the model and observations.
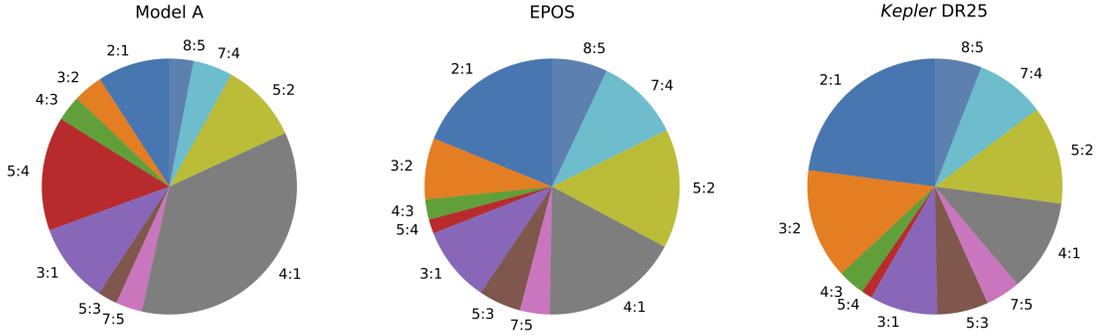
Figure 4.6: Relative occurrence of low-order commensurabilities within the Model A, EPOS, and observed populations.

We note that Model A in particular generates more pairs close to the 4:1 and 5:4 commensurabilities, and generally does not produce a population with similar occurrence rates as EPOS. This is due to how the area in period-ratio space considered "in proximity" to these resonances is large compared to other commensurabilties. In Fig. 4.7 we demonstrate this phenomenon across the two model populations. As the Model A population is agnostic to the presence of resonant configurations influencing the period ratio distribution of each planet pair in a system, the distribution in period ratio space appears more-or-less uniform compared to the observed population (Fig. 2.2) and the accompanying EPOS population. This is particularly notable for the 4:1 resonance, but a much larger fraction lies in proximity to the 5:4 resonance for Model A compared to the observed population as well. While the EPOS distribution of period ratios is less populated than Model A, there still appears to be a slight bias toward the regime of low-order commensurabilities–though not directly upon them as seen in the observed distribution–and a more subtle overabundance of pairs closest to the 4:1 chain compared to Model A. This is likely due to the lognormal sampling of D, enforcing compactness upon adjacent pairs in a multiplanet chain while still over-representing the 4:1 regime due to its large area in period ratio space.
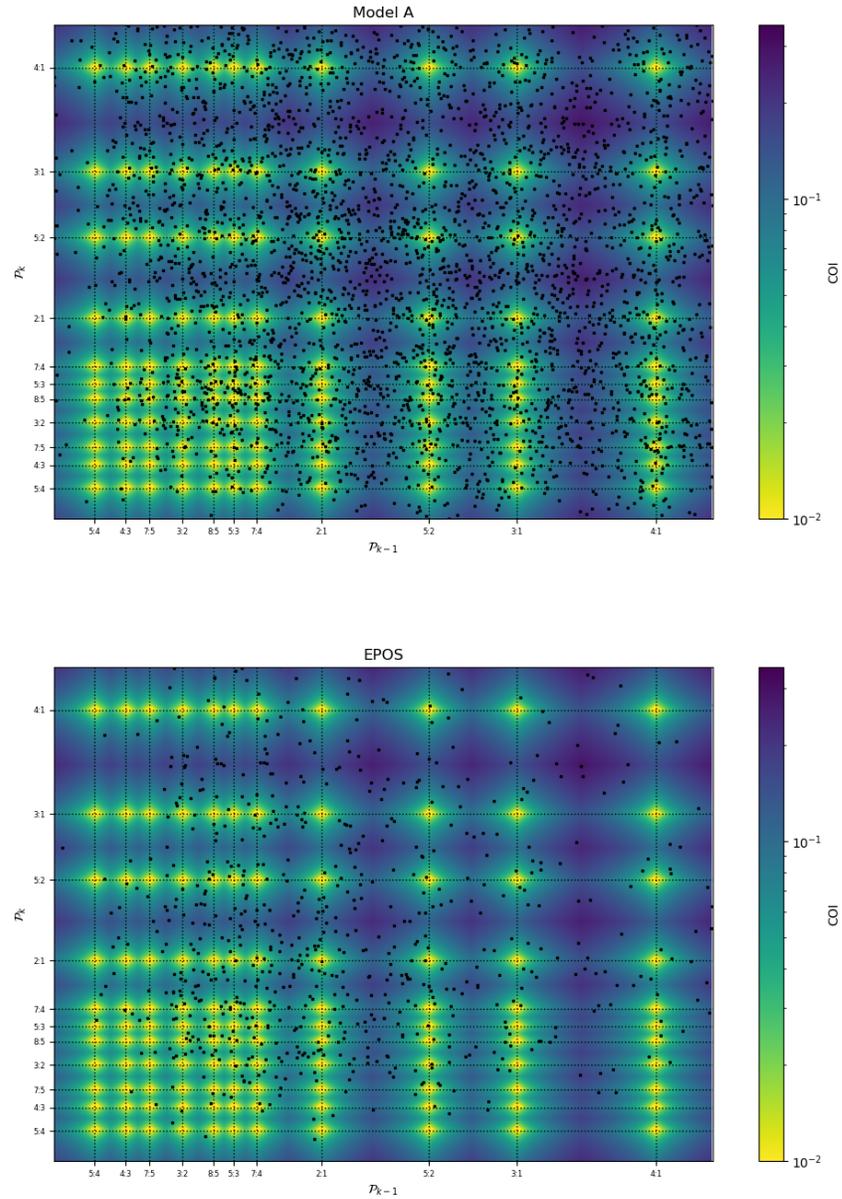
Figure 4.7: The same as Fig. 2.2, but for the Model A and EPOS populations. Lowest pairwise period ratios in each system, used in the calculation of COI for each system.

Chapter 5

Discussion

Here we discuss the significance of various aspects of the model populations described in Chapter 3 and discuss the relevance of our modeling to other related astrophysical populations. To this end, we compare the considered populations to the sample of angular substructures in planet-forming disks from the Disk Substructures at High Angular Resolution Project (DSHARP) catalog observed by the Atacama Large Millimeter/submillimeter Array (ALMA) presented in Huang et al. (2018), as well as various groupings of solar system objects in varying proximity to resonant configurations. We also discuss the possibility of using COI to help search for outer bodies in confirmed transiting $N = 2$ planet systems.

For one, we briefly discuss the prior assumption made in the EPOS population that successive planets should have similar planet masses, with the initial innermost planet mass determined by sampling a broken power law which is flat within $3M_\oplus$ (see Eqn. 3.3; Mulders et al. 2018, 2019), which may explain the weak period-radius relation in Fig. 1.5. This is referred to as the "peas in a pod" phenomenon and was first identified in Weiss et al. (2018). There is debate as to whether this pattern is intrinsic to the observed population or a bias of the decreased sensitivity to detect larger planets at larger orbital separations with the detection limits inherent to the *Kepler* survey (Zhu 2020). However, work by Weiss & Petigura (2020) suggests that this pattern is in fact physical and that

the presence of outer planets would emerge in survey observations with *Kepler*'s sensitivity. With the "peas in a pod" model assumed in this work, we plot the period-radius relation for the EPOS population in Fig. 5.1. We note the similarity between the model and the same plot for the observed in Fig. 1.5, including the slight upward tendency in the distribution indicating a weak period-radius relation.

In the following subsections, we discuss the various applications of our analysis to other relevant multiplanet samples, including the ALMA angular substructures catalog from the DSHARP survey presented by Huang et al. (2018) and various groupings of solar system planets. We also speculate as to the possibility of using COI as a predictive summary statistic to find outer unseen bodies in exoplanet systems containing $N = 2$ confirmed planets.

5.1   Comparison to ALMA Angular Substructures

The DSHARP angular substructures catalog presented by Huang et al. (2018) is a catalog of 18 protoplanetary disks around nearby young stars observed by ALMA, a large-scale microwave interferometer located in Northern Chile. Specifically, this catalog probed disks with a spatial resolution of approximately 5 au for angular features indicating the presence of planet formation in the disks. Assuming these substructures correspond to a planetary body forming at the corresponding orbital distance, such observations would be caused by the nascent protoplanet "clearing the neighborhood" along the path of its orbit.

Young planetary systems in particular are an interesting sub-sample in which to probe resonant configurations. Current models of planet formation (see Sec. 2) suggest that planets are drawn toward resonant configurations during the early
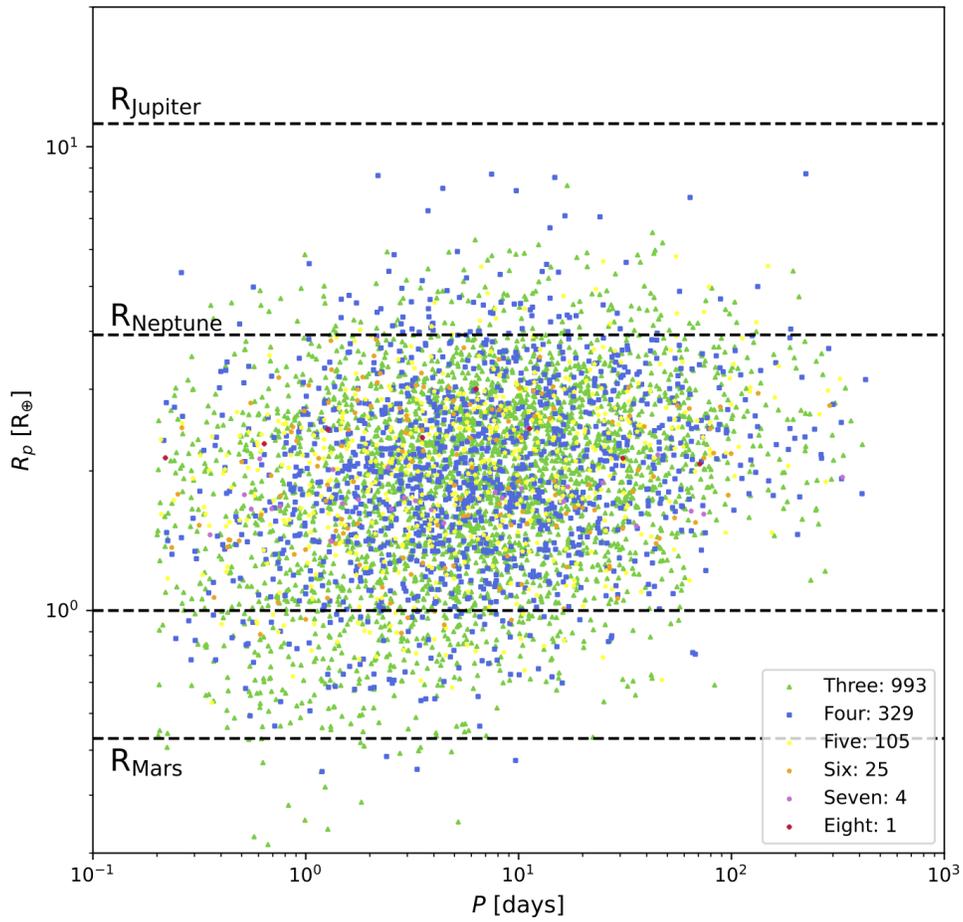
Figure 5.1: Remake of Fig. 1.5 with the simulated EPOS population.
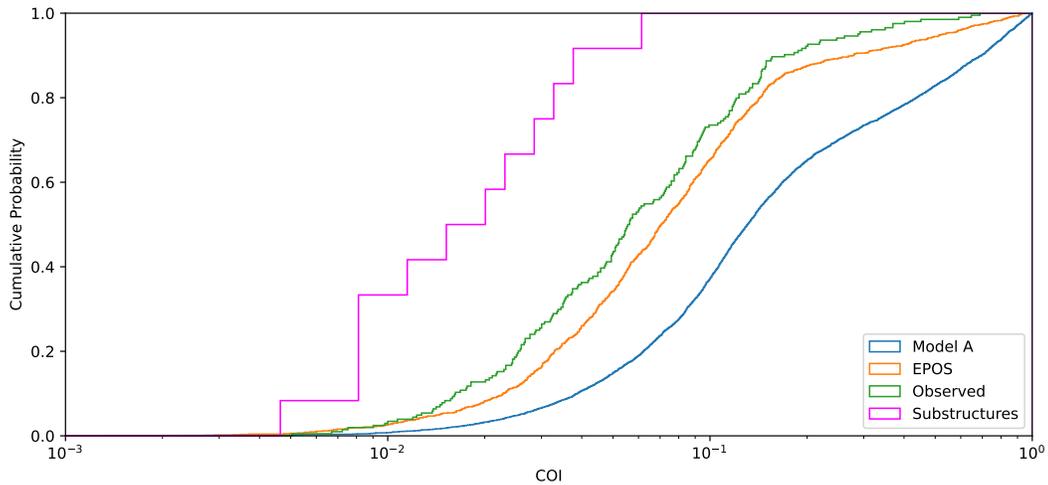
Figure 5.2: Cumulative distribution functions for model and observed multiplanet populations, including the DSHARP angular substructures sample presented in Huang et al. (2018).

disk stage, with torques from other forming planets and the disk itself causing protoplanets to converge upon resonance (see Armitage 2010). With the dissipation of protoplanetary disks, these planets may diverge from resonance with continued torques from companion planets and other gravitational torques. This sample, though, should theoretically correspond to a sample displaying a strong tendency toward resonant configurations.

In constructing a sample to probe the COI distribution of these sources, we exclude DoAr 33, Elias 27, SR 4, Sz 114, WaOph 6, WSB 52 from the angular substructure sample as these systems only contain two detected substructures, and our analysis requires a minimum of three "bodies" for a measured COI value. For the systems containing more than two detected substructures, we determine the COI by directly taking the ratio of the orbital periods and summing the offsets to the nearest commensurabilities according to Eqn. 2.2. We note that the catalog presents the orbital distance to the angular substructure as the semimajor axis,

a direct observable given a known distance to the star and inclination angle. As such, we directly take the period ratios $\mathcal{P}_i = \left(\frac{a_i}{a_{i-1}}\right)^{3/2}$ in accordance with Kepler's third law.

In Fig. 5.2 we provide the cumulative distribution functions for the substructures sample compared to the Model A, EPOS and observed transiting populations. While the sample size of the substructures catalog is small, their COI values exclusively cluster within the regions close to strict resonance within the transiting sample. This may be expected due to the tendency for planets to converge upon resonant configurations during the disk stage of planet formation. We cannot conclusively claim this, however, as we found many of these pairs in proximity to second-order resonances, which may be unstable in certain cases during the disk stage and beyond. In Sec. 5.3 we compare this distribution to various groupings of solar system objects.

5.2   Searching for outer unseen planets with COI

For systems with $N = 2$ planets in particular, this procedure motivates the prospect of finding possibly undetected outer planets by searching the orbital period space around resonant ratios. More than 500 $N = 2$ In Fig. 5.3 we plot the histogram of the individual pairwise offsets factoring into the COI calculation for each observed system with $N = 3$ planets. These distributions have mean values of 0.028 and 0.25 and variances of 0.0007 and 0.46, respectively.

We plot the distribution for $N = 3$ planet systems as the strictest pairwise offset in the COI calculation is often contributed by the closest-in pair. For example, if a pair in an $N = 2$ planet system is found to have an offset within the lower offset distribution, between COI $= 10^{-3}$ and COI $= 10^{-1}$, this may indicate
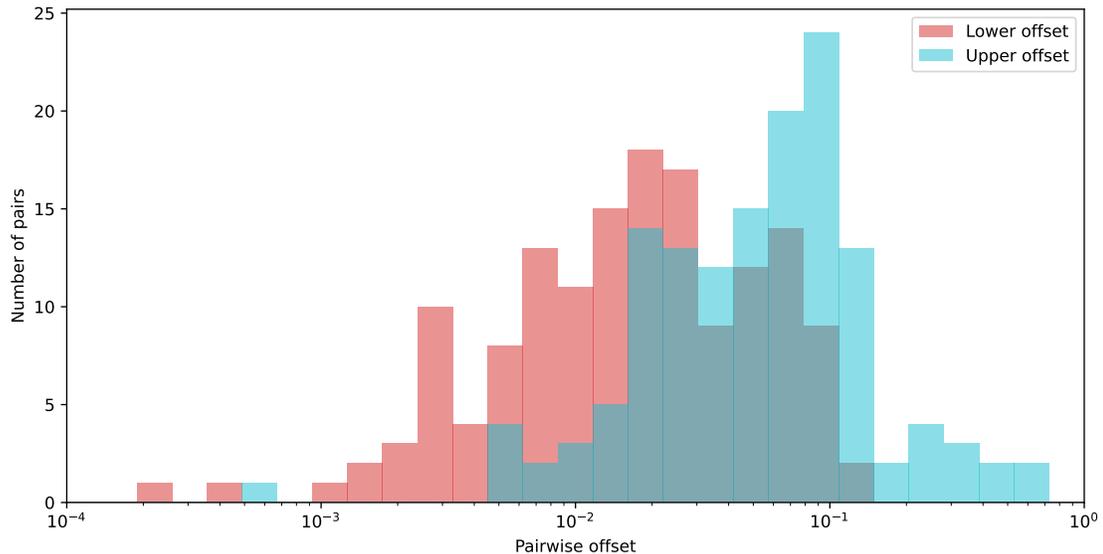
Figure 5.3: Distribution of individual pairwise offsets factoring into the COI for observed transiting systems with $N = 3$ planets. We provide the mean and variance of the distributions in Sec. 5.2.

the presence of a planet in the range of the upper offset distribution, with the orbital period solved for in the COI calculation (Eqn. 2.2). The relative occurrence rates of commensurabilities for the EPOS and *Kepler* populations may serve as priors in determining the particular resonance to which a pair is in proximity. Essentially, the relative abundances of commensurabilities in Fig. 4.6 may serve as a likely list of resonances for an outer planet to occupy, with the spread of the upper offset distribution in Fig. 5.3 serving as a tolerance range in which the position of outer bodies may be constrained.

This is particularly relevant to systems probed with low observing cadence where outer planets with lower detection probability (Eqn. 1.3) may not have occurred during survey observations. Such work could accompany transit timing variations (see Ballard et al. 2011) to constrain the location of such unseen planets. There are more than 500 transiting $N = 2$ planet systems to which this technique could apply.

5.3   Comparison to Solar System
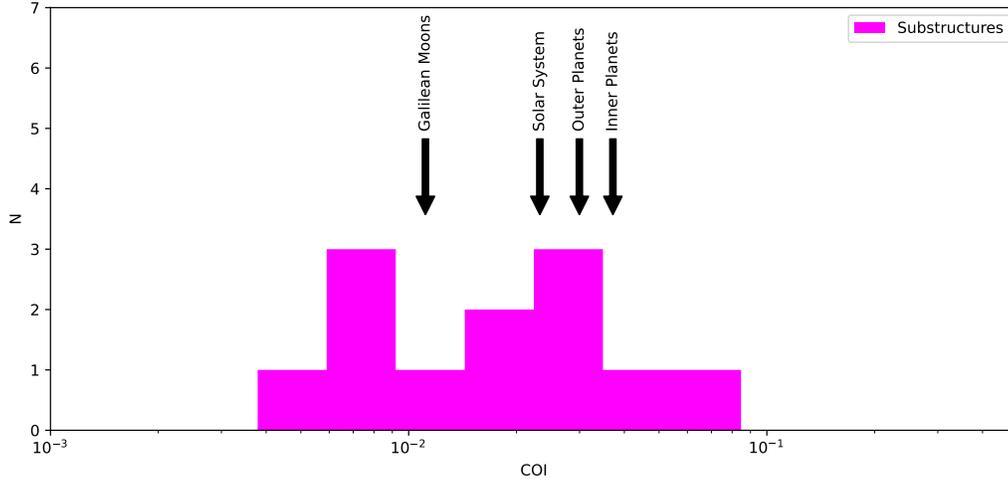
Figure 5.4: COI histogram for the Huang et al. (2018) substructures sample, compared to the calculated values for the solar system. We include the calculated COI value for the whole solar system as well as only the inner rocky planets and outer gas giants, and the Galilean moons of Jupiter. The details of these calculations are given in Table 1.

Here we present the calculation of COI for various groupings of the solar system planets, including the Galilean moons, and compare them to the distribution of COI for the angular substructures catalog from Huang et al. (2018) and to the COI distribution for transiting $N > 2$ planet systems. In Fig. 5.4 we plot the histogram of the COI distribution for the angular substructures catalog and include markers for the COI values of various groupings of solar system planets. We group the planets into the rocky planets and gas giants, taking the COI values for these subsets and the entire solar system. We also include the calculation for the Galilean moons presented in Chapter 2. We provide the calculated values for these groupings in Table 1, which includes the specific bodies factoring into each calculation.

Table 1: Calculated COI values for various solar system groupings, including the Galilean moons of Jupiter. We also include the specific orbital pairs factoring into each calculation.

	COI	Bodies Included
All	0.0232	Earth:Venus and Saturn:Jupiter
Rocky Only	0.0372	Venus:Mercury and Earth:Venus
Gas Only	0.0230	Saturn:Jupiter and Neptune:Uranus
Galilean Moons	0.0111	Europa:Io and Ganymede:Europa

It is interesting to note that these solar system scale COI calculations fall well within the range of values for the Huang et al. (2018) catalog. For one, the Galilean moons and inner planets both belong to similar classes of systems–that is rocky systems displaying the "peas in a pod" phenomenon–though the inner planets lie in a range of orbital periods too high to be effectively detected by the transit method (Eqn. 1.3) and too low to be detected through interferometric observations (that is, during their formation). Most of relevance, therefore, is the gas giants. These display the most similarity to the substructures sample and are most analogous in orbital separation and planet type.

These calculated values also lie on the lower end of the COI distribution in Fig. 2.3 and Fig. 4.4 for the observed transiting sample. For this sample the Galilean moons and rocky planets are the most relevant, so it is notable that the COI calculations for these groupings fall well within the range of the larger transiting sample. While the rocky planets would not be detected as an $N = 3$ planet chain using the transit method, the similar COI values between the populations may reflect the convergence on resonant configurations influencing the formation of the solar system, which some dynamical modeling has historically suggested (Goldreich 1965; Patterson 1987).

Chapter 6

Conclusions and Future Work

In this thesis we present a new metric, the commensurability offset index (COI), to determine the proximity of multiplanet exoplanet systems to containing a chain of mean motion resonances. We apply this metric to the observed transiting sample, dominated by systems detected by the *Kepler* space telescope, and compare this population to two model samples constructed using Monte Carlo methods. These models include a sample which generates orbital periods independently for each successive body in $N = 3$ planet systems ("Model A") and a sample which builds up successive orbits using a dimensionless orbital spacing parameter ("EPOS"). We simulate planet radii and masses for the catalogs constructed using these models resembling the "peas in a pod" phenomenon noted in Weiss et al. (2018) and Weiss & Petigura (2020). We then enforce stability for the model catalogs using N-body analysis. With these models informed by empirical trends we are able to mimic the transiting $N > 2$ planet sample using a Monte Carlo simulation.

We apply the COI metric to the model populations and find an overabundance of systems between the observed distribution and Model A, peaking at a value of COI $= 0.0211$. Model A is constructed to be agnostic to the gravitational effects causing orbital evolution into resonant configurations, so this overabundance at low COI for the observed sample evidences that proximity to orbital resonances

may be overrepresented in the observed sample. As such, these resonant effects may play a significant role in the orbital evolution of systems within the sample.

We discuss the possibility of using this metric to search for outer undetected planets in $N = 2$ planet systems, which are particularly numerous in the transiting catalog numbering more than 500. We also apply COI to relevant non-transiting populations, including the ALMA DSHARP catalog of angular substructures in protoplanetary disks from Huang et al. (2018) and several groupings of solar system objects, including the rocky and gas planets and the Galilean moons of Jupiter. Interestingly, the distribution of COI values for the angular substructures occupies a similar range to the solar system objects, with the angular substructures sample having COI < 0.085 and the solar system groupings having COI < 0.04. These occupy the lower end of the COI distributions for the transiting sample, which is particularly interesting for the angular substructures sample as resonances are expected to form during the planet formation stages probed by the DSHARP survey.

This is an exciting time for exoplanet population science as this field will continue to detect a larger sample through a growing variety of means, both with and beyond the transit method. As *TESS* matures, the population of confirmed systems will grow as it continues its observations across the celestial sphere. The mission will further explore the short-period sample limited to below 27.4 days, discovering new systems across the sky, and follow-up observations will extend these observations to build up a more robust multiplanet sample. Follow-up observation is crucial to this work as it serves both to help confirm the thousands of candidate planets through detecting additional transits, as well as to probe these candidate systems for additional outer planets. Small facilities like the Van Vleck Observatory here at Wesleyan are an important part of this work to extend the

observing cadence of these systems and uncover more multiplanet systems. Such work may be assisted by applying the COI metric for candidate $N = 2$ planet systems to constrain the orbital distances of possible outer planets.

There is also the GAIA astrometric population on the horizon, which will detect exoplanets through precise measurements of their host star as it orbits around their common center of gravity. As such, this sample will be more sensitive to distant orbits and is estimated to detect 21000 ± 6000 new exoplanets according to estimates from Perryman et al. (2014). Such a catalog, with multiple times more planets than currently confirmed, can probe resonances at much higher orbital separations, including analogs of outer solar system planets.

Such work on angular substructures is ongoing with the ALMA survey to Resolve exoKuiper belt Substructures (ARKS), an ALMA Large Program that will probe substructures in Kuiper belt analogs within $30 - 300$ au of their host stars. Radial and vertical substructures caused by the presence of planets with masses ranging from Jupiter mass and Pluto analogs will be detectable within these disks. The survey will probe the disks around 18 nearby stars, and it will be interesting to compare the COI distribution from ARKS to that calculated for DSHARP in this work, as well as the large future population from GAIA which probes a similar range of orbital distances.

With these future populations and an ongoing effort towards follow-up transit observations, it remains an open question to see if a much larger sample of non-transiting exoplanets would peak at the same COI value as the angular substructures catalog, and if that range would coincide with the ever-increasing transit catalog. Approaching from a different angle, it would be interesting for theoretical modeling to analytically reproduce the overabundance between a model assuming independent orbital periods (like Model A) and one with probabilistic

orbital spacings (like the observed sample, and EPOS) from first principles.

Population science serves to remind us of the minuscule window we possess on a large and changing universe. It allows us to take our momentary observations and contextualize them into a larger paradigm. But beyond the science, this is something we can feel if we stop and look. When gazing at the sky, day or night, think of all the things that you *don't* see. Think of all the bodies forming and dying, of all the planets transiting undetected, of all the worlds yet unseen by human inference. And think of the transient events that we do fortuitously detect, such as the total solar eclipse that passed a week before this thesis was due. How lucky we are to live in such a universe as this.

Bibliography

Almeida, A., et al. 2023, ApJS, 267, 44

Andreon, S., & Weaver, B. 2015, Bayesian Methods for the Physical Sciences. Learning from Examples in Astronomy and Physics., Vol. 4

Armitage, P. J. 2010, Astrophysics of Planet Formation

Babu, G. J., & Feigelson, E. D. 1996, Astrostatistics

Ballard, S., et al. 2011, ApJ, 743, 200

Borucki, W. J. 2016, Reports on Progress in Physics, 79, 036901

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

Chen, J., & Kipping, D. 2017, ApJ, 834, 17

Dai, F., et al. 2023, AJ, 165, 33

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

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

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

Fulton, B. J., et al. 2017, AJ, 154, 109

Gaia Collaboration et al. 2018, A&A, 616, A1

Galilei, G. 1610, Sidereus Nuncius (Apud Thomam Baglionum)

Ghosh, T., & Chatterjee, S. 2023, ApJ, 943, 8

Gillon, M., et al. 2017, Nature, 542, 456

Goldreich, P. 1965, MNRAS, 130, 159

Howard, A. W., et al. 2012, ApJS, 201, 15

Howell, S. B., et al. 2014, PASP, 126, 398

Huang, J., et al. 2018, ApJ, 869, L42

James, C. W. 2023, PASA, 40, e057

Johnson, J. A., et al. 2017, AJ, 154, 108

Lissauer, J. J., Rowe, J. F., Jontof-Hutter, D., Fabrycky, D. C., Ford, E. B., Ragozzine, D., Steffen, J. H., & Nizam, K. M. 2023, arXiv e-prints, arXiv:2311.00238

MacDougall, M. G., et al. 2023, AJ, 166, 33

Malhotra, R. 2015, ApJ, 808, 71

Mallonn, M., Poppenhaeger, K., Granzer, T., Weber, M., & Strassmeier, K. G. 2022, A&A, 657, A102

Mayor, M., & Queloz, D. 1995, Nature, 378, 355

McGregor, K., & Lorimer, D. R. 2024, ApJ, 961, 10

Mills, S. M., Fabrycky, D. C., Migaszewski, C., Ford, E. B., Petigura, E., & Isaacson, H. 2016, Nature, 533, 509

Mulders, G. D., Mordasini, C., Pascucci, I., Ciesla, F. J., Emsenhuber, A., & Apai, D. 2019, ApJ, 887, 157

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

Murray, C. D., & Dermott, S. F. 2000, Solar System Dynamics

Niraula, P., et al. 2017, AJ, 154, 266

Patterson, C. W. 1987, Icarus, 70, 319

Pereira, C. L., et al. 2023, A&A, 673, L4

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

Petigura, E. A., et al. 2017, AJ, 154, 107

Prieto-Arranz, J., et al. 2018, A&A, 618, A116

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

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

—. 2016, MNRAS, 459, 2275

Ricker, G. R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and
    Systems, 1, 014003

Robinson, T. D., & Reinhard, C. T. 2018, arXiv e-prints, arXiv:1804.04138

Shallue, C. J., & Vanderburg, A. 2018, AJ, 155, 94

Stassun, K. G., et al. 2018, AJ, 156, 102

Stovall, K., et al. 2014, ApJ, 791, 67

Struve, O. 1952, The Observatory, 72, 199

Thompson, S. E., et al. 2018, ApJS, 235, 38

Tremaine, S. 2023, Dynamics of Planetary Systems

Weiss, L. M., et al. 2018, AJ, 156, 254

Weiss, L. M., & Petigura, E. A. 2020, ApJ, 893, L1

Winn, J. N. 2010, in Exoplanets, ed. S. Seager, 55–77

Wisdom, J., & Holman, M. 1991, j, 102, 1528

Wolszczan, A., & Frail, D. A. 1992, Nature, 355, 145

Xu, W., & Lai, D. 2017, MNRAS, 468, 3223

Zhu, W. 2020, AJ, 159, 188