Statistics, Formation and Stability of Exoplanetary Systems

by

Ari Silburt

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Astronomy & Astrophysics
University of Toronto

Abstract

Statistics, Formation and Stability of Exoplanetary Systems

Ari Silburt
Doctor of Philosophy
Graduate Department of Astronomy & Astrophysics
University of Toronto
2017

Over the past two decades scientists have detected thousands of exoplanets, and collective properties of exoplanetary systems are now emerging. This thesis contributes to the exoplanet field by analyzing the statistics, formation and stability of exoplanetary systems.

The first part of this thesis conducts a statistical reconstruction of the radius and period distributions of *Kepler* planets. Accounting for observation and detection biases, as well as measurement errors, we calculate the occurrence of planetary systems, including the prevalence of Earth-like planets. This calculation is compared to related works, finding both similarities and differences.

Second, the formation of *Kepler* planets near mean motion resonance (MMR) is investigated. In particular, 27 *Kepler* systems near 2:1 MMR are analyzed to determine whether tides are a viable mechanism for transporting *Kepler* planets from MMR. We find that tides alone cannot transport near-resonant planets from exact 2:1 MMR to their observed locations, and other mechanisms must be invoked to explain their formation.

Third, a new hybrid integrator `HERMES` is presented, which is capable of simulating N-bodies undergoing close encounters. `HERMES` is specifically designed for planets embedded in planetesimal disks, and includes an adaptive routine for optimizing the close encounter boundary to help maintain accuracy. We find the performance of `HERMES` comparable to other popular hybrid integrators.

Fourth, the longterm stability of planetary systems is investigated using machine learning techniques. Typical studies of longterm stability require thousands of realizations to acquire statistically rigorous results, which can take weeks or months to perform. Here we find that a trained machine is capable of quickly and accurately classifying longterm planet stability.

Finally, the planetary system HD155358, consisting of two Jovian-sized planets near 2:1 MMR, is investigated using previously collected radial velocity data. New orbital parameters are derived using a Bayesian framework, and we find a high likelihood that the planets are in MMR. In addition, formation and stability constraints are placed on the HD155358 system.

"At the end of the day people won’t remember what you said or did, they will remember how you made them feel."

-Maya Angelou

Acknowledgements

This thesis would not be possible without my supervisor, Hanno Rein. Looking back to the beginning of my doctorate program, I realize now that, at the time I did not really know how to conduct quality research or even how to ask the right questions. Through your patience and hard work in answering my endless stream of questions, guiding me in fruitful research directions, and teaching me the right modes of thinking, I now feel more confident in my research abilities than I ever thought possible. You are the greatest academic mentor that I've had in my life, and I am both lucky and grateful to have worked with you.

Next, I would like to thank Dan Tamayo, who is one of the greatest humans I have met. Compassionate, intelligent and humble, you have done so much for me as well as the exoplanet community. From hosting machine learning workshops and dynamics discussion groups, to late-night collaborations for impending postdoc deadlines, to answering my endless stream of questions, you are a role model for the kind of person I would like to be.

Eric Gaidos and Yanqin Wu, you have been invaluable throughout my doctorate, and I owe you a world of thanks. Our publication together was my first real introduction to research, and it is only through your patience, intelligence and kindness that I was able to see the job through to the end. I'm grateful that you trusted in my abilities, wrote endless reference letters for me, and continually supported me throughout my doctorate.

Michael Reid you have been an incredible role model and mentor to me, having shared your amazing teaching practices through the numerous astronomy courses that I was a teaching assistant for. In addition, thank you for the reference letters, career advice, and my first introduction to the planetarium (which is now my favourite mode of science communication).

I would like to thank the Centre for Planetary Science (CPS) in Scarborough, Canada, as well as all its members, who have been an integral part of my development as a researcher. Through machine learning workshops, CPS planet days, lunchtime discussions and more, I have become a much more competent researcher than I otherwise would have ever been had CPS not been in my life.

I would like to thank the institutions and people that have funded me throughout this doctoral degree. In particular, the University of Toronto Department of Astronomy & Astrophysics for research equipment and stipends, the Natural Sciences and Engineering Research Council of Canada (NSERC) for CGS M and PGS D grants, the University of Toronto Reinhardt fund for travel reimbursement, the Walter C Sumner Memorial Fellowship Foundation, and my supervisor Hanno Rein for providing me the necessary funds and resources to complete my research.

Many colleagues at the department have made my time in Toronto very special. In particular, Max Millar-Blanchaer who has been an incredible friend and advisor, Yevgeni Kissin and Chenchong Zhu who have been my machine learning buddies over the past few years, and Lisa Esteves who I've had awesome chills, exoplanet discussions and philosophical conversations with during my time at U of T. I'd also like to thank my cohort year – APLEJIBS – who I've spent five years with in the trenches, and I'm really glad that we made the extra effort to hang out in our final year together.

My siblings have always been endlessly supportive of who I am, and especially so during my doctorate. Sam, Aviva, Hyla, Jane, thank you for not only being my family but also my friends. I feel accepted by you for who I am, regardless of my resume or list of accomplishments, and that is something I cherish dearly. To my brother Joey, thank you for the countless discussions about science, machine learning, life, etc., and all the fun times we've had together outside the academic realm (Oma's,

Age of Empires with Sam, etc.). I've cherished the years we've lived together in Toronto, and will miss you when I leave.

To my parents, now that I am older, I see just how much time, sacrifice and effort it truly was to raise six kids, including hockey and baseball practices/games, Hebrew school, music lessons, endless driving/cooking/cleaning/etc., working hard to provide for us and send us all to university – the list goes on. I am eternally grateful for what you have done for me, and love you very much. To my father, Bruce, I remember as a young child looking up at the stars with you in our backyard in Guelph, contemplating the biggest questions of the universe, and igniting my passion for science and the unknown. To my mother, Barbara, from circle time and homeschooling to your constant diligence over my scholastic career, you are responsible for my work ethic, drive, and academic success. Bruce and Barbara, this doctorate thesis is a realization of your synthesized combination of dream and determination; a flame that never burned out through all these years.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Exoplanet Observations

1.1.1 Detection Methods

Radial Velocity

A periodic wobble will be observed in the lightcurve of a star with an orbiting planet, due to their coupled motion around the common centre of mass. From the conservation of energy and angular momentum, a star's wobble (or radial velocity) due to a planetary companion can be calculated according to (Beaugé et al. 2007):

$$V_r = K \cdot [\cos(\nu + \omega) + e \cdot \cos(\omega) + \gamma] \tag{1.1}$$

where K is the semi-amplitude:

$$K = \frac{m_p \sin(i)}{m_* + m_p} \frac{2\pi a}{P\sqrt{1-e^2}} \tag{1.2}$$

a is the semi-major axis of the planet, P is the orbital period, m_p is the planet mass, m_* is the stellar mass, i is the inclination, e is the eccentricity, ν is the true anomaly, ω is the argument of periastron, and γ is the stellar drift in velocity. Since the probability of detecting a planet is related to K, radial velocity preferentially detects massive planets on short orbits.

From Equations 1.1 and 1.2, one can extract many orbital parameters of the planet including e, a, and $m_p \sin(i)$. However, successful extraction of these orbital parameters requires accurate and precise knowledge of the stellar mass, which is often difficult to obtain (e.g. Brown et al. 2011). Furthermore, only a lower limit estimate on planet mass can be calculated unless the inclination relative to Earth's line of sight is known. In addition, the radius of the planet cannot be determined from the radial velocity method, precluding information about size and density.

When multiple planets orbit a star, the radial velocity signal becomes more complicated. Assuming that the planets are well-separated such that their mutual gravitational interactions are weak, the radial velocity of the star can be approximated as the sum of planetary Keplerian orbits:

$$V_r = \gamma + \sum_{i=1}^{n_p} K_i \cdot [\cos(\nu_i + \omega_i) + e_i \cdot \cos(\omega_i)] \tag{1.3}$$

where n_p is the number of planets in the system. If however the planet-planet gravitational interactions are strong, Equation 1.3 is invalid and must be replaced with a more careful numerical treatment. Chapter 6 presents such a detailed numerical treatment, which builds off the work of Robertson et al. (2012) who used Equation 1.3 to extract the orbital parameters of the HD155358 system, which hosts two Jovian-sized planets near mean motion resonance.

Transit Method

When a planet passes (or transits) in front of its host star, a portion of the star's emitted flux F will be blocked. The fraction of blocked flux is proportional to the relative areas of the star and planet:

$$\frac{\Delta F}{F} \propto \left(\frac{r_p}{r_*}\right)^2 \tag{1.4}$$

where r_p and r_* are the planet and stellar radii, respectively. In addition, the probability of a transiting planet P_t is inversely proportional to its semi-major axis:

$$P_t = \frac{r_*}{a} \tag{1.5}$$

Thus, from Equations 1.4 and 1.5 we see that large, short-period planets are preferentially detected by the transit method.

The first transiting planet was detected by Henry et al. (1999) and independently by Charbonneau et al. (2000). Since then the transit method has become the most successful detection technique, with over 2,700 confirmed planets to date (Akeson et al. 2013). Most of these discoveries have come from the *Kepler Space Telescope* (hereafter *Kepler*). For reference, the next most successful detection technique is the radial velocity method (Section 1.1.1) with over 600 confirmed planets to date (Akeson et al. 2013).

From the transit method one can determine the period and radius of the planet. In addition, the atmospheric properties of exoplanets can also be probed (Kreidberg et al. 2014; Tsiaras et al. 2016; Stevenson et al. 2016). In total, the transit method is able to provide essential information about the composition, formation and habitability of a planet. However, the transit method has several weaknesses. For example, most orbital parameters required for dynamical studies, like eccentricity and mass, cannot be determined for individual systems via the transit method alone[1]. In addition, false positives from eclipsing binaries and background targets are a significant source of contamination (Fressin et al. 2013), and confirmation via a different method is often required. Smaller planets and/or planets belonging to multi-planet systems are much less likely to be false positives than large, single planets (Fressin et al. 2013; Lissauer et al. 2011).

1.1.2 Statistics of Kepler Planets

The *Kepler* mission is the most successful planet-finding mission to date, providing scientists for the first time a population of planets orbiting other stars. Due to the selection bias of the transit method (see Section 1.1.1), most of these planets have periods of less than 50 days and are larger than Earth. However, sufficient detection across the full range of Earth-to-Jupiter sized planets has enabled scientists to calculate the occurrence of planets around stars in our galaxy.

[1]When neighbouring planets exhibit Transit Timing Variations, information can be collected about eccentricity and mass (Holman & Murray 2005).

Two seminal works analyzing the *Kepler* population are Fressin et al. (2013) and Petigura et al. (2013). Fressin et al. (2013) found that the global false positive rate of the *Kepler* data is roughly 10%, the radius distribution peaks at mini-Neptune sized planets ($2 < r_p/r_\oplus < 2.8$, $r_\oplus$ is the radius of Earth), and the occurrence of Neptune and Jupiter sized planets is significantly lower than the occurrence of $r_p/r_\oplus < 2.8$ sized planets. Using a carefully vetted data sample, Petigura et al. (2013) found a similar result as Fressin et al. (2013), i.e. that the radius distribution peaks at mini-Neptune and the occurrence of Neptune and Jupiter sized planets is much lower.

A calculation of planet occurrence is usually performed by binning the planets in r–P space (where r is planet radius and P is planet period), and calculating planet occurrence for each bin. The occurrence $f(P_i, r_j)$ of planets in bin (i, j) for a population of planets is:

$$f(P_i, r_j) = \frac{1}{N_*} \sum_{k}^{n_p(i,j)} \left(\frac{a_k}{r_{*,k}} \frac{1}{\epsilon(i,j)} \right) \tag{1.6}$$

where N_* is the total number of stars surveyed in the sample, $n_p(i, j)$ are the number of planets falling into bin (i, j), $\frac{a_k}{r_{*,k}}$ is the geometric correction factor to account for missed non-transiting planets (inverse of Equation 1.5) and $\epsilon(i, j)$ is the detection completeness of bin (i, j) to account for missed transiting planets due to low signal-to-noise.

Of the ingredients that go into Equation 1.6 the most difficult to accurately measure is the detection completeness. Unlike Fressin et al. (2013) who used Combined Differential Photometric Precision (CDPP) values to calculate the detection completeness of known planets, Petigura et al. (2013) instead performed an injection and recovery of simulated light curves. Although the injection and recovery technique is certainly a more accurate way to calculate the detection completeness, Fressin et al. (2013) found that CDPP estimates were robust and could lead to accurate predictions.

One critical aspect of planet occurrence that Fressin et al. (2013) and Petigura et al. (2013) failed to incorporate into their calculations was the large error bars present in the *Kepler* data. The mean planetary radius error in the (Ramirez et al. 2014) *Kepler* dataset is 30%, meaning that, within 3 standard deviations a super-Earth sized planet could actually be Neptune or Earth-sized. These errors primarily stem from the fact that the radii of *Kepler* stars are not well known (Brown et al. 2011), also with a mean radius error of 30%. With such large error bars present in the *Kepler* data, ignoring them can affect the resulting radius and planet occurrence distributions. Chapter 2 performs a planet occurrence calculation similar to Fressin et al. (2013) and Petigura et al. (2013) whilst incorporating these large errors into the analysis.

1.2 Planet Formation

1.2.1 Minimum Mass Solar Nebula and Snowline

The initial conditions of the Solar System are still largely unknown. Hayashi (1981) and Weidenschilling (1977b) provided the first benchmarks for the initial mass of the Solar System by assuming a rocky core for all planets and calculating the initial surface density distribution Σ as a function of distance d required to produce the masses of the planets:

$$\Sigma(d) = \Sigma_0 \left(\frac{d}{1\mathrm{AU}} \right)^{-3/2} \tag{1.7}$$

where $\Sigma_0 = 1700\text{g/cm}^2$. This initial surface density distribution is known as the Minimum Mass Solar Nebula (MMSN), and is a benchmark against which most planet formation studies are compared to.

A second critical property of circumstellar disks is the location of the snowline. In standard formation theory the snowline marks the boundary where water ice begins to form, providing the additional solid material required to make large cores and form Jovian planets. Hayashi (1981) provided a benchmark temperature profile $T(d)$ for the Solar System:

$$T(d) = T_0 \left(\frac{d}{1\text{AU}} \right)^{-1/2} \tag{1.8}$$

where $T_0 = 280\text{K}$. In this model, the distance at which the temperature drops to below freezing is ~ 2.7 AU. More recent models of the snowline (Sasselov & Lecar 2000) include detailed radiative transfer physics and heating via accretion, and show that the primordial snow line in our Solar System could have been as close as 1 AU.

Until recently, scientists have lacked the tools to observationally test theories of early solar system formation. Telescopes like the Atacama Large Millimeter/submillimeter Array (ALMA) have begun providing observations of circumstellar disks around young stellar objects like HL Tau (ALMA Partnership et al. 2015). In total, about 80 circumstellar and debris disks have been observed to date (e.g. Schneider et al. 2014; Choquet et al. 2016), revealing unexpected features like strong debris disk asymmetries (Hines et al. 2007), concentric rings sculpted by possible Jovian-sized planets (Tamayo et al. 2015) and highly eccentric/unbound trajectories in debris disks that are difficult to reconcile with standard formation theories (Boccaletti et al. 2015).

1.2.2 Planetesimal Formation

Before planets can form there must be an abundance of kilometre-sized planetesimals. There are currently two leading models for forming such kilometre-sized planetesimals – coagulation (a.k.a. core-accretion) and gravitational instability. In the coagulation model, pairwise collisions between sticky dust particles lead to steady growth up to meter-sized and beyond (Weidenschilling 1977a; Armitage 2010). However, mean collision velocities are strongly coupled to size and destructive collisions tend to occur between meter-sized objects, hindering further planetesimal growth (Weidenschilling 1977a; Blum & Wurm 2008). In addition, meter-sized objects embedded in a protoplanetary disk will strongly couple to the gas and typically drift into the central star on a timescale of 10^3 years (Weidenschilling 1977a). Thus, a mechanism is required to quickly grow particles beyond a meter in size to avoid destruction. Solutions to this "meter-barrier problem" have been proposed, however no clear consensus has yet emerged. For example, Boley et al. (e.g. 2014) suggested that as meter-sized bodies drift closer to the central their they will become partially molten, increasing their stickiness and promoting further collisional growth.

In the gravitational instability model, as the solar nebula cools dust particles settle in the mid plane, becoming vulnerable to collapse (Goldreich & Ward 1973). These dust particles clump and eventually become gravitationally unstable, forming ∼100m sized particles (Goldreich & Ward 1973). This process then repeats, each time forming larger and larger planetesimals via gravitational instability. This model is attractive since it bypasses the scales most vulnerable to destructive collisions (i.e. the meter-barrier problem), and can form large planetesimals in $\sim 10^3$ years (Goldreich & Ward 1973; Armitage 2007).

However, in practice it is very difficult to collect dust particles in densities high enough for gravitational instability, and especially so in turbulent disks (Armitage 2007).

The formation mechanism of planetesimals is still unclear, however, *some* mechanism must exist since planetesimals are ubiquitous.

1.2.3 Formation of Protoplanets

The formation of protoplanets occurs shortly after the formation of large, kilometre-sized planetesimals. In a pioneering study, Greenberg et al. (1978) found that in the early stages of planet formation larger planetesimals grow more rapidly than smaller ones, resulting in the runaway growth of the largest planetesimal. As a result, 1–10 km sized planetesimals at 1AU can grow into 10^{22} – 10^{24} kg protoplanets in 10^5–10^6 years (Wetherill & Stewart 1989). This runaway process stems from the fact that 1–10 km sized planetesimals are large enough to gravitationally focus each other yet are still dynamically cool via damping from surrounding gas, resulting in a maximum collision cross section (Armitage 2010). Since proplanetary growth rates are related to orbital frequency, more distant regions will have slower growth timescales.

Eventually this runaway growth transitions into a slower, oligarchic growth. As protoplanets grow surrounding planetesimals are dynamically excited via gravitational stirring, decreasing the collision cross section and slowing the rate of protoplanetary growth (Kokubo & Ida 1998). By the end of oligarchic growth protoplanets have consumed most of the surrounding material and have reached their isolation masses (Schlichting 2014). Beyond the snowline the isolation mass is roughly the mass of Neptune, however inside the snowline it is a fraction of an Earth mass (Schlichting 2014). Since numerous *Kepler* planets larger than Earth reside inside the snowline of their host star, this suggests that either i) giant impacts between protoplanets, ii) inward drift of additional material, or iii) migration must have taken place (Schlichting 2014).

1.3 Planetary Dynamics

1.3.1 Mean Motion Resonance

Mean motion resonance (MMR) occurs when the orbital period of one planet is an integer ratio of another. Like other types of resonances occuring in nature, MMR results in the amplitude growth of various quantities characterizing the system like eccentricity, semi-major axis and the longitude of pericentre (Murray & Dermott 1999). As a result, the presence of MMR can strongly affect the formation, evolution and longterm stability of planetary systems in a diversity of ways. For example, Kirkwood gaps are unstable regions in the asteroid belt carved by MMRs with Jupiter, while Pluto and Neptune are protected from going unstable due to a 3:2 MMR.

For every $p : q$ MMR (where p and q are integers) there are two important resonant angles:

$$\begin{aligned} \phi_1 &= p\lambda_1 - q\lambda_2 + \varpi_1 \\ \phi_2 &= p\lambda_1 - q\lambda_2 + \varpi_2 \end{aligned} \tag{1.9}$$

where λ is the mean longitude and ϖ is the longitude of periapse. For planets to be in MMR, the time variation of one or both resonant arguments must be zero. As a result, MMRs can be modelled in terms of a pendulum oscillating about a stable, fixed point. After some algebra it can be shown that a MMR can be modelled as (Murray & Dermott 1999):

$$\ddot{\phi} = -\omega_0^2 \sin\phi \tag{1.10}$$

where ω_0 is the amplitude of libration and is dependent upon the orbital parameters of the system (mass, eccentricity, semi-major axis).

The pendulum model facilitates an understanding about certain properties of MMR. For energies larger than a critical energy, E_{crit}, the pendulum will circulate over all possible values of ϕ, while for energies smaller than E_{crit} the pendulum will be in MMR and librate about $\phi = 0$. The critical energy, E_{crit}, defines motion on the separatrix, which separates the circulation and libration regimes. In context of a pendulum, this would correspond to the pendulum suspended vertically in the air with an infinite period of libration.

The strength of a given MMR is related to its width, which in turn is related to the order of the resonance ($= p - q$) and the magnitude of p and q (Murray & Dermott 1999). More fundamentally, the strength of a MMR is related to the mass, eccentricity and mean motions of the planets involved (Murray & Dermott 1999). Stronger MMRs are associated with lower values of p, q and $p - q$, making the 2:1 and 3:2 MMRs the most probable resonant locations in nature. Figure 1.1 shows the distribution of period ratios for planets discovered by *Kepler*, along with the locations of first and second order MMRs. As can be seen, statistical excesses of planets exist wide of the 3:2 MMR and 2:1 MMR (Lissauer et al. 2011; Fabrycky et al. 2014; Steffen & Hwang 2015), and it is believed that these planets were transported away from MMR via dissipative mechanisms. The most popular of these dissipative mechanisms are tidal (Lithwick & Wu 2012; Batygin & Morbidelli 2013; Delisle et al. 2014), protoplanetary (Rein 2012; Baruteau et al. 2013; Goldreich & Schlichting 2014), and planetesimal (Moore et al. 2013; Chatterjee & Ford 2015). The formation implications for each mechanism are different, and no clear consensus has yet emerged. Chapter 3 critically examines the role of tidal forces in transporting *Kepler* planets from exact MMR to their observed locations, using analytical and numerical means.

1.3.2 Migration

Planetary migration is believed to be the most effective way of trapping planets in MMR (e.g. Lee & Peale 2002), making it a very important process for planet formation. Throughout this thesis (Chapters 3, 4 and 6) migration physics is utilized, and presented here are the two primary ways that planets can migrate.

Planetesimal-Driven Migration

Planetesimals passing through the Hill sphere of a planet will exchange angular momentum via gravity (Ida et al. 2000; Kirsh et al. 2009). If there is an asymmetry to the mass of planetesimals interacting with the planet on its near and far sides, a net force will migrate the planet. However, to guarantee a net migration planetesimal orbits must decouple from the planet. A massive enough planet (e.g. Jupiter) will directly eject and decouple planetesimals from the system, however if the planet is smaller (e.g. Neptune) planetesimals must decouple by interacting with a neighbouring planet. In addition, for sustained migration the planet must constantly encounter fresh, dynamically cold planetesimals (Gomes et al. 2004).

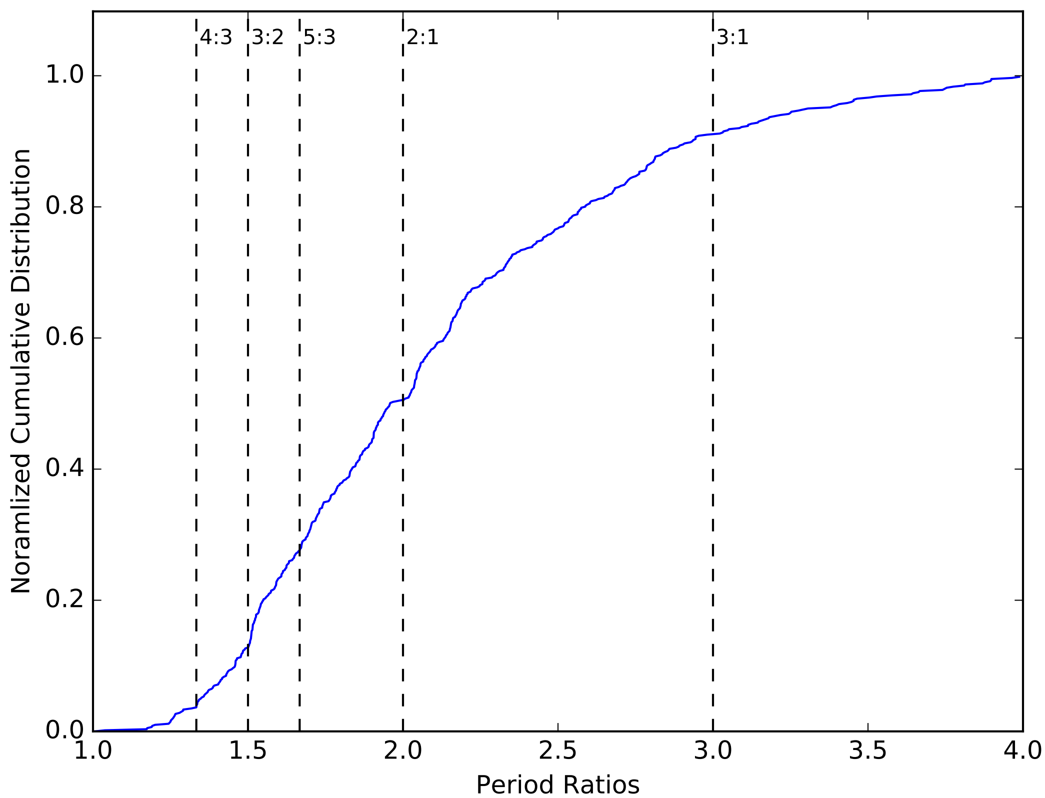

Figure 1.1: Cumulative distribution of period ratios of neighbouring planets in all known multi-planet systems. First and second-order mean motion resonances are displayed as dotted lines, and labelled at the top of the figure. Data from Akeson et al. (2013).

Since protoplanetary growth is proportional to orbital frequency (e.g. Rafikov 2003), planetesimal disks are most likely to exist in the outer reaches of planetary systems where fewer orbital cycles have occurred. In the Solar System, the primordial Kuiper belt is believed to have once been such a planetesimal disk, causing Neptune to migrate outwards into the Kuiper belt and shepherd planetesimals inwards to Jupiter, which subsequently ejected them from the Solar System (Fernandez & Ip 1984). This idea is well supported by observations of the outer Solar System, which show that Pluto, along with a host of smaller bodies, orbit in stable 3:2 MMRs with Neptune (Malhotra 1993; 1995).

Gas-Driven migration

Since the discovery of the first hot Jupiter (Mayor & Queloz 1995), gas-driven migration is believed to play an important role in shaping exoplanetary systems (Lin et al. 1996). Whenever a fully-formed planet is embedded in a protoplanetary disk, angular momentum can be exchanged via disk-planet torques (Goldreich & Tremaine 1980). The result of this exchange is planetary migration, and this scenario is believed to be common to all young planetary systems. Gas-driven migration comes in two main flavours – Type I and Type II.

Type I migration occurs when low-mass planets are fully embedded in a protoplanetary disk and do not significantly perturb the disk structure (Armitage 2010). At particular resonant locations, known as Linblad resonances, density waves are excited due to gravitational interactions between the planet and disk (Goldreich & Tremaine 1979). These density waves exchange angular momentum with the planet, and migration occurs when the inner and outer disk interact asymmetrically with the planet (Goldreich & Tremaine 1979). In general, the direction of Type-I migration tends to be inwards towards the central star (Ward 1997).

Type II migration occurs when high-mass Jovian planets significantly modify the structure of the surrounding protoplanetary disk, opening up a gap. This gap locks the planet in place, coupling the migration of the planet to the evolution of the disk (Lin & Papaloizou 1986). The viscous evolution of the disk causes the planet to slowly migrate inwards, at speeds typically one or two orders of magnitude slower than Type-I migration (Ward 1997).

In comparison to planetesimal migration, gas-driven migration is still not well understood. In particular, standard calculations of gas-driven migration are too quick by 1-2 orders of magnitude (Lin & Papaloizou 1986; Tanaka et al. 2002), causing planets to spiral into their central stars before the protoplanetary disk has dispersed. In contrast to this standard view, recent work (Fung & Chiang 2017) has suggested that planets actually do not migrate that much and tend to be better behaved than originally believed. A consensus on migration has yet to be established, but it is clear that some form of migration must occur in the universe due to the large number of planets in or near MMR (Lissauer et al. 2011; Fabrycky et al. 2014; Steffen & Hwang 2015).

1.3.3 Stability

The longterm stability of planetary orbits has been studied for hundreds of years by many famous scientists including Issac Newton, Joseph-Louis Lagrange and Carl Friederich Gauss. However, due to the chaotic and non-integrable nature of planetary systems it has been historically difficult to make progress on N-body problems. The chaos in planetary systems is caused by overlapping resonances (Chirikov 1979; Lecar et al. 2001), resulting in the divergence of near-identical systems on long timescales.

However, with the aid of computers the equations of motion governing planetary systems can be brute-force integrated into the future or past, allowing scientists to answer fundamental questions that have plagued humans for hundreds of years. For example, it is now known that the Solar System is marginally stable (Sussman & Wisdom 1988; Laskar 1994; Lecar et al. 2001), with Mercury having a 1% chance of colliding with Venus or the Sun within a couple billion years (Laskar & Gastineau 2009). It is also now well established that most known multi-planet systems are packed to capacity, and adding additional planets into these systems would result in dynamical instabilities (Fang & Margot 2013; Pu & Wu 2015).

Although most planetary systems cannot be analytically solved, constraints on these systems can still be derived from analytical means. For example, Wisdom (1980) and Duncan et al. (1989) showed that for small eccentricities in the Restricted 3-Body Problem (R3BP), chaotic orbits (leading to close encounters, collisions and ejections) will occur when the perturber and particle are separated by $\Delta a \leq 1.3\mu_p^{2/7} a_p$ (where subscript p indicates the perturber). Also associated with the R3BP is the Jacobi constant, which can be used to constrain the chaotic motion of a particle in parameter space. For two massive planets, Gladman (1993) showed that orbits are Hill stable if $\Delta a \geq 3.46 R_H$ (where R_H is the mutual Hill radius), forbidding close encounters for all time.

Since the discovery of numerous exoplanetary systems via *Kepler*, longterm stability has become a popular way to constrain orbital parameters (Lissauer et al. 2011; Steffen et al. 2013; Jontof-Hutter et al. 2014; Tamayo et al. 2015). If one assumes that an observed system is stable over billions of years, grids of N-body integrations can be used to find stable regions of parameter space, further narrowing the range of valid solutions constrained by observations. Although this brute-force method is certainly useful, it is not without its costs. A 10 billion year integration or the Solar System takes weeks to complete, and due to the chaotic nature of planetary systems hundreds to thousands of realizations must be simulated to acquire statistically rigorous results. Chapter 5 shows that using machine learning techniques, accurate predictions of longterm stability can be made. These predictions are orders of magnitude faster than direct N-body integrations.

1.4 Numerical Integration

1.4.1 Hamiltonian Dynamics

The Hamiltonian $\mathcal{H}$ encodes the kinetic and potential energy for a system of N bodies. In its most basic form, the Hamiltonian is:

$$\mathcal{H} = \sum_{i=0}^{N-1} \frac{\mathbf{p}_i^2}{2m_i} - \sum_{i=0}^{N-1} \sum_{j=i+1}^{N-1} \frac{G m_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.11}$$

where $\mathbf{r}_i$, $\mathbf{p}_i$ and m_i are the position, momentum and mass of body i, respectively. The first term in Equation 1.11 sums the kinetic energies of the system, while the second term sums the potential energies of the system.

The coordinates ($\mathbf{r}$,$\mathbf{p}$) used in Hamiltonian mechanics are canonical, obeying the fundamental Poisson bracket relations:

$$\{\mathbf{r}_i, \mathbf{r}_j\} = 0, \qquad \{\mathbf{p}_i, \mathbf{p}_j\} = 0, \qquad \{\mathbf{r}_i, \mathbf{p}_j\} = \delta_{ij} \tag{1.12}$$

where δ_{ij} is the Kronecker delta. A primary benefit of using this Hamiltonian framework is the ease of evolving a system of N-bodies into the future or past via Hamilton's equations:

$$\begin{aligned} \frac{d\mathbf{p}}{dt} &= -\frac{\partial H}{\partial \mathbf{r}} \\ \frac{d\mathbf{r}}{dt} &= \frac{\partial H}{\partial \mathbf{p}} \end{aligned} \tag{1.13}$$

Numerically this is straightward, with the accuracy of the result being inversely proportional to the size of the timestep, dt. A second benefit is the inherent area preserving or symplectic nature of Hamiltonian systems, with the energy error bound to a finite value (excluding e.g. numerical roundoff errors which grow with time).

Within the context of planetary systems, Equation 1.11 can be modified to make integration more efficient. For example, a single planet will orbit in a Keplerian fashion around a star, determined exactly by the planet's orbital parameters and stellar mass. In addition, for multi-planet systems gravitational interactions between planets will occur, perturbing planets off their original Keplerian trajectories. Thus, a new Hamiltonian can be constructed consisting of a Keplerian term, $\mathcal{H}_K$, and Interaction term, $\mathcal{H}_I$, according to (Wisdom & Holman 1991):

$$\mathcal{H} = \mathcal{H}_K + \mathcal{H}_I$$

If the planets are well-separated and the planet-planet gravitational interactions are small, the system can be numerically integrated by applying Keplerian and Interaction operators in a leapfrog manner:

$$E_{\mathcal{H}}(dt) = E_{\mathcal{H}_K}(dt/4) \cdot E_{\mathcal{H}_I}(dt/2) \cdot E_{\mathcal{H}_K}(dt/4) \tag{1.14}$$

where $E_X(Y)$ represents the evolution of the system under X for time Y. Equation 1.14 is a second order integration scheme, and is the most popular choice for simulating planetary systems (Wisdom & Holman 1991).

1.4.2 Coordinate Systems

As mentioned above, the most efficient way to solve the N-body problem is to split the Hamiltonian into Keplerian and Interaction components. In general, the Keplerian and Interaction Hamiltonians take the following form:

$$\mathcal{H}_K = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} - \frac{Gm_0m_i}{\mathbf{r}_i}, \qquad \mathcal{H}_I = \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} \frac{Gm_im_j}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.15}$$

However, there are numerous ways to perform these splits, with each coordinate system having relative strengths and weaknesses. The most popular splittings are are discussed below. In all cases $\mathbf{r}$ and $\mathbf{p}$ represent cartesian position and momenta, respectively, N is the total number of particles, and M is the total mass of the system.

Jacobi

Carl Jacobi worked out a coordinate system in which the planet positions are measured relative to the centre of mass of all bodies interior to it. Since a planet's semi-major axis (and hence position) is dependent upon the mass interior to it by Kepler's 3rd law, in this coordinate system the position of a particle is directly affected by all bodies interior to it.

The Jacobi position $\mathbf{r}'$ and momentum $\mathbf{p}'$ form a canonical set and are related to the cartesian position and momentum according to (Murray & Dermott 1999):

$$\mathbf{r}'_i = \mathbf{r}_i - \mathbf{R}_{i-1}, \qquad \mathbf{p}'_i = \frac{\eta_{i-1}}{\eta_i}\mathbf{p}_i - \frac{m_i}{\eta_i}\sum_{j=0}^{i-1}\mathbf{p}_j \tag{1.16}$$

where $\mathbf{R}_i = \frac{1}{\eta_i}\sum_{j=0}^{i} m_j\mathbf{r}_j$ is the centre of mass of all particles interior to body i and $\eta_i = \sum_{j=0}^{i} m_i$ is the sum of masses interior to body i. For the special case of $i = 0$:

$$\mathbf{r}'_0 = \mathbf{R}_N, \qquad \mathbf{p}'_0 = \sum_{j=0}^{N}\mathbf{p}_j \tag{1.17}$$

The inverse transformations from Jacobi coordinates to cartesian coordinates can be found in Chapter 9.5 of Murray & Dermott (1999).

The benefit of Jacobi coordinates is that the kinetic terms remain a sum of squares (Plummer 1918), making numerical integration a straightforward process according to Equation 1.14. In addition, these coordinates solve the Two Body and Restricted Three Body Problems exactly. However, the primary disadvantage is that a clear ordering of bodies is required for efficient integration.

Democratic Heliocentric

In Democratic Heliocentric coordinates, positions $\mathbf{Q}$ and momenta $\mathbf{P}$ form a canonical set with $\mathbf{Q}$ measured relative to the central body and $\mathbf{P}$ measuring barycentric momenta. These coordinates are related to the cartesian position and momenta according to (Duncan et al. 1998):

$$\mathbf{Q}_i = \mathbf{r}_i - \mathbf{r}_0, \qquad \mathbf{P}_i = \mathbf{p}_i - \frac{m_i}{M}\sum_{j=0}^{N}\mathbf{p}_j \tag{1.18}$$

For the special case of $i = 0$:

$$\mathbf{Q}_0 = \frac{1}{M}\sum_{j=0}^{N} m_j\mathbf{r}_j, \qquad \mathbf{P}_0 = \sum_{j=0}^{N}\mathbf{p}_j \tag{1.19}$$

The benefit of Democratic Heliocentric coordinates is that positions are not dependent upon other bodies (unlike Jacobi coordinates) and are conceptually easier to understand. However, these coordinates do not solve the Two Body Problem or Restricted Three Body Problems exactly. In addition, $\mathcal{H}_K$ cannot be cleanly written as shown in Equation 1.15 due to several cross terms that arise when substituting $\mathbf{Q}$ and $\mathbf{P}$. However, this problem can be rectified by transferring these cross terms into an additional Jump Hamiltonian:

$$\mathcal{H}_J = \frac{1}{2m_0}\left|\sum_{i=1}^{N}\mathbf{P}_i\right|^2 \tag{1.20}$$

Therefore, the evolution of $\mathcal{H}$ in Equation 1.14 must be modified to include an additional evolution operator under the Jump Hamiltonian, $E_{\mathcal{H}_J}$. Note that since $\{\mathcal{H}_J, \mathcal{H}_I\} = 0$, the ordering of $E_{\mathcal{H}_I}$ and $E_{\mathcal{H}_J}$ does not matter.

WHDS

The WHDS is a modification of the Democratic Heliocentric mapping and splits the kinetic energy slightly differently. The canonical coordinates **Q** and **P** have the same form, but the Keplerian, Interaction and Jump steps are now (Laskar & Robutel 1995; Wisdom 2006; Hernandez & Dehnen 2016):

$$\begin{aligned}
\mathcal{H}_K &= \sum_{i=1}^{N} \frac{\mathbf{P}_i^2}{2\mu_i} - \frac{G(m_0+m_i)\mu_i}{\mathbf{Q}_i} \\
\mathcal{H}_I &= \sum_{i=1}^{N} \sum_{j=1,j\neq i}^{N} \frac{G m_i m_j}{|\mathbf{Q}_i - \mathbf{Q}_j|} \\
\mathcal{H}_J &= \sum_{i=1}^{N} \sum_{j=1,j\neq i}^{N} \frac{\mathbf{P}_i \cdot \mathbf{P}_j}{m_0}
\end{aligned} \tag{1.21}$$

where $\mu \equiv m_i m_0/(m_0+m_i)$.

The benefit of these coordinates is that, unlike Democratic Heliocentric coordinates, the Two Body and Restricted Three Body Problems are now solved exactly. In addition, (like Democratic Heliocentric coordinates) a particle's position are not dependent upon any other bodies. One price to pay for these benefits is that $\{\mathcal{H}_J, \mathcal{H}_I\} \neq 0$, and in Equation 1.14 $E_{\mathcal{H}_J}$ must be exactly nestled in between the $E_{\mathcal{H}_K}$ and $E_{\mathcal{H}_I}$ operators. In addition, the standard symplectic correctors of Wisdom (2006) cannot be used in this coordinate system, however other correctors may be possible to construct.

1.4.3 Integrator Types

The three main classes of N-body integrators are symplectic, non-symplectic and hybrid. Symplectic integrators employ a fixed timestep, bounded energy error (excepting bias and roundoff errors which grow with time) and fast integration time. These integrators are best suited for well-separated bodies where $\mathcal{H}_I \ll \mathcal{H}_k$, otherwise the energy error increases dramatically or the timestep must be changed, both of which break symplecticity. The most popular symplectic integrator is the Wisdom-Holman mapping (Wisdom & Holman 1991).

Non-symplectic integrators do not require a fixed timestep and do not have a bounded energy error over time. However, these integrators can employ very accurate and precise numerical convergence schemes, for example the Burlish-Stoer or predictor-corrector algorithms (Press et al. 2002). As a result, non-symplectic integrators are typically more accurate but slower than their symplectic counterparts, and can solve most classes of planetary physics problems. A popular non-symplectic integrator is `IAS15` (Rein & Spiegel 2015).

Hybrid integrators typically mix the symplectic and non-symplectic schemes, applying the symplectic component on bodies that are distant and non-symplectic component on bodies that are close. However, there are hybrid schemes that purely rely on symplectic principles (e.g. Duncan et al. 1998). As a result, hybrid integrators are often an optimal balance between speed and accuracy, especially for problems involving numerous close encounters. The most popular hybrid integrator is `Mercury` (Chambers 1999). Chapter 4 presents `HERMES`, a new hybrid integrator designed for the migration of planets embedded in planetesimal disks.

1.5 This Work

The work presented in this thesis focuses on the statistics, formation and stability of planetary systems. Chapter 2 deals with a statistical analysis of *Kepler* planets while incorporating the large radius errors present in the *Kepler* data. As mentioned in Section 1.1.2, the analyses by Fressin et al. (2013) and Petigura et al. (2013) ignored these large errors in planetary radius, potentially affecting their calculations of planet occurrence.

Chapter 3 investigates the role of tides in transporting planets from exact MMR commensurability to a few percent wide. From Figure 1.1, an excess of planets is seen wide of the 3:2 MMR and 2:1 MMR. As mentioned in Section 1.3.1, it is believed that these planets were once closer to MMR and some mechanism transported them a few percent wide. Other studies have investigated the role of tides in transporting planets from MMR using analytical means (Lee et al. 2013; Delisle et al. 2014), but no one has yet investigated this question numerically. Although computationally expensive, N-body simulations are very accurate and avoid approximations.

Chapter 4 presents `HERMES`, a hybrid integrator for simulating planetesimal migration, built from `IAS15` (Rein & Spiegel 2015) and `WHFAST` (Rein & Tamayo 2015). An adaptive switching radius is also available for `HERMES`, optimizing the close-encounter boundary and ensuring accurate integration. In addition, `HERMES` is compared to other popular hybrid integrators. As outlined in Section 1.4.3, hybrid integrators are best for simulating planets embedded in planetesimal disks, which I plan to investigate in future works (Chapter 7.2).

Chapter 5 investigates the role of machine learning in predicting longterm planet stability. As stated in Section 6.5, due to the non-integrable and chaotic nature of multi-planet systems, classification of planetary stability is difficult without performing numerous costly N-body simulations. However, using information from the initial conditions and early evolution, we show that a trained machine learning algorithm can accurately predict whether a planetary system will be longterm stable. This tool will be useful to the scientific community due to its speed and accuracy in classifying systems over direct N-body integrations.

Finally, in Chapter 6 the HD155358 system is analyzed, which consists of two Jovian-sized planets orbiting a few percent from the 2:1 MMR. Radial velocity data for HD155358 was previously collected and analyzed by Robertson et al. (2012), but they did not account for gravitational planet-planet interactions when deriving the orbital parameters. As mentioned in Section 1.1.1, neglecting planet-planet interactions in the orbital fit is only valid when the interactions are weak. We derive new orbital parameters and find a high likelihood that the planets are in MMR. In addition, as mentioned in Section 1.3.2 planets near MMR are likely to have formed via migration, and we analyze the formation as well as the longterm stability of the system.

Chapter 2

A Statistical Reconstruction of the Planet Population around Kepler Solar-Type Stars

2.1 Chapter Overview

The work, tables, and figures from this Chapter are based off the following publication:

Ari Silburt, Eric Gaidos, Yanqin Wu
The Astrophysical Journal, Volume 799, Issue 2 - article id. 180, 12 pp., 2015
(Silburt et al. 2015).

Using the most recent catalog of (candidate) planets detected by the NASA *Kepler* mission, we reconstruct the intrinsic occurrence of Earth- to Neptune-size planets and their distributions with radius and orbital period. We analyze 76,711 solar-type ($0.8 < R_*/R_\odot < 1.2$) stars with 430 planets on 20-200 d orbits, excluding close-in planets that may have been affected by the proximity to the host star. Our analysis considers errors in planet radii and includes an "iterative simulation" technique that does not bin the data. We find a radius distribution that peaks at 2-2.8 Earth radii, with lower numbers of smaller and larger planets. These planets are uniformly distributed with logarithmic period, and the mean number of such planets per star is 0.46 ± 0.03. The occurrence is ~ 0.66 if planets interior to 20 d are included. We estimate the occurrence of Earth-size planets in the "habitable zone" as $6.4^{+3.4}_{-1.1}$%. Our results largely agree with those of Petigura et al. (2013), although we find a higher occurrence of 2.8–4 Earth-radii planets. The reasons for this excess are the inclusion of errors in planet radius, updated Huber et al. (2014) stellar parameters, and likely also the exclusion of planets which may have been affected by proximity to the host star.

2.2 Introduction

Anaximander of Miletus proposed that there were many Earth-like worlds [1]. In the twenty five centuries since the Ionian philosopher's speculation, we have seen an accelerating convergence towards an answer: many if not most stars have planets, Earth-sized (and presumably rocky) bodies are more common than Jupiter-sized gaseous bodies, and some Earth-sized planets are on orbits in the so-called "habitable zones" of their stars (Schneider et al. 2011). The most recent advances in this field have come from data collected by the *Kepler* transiting-planet mission (Borucki et al. 2010). Four years of observations of a $\sim$ 100 sq. deg. field have led to the identification thus far of 4234 confirmed or candidate transiting planets, about 84% of the candidates appear to be smaller than Neptune ($3.88R_\oplus$) (NASA Exoplanet Archive 2017).

Beyond simple human curiosity about the prevalence of Earth-like planets and possibility of life elsewhere, planet statistics provide an important test of planet formation models (e.g. Benz et al. 2014). For example, the distribution of planets with respect to orbital period and mean motion resonances test models of planet formation and early migration (e.g., Hansen & Murray 2013; Baruteau et al. 2013). Planet radius and period distributions can be combined to reconstruct the distribution of solid mass in disks (e.g., Chiang & Laughlin 2013; Raymond & Cossou 2014).

There have been many previous studies that use the *Kepler* data to infer the intrinsic population of planets around *Kepler* target stars, or subsets of those stars (e.g. Catanzarite & Shao 2011; Youdin 2011; Traub 2012; Howard et al. 2012; Fressin et al. 2013; Dressing & Charbonneau 2013; Gaidos 2013; Dong & Zhu 2013; Kopparapu 2013; Petigura et al. 2013). These works differ in their samples and methods, but are broadly consistent in estimating that the occurrence of planets per star on orbits of days to months to be of order unity. Some of these works have estimated $\eta_\oplus$, the occurence of Earth-size planets in a circumstellar "habitable zone" (where stellar irradiation is similar to the solar constant), and find that $\eta_\oplus$ is of order tens of percent. Of particular interest to us is the work of Petigura et al. (2013, hereafter PHM13) who performed an independent analysis of the *Kepler* photometric lightcurves, both identifying candidate planet transits and determining the detection efficiency by injecting synthetic transit signals into *Kepler* lightcurves and recovering them. Among the salient conclusions from PHM13 are that the distribution of planets peaks at a radius of 2–$2.8R_\oplus$ and that, $\eta_\oplus$ is about 22% (using a liberal definition of the habitable zone).

We identify two reasons to revisit the derivation of the *Kepler* planet population. Firstly, we want to more fully account for the uncertainties and biases in the *Kepler* data and related observations of the target stars. Secondly, we wish to consider a "primordial" planet population, and restrict our analysis to planets far enough from their host stars such that their properties have not been altered by proximity since their formation.

Precise statements on the occurrence of planets requires rigorous statistical methods, full accounting of errors, and adequate assessment of potential biases. First, while the overall rate of "false positives" among *Kepler* candidate planets appears to be low (Morton & Johnson 2011; Fressin et al. 2013), it is not uniform across all periods and all sizes (Santerne et al. 2013). Second, determination of planet occurrence from transit surveys requires accurate estimates of detection efficiency (also known as "completeness"), which depends on the parameters (i.e. density and/or radius) of not only the planet host stars but of the entire target catalog. The parameters of *Kepler* stars were first determined by

[1] *Dissertation on the Philosophy of Aristotle, in which his Principal Physical and Metaphysical Dogmas are Unfolded*, transl. by T. Taylor, London, 1812

combining multi-wavelength photometry, stellar models, and Bayesian inferrence (Brown et al. 2011). Colors of solar-type stars depend only weakly on gravity and metallicity, and parameter values based on photometry have large random and systematic errors in both effective temperature (Pinsonneault et al. 2012), gravity and luminosity class (Mann et al. 2012). Spectroscopy, asteroseismology, and improved stellar models have yielded more reliable parameters (Huber et al. 2014), especially for *Kepler* planet-hosting stars (*Kepler* Objects of Interest or KOIs). Nevertheless, 70% of all stars in the Huber et al. (2014) catalog have assigned parameters based on KIC photometry; the median upper and lower fractional errors in stellar radius among these solar-type stars is 40% and 10%, respectively. Because the estimated radius of a planet detected by transit depends on the stellar radius, and the probability of transit depends on stellar density, these errors need to be taken into account when computing occurrence rates. Errors in luminosity also affect the certainty with which a planet can be assigned to a habitable zone described by a range of stellar irradiance (Gaidos 2013; Mann et al. 2013).

Third, biases, if uncorrected or unaccounted for, will distort our perspective on planet populations. Mann et al. (2012) found that up to 96% of the reddest *Kepler* target stars are giants, even though virtually all red KOI hosts are dwarfs for the simple reason that it is extremely difficult to detect a transiting planet around a giant star. They showed that dilution of the target catalog by giants had led to an underestimate of the occurrence rate and an incorrect claim that M dwarf hosts of detected planets are redder and thus more metal-rich than those without detected planets.

Gaidos & Mann (2013) showed that because the *Kepler* target catalog is essentially magnitude-limited, Malmquist bias combined with uncertainties in stellar parameters means that stellar distances are underestimated and many stars are likely to be more luminous, evolved, and larger than their nominal values. Follow-up observations and analysis thus far seem to confirm this (e.g. Bastien et al. 2014; Everett et al. 2013; Verner et al. 2011). Because transiting planet radius scales with increasing stellar radius, this means that planet radii are underestimated. Moreover, the rate of planet detection decreases with increasing stellar radius or density, thus for a given planet radius, the detection rate is overestimated and thus the occurrence is underestimated. Detection bias again means that any estimate of this effect based on the host stars of transiting planets is an *underestimate*: the effect will be greater among stars in the overall target catalog. Another bias is Eddington bias: scatter by error from more populated regions of parameter space into less populated regions will produces the opposite effect, i.e. occurrence will be overestimated (Gaidos & Mann 2013).

Previous analyses of the *Kepler* planet population have sometimes not taken these errors or biases into account, and instead have considered only Poisson (counting) statistics (e.g. Petigura et al. 2013; Howard et al. 2012). Finally, many analyses were performed by binning the data into discrete bins of planet radius R_p and orbital period P. While simple and readily explicable, the binning method runs the risk of masking details of a distribution, especially that of radius, which may be important for testing theoretical models.

In this Chapter we consider a "primordial" population of planets, as opposed to one that has evolved under the influence of the host star. Effects of the latter, including tidal heating (Jackson et al. 2008), atmospheric escape (Tian et al. 2005), ohmic heating (Batygin et al. 2011), and impact erosion (Marcus et al. 2009), act with an efficiency that is inversely proportional to the distance to the host star. In particular, Owen & Wu (2013) proposed that photoevaporation by stellar X-ray and ultraviolet irradiation have effectively removed the hydrogen envelopes of close-in planets ($P \leq 10$ d), leading to the observed paucity of super-Earth sized planets in that neighbourhood. This process was also investigated for a few *Kepler* systems by Lopez et al. (2012). Regardless of the mechanism, the distinctiveness of the $P < 20$ d and $P > 20$ d populations (see, e.g. Youdin 2011) suggests that any analysis treat these separately. In this study, we focus exclusively on the latter population as we believe it is more likely to represent the "primordial" state. On the other hand, because of *Kepler*'s low efficiency at detecting long-period planets (see §2.4.1), we are forced to limit our consideration to planets with $P < 200$d.

Practical reasons also limit the range of planet radius R_p considered. Although *Kepler* can readily detect a transiting giant planet, the occurrence of these objects is indubitably much lower than that of smaller planets. The distribution with planet radius falls to a very low level beyond Neptune-size objects: only 8% of *Kepler* candidate planets have nominal radii $> 8R_\oplus$, and the false-positive rate increases as well (Santerne et al. 2012; Colón et al. 2012). Conversely, *Kepler* can detect planets smaller than $1R_\oplus$ for only a tiny fraction of stars, mostly M dwarfs. For these reasons we restrict our analysis to a radius range of 1–$4R_\oplus$ over which statistically rigorous analyses can be performed.

In this Chapter, we infer the intrinsic distribution of planets with $20 < P < 200$ d (equivalent to 0.16–0.67 AU) and $1R_\oplus < R_p < 4R_\oplus$ around solar-type stars as observed by *Kepler* over its entire mission (Quarters 1–16). This analysis includes the effects of errors in stellar and planet radius, and takes into account some of the biases that may affect previous works. We introduce a method of iterative simulation to determine the radius distribution of planets without resorting to binning. We compare our results with those of PHM13 and also carry out a detailed comparison of the two methods to understand the source of any discrepancies. Finally, we use our simulations to assess the effect of systematic bias, namely an overall underestimate of stellar radius, on inferences of a planet population from the *Kepler* catalog.

2.3 Methods

2.3.1 Catalogs of Stars and Planets

We construct a stellar sample from the Huber et al. (2014) catalog of 196,468 stars observed during the *Kepler* mission (Quarters 1–16), selecting "Sun-like" stars with radii $0.8R_\odot < R_* < 1.2R_\odot$. We restrict the sample to stars with a *Kepler* magnitude $K_p < 15.5$ to avoid faint stars with uncertain properties and noisy lightcurves. This leaves a sample of 76,711 solar-type stars, hereafter known as the "Solar76k" sample. Our planet sample is constructed from the 26 February 2014 version of the KOI catalog (Ramirez et al. 2014). Planet and stellar parameters are updated with values from Huber et al. (2014) where available. In addition to the cuts made on our stellar sample we also require $20 < P < 200$ d, $0.5R_\oplus < R_p < 6R_\oplus$ and SNR>12. Although we are only interested in the occurrence of planets $1R_\oplus < R_p < 4R_\oplus$, we use a larger radius range for our analysis since these planets have a non-zero probability of being in our region of interest after accounting for radius errors (§2.3.3). This leaves us with 430 candidate planets, hereafter known as the "430KOI" sample. We also construct a second planet sample (used only in §2.4.2) using the same cuts above except relaxing the period restriction to $5 < P < 200$. This leaves us with 1052 KOIs, hereafter known as the "1052KOI" sample.

To ease comparison with PHM13, we retrieve their vetted sample of 603 planet candidates that fall within $5 < P < 100$ d. We update the stellar parameters where possible, using Huber et al. (2014). We hereafter refer to this planet sample as the "603PHM" sample.

2.3.2 Simulated Planet Detections

Our simulator synthesizes single planet-star pairs[2], drawing stars from one of the catalogs described above, and planet parameters from a large "master" population as we describe in §2.3.3. We calculate whether each planet transits its host star in a probabilistic manner, and then determine whether *Kepler* could have detected it. We compare the properties of these simulated detections with the observed candidate *Kepler* planets, and modify the master population using two different techniques: iterative Monte Carlo Markov Chain (MCMC) and Iterative Simulation (IS). In implementing our simulations we make two critical assumptions. First, that the orbital periods and radii of *Kepler* planets (as well as orbital eccentricities), are independently distributed, i.e. the occurrence is a separable function of period and radius:

$$\frac{d^2N}{d\log P\, d\log R} = p(P)r(R)\,, \tag{2.1}$$

where p and r are some yet-to-be-determined functions, and N is the total number of planets. Second, we assume that these distributions do not vary over the range of stellar parameters considered. We discuss these assumptions in §2.6.1. In this study, we further specify that the period distribution is a power-law:

$$\frac{dN}{dlogP} = CP^{\alpha}\,, \tag{2.2}$$

where C is a normalization constant.

The geometric probability that a planet transits its host star is (Winn 2010),

$$p = \frac{R_*}{a}\frac{1+e\sin\omega}{1-e^2}, \tag{2.3}$$

where a is the semimajor axis, e the orbital eccentricity, and ω the argument of periastron. While a can be calculated from P and the estimated mass of the host star, the orbital eccentricity of *Kepler* planets are unknown and must be estimated statistically. Assuming a Rayleigh distribution with dispersion σ_e, Moorhead et al. (2011) estimated $\sigma_e = 0.2$ by studying the distribution of transit durations. This is likely affected by uncertain stellar radii and may be an overestimate. TTV studies have led to much smaller eccentricity dispersion ($\sigma_e \sim$ a few percent), at least in multiple planet systems (Wu & Lithwick 2013; Hadden & Lithwick 2014a). Here, we choose $\sigma_e = 0.18$ and show that our results are not sensitive to the exact value of σ_e (§2.6.1). The underlying distribution of ω can be safely assumed to be uniform over $[0,2\pi]$. Integrating p over the distributions of e and ω, Eq. 2.3 becomes $p = p_0 R_*/a$, with $p_0 = 1.073$. As in previous works, we require that at least three transits have been observed. We scale every transit probability by $(1/p_0)(a/R_*)_{max}$, the inverse of the max transit probability. Since transiting planets occur only for < 5 degrees of inclination, this scaling is used to speed up the rate of transiting planets. Otherwise, we would have to wait long periods of time in order acquire the large numbers of transiting planets we require to conduct this analysis.

We now proceed to assign a transit duration, T, to a given transiting planet. We follow the procedure of Gaidos (2013) by setting

$$T = \tau^{2/3}P^{1/3}\Delta\,, \tag{2.4}$$

where $\tau = 2\sqrt{R_*^3/(\pi G M_*)}$ is the stellar free-fall time, G the gravitational constant, M_* the stellar mass,

[2] We ignore the occurrence of multiple systems and assume that each planet has an independent occurrence.

and

$$\Delta = \frac{\sqrt{(1-e^2)(1-b^2)}}{1+e\cos\omega}\,, \tag{2.5}$$

with b being the impact parameter. For $a \gg R_*$, the impact parameter b is uniformly distributed in the range [0,1]. We then calculate $dN/d\Delta$, the likelihood of drawing a given Δ, or rather, its cumulative distribution, $\overline{N(\Delta)} \equiv \overline{\int_0^\Delta \frac{dN}{d\Delta'}} d\Delta'$, with the overbar indicating marginalization over e and ω. Using the chain rule, we find

$$\frac{dN}{d\Delta}d\Delta = \frac{dN}{db}\frac{db}{d\Delta}d\Delta = \frac{db}{d\Delta}d\Delta, \tag{2.6}$$

where we have used the fact that $dN/db = 1$ (i.e. b is uniformly distributed) for transiting systems. As a result, $\overline{N(\Delta)} = \overline{b} = \overline{\int_0^\Delta \frac{db}{d\Delta'}d\Delta'}$. Inverting Eqn. 2.5 then yields:

$$\overline{N(\Delta)} = \int_0^1 \eta(e)de \int_0^{2\pi} \sqrt{1 - \frac{\Delta^2\left(1+e\cos\omega\right)^2}{(1-e^2)}}d\omega, \tag{2.7}$$

where $\eta(e)$ is the assumed eccentricity distribution.

We also need to assign a radius to each trial planet: this process differs between our MCMC and IS methods and is described in their respective sections. Moreover, in comparing the radius distribution of trial planets to the observations, we must take into account significant uncertanties in the radius of KOIs. We describe how we do this in the next section.

Our detection criterion is based on a comparison between the transit signal, $(R_p/R_*)^2$, and the effective noise over the transit duration. Fressin et al. (2013) established that at signal-to-noise SNR > 12 the false-positive rate among *Kepler* KOIs is very low. PHM13 used this criterion for their analysis and we follow suit. Noise in *Kepler* lightcurves is derived from photon (shot) noise, measurement error (e.g. pointing error and instrument noise) and stellar variability (Koch et al. 2010). The *Kepler* team encapsulates the total noise of each star into quarterly transit durations of 3-hr, 6-hr and 12-hr, known as "CDPP" (Combined Differential Photometric Precision, Christiansen et al. 2012) values. For a given star in a given quarter, we generate the appropriate noise for transit duration T, by interpolating among the various CDPP values using a power-law relation. Because sources of noise (e.g., stellar variability) are not neccesarily "white", the power-law index can and often does depart from -0.5, the white noise value.

We then calculate the total SNR of a model star-planet pair as

$$SNR = \left(\frac{R_p}{R_*}\right)^2 \left[\sum_{j=1}^{16} \frac{n_j}{\left(\mathrm{CDPP}_j\right)^2}\right]^{1/2}, \tag{2.8}$$

where n_j is the number of transits in quarter j, and CDPP_j is the interpolated CDPP value for that quarter. n_j is found by first assigning a randomly drawn phase for each planet, and then counting the number of transits in each quarter. The system is proclaimed detectable if SNR > 12. The SNR threshold does not account for noise that is non-Gaussian or non-stationary on a timescale shorter than one observing quarter (90 d). However, the conservative requirement that SNR > 12 for detection partially addresses this limitation and we consider the possible impact of this simplification in §2.5 when we compare our analysis to PHM13. Other non-stationary effects unaccounted for in Eq. 2.8 include thermal settling events, sudden pixel sensitivity drop offs, and cosmic rays. Eq. 2.8 also does not account for gaps in the data.

2.3.3 Uncertainties in Planet Radii

As described in §3.2, there are significant uncertainties in the radii of most KOIs (median uncertainty = 33%), primarily due to our limited knowledge of the host star. For example, this means that there is a non-negligible chance that a planet with a cataloged radius value of $R_p = 2.5R_\oplus$ is actually Earth-sized or Neptune-sized. To clarify, we are not addressing the issue that some dwarf stars are actually giants (which is addressed in Section 2.4.3 when we consider that all stars are 25% larger). Instead, we are assuming that all claimed dwarf stars are truly dwarfs, and are accounting for the fact that the exact radii of these stars are still uncertain (on average) to $\sim$ 30%.

It is important that uncertainties of such magnitude be considered, and we do this by replacing each nominal radius by a distribution of radii governed by Bayesian statistics. The probability that a planet with a reported radius R actually has a true radius R' is given by $p(R'|R) = q(R|R')r(R')$, where $r(R')$ is a normalized prior and is the probability that a planet of radius R' (with same period P) would be detected by *Kepler* around a given star. Put another way, $r(R')$ is essentially the survey completeness (§2.4.1) of planet R' (having period P) with respect to the entire Solar76k catalog.

In our treatment, we assume that errors in R_* and hence R_p are normally distributed. This means that $q(R|R') = q(R'|R)$ because the Gaussian only depends on the square of the difference $R - R'$. We also assume that errors in stellar radius are uncorrelated. This latter assumption means that a planet with a radius that has been over/underestimated would, on average, produce a weaker/stronger transit signal among the ensemble of target stars and that such a planet would become less/more detectable. If errors in R_* were exactly correlated, then errors in R_p would be unaffected by considerations of detection; if all stars are smaller then their planets will also be smaller but by the same proportion, and thus produce transit signals of the same depth.

Provided these assumptions hold, a planet cannot be arbitrarily small, even if the errors in radius are large, because it would never have been detected in the first place. The $r(R')$ factor accounts for this fact. Our prescription for handling radius errors also accounts for the fact that the cataloged radius is more likely to be an underestimate, rather than an overestimate, of the true radius. This effect becomes most pronounced among KOIs with small cataloged radii and large uncertainty. For these cases the result is an error distribution that is no longer a Gaussian but is strongly asymmetric, with a cutoff just below the cataloged radius and an extended tail to larger radii. An example probability distribution in radius for KOI-1338.01 (solid) and KOI-1925.01 (dotted) are shown in Figure 2.1. The vertical red lines represents the catalogued radius values. As seen in Figure 2.1, the most likely radius differs from the catalog value.

We implement radius errors into our analysis by replacing each KOI with a probability distribution function (PDF) that is the product of a Gaussian times a prior detection function which is the fraction of stars around which the planet would be detected (i.e. completeness). We represent the PDF by a large number of Monte Carlo planets drawn from a Gaussian distribution with mean equal to the nominal value of R_P and standard deviation equal to the cataloged error. We calculate the fraction of stars F around which each Monte Carlo planet could be detected.

We then compute a normalized CDF of F with R_p for each Monte Carlo set. We can then draw a radius value from each corrected error distribution by comparing the CDF to a unit random deviate. We create a "master" radius distribution by randomly drawing 2 million values from all of these distributions according to their CDFs. We use this distribution to represent the inferred radius distribution of observed candidate planets after errors have been accounted for.

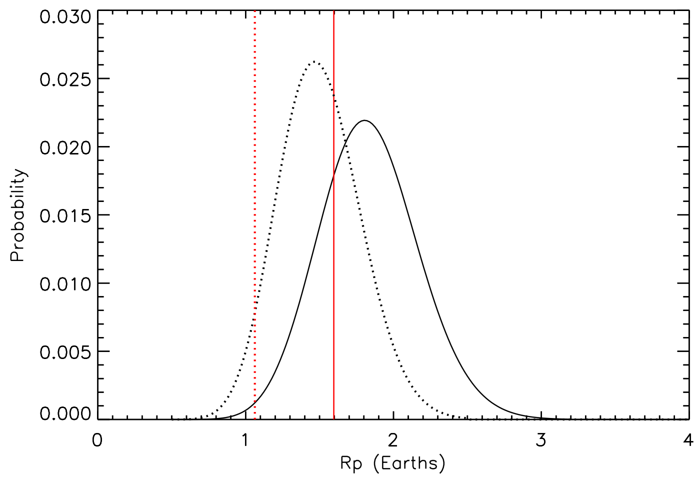

Figure 2.1: Example probability distributions in radius for KOI-1338.01 (solid) and KOI-1925.01 (dotted). The probability distributions are shown in black, while the vertical red lines represents the quoted radius values. As seen in the figure, the most likely radius value can differ from the catalogue value, and sometimes significantly so.

2.3.4 Monte Carlo Markov Chain

We implement the Monte Carlo Markov Chain according to the algorithm by Gelman & Rubin (1992). We discretize the radius-period plane into 16 bins, 4 period bins equally spaced in $\log P$ (P = [20–40, 40–80, 80–160, 160–200] d) and 4 radius bins equally spaced in $\log R$ (1–1.4$R_\oplus$, 1.4–2$R_\oplus$, 2-2-.8$R_\oplus$, 2.8–4$R_\oplus$). We parametrize the planet population by a set of 4 parameters: α (from Eq.2.2), and $\kappa_1, \kappa_2, \kappa_3$, where the latter 3 parameters are the relative numbers of planets in the first 3 radius bins to the last bin. For each set of parameters, we generate a mock catalog by simulating 10^5 transiting pairs around a given stellar sample (§2.3.2) while properly taking into account errors in planet radius (§2.3.3). We then compare our mock population against the 430KOI sample by first binning the KOIs in the same manner and then removing a portion from each bin to account for false positives (Table 1 from Fressin et al. 2013). We then scale our mock catalog down from 10^5 to match the total number of remaining KOIs. The goodness of fit is measured by comparing the simulated number of planets in each of the 16 bins, S_i, versus that of the observed, D_i,

$$\chi^2 = \sum_{i=1}^{16} \frac{(\mathrm{D}_i - \mathrm{S}_i)^2}{\mathrm{S}_i}, \tag{2.9}$$

Here, we have assumed that the error in each bin is dominated by Poisson error. Our Markov chain is run for 4000 iterations, with a "burn in" of 200 steps that are excluded from subsequent statistical analysis. A new set of parameters are accepted if $\chi_n^2 < \chi_{n-1}^2$ (also known as Gibbs sampling), where the index refers to the Markov step. If $\chi_n^2 > \chi_{n-1}^2$, the algorithm accepts the new parameter set with probability $e^{-(\chi_n^2 - \chi_{n-1}^2)/2}$. We adopt the medians of the accepted steps as the best-fit set, and we calculate both upper and lower standard errors using the 16th and 84th percentile values. The error of the $2.8 < R_p < 4$ bin is calculated from the standard deviation of $1/\kappa_3$, i.e. the relative occurrence of the 2.8-4$R_\oplus$ bin with respect to the 2-2.8$R_\oplus$ bin.

2.3.5 Iterative Simulation (IS)

We use the method of sequential Monte Carlo with bootstrap filter described in Olivier et al. (2007), which we hereafter refer to as Iterative Simulation (IS), to infer the intrinsic planet radius distribution without resorting to binning. In this technique we first generate a trial population of planets by simulating detections (§2.3.2). The radii of these simulated detections are then replaced by actual KOIs, and the process repeats until the simulated detections converge on the observations. Convergence is established when the radius distribution of our simulated detected planets (see dotted line in Figure 2.2) matches the observed 430KOI distribution. The radius distribution of the trial population then reflects that of the intrinsic population of planets. With a sufficiently large trial population, the resolution of the radius distribution is limited only by the amount of information in the observations (i.e. KOIs). In principle the periods can also be replaced by actual KOI values, but we instead choose to fix the distribution of period values using Eq. 2.2 and α, determined from our MCMC analysis. Since we find an excellent α fit to the period distribution (§2.4.2) over $20 < P < 200$, applying this distribution instead reduces additional noise in our measurement.

Our IS simulates 10^6 transiting planet-star pairs, drawing stars from the selected catalog with replacement (§2.3.2). Eccentricities and arguments of periastron of trial planets are drawn from Rayleigh and uniform distributions, respectively, and periods are drawn from a power-law with the index of the best-fit MCMC model (§2.3.4), $\alpha = -0.04$. Planet radii are initially drawn from a uniform distirbution over 0.5–$6R_\oplus$. Detections are simulated as described in §2.3.2. We randomly replace the radii of all simulated detected planets with values drawn from the "master" radius distribution (§2.3.3), after correcting for false positives using the rates in Table 1 of Fressin et al. (2013). For a discussion of our false positive treatment, see §2.6.1. We then redraw new values for all the other planet parameters besides radius and period, and reshuffle the planets among the stars. We repeat this process until acceptable convergence is achieved, usually within 100 iterations. At this point, the trial population is used to calculate the intrinsic radius distribution.

Errors are calculated by constructing 50 bootstrapped samples of the detected planet catalog. The size of each sample is a random Poisson deviate with expectation equal to the size of the actual sample. The bootstrapped samples are drawn with replacement from the actual KOI sample. Planets are randomly removed according to the false positive probabilities of Fressin et al. (2013) and then intrinsic radius distributions are calculated as in §2.3.3. For our bootstrapped samples we make the false positive correction before constructing the much larger intrinsic distribution to capture the contribution of false positives to the "noisiness" of each bootstrapped sample. We analyze each sample using the IS technique and compute standard deviations of the ensemble of bootstrapped planet populations to represent 1σ uncertainties.

Figure 2.2 shows the results of an artificial test case of the reconstruction of a planet radius distribution using the IS technique. The intrinsic distribution (dashed line) is the sum of a Gaussian plus a rising slope. The dotted line is the distribution of 364 simulated observations, which is similar in scale to our 430KOI sample. We apply the IS technique on these simulated observations to recover the actual distribution: the result is plotted as the solid line, with error bars determined from 25 bootstrapped runs. The ability of IS to reconstruct an intrinsic distribution is limited by the information available in any region of a distribution, i.e. it will fail where the number of planets or rate of detection is too low. In Fig. 2.2, errors or large uncertainties appear in the reconstructed at $R_p \sim 1R_\oplus$ where the detection efficiency is low. Also, in this simple demonstration we ignore the effect of planet radius errors. Adding planet errors tends to broaden and smooth features.

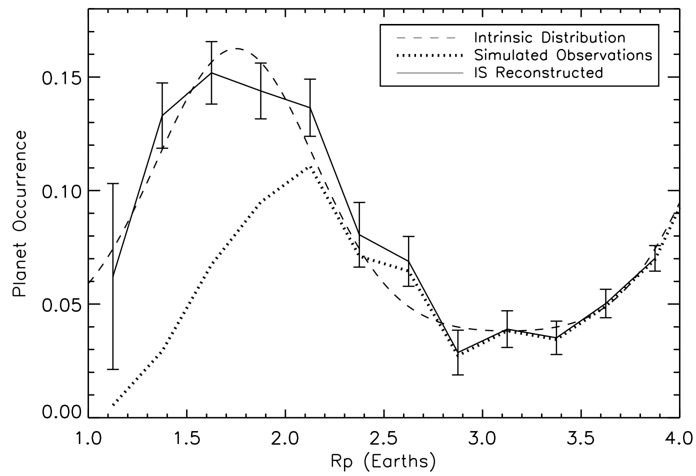

Figure 2.2: A test case demonstating the recovery of a known, artificial radius distribution (dashed line) using the method of iterative simulation. The dotted line represents the simulated observations, which consist of 364 detections, and the solid line is the reconstructed distribution (here binned for display). Error bars are determined from 25 bootstrapped simulations.

2.3.6 Calculation of Occurrence

After obtaining best-fit distributions from our MCMC and IS analyses, we calculate the rate of planet occurrence as a function of P and R_p. We generate 10^6 mock star-planet pairs and the corresponding simulated detections (§2.3.2). These planets are binned in a logarithmic grid of P (index i) and R_p (index j). The occurrence $f(i,j)$ in the bin (i,j) is then

$$f(i,j) = \frac{K_{ij}}{N_*}\frac{N_{*,S}}{S}, \tag{2.10}$$

where K_{ij} is the false positive-corrected (Table 1, Fressin et al. 2013) number of KOIs falling into the bin (i,j), N_* (=76,711) is the total number of *Kepler* target stars in the Solar76k sample, $N_{*,S}(=10^6)$ is the total number of mock pairs, and S is the number of simulated detections. The ratio $S/N_{*,S}$, the fraction of mock pairs that should be detected, and is the product of the geometric factor R_*/a as well as the detection completeness in bin (i,j), hereafter known as $C(i,j)$ (see §2.4.1). We then sum over all bins to obtain the total occurrence, f.

To demonstrate the importance of accounting for radius errors (§2.3.3) when calculating occurrence, we also conduct a separate analysis which excludes radius errors. Observed planets are binned as above and we calculate the occurrence of each bin, f according to:

$$f(P_i, R_i) = \frac{1}{N_*}\sum_{k}^{n_p(i,j)}\left(\frac{a_k}{R_{*,k}}\frac{1}{C(i,j)}\right) \tag{2.11}$$

Where $C(i,j)$ is the average completeness of the bin (see §2.4.1), $n_p(i,j)$ is the number of planets in bin (i,j), N_* (=76,711) is the total number of stars in the sample and $a_k/R_{*,k}$ is the geometric correction

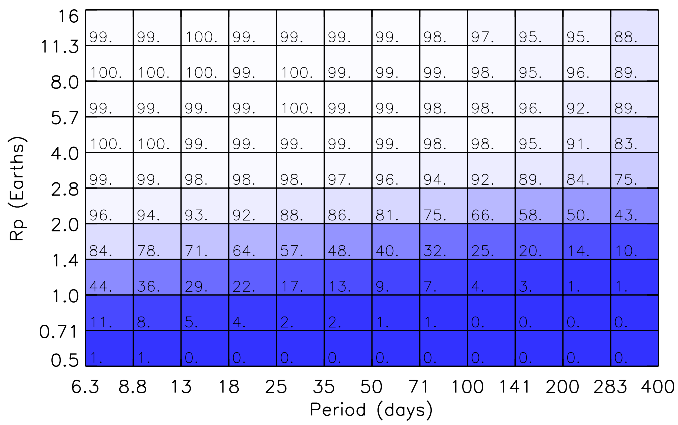

Figure 2.3: Completeness values for *Kepler* planet detection around stars in the Solar76k catalog. The numbers within each grid cell indicate the completeness percentage, and each grid has been colour coded from low (blue) to high (white) completeness.

factor for planet k.

2.4 The Primordial Population of *Kepler* Planets

2.4.1 Completeness

For a transit survey, completeness is the fraction of transiting planets of a given P and R_p that are actually detected, i.e. not including the geometric transit probability. Accurately capturing the dependence of completeness on R_p and P is crucial to a robust determination of planet occurrence.

We emphasize that survey completeness depends not only on the properties of the planet host stars, but also on the stellar and noise properties of the entire catalog. We calculate the completeness $C(i, j)$ of bin (i, j) by inserting 2×10^4 planets randomly distributed around the Solar76k sample of stars. The fraction of "detected" planets, modulo the transit probability factor, yields the completeness in this bin. The results are displayed in Figure 2.3.

Figure 2.3 shows that *Kepler* completeness is nearly 100% for planets larger than Neptune ($3.8R_\oplus$) and for nearly the full range of periods shown here. This falls rapidly beyond $P \sim 500$ d (not shown) since some systems no longer have the required three transits during the four-year *Kepler* mission. The completeness drops rapidly with decreasing radius where $R < 2R_\oplus$: Earth-sized planets are readily detected by *Kepler* only if they have orbital periods of a few days, and beyond $P = 200$ d, even the completeness of $2R_\oplus$ planets falls below 50%. For these reasons we restrict our analysis to $P < 200$d. As true of any completeness study, we also note that our completeness calculations are based on the SNR criterion as stated. For example, if the *Kepler* pipeline was uniformly missing 25% of all transiting planets (regardless of the SNR), our completeness calculations would not account for these missing planets.

2.4.2 Period Distribution

Our MCMC study yields a best-fit distribution for the 430KOI sample of $\alpha = -0.04 \pm 0.09$ (see Eq. 2.2), with a reduced chi-squared χ^2_ν of 1.07. Our value of α is consistent with zero (a flat logarithmic distribution) within errors, confirming previous determinations (e.g., Youdin 2011; Howard et al. 2012; Petigura et al. 2013; Fressin et al. 2013).

Figure 2.4 compares this best-fit period distribution with the observed sample, using the 4 period bins from Section 2.3.4 as well as an additional 2 bins to include to include planets inward of 20 d, i.e., the 1052KOI sample (§2.3.1). Inside of $P = 20$ d, *Kepler* planets deviate from a simple power-law distribution (also see Youdin 2011; Howard et al. 2012).

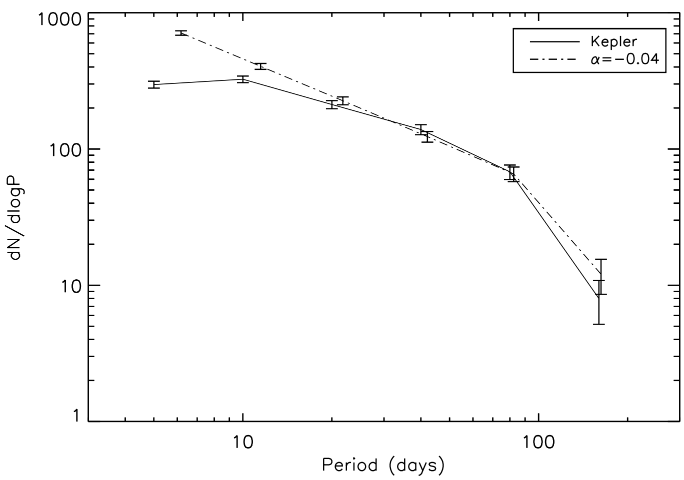

Figure 2.4: Period distribution of *Kepler* small planets. The observed distribution (solid line) includes planets inward of 20 d (i.e. the 1052KOI sample), while the simulated distribution from the MCMC best fit ($\alpha = -0.04$, see Eq. 2.2) is plotted as a dashed-dotted curve. This best fit is obtained for planets in the $20 < P < 200$ d range, but is extended here to shorter periods to demonstrate that the observed population deviates significantly from a single power-law shortward of 20 d. The location of each data point corresponds to the lowest period value in the bin, e.g. the first data point is $5 < P < 10$ d. Slight horizontal offsets have been applied to each curve for clarity.

2.4.3 Radius Distribution

Figure 2.5 displays our best-fit radius distributions for both the MCMC (solid, black) and IS (dotted, red) techniques, where we have binned the IS result for ease of comparison. Both the IS and MCMC distributions peak at 2–2.8$R_\oplus$ and decrease towards smaller radii. We have also plotted two additional distributions in Figure 2.5, a "No Error" case (dashed, green) constructed from Eq. 2.11 and a "25% Larger" IS case (dashed-dotted, blue) where it is assumed that both planet and stellar radii are 25% larger than their catalog values.

As we discuss in §2.3.3, errors in the radius of candidate *Kepler* planets, primarily due to uncertainties in stellar radius, are large and detection bias against small planets means that a planet's cataloged radius is likely an underestimate. Comparing our IS and MCMC results (which include radius errors) with our "No Error" case (which doesn't include radius errors) we see that the latter exhibits a significant excess of 1–1.4$R_\oplus$ planets. This is as expected. Correcting for radius error in a Bayesian way (§2.3.3) tends to promote small planets to larger size bins, and de-populates the smallest radius bin. However,

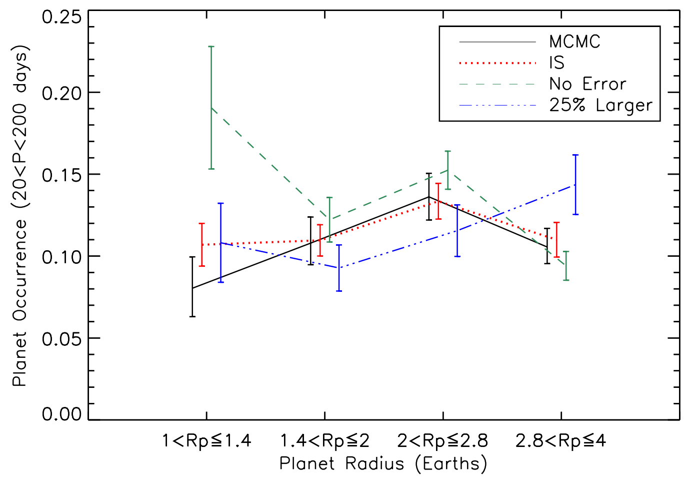

Figure 2.5: Size distribution of planets between 1 and $4R_\oplus$, obtained using the MCMC (solid black line) and the IS (dotted red line) techniques. Both show that planet occurrence peaks in the bin 2–$2.8R_\oplus$. Earth-sized planets are less common, though the statistical significance of this result is still low. If we assume that the currently determined planet radii carry no uncertainty, or that all stars (and hence planets) have 25% larger radii than their cataloged values, we obtain rather different radius distributions. The error bars for the "No Error" case account for poisson error only, while for the IS and 25% Larger cases, error bars are calculated from 50 bootstrapped simulations of the data (see §2.3.3). The MCMC error bars are calculated in the standard manner. Planet occurrence at each logarithmic radius bin is obtained by summing over all period bins. Slight horizontal offsets have been applied to each curve for clarity.

planet occurrence of the two largest bins does not increase significantly since the survey completeness is substantially higher in these bins compared to the 1–$1.4R_\oplus$ bin. We conclude that not accounting for this detection bias on radius leads to an erroneously high (by a factor of ~ 2) value of occurrence for the 1–$1.4R_\oplus$ bin.

If we assume the extreme scenario that the true radii of all stars (and therefore their planets) are 25% larger than the KOI values ("25% Larger" case in Fig. 2.5), as is shown to be the case for at least a subset of the *Kepler* stars (§2.6.1), we observe that the 2–$2.8R_\oplus$ peak is now shifted to 2.8–$4R_\oplus$. However, we do not see a significant change in the bin 1–$1.4R_\oplus$ because the depopulation of this region (due to increased planet radii) is roughly balanced by a decrease in completeness as the stars have also become larger. See §2.6.1 for a more detailed discussion.

Our treatment for the radius error is far from perfect. Some genuinely small planets may, under our procedure, be wrongly inferred to have larger radii. A better treatment will require improved errors in stellar/planet radius (see, e.g. §2.6.3).

There is a small and statistically insignificant discrepancy between the MCMC and IS results at the smallest bin ($1 < R_p < 1.4R_\oplus$). Since the two methods use identical input catalogs (§2.3.1) and detection algorithms (§2.3.2), the difference could be due to the intrinsic binning in the MCMC method.

We test the effect of different bin sizes on our MCMC results by varying the bin sizes in both radius and period space. Using the logarithmic binning scheme to describe the width, w, of each bin:

$$w(n) = 10^{Kn} \tag{2.12}$$

where n is the bin number and K is a constant, we investigate the effect of varying K (and thus bin

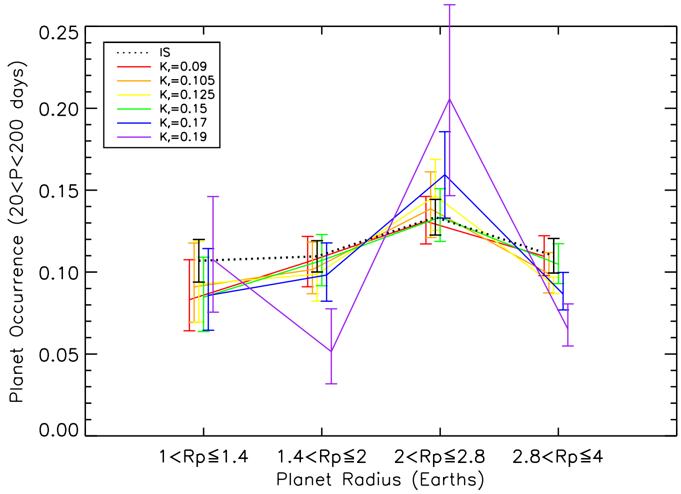

Figure 2.6: Investigating the effect of binning on the MCMC results. The IS result from Figure 2.5 is shown as a dotted, black line, while our MCMC results using various bin sizes are the coloured curves. Using Equation 2.12, we vary K_r from 0.09 to 0.19, keeping K_p fixed. As can be seen, the results can change noticeably depending on the binning choice. This illustrates the usefulness of the IS method, which does not require any binning. To clarify, we do not change the number of free parameters in our MCMC analysis (i.e. $\kappa_1, \kappa_2, \kappa_3, \alpha$), merely the size of the bins used to calculate χ^2 values. Horizontal offsets have been added for clarity.

size) on occurrence. When varying period bin size, K_p, and keeping the radius bin size, K_r, constant we find no effect on occurrence, even when K_p is varied by a factor of 2. However, when K_p is kept constant and K_r is varied, we do find an affect on occurrence rates shown in Figure 2.6 (where our IS result from Figure 2.5 is plotted as a dotted black line). As K is increased (larger bins), we see that the occurrence of $2 < R_p < 2.8 R_\oplus$ planets increases while the occurrence of $2.8 < R_p < 4 R_\oplus$ planets decreases, becoming significantly different from our IS result for extreme cases. In contrast, there is no statistically significant difference in the smallest bin. We find that the error bars increase with more extreme bin sizes, indicating that the MCMC algorithm has a harder time converging. The default bin sizes for the MCMC result in Fig. 2.5 are K_p=0.301 and K_r=0.150515.

Lastly, we display the IS radius distribution for a smaller logarithmic bin size in Figure 2.7. As explained in §2.3.5, since the IS technique requires no binning, the resolution of the result is limited only by the data and its errors. This improved resolution can reveal finer detail about the intrinsic radius distribution. We observe a slight excess of planets in the now smallest bin (1–1.15$R_\oplus$), over that in larger bins, though improved statistics are required to confirm this upward turn. We expect that, with its independence on binning, the IS technique will become central to future analysis.

2.4.4 Total Occurrence of Small Planets

In Table 2.1 we report our estimates for the total planet occurrence within $20 < P < 200$ d and $1 < R_P < 4 R_\oplus$, for the four curves in Figure 2.5. There is excellent agreement between the MCMC and the IS results. Even cases with different assumption about the radius error yield statistically consistent occurrence rates.

The total occurrence rate we calculate here is defined to be the average number of planets per star (Youdin 2011; Fressin et al. 2013; Petigura et al. 2013). Such a definition ignores the complication that

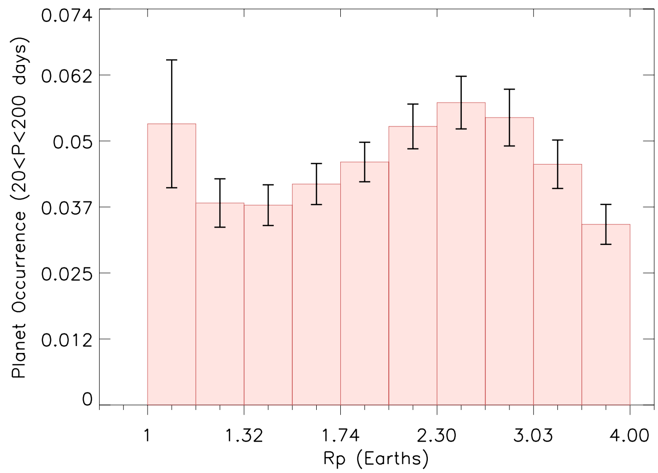

Figure 2.7: Our IS distribution displayed for a smaller logarithmic bin size. This finer resolution reveals more information about the intrinsic distribution, and specifically we see a potential rise in the number of 1–1.15$R_\oplus$ planets. This bin has large error however, and thus more statistics are required to confirm this conclusion.

many of the *Kepler* systems are multiple systems (e.g. Lissauer et al. 2011). A quantity perhaps more relevant for studies of planet formation is the occurrence rate of planetary systems, or the average number of planetary systems a star has. However, this requires knowledge of the system architecture, a task not yet attempted.

Lastly, if we include small planets on orbits interior to 20 d, the total occurrence rate is raised to $\sim 66\%$. It will rise by another $\sim 15\%$ if we include planets larger than $4R_\oplus$.

Table 2.1: Planet Occurrence for $20 < P < 200$ d, $1 < R_p < 4R_\oplus$.

Technique	Planet Occurrence (%)
MCMC	43 ± 3
IS	46 ± 3
No Error	56 ± 10
25% Larger Stars	46 ± 3

Eta-Earth

We estimate $\eta_\oplus$, the occurrence of Earth-like planets in the "habitable zone" of solar-type stars. By Earth-like, we are referring to planets between 1–2$R_\oplus$. We adopt the inner and outer boundaries of the habitable zone to be those calculated by 1-D, cloud-free, climate models (Kasting et al. 1993; Kopparapu et al. 2013). For a Sun-like star, these boundaries lie at 0.99 and 1.70 AU respectively (or orbital periods of 350 and 810 days); for other stellar spectral types, the boundaries are as tabulated in Kopparapu et al. (2013).

Such a habitable zone, however, lies outside the 200 d limit of our study. At these very long periods, the low detection efficiency of *Kepler* engenders inaccuracies in estimating $\eta_\oplus$. So instead, we have opted to calculate $\eta_\oplus$ by extrapolation, according to:

$$\eta_\oplus = \frac{f}{N_*}\sum_{i=1}^{N_*} h_i \tag{2.13}$$

where N_* (=76,711) is the number of stars in the Solar76k sample and h_i is the relative occurrence of planets (per star) within the habitable zone to some reference zone having absolute occurrence f. We choose this reference zone to be our standard $20 < P < 200$d, $1 < R_p < 4R_\oplus$ bound, with $f = 0.46 \pm 0.03$ (IS value). We calculate h_i for each star by adopting the IS radius distribution (Figure 2.5) and integrating the MCMC best-fit power-law (Eq. 2.2 with $\alpha = -0.04$) over each star's habitable limits (Kopparapu et al. 2013). Finally, we obtain

$$\eta_\oplus = 6.4^{+3.4}_{-1.1}\% \tag{2.14}$$

The error is calculated from error propagation of the IS radius distribution, occurrence of our reference zone, habitable zone limits (based on R_*, M_*, and T_*) and α. This value is consistent within errors with the analysis done by PHM13 for the same Kopparapu et al. (2013) limits, 8.6%. The reader is reminded that our calculation of $\eta_\oplus$ is an extrapolation, and depends crucially on the assumptions made.

2.5 Comparison with PHM13

We now compare PHM13 - an analysis which is similar to ours in terms of scope but which obtains their results of the *Kepler* data using the TERRA pipeline (Petigura et al. 2013). The TERRA pipeline is an analysis tool independent of the *Kepler* project pipeline and its products on which our work relies.

We first compare our estimates of detection completeness C with that of PHM13. For this completeness comparison, we re-compute C using the Best42k stars from PHM13 and compare these results to the values in Figure S11 from PHM13. We calculate the fractional difference ($2(T-P)/(T+P)$, where T=This Work and P=PHM13) and display as percentages in Figure 2.8. With the exception of a single cell all values of C in the range P = 20–200 d, and R_p= 1–4$R_\oplus$ are within 20% of PHM13 values. This shows that even a comparatively simple description of *Kepler* planet detection can account for most of the statistics. The single exception is for the 1–1.4$R_\oplus$ and 71–100 d bin where our estimate of C is 32% lower than that of PHM13. This bin includes the 90 d roll-period of *Kepler*. Large systematics appear in raw *Kepler* lightcurves at this period because the stars change positions on the detector array. It might be expected that the planets with P near 90 d would be more difficult to detect than our naive criteria and that actual completeness would be lower. If the PHM13 values are more realistic, then the opposite appears to be the case. Elsewhere in P-R_p space our completeness values are slightly and systematically higher than those of PHM13, and the discrepancy increases with increasing P and decreasing R_p. This is to be expected because PHM13 determine detection efficiency using actual lightcurves rather than representations of noise a la CDPP values. At $P > 200$ d our values of C become significantly higher than PHM13 for nearly all values of R_p. This discrepancy motivates our restriction to $P < 200$ d. One possible explanation for this difference is that detection of signals by phase-folding in the *Kepler* detection pipeline becomes inefficient at long periods.

We note that PHM13 calculates completeness based on a finite number (4×10^4) of systems of which very few detections are in the Earth-sized bins, leading to large counting (Poisson) error in

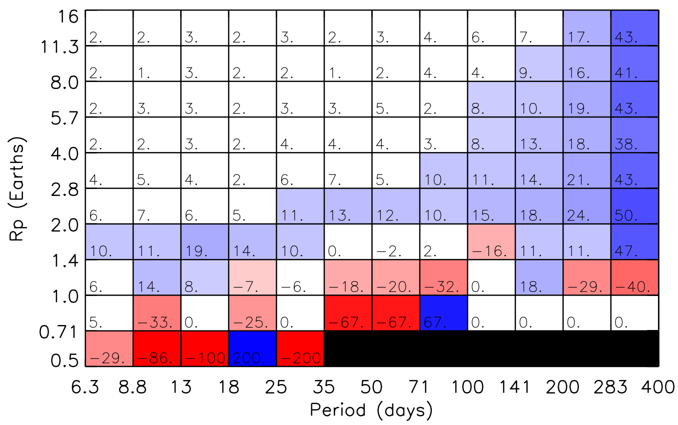

Figure 2.8: Comparison of completeness values computed with our methods to those reported by PHM13 in Figure S11, expressed as a fractional difference (i.e. in percentages). Our values are generated using the "Best42k" catalog from PHM13 and our detection criteria. Bins where both our completeness and those of PHM13 are zero have been blacked out.

completeness values. Since we use a different approach, simulating large numbers (2×10^6 total, 2×10^4 per bin) of test planets and giving high drawing probability to Earth-sized planets, we are able to simulate a much larger number of detections for each bin and thus have a more precise (although not necessarily more accurate) value of completeness.

We next compute the impact of these differences in completeness on occurrence over P=5–100 d, shown in Figure 2.9. We first calculated occurrence using the "603PHM" dataset (§2.3.1), Eqn. 2.11 along with our own detection criteria (§2.3.2) and completeness values (calculated from the Best42k sample), shown as the solid black curve in Figure 2.9. We then re-calculated planet occurrence using the 603PHM dataset, Eqn. 2.11 along with our own detection criteria but substituting in the completeness values from Figure S11 of PHM13 for ours, shown as the dashed green line in Fig. 2.9. Lastly, the direct results of PHM13 are shown as a dotted, red line. The differences between all curves in Figure 2.9 are small and, within errors, agree with each other. The residual differences in completeness seen in Fig. 2.8 do not appear to play a significant role in the comparative occurrence of our work and that of PHM13 but could be responsible for some of the minor (and statistically insignificant) differences that we find. *We conclude that simple detection criteria and noise model can be used in planet occurrence studies to achieve accurate and precise results.*

Comparing PHM13's intrinsic radius distribution (red curve in Figure 2.9) to our own (black curve in Figure 2.5), we notice a difference in the occurrence of large ($2 < R_p < 4$) planets. Specifically, PHM13 calculated an occurrence of $18.5 \pm 1.5\%$ and $6 \pm 0.5\%$ for $2 < R_p < 2.8$ and $2.8 < R_p < 4$ planets, respectively, while we find an occurrence of $13.0 \pm 1.1\%$ and $10.5 \pm 1.0\%$ for the same bins. This results in a statistical difference of about 2% and 3% between works for the $2 < R_p < 2.8$ and $2.8 < R_p < 4$ bins, respectively. We see a few possible reasons for this. Table 2.2 displays the major differences in raw samples between our work and PHM13, displaying the raw occurrence of large ($2 < R_p < 4R_{\oplus}$) planets for their nominal period ranges, $P > 20$ d and $P < 20$ d. The bottom two rows of "This Work" are empty in Table 2.2 since all of the planets in our sample are $P > 20$, and thus all the

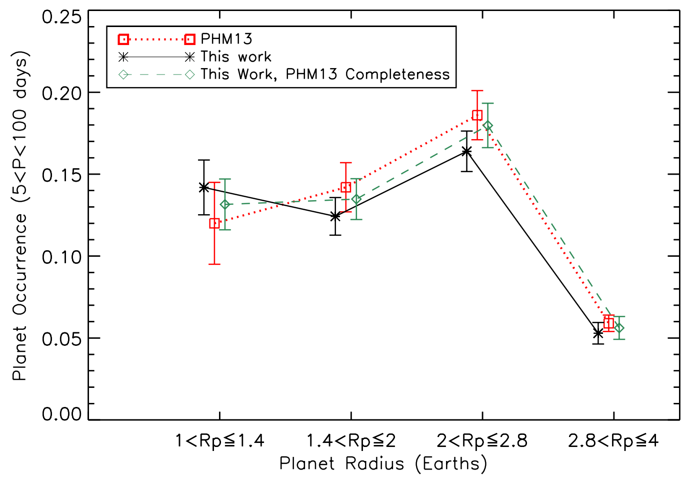

Figure 2.9: A summary of our comparison to PHM13, plotted in percentages along the y-axis. All curves are constructed using the "603PHM" dataset and "Best42k" sample. The results of PHM13 are plotted as the dotted curve in red with squares, "This work" uses our own completeness values, and is the black solid curve in black with stars, "This work, PHM13 Completeness" uses PHM13's Figure S11 completeness values and is the green dashed curve with diamonds. Slight horizontal offsets have been applied to each curve for clarity.

relevant information is present in the first row.

Table 2.2: Differences in raw counts between our work and PHM13

Range	This Work	PHM13
Nominal P range:		
$1 < R_p < 4R_\oplus$ (Total)	394	495
$2 < R_p < 2.8R_\oplus$	166 (42%)	191 (39%)
$2.8 < R_p < 4R_\oplus$	120 (30%)	87 (18%)
$P > 20$ d:		
$1 < R_p < 4R_\oplus$ (Total)	–	201
$2 < R_p < 2.8R_\oplus$	–	89 (44%)
$2.8 < R_p < 4R_\oplus$	–	45 (22%)
$P < 20$ d:		
$1 < R_p < 4R_\oplus$ (Total)	–	294
$2 < R_p < 2.8R_\oplus$	–	102 (35%)
$2.8 < R_p < 4R_\oplus$	–	42 (14%)

To summarize Table 2.2, of the 394 KOIs in our raw planet sample between $1 < R_p < 4$, 166 (42%) and 120 (30%) are between $2 < R_p < 2.8R_\oplus$ and $2.8 < R_p < 4R_\oplus$, respectively, while for PHM13, of the 495 KOIs between $1 < R_p < 4R_\oplus$ only 191 (39%) only 87 (18%) fall into the same limits, respectively. This is a significantly lower fraction, and it appears the two samples are markedly different. There are only a couple options available to explain this difference – either a disproportionate number of large ($2 < R_p < 4R_\oplus$) planets in our sample are false positives, or the two datasets are not subsamples from the same population.

If the two samples are from different populations, this difference could be due to the photoevaporation (Owen & Wu 2013) of PHM13's close-in planets, which acts to convert large ($2 < R_p < 4R_\oplus$) planets into smaller ones. Analyzing he PHM13 dataset adds support to this theory, since as we move from $P > 20$ d to $P < 20$ d planets, raw occurrence drops proportionally (i.e. relative to the entire $1 < R_p < 4R_\oplus$ sample) by $(44\% - 35\%)/44\% \sim 20\%$ and $(22\% - 14\%)/22\% \sim 40\%$ for $2 < R_p < 2.8R_\oplus$ and $2.8 < R_p < 4R_\oplus$ planets, respectively.

In explaining the difference between the occurrence of large planets between this work and PHM13, one must also consider the improved treatment of radius errors and updated Huber et al. (2014) stellar parameters used in this work. Both tend to increase planet size, pushing smaller planets into larger bins. Comparing the "No error" case to the "IS" case in Fig. 2.5, we can see that the former effect increases the occurrence of $2 < R_p < 2.8R_\oplus$ and $2.8 < R_p < 4R_\oplus$ planets by about 2% and 1.5%, respectively. Thus, it appears that the discrepancy of our results with the PHM13 in the $2 < R_p < 2.8R_\oplus$ bin can be explained by our incorporation of radius errors. However this does not account for the total discrepancy in the $2.8 < R_p < 4R_\oplus$ bin, leaving about 1.5% of discrepancy between our work and PHM13.

Finally, PHM13 reported an occurrence of $37 \pm 3.4\%$ for planets with $25 < P < 200$ d and $1 < R_p < 4R_\oplus$. Including planets from $20 < P < 25$ d raises this value to $42 \pm 3.6\%$. So the overall occurrence rates are consistent among studies that are based on different detection criteria and different model assumptions.

2.6 Discussion

2.6.1 Sensitivities and Systematics

To investigate sensitivities to some of the assumptions and parameters in our analysis, we repeat our IS simulations varying σ_e and α. First, we varied the Rayleigh distribution of orbital eccentricities between 0.1 and 0.3, and second varied the value of α between -0.15 and 0.15. In all cases we found no significant difference in the radius distributions. We also investigate our separability assumption (Eqn. 2.1) by splitting our 430KOI dataset into two equal sized subsets corresponding to $20 < P < 50$ and $50 < P < 200$ d planets, and performing an MCMC analysis on each. The resulting distributions are consistent with each other as well as with our main IS and MCMC results, indicating that there is no significant correlation between planet radius and period in our sample, and that Eqn. 2.1 is a reasonable assumption.

In this analysis, we address the fact that many *Kepler* planets have large errors in radius, driven primarily by uncertainties in the radii of their host stars. But we have not addressed the issue of systematic errors in stellar radii. For example, stellar effective temperatures based on photometry from KIC (Brown et al. 2011) are systematically ~200 K hotter than more reliable estimates based on the infrared flux method (Pinsonneault et al. 2012) or spectroscopy (Gaidos 2013). The combination of uncertainties in stellar parameters and Malmquist bias in the magnitude-limited *Kepler* target catalog means that the sample is biased towards the most luminous, hottest, and largest stars (Gaidos & Mann 2013). There is increasing evidence that many *Kepler* target stars, including planet hosts, are subgiants (e.g., Verner et al. 2011; Everett et al. 2013; Bastien et al. 2014). For fixed values of R_p/R_*, systematically larger stellar radii means the planets are also systematically larger and that the geometric transit probability is higher than presumed (transit probability depends inversely on stellar density and hotter, more evolved stars are less dense). The detection completeness of small planets is also smaller than presumed and thus, for fixed number of detections, the occurrence is higher.

We have explored the possible impact of these effects by assuming that all stellar radii in our Solar76k catalog, as well as all the planet radii in the corresponding 450KOI catalog, are 25% larger than their nominal values (dashed-dotted blue curve, Fig.2.5). The distribution differs markedly from our IS and MCMC distributions. The peak in the distribution at 2–2.8$R_\oplus$ has shifted towards larger radii. Surprisingly, the occurrence of planets with R_p = 1–1.4$R_\oplus$ has not changed. This is understood. As stellar radii increase, completeness of small planets decrease leading to an increase in planet occurrence. However, the number of 1–1.4$R_\oplus$ planets in our sample also decreases (from 25 to 8 after a 25% radius increase), reducing the raw planet occurrence in this radius bin. It appears that these two competing effects roughly cancel, resulting in no significant change in planet occurrence of the smallest radius bin.

We now comment on our treatment of false positives in this analysis. Our work makes use of the Fressin et al. (2013) false positive rates based on the Q1–Q6 *Kepler* data, while other works (e.g. PHM13) use custom methods to detect false positives. However since we use the latest disposition of the *Kepler* catalog in this analysis, the occurrence of false positives in our Solar76k sample could be significantly different. We calculate the ratio of false positives ("FP") vetted by the *Kepler* science team to planet candidates ("CAN") for the Q1–12, Q1–16 and cumulative *Kepler* catalogs according to FP/(FP+CAN). In addition, we organize these false positive ratios by radius, using the same radius bins as our analysis. These false positive ratios are shown in Table 2.3, as well as the Fressin et al. (2013) values for reference. It should be noted that the fraction of planets yet to be dispositioned (calculated according to DISP/(DISP+CAND+FP), where DISP is the number of planets yet to be dispositioned) in the Q1–12 and Q1–16 datasets are quite high (over 50 %), and may reflect the significant difference between their false positive rates and the cumulative KOI dataset rates.

Table 2.3: False positive ratios, calculated according to FP/(FP+CAND)

Dataset	FP→[1-1.4, 1.4-2, 2-2.8, 2.8-4]$R_\oplus$
Fressin et al. (2013)	0.088, 0.088, 0.067, 0.067
Q1–12	0.37, 0.39, 0.27, 0.24
Q1–16	0.28, 0.31, 0.36, 0.53
Cumulative	0.22, 0.24, 0.16, 0.19

We use the false positive fractions in table 2.3 to estimate the uncertainty of using Fressin et al. (2013) false positive rates in our analysis. For Q1–12, Q1–16 and the cumulative list, we carry out separate MCMC analyses substituting in the false positive rates from Table 2.3. Only the cumulative list remains consistent with our main IS and MCMC results, with overall occurrence rates of the Q1–12, Q1–16 and cumulative datasets being $0.30 \pm 3\%$, $0.32 \pm 3\%$ and $0.38 \pm 3\%$, respectively, i.e. much lower occurrence values. Taking the standard deviation of each radius bin for our main IS result (Fig. 2.5) plus these three new analyses using Q1–12, Q1–16 and cumulative false positive rates, we estimate the uncertainty in our false positive rates on the occurrence of planets for $1 < R_p < 1.4$, $1.4 < R_p < 2$, $2 < R_p < 2.8$, and $2.8 < R_p < 4R_\oplus$ to be 2.1%, 1.6%, 1.8% and 2.4%, respectively.

2.6.2 Astrophysical and Astrobiological Implications

The period distribution of *Kepler* small planets contains two distinct parts. The first is a rise from $\sim$ 5–10 d (e.g. Youdin 2011; Howard et al. 2012), the second is a logarithmically flat distribution extending from $\sim$ 10 d out to at least 200 d (Fig. 2.4 here, Petigura et al. 2013; Fressin et al. 2013). The origin of both features are unclear. But we speculate on one origin for the logarithmically flat feature. Imagine a set of planetary systems comprised of closely-packed, equal-mass planets. Dynamical stability requires that neighbouring planets be spaced apart by more than a few Hill radii (Chambers et al. 1996b; Smith

& Lissauer 2009). Since the Hill radius scales linearly with orbital semi-major axis, this means the separation between neighbouring planets grows linearly with their orbital span. This would then translate into a period distribution that is flat in logarithmic period. In other words, it is possible that most or all of our planets are actually in multiple systems, and that the flat feature is a result of the stability requirement. Although Fang & Margot (2012) have quoted that $75-80\%$ of planetary systems have one or two planets with orbital periods less than 200 d (suggesting that most systems are not close to the Hill stability limit), this result is based on the occurrence of *observed* systems. Most Earth-sized planets (or smaller) around Sun-like stars are undetectable by *Kepler*, and it is possible that the multiplicity of such systems are much higher than we currently believe due to these unseen planets.

Alternatively, the flat feature can arise from the primordial mass distribution in the disk. Assuming that all planets have comparable masses, are in multiple systems, and are formed where they are found today, a logarithmically flat spacing would suggest that the disk surface density Σ scales with the orbital separation a as,

$$\Sigma \propto a^{-2}\,. \tag{2.15}$$

This is not vastly different from the theoretical MMSN profile: $\Sigma \propto a^{-3/2}$ (Hayashi 1981; Weidenschilling 1977b), a useful benchmark to study proto-planetary disks.

The radius distribution is equally intriguing. The radius of a Earth- to Neptune-sized planet mostly reflects the expanse of its hydrogen envelope (Wolfgang & Lopez 2014). By focusing on planets outward of 20 d, we discard candidates that may have had their atmospheres eroded by stellar irradiation Owen & Wu (2013). The distribution shown in Fig.2.5 is therefore likely "primordial". Compared to planets inward of 10 d that have radii $\leq 1.5R_\oplus$, this "primordial" population appears to prefer a size of $\sim 2.5R_\oplus$. Such a size corresponds to a fractional mass in the hydrogen envelope of $\sim 1\%$ (assuming a rocky core roughly in the $10M_\oplus$ range, see, e.g. Wu & Lithwick 2013). What is the reason behind this preferrence for 1%? A planet embedded in a proto-planetary disk can accrete a hydro-static atmosphere. Rafikov (2006) calculated that this atmosphere has a mass of a few $M_\oplus$ for a $10M_\oplus$ planet at 0.1 AU in a MMSN disk. This lies much above the 1% value but it depends on disk parameters and its evolution history. In future works, the observed radius distribution should be used to decipher formation history.

Moreover, the gradual decline toward smaller sizes in logarithmic space has implication for the formation of bare-core planets, the norm in the inner Solar system. Our terrestrial planets are thought to have formed in a gas-free environment by conglomeration of solid materials. The relative shortage of bare-core planets may suggest that the observed *Kepler* planets may have followed different formation path than that of the terrestrial planets.

Lastly, we turn to the issue of $\eta_\oplus$. We calculate $\eta_\oplus$ more out of respect for tradition than with any conviction that there is additional accuracy to be assigned to our calculation. The limits of the habitable zone depend on important assumptions regarding the climate state of Earth-like planets (Kopparapu et al. 2013), mass (Kopparapu et al. 2014), and the composition of the atmosphere (Pierrehumbert & Gaidos 2011). Nevertheless, the search for life elsewhere can take heart in the fact that multiple investigations point to an occurrence of Earth-size planets in habitable zones of $\mathcal{O}(0.1)$ or more. Indeed, studies of M dwarfs suggest that $\eta_\oplus \sim 0.5$ (Bonfils et al. 2013; Kopparapu 2013; Gaidos 2013). M dwarfs comprise about 70% of all stars and hence weigh heavily in the census for Earth-like planets.

2.6.3 Improvements in Occurrence will Happen

The errors associated with most *Kepler* planets are dominated by the uncertainty in the parameters of their host stars. Thus, in order to improve planet occurrence calculations for the future we must first understand *Kepler* stars better. The *Gaia* (Global Astrometric Interferometer for Astrophysics) mission, launched in December 2013, will measure the parallaxes of 1 billion stars in the local group with accuracies approaching 10 μas, as well as obtain multi-band photometry measurements (de Bruijne 2012). Liu et al. (2012) estimate that for stars in the KIC, *Gaia* will be able to estimate $T_{\rm eff}$ to 1%, $\log g$ to within 0.1-0.2 dex, and [Fe/H] to within 0.1-0.2 dex. The combinations of these data should dramatically improve our knowledge of the properties of *Kepler* target stars and hence reconstructions of the *Kepler* planet population.

Other advances include improved maps of interstellar reddening in the *Kepler* field based on the colors of oscillating red giants with established properties, as well as WISE infrared photometry (Huber et al. 2014). The advent of multiplexed, multi-object spectrographs capable of simultaneously measuring thousands of stars (Hill et al. 2010) should, combined with *Gaia* parallaxes, allow stellar parameter estimation with unprecedented scale and precision. In addition, measurement of photometric noise due to stellar granulation ("flicker") is a promising technique for estimating the $\log g$ and hence radius of bright *Kepler* stars to within 0.1-0.2 dex (Bastien et al. 2014), although its calibration and applicability to fainter *Kepler* stars – the majority of targets, with lower photometric precision – remains to be seen.

2.7 Conclusions

In this work we have developed a population simulator to extract the underlying period and radius distributions of Earth- to Neptune-sized planets detected by *Kepler*. We focus on a "primordial" population of planets outside 20 d to exclude the impact of, e.g. photoevaporation. We find that the adoption of a simple model of photometric noise and transit signal detection allow us to accurately estimate the survey completeness of *Kepler*. We have accounted for radius errors in our analysis, and have found that doing so is important for reconstructing the intrinsic radius disitribution. We apply the iterative simulation technique to reconstruct the planet distribution with radius. This does not require binning and allows radius errors to be readily accounted for. Lastly, we are the first to use the updated Huber et al. 2014 parameters along with all 16 quarters of *Kepler* data, representing the most up to date analysis. The main results are as follows:

1. The distribution of planets with $20 < P < 200$ days is roughly uniform with logarithmic period (power-law index $\alpha = -0.04 \pm 0.09$).

2. The (likely primordial) radius distribution for *Kepler* planets with $20 < P < 200$d peaks in the radius bin 2–2.8$R_\oplus$.

3. The overall occurrence of planets within $20 < P < 200$ d and $1 < R_p < 4R_\oplus$ is $46\% \pm 3\%$. This represents the average number of planets per solar-type star in the *Kepler* field.

4. Extrapolating our radius and period distributions out to the habitable zone for solar-type stars, we find $\eta_\oplus = 6.4^{+3.4}_{-1.1}\%$.

5. While our results confirm those from earlier studies, there is a discrepancy in the occurrence of planets for $2.8 < R_p < 4R_\oplus$ planets between our work ($10.5 \pm 1.0\%$) and PHM13 ($6.0 \pm 0.5\%$). Our incorporation of radius errors and updated Huber et al. (2014) stellar parameters account for about half of this discrepancy, while the difference in raw samples account for the remainder. PHM13 includes $P < 20$ d planets into their analysis which likely contains photoevaporated planets (see Table 2.2), decreasing the overall occurrence in the $2.8 < R_p < 4R_\oplus$ bin. Although inconclusive at this time, there is a good chance that the increase in the occurrence of $2.8 < R_p < 4R_\oplus$ planets in our analysis is due to the exclusion of planets altered by proximity to their host stars.

6. In a detailed comparison with PHM13 we find that using CDPP values can effectively reproduce the detection completeness found by the more sophisticated analysis of PHM13.

7. Large radius errors are present in the *Kepler* data, and failing to account for these properly can lead to a different radius distribution. Specifically, this tends to result in a large excess of earth-sized planets. Increasing the size of *Kepler* stars by 25% increases the frequency of large planets while keeping the occurrence of small planets roughly constant. Many stellar radii in the *Kepler* catalog are suspected to be underestimated, and GAIA will improve these stellar radius errors and resolve this issue.

Chapter 3

Tides Alone Cannot Explain Kepler Planets Close to 2:1 MMR

3.1 Chapter Overview

The work, tables and figures from this Chapter are based off the following publication:

Ari Silburt, Hanno Rein
Monthly Notices of the Royal Astronomical Society, Volume 453, Issue 4, p.4089-4096, 2015 (Silburt & Rein 2015).

A number of *Kepler* planet pairs lie just wide of first-order mean motion resonances (MMRs). Tides have been frequently proposed to explain these pileups, but it is still an ongoing discussion. We contribute to this discussion by calculating an optimistic theoretical estimate on the minimum initial eccentricity required by *Kepler* planets to explain the current observed spacing, and compliment these calculations with N-body simulations. In particular, we investigate 27 *Kepler* systems having planets within 6% of the 2:1 MMR, and find that the initial eccentricities required to explain the observed spacings are unreasonable from simple dynamical arguments. Furthermore, our numerical simulations reveal *resonant tugging*, an effect which conspires against the migration of resonant planets away from the 2:1 MMR, requiring even higher initial eccentricities in order to explain the current *Kepler* distribution. Overall, we find that tides alone cannot explain planets close to 2:1 MMR, and additional mechanisms are required to explain these systems.

3.2 Introduction

The NASA *Kepler* mission has been immensely successful for detecting planets outside of our solar system. To date, it has discovered over 4,500 exoplanet candidates along with 466 multi-planet systems (Akeson et al. 2013; Rowe et al. 2014). A number of these systems are just wide of a mean motion resonance (MMR), which occurs when the period of one planet is an integer ratio of another. In

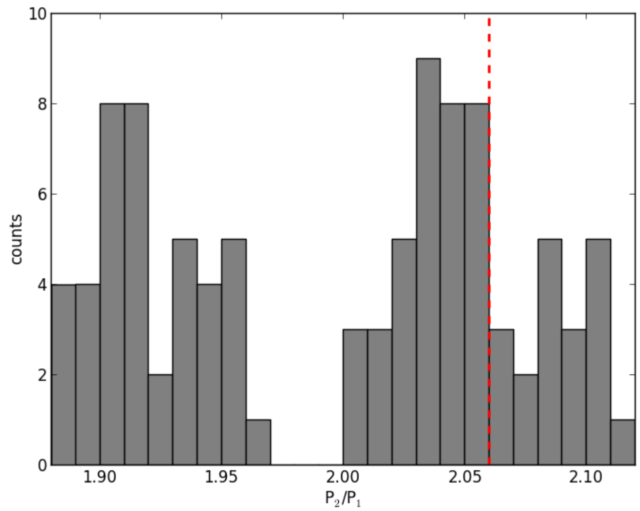

Figure 3.1: *Kepler* systems close to 2:1 MMR. A statistical excess is present just wide of the 2:1 MMR, and appears to decline beyond 6% of the resonance, as marked by a red dotted line.

particular, statistical excesses in the period distribution of *Kepler* planets have been detected just wide of the 3:2 and 2:1 MMR (Lissauer et al. 2011; Fabrycky et al. 2014; Steffen & Hwang 2015). It is believed that in the past these planetary systems migrated into resonance via convergent migration (Lee & Peale 2002), and a number of dissipative mechanisms have been proposed to slowly bring these planets out of MMR. The most popular dissipative mechanisms to explain the observed near-resonant systems are tidal (Lithwick & Wu 2012; Batygin & Morbidelli 2013; Delisle et al. 2014), protoplanetary (Rein 2012; Baruteau et al. 2013; Goldreich & Schlichting 2014), and planetesimal (Moore et al. 2013; Chatterjee & Ford 2015).

In this Chapter we focus exclusively on tidal dissipation, for which there is no clear consensus on whether this mechanism alone can successfully explain the excess of near-resonant pairs near first order MMRs. Several authors (Lithwick & Wu 2012; Batygin & Morbidelli 2013) have argued for (Delisle et al. 2014) tidal dissipation, whereas others (Lee et al. 2013) have argued against it. In particular, Lithwick & Wu (2012) first introduced the mechanism of *resonant repulsion*, and showed that in the limit of low eccentricities for near-resonant planets, the space between planets grows as $t^{1/3}$. Batygin & Morbidelli (2013) confirmed this result. Lee et al. (2013) then used this relationship to show that most near-resonant planet pairs cannot be explained via this mechanism, the few exceptions being small rocky planets for which tidal dissipation is particularly effective. Delisle et al. (2014) then suggested a high eccentricity mechanism by which planets may still be able to evolve to their current positions via tides alone.

We contribute to this debate by developing optimistic theoretical estimates for the evolution of planets away from resonance, and compare these estimates to N-body simulations. We then make a statement about the likelihood of tides to explain near-resonant pairs. We focus on *Kepler* systems within 6% of the 2:1 MMR, which appears to be the natural cutoff for this statistical excess, as shown by a dotted line in Figure 3.1. Using this sample we will calculate the minimum required initial eccentricity to explain their current positions, given that they started in MMR and evolved under the influence of tides alone. In addition, we present numerical findings of *resonant tugging*, an effect which prevents the evolution of planets away from MMR when eccentricity is high, making it significantly more difficult to achieve the observed spacings seen today. Resonant tugging does not appear to have been extensively studied/accounted for in this field due to the fact that most analysis of planets in resonance have worked in the $e \ll 1$ limit, and we find that resonant tugging exclusively affects planets in MMR with moderate to high e.

This Chapter is organized as follows – in Section 3.3 we outline our theoretical and numerical framework, in Section 3.4 we present the main findings of this Chapter, in Section 3.5 we present a discussion and conclude in Section 4.6.

3.3 Methods

3.3.1 Theory

The equations widely used to describe the evolution of planets under the influence of tides are (e.g. Barnes et al. 2008; Lithwick & Wu 2012; Lee et al. 2013):

$$\dot{e} = -\frac{9}{2}\pi\frac{k}{Q}\frac{1}{m_p}\sqrt{\frac{GM^3}{a^3}}\left(\frac{r_p}{a}\right)^5 e \tag{3.1}$$

$$\dot{a} = -9\pi\frac{k}{Q}\frac{1}{m_p}\sqrt{\frac{GM^3}{a^3}}\left(\frac{r_p}{a}\right)^5 e^2 a, \tag{3.2}$$

where a is the semi-major axis, e is the eccentricity, k is the planet's Love number, Q is the planet's *tidal quality factor* (Goldreich & Soter 1966), m_p and r_p are the planet's mass and radius, respectively, M is the stellar mass, and G is the gravitational constant. From Eqs. 3.1 and 3.2 it follows that $\dot{e}$ and $\dot{a}$ are related by:

$$\frac{\dot{a}}{a} = 2e\dot{e} \tag{3.3}$$

Which arises from conservation of orbital angular momentum. We are interested in finding a relationship between the total migration of the system and its state variables (such as a and e), which can be obtained by integrating Eq. 3.3:

$$\int_{a_f}^{a_i}\frac{da}{a} = 2\int_{e_f}^{e_i} ede$$

$$a_f = a_i \exp(-e_i^2 + e_f^2), \tag{3.4}$$

where subscripts i and f refer to the initial and final states, respectively.

Eq. 3.4 is surprisingly simple – if we know the initial states e_i and a_i of a planetary body, as well as the final eccentricity e_f, we can predict its final position. Eq. 3.4 is independent of tidal response parameters (k, Q), the length of time considered, the mass and radii of the star/planet, etc. These additional factors affect the timescale by which the body arrives at its final position, a_f, but not the final position of the body itself.

Let us measure the spacing of a planet pair close to a $j : j+1$ MMR by defining

$$\Delta \equiv \frac{P_{out}}{P_{in}} - \frac{j+1}{j}, \tag{3.5}$$

where P is the orbital period, and the subscripts *in* and *out* refer to the inner and outer planet, respectively. For the 2:1 MMR, $j = 1$. This definition of Δ is the same as Lee et al. (2013), and twice the value used by Lithwick & Wu (2012).

Using the fact that $P \propto a^{3/2}$, and substituting Eq. 3.4 into Eq. 3.5, we recast Δ in terms of a_i and e_i for the near resonant pair (setting $j = 1$):

$$\Delta = \left(\frac{a_{out,f}}{a_{in,f}}\right)^{3/2} - 2 \tag{3.6}$$

$$\Delta = \left(\frac{a_{out,i}\exp(-e^2_{out,i} + e^2_{out,f})}{a_{in,i}\exp(-e^2_{in,i} + e^2_{in,f})}\right)^{3/2} - 2$$

We make the assumption that $e_{out,i} \approx e_{out,f}$ (see Section 3.5 for a discussion). After simplifying, we get:

$$\Delta = \left(\frac{a_{out,i}\exp(e^2_{in,i} - e^2_{in,f})}{a_{in,i}}\right)^{3/2} - 2. \tag{3.7}$$

Thus, Eq. 3.7 relates the final spacing of the system, Δ, to the initial conditions of the system (a_i and e_i) and the final eccentricity of the inner planet, $e_{in,f}$. Eq. 3.7 assumes that the angular momentum of each individual planet is conserved, which in general is not true for multi-planet systems (only total angular momentum is). Other works (Lithwick & Wu 2012; Batygin & Morbidelli 2013) have accounted for this fact, resulting in more accurate equations for the evolution of a planet pair. We now aim to compare Eq. 3.7 to numerical simulations, and estimate its accuracy for *Kepler* systems with $\Delta < 0.06$ of the 2:1 MMR. We first outline our experimental setup.

3.3.2 Experimental Setup

Numerical simulations are carried out using the Wisdom & Holman integration scheme (Wisdom & Holman 1991), implemented via REBOUND (Rein & Liu 2012). Our sample consists of *Kepler* systems near the 2:1 MMR with $\Delta < 0.06$, which we call *near-resonant pairs*. Our choice of $\Delta < 0.06$ stems from a natural cutoff where the excess of near-MMR ends, as shown in Figure 3.1. We exclude near-resonant pairs interior to the 2:1 MMR, since tidal forces appear only to increase planet separation with time (e.g. Lithwick & Wu 2012). In addition, we also exclude near-resonant pairs that are *complex*. By complex, we mean dynamically involved in an additional (near) resonance (e.g. Laplace resonance), and/or containing an additional planet orbiting between the near-resonant pair. Our exclusion of complex resonant systems decreases the number of *Kepler* systems by 6. Lastly, we also remove the Kepler-11 system, which does not contain a complex resonance, but does go unstable on short (∼Myr) timescales when placed into resonance. This leaves us with 27 *Kepler* systems, and their properties are displayed in Table A.1. We remind the reader that for each *Kepler* system we simulate the entire system, not just the near-resonant planets.

Many *Kepler* planets do not currently have measured mass values, so we assign planet masses using Eq. 3 from Weiss & Marcy (2014) for planets $r_p < 4r_\oplus$: $(m_p/m_\oplus) = 2.69(r_p/r_\oplus)^{0.93}$, and assume a density of Jupiter for planets $r_p > 4r_\oplus$. In addition, we also input mass values from the transit-timing variation study of Hadden & Lithwick (2014b) where applicable. For stars without measured stellar masses, we assume $M/M_\odot = (R/R_\odot)^{1.25}$, derived from Demircan & Kahraman (1991). For simplicity, we also assume that the inclination of our *Kepler* planets is zero.

We assign k/Q values following a similar prescription as Lee et al. (2013) by assigning the most generous values possible. For Earth-like rocky planets, k/Q($r_p/r_\oplus < 2$) = 1/40, for planets smaller than Jupiter, k/Q($2 < r_p/r_\oplus < 10$) = 1/22000, and for Jupiter-sized giant gaseous planets, k/Q($r_p/r_\oplus > 10$) = 1/54000.

To speed up simulation time, we increase k/Q by a factor of $A_{k/Q}$ (or alternatively this could also be interpreted as increasing tidal strength). This tactic has been used by other scientists (e.g. Delisle et al. 2014), and is valid as long as τ_e is much longer than the planet's eccentricity libration time. We simulate our *Kepler* sample for 50 Myr, and use $A_{k/Q} = 200$, giving a total simulation time of $T = 10$ Gyr.

To begin our simulations, we place each planet at a distance of $1.15a_{obs}$, where a_{obs} is the observed semi-major axis value from the *Kepler* catalog, and assign $e_i = 0.01$. We then migrate each planet (in a type-I fashion) back to its original starting position a_{obs} except for of the outer near-resonant planet, which instead migrates a distance of $a_{obs} + \Delta$, forcing the near-resonant pair into a 2:1 MMR.

Each planet migrates for time t_{mig} at rate $\dot{a} = an\mu^{4/3}/C$, where n is the mean motion of the inner planet, μ is the planet/star mass ratio, and C is a constant. Lower values of C cause the outer planet to encounter the MMR sooner, allowing time for both planets to migrate together in resonance, which increases eccentricity to a desired value (Lee & Peale 2002). For the restricted 3-body problem Goldreich & Schlichting (2014) guarantee capture into resonance if $C_{out} > 3.75$, $e_{in} < (\mu_{out}/j)^{1/3}$ but we use a more conservative value of $C_{out} = 6$ as well as perform numerical tests to ensure that overstabilities do not occur on Myr timescales.

Defining $K \equiv \frac{\dot{e}_i/e_i}{\dot{a}_i/a_i}$, we use a default $K = 100$ when migrating planets into resonance but also experiment with $K = 10$. K (along with m_p/M) affects the resulting equilibrium eccentricity, but does not affect tidal evolution, and thus does not affect our main conclusions. At this point, initial eccentricities of our simulated *Kepler* systems range from $0.05 < e_i < 0.25$, depending on the value of C, K, etc.

After time t_{mig}, migration is quenched over a timescale of $t_{mig}/3$ by letting $\tau_a \to \infty$ and $\tau_e \to \infty$, where $\tau_a \equiv -a/\dot{a}$ and $\tau_e \equiv -e/\dot{e}$. It is at this point that tides are turned on, and the system evolves under the influence of Eq. 3.1 and 3.2 for the remainder of the simulation.

3.4 Results

3.4.1 Theory vs. Numerics

We compare our theoretical predictions of tidal evolution, Δ_{th} (Eq. 3.7), to our numerical simulations, Δ_{num} (Eq. 3.6). The difference, $\Delta_{num} - \Delta_{th}$, is displayed as a solid line in Figure 3.2, and expressed as a cumulative distribution (CDF). For all but 2 systems we see that $\Delta_{num} - \Delta_{th} < 0$, indicating that our theoretical predictions Δ_{th} consistently over-predict our numerical results, Δ_{num}. These 2 exceptions, Kepler-32 and Kepler-221, are expected due to the fact that $T/\tau_e \sim 1000$, which allows extensive $\Delta \propto t^{1/3}$ resonant repulsion growth (Lithwick & Wu 2012) in time T which is not accounted for in Eq. 3.7. Otherwise, we find that Eq. 3.7 consistently overestimates the true evolution of the system.

The reason for this $\Delta_{num} - \Delta_{th} < 0$ trend is due to *resonant tugging*, an effect present in the numerics but not captured by our theoretical predictions. Resonant tugging acts to keep planets closer together than theory would predict. We explore resonant tugging in the next section.

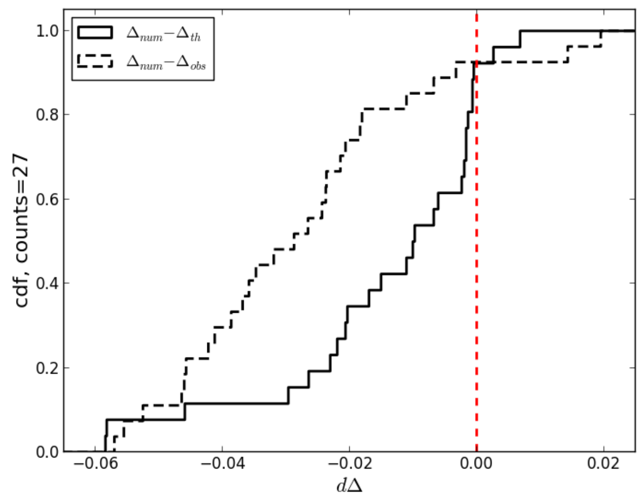

Figure 3.2: Cumulative distribution function (CDF) of our results. The solid line shows $\Delta_{num} - \Delta_{th}$, the difference between the theoretical and simulated planet separations after T = 10 Gyr. The dashed line shows $\Delta_{num} - \Delta_{obs}$, the difference between our numerical simulations and the observed *Kepler* spacing.

For the same simulations we also plot $\Delta_{num} - \Delta_{obs}$ as a dashed line in Figure 3.2, which is the difference between our numerical results, Δ_{num}, and the observed spacing of *Kepler* planets seen today, Δ_{obs} (Eq. 3.5). As is clearly shown, $\Delta_{num} - \Delta_{obs} < 0$ for all but two systems, suggesting that tides alone cannot explain the observed spacing of *Kepler* planets. These two exceptions are again, Kepler-32 and Kepler-221, and are exceptions for the same reason as above. Since $\Delta_{obs} = 0.038$ and 0.035 for Kepler-32 and Kepler-221, respectively, using a lower Δ cutoff for our *Kepler* sample (e.g. $\Delta_{obs} < 0.03$) would not have changed our result that tides cannot explain near-resonant pairs.

Although highly suggestive, this result does not conclusively disprove tides as the primary evolving mechanism, since higher e_i could cause more migration in time T (see Eq. 3.4), and the median eccentricity for our simulations is $e_{in,i} = 0.14$. Higher values of $e_{in,i}$ are possible. Instead of numerically exploring every possible $e_{in,i}$ value, however, we instead reverse the argument in Section 3.4.3 and calculate the minimum $e_{in,i}$ required to explain Δ_{obs}. First, however, we explore resonant tugging.

3.4.2 Resonant Tugging

Resonant tugging affects planets in MMR subjected to energy dissipation (e.g. tidal), with moderate to high eccentricity. When these conditions are met and the inner planet migrates inwards (trying to leave the resonance) it tugs the outer planet inwards along with it, transferring dissipative forces from the inner to the outer planet. The result is that the inner planet migrates less than expected, the outer planet migrates more than expected, and the planets are closer together than theory would have predicted.

Figure 3.3 illustrates resonant tugging for a pair of $m = 10^{-4}M$ planets in MMR. For the black curve $e_{in,i} = 0.125$, for the grey curve $e_{in,i} = 0.018$, and otherwise the initial conditions of each test case are the same. In both cases we allow *only the inner planet* to evolve under the influence of tides. The top and bottom panels show the period evolution of the outer and inner planets, respectively, and the dotted curve in the bottom panel shows evolution of the inner planet in the absence of the outer planet (also $e_{in,i} = 0.125$).

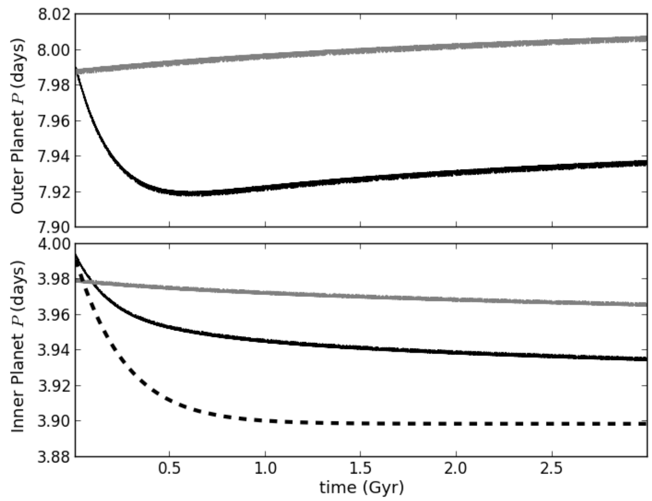

Figure 3.3: Two test cases illustrating resonant tugging and repulsion. The top and bottom panel shows the period evolution of the inner and outer planet, respectively. For the black curve $e_{in,i} = 0.125$, while for the grey curve $e_{in,i} = 0.018$. The dotted black curve shows the numerical trajectory of the inner planet ($e_{in,i} = 0.125$) in the absence of the outer planet.

Resonant tugging is exhibited in the first 0.5 Gyr of evolution for the black curve, showing how tidal forces affecting the inner planet also affect the outer planet by dragging it inwards too. Comparing the solid and dotted black curves in the bottom panel of Figure 3.3, we see that in the presence of the outer planet, the inner planet migrates much less than expected due to resonant tugging. Since the outer planet has also migrated inwards more than expected, the result is that Δ_{black} has grown very little over the first 0.5 Gyr of evolution ($\Delta_{black} = 0.004$ after 0.5 Gyr).

Within the framework of resonant tugging, the outer planet can be thought of as a massive anchor – as $m_{out}/m_{in} \to \infty$ the inner planet has an increasingly difficult time migrating both bodies inwards, leading to a pair of (relatively) stationary planets. Conversely, as $m_{out}/m_{in} \to 0$ the inner planet has an easier time migrating both bodies inwards, and the trajectory of the inner planet will approach its single-planet trajectory (i.e. the dotted black line in Figure 3.3).

We now briefly contrast resonant tugging from resonant repulsion (first described by Lithwick & Wu 2012). Qualitatively, the two main differences between resonant tugging and resonant repulsion are:

1. Resonant tugging exclusively affects planets in resonance (resonant angle(s) librating) with moderate to high eccentricity, while resonant repulsion affects planets both in and close to resonance, and is most noticeable in the $e \ll 1$ limit.

2. Resonant tugging decreases a_{in} and decreases a_{out}, while resonant repulsion decreases a_{in} and increases a_{out}.

These two differences are illustrated in Figure 3.3. The grey curve, which has low initial eccentricity ($e_{in,i} = 0.018$) exhibits pure resonant repulsion – a decrease in a_{in} and an increase in a_{out}, while the black curve first exhibits a shorter period of resonant tugging followed by resonant repulsion. The transition from resonant tugging to resonant repulsion for the black curve occurs after 0.5 Gyr, when the eccentricity of the inner planet has dropped to a low ($e_{in} = 0.035$) value.

In Figure 3.3, after ~ 1 Gyr, $\Delta_{black} < \Delta_{grey}$, but $\dot{\Delta}_{grey} = \dot{\Delta}_{black}$, showing how resonant tugging can permanently stunt the growth of Δ. We also see for the black curve that $a_{out,f} < a_{out,i}$, and it is only after many τ_e timescales that the outer planet can recover (or exceed) its initial position via resonant repulsion. Furthermore, since to first order (when $T/\tau_e \ll 1000$) $\dot{\Delta} \propto \tau_{a,in} \propto e^2_{in,i}$ one would naively expect the black curve to experience e^2 more Δ growth in time T, yet we actually find $\Delta_{black} < \Delta_{grey}$, illustrating just how significant resonant tugging can be. This means that the high eccentricity tidal mechanism suggested by Delisle et al. (2014) does not work for planets in resonance, since (contrary to expectation) high eccentricity actually stunts the growth of Δ, not accelerates it.

The results shown in Fig. 3.2 are due to resonant tugging, and is supported by the fact that for every simulated *Kepler* system we find $a_{out,f}/a_{out,i} < 1$. Since resonant repulsion can only increase a_{out} (as shown in Fig. 3.3), $a_{out,f}/a_{out,i} < 1$ can only be due to resonant tugging since inward tidal migration is a negligible contribution for the outer planet. The analytic aspects of resonant tugging will be studied in more depth in future works.

3.4.3 Minimum Eccentricity to Explain Δ_{obs}

Since we found in Section 3.4.1 that Δ_{th} consistently over-predicts the amount of tidal migration (due to resonant tugging), we can use it as an upper limit predictor of tidal evolution, assuming that near-resonant pairs started in 2:1 MMR and evolved to their present locations. Starting from Eq. 3.3, and using the same logic as Section 3.4.1 we calculate the minimum eccentricity required by the inner planet to achieve the observed spacing seen today, Δ_{obs}, after T years (see Appendix A.1 for detailed calculations):

$$e_{in,i} = \sqrt{\frac{\ln[(\frac{a_i}{a_f})_{out}(\frac{\Delta_{obs}+2}{2})^{2/3}]}{1-\exp(-2T/\tau_{e,in})}} \tag{3.8}$$

Thus, for a pair of planets starting in 2:1 MMR, if we know the observed spacing today, Δ_{obs}, the number of $\tau_{e,in}$ damping timescales in time T, and estimate the amount of migration done by the outer planet in time T, $(a_i/a_f)_{out}$, we can calculate the *minimum* initial eccentricity that the inner planet must have in order to arrive at the current spacing, $e_{in,i}$. Since we assume a starting position of exact 2:1 commensurability, Equation 3.8 is only tailored for $\Delta_{obs} > 0$. Tides cannot decrease planet spacing over time.

Figure 3.4 shows CDFs of Eq. 3.8 applied to our *Kepler* sample for T = 1 (solid), 5 (dashed), and 10 (dotted) Gyr with the outer planet remaining stationary, i.e. $(a_i/a_f)_{out} = 1$ (see Section 3.5 for a discussion). In addition we plot a $T \to \infty$ curve as a dash-dotted line, which the other curves converge to. The shaded red region marks where $e_{in,i} \geq 1.0$, while the blue marks an unstable region where eccentricities are unlikely to exist. We construct the blue region by numerically finding the maximum eccentricity allowed before $> 50\%$ of our *Kepler* systems go unstable within 2 Myr (see Appendix A.2 for details). We find this maximum eccentricity to be ≈ 0.3 (the left boundary of the blue shaded region in Figure 3.4).

The three $T < \infty$ curves in Figure 3.4 lie largely in the red and blue shaded regions, indicating that the required eccentricities to explain Δ_{obs} are unreasonable. We thus conclude that most *Kepler* systems cannot be explained by tides alone. Even for the $T = 10$ Gyr curve, an optimistic estimate for the age of many *Kepler* systems, about 35% of systems still cannot be explained due to tides alone. Clearly, another mechanism is needed to explain the near-resonant 2:1 MMR pairs.

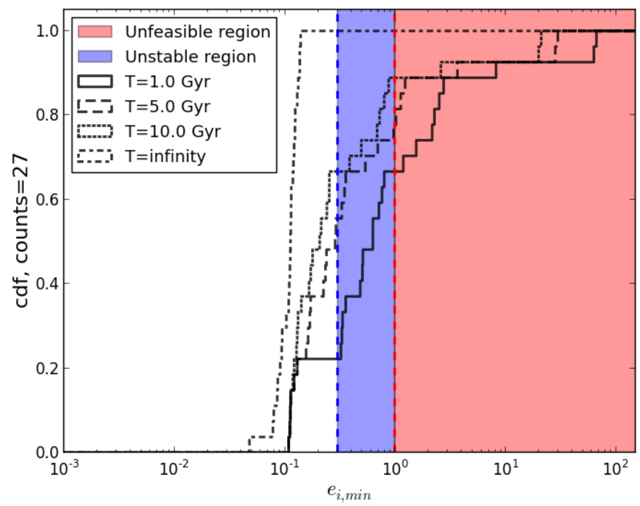

Figure 3.4: Three CDFs showing the theoretical minimum eccentricity required by the inner planet in order to achieve the observed Δ spacing seen by *Kepler* planets today. The solid, dashed, dotted lines represent $T = 1, 5, 10$ Gyr tracks, respectively, while the dash-dotted line represents $T \to \infty$. In all calculations we assume the outer planet remains stationary, i.e. $(a_i/a_f)_{out} = 1$. The red shaded region marks the unphysical region where the eccentricity is larger than unity. The blue region marks the region where most systems undergo a dynamical instability.

3.5 Discussion

A number of assumptions have been made in constructing Figure 3.4 which only strengthen our conclusion that planets close to the 2:1 MMR cannot be explained due to tides alone.

First and foremost, from Sections 3.4.1 and 3.4.2 we have shown that resonant tugging causes our theoretical predictions of Δ to be overestimates, and thus Eq. 3.8 underestimates the minimum $e_{in,i}$ required to explain the current observed spacing. This discrepancy becomes most pronounced when $m_{out}/m_{in} > 1$, which is the case for many *Kepler* systems in our sample. In particular we showed that resonant tugging can stunt the evolution of planets away from MMR by roughly e^2 (see Sec 3.4.2), and this stunted evolution is not accounted for in Figure 3.4.

Second, we have assumed optimistic k/Q values which allow for more migration in time T. Although *some* planets could have such generous values, it is unlikely that all of them do. Since $\dot{a} \propto (k/Q)e^2$, as k/Q decreases $e_{in,i}$ must increase in order for the planets to achieve the same observed Δ in time T.

Third, our estimate of $e_{max} = 0.3$ (blue-shaded region in Figure 3.4) is very likely an overestimate. For example, Pu & Wu (2015) found that the mean eccentricity of high-multiple Kepler planets must not exceed $e_{mean} = 0.02$ to guarantee long-term dynamical stability. Also, from simple orbit crossing arguments when $e > 0.23$ the perihelion of the outer planet crosses the aphelion of the inner planet for a 2:1 MMR, and long-term stability can no longer be guaranteed. Even if it were possible for *Kepler* systems to remain stable with high (> 0.3) eccentricities, it is unclear what kind of mechanism could consistently generate them for our *Kepler* sample.

There are other assumptions we made throughout this Chapter which we now justify. For the results presented in Fig. 3.4 we assumed that the outer planet remains stationary, i.e. $(a_i/a_f)_{out} = 1.00$. This is a reasonable assumption since, referring to our $T = 10$ Gyr numerical simulations as a benchmark, the median value of $(a_i/a_f)_{out} = 1.001 \approx 1$. Increasing the value of $(a_i/a_f)_{out}$ shifts our CDFs in Fig. 3.4 further into the blue/red instability region.

In Eq. 3.7 we assumed that $e_{out,i} \approx e_{out,f}$ which essentially states that the the initial and final positions of the outer planet are the same. As stated in the previous paragraph, since we found numerically from our simulations that the median $(a_i/a_f)_{out} \approx 1$, this assumption is reasonable.

In constructing Fig. 3.4 we have used Eq. 3.8, which is Eq. 89 from Delisle et al. (2014), who argued that for moderate to high eccentricities ($e_{in,i} \geq 0.15$), many of the near-resonant pairs could in fact be explained by tides. There are a number of differences between our analysis and theirs however. First (and most importantly), their estimates are based on theoretical predictions (i.e. resonant tugging unaccounted for), and we have shown in Section 3.4.2 that the growth of Δ is significantly stunted when resonant tugging is accounted for, especially when $m_{out}/m_{in} > 1$. Second, their analysis assumes that $T \to \infty$, while we restrict to $T = 1, 5$ and 10 Gyr. Looking at Fig 3.4, we see that the $T \to \infty$ curve tells a very different story than the $T = 1, 5$ and 10 Gyr curves, and the conclusion of whether or not tides can explain Δ_{obs} is certainly time dependent. Lastly, Delisle et al. assumes $\Delta = 0.03$ for all systems, while we use the system specific Δ values.

As a consistency check for our results, we perform the same set of experiments (including our tests of resonant tugging) using a different version of tides, implementing them in terms of forces (as opposed to orbital elements like in Eq. 3.1 and 3.2) according to Papaloizou & Larwood (2000):

$$\vec{a}_{damp} = -2\frac{(\vec{v}\cdot\vec{r})\vec{r}}{r^2\tau_e} \tag{3.9}$$

where $\vec{a}_{damp}$ is the damping acceleration, $\vec{v}$ is the velocity, $\vec{r}$ is the position, r is the scalar position and $\tau_e \equiv -e/\dot{e}$ as before. Whenever a planet receives a "kick" in the WH integration scheme, an additional kick of a_{damp} is supplied to account for tides. We find our overall conclusions unaffected using this implementation of tides.

We have omitted complex resonances from our analysis because their behaviour is much more unpredictable. For simple resonances we found that $\Delta_{th} > \Delta_{num}$, but some complex resonances violated this relationship. The reasons are currently unknown, but will be more thoroughly investigated in future works.

Lastly, it should be mentioned that spin tides were omitted from this analysis, which arise when the spin rate of the host star Ω_* is different than the mean motion n of the orbiting planet (responsible for the Moon's recession from Earth over time). Spin rates of *Kepler* stars are largely unknown, as well as the evolution of these spin rates, $d\Omega_*/dt$. Depending on the sign of $(\Omega_* - n)$, spin tides can induce inward or outward migration. It is thus a non-trivial process to determine what the affect of spin tides might be on the evolution of a system. The equation governing spin tide migration is (Murray & Dermott 1999):

$$\dot{a}_p = \mathrm{sign}(\Omega_* - n_p)\frac{3k_*}{Q_*}\frac{m_p}{M}\left(\frac{R}{a_p}\right)^5 n_p a_p \tag{3.10}$$

where the subscripts p and $*$ refer to the planet and star, respectively. We can estimate the relative strength of eccentricity tides (Eq. 3.2) to spin tides (Eq. 3.10):

$$\frac{\dot{a}_{ecc}}{\dot{a}_{spin}} = \frac{9\pi\frac{k_p}{Q_p}\sqrt{\frac{GM^3}{a_p^3}}\frac{e_p^2 a_p}{m_p}}{3\frac{k_*}{Q_*}\frac{m_p}{M}\left(\frac{R}{a_p}\right)^5 n_p a_p} = 3\pi\frac{k_p}{k_*}\frac{Q_*}{Q_p}\left(\frac{M}{m_p}\right)^2\left(\frac{r_p}{R}\right)^5 e^2$$

Interestingly enough, the relative strength of spin vs. eccentricity tides is independent of the semi-major axis. Assigning typical values from Wu & Murray (2003) for Ω_*, $(k/Q)_*$, M and R, and assuming a typical $\sim 4m_\oplus$ planet we get:

$$\frac{\dot{a}_{ecc}}{\dot{a}_{spin}} \sim 30 \tag{3.11}$$

Combining this with the fact that Ω_* and $d\Omega_*/dt$ are largely unknown for *Kepler* stars, we felt justified omitting spin tides from our analysis.

3.6 Conclusion

In conclusion, we have investigated 27 *Kepler* systems containing 2:1 near-resonant pairs, and find that tides alone cannot explain their current observed spacing, Δ_{obs}. In Figure 3.4 we calculated the minimum theoretical eccentricity required by the inner planet to explain Δ_{obs} and found that for a large number of systems $e_{in,i} > 0.3$, which from simple dynamical arguments is not a reasonable eccentricity for *Kepler* planets to have. Furthermore, our numerical study of resonant tugging reveals that our theoretical predictions of $e_{in,i}$ are optimistic estimates, and in cases where $m_{out}/m_{in} > 1$, significantly so. A number of other assumptions made throughout the Chapter contribute to these optimistic estimates.

As a numerical compliment to our theoretical investigation, we simulated our *Kepler* sample for 10 Gyr with a median eccentricity of $e_{in,i} = 0.14$, and found only two systems, Kepler-32 and Kepler-221, that migrated to Δ_{obs} (dotted line in Figure 3.2). Clearly, another mechanism is required to explain the excess of *Kepler* systems exterior to the 2:1 MMR.

Chapter 4

HERMES: a hybrid integrator for simulating close encounters and planetesimal migration

4.1 Chapter Overview

The work and figures from this Chapter have been done in collaboration with Hanno Rein and Daniel Tamayo. A publication is currently in preparation.

We present HERMES, a new hybrid integration scheme for long-term simulations of planetary systems undergoing close encounters and planetesimal-driven migration. Particles are integrated using WHFAST, a fast and accurate symplectic integrator, unless a close encounter occurs. During a close encounter, a subset of particles is integrated with the high-order integrator IAS15, while the rest of the particles continue to be integrated with WHFAST. We created an adaptive routine for optimizing the close encounter boundary to help maintain accuracy whilst close encounters are occurring.

Like most hybrid integrators, the switching between integrators leads to an additional, finite energy error above the standard oscillatory energy error arising from symplectic integration. HERMES takes a more direct approach when switching between integrators than previous schemes in the literature, allowing us to analytically estimate the numerical error of our algorithm. Since WHFAST is symplectic, IAS15 is accurate to machine precision and both of them are unbiased, the energy error grows sub-linearly with time under the assumption that either impact parameters are randomly distributed or close encounters are rare.

We find that HERMES provides a good balance between speed and accuracy, neither achieved by the individual symplectic or non-symplectic integrators alone. In this Chapter, we describe the details of implementation, accuracy and performance, as well as its incorporation within the larger framework of the N-body package REBOUND.

4.2 Introduction

Over the last 25 years scientists have made considerable progress integrating N gravitationally interacting particles (an N-body system) using computational techniques. The most widely used integrator today for solving Solar System type problems is the Wisdom-Holman integrator (Wisdom & Holman 1991, hereafter WH), which decomposes the system's Hamiltonian, H, into a Keplerian and an interaction component, H_K and H_I. Symplectic integrators which split the Hamiltonian in this way are known as mixed-variable symplectic integrators (Wisdom & Holman 1991; Saha & Tremaine 1992). The system is then evolved in a second-order leapfrog manner, taking the form of $K(\mathrm{dt}/2)I(\mathrm{dt})K(\mathrm{dt}/2)$, where K represents evolution under H_K, I represents evolution under H_I, and dt is the timestep. Although higher-order algorithms are possible (e.g. Yoshida 1990), the second-order WH method is an ideal balance of speed and accuracy, since the marginal increase in accuracy of higher-order methods comes at the cost of significant additional calculation. The evolution under the interaction Hamiltonian is trivial to solve exactly in Cartesian coordinates, whereas the evolution under the Keplerian Hamiltonian is easy to solve exactly using orbital elements. This algorithm therefore converts between the two coordinate systems each timestep.

Since the WH scheme breaks the evolution into operators that both derive from Hamiltonians, the algorithm is symplectic (for a review on symplectic algorithms see Yoshida (1993)). This implies that the numerical solution conserves quantities closely related to the integrals of motion, such as the total energy. In practice these integrals of motion are not constant, but oscillate in a bounded manner. The relative energy error scales as $O(\epsilon dt^2)$ if the magnitude of H_I remains $O(\epsilon)$ smaller than H_K, where $\epsilon \ll 1$ (Saha & Tremaine 1994). For distant particles in non-overlapping orbits ϵ is typically much less than unity. This is one motivation for splitting H into H_K and H_I, as it allows for longer timesteps (and thus shorter integration times) than conventional integration schemes. However, during close encounters, H_I becomes comparable to or larger than H_K causing ϵ, and thus the energy error, to grow substantially. Therefore, despite their brief duration, close encounters typically dictate an unacceptably short timestep for the entire simulation. Note that it is not possible to dynamically change the timestep in the standard WH integrator as it would break time symmetry and symplecticity[1].

For very high accuracy integrations (with relative errors of order the machine precision), non-symplectic integrators are as good and as fast as or faster than symplectic integrators (Rein & Spiegel 2015). But in most integrations, medium to low accuracy is enough to capture the qualitative evolution of a system. In such a case a symplectic integrator provides an advantage as the timestep can be large while keeping the numerical errors bounded. This advantage of symplectic integrators, together with the common need to accurately resolve close encounters motivates the development of hybrid integrators that can make use of both symplectic and non-symplectic integrators.

Several hybrid integrators that make use of a (modified) WH integrator have been developed. The two most popular ones are `SyMBA` (Duncan et al. 1998) and the hybrid integrator from the `Mercury` package (Chambers 1999, hereafter referred to as `Mercury`).

`SyMBA` decomposes the interaction potential into a series of shells around each body and uses progressively smaller timesteps for each shell to increase time resolution. If the particles are well separated, there is only one contributing shell to the interaction term and the integrator is effectively

[1]If one can make the timestep choice independent of the current state, for example by using a predefined sequence of timesteps, then the integrator remains symplectic.

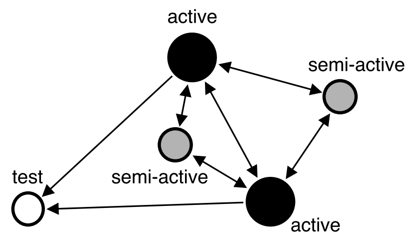

Figure 4.1: A diagram illustrating how different particle types (active, semi-active, test) affect each other. Arrows indicate directions of gravitational influence.

WH. During a close encounter the inner shells contribute to the integration using smaller timesteps to resolve the encounter.

Mercury handles close encounters by using a smooth changeover function to transfer large terms from H_I to H_K. When particles are distant, interactive forces between particles are small and evaluated during H_I. When particles are close, interactive forces become large and are transferred to H_K, thus keeping H_I small. This makes H_K a three body problem (central body plus two particles undergoing a close encounter) which cannot be solved analytically but is straightforward to integrate numerically to high precision using a Bulirsch-Stoer routine (Press et al. 1988).

In this Chapter we present a new integration method, HERMES, which borrows ideas from the integrators mentioned above, but takes a more direct approach to handling close encounters. It combines two existing integrators, WHFAST (Rein & Tamayo 2015) which is a fast and unbiased implementation of the WH method, and the high-order IAS15 integrator (Rein & Spiegel 2015). HERMES has been seamlessly incorporated into REBOUND (Rein & Liu 2012), adding further flexibility to the modular N-body package.

The outline for the Chapter is as follows: Section 4.3 describes the algorithm for HERMES, Section 4.4 characterizes error, Section 4.5 shows standard tests of HERMES as well as a comparison to SyMBA and Mercury, and we conclude in Section 4.6.

4.3 Methods

4.3.1 Particle Classification

First we define the three different types of particles handled by HERMES: active particles, semi-active particles and test particles. Active particles can gravitationally affect all other types of particles, and are typically stars or planets. Semi-active particles can affect active particles only (not other semi-active particles), and are typically asteroids, planetesimals, and other smaller objects. Test particles are only affected by active particles and cannot affect any other particle, and are typically dust grains, rocks, small asteroids or spacecrafts. Figure 4.1 summarizes these interactions, where arrows represent directions of gravitational influence.

4.3.2 Algorithm

The HERMES integrator is composed of two parts, a *global* simulation which contains all particles, and a *mini* simulation which contains all active particles plus any semi-active or test particles involved in a close encounter. The global simulation is integrated using WHFAST, while the mini simulation is integrated using IAS15. We first outline the overall algorithm for one timestep[2] of length dt and then describe the individual steps in more detail.

To evolve the system for a single timestep HERMES performs the following steps:

1. **Check for close encounters.** If any particle (active, semi-active, or test) has a close encounter with an active particle, then copy the particles involved in the close encounter plus all active particles to the mini simulation.

2. **Integrate the global simulation** using the WHFAST integrator for one timestep, dt.

3. **Integrate the mini simulation** using the IAS15 integrator if a close encounter was identified in step 1.

4. **Update the particles in the global simulation** using the mini simulation if the mini simulation was active this timestep.

Although the algorithm is simple to write down in the above form, there are several caveats to point out. For all particles excluding the central body, we define the parameter f_{H}, which we dub the Hill Switch Factor. A spherical shell is of radius $f_{\mathrm{H}} r_H$ is constructed around each body, where r_H is the Hill radius and defines the region surrounding each body where its local gravity dominates that of central body. If the shells of any two particles overlap it is deemed a close encounter. Since all semi-active and test particles are invisible to each other they cannot be involved in close encounters with one another. Only particles that gravitationally interact with each other can participate in close encounters (see Fig. 4.1). Whenever there is at least one close-encounter the mini simulation is integrated, if no close-encounters occur, the mini simulation is not active and the integrator defaults to WHFAST. Since the central object has no Hill sphere this motivates us to define the Solar Switch Factor, $f_{\odot}$, which only applies to the central body. Like f_{H} it also defines a spherical shell except is in units of the central object's physical radius instead of Hill radii. Therefore, when a body passes within $f_{\odot} r_{\odot}$ of the central object it is deemed a close encounter, where $r_{\odot}$ is the radius of the central object.

During a close encounter, WHFAST still integrates all particles (including those involved in the close encounter) leading to momentarily large errors for the particles involved in the close encounter. One might expect that this poses a real problem for the accuracy, but that is not the case, since all particles involved in the close encounter plus all active particles are overwritten at the end of the timestep using the accurate results from the mini simulation.

Unlike the global simulation, the mini simulation may take many timesteps to get from t to $t + dt$. The length of the timestep in the mini simulation is automatically determined by the IAS15 integrator. During each sub-timestep the mini simulation also checks for physical collisions between overlapping particles.

As an example of how the mini and global simulations integrate through time, consider a 2 planet, 2 planetesimal system. The planets are active particles and the planetesimals are semi-active particles. Figure 4.2 shows the distance of the planetesimals and planet 2 from planet 1 as a function of time. A point is plotted after every timestep in both the global and mini simulation. After 0.2 years, planetesimal 1 (a semi-active body) has a close encounter with planet 1 (an active body). At that time,

[2] We refer to the timestep that WHFAST takes as dt. The IAS15 integrator chooses its own timestep which is typically smaller than dt.

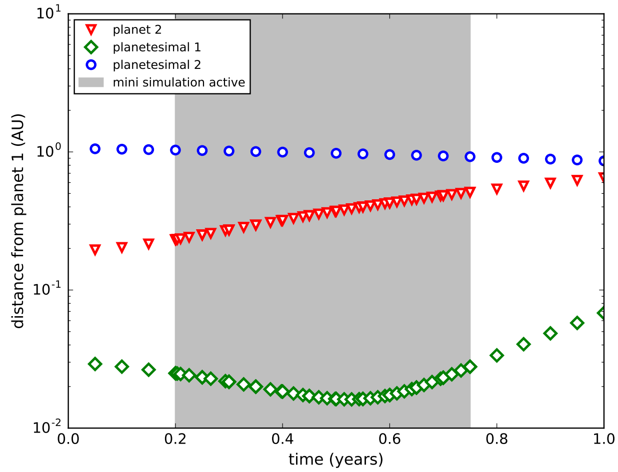

Figure 4.2: A short simulation displaying the HERMES integrator for a 2 planet, 2 planetesimal system orbiting a central star. When active, the mini simulation takes many sub-timesteps for each dt and integrates planets 1, 2 and planetesimal 1 during the close encounter between planet 1 and planetesimal 1.

the mini simulation is turned on and planetesimal 1, planet 1 the central star (not shown) and planet 2 are added to the mini simulation and integrated until 0.75 years, at which point the close encounter between planetesimal 1 and planet 1 is complete. Planetesimal 2 on the other hand continues to be solely integrated by the global simulation (using WHFAST) throughout the close encounter. By comparing the outputs of planetesimal 2 and the other particles in Fig. 4.2, one can see that the mini simulation takes numerous sub-timesteps compared to the global simulation. Since IAS15 is an adaptive method, it automatically chooses the appropriate timestep to resolve the close encounter with machine precision accuracy.

We briefly discuss the speed of the algorithm. If no particles are integrated with IAS15, then the speed is effectively that of WHFAST with a small overhead due to collision checks. If all particles are integrated with IAS15, then the speed is that of a simulation running only IAS15, again with a small and in general negligible overhead due to collision checks. Consider a typical simulation of multiple active particles undergoing planetesimal migration from a large number of semi-active particles with a reasonable f_H value. As the number of semi-active particles is increased, the ratio of the number of particles in the mini simulation, N_{mini}, to the number of particles in the global simulation, N_{global}, approaches a constant. In this limit the elapsed simulation time is linearly proportional to the number of semi-active particles in the simulation.

4.3.3 Perturbative Forces in the Mini Simulation

One must carefully treat the forces perturbing the motions of active particles in the mini simulation. In the global simulation, active particles receive perturbative kicks from all semi-active particles, and it is important to reproduce these forces in the mini simulation, which only evolve a subset of particles. To further complicate this issue, the mini simulation takes numerous sub-timesteps for each global timestep. In HERMES, we linearly interpolate the forces of all semi-active particles absent from the mini simulation using the initial and final values from the global simulation. We find that interpolating

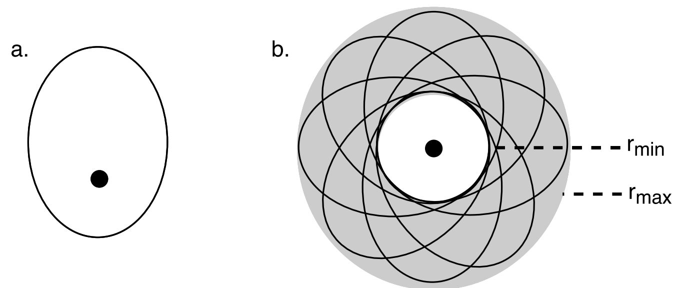

Figure 4.3: Panel a. shows a regular orbit in 2D, while panel b. shows the construction of a ring by rotating the orbit's pericenter by 2π.

the forces (rather than positions) leads to a significant speed gain without compromising noticeable accuracy.

One could argue that this process should be iterated – the active particles will arrive at slightly different (and more accurate) final positions when integrated with the mini simulation, and thus the perturbative forces they would have induced on the semi-active particles in the global simulation would be slightly different too. Thus if one wanted to improve the accuracy of the algorithm an iterative process could be constructed where the global and mini simulations take turns integrating for a timestep dt making use of the updated positions from the previous iteration.

However, as long as the semi-active particles have a much smaller mass than the active particles, we have found that this iterative process is unnecessary, allowing us to reduce both the computation time and algorithmic complexity. We note that in our current non-iterative scheme, interpolating the forces between pairs of *active* particles introduces non-negligible numerical errors. For that reason, all active particles are automatically added to the mini simulation during any close encounter, even if they are not involved in the close encounter themselves.

4.3.4 Adaptive $f_{\rm H}$ Algorithm

$f_{\rm H}$ and dt are the most important parameters to consider when simulating a system with HERMES. For a system free of close encounters, dt alone determines the precision of the algorithm. However during close encounters $f_{\rm H}$ and dt together determine the algorithm's precision (see Section 4.4). Specifically, if a particle moves a distance $\sim f_{\rm H} r_H$ per timestep dt the algorithm could miss a close encounter, introducing large errors into the simulation. In addition, an initial choice of $f_{\rm H}$ and dt can become non-optimal if a system evolves significantly from its initial state.

To aid the user in making the correct parameter choices, we have developed a simple algorithm that, given a timestep dt, conservatively estimates the smallest value of $f_{\rm H}$ under the condition that no close encounter is missed. Although $f_{\rm H}$ and dt both determine the precision of HERMES during a close encounter, we only optimize $f_{\rm H}$ since constantly changing dt would result in non-negligible numerical errors for a symplectic integrator like WHFAST. We calculate the optimal $f_{\rm H}$ each iteration, ensuring that $f_{\rm H}$ adapts to an evolving system and guarantees that close encounters are continuously resolved. The user is therefore only required to set the timestep dt for a standard integration ($f_\odot$ is set to a default value serving most purposes). The full algorithm works as follows.

For each body, we ignore the inclination, and marginalize over the phase and longitude of periapsis from 0 to 2π. As illustrated in Figure 4.3, this smears out the orbit into a ring in the reference plane.

Each ring has a maximum and minimum distance from the central object, r_{min} and r_{max}. We then check if any two interacting particles can possibly have a close encounter by comparing their r_{min} and r_{max} values. If an intersection of two rings occurs, say for particles i and j, we calculate their maximum relative velocity $\Delta v_{ij,\text{max}}$ in the overlapping interval between the two particles. We calculate $\Delta v_{ij,\text{max}}$ as a simple overestimate rather than the true maximum in order to speed up the calculation. As a result of ignoring the inclination and marginalizing over the phase and longitude of periastron, Δv_{max} can be calculated from just the semi-major axis and the eccentricity. We note that we do not need to solve Kepler's equation in estimating Δv_{max}.

We can then calculate f_{H} by taking the maximum over all interacting particle pairs (see Fig 4.1),

$$f_{\text{H}} = 4 \max_{i,j} \frac{\Delta v_{ij,\text{max}} dt}{r_{H,i} + r_{H,j}}.$$

The numerical constant 4 ensures that two particles move at most one quarter of $f_{\text{H}}(r_{H,i} + r_{H,j})$ in one timestep and therefore no close encounters are missed.

In practice, we also round up f_{H} to the nearest 1.25^x where x is an integer to avoid continuous fluctuations in f_{H}. By default, our adaptive f_{H} algorithm is enabled in HERMES. Since the algorithm assumes particles move on approximately coplanar orbits, it may therefore fail at very high mutual inclinations. We note that the above algorithm chooses the smallest value of f_{H} that captures all close encounters; however, small values of f_{H} introduce numerical errors when switching between integrators (see Section 4.4). Therefore, to avoid large errors we set a default lower limit of $f_{\text{H}} = 3$, which is close to the default close encounter boundary for Mercury and SyMBA. The user can specify their own lower limit for f_{H} by setting `ri_hermes.hill_switch_factor` at runtime.

The adaptive f_{H} algorithm can be switched off by setting the variable `ri_hermes.adaptive_hill_switch_factor` to zero. If the adaptive f_{H} algorithm is switched off, setting `ri_hermes.hill_switch_factor` simply defines a constant value of f_{H} for the duration of the simulation (analogous to Mercury and SyMBA).

We decided against devising a similar algorithm for $f_{\odot}$ due to the additional difficulties that can arise. For example, an object in a circular orbit around a planet would be confused as a heliocentric orbit with a very high eccentricity, leading to large relative velocities and an excessive $f_{\odot}$ value.

4.4 Error

Several terms contribute to the relative energy error of an integrator (e.g. Rein & Spiegel 2015):

$$E = E_{\text{floor}} + E_{\text{round}} + E_{\text{bias}} + E_{\text{scheme}}. \tag{4.1}$$

E_{floor} is a constant due to the inability to represent numbers with arbitrary precision on a computer. Here we work exclusively in IEEE754 double floating point precision and thus have $E_{\text{floor}} \sim 10^{-16}$.

E_{round} arises when a computation is performed on two floating point numbers. Almost all operations (addition, multiplication, square roots) lead to a roundoff error at the level of the machine precision. The IEEE754 standard guarantees that the round-off error in consecutive floating point operations is random, thus leading to a $\propto t^{1/2}$ growth of E_{round} with time.

E_{bias} is the error from any biased operations and grows at least as $\propto t$. Biased operations can originate from poor implementations[3] or from library functions that the IEEE754 standard does not guarantee will return unbiased results[4] .

E_{scheme}, the final term in Eq. 4.1, is the error introduced by the algorithm itself. Typically, this quantity is bound for symplectic integrators but grows linearly with time for non-symplectic integrators.

The important question is which error term dominates, and the answer will depend on the problem at hand. For example, if a three-body system (star and two planets) is integrated with IAS15, $E^{ias}_{\text{scheme}} \approx 10^{-28}$ and the dominant error term will be E_{round}, starting at 10^{-16} and growing as $t^{1/2}$ (see Rein & Spiegel 2015). If we instead integrate the system with WHFAST, the dominant error term will be E_{scheme}, which is determined both by the mass ratio in the system and the timestep. For typical parameters $E^{WH}_{\text{scheme}} \sim 10^{-9}$, and only for very long simulation times ($\sim 10^{14}$ timesteps) will the growth of E_{round} dominate over E^{WH}_{scheme}. For biased implementations, E_{bias} will dominate the error budget at earlier times.

For typical simulations integrated with HERMES, E_{scheme} will dominate. The WH algorithm integrates a slightly different Hamiltonian from the true Hamiltonian described by the system, leading to an error that oscillates in a bounded manner as long as the integrated Hamiltonian remains constant. However each time a particle is transferred to or from the global simulation (see Section 4.3.2), the WH-integrated Hamiltonian changes, and thus the error will change too. For a typical WH integration, working in democratic heliocentric coordinates[5], E_{scheme} is (e.g. Saha & Tremaine 1994; Wisdom 2006):

$$E^{\text{WH}}_{\text{scheme}} = \frac{dt^2}{12} \left\{ \left\{ H_K, H_\beta \right\}, 0.5 H_K + H_\beta \right\} + O(dt^4) \tag{4.2}$$

Here H_K is the Keplerian Hamiltonian, $H_\beta = H_C + H_I$ is the summed momentum cross-term and interaction Hamiltonians, respectively, and {} are Poisson brackets (the quantities are explicitly defined for a test case with three particles below). For large N-body systems, Eq. 4.2 quickly becomes very difficult to evaluate analytically. However, we can gain some insight by applying Eq. 4.2 to a simple system.

We turn to a three body problem consisting of a star (active body), planet (active body) and planetesimal (semi-active body), shown in Figure 4.4. The planetesimal is initially placed inside f_{H} with sufficient velocity such that the distance between the planet and planetesimal grows over time. While the planetesimal is inside f_{H} (panel a. in Fig. 4.4) the system is integrated to machine precision by IAS15. However once the planetesimal leaves f_{H} (panel b. in Fig. 4.4) the system switches to being integrated by WHFAST, and an error of size $E^{\text{WH}}_{\text{scheme}}$ is introduced.

To estimate $E^{\text{WH}}_{\text{scheme}}$, we start from the general Hamiltonian for an N-body system in democratic heliocentric coordinates:

$$H = H_0 + H_K + H_C + H_I$$

where H_0 is a constant describing the motion of the centre of mass along a straight line, and disappears

[3]For example, the expression `x*(2./3.)` multiplies x by a number that is consistently slightly too big or too small when represented in binary. By contrast, the expression `2.*(x/3.)` multiplies and divides x by numbers that are exactly representable in binary and is unbiased.

[4]For example, the standard library function `sqrt()` returns an unbiased result whereas `sin()` returns a biased result.

[5]Although Figures 4.5, 4.6, and 4.7 along with the results in Section 4.5 are all performed using HERMES with WHFAST in Jacobi coordinates, we derive the error using democratic heliocentric coordinates. This is because, a) it is much simpler to do, and b) as discussed in Section 7.2 we believe that HERMES with WHFAST in heliocentric coordinates is ultimately a better choice. At the time of this thesis, the transition from HERMES with WHFAST in Jacobi to heliocentric coordinates was underway. However, as shown in Figures 4.5, 4.6 and 4.7 the error derived from democratic heliocentric coordinates (Equation 4.8) agrees well with the numerical experiments performed with WHFAST in Jacobi coordinates.

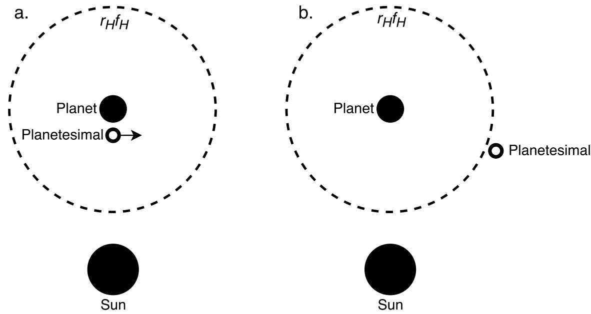

Figure 4.4: Three body problem, in the reference frame of the planet. In a. the initial setup is shown, where the planetesimal starts near the planet, inside a sphere of radius $r_H f_\mathrm{H}$ and the entire system is integrated purely by `IAS15`. Here the arrow indicates the initial direction of the planetesimal. In b. the planetesimal exits the sphere with radius $r_H f_\mathrm{H}$ and the system is integrated purely via `WHFAST`, introducing a numerical error of $E^{\mathrm{WH}}_{\mathrm{scheme}}$.

when we evaluate Eq 4.2. The remaining terms in Eq. 4.4 take the form (Duncan et al. 1998):

$$H_K = \sum_{i=1}^{N-1} \frac{\mathbf{P_i^2}}{2m_i} - \sum_{i=1}^{N-1} \frac{Gm_0 m_i}{|\mathbf{Q_i}|} \tag{4.3}$$

$$H_C = \frac{1}{2m_0} \left| \sum_{i=1}^{N} \mathbf{P_i} \right|^2 \tag{4.4}$$

$$H_I = - \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{Gm_i m_j}{|\mathbf{Q_i} - \mathbf{Q_j}|} \tag{4.5}$$

where the canonical coordinates $\mathbf{Q}$ and $\mathbf{P}$ are:

$$\mathbf{Q_i} = \begin{cases} \mathbf{r_i} - \mathbf{r_0} & \text{if } i \neq 0 \\ \frac{1}{m_{\mathrm{tot}}} \sum_{j=0}^{N} m_j r_j & \text{if } i = 0 \end{cases} \tag{4.6}$$

$$\mathbf{P_i} = \begin{cases} \mathbf{p_i} - \frac{\mathbf{m_i}}{\mathbf{m_{tot}}} \sum_{\mathbf{j=0}}^{\mathbf{N}} \mathbf{p_j} & \text{if } i \neq 0 \\ \sum_{j=0}^{N} \mathbf{p_j} & \text{if } i = 0 \end{cases} \tag{4.7}$$

Here $\mathbf{p}$, $\mathbf{r}$, m are a particle's momentum, position and mass in any inertial frame respectively, while G is the gravitational constant and $m_{\mathrm{tot}} = \sum_{j=0}^{N} m_j$ is the total mass of the system. $\mathbf{Q_i}$ are therefore heliocentric positions (with $\mathbf{Q_0}$ the centre of mass), while $\mathbf{P_i}$ are barycentric momenta (with $\mathbf{P_0}$ the momentum of the centre of mass).

We also further simplify the system to two dimensions by considering motion in a plane. One can then straightforwardly, albeit tediously, evaluate the Poisson bracket in Eq. 4.2 by plugging in Eqs. 4.3–4.5 and taking derivatives with respect to all three particles. We make the further simplifying assumptions that $m_0 \gg m_2 \gg m_1$ and that $v_1 \approx v_2 \approx \sqrt{Gm_0/a}$, where v denotes particle velocities and a_2 is the semi-major axis of the planet. In addition, after solving Eq. 4.2 we set the distance of the planetesimal from the planet to $r_H f_\mathrm{H}$, where r_H is the Hill radius of the planet. This is true at the moment the integration method is switched. We are then left with a single dominating term,

$$E^{\mathtt{HERMES}}_{\mathrm{scheme}} \approx \frac{dt^2}{12} \frac{G^2 m_0 m_1 m_2}{a_2 (r_H f_\mathrm{H})^3}, \tag{4.8}$$

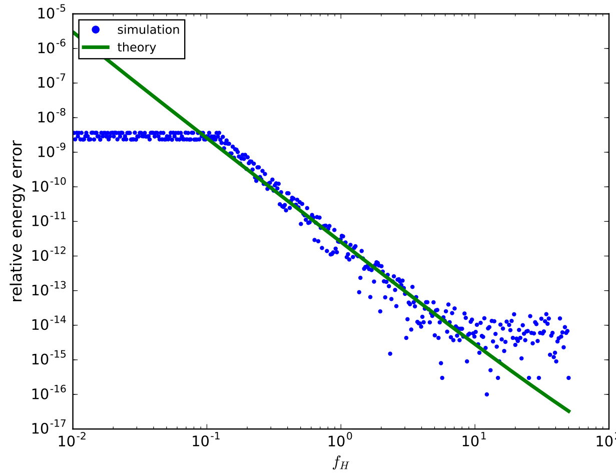

Figure 4.5: Final relative energy error as a function of f_H for a star-planet-planetesimal system. Blue dots are numerical simulations, the green curve is the theoretical prediction of Eq. 4.8.

where we have ignored numerical constants of order unity. An IPython notebook with a computer derivation is available at `https://github.com/silburt/hermes_ipython`.

We compare our theoretical predictions in Eq. 4.8 to numerical tests in Figures 4.5 and 4.6. Our numerical setup consists of a star with mass $1M_\odot$, a Neptune mass planet on a circular orbit at 1 AU, and a planetesimal with mass $10^{-8}M_\odot$ placed at 0.001 AU from the planet. To marginalize over the phase of the encounter when sampling the energy error, the initial position of the planetesimal is randomized for each realization. In addition, the planetesimal is given a small kick equal to the escape velocity of the planet to ensure that the planetesimal-planet distance increases with time. Each realization is simulated for 7 years. When varying the timestep in Fig. 4.6 we use a constant $f_H = 6$, and when varying f_H in Fig. 4.5 we use a constant timestep of $dt = 0.058$ days.

In Figure 4.5 one can see that the numerical tests agree with the theoretical predictions of Eq. 4.8. In addition, one can see the two extreme regimes on each end of the figure. For particles starting outside the $r_H f_H$ boundary the integrator always uses `WHFAST` (like panel b. of Fig. 4.4), while for particles inside the $r_H f_H$ boundary the simulation uses `IAS15` (like panel a. of Fig. 4.4). In Figure 4.6 one can see that the relative energy error is proportional to dt^2, again matching the predictions of Eq. 4.8. For both Fig. 4.5 and Fig. 4.6, we have not performed any kind of fit, we simply over-plotted Eq. 4.8 with the data from our numerical experiments. We have performed other suites of simulations testing how the energy error scales with all other relevant quantities (semi-major axis, planetesimal mass, planet mass, stellar mass) and find Eq. 4.8 in good agreement. An IPython notebook for these experiments (including tests of the other relevant quantities) is available at `https://github.com/silburt/hermes_ipython`.

We now extend the characterization of the error to a more realistic case with N particles. We refer to the total relative energy error for an integration as $E^{\texttt{HERMES}}_{\text{scheme,tot}}$. Since an error of size $E^{\texttt{HERMES}}_{\text{scheme}}$ is introduced each time a particle leaves/enters the global simulation, the total errors should be related to the number of close encounters, N_{CE}. In the ideal case where the integrator is unbiased, the error $E^{\texttt{HERMES}}_{\text{scheme}}$ introduced by each close encounter is random, and $E^{\texttt{HERMES}}_{\text{scheme,tot}}$ will grow as a $N_{\text{CE}}^{1/2}$ random walk. However, if `HERMES` is biased (i.e. $E^{\texttt{HERMES}}_{\text{scheme}}$ is not random), then $E^{\texttt{HERMES}}_{\text{scheme,tot}}$ will grow faster than $N_{\text{CE}}^{1/2}$.

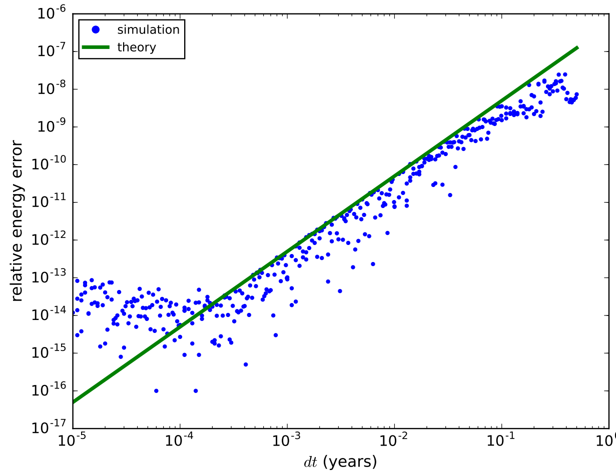

Figure 4.6: Final energy error as a function of dt for a star-planet-planetesimal system. Blue dots are numerical simulations, the green curve is the theoretical prediction of Eq. 4.8.

Assuming the unbiased case, $E^{\texttt{HERMES}}_{\text{scheme,tot}}$ is equal to:

$$E^{\texttt{HERMES}}_{\text{scheme,tot}} = K \cdot E^{\texttt{HERMES}}_{\text{scheme}} \cdot \sqrt{N_{\text{CE}}} \tag{4.9}$$

where K is a constant of proportionality. We test Eq. 4.9 against numerical tests for a Solar mass star, a Neptune mass planet on a circular orbit at 1 AU, and a disk of 200 planetesimals located between $0.98 - 1.02$ AU. The initial inclinations and eccentricities of the planetesimals in the disk are set to 0, while the argument of perihelion and true anomaly are drawn from a uniform distribution. In addition, for these simulations we set $f_{\text{H}} = 6$ and $dt = 0.015$ years. We performed numerous integrations, integrating each realization for a randomly chosen number of orbital periods between 10-1000, yielding different numbers of close encounters.

The results are shown in Figure 4.7, the x-axis showing the number of close encounters during a simulation and the y-axis showing the final relative energy error for each simulation. We fit a power-law distribution to the data using Python's Scipy Optimize Curve Fit package (Peterson 2009), displayed as a green line in Fig. 4.7. The resulting fit is $E^{\texttt{HERMES}}_{\text{scheme,tot}} = 15.5 \cdot E^{\texttt{HERMES}}_{\text{scheme}} \cdot N^{0.53}_{\text{CE}}$, so the energy growth is well approximated by Eq. 4.9. For reference, we plot the biased prediction of $E^{\texttt{HERMES}}_{\text{scheme,tot}} \propto N_{\text{CE}}$ as a red line. We conclude that HERMES is unbiased for this setup. Thus, Eq. 4.9 provides an intuitive way of understanding how the error of HERMES grows without having to analytically solve Eq. 4.2 in three dimensions for N particles.

The results from Fig. 4.7 did not allow for physical collisions between particles. When physical collision are enabled a systematic bias can be introduced.

To see why, note that for each close encounter two contributions to the energy error arise according to Eq. 4.8; one when the particles are transferred from the global to the mini simulation (i.e. the ingress of the close encounter) and one when the particles are transferred back from the mini to the global simulation (i.e. the egress of the close encounter). The precise energy change depends on the specific properties of the system (phases of the orbits, angles of approach, etc.), and typically the energy changes associated with the ingress and egress of the close encounter are anticorrelated. As a result,

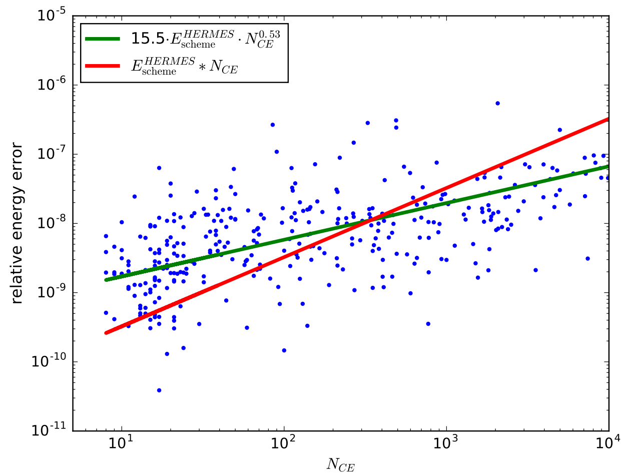

Figure 4.7: Final relative energy error as a function of the number of close encounters, for a system composed of a star, planet and 200 planetesimals. Blue dots are numerical simulations, the green line is our unbiased theoretical prediction of Eq. 4.9 with $K = 15$, while the red line is the biased theoretical prediction.

no appreciable energy error is introduced by close encounters, and the energy over the course of a simulation grows as expected according to Eq. 4.9.

However, when a physical collision occurs in HERMES, the close encounter only has an ingress, resulting in a biased growth in the energy error. In practice this energy bias is orders of magnitude smaller than the physical energy lost during a collision, and therefore should not interfere with the longterm evolution of a system. We plan to study this issue in more detail in the future.

4.5 Examples

HERMES is well suited for a number of astrophysical problems, including planets embedded in planetesimal disks, high eccentricity comets, planet-planet scattering, star clusters orbiting a central potential, and more. Many example problems can be found at `https://github.com/hannorein/rebound/tree/master/ipython_examples`. Below, we highlight just a few example problems that can be simulated with HERMES.

4.5.1 Massive Outer Solar System

Both Duncan et al. (1998) and Chambers (1999) simulated the outer Solar System, but increased the masses of all planets by a factor of 50 to trigger close encounters between planets. Analogous to Chambers (1999) and Duncan et al. (1998) we use $f_{\mathrm{H}} = 3$, $dt = 0.03$ yrs, and integrate the system for 1000 years.

We perform a number of simulations of the massive outer Solar System, and find that the relative energy error stays bounded at $\sim 10^{-7}$ for all simulations, matching the results of Chambers (1999) and Duncan et al. (1998).

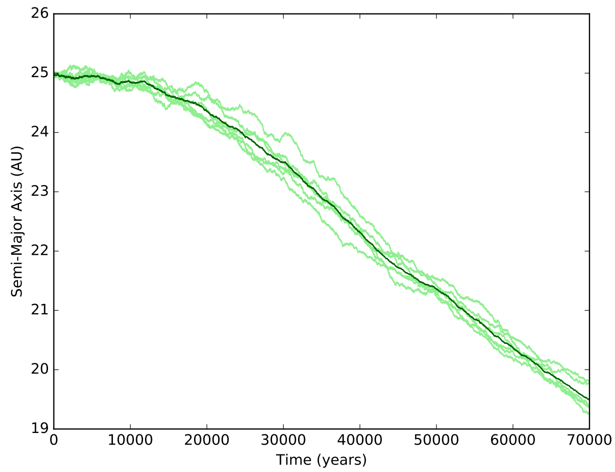

Figure 4.8: Planet's semimajor axis vs. time, analogous to the numerical experiment in the lower right panel of Figure 3 from Kirsh et al. (2009). Light green lines represent individual runs, while the dark thicker green line represents the average of the individual runs.

4.5.2 Migration of a Planet in Planetesimal Disk (Kirsh et al. 2009)

Here we reproduce the results of Kirsh et al. (2009) for the migration of a single planet embedded in a planetesimal disk. In this study a $2.3M_\oplus$ planet orbits a Solar mass star at 25 AU, embedded in a disk of $\sim 6 \cdot 10^4$ planetesimals. The planetesimal disk extends 10.5 AU on each side of the planet, each planetesimal has a mass 1/600th of the planet, and an overall surface density profile proportional to a^{-1} is used. The radii of all orbiting particles were determined assuming a constant density of 2 g/cm^3. The eccentricities and inclinations were drawn from a Rayleigh distribution, which is parameterized by the scale parameter σ. For this experiment we use $\sigma_e = 0.01$, and $\sigma_i = 0.005$, where inclination is in radians. The argument of periapse, true anomaly, and longitude of ascending node were all randomly drawn from uniform distributions over $[0, 2\pi]$. Adopting the default settings of Kirsh et al. (2009), we set `HERMES` to merge particles inelastically, set $dt = 2$ years, $f_\mathrm{H} = 5$, $f_\odot = 15$, and track the energy lost due to collisions or ejections. We also set `r->ri_hermes.adaptive_hill_switch_factor` = 1, activating the adaptive Hill switch routine (Section 4.3.4). An IPython notebook containing the code to run this example can be found at `https://github.com/silburt/hermes_ipython`.

We run 6 separate simulations, each for 70,000 years, and plot the results in Fig. 4.8. The relevant comparison plot in Kirsh et al. (2009) is the lower right panel in Figure 3. For all runs in both Kirsh et al. (2009) and our work the final position of the planet is between $18.5 < a < 20$ AU, and show similar evolution tracks throughout the simulation. The relative energy error over the course of all our simulations do not exceed $2 \cdot 10^{-7}$ after accounting for the energy lost in inelastic collisions.

4.5.3 Comparison to `Mercury` and `SyMBA`– Long Simulations

Here we perform a set of long simulations and compare our results to `Mercury` and `SyMBA`. These integrators are also capable of integrating complex N-body systems with close encounters, and also make use of semi-active particles. The simulations for these tests contain a Solar-mass star, a Neptune-mass planet at $a = 1$ AU and a disk of 100 semi-active planetesimals distributed according to a powerlaw

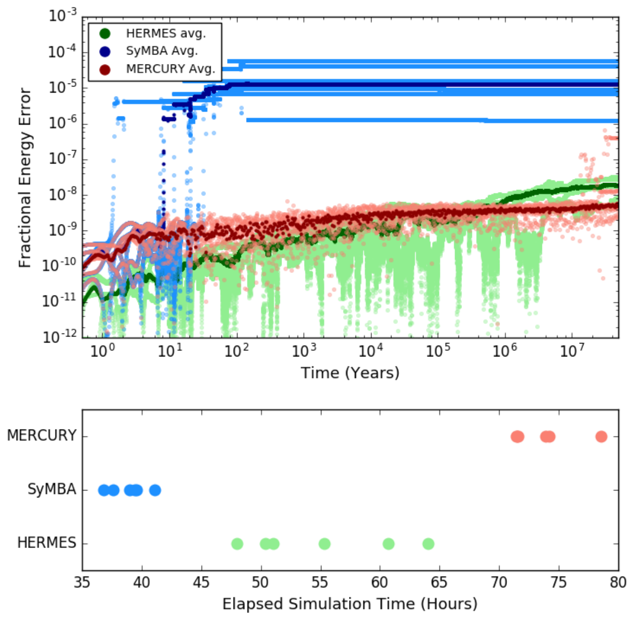

Figure 4.9: A test of HERMES, Mercury and SyMBA for collections of 50Myr, 50 planetesimal runs. Top panel shows the relative energy error over time, with individual runs in lighter shades and averaged values in dark shades. Bottom panel shows the elapsed simulation times of individual runs, using the same colour scheme as the top panel.

between $0.8 - 1.2$ AU. The mass of each planetesimal is a third of a lunar mass, the eccentricities and inclinations are set to 0, and the argument of periapse, true anomaly, and longitude of ascending node were all randomly drawn from uniform distributions over $[0, 2\pi]$. We set $dt = 0.01$, $f_{\mathrm{H}} = 3$, merge particles inelastically, and use our adaptive f_{H} routine (Section 4.3.4). In addition, for all three integrators we track the energy lost due to inelastic collisions and ejections so that we can isolate the numerical energy error. We accomplish this by using the `eoffset` variable in SyMBA and `EN(3)` variable in Mercury, and calculate the relative energy error according to $(E_i + E_{\mathrm{off}} - E_0)/E_0$, where E_i is the total energy at iteration i, E_0 is the initial total energy and E_{off} is the energy lost due to collisions and ejections, i.e. `eoffset` in SyMBA and `EN(3)` in Mercury.

We feed identical initial conditions to SyMBA, Mercury and HERMES, and evolve all simulations for 50 Myr. We run 6 simulations per integrator that differ only by the seed of the random number generator. We average the relative energy error for these simulations to smooth out variations between individual runs. The results are presented in Figure 4.9. The top panel displays the relative energy error with each integrator, while the bottom panel shows the elapsed time for each individual run.

Looking at the top panel, SyMBA incurs significant energy jumps early in the simulation. We suspect these energy jumps are due to inadequately resolved close encounters, since the energy jumps are uncorrelated with particle collisions. We simulated runs using different combinations of RHSCALE and RSHELL, i.e. the parameters which define close encounter regions, but were unable to resolve these energy jumps. By the end of the simulation, SyMBA is orders of magnitude less accurate than Mercury and HERMES.

We see that HERMES initially has the lowest energy error, but grows faster than Mercury. The final energy error for an integration therefore depends on the particular problem (length of simulation, timestep, number and severity of close encounters etc.). Furthermore, we find that some Mercury simulations undergo significant energy jumps, probably also due to very close encounters that are not properly resolved. Our adaptive $f_{\rm H}$ algorithm (Section 4.3.4) protects against these situations from occurring, and for a properly chosen $f_{\rm H}$ and dt combination we find the energy growth of HERMES to be well behaved.

The bottom panel shows that for these simulations HERMES is slower than SyMBA but faster than Mercury. We see that HERMES exhibits a larger variance in elapsed simulation times than Mercury and SyMBA. This is because when the adaptive $f_{\rm H}$ routine is engaged, $f_{\rm H}$ can be enlarged considerably during severe encounters, slowing down the integrator. In summary, these simulations show that HERMES provides a good balance between speed and accuracy.

4.6 Conclusion

In this Chapter we have presented HERMES, a hybrid integrator capable of integrating close encounters and collisions. HERMES integrator is composed of two parts, a global simulation which contains all particles and a mini simulation which contains all active particles (e.g. stars and planets) plus any semi-active/test particles (e.g., planetesimals) involved in a close encounter. The global simulation is integrated with WHFAST while the mini simulation is integrated with IAS15. Comparing HERMES to the openly available Mercury and SyMBA we find that HERMES provides a good balance between accuracy and speed.

HERMES takes a more direct approach when integrating close encounters over other methods. However this has enabled us to characterize the error of HERMES, and we find that the switching error from a single close encounter is well described by Eq. 4.8, which we calculate from first principles. The total energy error of HERMES for an N-body simulation is well described by a random walk, (see Eq. 4.9) with the step size being the switching error.

We have also developed an adaptive algorithm that chooses the optimal Hill switch factor, $f_{\rm H}$, which governs the size of the close encounter region surrounding each particle. This frees the user from optimizing integrator parameters for typical cases. Finally, we have introduced a number of new features in REBOUND, including semi-active particles (see Fig. 4.1) and inelastic collisions.

We have showcased a number of problems well-suited for HERMES, and compared our integrator's performance to similar integrators in the literature. In particular, we integrated the outer Solar System with planetary masses increased by a factor of 50, and we simulated a planet migrating through a disk of planetesimals, both in the limit of many planetesimals ($\sim 10^5$) and short times ($\sim 10^3$ orbits), and few planetesimals (100) and long times ($\sim 10^8$ orbits). Many more types of problems are possible with HERMES, and additional examples can be found at https://github.com/hannorein/rebound.

Chapter 5

A Machine Learns to Predict the Stability of Tightly Packed Planetary Systems

5.1 Chapter Overview

The work, tables and figures from this Chapter are based off the following publication:

Daniel Tamayo, Ari Silburt, Diana Valencia, Kristen Menou, Mohamad Ali-Dib, Cristobal Petrovich, Chelsea X. Huang, Hanno Rein, Christa van Laerhoven, Adiv Paradise, Alysa Obertas, Norman Murray.
The Astrophysical Journal Letters, Volume 832, Issue 2, article id. L22, pp. 5, 2016 (Tamayo et al. 2016).

The requirement that planetary systems be dynamically stable is often used to vet new discoveries or set limits on unconstrained masses or orbital elements. This is typically carried out via computationally expensive N-body simulations. We show that characterizing the complicated and multi-dimensional stability boundary of tightly packed systems is amenable to machine learning methods. We find that training a state-of-the-art machine learning algorithm on physically motivated features yields an accurate classifier of stability in packed systems. On the stability timescale investigated (10^7 orbits), it is 3 orders of magnitude faster than direct N-body simulations. Optimized machine learning classifiers for dynamical stability may thus prove useful across the discipline, e.g., to characterize the exoplanet sample discovered by the upcoming Transiting Exoplanet Survey Satellite (TESS).

5.2 Introduction

In order to characterize planetary systems, it is common practice to assume long-term stability in order to set upper limits on planetary masses and orbital eccentricities (e.g. Lissauer et al. 2011; Steffen et al.

2013; Tamayo 2014; Tamayo et al. 2015). This involves running grids of direct N-body integrations over the large multi-dimensional parameter space of initial conditions that are consistent with observational error. In practice, however, one can often explore only a minute fraction of the phase space; in the case of many systems discovered by the Kepler mission, each integration requires several weeks of computation to simulate timescales comparable to the star's age ($\gtrsim 10^{11}$ planetary orbits) with current hardware. An efficient classifier of dynamical stability would thus be invaluable.

In this investigation, we specialize to the case of tightly packed systems, where the interplanetary separations are less than 10 mutual Hill radii (R_H), where

$$R_H \equiv \left(\frac{M_1 + M_2}{3M_\star}\right)^{1/3} a_1, \tag{5.1}$$

M_1, M_2 and $M_\star$ are the masses of a pair of planets and the central star, and a_1 is the semimajor axis of the inner planet. This regime has long been recognized as important in the early stages of planetary systems when bodies are still merging, and has received much attention.

For the special case of two-planet systems, one can prove that there exists a bifurcation in the dynamics that forbids planets from undergoing close encounters for all time (Marchal & Bozis 1982; Milani & Nobili 1983). In the case of co-planar, circular orbits, the transition occurs at a planetary separation of 3.46 R_H (Gladman 1993). If the planets are more widely separated than this threshold, the system is Hill stable, i.e., close encounters are forbidden for all time. While there are ways for planets to escape or collide with the star without undergoing close encounters (e.g., Deck et al. 2012; Veras & Mustill 2013), and eccentric planets that fail this criterion can nevertheless be long-lived (e.g., Gladman 1993), the Hill criterion provides good guidance, particularly in the low-eccentricity regime (Barnes & Greenberg 2006).

In the general case with more than two planets, however, the additional degrees of freedom preclude a topological criterion for Hill stability. This has led many authors to perform suites of N-body simulations and fit empirical curves to the results. Because of the prohibitively large phase space of possible initial conditions, several authors (Chambers et al. 1996a; Faber & Quillen 2007; Smith & Lissauer 2009, Obertas et al., submitted) have considered the special case of initially circular, planar orbits with all planet pairs having equal Hill separations Δ, defined as the difference in semimajor axis divided by R_H. They find that even in cases where each pair of planets satisfies the Hill criterion, the system nevertheless destabilizes, and that the associated timescale seems to grow exponentially with Hill separation. However, the fitted coefficients vary as the number of planets and planetary masses are varied, and introducing inclinations (Marzari & Weidenschilling 2002), eccentricities (Ito & Tanikawa 1999; Chatterjee et al. 2008; Pu & Wu 2015), or unequal spacings between planets (Marzari 2014) change the instability timescales substantially. For a given planetary system, it is therefore not always clear which scaling law is appropriate to apply, and what confidence one can have in the resulting estimate.

For this investigation we take a machine learning approach. High dimensional classification tasks are ubiquitous across industry and data science, and sophisticated machine learning algorithms have been developed to tackle these problems. Such techniques have been highly successful in an astronomical context for several image classification tasks, e.g., assigning morphological types to galaxies (Collister & Lahav 2004); however, they have seen little use in dynamical classification to date(see Petrovich 2015, for a recent counterexample).

5.3 Methods

We choose to frame the problem as a binary classification task, i.e., predicting whether or not a given planetary system is stable (over a given timescale). Each “example" (planetary system) is described by a set of “features" that the algorithm uses to predict stability, in the form of a probability between 0 and 1. In supervised machine learning, an algorithm is first trained on examples where it is told the correct answer (stable or not stable). The trained algorithm can then be used to predict on new examples.

5.3.1 Dataset

In order to train our algorithms, we generated a dataset of 5000 N-body integrations of 3-planet systems over 10^7 orbits of the innermost body. We focus on 3-planet systems since there exists an analytic criterion for the case of two planets, and systems with more planets exhibit qualitatively similar behavior. The number of simulations and length of integration were chosen to generate a dataset at limited computational cost, and assess the value of investing significant computing time to train classifiers on astrophysically relevant timescales of $\sim 10^9$ orbits. Because we expect that instability timescales of 10^7 and 10^9 orbits are both physically driven by Chirikov diffusion due to the overlap of mean-motion resonances (see Fig. 2 in Obertas et al., submitted), we expect the performance of models trained on this dataset to be comparable to that of similar algorithms trained on longer datasets.

With a view toward applying these models to Kepler discoveries, we adopted a solar-mass star, 5 $M_\oplus$ (Earth-mass) planets, and drew the innermost planet’s semimajor axis randomly between 0.04 and 0.06 AU. We note, however, that our results are strictly scale-free, and can be applied to comparable systems with masses, orbital periods and semimajor axes expressed in terms of the star’s mass, innermost planet’s orbital period and semimajor axis, respectively. The second planet’s semimajor axis was separated from the first by a number of mutual Hill radii drawn from a uniform distribution in the range $[5, 9]$. The third planet’s separation was then independently drawn from the same distribution, yielding unevenly spaced planets. The particular range of Hill-radius separations was chosen to capture the regime of interest on our adopted timescale of 10^7 orbits and roughly generate a balanced dataset of stable and unstable systems (this yielded 1479 stable systems out of 5000). Eccentricities and inclinations were drawn independently for each planet from uniform distributions over [0, 0.02] and [0, 1°], and the remaining angles were drawn randomly over [0,2π].

All integrations were performed using the `WHFAST` integrator (Rein & Tamayo 2015) in the open-source `REBOUND` N-body package (Rein & Liu 2012), which is written in `C99` and comes with an optional `python` interface. We adopted a timestep of 1% of the innermost planet’s orbital period, and classified systems as unstable if any pair of planets came within 1 Hill radius of each other during the simulation.

5.3.2 Metrics

Binary classifiers are often evaluated on their *precision* (here the fraction of systems that are actually stable when the model predicts stability) and *recall* (the fraction of systems the model predicts are stable out of the truly stable cases). For typical algorithms that predict probabilities of class membership, one can trade off between precision and recall by varying the threshold probability for classification. For example, a conservative model that only classifies systems as stable if it predicts a probability of stability greater than 0.99 will be right most of the times that it predicts stability (high precision), but will miss all the stable systems that were assigned slightly lower probabilities (low recall). The appropriate threshold depends on the application (e.g., if predicting DNA matches for crime cases, one might set a high threshold as above to have confidence in predicted matches).

When comparing two models, one can plot pairs of precision and recall scores for all possible thresholds to generate a precision-recall curve. A common scalar metric for comparing classifiers is the area under this curve (AUC), which would be unity for a perfect model.

5.3.3 Algorithm Training

After experimenting with several machine learning algorithms (random forest and support-vector machine implementations in the `Python scikit-learn` library), we found that gradient-boosted decision trees (GBDT[1]) `XGBoost v0.6` (Chen & Guestrin 2016) gave the best results for our dataset.

A recurring theme in machine learning is that of "overfitting," an algorithm's tendency to latch onto irrelevant idiosyncrasies in the training set that cause it to predict poorly on unseen examples. Different algorithms therefore try to penalize overly complicated models in an effort to retain only the broad features that are likely to generalize well. In practice, the user navigates this balance between simplicity and complexity empirically, by tuning an algorithm's "hyperparameters" that mediate this tradeoff, training it, and checking performance (Sec. 5.3.2) on unseen data; this process, together with trying different features to maximize performance, is known as cross-validation. In order to rule out the possibility of (sometimes subtle) mistakes in cross-validation yielding overly optimistic performance metrics, it can be good practice to assign a subset of the data to a holdout (test) set that is never seen by the algorithm during cross-validation. Evaluation of the final model on the holdout set therefore provides robust metrics of the trained algorithm's expected performance on unseen examples; consistency between the cross-validation and test scores also suggest a robust cross-validation methodology. In our case, we randomly assigned 1500 systems to a holdout test set, and used only the remaining 3500 for cross-validation.

A typical technique to reduce statistical fluctuations when comparing the performance of different sets of hyperparameters is k-fold cross validation. One begins by splitting the training examples into k evenly sized groups; then, for each group, one trains the model on the remaining $k-1$ chunks, and uses the remaining group to evaluate performance. The scores from the k folds are then averaged, reducing the variance in the estimate. Finally, it is generally good practice to use stratified cross-validation, whereby one ensures that each of the k folds is assigned an approximately equal number of samples from each class (stable and unstable).

`XGBoost` has several hyperparameters, so we sequentially performed grid searches through 2-dimensional cuts through the parameter space, evaluating performance through the precision-recall AUC (Sec. 5.3.2) using stratified, 5-fold cross-validation on the training set of 3500 examples. Values of the final adopted hyperparameters for the algorithm are discussed in Sec. 5.4.1 and 5.4.2 can be found in Table 5.1.

5.4 Results

5.4.1 Model 1: Learning from Initial conditions

We begin by considering as features the initial orbital elements and orbital period of each planet, and the interplanetary separations between adjacent planets in units of mutual Hill radii. We then trained

[1]GBDT algorithms create and combine large numbers of individually weak but complementary classifiers to yield a robust estimator (Friedman et al. 2001)

Table 5.1: Hyperparameters used for the initial-conditions (IC) model (Sec. 5.4.1) and short-integrations (SI) model (Sec. 5.4.2), and their associated performance.

	IC Model	SI Model
`base_score`	0.5	0.5
`colsample_bylevel`	1	1
`colsample_bytree`	1	1
`gamma`	0	0
`learning_rate`	0.001	0.00359
`max_delta_step`	0	0
`max_depth`	6	8
`min_child_weight`	1.0	1.2
`missing`	None	None
`n_estimators`	5000	5000
`objective`	binary:logistic	binary:logistic
`reg_alpha`	0	0
`reg_lambda`	1	1
`scale_pos_weight`	1	1
`seed`	27	27
`subsample`	0.4	0.5
AUC (Cross-Validation)	0.842 ± 0.012	0.911 ± 0.012
AUC(Test Dataset)	0.843	0.898
Recall (At 90% Precision)	0.52	0.68

an `XGBoost` classifier on these features (Section 5.3.3), allowing us to predict the stability of each system in the test set in the form of an estimated probability.

As discussed in Section 5.3.2, a threshold probability is required for classifying a system as stable/unstable, and is a subjective choice that depends on the desired qualities of the classifier. For our purposes we argue it is logical to adopt a conservative threshold, in the sense that if the model predicts stability, there is a strong likelihood that the system is actually stable (high precision). This follows from the fact that it is computationally much faster to verify that a system is unstable (on short timescales) than it is to check that it is stable (on long timescales). We choose to require a precision of 90% on our test dataset, which corresponded to the model only classifying systems as stable if their predicted stability probability is larger than a 0.785 threshold.

Since previous works have identified the Hill separations between adjacent planets as important features (Chambers et al. 1996a; Marzari 2014), we plot the performance of the model projected onto this 2D plane (Fig. 5.1).

Looking at the dashed, black line in Figure 5.1, one can see that to first order, the model's prediction boundary roughly obeys the relation $\Delta_1 + \Delta_2 > 16.1$. This is in fact the form of the simple criterion suggested by Lissauer et al. (2011), who quote $\Delta_1 + \Delta_2 > 18$ for stability on timescales of 10^9 orbits. Because we consider stability on shorter timescales, the threshold number of Hill radii should be adjusted for a fair comparison. We term models of the form $\Delta_1 + \Delta_2 > x$ "Lissauer-family" models, and find 16.1 is the threshold for Lissauer-family models that yields a precision of 90% on this dataset.

As stated above, a conservative probability threshold has the disadvantage that the model will misclassify many stable systems as unstable (low recall). This is easily seen by considering different Lissauer-family models (dashed black line in Fig. 5.1), i.e. imposing different threshold values than 16.1 and generating lines parallel to the one plotted. Larger threshold values ensure that a larger fraction of

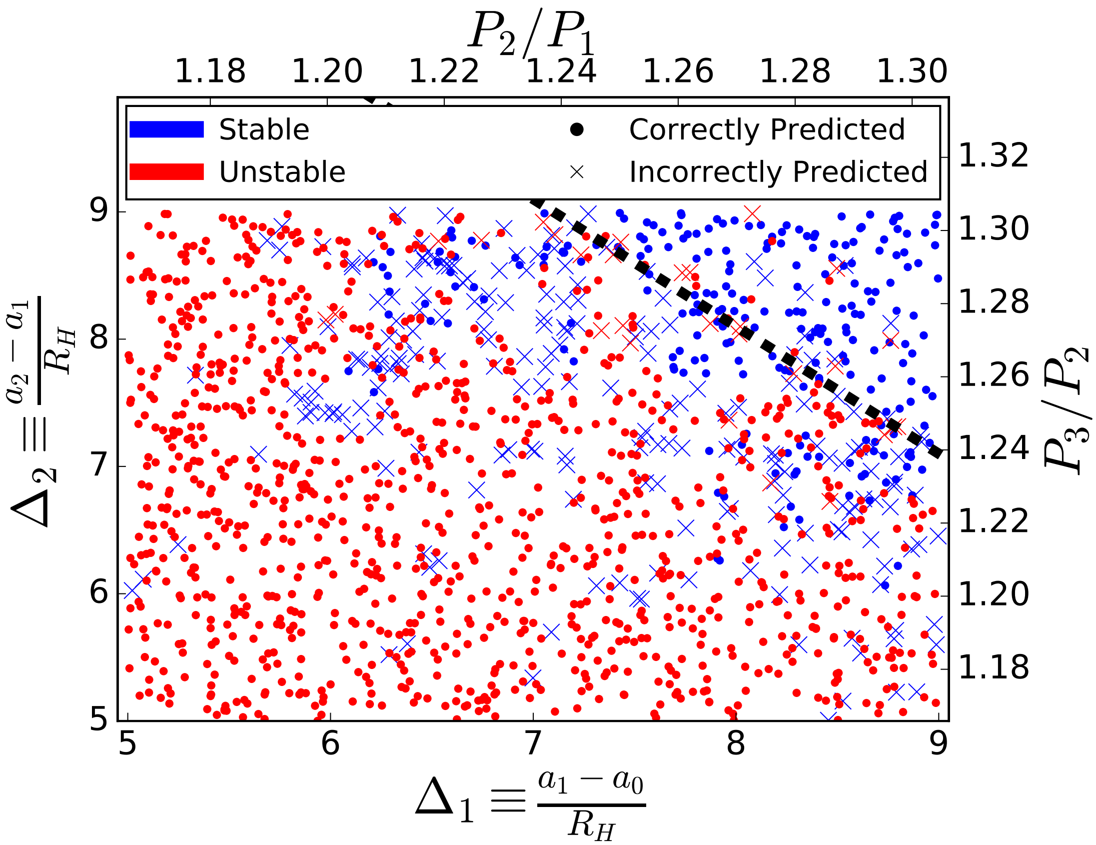

Figure 5.1: Performance on the test dataset using the machine learning model trained on system's initial conditions. Stable systems are marked blue, unstable systems are marked red. Correctly classified systems are plotted as circles, incorrect predictions are marked as crosses. The bottom and left axes show Hill-sphere separations Δ for the inner and outer planet pairs, respectively. The top and right axes correspond to period ratios between the planet pairs. The dashed black line corresponds to the Lissauer-family model $\Delta_1 + \Delta_2 > 16.1$.

the systems to the right of the boundary (i.e., those predicted stable) are in fact stable (blue), leading to higher precision. However, this lowers the recall, since now fewer of the stable systems (blue) are predicted to be stable by the model (i.e., lie to the right of the line). For a fixed precision of 90%, the machine learning model has a significantly higher recall (52%) than the Lissauer-family model (30%). This is because the machine learning model can use information in the features not visible in this 2D projection to make better predictions.

5.4.2 Model 2: Generating Features from Short Integrations

An important factor determining the performance of a machine learning algorithm is the quality of the features it is provided for each of the training examples. To this end, we improved upon the previous model (Section 5.4.1) by generating new features from short N-body integrations. To create the new features, we performed simulations over 10^4 orbits (0.1% of the instability timescale probed) for each of the 5000 systems in the dataset, and recorded each planet's orbital elements and the current Lyapunov timescale every 5 orbits.

Because we suspect that the instability is driven by overlapping mean-motion resonances, we first generated features that capture the variation in semimajor axes, which would vanish if the dynamics were purely secular (Murray & Dermott 1999). In particular, we generated features for the standard deviation and maximum value of each planet's semimajor axis over the 10^4 orbits, normalized to the mean value over the same period (`std_ai` and `max_ai`, where `i` denotes the planet number) . In addition, we generated features for the same quantities over only the first 50 orbital periods in order to capture the variations on orbital timescales (`std_window_ai` and `max_window_ai`). Furthermore, we noticed in early attempts that some misclassified systems exhibited drifts in semimajor axis, so we generated slope features from linear fits to each of the three planets' semimajor axes, normalized to the mean semimajor axis divided by the integration time (`slope_ai`). For the eccentricities of each planet, we generated features for the mean, standard deviation, maximum and minimum values over the full 10^4 orbits, and normalized them to the eccentricity that planet would require to reach its nearest neighbor's semimajor axis (`avg_ei`, `std_ei`, `max_ei`, `min_ei`). For the Lyapunov time we generated a single feature corresponding to the value measured at the end of the integration, normalized to the innermost orbital period. Finally, we added features for the two pairs of initial Hill-radius separations, and for the minimum and maximum initial Hill-radius separations. We experimented with adding features involving the planetary inclinations, but they did not significantly improve the models.

A summary of all the features used can be found in Table 5.2, which are ordered by their importance. We quantify the importances through the `Gain` value recorded by `XGBoost`, which corresponds to the gain in accuracy that a given feature provides when it is introduced into the underlying decision trees used by `XGBoost`. The units are normalized so that the gains sum to 100. We find that the variations in the middle planet's semimajor axis are the most informative in this sense. We interpret this as suggesting that the instabilities in these closely packed systems are driven by the overlap of mean motion resonances (which change the semimajor axes), rather than secular effects (which would keep the semimajor axes constant).

Finally, we compare the performance of this 'short-integration' model to the previous 'initial-condition' model. Fig. 5.2 shows that while both models often assign unstable systems (blue bins) a low predicted probability, the 'initial-condition' model assigns stable systems (green bins) a wide range of predicted probabilities, translating to a lower recall. In contrast, the 'short-integration' model more

Table 5.2: 'Short-Integration' Model Feature Importances. See text for a description of the gain and of the features.

Feature	Gain
`max_a2`	20.3
`std_a2`	8.3
`mindaOverRH`	2.6
`maxdaOverRH`	2.6
`std_a3`	2.3
`max_e2`	2.3
26 more features...	

confidently assigns high predicted probabilities to stable systems, better separating the two classes. Again setting the predicted probability threshold so as to obtain 90% precision, the recall improves to 68%.

5.5 Discussion & Conclusion

In this investigation, we numerically integrated a dataset of 5000 three-planet systems over 10^7 orbits. We then trained machine learning algorithms to classify systems' orbital stability on this timescale. In particular, we trained two models using the gradient-boosted decision trees algorithm `XGBoost`, an 'initial-conditions' model (Sec. 5.4.1) that learned only from the system's initial orbital elements, and a 'short-integration' model (Sec. 5.4.2) that generated features from short N-body integrations. We then compared their performance to 'Lissauer-family models' that require the sum of the interplanetary separations (expressed in mutual Hill radii) to be greater than a particular threshold.

We summarize the investigated models' performances in Fig. 5.3, which plot values for the respective classifier's precision and recall for all possible values of the probability threshold above which the model labels a system as stable. As discussed above, the appropriate choice of this threshold depends on the desired qualities of the classifier. Above we advocated for conservative models that are correct 90% of the time when a system is predicted stable (90% precision), but different applications might consider different criteria.

By generating dynamically relevant features from short integrations, our best machine-learning model:

- Dominates other models at all threshold values.
- Recovers 2.25 as many of the truly stable systems (has 2.25 times higher recall) as a Lissauer-family model at a fixed precision of 90%.
- Is three orders of magnitude faster than direct N-body integration. Performing the short integration, generating the features and evaluating the model for a given system takes ~ 1 second with current technology.

An important limitation of this work are the fixed masses (5 Earth masses around a solar-mass star) and the comparatively short integration timescales (10^7 orbits). Our results strongly motivate investing computational time to generate datasets over longer timescales with a range in masses. Running these integrations in a reproducible and openly accessible manner (Rein & Tamayo, in prep.) would allow

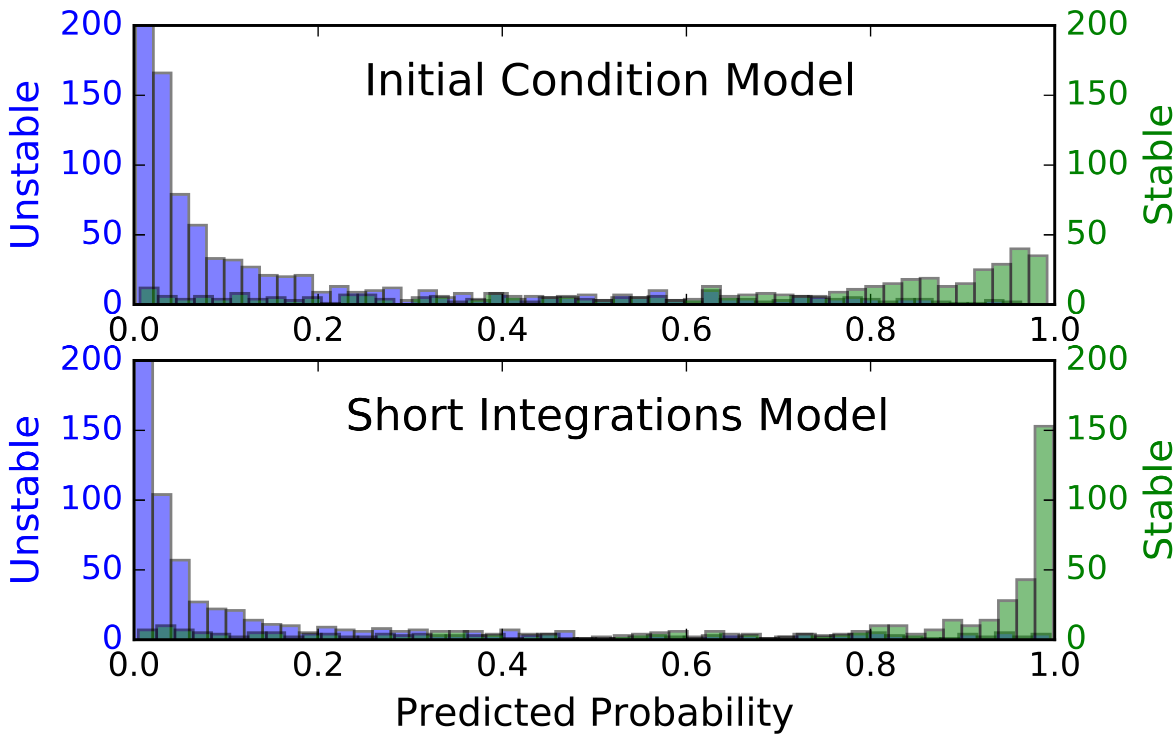

Figure 5.2: Comparison of predictions of the initial-condition and short-integration models on the test set. Stable systems are shown in green, unstable systems are shown in blue, and the model-predicted probability of stability for each system is shown along the x-axis. The leftmost blue bin is cut off to render smaller bins visible–in the top panel it reaches 395, and in the bottom panel 640.

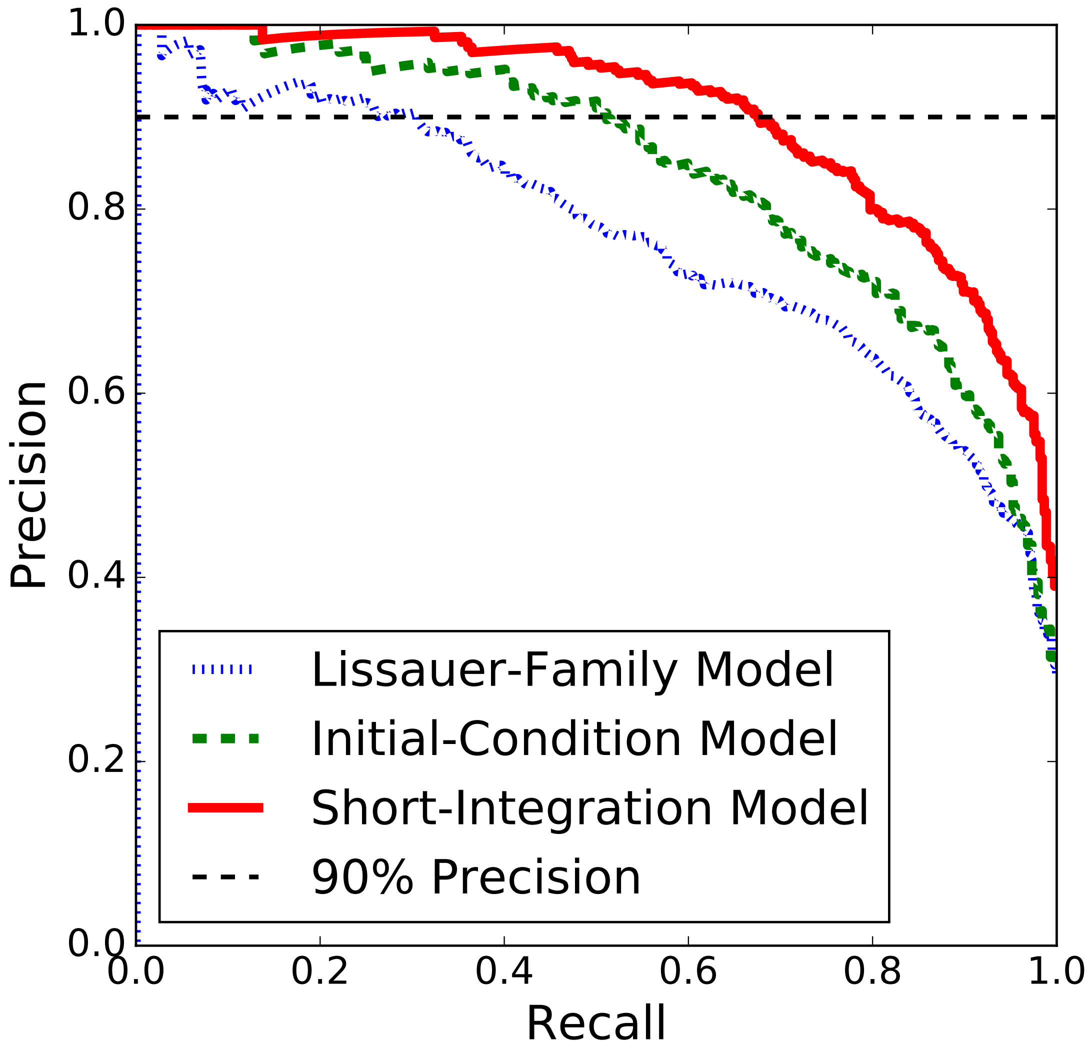

Figure 5.3: Precision-recall curves for each model in this Chapter, generated from all possible values of the model's threshold classification probability. The Lissauer-family models predict stability if the sum of the Hill-sphere separations is greater than a particular threshold, and the corresponding curve was generated by considering all possible threshold values. The horizontal black-dashed line is the 90% precision requirement we imposed on all models in this Chapter.

this dataset to be used as a standard on which current and future models could be evaluated and compared, and as a database of integrations to test theoretical investigations.

Such classifiers could be used in many of the ways that direct N-body integrations currently are employed, but these models' dramatically improved efficiency would allow much faster and complete explorations of parameter space, e.g.:

- Mapping out the stability boundary in mass-eccentricity space for observed transiting systems.
- Mapping out the parameter space that unseen planets could stably inhabit in a given system to guide observational follow-up strategies or theoretical considerations.
- Vetting low signal-to-noise detections through stability constraints.
- Generating stable or unstable systems for theoretical population studies.
- As a stopping condition for simulations once they achieve a dynamically long-lived configuration.

Such tools may be of particular interest for the upcoming Transiting Exoplanet Survey Satellite (TESS). While transit-timing variations (TTVs) have been powerful tools for constraining the masses and eccentricities of near-resonant Kepler systems (e.g., Ford et al. 2012; Steffen et al. 2013; Deck & Agol 2015), TESS' shorter time baselines and planets' smaller semimajor axes will render such analyses difficult or impossible. It is therefore likely that in many systems, stability considerations will provide the strongest constraints on planetary masses and eccentricities, and this will be important for guiding the substantial radial-velocity follow-up efforts from the ground.

More broadly, we have shown that machine learning can be a powerful tool for high-dimensional classification problems in dynamics and other fields. But in addition to their predictive power, our models also revealed new insights into the underlying dynamics. In particular, the most informative features in our model based on short integrations were the variations in the middle planet's semimajor axis. This suggests that the orbital instabilities are driven by the overlap of mean motion resonances (which vary the semimajor axes) rather than the secular chaos at work in our own solar system (Lithwick & Wu 2011; Batygin et al. 2015), which proceeds with constant semimajor axes.

Chapter 6

Resonant structure, formation and stability of the planetary system HD155358

6.1 Chapter Overview

The work, figures and tables from this Chapter are based off the following publication:

Ari Silburt, Hanno Rein
Accepted for publication in MNRAS, pp.8, 2017 (Silburt & Rein 2017).

Two Jovian-sized planets are orbiting the star HD155358 near exact mean motion resonance (MMR) commensurability. In this work we re-analyze the radial velocity (RV) data previously collected by Robertson et al. (2012). Using a Bayesian framework we construct two models – one that includes and one that excludes gravitational planet-planet interactions (PPI). We find that the orbital parameters from our PPI and noPPI models differ by up to 2σ, with our noPPI model being statistically consistent with previous results. In addition, our new PPI model strongly favours the planets being in MMR while our noPPI model strongly disfavours MMR. We conduct a stability analysis by drawing samples from our PPI model's posterior distribution and simulating them for 10^9 years, finding that our best-fit values land firmly in a stable region of parameter space.

We explore a series of formation models that migrate the planets into their observed MMR. We then use these models to directly fit to the observed RV data, where each model is uniquely parameterized by only three constants describing its migration history. Using a Bayesian framework we find that a number of migration models fit the RV data surprisingly well, with some migration parameters being ruled out.

Our analysis shows that planet-planet interactions are important to take into account when modelling observations of multi-planetary systems. The additional information that one can gain from interacting models can help constrain planet migration parameters.

6.2 Introduction

Many Jovian planets have been detected in exoplanetary systems, and the range of conditions under which they can form is uncertain. In particular, many Jovian planets are observed inside the snowline of their host star (e.g. Hayashi 1981), and it is unclear whether they formed in-situ (Boley et al. 2016; Huang et al. 2016; Batygin et al. 2016) or beyond the snowline and migrated in afterwards (Mayor & Queloz 1995; Lin et al. 1996; Pollack et al. 1996). For systems near mean motion resonance (MMR), planetary migration offers a natural formation mechanism (e.g. Lee & Peale 2002; Rein & Papaloizou 2009). Since roughly 30% of Jovian-sized planets with a detected planet companion are near a first-order MMR (Akeson et al. 2013), at least a substantial fraction of planetary systems are likely to have formed via migration.

The star HD155358 hosts two observed Jovian-sized planets orbiting inside the snowline near a MMR. The star has a mass of 0.92 $M_{\odot}$, and the planets have orbital periods $P_1 = 194$ and $P_2 = 392$ days, respectively, about 2% away from the exact 2:1 commensurability (Robertson et al. 2012, hereafter R2012). In addition, HD155358 is also among the lowest metallicity stars known to host Jovian-sized planets (Cochran et al. 2007). Since the presence of MMR can provide additional constraints on formation and evolution, HD155358 provides an opportunity to better understand the formation of Jovian planets in exoplanetary systems.

The HD155358 system has been previously studied by many scientists (Cochran et al. 2007; Fuhrmann & Bernkopf 2008; Robertson et al. 2012; André & Papaloizou 2016). Of particular interest is the work by R2012, who updated the orbital parameters initially reported by Cochran et al. (2007) after collecting additional radial velocity (RV) observations. In R2012 the orbital properties were derived from GAUSSFIT (Jefferys et al. 1988) and SYSTEMIC (Meschiari et al. 2009), and planet-planet interactions were not included in their analysis (Robertson 2016, private communication). For well-separated orbits this assumption is valid and simplifies the analysis. However, near a MMR this approximation can lead to different orbital solutions and therefore different formation constraints and evolutionary predictions. In particular, without planet-planet interactions one cannot draw any conclusion as to whether the system is in resonance or not.

In this Chapter we re-analyze HD155358 using the RV observations collected by R2012. In Sect. 6.3, we derive new orbital parameters using a Bayesian framework coupled to direct N-body integrations and assess the probability that the system is in resonance. In Sect. 6.4, we perform simulations constraining the migration history, and in Sect. 6.5 conduct a stability analysis. We conclude with a discussion in Sect. 6.6.

6.3 Best-Fit Parameters and Resonance Analysis

6.3.1 Methods

The RV data used for this analysis comes from Table 2 of R2012. We model this system using `REBOUND` (Rein & Liu 2012), an open-source N-body code for simulating problems in planetary dynamics. Using `REBOUND` we can extract the motion of the central star over the course of our simulation, comparing it directly to the RV data.

In this Chapter we explore two different models – one that includes the planet-planet gravitational interactions, abbreviated as PPI model, and one that excludes planet-planet gravitational interactions, abbreviated as noPPI model. This is achieved in REBOUND by initializing each planet as either a massive body (PPI model) or a semi-active particle (noPPI model). Semi-active particles can gravitationally interact with massive bodies (e.g. the central star), but cannot interact with other semi-active particles. For reference, R2012 does not include the gravitational interactions between the planets, and is analogous to our noPPI model.

We construct our models using a Bayesian framework, where the quantity of interest is the posterior probability density function. Benefits of conducting an analysis within a Bayesian framework include marginalization over nuisance parameters, preservation of correlations between parameters, and a natural Occam's razor when comparing models (see e.g. Gregory 2005, for a full discussion). Using Bayes' theorem we calculate the (unnormalized) posterior probability according to:

$$p(\theta|D) \propto p(D|\theta)p(\theta),$$

where θ are the parameters in our model, D is our data, $p(\theta)$ is our prior probability on θ, $p(D|\theta)$ is the probability of D given θ (i.e. the likelihood function), and $p(\theta|D)$ is our posterior probability. As with most models that employ Bayes' theorem, the posterior probability cannot be analytically computed, requiring the use of numerical methods for approximating it. We use EMCEE, an open-source affine-invariant Markov chain Monte Carlo (MCMC) routine (Foreman-Mackey et al. 2013).

For our models we assume that the system is co-planar, and use the known value of stellar mass, $m_* = 0.92M_\odot$. For each planet we have the following parameters: the reduced mass, $m\sin(i)$, the semi-major axis, a, the eccentricity, e, the argument of periapsis, ω, and the mean anomaly M. To avoid singularities for ω (which is ill-defined when $e \sim 0$), we fit $h = e\sin(\omega)$ and $k = e\cos(\omega)$ instead of e and ω. We also add an offset parameter γ to account for any stellar drift along the line of sight, a jitter parameter J to account for any stellar noise and unreported instrument variability, and a viewing angle parameter $\sin(i)$. This yields a total of 13 parameters to be sampled by the MCMC.

We initialize our MCMC chain with values similar to the best fit values of R2012 to speed up convergence, and use the following priors on our parameters: $0.4M_J < m_1\sin(i) < 2M_J$, $0.4M_J < m_2\sin(i) < 2M_J$, $0.2\text{AU} < a_1 < 0.8\text{AU}$, $0.8\text{AU} < a_2 < 1.4\text{AU}$, $1 < h_1, h_2, k_1, k_2 < -1$, $0 < \omega_1, \omega_2, M_1, M_2 < 2\pi$, $-40\text{m/s} < \gamma < 40\text{m/s}$, and $0\text{m/s} < J^2 < 50\text{m/s}$. Our MCMC chain consists of an initial burnin phase of 1000 steps with 400 walkers, after which the MCMC walkers are resampled near the best solution and run for 5000 steps. In REBOUND we use the WHFAST integrator (Rein & Tamayo 2015) with a timestep of $P_1/200$, leading to a bounded relative energy error of $< 10^{-7}$.

6.3.2 Best Fit Parameters

Our maximum a-posteriori (MAP) values for each parameter are displayed in Table 6.1. For ease of comparison between our results and R2012, we convert back into e and ω variables from h and k variables. Since we find $\sin(i)$ unconstrained in our MCMC analysis (see Figure 6.4) it is not listed in Table 6.1. Parameters with subscript 1 relate to the inner planet, while parameters with subscript 2 relate to the outer planet. A random sample of 2000 draws from our PPI model's posterior distribution is shown in Figure 6.4.

From Table 6.1 there are a few statistically significant differences between our PPI and noPPI models. In particular, e_2 is inconsistent at the 1-σ level, while $m_1\sin(i)$, ω_1 and M_1 are inconsistent at the 2-σ level. Furthermore, every parameter in our noPPI model is statistically consistent with R2012.

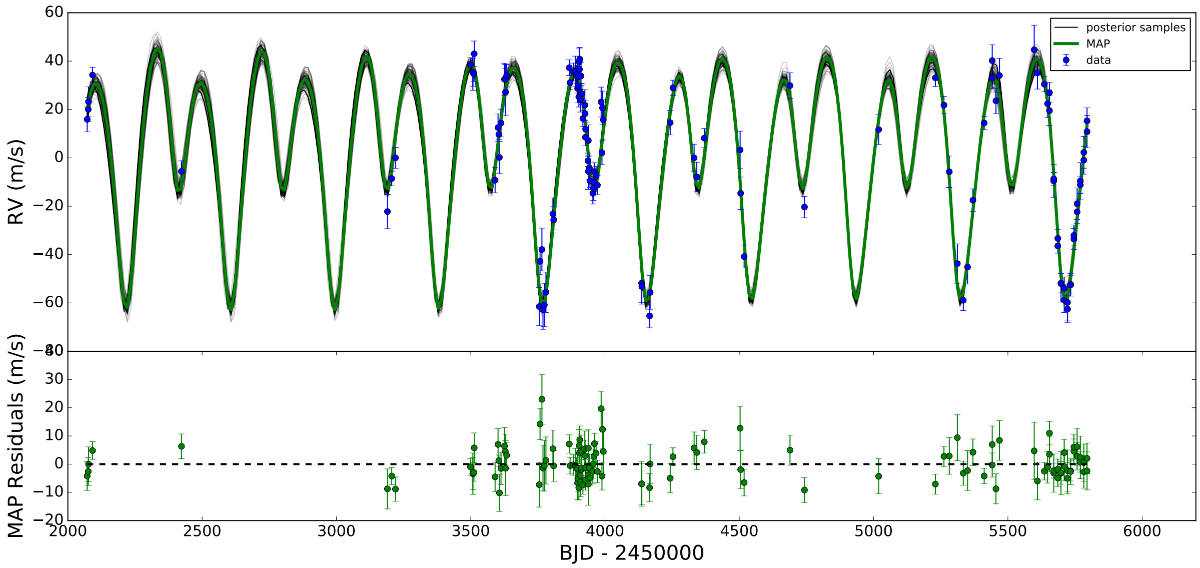

Figure 6.1: RV data and PPI model fit. Top panel shows the RV data points, MCMC MAP value (green), and 80 randomly plotted samples from the posterior (black). Bottom panel shows the residuals for the MAP fit to the RV data.

Parameter	PPI model	noPPI model	R2012
$m_1 \sin(i)$ (M_J)	0.92 ± 0.04	0.82 ± 0.04	0.85 ± 0.05
$m_2 \sin(i)$ (M_J)	0.85 ± 0.03	0.87 ± 0.03	0.82 ± 0.07
a_1 (AU)	0.641 ± 0.002	0.643 ± 0.004	0.64 ± 0.01
a_2 (AU)	1.017 ± 0.005	1.015 ± 0.007	1.02 ± 0.02
e_1	0.17 ± 0.03	0.18 ± 0.03	0.17 ± 0.03
e_2	0.10 ± 0.04	0.20 ± 0.05	0.16 ± 0.1
ω_1	$178^{\circ}{}^{+14^{\circ}}_{-17^{\circ}}$	$142^{\circ} \pm 13^{\circ}$	$143^{\circ} \pm 11^{\circ}$
ω_2	$241^{\circ}{}^{+72^{\circ}}_{-65^{\circ}}$	$189^{\circ}{}^{+10^{\circ}}_{-16^{\circ}}$	$180^{\circ} \pm 26^{\circ}$
M_1	$91^{\circ}{}^{+15^{\circ}}_{-18^{\circ}}$	$124^{\circ} \pm 15^{\circ}$	$129^{\circ} \pm 0.7^{\circ}$
M_2	$178^{\circ}{}^{+73^{\circ}}_{-66^{\circ}}$	$244^{\circ}{}^{+12^{\circ}}_{-17^{\circ}}$	$233^{\circ} \pm 0.9^{\circ}$
γ (m/s)	3.76 ± 0.71	3.70 ± 0.70	—
J (m/s)	$2.90^{+0.78}_{-0.64}$	$2.92^{+0.78}_{-0.62}$	2.49

Table 6.1: Best-fit model parameters from our MCMC samples. Our PPI model includes planet-planet interactions, while our noPPI model excludes planet-planet interactions. The derived parameters from R2012 are also listed for ease of comparison.

	PPI model	noPPI model
ϕ_1 librating	$98\% \pm 6\%$	$1\% \pm 6\%$
ϕ_2 librating	$5\% \pm 6\%$	$1\% \pm 6\%$

Table 6.2: The percent of randomly drawn samples with ϕ_1 and ϕ_2 librating, for our PPI and noPPI models. Error bars calculated from simple Poisson statistics.

In the top panel of Figure 6.1 we plot the MAP estimate for our PPI model, along with 80 randomly drawn samples from our posterior distribution. In the lower panel the residuals between our MAP estimate and the RV data are shown. As can be seen, our PPI model represents a good fit to the data.

We also try adding two additional parameters to our PPI model to account for mutual inclination between the planets. For this inclination model we add $i_{x,2} = 2\sin(i_2/2)\cos(\Omega_2)$ and $i_{y,2} = 2\sin(i_2/2)\sin(\Omega_2)$, where i_2 is the inclination of the outer planet (with respect to the inner planet and star) and Ω_2 is the longitude of ascending node of the outer planet (Pál 2009). The mutual inclination returned by this model is $i_2 = 30^{\circ}{}^{+22^{\circ}}_{-37^{\circ}}$, being consistent with 0 and ruling out large mutual inclinations. Comparing this inclination model to our original PPI model using a posterior odds ratio (see Section 6.4.2) yields a value of 0.4, indicating a slight preference over the inclination model. Since our original PPI model is simpler and the model parameters between models are similar, we choose to stick with our PPI model for the remainder of our analysis.

6.3.3 Resonance Analysis

Using our posterior samples found in Section 6.3.2, we can assess the likelihood that the planets orbiting HD155358 are in 2:1 MMR. More specifically, we can draw samples from the posterior distribution and calculate what fraction of these systems have librating resonant angles. The resonant angles are calculated according to:

$$\begin{aligned}\phi_1 &= (j-1)\lambda_1 - j\lambda_2 + \omega_1 \\ \phi_2 &= (j-1)\lambda_1 - j\lambda_2 + \omega_2\end{aligned}$$

where λ is the mean longitude, ω is the argument of periapsis, and j is the order of the resonance ($j = 2$ for this work). Our aim is to use the fraction of systems with librating resonant angles as a measure of how likely the system is to be in MMR.

For each set of parameters drawn from our posterior we simulate the corresponding system for 4000 years with a timestep equal to $P_1/50$. Over the course of a given simulation we generate 1000 equally spaced outputs for each resonant angle, and classify a resonant angle to be librating if the difference between the maximum and minimum is less than a threshold value κ. We set $\kappa = 7\pi/4$ since this allows for large resonant libration amplitudes while also ensuring that misclassification of librating/non-librating systems is low. Our results are insensitive to nearby values of κ (e.g. $5\pi/3$, $9\pi/5$, etc.). We check for librations around both 0 and π.

Our results are presented in Table 6.2 for our PPI and noPPI models, for 300 samples randomly drawn from each posterior distribution. The error bars are calculated from simple poisson statistics. For our PPI model, almost every drawn sample has ϕ_1 librating, strongly suggesting that the system is in MMR. In contrast, our noPPI model shows no evidence of librating resonant angles.

Given our analysis, there is a high likelihood that the HD155358 system is in MMR, and is therefore likely to have formed via planet migration, as we show in the next section.

6.4 Formation via Migration

6.4.1 Methods

We now use the results found in the previous section to explore the formation of HD155358 via migration. We use a simple model by which the outer planet is migrated into 2:1 MMR using a parametric model, mimicking the presence of a protoplanetary disk. Specifically, we introduce semi-major axis and eccentricity damping timescales, $\tau_a = a/\dot{a}$ and $\tau_e = e/\dot{e}$, respectively, where the two are related by a constant $K = \tau_a/\tau_e$. We refer to this class of convergent-migration models as "CM models".

We follow the implementation by Papaloizou & Larwood (2000) and add an additional acceleration on each planet according to:

$$\mathbf{a}_{\text{mig}} = -\frac{\mathbf{v}}{\tau_a}$$
$$\mathbf{a}_{\text{damp}} = -2\frac{(\mathbf{v}\cdot\mathbf{r})\mathbf{r}}{r^2\tau_e},$$

where $\mathbf{v}$ is the velocity and $\mathbf{r}$ is the position of the planet relative to the star. The migration prescriptions of Papaloizou & Larwood (2000) were designed with Type-I migration in mind, yet our planets likely lie in the Type-II migration regime. However, since we are interested primarily in seeing whether a simple migration model can reproduce the data, this model choice will suffice. More accurate, 3-D hydrodynamical simulations of Type-II planetary migration are significantly more complex and beyond the scope of this chapter.

For each simulation three parameters are independently varied – the eccentricity damping timescale of the inner planet τ_{e_1}, the eccentricity damping timescale of the outer planet, τ_{e_2}, and the migration timescale of the outer planet, τ_{a_2}. The inner planet does not undergo migration by itself, i.e. $\tau_{a_1} = \infty$. Initial values for τ_{a_2} are drawn from a uniform distribution in log-space between $10^{2.5}$ and 10^7 years, while initial values for K_1 and K_2 are drawn between 10^{-1} and 10^3. Planet masses are drawn from our posterior distribution (Section 6.3.2), all eccentricity and inclination values are initialized to zero, and the remaining orbital parameters are randomly drawn from uniform distributions.

When planets migrate into MMR in the presence of a protoplanetary disk an equilibrium eccentricity e_{eq} is reached, representing a stable balance between migration excitation and protoplanetary disk damping (e.g. Goldreich & Schlichting 2014). For each simulation we migrate the outer planet inwards for a time of $T_{\text{mig}} = 5\tau_a$ years, after which e_{eq} has reached a stable value. In cases where $T_{\text{mig}} < 5000$ years we extend the simulation time to $T_{\text{mig}} = 5000$ to ensure a stable solution. The migration parameters τ_{a_1}, τ_{e_1} and τ_{e_2} are then logarithmically increased to 10^7 years over the same amount of time (T_{mig}), mimicking the dispersal of the protoplanetary disk. Simulations are independently checked to ensure that a) an equilibrium eccentricity is reached over T_{mig}, and b) the disk dispersal is done adiabatically, ensuring that the resonance is not abruptly changed during the increase of τ_{a_1}, τ_{e_1} and τ_{e_2}. Simulations that do not fulfill this criteria are discarded.

Once migration is turned off, a RV curve is generated from the simulation and fit to the original RV data using `EMCEE` as before, but now with only five free parameters: t_s, a parameter allowing the rescaling of the orbital periods, t_t, a parameter accounting for an offset in time, y_s, a RV-stretch parameter to account for amplitude offsets (effectively rescaling the masses of both the planets and the star, while keeping the mass ratios and periods constant). As in the original analysis, we also include γ, an offset parameter to account for any stellar drift along the line of sight, and J, a jitter parameter to account for underestimated measurement noise and intrinsic stellar noise in the RV data. Each `EMCEE` fit is run for an initial burnin phase of 200 steps with 100 walkers, after which the walkers are resampled near the best solution and run for 2000 steps.

6.4.2 Model Comparison

In a Bayesian sense, each simulation represents a different model M characterized by $\theta_{\rm model}$ = {τ_{a_1}, K_1, K_2}, with each model having a unique MAP estimate for its parameters θ_{fit} = {t_s, t_t, y_s, γ, J}. Note that our analysis thus incorporates the fact that many different models might explain the data equally well. A single model with parameters θ = {τ_{a_1}, K_1, K_2, t_s, t_t, y_s, γ, J} on the other hand would make MCMC convergence more difficult.

When multiple models can explain the data, the posterior odds ratio can be used to determine if there is any preference for one model over another. The posterior odds ratio, P_{ij}, is calculated according to (Gregory 2005):

$$P_{ij} = \frac{p(M_i)}{p(M_j)}\frac{p(D|M_i)}{p(D|M_j)}, \tag{6.1}$$

where $p(M)$ is the prior odds for model M, and $p(D|M)$ is the marginal (or global) likelihood for model M. The ratio of marginal likelihoods for two competing models is also known as Bayes' factor. In our case, the ratio of prior odds is unity since we have no prior model preference. Thus, the posterior odds ratio is equivalent to calculating Bayes' factor.

Formally, the marginal likelihood is calculated by marginalizing over the parameters θ of a model according to (Gregory 2005):

$$p(D|M) = \int p(\theta|M)p(D|\theta, M)d\theta = \mathcal{L}(M) \tag{6.2}$$

where $p(D|\theta, M)$ is the likelihood and $p(\theta|M)$ is the prior. The marginal likelihood represents the probability of the data given the model.

When the marginal likelihood is normally distributed with flat priors, Eq. 6.2 can be approximated as (Kass & Raftery 1995; Gregory 2005):

$$\mathcal{L}(M) \approx \mathcal{L}(\hat{\theta})(2\pi)^{d/2}|\Sigma|^{1/2}\Delta\theta \tag{6.3}$$

where $\mathcal{L}(\hat{\theta})$ is the the maximum likelihood, d is the number of parameters in model M, Σ is the covariance matrix of the posterior distribution, and $\Delta\theta = \prod_i^d 1/\delta\theta_i$ is the product of prior probabilities normalized by their lengths $\delta\theta$. Since all models have the same flat priors for each parameter, Eq. 6.1 becomes:

$$P_{ij} = \frac{\mathcal{L}_i(\hat{\theta})|\Sigma_i|^{1/2}}{\mathcal{L}_j(\hat{\theta})|\Sigma_j|^{1/2}} \tag{6.4}$$

Since `EMCEE` returns both the samples and log-likelihoods at each step in the MCMC chain, it is straightforward to calculate the posterior odds ratio using Eq. 6.4.

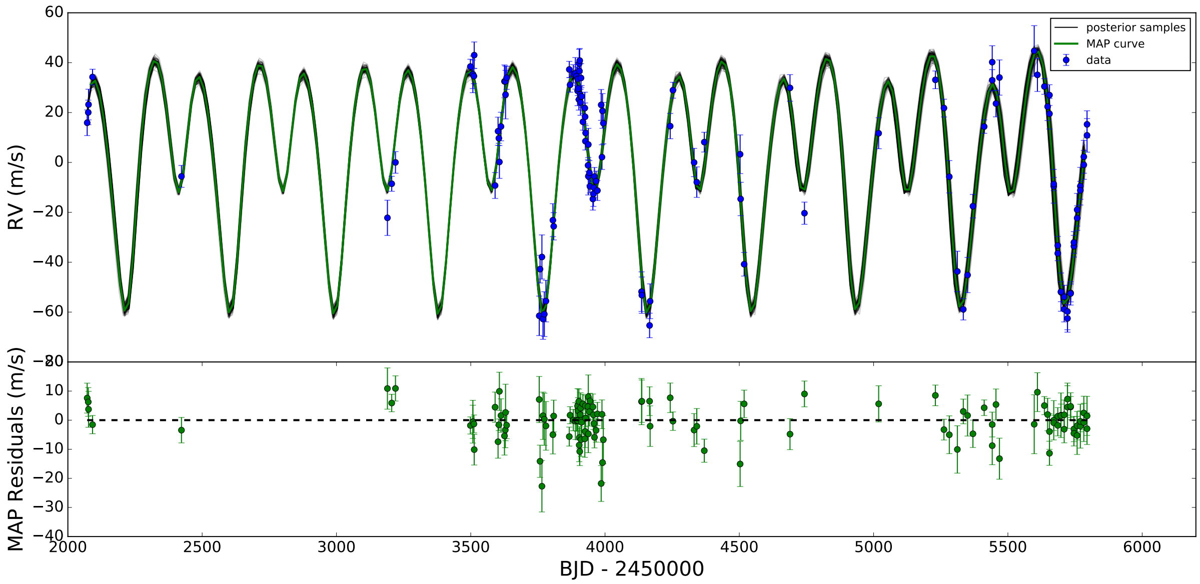

Figure 6.2: RV data and $\theta_{\text{mig,best}}$ fit. Top panel shows the RV data points, MCMC MAP value (green), and 80 randomly plotted samples from the posterior (black). Bottom panel shows the residuals for the MAP fit to the RV data.

6.4.3 Results

Our best model from a sample of 3000 simulations is plotted in Figure 6.2, corresponding to $\theta_{\text{model,best}}$ = (τ_{a_2}, K_1, K_2) = (4000 years, 7, 265). In the top panel we plot the MAP estimate along with 80 randomly drawn samples from our posterior distribution. In the lower panel the residuals between our MAP estimate and the RV data are shown. As can be seen, our model represents a good fit to the data.

Following Kass & Raftery (1995), a posterior odds ratio lower than 100 indicates no decisive preference for one model over another, and in Figure 6.3 we plot competing models which, when compared to $\theta_{model,best}$, yield a posterior odds ratio of less than 100. The top panel in Figure 6.3 presents the valid models in (K_1, K_2, τ_{a_2}) space, while the bottom panel presents them in $(\tau_{e_1}, \tau_{e_2}, \tau_{a_2})$ space. Although many different migration models are able to explain the data, regions of parameter space appear to be ruled out. In the top panel of Figure 6.3, $(K_1 < 3, K_2 < 100)$ and $K_1 > 20$ are two regions that appear to be ruled out by the data. In the bottom panel, positive correlation can be seen between τ_{a_2} and $\tau_{e_1} + \tau_{e_2}$.

6.5 Stability

Following Marshall et al. (2010) and Horner et al. (2011), R2012 conducted a stability analysis by holding the inner planet constant at the best-fit values and varying the outer planet's orbital parameters randomly over a 3σ range in a, e, ω and M. However, drawing parameters in this manner does not preserve their correlations, and it is possible to draw parameter combinations that are inconsistent with the RV data. R2012 also kept the planet masses fixed at their minimum $m \sin(i)$ values, precluding the possibility of further constraining the masses from the stability analysis.

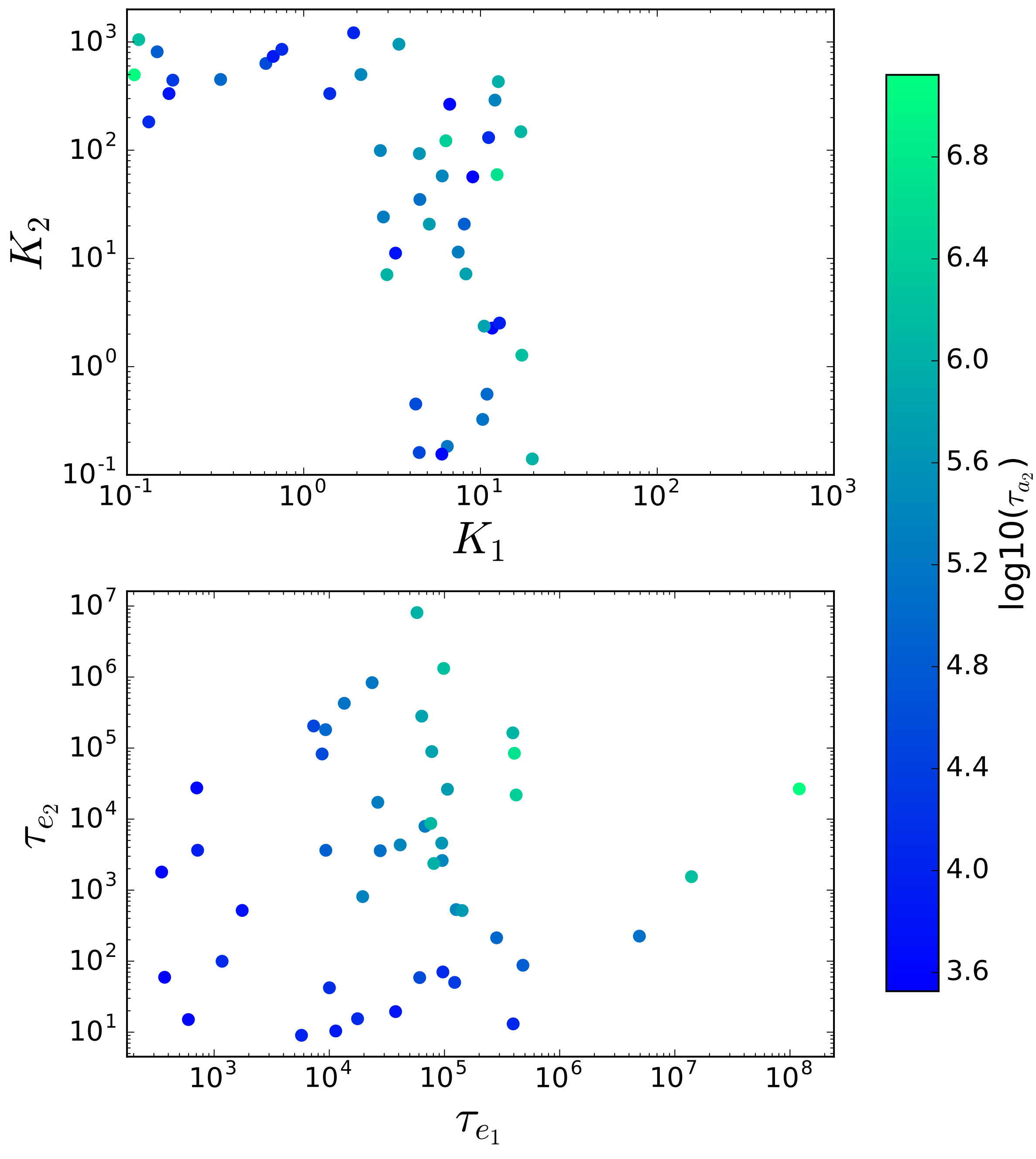

Figure 6.3: The best 45 models $\theta_{\text{model}} = \{\tau_{a_1}, K_1, K_2\}$ with Bayes' factors less than 100 when compared against the best model in our sample of 3000 simulations. Top panel presents the data in (K_1, K_2) space, and the parameters from our 3000 simulations were uniformly sampled from the region displayed. Bottom panel displays the same data in (τ_{e_2}, τ_{e_2}) space. Colour indicates the log value of τ_{a_1}.

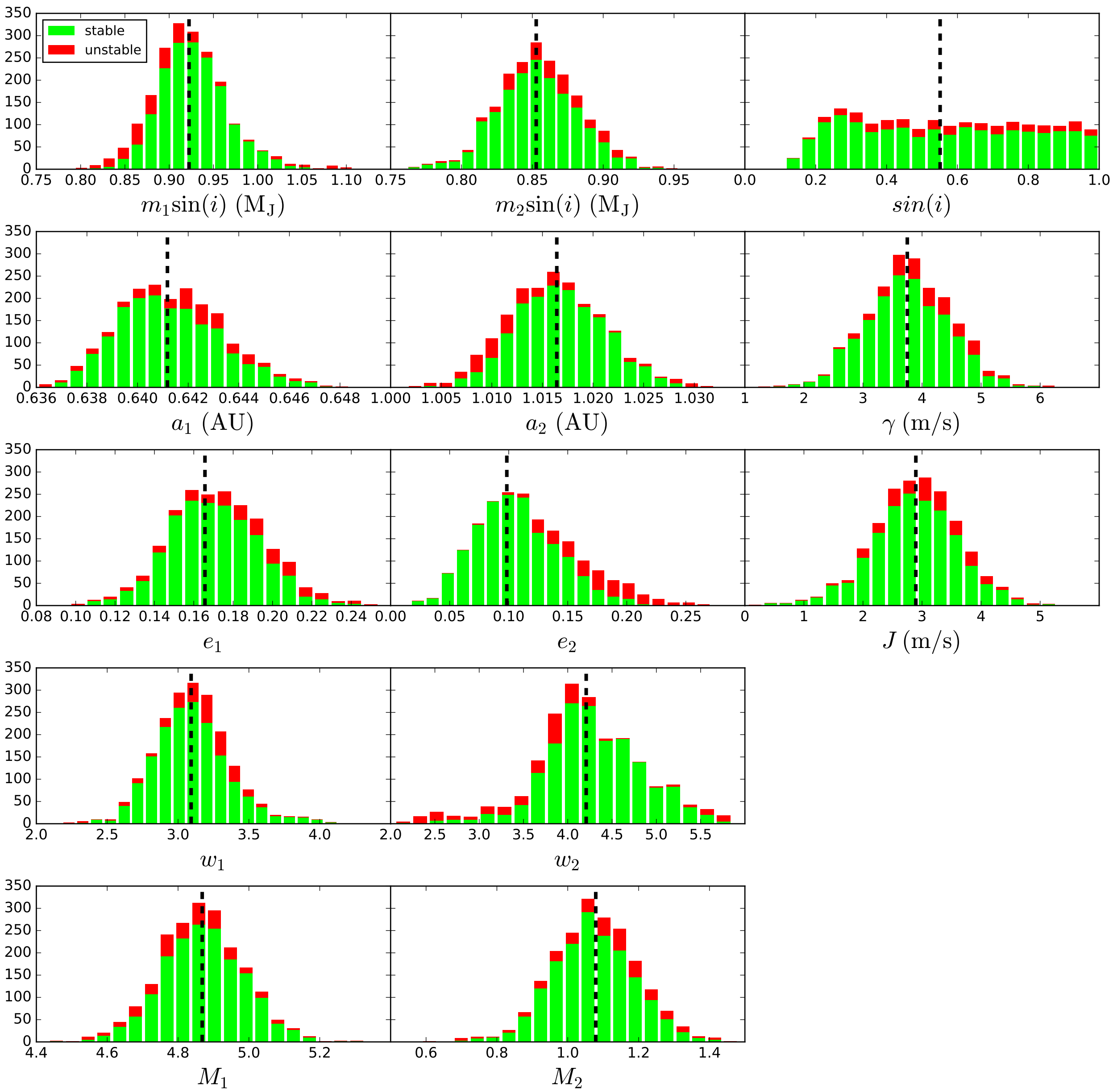

Figure 6.4: Histograms of each parameter showing the distribution of 2000 stable (green) and unstable (red) systems drawn from our PPI model's posterior distribution and simulated for 10^9 years. The black vertical dashed line in each histogram represents our MAP estimate for that parameter.

In a Bayesian framework, we can simply draw parameters from the posterior distribution and simulate them for a desired length of time. Parameter correlations are naturally preserved in the posterior distribution, and a probability of longterm stability can be easily found.

In this study we draw 2000 random samples from our PPI model's posterior distribution and simulate the subsequent evolution for 10^9 years using `WHFAST` with a timestep of $P_1/200$. We also make use of the `Simulation Archive` (Rein & Tamayo 2017) for stopping, restarting and analyzing simulations. The relative energy error for all stable simulations remains bounded at $< 10^{-7}$.

The results are presented in Figure 6.4 as a histogram for each parameter. Stable and unstable systems over 10^9 years are marked in green and red, respectively, while our MAP values from our PPI model is shown as vertical, black, dashed lines. $83\% \pm 2\%$ of our samples are stable over 10^9 years and our MAP values land in highly stable regions, indicating that our PPI model is longterm stable.

Unstable systems appear to be clustered in certain regions of parameter space. As can be seen in Figure 6.4, $m_1 \sin(i) < 0.87$, $e_2 > 0.15$, $a_2 < 1.01$ AU and $\omega_2 < 4$ are regions where there are a significant fraction of unstable systems. It is therefore unlikely that the true planet parameters lie in these regions.

These unstable regions are consistent with the results found in R2012, however, unlike R2012 our best-fit values are centred in highly stable regions (see Figures 9 and 10 in R2012 for a comparison). Although not shown, we find no clustering of unstable systems in (m_1, m_2) space, and thus the planet masses cannot be further constrained via stability analysis.

6.6 Discussion and Conclusion

In this Chapter we used three different model types to analyze the RV curve of HD155358 – the PPI, noPPI and CM models. Our PPI (planet-planet interactions) and noPPI (no planet-planet interactions) models fit the RV curve to a model parameterized by the orbital elements of each planet (Section 6.3), while our CM model was parameterized by migration parameters and fit this result to the RV curve (Section 6.4). The best-fit parameters from our PPI and noPPI models are shown in Table 6.1, while Table 6.2 shows that there is a high likelihood that the planets of HD155358 are in MMR.

The noPPI model is an approximation to the PPI model, and we have shown in Section 6.3.2 that planet-planet interactions are strong enough in the HD155358 system to yield differing conclusions about the resonant structure. More specifically, our PPI model predicts a high likelihood that the planets are in MMR, while our noPPI model predicts a low likelihood that the planets are in MMR. Since planets in MMR are highly likely to have formed via migration, by extension the PPI and noPPI models have different formation implications. Quantitatively, the orbital parameters returned by our PPI and noPPI models differ by up to 2σ.

Our CM model has demonstrated that formation models of this style can be used to fit RV curves and constrain the initial conditions of a given system. One possibility for future work is to make the CM model more sophisticated, for example by modelling the planet-disk interactions in more realistic hydrodynamic simulations (Rein et al. 2010). With such a model, one could place further constraints on the formation of HD155358.

From a frequentist's point of view, the reduced χ^2 values for our PPI, noPPI and best CM model are 1.4, 1.6 and 0.7 respectively, indicating that all three models are capable of fitting the data well. However, our PPI model should be considered the only valid model when it comes to determining the true values of the observed system. We also performed a stability analysis on our PPI model, finding that $83\% \pm 2\%$ of samples drawn from our posterior are stable over 10^9 years. We find regions of stable and unstable parameter space similar to R2012, however unlike R2012 our best-fit model solution is centred in a stable region.

Chapter 7

Conclusions & Future Work

7.1 Summary

The number of discovered exoplanets has dramatically increased over the past decade, allowing scientists to study planetary systems outside the Solar System. Broadly speaking, scientists are interested in understanding three things about planetary systems – their past formation, present archetechtures and future evolution. This thesis has contributed to all three domains by analyzing the statistics, formation and stability of exoplanetary systems.

In Chapter 2, I performed a statistical analysis of the *Kepler* catalog using the Ramirez et al. (2014) dataset to calculate the occurrence rate of small planets, accounting for detection biases and radius errors. Here I found that the occurrence of Kepler planets are preferentially peaked at $2 - 2.8R_{\oplus}$, with their numbers decreasing gradually toward smaller sizes, with a roughly log-uniform period distribution. These results were compared to Petigura et al. (2013), who finds a similar peak but different occurrence values for Neptune-sized planets. I calculated an average number of planets per star with periods between 20 and 200 days and radii between 1 and $4R_{\oplus}$ of 0.46 ± 0.03. Such planets have likely experienced little photoevaporation, and may reflect the "primordial" planet population. Upon extrapolation I obtained an occurrence rate for Earth-like planets within the "habitable zone" (as calculated by 1-D climate models) of $6.4^{+3.4}_{-1.1}\%$. As mentioned in Section 1.1.2, the most difficult aspect of calculating planet occurrence is accurately measuring the detection completeness, and our method was validated by successfully reproducing the work of Petigura et al. (2013) through applying our detection completeness calculations to their dataset.

The analysis in Chapter 2 revealed additional structures in the *Kepler* data, particularly the statistical excesses of planets wide of MMRs. It is believed that these planets were originally closer to MMR and a dynamical mechanism transported them to larger distances. The specific mechanism for transporting these planets is still unclear, motivating my work in Chapter 3. Here I used optimistic theoretical estimates for the minimum initial eccentricity required to explain the offset of *Kepler* planets from MMR due to tidal forces alone, complimenting these calculations with N-body simulations. Analyzing 27 *Kepler* systems with period ratios within 6% of the 2:1 MMR, I found that the initial eccentricities required to explain the observed spacings due to tides alone are unreasonable from simple dynamical arguments. Furthermore, our numerical simulations revealed "resonant tugging", an effect which conspires against the migration of resonant planets away from the 2:1 MMR, requiring even higher initial eccentricities in order to explain the current *Kepler* distribution. In summary, I found that tides alone cannot explain planets close to 2:1 MMR, and additional mechanisms are required to explain these systems.

After my work in Chapter 3, I sought other mechanisms capable of transporting planets from MMR. Planetesimals are a promising candidate for transporting planets from MMR, motivating my work in Chapter 4 where I presented `HERMES`, a new hybrid integration scheme for long-term simulations of planetary systems undergoing close encounters and planetesimal-driven migration. Distant particles are integrated using `WHFAST`, while close particles are integrated with `IAS15`. In addition, I created an adaptive routine for optimizing the close encounter boundary to help maintain accuracy whilst close encounters are occurring. Since `WHFAST` is symplectic, `IAS15` is accurate to machine precision and both of them are unbiased, the energy error grows sub-linearly with time under the assumption that either impact parameters are randomly distributed or close encounters are rare. I found that `HERMES` provides a good balance between speed and accuracy, neither achieved by the individual symplectic or non-symplectic integrators alone.

In addition to my work analyzing *Kepler* planets, I have also been taking part in machine learning workshops held frequently at the Centre for Planetary Science. From these workshops, I became interested in applying machine learning to the problem of planetary stability. In Chapter 5, I showed that characterizing the complicated and multi-dimensional stability boundary of tightly packed systems is amenable to machine learning methods. In particular, training a state-of-the-art machine learning algorithm on physically motivated features yields an accurate classifier of stability in packed systems. On the stability timescale we investigated (10^7 orbits), our trained machine was 3 orders of magnitude faster than direct N-body simulations. Optimized machine learning classifiers for dynamical stability may thus prove useful across the discipline, e.g., to characterize the exoplanet sample discovered by the upcoming Transiting Exoplanet Survey Satellite (TESS).

Finally, my previous research experience motivated me to work on the planetary system HD155358, since it drew on a number of familiar topics like MMR, formation and stability. In Chapter 6, I analyzed the RV data for the star HD155358, which hosts two Jovian-sized planets near 2:1 MMR. Using a Bayesian model parameterized by the orbital elements of each planet, I showed that excluding planet-planet interactions can yield statistically different orbital solutions, leading to different implications for formation and stability. From my updated orbital solutions I calculated a high likelihood that the planets are in MMR. In addition, I conducted a stability analysis by drawing samples from our posterior distribution and simulating them for 10^9 years, finding that our best-fit values land firmly in a stable region of parameter space. Furthermore, I also explored a series of formation models that migrate the planets into MMR and generated synthetic RV curves to fit directly to the observed data. I found that a number of formation models fit the RV data surprisingly well, with some migration parameters being ruled out.

7.2 Future Work and Directions

From Chapter 3, I found that a primary reason why tides cannot transport planets from exact MMR to a few percent wide is due to "Resonant Tugging" (Section 3.4.2). This effect was seen using two different tidal prescriptions, and is believed to be a real effect vs. a numerical artifact. However, this effect was not analytically derived and is poorly understood. Additional work is needed to understand the physics behind this process, as well as its range of applicability. Since the effect was seen for moderate-to-high eccentricities, a good starting point would be to expand the Resonant Hamiltonian to include higher-order eccentricity terms.

Chapter 4 presented `HERMES`, a hybrid integrator for planetesimal migration. The base integrator used in `HERMES` is `WHFAST` which employs Jacobi coordinates to integrate particles. However, as mentioned in Section 1.4.2 this is not the best choice since a clear ordering of bodies is required for efficient integration. Instead, democratic heliocentric coordinates or the WHDS coordinate system (Section 1.4.2) are better choices. As mentioned in Section 1.4.2, since the WHDS coordinate system relies on heliocentric coordinates for integrating bodies *and* solves the two-body and restricted three-body problems exactly, it might ultimately be the best choice of coordinate system for `HERMES`. However, at the time of writing this thesis this goal could not be fully realized, but will be completed in the future.

In addition, `HERMES` was originally designed to investigate the role of planetesimals in transporting planets from MMR, however this was never completed during my thesis. Most recently, Chatterjee & Ford (2015) investigated planetesimal migration for two Neptune mass planets, finding that $0.2m_N$ (where m_N is the mass of Neptune) of nearby, dynamically cold, planetesimals was required to break the resonance. However, this result warrants further investigation. For example, given the process of planetary formation (Section 1.2) is it reasonable to assume that $0.2m_N$ of nearby, dynamically cold planetesimals would be available to migrate planets? In addition, how would the initial migration phase of planets into MMR excite, eject and collide nearby planetesimals? In short, a self-consistent model of MMR planets embedded in a planetesimal disk has yet to be researched, and such a project can be accomplished with `HERMES` in the future.

Chapter 5 showed that a machine trained on sample N-body integrations can accurately predict the longterm stability of new systems based off their initial conditions and early evolution. The training set consisted of equal mass planets on circular orbits and simulated for 10^7 orbits. Substantial work can still be done to make this a valuable asset to the exoplanet community. Primarily, this work can be extended to planetary systems more representative of the *Kepler* population, with unequal mass planets on non-circular orbits simulated for an order of magnitude longer. In addition, for other scientists to use this tool, an easy-to-use pipeline must be created that can take a given planetary system as input and output a probability of longterm stability.

Finally, in Chapter 6 I used a simple formation model to fit the RV curve of HD155358 and determine the initial conditions of the system. This simple model could be replaced with a full 3D hydrodynamical model, further constraining the initial conditions of the system.

Bibliography

Akeson, R. L., Chen, X., Ciardi, D., & Crane, M. 2013, PASP, 125, 10

ALMA Partnership et al. 2015, ApJ, 808, L3

André, Q., & Papaloizou, J. C. B. 2016, MNRAS, 461, 4406

Armitage, P. J. 2007, ArXiv Astrophysics e-prints

—. 2010, Astrophysics of Planet Formation, 294

Barnes, R., & Greenberg, R. 2006, ApJ, 647, L163

Barnes, R., Raymond, S. N., Jackson, B., & Greenberg, R. 2008, Astrobiology, 8, 557

Baruteau, C., et al. 2013, ArXiv e-prints

Bastien, F. A., Stassun, K. G., & Pepper, J. 2014, ArXiv e-prints

Batygin, K., Bodenheimer, P. H., & Laughlin, G. P. 2016, ApJ, 829, 114

Batygin, K., & Morbidelli, A. 2013, AJ, 145, 10

Batygin, K., Morbidelli, A., & Holman, M. J. 2015, ApJ, 799, 120

Batygin, K., Stevenson, D. J., & Bodenheimer, P. H. 2011, ApJ, 738, 1

Beaugé, C., Ferraz-Mello, S., & Michtchenko, T. A. 2007, Planetary Masses and Orbital Parameters from Radial Velocity Measurements, 1

Benz, W., Ida, S., Alibert, Y., Lin, D. N. C., & Mordasini, C. 2014, ArXiv e-prints

Blum, J., & Wurm, G. 2008, ARA&A, 46, 21

Boccaletti, A., et al. 2015, Nature, 526, 230

Boley, A. C., Granados Contreras, A. P., & Gladman, B. 2016, ApJ, 817, L17

Boley, A. C., Morris, M. A., & Ford, E. B. 2014, ApJ, 792, L27

Bonfils, X., et al. 2013, A&A, 549, A109

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

Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, AJ, 142, 112

Catanzarite, C., & Shao, M. 2011, ApJ

Chambers, J. E. 1999, MNRAS, 304, 793

Chambers, J. E., Wetherill, G. W., & Boss, A. P. 1996a, Icarus, 119, 261

—. 1996b, Icarus, 119, 261

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

Chatterjee, S., & Ford, E. B. 2015, ApJ, 803, 33

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

Chen, T., & Guestrin, C. 2016, arXiv preprint arXiv:1603.02754

Chiang, E., & Laughlin, G. 2013, MNRAS, 431, 3444

Chirikov, B. 1979, Physics Review, 52, 263

Choquet, É., et al. 2016, ApJ, 817, L2

Christiansen, J. L., et al. 2012, ApJ, 124

Cochran, W. D., Endl, M., Wittenmyer, R. A., & Bean, J. L. 2007, ApJ, 665, 1407

Collister, A. A., & Lahav, O. 2004, Publications of the Astronomical Society of the Pacific, 116, 345

Colón, K. D., Ford, E. B., & Morehead, R. C. 2012, MNRAS, 426, 342

de Bruijne, J. H. J. 2012, Ap&SS, 341, 31

Deck, K. M., & Agol, E. 2015, The Astrophysical Journal, 802, 116

Deck, K. M., Holman, M. J., Agol, E., Carter, J. A., Lissauer, J. J., Ragozzine, D., & Winn, J. N. 2012, The Astrophysical Journal Letters, 755, L21

Delisle, J.-B., Laskar, J., & Correia, A. C. M. 2014, AA, 566, 14

Demircan, O., & Kahraman, G. 1991, ApSS, 181, 9

Dong, S., & Zhu, Z. 2013, ApJ, 778, 53

Dressing, C. D., & Charbonneau, D. 2013, ApJ, 767, 95

Duncan, M., Quinn, T., & Tremaine, S. 1989, Icarus, 82, 402

Duncan, M. J., Levison, H. F., & Lee, M. H. 1998, AJ, 116, 2067

Everett, M. E., Howell, S. B., Silva, D. R., & Szkody, P. 2013, ApJ, 771, 107

Faber, P., & Quillen, A. C. 2007, Monthly Notices of the Royal Astronomical Society, 382, 1823

Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., Rowe, J. F., & Steffen, J. H. 2014, ApJ, 790, 12

Fang, J., & Margot, J.-L. 2012, The Astrophysical Journal, 761, 92

Fang, J., & Margot, J.-L. 2013, ApJ, 767, 115

Fernandez, J. A., & Ip, W.-H. 1984, Icarus, 58, 109

Ford, E. B., et al. 2012, The Astrophysical Journal, 750, 113

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

Fressin, F., et al. 2013, ApJ, 766, 81

Friedman, J., Hastie, T., & Tibshirani, R. 2001, The elements of statistical learning, Vol. 1 (Springer series in statistics Springer, Berlin)

Fuhrmann, K., & Bernkopf, J. 2008, MNRAS, 384, 1563

Fung, J., & Chiang, E. 2017, ArXiv e-prints

Funk, B., Wuchterl, G., Schwarz, R., Pilat-Lohinger, E., & Eggl, S. 2010, A&A, 516, 6

Gaidos, E. 2013, ApJ, 770, 90

Gaidos, E., & Mann, A. W. 2013, ApJ, 762, 41

Gelman, A., & Rubin, D. 1992, Statistical Science, 7, 457, http://www.stat.columbia.edu/~gelman/research/published/itsim.pdf

Gladman, B. 1993, Icarus, 106, 247

Goldreich, P., & Schlichting, H. E. 2014, AJ, 147, 32

Goldreich, P., & Soter, S. 1966, Icarus, 5, 14

Goldreich, P., & Tremaine, S. 1979, ApJ, 233, 857

—. 1980, ApJ, 241, 425

Goldreich, P., & Ward, W. R. 1973, ApJ, 183, 1051

Gomes, R. S., Morbidelli, A., & Levison, H. F. 2004, Icarus, 170, 492

Greenberg, R., Hartmann, W. K., Chapman, C. R., & Wacker, J. F. 1978, Icarus, 35, 1

Gregory, P. C. 2005, Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with 'Mathematica' Support (Cambridge University Press)

Hadden, S., & Lithwick, Y. 2014a, ApJ, 787, 80

—. 2014b, ApJ, 787, 7

Hansen, B. M. S., & Murray, N. 2013, ApJ, 775, 53

Hayashi, C. 1981, Progress of Theoretical Physics Supplement, 70, 35

Henry, G. W., Marcy, G., Butler, R. P., & Vogt, S. S. 1999, IAU Circ., 7307

Hernandez, D. M., & Dehnen, W. 2016, ArXiv e-prints

Hill, G. J., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Hines, D. C., et al. 2007, ApJ, 671, L165

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

Horner, J., Marshall, J. P., Wittenmyer, R. A., & Tinney, C. G. 2011, MNRAS, 416, L11

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

Huang, C., Wu, Y., & Triaud, A. H. M. J. 2016, ApJ, 825, 98

Huber, D., et al. 2014, ApJS, 211, 2

Ida, S., Bryden, G., Lin, D. N. C., & Tanaka, H. 2000, ApJ, 534, 428

Ito, T., & Tanikawa, K. 1999, Icarus, 139, 336

Jackson, B., Barnes, R., & Greenberg, R. 2008, MNRAS, 391, 237

Jefferys, W. H., Fitzpatrick, M. J., & McArthur, B. E. 1988, Celestial Mechanics, 41, 39

Jontof-Hutter, D., Lissauer, J. J., Rowe, J. F., & Fabrycky, D. C. 2014, ApJ, 785, 15

Kass, R. E., & Raftery, A. E. 1995, Journal of the American Statistical Association, 90, 773

Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108

Kirsh, D. R., Duncan, M., Brasser, R., & Levison, H. F. 2009, Icarus, 199, 197

Koch, D. G., Borucki, W. J., Basri, G., Batalha, N. M.and Brown, T. M., Caldwell, D., Christensen-Dalsgaard, J., & Cochran, W. D. amd DeVore, E. 2010, ApJ, 713

Kokubo, E., & Ida, S. 1998, Icarus, 131, 171

Kopparapu, R. K. 2013, ApJ, 767, L8

Kopparapu, R. k., Ramirez, R. M., SchottelKotte, J., Kasting, J. F., Domagal-Goldman, S., & Eymet, V. 2014, ArXiv e-prints

Kopparapu, R. K., et al. 2013, ApJ, 765, 131

Kreidberg, L., et al. 2014, Nature, 505, 69

Laskar, J. 1994, A&A, 287, L9

Laskar, J., & Gastineau, M. 2009, Nature, 459, 817

Laskar, J., & Robutel, P. 1995, Celestial Mechanics and Dynamical Astronomy, 62, 193

Lecar, M., Franklin, F. A., Holman, M. J., & Murray, N. J. 2001, ARA&A, 39, 581

Lee, H., Fabrycky, D., & Lin, D. N. C. 2013, ApJ, 774, 8

Lee, M. H., & Peale, S. J. 2002, ApJ, 567, 596

Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. 1996, Nature, 380, 606

Lin, D. N. C., & Papaloizou, J. 1986, ApJ, 309, 846

Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., Steffen, J. H., & Ford, E. B. 2011, ApJ, 197, 26

Lithwick, Y., & Wu, Y. 2011, ApJ, 739, 31

—. 2012, ApJ, 756, 5

Liu, C., Bailer-Jones, C. A. L., Sordo, R., Vallenari, A., Borrachero, R., Luri, X., & Sartoretti, P. 2012, ApJ, 426

Lopez, E. D., Fortney, J. J., & Miller, N. 2012, ApJ, 761, 59

Malhotra, R. 1993, Nature, 365, 819

—. 1995, AJ, 110, 420

Mann, A. W., Gaidos, E., & Ansdell, M. 2013, ApJ, 779, 188

Mann, A. W., Gaidos, E., Lépine, S., & Hilton, E. J. 2012, ApJ, 753, 90

Marchal, C., & Bozis, G. 1982, Celestial Mechanics, 26, 311

Marcus, R. A., Stewart, S. T., Sasselov, D., & Hernquist, L. 2009, ApJ, 700, L118

Marshall, J., Horner, J., & Carter, A. 2010, International Journal of Astrobiology, 9, 259

Marzari, F. 2014, MNRAS, 442, 1110

Marzari, F., & Weidenschilling, S. J. 2002, Icarus, 156, 570

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

Meschiari, S., Wolf, A. S., Rivera, E., Laughlin, G., Vogt, S., & Butler, P. 2009, PASP, 121, 1016

Milani, A., & Nobili, A. M. 1983, Celestial Mechanics, 31, 213

Moore, A., Hasan, I., & Quillen, A. C. 2013, MNRAS, 432, 7

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

Morton, T. D., & Johnson, J. A. 2011, ApJ, 738, 170

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

NASA Exoplanet Archive. 2017, Nasa Exoplanet Archive

Olivier, C., Simon, G., & Eric, M. 2007, SPIE

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

Pál, A. 2009, MNRAS, 396, 1737

Papaloizou, J. C. B., & Larwood, J. D. 2000, MNRAS, 315, 823

Peterson, P. 2009, Int. J. of Computational Science and Engineering, 4, 296

Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, Proceedings of the National Academy of Science, 110, 19273

Petrovich, C. 2015, The Astrophysical Journal, 808, 120

Pierrehumbert, R., & Gaidos, E. 2011, ApJ, 734, L13

Pinsonneault, M. H., An, D., Molenda-Żakowicz, J., Chaplin, W. J., Metcalfe, T. S., & Bruntt, H. 2012, ApJS, 199, 30

Plummer, H. C. K. 1918, An introductory treatise on dynamical astronomy

Pollack, J. B., Hubickyj, O., Bodenheimer, P., Lissauer, J. J., Podolak, M., & Greenzweig, Y. 1996, Icarus, 124, 62

Press, W., Flannery, B., Teukolsky, S., & Vetterling, W. 1988, Numerical Recipes in C (Cambridge University Press)

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2002, Numerical recipes in C++ : the art of scientific computing

Pu, B., & Wu, Y. 2015, ApJ, 807, 44

Rafikov, R. R. 2003, AJ, 125, 942

—. 2006, ApJ, 648, 666

Ramirez, S., Akeson, R. L., Ciardi, D. R., Chen, X., Christiansen, J., & Plavchan, P. 2014, Å

Raymond, S. N., & Cossou, C. 2014, MNRAS

Rein, H. 2012, MNRAS, 427, 5

Rein, H., & Liu, S.-F. 2012, A&A, 537, 10

Rein, H., & Papaloizou, J. C. B. 2009, A&A, 497, 595

Rein, H., Papaloizou, J. C. B., & Kley, W. 2010, A&A, 510, A4

Rein, H., & Spiegel, D. S. 2015, MNRAS, 446, 1424

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

—. 2017, MNRAS

Robertson, P., et al. 2012, ApJ, 749, 39

Rowe, J. F., Bryson, S. T., Marcy, G. W., Lissauer, J. J., & Jontof-Hutter, D. 2014, ApJ, 784, 20

Saha, P., & Tremaine, S. 1992, AJ, 104, 1633

—. 1994, AA, 108, 1962

Santerne, A., Fressin, F., Díaz, R. F., Figueira, P., Almenara, J.-M., & Santos, N. C. 2013, A&A, 557, A139

Santerne, A., et al. 2012, A&A, 545, A76

Sasselov, D. D., & Lecar, M. 2000, ApJ, 528, 995

Schlichting, H. E. 2014, ApJ, 795, L15

Schneider, G., et al. 2014, AJ, 148, 59

Schneider, J., Dedieu, C., Le Sidaner, P., Savalle, R., & Zolotukhin, I. 2011, AA

Silburt, A., Gaidos, E., & Wu, Y. 2015, ApJ, 799, 180

Silburt, A., & Rein, H. 2015, MNRAS, 453, 4089

—. 2017, ArXiv e-prints

Smith, A. W., & Lissauer, J. J. 2009, Icarus, 201, 381

Smith, A. W., & Lissauer, J. J. 2009, Icarus, 201, 381

Steffen, J. H., & Hwang, J. A. 2015, MNRAS, 448, 16

Steffen, J. H., et al. 2013, MNRAS, 428, 1077

Stevenson, K. B., Bean, J. L., Seifahrt, A., Gilbert, G. J., Line, M. R., Désert, J.-M., & Fortney, J. J. 2016, ApJ, 817, 141

Sussman, G. J., & Wisdom, J. 1988, Science, 241, 433

Tamayo, D. 2014, MNRAS, 438, 3577

Tamayo, D., Triaud, A. H. M. J., Menou, K., & Rein, H. 2015, ApJ, 805, 100

Tamayo, D., et al. 2016, ApJ, 832, L22

Tanaka, H., Takeuchi, T., & Ward, W. R. 2002, ApJ, 565, 1257

Tian, F., Toon, O. B., Pavlov, A. A., & De Sterck, H. 2005, ApJ, 621, 1049

Traub, W. A. 2012, ApJ, 745, 20

Tsiaras, A., et al. 2016, ApJ, 820, 99

Veras, D., & Mustill, A. J. 2013, Monthly Notices of the Royal Astronomical Society: Letters, slt067

Verner, G. A., et al. 2011, ApJ, 738, L28

Ward, W. R. 1997, Icarus, 126, 261

Weidenschilling, S. J. 1977a, MNRAS, 180, 57

—. 1977b, Ap&SS, 51, 153

Weiss, L. M., & Marcy, G. W. 2014, ApJL, 783, 7

Wetherill, G. W., & Stewart, G. R. 1989, Icarus, 77, 330

Winn, J. N. 2010, ArXiv e-prints

Wisdom, J. 1980, AJ, 85, 1122

—. 2006, AJ, 131, 2294

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

Wolfgang, A., & Lopez, E. 2014, ApJ

Wu, Y., & Lithwick, Y. 2013, ApJ, 772, 74

Wu, Y., & Murray, N. 2003, ApJ, 589, 605

Yoshida, H. 1990, Science Direct, 150, 6

—. 1993, AA, 56, 16

Youdin, A. N. 2011, ApJ, 742, 38

Appendix A

Appendix for Chapter 3

A.1 Calculation of Minimum Eccentricity Required for Inner Planet Given Δ_{obs} and T Years

Starting from Eq. 3.3, we now derive Eq. 3.8:

$$\frac{\dot{a}}{a} = 2e\dot{e}$$

$$\frac{da}{a} = 2e^2\frac{\dot{e}}{e}dt \tag{A.1}$$

Since $\tau_e \equiv -\dot{e}/e$ is a constant, the only quantity with a time dependence on the right hand side is e. Integrating Eq. 3.1 gives us an expression for $e(t)$:

$$\frac{de}{e} = -\frac{9}{2}\pi\frac{k}{Q}\frac{1}{m_p}\sqrt{\frac{GM^3}{a^3}}\left(\frac{r_p}{a}\right)^5 dt = \frac{\dot{e}}{e}dt = -\frac{1}{\tau_e}dt$$

Since (again) we assume $\tau_e \equiv -e/\dot{e}$ is a constant, the integration is straightforward to yield:

$$e(t) = e_i \exp\left(-\frac{1}{\tau_e}t\right) \tag{A.2}$$

Now that we have eccentricity as a function of time, we can plug Eq. A.2 into Eq. A.1 and integrate to get:

$$\int_{a_i}^{a_f}\frac{da}{a} = 2e_i^2\int_0^T\frac{-1}{\tau_e}\exp\left(-2\frac{1}{\tau_e}t\right)dt$$

$$\ln(a_i/a_f) = e_i^2\left(1-\exp(-2\frac{1}{\tau_e}T)\right) \tag{A.3}$$

We rearrange Eq. A.3 for e_i to get:

$$e_i = \sqrt{\frac{\ln(a_i/a_f)}{1-\exp(-2T/\tau_e)}} \tag{A.4}$$

Where all quantities refer to a single planet (e.g. the inner planet). To connect Eq. A.4 to a 2:1 MMR

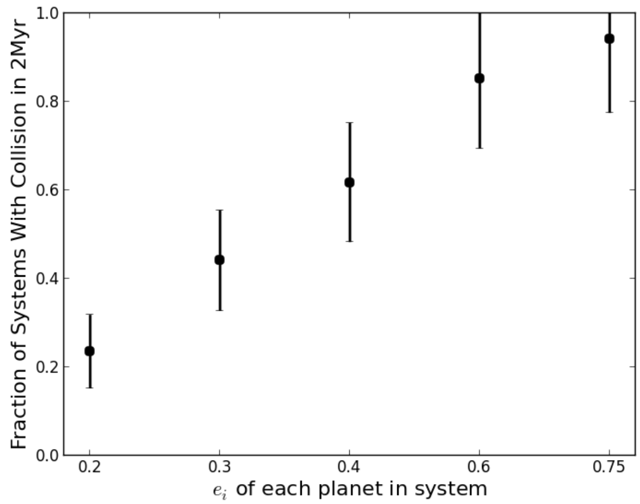

Figure A.1: The fraction of *Kepler* systems in our sample that go unstable within 2Myr if each planet is given an initial eccentricity of e_i. The error bars derived from Poisson statistics.

pair, we plug in Eq. 3.6 along with the fact that $a_{in,i} = a_{out,i}/2^{2/3}$ (for 2:1 MMR) to get:

$$e_{in,i} = \sqrt{\frac{\ln(a_i/a_f)_{in}}{1-\exp(-2T/\tau_{e,in})}}$$

$$e_{in,i} = \sqrt{\frac{\ln[(\frac{a_{out,i}}{2^{2/3}})(\frac{(\Delta_{obs}+2)^{2/3}}{a_{out,f}})]}{1-\exp(-2T/\tau_{e,in})}}$$

$$e_{in,i} = \sqrt{\frac{\ln[(\frac{a_i}{a_f})_{out}(\frac{\Delta_{obs}+2}{2})^{2/3}]}{1-\exp(-2T/\tau_{e,in})}} \tag{A.5}$$

Which is the result displayed in Eq. 3.8.

A.2 Maximum Eccentricity From Dynamical Simulations

In constructing the blue shaded region in Figure 3.4, we have performed numerical simulations of stability for each Kepler system. We simulate each Kepler system for 2 Myr, and assign the same initial eccentricity e_i to each planet in the system. We do not migrate these planets into resonance, but since (in general) resonance can add a destabilizing effect (Chambers et al. 1996b; Funk et al. 2010; Pu & Wu 2015), these simulations can be used as conservative upper limits for the maximum eccentricity that a system can have and remain stable. The results are presented in Figure A.1, and error bars are derived from Poisson statistics.

Each point in Figure A.1 represents the fraction of systems in our sample having a collision within 2 Myr. We can see that for $e_i = 0.3$, about half of the Kepler systems have had a collision. Thus from stability arguments, the maximum initial eccentricity that planets in our Kepler sample can have in order for $\sim 50\%$ of them to survive at least 2 Myr is $e_{max} \approx 0.3$.

A.3 Planet Sample

Table A.1: Kepler Systems Used In This Analysis

System	Planet	P (days)	near MMR	$m_p/m_\oplus$	$r_p/r_\oplus$	N_p	$M/M_\odot$	$R/R_\odot$
KOI-142	b	10.95	✓	8.7 ± 2.5	3.82 ± 0.44	2	0.96 ± 0.04	0.88 ± 0.03
	c	22.34	✓	198.8 ± 9.2	6.82 ± 1.09			
Kepler-120	b	6.31	✓		2.18 ± 0.22	2		0.53 ± 0.03
	c	12.79	✓		1.53 ± 0.11			
Kepler-127	b	14.44	✓		1.42 ± 0.11	3		1.36 ± 0.04
	c	29.39	✓		2.62 ± 0.11			
	d	48.63			2.62 ± 0.11			
Kepler-176	b	5.43			1.42 ± 0.76	3		0.89 ± 0.46
	c	12.76	✓		2.62 ± 1.31			
	d	25.75	✓		2.51 ± 1.31			
Kepler-183	b	5.69	✓		2.07 ± 0.87	2		0.96 ± 0.41
	c	11.64	✓		2.29 ± 0.98			
Kepler-221	b	2.8	✓		1.75 ± 0.22	4	0.72 ± 0.05	0.82 ± 0.07
	c	5.69	✓		2.95 ± 0.33			
	d	10.04			2.73 ± 0.22			
	e	18.37			2.62 ± 0.22			
Kepler-244	b	4.31			2.73 ± 1.2	3		0.8 ± 0.34
	c	9.77	✓		2.07 ± 0.87			
	d	20.05	✓		2.29 ± 0.98			
Kepler-25	b	6.24	✓	9.0 ± 2.4	2.62 ± 0.0	3	1.19 ± 0.06	1.31 ± 0.02
	c	12.72	✓	14.3 ± 2.7	4.48 ± 0.0			
	d	123.0		89.9 ± 13.7	5.46 ± 0.0			
Kepler-267	b	3.35	✓		1.97 ± 0.11	3	0.56 ± 0.05	0.56 ± 0.02
	c	6.88	✓		2.07 ± 0.11			
	d	28.46			2.29 ± 0.11			
Kepler-27	b	15.33	✓	41.8 ± 5.0	4.04 ± 0.0	2	0.65 ± 0.16	0.59 ± 0.15
	c	31.33	✓	21.2 ± 3.2	4.91 ± 0.0			
Kepler-272	b	2.97	✓		1.42 ± 0.76	3	0.79 ± 0.05	0.93 ± 0.5
	c	6.06	✓		1.75 ± 0.98			
	d	10.94			2.29 ± 1.2			
Kepler-30	b	29.33	✓	11.3 ± 1.4	3.93 ± 0.22	3	0.99 ± 0.08	0.95 ± 0.12
	c	60.32	✓	640.0 ± 50.0	12.34 ± 0.44			
	d	143.34		23.1 ± 2.7	8.84 ± 0.55			
Kepler-305	b	5.49		10.5 ± 2.6	3.6 ± 0.87	3	0.76 ± 0.13	0.79 ± 0.05
	c	8.29	✓	6.0 ± 2.4	3.28 ± 0.76			
	d	16.74	✓		2.73 ± 0.44			

Kepler-32	f	0.74			0.76 ± 0.11	5	0.54 ± 0.02	0.53 ± 0.02
	e	2.9	✓		1.53 ± 0.11			
	b	5.9	✓	9.4 ± 3.6	2.18 ± 0.22			
	c	8.75		7.7 ± 5.0	1.97 ± 0.22			
	d	22.78			2.73 ± 0.11			
Kepler-326	b	2.25	✓		1.53 ± 0.22	3	0.98 ± 0.05	0.8 ± 0.05
	c	4.58	✓		1.42 ± 0.11			
	d	6.77			1.2 ± 0.11			
Kepler-327	b	2.55	✓		1.09 ± 0.11	3	0.55 ± 0.05	0.49 ± 0.02
	c	5.21	✓		0.98 ± 0.11			
	d	13.97			1.75 ± 0.11			
Kepler-328	b	34.92	✓	28.5 ± 12.9	2.29 ± 0.98	2	1.15 ± 0.22	1.06 ± 0.44
	c	71.31	✓	39.4 ± 13.6	5.46 ± 2.29			
Kepler-384	b	22.6	✓		1.09 ± 0.33	2	0.76 ± 0.05	0.88 ± 0.25
	c	45.35	✓		1.09 ± 0.33			
Kepler-386	b	12.31	✓		1.42 ± 0.76	2	0.74 ± 0.05	0.77 ± 0.43
	c	25.19	✓		1.64 ± 0.87			
Kepler-396	b	42.99	✓	75.5 ± 11.8	3.49 ± 1.31	2	0.85 ± 0.13	1.06 ± 0.39
	c	88.5	✓	17.9 ± 2.8	5.35 ± 1.97			
Kepler-48	b	4.78	✓	14.3 ± 4.3	2.18 ± 0.0	3	0.88 ± 0.06	0.89 ± 0.05
	c	9.67	✓	9.8 ± 3.3	3.17 ± 0.0			
	d	42.9		7.93 ± 4.6	2.07 ± 0.11			
Kepler-56	b	10.5	✓	22.1 ± 3.9	6.55 ± 0.33	2	1.32 ± 0.13	4.23 ± 0.15
	c	21.4	✓	181.0 ± 21.0	9.83 ± 0.44			
Kepler-57	b	5.73	✓	118.1 ± 24.1	2.18 ± 0.0	2	0.83 ± 0.05	0.73 ± 0.0
	c	11.61	✓	7.4 ± 9.4	1.53 ± 0.0			
Kepler-79	b	13.48	✓		2.62 ± 0.76	4	1.1 ± 1.63	1.4 ± 0.25
	c	27.4	✓		2.73 ± 0.87			
	d	52.09			7.64 ± 1.42			
	e	81.07			3.38 ± 0.66			
Kepler-81	b	5.96	✓		2.4 ± 0.44	3	0.64 ± 0.38	0.59 ± 0.03
	c	12.04	✓		2.4 ± 0.33			
	d	20.84			1.2 ± 0.33			
Kepler-83	d	5.17			1.97 ± 0.11	3	0.66 ± 0.41	0.59 ± 0.03
	b	9.77	✓		2.84 ± 0.44			
	c	20.09	✓		2.4 ± 0.33			
Kepler-9	d	1.59			1.64 ± 0.22	3	1.07 ± 0.05	1.02 ± 0.05
	b	19.24	✓	80.09 ± 4.13	9.5 ± 0.76			
	c	38.91	✓	54.35 ± 4.13	9.28 ± 0.76			