The Shapes of Planet Transits and Planetary Systems

Emily Ruth Sandford

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2020

ABSTRACT

The Shapes of Planet Transits and Planetary Systems

Emily Ruth Sandford

In this Thesis, I explore transiting exoplanets: what we can learn from modeling their light curves, and what we can learn from their arrangement in planetary systems. I begin in Chapter 1 by briefly reviewing the history of transit modeling, from the earliest theoretical models of eclipsing binary stars to the models in current widespread use to model exoplanet transits. In Chapter 2, I model the transits of a sample of *Kepler* exoplanets with strong prior eccentricity constraints in order to derive correspondingly strong constraints on the density of their host stars, and find that the density constraints I derive are as precise as density constraints from asteroseismology if the transits are observed at high signal-to-noise. In Chapter 3, I apply the same methodology in reverse: using prior knowledge of the stellar density based on *Gaia* parallax measurements, I model the transits of twelve singly-transiting planets observed by K2 and derive constraints on their periods. In Chapter 4, I consider the general problem of deducing the shape of a transiting object from its light curve alone, which I term "shadow imaging;" I explore the mathematical degeneracies of the problem and construct shadow images to explain Dips 5 and 8 of Boyajian's Star.

I next turn to multi-planet systems: in Chapter 5, I investigate the underlying multiplicity distribution of planetary systems orbiting FGK dwarfs observed by *Kepler*. I find that we can explain the multiplicities of these systems with a single Zipfian multiplicity distribution, without invoking a dichotomous population. In Chapter 6, I consider the arrangement of planets in those systems, and use neural networks inspired by models used for part-of-speech tagging in computational linguistics to model the relationship between exoplanets and their surrounding "context," i.e. their host star and sibling planets. I find that our trained regression model is able to predict the period and radius of an exoplanet to a factor of two better than a naive model which only takes into account basic dynamical stability. I also find that our trained classification model identifies consistent classes of planets in the period-radius plane, and that it is rare for multi-planet systems to contain a neighboring pair of planets from non-contiguous classes.

In Chapter 7, I summarize these results and briefly discuss avenues for future work, including the application of our methods to planets and planetary systems discovered by TESS.

Contents

List of Tables

List of Figures

Acknowledgements

Many people have looked after me these past six years, which have been (like many exoplanets) rocky but happy. I am very grateful.

First, thanks to my PhD advisor, David Kipping, the most imaginative and enthusiastic scientist I know. I have learned so much from him about how to have ideas and how to see them through. I feel very lucky to have had such a dedicated and encouraging advisor, and it has been a privilege to be part of the Cool Worlds group.

Thanks to my first-year project advisor, Kathryn Johnston, who has been not only extremely supportive to me from my first days at Columbia, but also deeply inspirational as a leader in the department and the field. Thanks also to my co-advisor Andreas Küpper for helping me find my feet throughout my first two years.

Thanks to my collaborators Néstor Espinoza, Rafael Brahm, and Andrés Jordán, for their mentorship throughout our project and for their hospitality during my visit to Santiago. This was my first project independent of an advisor, and they were immensely supportive and patient as I figured out what I was doing.

Thanks to my collaborator Michael Collins for his guidance over the past two projects and enthusiasm for exoplanet science.

Thanks to Didier Queloz and the Cambridge Exoplanets Group for hosting me as a visiting researcher this past year and for teaching me (already!) an incredible amount about stars, RVs, and ground-based observations. Thanks especially to Annelies Mortier for helping me begin to learn about the Sun.

Thanks to my undergraduate advisors, Marla Geha and Joanne Bibby, for helping me

decide to be an astronomer.

Thanks to the other members of my thesis committee, David Schiminovich and Greg Bryan, for their guidance throughout my dissertation work. Thanks to my external examiner, David Hogg, for his thoughtful comments during my thesis defense.

Thanks to the rest of the Columbia astronomy faculty, especially: Marcel Agüeros for his support before, during, and after my trip to Chile; Jacqueline van Gorkom and David Schiminovich for their advice as I prepared to move to the UK; and Frits Paerels and Mary Putman for their support.

Thanks to Millie Garcia and Ayoune Payne, without whom I would not have been able to navigate Columbia at all.

Thanks to the wonderful graduate students I have overlapped with at Columbia—denizens of upstate cabins, Appletree Deli, the Pyramid Club, the ferry to Rockaway, Mona's, and zappy hour; talkers of a big walkie-talkie talk without walking any kind of walkie-talkie walk; most of all, dear, dear friends. Thanks especially to my cohort-siblings, Andrea Derdzinksi, José Manuel Zorrilla Matilla, and Adrian Lucy, and to my near-cohort-siblings, Nicole Melso, Alex Teachey, Daniel DeFelippis, Mihir Kulkarni, Julia Blue Bird, Kirsten Blancato, Moiya McTier, Rose Gibson, Tiffany Jansen, Jorge Cortés, and Steven Mohammed. Thanks to my officemates Aleksey Generozov, Andrew Emerick, Sarah Pearson, Tomer Yavetz, Rayna Rampalli, Shifra Mandel, Alejandro Núñez, and Ron Tso. Thanks to Sarah and Andrea (and Lilith!) for being such good roommates. Thanks to Adrian Price-Whelan for bringing me along to MDM. Thanks to my grad mentor, Miao Li. Thanks to my fellow Cool World, Jingjing Chen.

Thanks to my undergrad astronomy compatriots, Chris Cappiello and Liang Yu, and my friends from the AMNH astronomy REU. Thanks to the Astrobites team.

Thanks to my students, Catherine Atalig, Christopher Ambrus, James Hamue, and Caroline Klewinowski, who taught me much more than I taught them.

Thanks to Anne Goldfield, who helped me through the low point.

Thanks to my dear friends Lauren Walters and Sarah Connor, the greatest joys of living in New York City. Thanks to Lauren for making two years living in the worst-managed apartment in northern Manhattan so much fun and for once making a Gristedes vegetable joke so funny that it kept me going through about three months of grad school. Thanks to Sarah for coming with me on so many adventures, from the Women's March to the rainiest imaginable Shakespeare in the Park to a two-hour concert in Brooklyn that it took three hours to get home from.

Thanks to my far-flung friends from home: to Jackie Li, trailblazer extraordinaire, who introduced me to scientific research in the first place, way back when, and whose adventures in Finland (a real place?) gave me the courage to move abroad; to Katie Holliday, my pen pal, whose letters have kept me going through the lockdown; and to Amy Schaffman, who makes me feel at home even 3600 miles away.

Thanks to everyone who made New York an (intermittent) home over the past six years. Thanks to Danielle Trubow, the kindest friend and roommate. Thanks to Nicole Melso, who let me crash at her apartment more times than I can count. Thanks to my grandmother Eleanor Hochuli, who welcomes me for weeks at a time, and her wonderful carers, especially Roxana Roca.

Thanks to my mother, Janelle Sandford, for her boundless love and support, and for her example in never, ever giving up. Thanks to my father, Jim Sandford, for sparking my interest in math from the very beginning, and for calm reassurance through the worst of grad school. Thanks to my brother, Evan Sandford, for setting an example of learning, growing, and figuring things out, and for being so constantly, unflappably funny. Thanks to Kate Miller and Greg Penoyre for their support and kindness. Thanks to Magdalena Svirlova, my teacher and friend.

Thanks to the many dogs we looked after against the terms of our lease: Perlie, Vali, Oscar, Jess, Doug, Ella, Sesame, Neo, Luna, and Coco.

Lastly and mostly, thanks to Zephyr, who thinks about everything.

1. Introduction

Work stops at sunset. Darkness falls over the building site. The sky is filled with stars. "There is the blueprint," they say.

– Italo Calvino

As astronomical objects go, planets are small ($\lesssim 30R_\oplus$, or $0.3R_\odot$) and dim (absolute magnitude $\gtrsim 15.5$, where the limit is calculated for a perfectly reflective hot Jupiter orbiting a Sun-like star). Planet seekers thus find themselves in an unusual situation: it is much easier to detect a planet by an absence of light—an eclipse of the planet's host star—than a presence.

The eclipse, or transit, method of exoplanet detection was first suggested in the literature in 1952 by Otto Struve, who noted at the end of a short monograph on the possibility of detecting planets with high-precision stellar radial velocity measurements that, "There would, of course, also be eclipses" (Struve, 1952). This was a natural suggestion, given that eclipses, or transits, of Venus and Mercury across the Sun were first observed centuries ago, perhaps as early as the transit of Venus in late 1153 AD: frescoes in the Maya city of Mayapan, carbon-dated to 1200-1350 AD, depict yellow Sun-circles, occulted by richly dressed human figures, which may represent priests or gods (Galindo Trejo & Allen, 2005). Galindo Trejo & Allen note first that the occulting figures are significant, as no other extant Maya depictions of the Sun have them, and second, that Maya astronomers kept careful track of the position of Venus, so they would have known to look out for the transit.

The first recorded observation of a planetary transit by telescope was the 1631 transit of Mercury by Pierre Gassendi, which had been predicted four years prior by Johannes Kepler

(Gassendi 1632; Kepler 1630; see van Helden 1976 for a historical account). Eight years later, Jeremiah Horrocks and William Crabtree observed the 1639 transit of Venus, which Kepler had also predicted (Kepler 1630, see Chapman 1990; Kollerstrom 2005 for historical accounts).

In 1999, Charbonneau et al. detected the first exoplanet transit, of the radial velocity-discovered hot Jupiter HD 209458b (observed independently in the same year by Henry et al.). Since then, despite the intrinsic rarity of transiting alignment, more than 3000 exoplanets have been discovered in transit, comprising roughly 75% of all confirmed exoplanets.[1]

Any individual transit observation reveals much about the transiting planet, its orbit, and its host star. Together, the population of transiting planets reveals—albeit noisily, through the grimy window of observational bias—deeper truths, about the ubiquity of planets in the Universe, how they formed in their planetary systems, and how these systems have evolved. A rich body of theory has evolved alongside transit observations: from transit modeling, which predicts the shape of transit light curves as a consequence of the underlying physical properties of planets and stars, to modeling of hypothetical planetary populations to match the observed occurrence rates of planets, their grouping into systems, and the distributions of their properties.

This thesis concerns two main questions. First, what are the limits of transit modeling? What information belongs to the light curve of a transiting planet, and what can we learn by fitting the shape of that light curve as precisely as possible? Second, what information belongs to the arrangement of planets in a system, and what can the transiting planets tell us about their unobserved neighbors?

1.1 A brief history of transit modeling

I begin by briefly outlining the history of transit modeling, or the prediction of the light curve generated when one astronomical body eclipses another. Transit modeling is

[1] NASA Exoplanet Archive confirmed planets table, accessed 15 June 2020.

conceptually simple—how much light is visible given the configuration of the two bodies from a particular vantage point at a particular time?—but far-reaching in its influence on exoplanet theory.

1.1.1 Eclipsing binary stars

Because planets are generally so small relative to their host stars, and their eclipses so shallow (of order 1%, for a Jupiter-sized planet around a Sun-sized star), eclipse modeling preceded exoplanet transit observations by nearly a century. Transit modeling—or specifically, modeling of the disk-integrated time series photometry of transits—was instead driven by observations of eclipsing binary stars, which have much more extreme eclipses, some visible to the naked eye. Algol, for example, dims by nearly 1.5 magnitudes during eclipse, or $\sim$ 70%; the first light curve of this system was published in 1783, by John Goodricke (Goodricke, 1783). An excellent history of eclipsing binary light curve modeling, from which I draw here, is given in Wilson (1994).

(That eclipsing binaries were the earliest eclipsing systems to be modeled is somewhat ironic, because stellar binaries, in which the primary and secondary bodies may (1) glow in similar wavebands; (2) be of similar sizes; (3) measurably irradiate each other; (4) exhibit limb darkening; and (5) be tidally distorted into non-spherical shapes, are generally much more complicated to model than planet-star systems, demanding a more sophisticated treatment of stellar atmospheric physics and tidal theory.)

The first eclipsing binary light curve model was published in 1912, in a series of three papers by Henry Norris Russell (Russell, 1912a,b,c). In Russell (1912a), Russell addresses the problem of deducing from an eclipsing binary light curve four quantities—the radius of the primary, the radius of the secondary, the luminosity of the primary (which can be subtracted from the total luminosity to immediately yield the luminosity of the secondary), and the orbital inclination—based on the assumptions that (1) the period is known; (2) the orbits are circular; (3) both stars are spherical; and (4) both stars are uniformly illuminated,

i.e. there is no limb darkening. He considers, furthermore, both grazing and non-grazing configurations.

In Russell (1912b), he relaxes these assumptions, first considering stars on slightly eccentric orbits (eccentricity $e \lesssim 0.1$); then stars which have been slightly tidally elongated into prolate spheroids; stars which are irradiated by their binary companions, such that the stellar hemispheres facing each other are brighter than those facing away; and linearly limb-darkened stars.

It is worth noting here, as Russell notes, that all of these complications are imprinted in the light curve. Let us label the primary eclipse with an orbital phase of 0, such that the (normalized) orbital phase of the next primary eclipse is 1. Eccentricity shows itself as an asymmetry, where the secondary eclipse falls somewhere other than a phase of exactly 0.5 (neglecting the light travel time effect and the edge case where periapse is directly aligned with our line of sight). Ellipticity of the stars causes peaks in brightness halfway between the primary and secondary eclipses, where the elongated hemisphere of the star faces us; spherical stars have a flat light curve between eclipses. If the hemisphere of each star which faces its companion is brighter than the hemisphere which faces away, that imprints a slight excess in light for each body, peaking at the moment of that body's eclipse by the other star. Finally, limb darkening affects the shape of the ingress and egress phases of each eclipse, such that the decrement in light is no longer strictly proportional to the area of overlap between the sky-projected shapes of the stars.

In Russell (1912c), Russell applies his models to three real eclipsing binary systems: ω Delphini, ω Ursae Majoris, and ω Crucis. This is therefore the first study which deduced certain fundamental physical properties of binary systems directly from their light curves, nearly 90 years before the same would be done for exoplanet transit light curves.

Russell made several improvements to these analytical models in subsequent decades (summarized in Russell & Merrill 1952), but the basic ingredients remained the same. As Russell himself noted in 1948, a major outstanding deficit of his model was its inability to

capture tidal distortion (beyond the first-order ellipticity of the prolate spheroid approximation) and the attendant gravity darkening of stars in close binaries (Russell, 1948).

Zdeněk Kopal made major contributions on this front throughout the 1940s and 1950s, summarized in Kopal (1959). Kopal studied the shapes of equipotential surfaces in close binaries, which he described in series of spherical harmonics; he studied how fluid bodies with given internal density profiles occupy those equipotential surfaces; and he calculated the photometric signatures their distorted, limb-darkened, gravity-darkened, mutually-irradiated shapes imprint in binary light curves. Beyond these considerable mathematical developments, he further introduced other physical effects not included at all in the Russell model, including stellar rotation. It is not, I think, an exaggeration to say that Kopal developed eclipsing binary theory essentially to its analytic limit, and that further progress required fast computation.

Lucy (1968) introduced the first numerical technique for calculating an eclipsing binary light curve, for the special case of a ω Ursae Majoris-like contact binary with a common convective envelope, based on calculating the intensity of the binary across a fine sky-projected grid, taking into account gravity and limb darkening, and then integrating over this grid to produce the light curve. Lucy compares a theoretical light curve produced this way to certain observed light curves of ω UMa systems, and notes the ways in which the model does not capture all the observed behavior, but does not attempt to fit any observed data with the model, in part because (he notes) the computation is "grossly inefficient."

Soon thereafter, numerical techniques improved to the point that direct fits to observational data could be made. Hill & Hutchings (1970) fit the light curve of the Algol hierarchical triple system to obtain the effective temperatures and (relative) radii of all three stellar components. To do this, they used a computational method first described in Hutchings (1968): defining the intersection of the line of sight with the surface of the star as a sort of observational "pole," they then divide the rest of the visible hemisphere into zones by "latitude" with respect to this pole. Projected onto the sky, the zones would appear as annuli on the

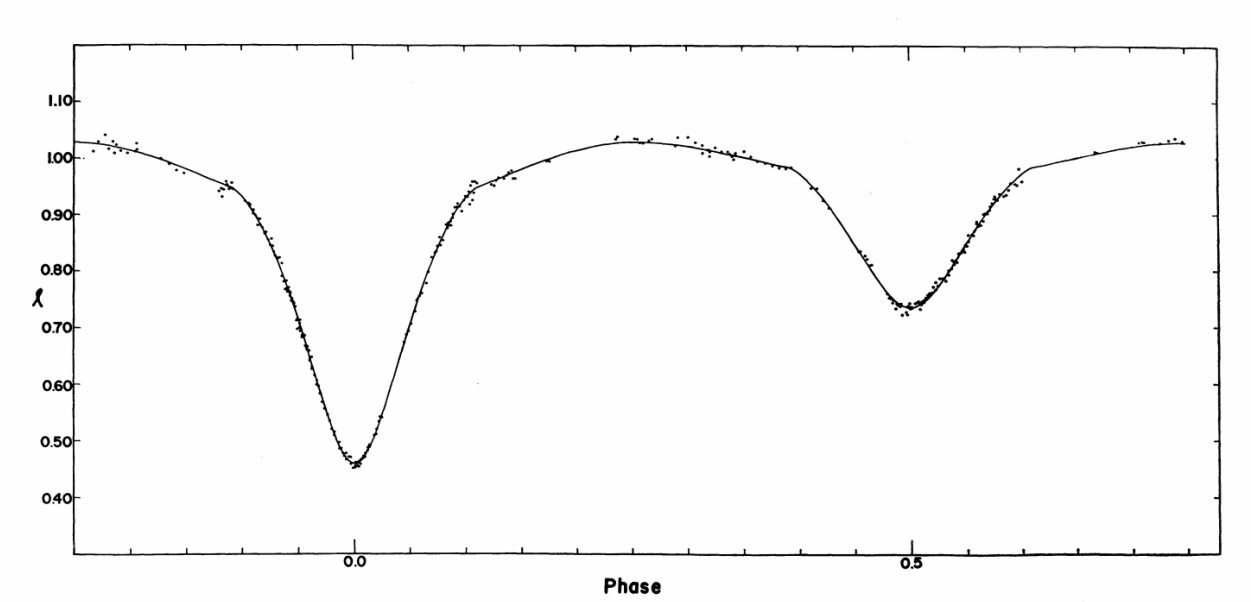

Figure 1.1: The light curve of the eclipsing binary MR Cygni, modeled by Wilson & Devinney (1971), overplotted on observations. Figure reproduced from Figure 3 of Wilson & Devinney (1971).

stellar surface, thick near the center but thinner and thinner toward the limb. They calculate the light curve by summing the intensity contributions of the zones. (For tidally distorted stars, they do the same, but take as their "pole" the substellar point, i.e. the point closest to the binary companion.)

Wilson & Devinney (1971) improve on this initial technique in several respects, including a better treatment of the shapes of tidally distorted stars, by writing a much more general method of summing over the flux emitted by the stellar surface elements that face the observer. They are also the first to use least-squares fitting to minimize the difference between their model prediction and the observed light curve data points, and the first to estimate uncertainties on a subset of their best-fit model parameters, by tweaking them as far as possible while still maintaining a "subjectively satisfactory fit to the data." Recognizing the limitations of this by-eye uncertainty estimation, they also foresee the need for later model parameter sampling techniques: "Only a rigorous adjustment procedure...is capable of properly accounting for correlations among various parameters of the model. In contrast,

the estimating technique [used here] does not reliably account for correlations, and its use will always result in error estimates which are too small." I reproduce their fit to the light curve of MR Cygni in Figure 1.1.

Other eclipsing binary light curve models of this era, each of which improved on some aspect of binary physics or computational efficiency, include Wood (1971), Nelson & Davis (1972), Mochnacki & Doughty (1972), and Budding (1977).

1.1.2 Exoplanets

Once eclipsing binary light curve modeling was on firm theoretical and numerical ground, the extension of the theory—in many ways, simplification of the theory—to modeling exoplanet light curves was inevitable, even as the observation of an exoplanet transit was still widely considered wildly unlikely. The barrier was not photometric precision, as a Jupiter-analog around a Sun-analog would produce an eclipse depth of $\sim 1\%$, or 0.01 mag, an observation "considered feasible" long before it was made, to quote Rosenblatt (1971).

Rather, before the discovery of 51 Pegasi b (period 4.22 days; semi-major axis 0.05 AU) in radial velocity data by Mayor & Queloz (1995), no one anticipated the discovery of *close-in* Jupiter-sized exoplanets. As Mayor & Queloz note, "The very small distance between the companion and 51 Peg is certainly not predicted by current models of giant planet formation." Because transit probability scales $\sim R_*/a$, Jupiter-sized planet like 51 Peg b, at 0.05 AU, is 100 times more likely to be in transiting alignment than a true Jupiter analog, at ~ 5 AU; there is the additional problem that a true Jupiter analog would have a nearly 12-year orbital period, compared to a transit duration of only ~ 30 hours, so the probability of catching such a transit is extremely small. Rosenblatt summarized the outlook as of 1971: "the probability of observing an eclipse of a planet in Jupiter's orbit, seen from a random direction and at a random time, is negligibly small, so this method has also been dismissed as impractical."

Borucki & Summers (1984) were the first to seriously suggest an observational program

capable of detecting exoplanet transits, despite their presumed rarity. Assuming an occurrence rate of approximately one Jupiter-analog per Sun-like star, they concluded that it would be necessary to monitor of order 10^4 stars, continuously, with a photometric precision of 0.1% (1 part per thousand), and hoped for a "detection rate of one planet per year of observation." They noted that photometers were already precise enough in general to detect Jupiter-sized planets from the ground, but that terrestrial planets would likely be lost in atmospheric scintillation noise, so a space telescope would be necessary to find them. A space telescope would also be capable of watching for the entire 30-hour transit duration of a Jupiter analog, where ground-based observations would not. (This proposal became the *Kepler* space telescope, launched 25 years later to monitor $\sim 10^5$ stars at ~ 10 ppm precision (Borucki et al., 2010). *Kepler* is responsible for ~ 2400 confirmed exoplanet discoveries per the NEA at the time of writing,[1] which averages out to roughly 600 planets per year of its four-year primary mission lifetime.)

Still, it took the discovery of 51 Peg b in 1995, and the discovery of another two dozen Jupiter-mass planets in RV over the next four years, to awaken serious interest in modeling exoplanet transit light curves. Sackett (1999) introduces explicitly the equations to calculate the light curve of a star of known radius with arbitrary limb darkening, occulted by a dark planet of known radius. In the following year, Charbonneau et al. (2000) used these equations to fit the first observed exoplanet transit light curve, that of HD 209458b. I reproduce the model light curves calculated by Sackett for Earth-, Jupiter-, and 51 Peg b-analogs orbiting a uniformly bright Sun-analog, showing the characteristic flat-bottomed, near-trapezoidal shape of a transit across a uniformly bright star, in Figure 1.2. (The light curves are not true trapezoids, however, because the planet and star are sky-projected circles, and their area of overlap does not change linearly during ingress and egress.)

Seager & Mallén-Ornelas (2003) further develop the theory of transit modeling in the idealized case of a small ($M_p \ll M_*$), opaque, spherical planet on a circular orbit around a uniformly bright spherical star. They demonstrate that under these assumptions, there

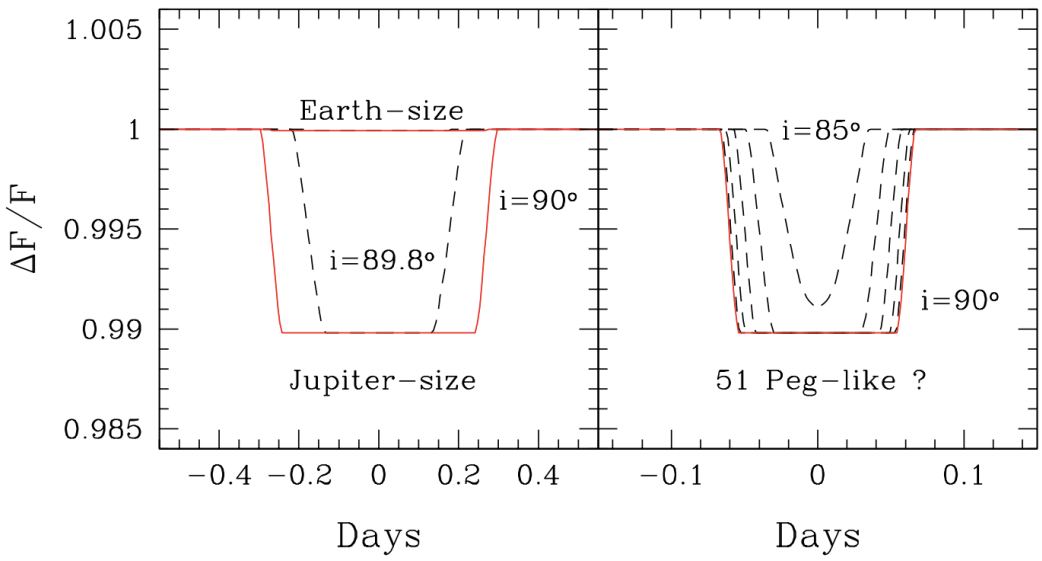

Figure 1.2: Left panel: Light curve models of Earth- and Jupiter-analogs orbiting a Sun-analog with no limb darkening at 1 AU, for inclination $i = 90°$ (solid red line) and $i = 89.8°$ (dashed black line). Right: A Jupiter-sized planet orbiting a Sun-analog with no limb darkening at 0.05 AU, analogous to 51 Peg b, at a range of inclinations. Figure reproduced from Figure 8 of Sackett (1999).

exists a unique, analytic solution for the ratio-of-radii R_p/R_*, the impact parameter b, the semi-major axis in units of stellar radii a/R_*, and the stellar density ρ_*, directly from four observables of the transit light curve: the transit depth ΔF, the total transit duration t_T, the duration of the flat-bottomed part of the transit (during which the planet is completely in front of the star) t_F, and the period P. They further demonstrate that these analytically-derived parameters can be translated into the physical parameters M_*, R_*, i, a, and R_p if a mass-radius relationship is assumed for the host star.

Mandel & Agol (2002) derive analytic expressions for the transit light curve in the case of uniform, quadratic, and four-parameter nonlinear limb darkening. Light curves calculated by these formulae, with limb darkening, assume a rounded U-shape, rather than the flat-bottomed shape of a transit without limb darkening, because the amount of light occulted by the planet now depends on the planet's position on the stellar disk instead of strictly on the area of planet-star overlap, so the light curve changes continuously as the planet crosses

the disk of the limb-darkened star. Their expressions are valid for any exoplanet orbit, as long as the sky-projected separation of centers of the star and planet over the course of the transit can be expressed as a function of time. Mandel & Agol also calculate a very useful small-planet approximation to their equations, in which case the stellar surface brightness in the annulus spanned by the planet disk is assumed to be constant, for the case when $R_p/R_* \lesssim 0.1$. I reproduce their results in Figure 1.3.

Around this time, exoplanet researchers were increasingly interested in how these developments in transit modeling could be used for automated transit detection, vetting, and prioritization of candidates for follow-up radial velocity observation. Both of the above works were explicitly concerned with large-scale transit searches. In addition to fitting of observed transit candidates, Mandel & Agol (2002) intended their formulae to be used for injection/recovery simulations of exoplanet light curves. Seager & Mallén-Ornelas (2003) highlight, in their paper, several immediately useful applications of their analytic work to transit surveys: first, deriving ρ_* directly from the light curve meant that low-density giant stars hosting similarly giant companions could be identified and discarded as interesting follow-up candidates; second, comparing this transit-derived ρ_* to a value estimated from spectroscopy plus isochrone modeling, and identifying discrepancies, could pick out sources that were likely to be blended with a background star; and third, solving for P using their formulae plus an independent estimate of ρ_* to their formulae could constrain the period of of a planet only seen to transit once, or a "single transiter."

Giménez (2006) derives another set of analytic expressions for transit light curves, based on later eclipsing binary star theory developed in Kopal (1979). These equations assume only that the sky-projected star and planet are both well-approximated as circular disks (an assumption that will only be violated in the case of extreme tidal distortion or extreme stellar rotation), and can be used with any limb darkening law, for any choice of exoplanet orbital elements. The relationship between the physical parameters of the system and the transit observables in this formulation are transcendental, i.e. not analytically invertible,

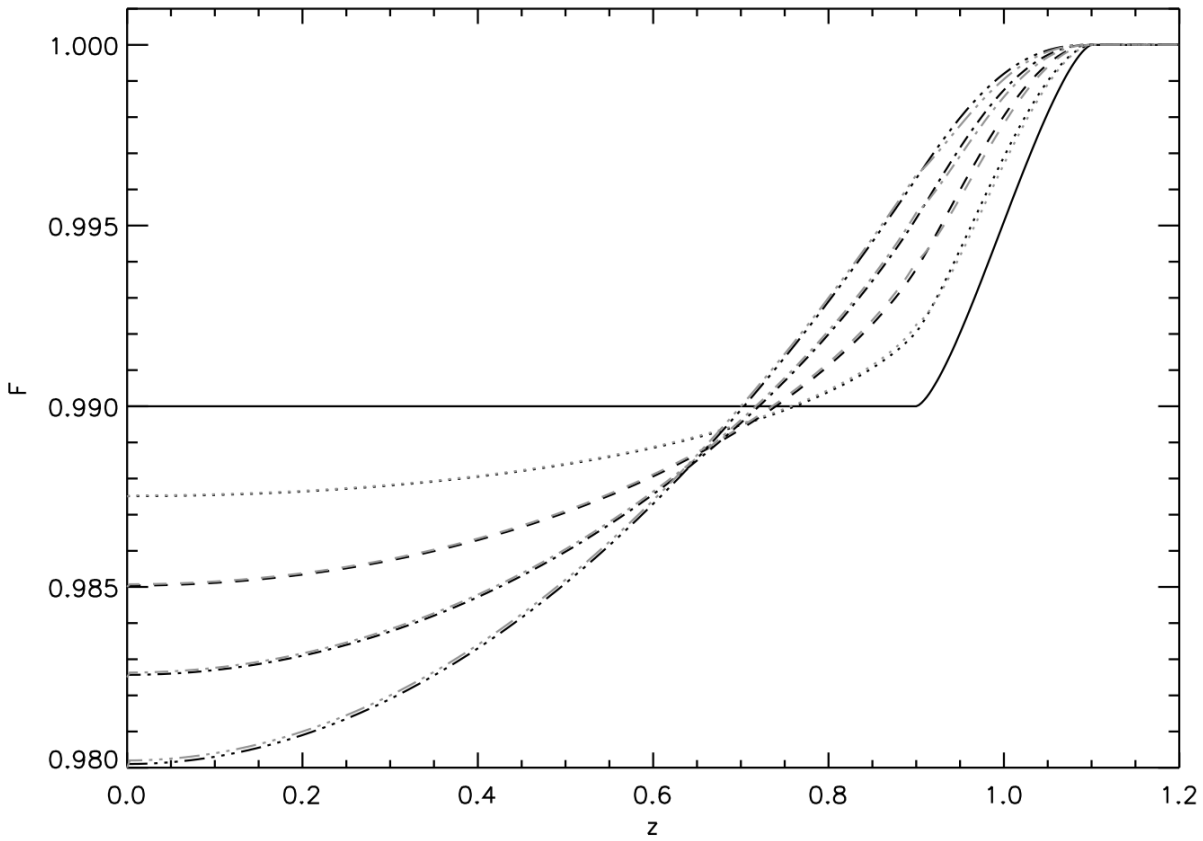

Figure 1.3: Analytically-calculated light curves for a planet with $R_p/R_* = 0.1$ by Mandel & Agol (2002). The y-coordinate F is the ratio of the in-transit flux to the out-of-transit flux; the x-coordinate z is the normalized on-sky separation of centers of the planet and star, d/R_*. Various choices of limb darkening are plotted: uniform (solid line); four-parameter nonlinear with all coefficients equal to 0 except $c_1 = 1$ (dotted line), $c_2 = 1$ (dashed line), $c_3 = 1$ (dash-dot line), and $c_4 = 1$ (dash-triple dot line). In light gray are the corresponding light curves under their small-planet approximation, where the stellar surface brightness is assumed to be constant in the annulus spanned by the planet disk. Figure reproduced from Figure 2 of Mandel & Agol (2002).

so the observed transit light curve needs to be fit (as has since become common practice regardless of the choice of transit model formulation).

Since then, further work has been done on modeling more exotic transit configurations—Pál (2012), for example, considers the general case of an arbitrary number of transiting bodies, in which mutual transits (i.e., multiple bodies transiting at the same time) may occur. Many studies have also modeled second-order effects in light curves, such as moons (Kipping, 2011), rings (Barnes & Fortney, 2004), nightside emission (Kipping & Tinetti, 2010), planet oblateness (Barnes & Fortney, 2003), atmospheric refraction (Hubbard et al., 2001), and exomountains (McTier & Kipping, 2018). Significant progress has also been made on modeling phase curves, which are light curves folded on the period of a known exoplanet that may reveal subtler photometric signals, such as those due to tides raised on the host star by the planet, Doppler beaming of the starlight, or reflection or thermal re-emission of starlight by the planet (see e.g. Faigler & Mazeh 2011 for a circular-orbit model, or Penoyre & Sandford 2019 for an eccentric-orbit model), and are in many ways a re-introduction of the ideas explored initially in eclipsing binary star modeling to exoplanet science. For the rest of this Thesis, however, we will concern ourselves primarily with transits.

1.2 Transiting planets in systems

Many individually interesting planets have been discovered and modeled in transit—evaporating super-Earths (Charbonneau et al., 2009), circumbinary planets (Doyle et al., 2011b), hot Jupiters misaligned with the spin of their host stars (Winn et al., 2010). As a population, the thousands of transiting exoplanets contain much information about the underlying occurrence rates (see e.g. Dressing & Charbonneau 2013; Foreman-Mackey et al. 2014b; Burke et al. 2015) and distributions of properties of exoplanets (see e.g. Borucki et al. 2011 and Fulton et al. 2017 for investigations of the radius distribution; Wright et al. 2011 for a database of orbital properties; Cowan & Agol 2011 for the distribution of albedos and

heat redistribution efficiencies, etc.).

Beyond this, the transit method has revealed more than 500 confirmed multi-planet systems[1] ($\sim$ 80% of all known multi-planet systems, as of writing), where two or more planets transit the same host star. Multi-planet systems, and the comparison of multi-planet systems to ostensibly single-planet systems, have much to teach us about how planetary systems form and dynamically evolve (see Winn & Fabrycky 2015 for a review), and about how (un)usual our Solar System is.

One of the most basic questions to ask of multi-planet systems is: what is the intrinsic distribution of multiplicities? How common are systems with a single planet, or two, or three, and do systems of different multiplicity represent separate populations? The question is interesting because the multiplicity distribution is an important test of theories of planetary formation and evolution. If high-multiplicity systems like Kepler-11 (Lissauer et al., 2011b) or TRAPPIST-1 (Gillon et al., 2017) are common, it would indicate that many planetary systems lead dynamically quiet lives, where the present-day arrangement of planets results from secular processes like disk migration, whereas if these systems are intrinsically rare, it would indicate that disruptive events like close encounters and planet-planet scatterings are common. Several authors have interpreted the observed multiplicities of *Kepler* systems to indicate that planetary systems fall into two underlying populations (the "*Kepler* dichotomy"): one population of dynamically cold, near-coplanar multi-planet systems, and one of single-planet or highly-mutually-inclined, dynamically hot multi-planet systems (Lissauer et al., 2011a; Ballard & Johnson, 2016).

The architectures of multi-planet systems contain further information about their formation and history, and the arrangement of planets in multi-planet systems has been studied from a number of angles. As Gilbert & Fabrycky (2020) note, the most common approach to understanding system arrangement is to compare pairs of planets within systems: what do the period or size ratios of adjacent planets in a system tell us, for example? This approach has revealed several interesting features of the planet population—for example, both

Lissauer et al. (2011a) and Fabrycky et al. (2014) find evidence of an excess of planets just outside of low-order mean-motion resonance (compared to a random distribution of period ratios), for example, which could indicate (among other hypotheses) that resonant trapping is common in planetary systems, but that the inner planet of a resonant pair often sinks closer to the star because of tidal dissipation (Terquem & Papaloizou, 2007). Weiss et al. (2018), meanwhile, find that planets tend to be similar in size to their immediate neighbors ("peas in a pod"), and that the period ratios of adjacent pairs of planets are correlated with each other, meaning that planets in high-multiplicity systems are regularly spaced. (Zhu (2020) calls these findings of intra-system regularity into question, claiming that they result instead from *Kepler*'s detection biases. However, Weiss & Petigura (2020) counter that Zhu's analysis begs the question: Zhu simulates a distribution of planet radii by sampling transit signal-to-noise, then analyzes the effect of detection bias on that population as if the radius draws were independent, which they are not.)

by generating a distribution of planet radii by sampling transit signal-to-noise and then studying the effects of detection bias on these radii as if they were independent draws from the underlying distribution.)

Gilbert & Fabrycky (2020) take a different approach, advocating for a view of exoplanet science where the planetary system, rather than the individual planet, is the fundamental unit of study. Instead of comparing pairs of planets within a system, they define a number of scalar statistics, such as degree of mass partitioning, that each captures some property of the system overall.

I advocate here for a similar view: that a planetary system is more than the sum of its constituent planets, and that the arrangement, in and of itself, deserves our attention.

1.3 Outline of thesis

In what follows, we address several of the questions introduced above. Broadly, chapters 2-4 explore the limits of what we can learn from precise transit modeling, and chapters 5-6 explore the arrangements of *Kepler*'s multi-planet systems, and what we can learn by applying techniques from computational linguistics to these systems.

In Chapter 2 (published as Sandford & Kipping 2017), we measure the stellar density ρ_* of 66 *Kepler* host stars directly from the light curves of their transiting planets, using the method originally set forth in Seager & Mallén-Ornelas (2003) but extended from the circular-orbit case to the general case of any transiting planet with strongly constrained eccentricity. We select target planets whose eccentricities are constrained by either secondary eclipse observations, short tidal circularization timescales, or belonging to compact multi-planet systems. For 62% of our targets, we measure ρ_* to a fractional uncertainty $\leq 5\%$, comparable to the precision achievable by asteroseismology, and further find that the precision of our ρ_* constraints depends primarily on the signal-to-noise ratio of the transit light curve.

In Chapter 3 (published as Sandford et al. 2019a), we apply the same logic in reverse, as first suggested by Yee & Gaudi (2008), to constrain the periods of 12 singly-transiting planets observed by K2. We combine Gaia parallax measurements with broad-band photometry and isochrone modeling to constrain the stellar densities of the single transiters' host stars, then fit a transit model with this stellar density prior to derive a constraint on period. When we treat eccentricity as a free parameter, we achieve a fractional period uncertainty of $94^{+87}_{-58}\%$, and when we fix $e = 0$, we achieve fractional period uncertainty $15^{+30}_{-6}\%$, a roughly threefold improvement over typical period uncertainties of previous studies of single transiters.

In Chapter 4 (published as Sandford & Kipping 2019), we consider the problem of how to infer the shape of a general transiting object from the shape of its light curve, without assuming a physical model for the object. This work was motivated by the discovery of Boyajian's Star (Boyajian et al., 2016), which exhibits deep, aperiodic transits that do not

resemble the light curves of transiting planets. We explore the degeneracies of this problem, find that we are indeed able to recover informative "shadow images" of arbitrary transiting objects, and apply our modeling to the triple transit of TRAPPIST-1c,e,f (Gillon et al., 2017) and to Dips 5 and 8 of Boyajian's Star.

In Chapter 5 (published as Sandford et al. 2019b), we investigate the underlying multiplicity distribution of *Kepler's* FGK planetary systems. We simulate catalogs of "observed" planets according to ten different multiplicity distributions—five single-population, and five dichotomous—and use approximate Bayesian computation to compare these simulated catalogs with the real *Kepler* systems. We find that the *Kepler* multiplicities are best fit by a single-population Zipfian (power-law) model, in contrast with previous work that invoked a dichotomous population to explain the high fraction of observed single-planet systems. We find that according to the best-fit Zipfian model, approximately one in two ostensibly single *Kepler* systems is expected to host at least one unseen additional planet.

In Chapter 6, we consider the arrangement of the *Kepler* multi-planet systems. We train a neural network model to infer the period and radius of a planet from its surrounding "context," that is, the properties of its host star and its sibling planets, and find that the trained model predicts both period and radius to a factor of two better than a "naive" model which considers only dynamical stability. We also adapt a model used for part-of-speech tagging in computational linguistics to classify planets into categories, and find that (1) the model identifies consistent classes in the period-radius plane, but learns nothing from the stellar dimensions; and (2) it is rare for neighboring planets in a multi-planet system to come from non-contiguous classes, which is superficially consistent with a "peas in a pod" picture, where planets in the same system resemble each other.

In Chapter 7, I conclude and suggest some directions for future work.

2. Know the planet, know the star: precise stellar densities from *Kepler* transit light curves

2.1 Introduction

Since its launch in 2009, the *Kepler* mission has discovered over 4500 transiting exoplanet candidates, nearly 2000 of which have been validated at $> 99\%$ confidence (Morton et al., 2016). Furthermore, *Kepler* transit light curve modeling (e.g., Batalha et al. 2013) has yielded precise constraints on the characteristics of these planet candidates and their orbits; the transit depth, for example, reveals the size of the planet relative to its host star.

Encoded in each transit light curve, however, is not just the character of the transiting planet, but also properties of the host star. In particular, the stellar density (ρ_*) can be derived analytically from the transit duration using Kepler's third law (Seager & Mallén-Ornelas 2003; see section 2.2.1), provided the eccentricity of the planet's orbit is well-constrained. Furthermore, the star's limb darkening behavior influences the shape of the transit light curve during planetary ingress and egress (see e.g. Knutson et al. 2007b). In other words, the star's interior and atmospheric properties manifest themselves in the shape of the transit light curve.

By fitting transit models to the *Kepler* light curves, we can measure these stellar properties very precisely. We have several motivations to measure stellar properties for a large

sample of *Kepler* hosts in this way. First, transit modeling serves as an independent check on other means of measuring stellar properties. In the case of ρ_*, such methods include asteroseismology, as well as spectroscopy plus isochrone modeling. These methods rest on different assumptions and, often, different input data.

Transit modeling also offers an independent test of stellar atmosphere theory, particularly with regard to limb darkening behavior. Such behavior is usually expressed in the form of an analytic stellar intensity profile weighted by limb-darkening coefficients (LDCs). Traditionally, LDCs are adopted from the theoretical predictions of stellar atmosphere modeling codes (see e.g. Sing 2010; Claret 2000). This practice is known to introduce biases in exoplanet parameters subsequently derived from the light curve (Espinoza & Jordán, 2015). Measuring the LDCs directly from the light curve enables an empirical check of these stellar atmosphere model predictions.

Furthermore, the commonly used quadratic limb darkening law, which has two LDCs, is known to be less accurate than laws with three or four LDCs (Kipping, 2016; Sing, 2010). By adopting a three-parameter nonlinear limb darkening law in transit light curve modeling and building up an empirical catalog of the fitted LDCs, we may address some of these inaccuracies.

We may also use transit fitting to derive stellar properties for stars that are not amenable to traditional analysis. For example, asteroseismology, which yields extremely precise constraints on ρ_* (typical fractional uncertainties $\lesssim 5\%$), is only possible for stars which are bright (Kepler-band magnitude $\lesssim 12$) and massive ($\gtrsim 1 M_\odot$) (Huber et al., 2013). Most stars are smaller and dimmer than this.

Finally, measuring stellar properties from transit light curves alone allows us to characterize planet-hosting stars without committing telescope time to obtain follow-up observations. In an era of large-scale surveys of transiting exoplanets, such efficiency will be crucial. The NASA Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2014), scheduled to launch in 2018, is expected to discover thousands of transiting planets orbiting stars observed at

two-minute cadence, but potentially tens of thousands more around other stars in its field of view (Sullivan et al., 2015). The Large Synoptic Survey Telescope (LSST; Ivezić et al. 2019), expected to begin full-scale science operations in 2023, will discover thousands more.

In this work, we fit transit models to a large sample of *Kepler* host stars to build an empirical catalog of transit-derived stellar densities and limb darkening coefficients and demonstrate that this method is capable of delivering precise constraints on these stellar parameters. In Section 2.2, we describe our host star target selection and detail our data analysis, including data processing, detrending, and Markov chain Monte Carlo (MCMC) approach to fitting the transit model. In Section 2.3, we present results of this analysis, including the full posterior distributions of the stellar density and LDCs. We specifically discuss the types of planet-star systems for which this method succeeds in producing high-precision constraints on stellar density in Section 2.3.3.1. We conclude and highlight this approach's potential to aid in the characterization of singly-transiting planets discovered by the upcoming NASA TESS mission in Section 2.4.

2.2 Methods

2.2.1 How to measure ρ_* from a transit light curve

Seager & Mallén-Ornelas (2003) demonstrated that the mean stellar density ρ_* can be measured from a transit light curve without any direct measurement of the stellar mass M_* or radius R_* as a result of Kepler's third law. Figure 2.1 offers some intuition about this procedure in the case of a circular orbit, and we sketch the analytic derivation of the circular-orbit case here.

We begin with Kepler's third law:

$$\frac{P^2}{4\pi^2} = \frac{a^3}{G(M_* + M_p)} \simeq \frac{a^3}{GM_*}, \tag{2.1}$$

where the right-hand side assumes that $M_p \ll M_*$. Dividing and multiplying the right-hand side of this equation by the stellar volume, $\frac{4}{3}\pi R_*^3$, we obtain:

$$\frac{P^2}{4\pi^2} = \frac{3(a/R_*)^3}{4\pi G\rho_*} \tag{2.2}$$

Rearrangement yields:

$$\rho_* = \frac{3\pi(a/R_*)^3}{GP^2} \tag{2.3}$$

Therefore, to measure ρ_*, we need only know the orbital period P and normalized semi-major axis a/R_* of a planet orbiting the star. (In particular, neither M_* nor R_* is necessary to obtain ρ_*.) Both P and a/R_* are directly measurable from the transit light curve: P is the interval between successive transits, and a/R_* can be derived from the transit duration. In the case of a circular orbit, a/R_* follows trivially from the transit duration and P (see Figure 2.1):

$$T = \frac{2R_*}{(2\pi a/P)} \tag{2.4}$$

Rearrangement of this equation yields the normalized semimajor axis a/R_*:

$$\frac{a}{R_*} = \frac{P}{\pi T} \tag{2.5}$$

However, in general, the eccentricity e of the transiting planet's orbit also influences the transit duration T. The exact solution for T in the case of an eccentric orbit involves solving a quartic equation in $\cos f$, where f is the true anomaly (see Kipping 2008, 2010a for details). However, Kipping (2010a) found the following approximate expression for T under the simplifying assumption that the planet-star separation does not change during the

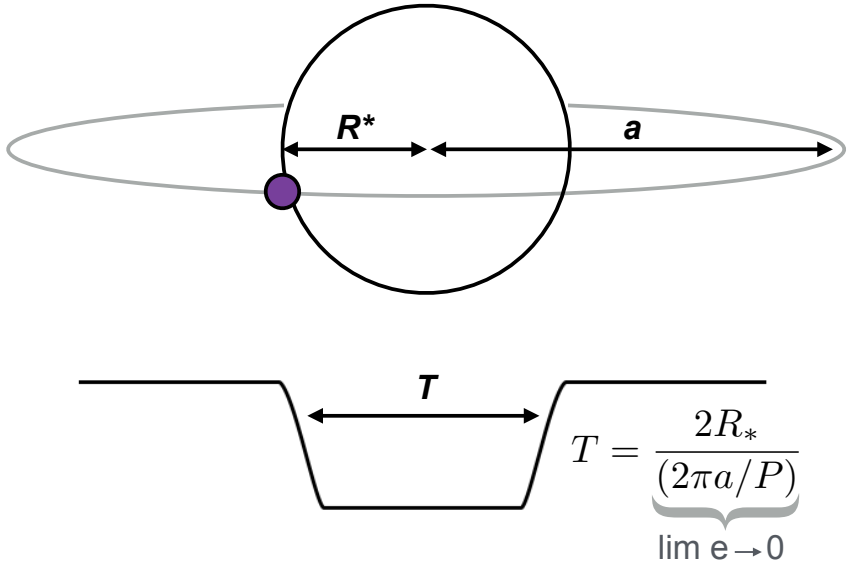

Figure 2.1: The transit duration T is equal to the stellar diameter divided by the mean orbital velocity, which is equal to $2\pi a/P$ in the case of a circular orbit. Rearrangement of the equation in the lower panel yields the normalized semimajor axis a/R_*. An analogous calculation is possible for planets on eccentric orbits (for which orbital velocity varies with phase), provided the eccentricity is known.

transit:

$$T \simeq \frac{P}{\pi}\frac{\varrho_c^2}{\sqrt{1-e^2}}\arcsin\left(\frac{\sqrt{1-(a/R_*)^2\varrho_c^2\cos^2 i}}{(a/R_*)\varrho_c \sin i}\right) \tag{2.6}$$

where ϱ_c is the separation between the planet and star at mid-transit, in units of stellar radii.

Since e and ρ_* both influence the transit duration T, it is necessary to have a precise constraint on the e in order to derive a precise constraint on ρ_* (Kipping, 2010a). For some planets, such as planets with observed secondary eclipses, e is directly measurable (e.g. Knutson et al. 2007a); for others, such as planets on very short-period orbits which are expected to tidally circularize quickly or planets in compact multi-planet systems, dynamical stability constrains e to low values. For each of these categories of planet—secondary-eclipse planets, tidally circularized planets, and multi-planet systems—we may express the existing eccentricity constraint as a Bayesian prior on e. In sections 2.2.2.1-2.2.2.3, we describe how

we select a sample of *Kepler* Objects of Interest (KOIs) belonging to each category for transit modeling.

We note that, in principle, ρ_* could also be measured from the transits of planets with radial velocity-measured eccentricities. However, analyzing such planets requires jointly fitting the radial velocity curves, including accurate treatment of stellar activity effects. This is beyond the scope of the present study, and we defer analysis of planets with radial velocity-measured e to later work.

Assuming, then, that we have a strong e prior, all we must do to measure ρ_* from a transit is fit a transit model, comprising ten parameters: the transit epoch t_0, the orbital period P, the impact parameter b, the stellar density ρ_*, the ratio-of-radii R_p/R_*, the orbital eccentricity e, the argument of periastron ω, and three coefficients of a modified nonlinear limb darkening law (transformed to allow for efficient sampling as described in Kipping 2016), α_r, α_h, and α_θ. In other words, we must explore this ten-dimensional parameter space and find a region that matches the *Kepler* transit data.

We use the transit-modeling code `BATMAN` (Kreidberg, 2015) to compute the light curve of a given set of ten transit model parameters, compare this model to the *Kepler* data, and evaluate the likelihood of the parameters. We step through the ten-dimensional parameter space and derive posterior distributions for the model parameters with `emcee` (Foreman-Mackey et al., 2013), an affine-invariant ensemble Markov chain Monte Carlo (MCMC) sampler. Details of this procedure are given in section 2.2.4.

2.2.2 Sample selection

Here, we describe how we select a sample of KOIs with strong eccentricity priors for transit modeling. We furthermore select the host stars of these KOIs to span a broad range in *Kepler*-band magnitude and effective temperature, as shown in Figure 2.2, in order to investigate the efficacy of this method across a wide range of stellar types.

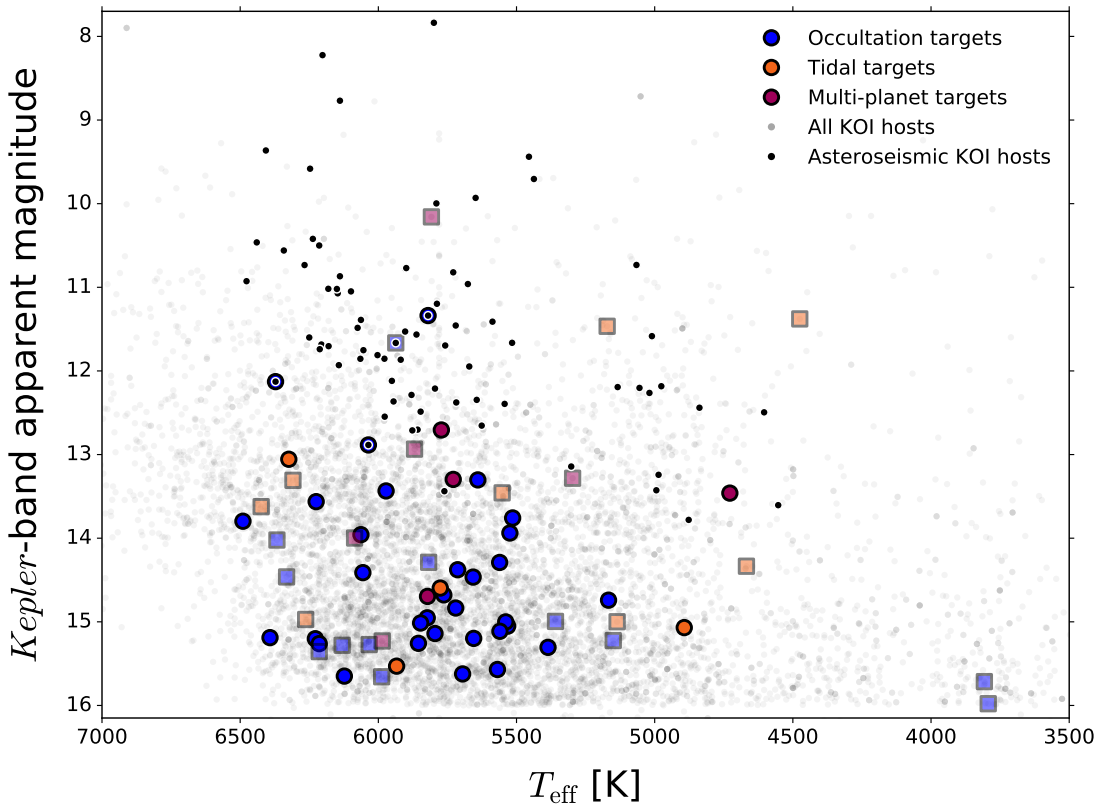

Figure 2.2: The distribution of our target stars, compared to all KOI-hosting stars and KOI-hosting stars with asteroseismic density measurements, in $T_{\rm eff}$-magnitude space. Opaque circles represent stars for which we achieve comparable ρ_* precision to asteroseismology (fractional uncertainty $\leq 5\%$); transparent squares represent stars for which we do not. Four of our targets overlap with the Huber et al. (2013) asteroseismic sample; see Figure 2.7 for details.

2.2.2.1 Secondary eclipse targets

Certain exceptional transiting planets are bright enough relative to their host star that the flux from the planet-star system drops perceptibly when the planet passes behind the star. Such planets are thus detected both when they pass in front of their host stars (transit) and when they pass behind (occultation, or secondary eclipse). Clocking the planet at two points in its orbit, rather than just at transit, makes it possible to precisely constrain the

eccentricity and argument of periastron (ω) of the orbit (e.g., Winn 2010); in other words, it places a strong prior on e, which allows us to measure ρ_* from the planet's transit.

More specifically, we may derive constraints on $e\cos\omega$ and $e\sin\omega$ by measuring the time elapsed between mid-transit and the subsequent mid-occultation (Δt) and the relative duration of the transit compared to the occultation, $T_{\rm transit}/T_{\rm occultation}$. Approximate expressions for these constraints may be found in Winn (2010):

$$e\cos\omega \simeq \frac{\pi}{4}\left(\frac{2\Delta t}{P} - 1\right) \tag{2.7}$$

$$e\sin\omega \simeq \frac{T_{\rm transit}}{T_{\rm occultation}} - 1 \tag{2.8}$$

We draw KOIs with observed secondary eclipses (hereafter, "occultation targets") from catalogs compiled by Coughlin & López-Morales (2012) and Shabram et al. (2016). Shabram et al. (2016) measured the eccentricities of 50 KOIs with detected secondary eclipses. Of these, five (KOI-774.01, KOI-805.01, KOI-895.01, KOI-1227.01, and KOI-1391.01) were subsequently identified as false positives in the NASA Exoplanet Archive[1] (Akeson 2015, hereafter "NEA"), leaving 45 targets with measured eccentricities. To this list, we add a further 10 KOIs detected in secondary eclipse by Coughlin & López-Morales (2012). Of these, four (KOIs 1.01, 5.01, 10.01, and 412.01) are not counted as significant SE detections by Coughlin & López-Morales 2012, so we advise caution in adopting our transit parameter posteriors for these targets.

We remove one of these targets (KOI-203.01) from the list due to stroboscopic starspot activity (Désert et al., 2011) and another four (KOI-202.01, KOI-760.01, KOI-883.01, and KOI-1781.01) due to detected transit timing variations (TTVs; Holczer et al. 2016). Modeling the transits of a planet with detected TTVs is prohibitively computationally expensive, because it requires adding a new model parameter to describe every successive interval

[1] https://exoplanetarchive.ipac.caltech.edu/, accessed 1 August 2017.

between transits (Teachey et al., 2018), and some of our target planets undergo hundreds of transits over *Kepler*'s 4-year observational baseline.

Finally, we remove 6 occultation targets because their MCMC analysis was prohibitively slow (see Section 2.2.4) as a result of their unusually high number of data points ($\sim 10^5 - 10^6$, compared to $\sim 10^3 - 10^5$ for successfully analyzed occultation targets). The resulting occultation target list, comprising 44 KOIs (the majority of our targets), is presented in Table 2.1.

2.2.2.2 Tidally circularized targets

Our second target population consists of KOIs with short tidal circularization timescales τ_{circ} ("tidal targets"). In general, we expect such KOIs to have approximately circular orbits (e close to 0); more precisely, Wang & Ford (2011) found that the e distribution for single-planet systems with short τ_{circ} is consistent with an exponential distribution, $P(e, \lambda) = \frac{1}{\lambda} \exp \frac{-e}{\lambda}$, with scale parameter $\lambda = 0.00796$. Similarly, Kipping (2013) found strong evidence that the short-period sample of RV-observe planets reside on less elliptical orbits than their longer-period counterparts, at a confidence of 11.6σ.

Table 2.1: Occultation targets, with measured transit parameters, stellar densities, and limb darkening coefficients. A machine-readable version of this table with more significant figures and the complete target list is available online.

KOI	t_0 [BKJD]	P [days]	b	$\log_{10}\rho_*$ [kg m^{-3}]	R_p/R_*	$\sqrt{e}\cos\omega$	$\sqrt{e}\sin\omega$	α_r	α_h	α_θ
1.01	$122.7625^{+0.0003}_{-0.0006}$	$2.47061338^{+1e-08}_{-1e-08}$	$0.79^{+0.01}_{-0.01}$	$3.3^{+0.07}_{-0.03}$	$0.129^{+0.001}_{-0.001}$	$-0.4^{+0.1}_{-0.1}$	$-0.21^{+0.05}_{-0.08}$	$0.01^{+0.07}_{-0.008}$	$0.78^{+0.05}_{-0.06}$	$0.46^{+0.08}_{-0.31}$
5.01	$132.972^{+0.001}_{-0.002}$	$4.780332^{+2e-06}_{-2e-06}$	$1.0^{+0.7}_{-0.3}$	$2.3^{+1.0}_{-0.3}$	$0.07^{+0.68}_{-0.05}$	$-0.1^{+0.2}_{-0.1}$	$-0.15^{+0.21}_{-0.07}$	$0.6^{+0.3}_{-0.4}$	$0.2^{+0.7}_{-0.2}$	$0.3^{+0.6}_{-0.3}$
10.01	$121.1195^{+0.0001}_{-0.0001}$	$3.5224985^{+1e-07}_{-1e-07}$	$0.72^{+0.01}_{-0.03}$	$2.68^{+0.05}_{-0.02}$	$0.0981^{+0.0003}_{-0.0004}$	$0.01^{+0.09}_{-0.09}$	$-0.1^{+0.1}_{-0.1}$	$0.1^{+0.01}_{-0.02}$	$0.9^{+0.08}_{-0.16}$	$0.39^{+0.05}_{-0.04}$
17.01	$121.48659^{+2e-05}_{-3e-05}$	$3.23469935^{+8e-08}_{-8e-08}$	$0.15^{+0.07}_{-0.03}$	$2.85^{+0.03}_{-0.09}$	$0.0937^{+0.0006}_{-0.0002}$	$0.0^{+0.1}_{-0.2}$	$0.1^{+0.1}_{-0.2}$	$0.56^{+0.05}_{-0.02}$	$0.8^{+0.1}_{-0.3}$	$0.594^{+0.028}_{-0.007}$
22.01	$177.25002^{+6e-05}_{-6e-05}$	$7.8914479^{+3e-07}_{-3e-07}$	$0.47^{+0.01}_{-0.02}$	$2.89^{+0.05}_{-0.03}$	$0.0967^{+0.0002}_{-0.0002}$	$0.0^{+0.1}_{-0.2}$	$-0.1^{+0.2}_{-0.1}$	$0.94^{+0.05}_{-0.11}$	$0.33^{+0.06}_{-0.03}$	$0.647^{+0.005}_{-0.01}$
97.01	$134.27692^{+0.00012}_{-9e-05}$	$4.8854886^{+2e-07}_{-2e-07}$	$0.53^{+0.02}_{-0.02}$	$2.42^{+0.08}_{-0.05}$	$0.083^{+0.0002}_{-0.0002}$	$0.01^{+0.08}_{-0.09}$	$-0.2^{+0.2}_{-0.1}$	$0.6^{+0.2}_{-0.2}$	$0.3^{+0.11}_{-0.06}$	$0.63^{+0.02}_{-0.04}$
98.01	$138.0881^{+0.0003}_{-0.0006}$	$6.7901219^{+7e-07}_{-5e-07}$	$0.61^{+0.04}_{-0.04}$	$2.21^{+0.1}_{-0.06}$	$0.0461^{+0.0001}_{-0.0001}$	$-0.0^{+0.2}_{-0.2}$	$-0.0^{+0.2}_{-0.2}$	$0.7^{+0.2}_{-0.3}$	$0.18^{+0.1}_{-0.04}$	$0.65^{+0.01}_{-0.06}$
127.01	$134.03056^{+5e-05}_{-7e-05}$	$3.5787806^{+1e-07}_{-1e-07}$	$0.42^{+0.07}_{-0.04}$	$2.97^{+0.09}_{-0.07}$	$0.0996^{+0.0007}_{-0.0004}$	$0.35^{+0.06}_{-0.07}$	$0.23^{+0.06}_{-0.11}$	$0.53^{+0.07}_{-0.02}$	$0.6^{+0.1}_{-0.1}$	$0.614^{+0.008}_{-0.01}$
1793.01	$131.7787^{+0.0005}_{-0.001}$	$3.26176^{+1e-06}_{-1e-06}$	$1.2^{+0.5}_{-0.9}$	$2.8^{+1.0}_{-0.1}$	$0.4^{+0.4}_{-0.3}$	$-0.1^{+0.1}_{-0.1}$	$-0.1^{+0.1}_{-0.1}$	$0.6^{+0.3}_{-0.3}$	$0.6^{+0.3}_{-0.4}$	$0.8^{+0.2}_{-0.4}$

To identify circularized KOIs, we adopt a theoretical upper limit for τ_{circ} from Haswell (2010), based on an upper limit for planet mass $M_{p,\ max} = 25M_J = 0.025M_\odot$ chosen to exceed the mass of any confirmed exoplanet in the exoplanets.org database:

$$\tau_{circ} \leq \frac{P}{21\pi}\left(\frac{K_{dP}}{Q_P}\right)^{-1}\frac{0.025M_\odot}{M_*}\left(\frac{a}{R_*}\right)^5\left(\frac{R_P}{R_*}\right)^{-5} \tag{2.9}$$

Here, P is the orbital period of the planet, M_* is the host star's mass, R_* is the stellar radius, a is the planet's semi-major axis, and R_p is the planet's radius. K_{dP} is the planet's dynamical Love number, a dimensionless parameter which expresses the ratio of the additional gravitational potential produced by tidal redistribution of the planet's mass to the gravitational potential before redistribution (Poulsen, 2009; Love, 1934). Q_P is the planet's tidal quality factor, another dimensionless parameter which quantifies the efficiency of tidal dissipation in the planet (Ogilvie & Lin, 2004).

Using Kepler's third law, we may express a/R_* in terms of M_* and R_*, which are more reliably reported in the *Kepler* catalog because they do not depend on transit modeling. This conversion yields

$$\tau_{circ} \leq \frac{P}{21\pi}\left(\frac{K_{dP}}{Q_P}\right)^{-1}\frac{0.025M_\odot}{M_*}\left(\frac{R_P}{R_*}\right)^{-5}\left(\frac{P^2GM_*}{4\pi^2R_*^3}\right)^{5/3} \tag{2.10}$$

We apply a linear interpolation to Solar System values to obtain the following equation for $\frac{K_{dP}}{Q_P}$ (Teachey et al., 2018):

$$\frac{K_{dP}}{Q_P} = 10^{-2.90-20.33\frac{R_P}{R_*}\frac{R_*}{R_\odot}} \tag{2.11}$$

To assemble our tidal target list, we select every KOI with τ_{circ} less than 10^8 years according to these equations. There are 19 such KOIs; the maximum orbital period of these is 1.6 days (KOI-809.01). Of these, we remove five from the target list: KOI-203.01, again due to its stroboscopic starspot activity (Désert et al., 2011); KOI-1546.01, for detected transit timing variations (Holczer et al., 2016); KOI-3156.01, an identified hierarchical quintuple star

system (Shibahashi & Kurtz, 2013; Rappaport et al., 2016); KOI-5804.01, a highly active star where NEA-identified "transits" correspond to alternating minima in the stellar light curve; and KOI-6534.01, which has no visible transits in its light curve at the NEA-determined transit epoch and period.

Finally, we remove 1 tidal target because because its MCMC analysis was prohibitively slow (see Section 2.2.4) as a result of its unusually high number of data points ($\sim 10^5$, compared to $\sim 10^3 - 10^4$ for successfully analyzed tidal targets). The remaining 13 tidal targets are listed in Table 2.2.

Table 2.2: Tidally circularized ("tidal") targets, with measured transit parameters, stellar densities, and limb darkening coefficients. A machine-readable version of this table with more significant figures and the complete target list is available online.

KOI	t_0 [BKJD]	P [days]	b	$\log_{10}\rho_*$ [kg m^{-3}]	R_p/R_*	$\sqrt{e}\cos\omega$	$\sqrt{e}\sin\omega$	α_r	α_h	α_θ
809.01	$170.64828^{+7e-05}_{-0.00186}$	$1.5947455^{+1e-07}_{-1e-07}$	$0.67^{+0.04}_{-0.02}$	$3.1^{+0.03}_{-0.02}$	$0.13^{+0.002}_{-0.002}$	$-0.02^{+0.07}_{-0.56}$	$-0.03^{+0.08}_{-0.15}$	$0.3^{+0.4}_{-0.2}$	$0.5^{+0.3}_{-0.2}$	$0.22^{+0.31}_{-0.08}$
1064.01	$133.49^{+0.21}_{-0.09}$	$1.187^{+0.005}_{-0.005}$	$0.8^{+0.3}_{-0.1}$	$1.97^{+0.55}_{-0.08}$	$0.5^{+0.2}_{-0.3}$	$0.0^{+0.3}_{-0.2}$	$-0.1^{+0.2}_{-0.2}$	$0.7^{+0.3}_{-0.2}$	$0.6^{+0.3}_{-0.3}$	$0.25^{+0.34}_{-0.09}$
1075.01	$133.2764^{+0.0004}_{-0.0007}$	$1.3437661^{+6e-08}_{-7e-08}$	$1.91^{+0.02}_{-0.04}$	$2.437^{+0.007}_{-0.011}$	$0.98^{+0.01}_{-0.04}$	$-0.01^{+0.03}_{-0.04}$	$-0.1^{+0.02}_{-0.02}$	$0.8^{+0.2}_{-0.4}$	$4e-05^{+0.00094}_{-4e-05}$	$0.07^{+0.86}_{-0.05}$
1658.01	$133.54889^{+0.00016}_{-6e-05}$	$1.5449288^{+1e-07}_{-3e-07}$	$0.495^{+0.053}_{-0.001}$	$3.673^{+0.002}_{-0.819}$	$0.0687^{+0.002}_{-0.0002}$	$-0.08^{+0.02}_{-0.04}$	$0.424^{+0.0007}_{-0.1442}$	$8e-05^{+0.35166}_{-8e-05}$	$0.513^{+0.003}_{-0.144}$	$0.57^{+0.03}_{-0.09}$
2925.01	$131.721^{+0.006}_{-0.003}$	$0.716532^{+2e-06}_{-3e-06}$	$1.0^{+0.4}_{-0.7}$	$3.7^{+0.7}_{-1.5}$	$0.017^{+0.396}_{-0.009}$	$0.01^{+0.05}_{-0.06}$	$-0.02^{+0.07}_{-0.07}$	$0.5^{+0.3}_{-0.3}$	$0.5^{+0.3}_{-0.3}$	$0.6^{+0.3}_{-0.4}$
3913.01	$131.621^{+0.0007}_{-0.0006}$	$0.5828955^{+3e-07}_{-3e-07}$	$1.1^{+0.6}_{-0.8}$	$2.8^{+1.1}_{-0.4}$	$0.1^{+0.6}_{-0.1}$	$0.0^{+0.06}_{-0.06}$	$0.0^{+0.06}_{-0.06}$	$0.6^{+0.3}_{-0.4}$	$0.6^{+0.3}_{-0.3}$	$0.7^{+0.2}_{-0.5}$
4351.01	$131.9546^{+0.0009}_{-0.0008}$	$0.6443225^{+2e-07}_{-2e-07}$	$1.4^{+0.4}_{-0.4}$	$2.8^{+0.3}_{-0.3}$	$0.5^{+0.4}_{-0.4}$	$0.05^{+0.11}_{-0.08}$	$-0.04^{+0.1}_{-0.09}$	$0.7^{+0.3}_{-0.5}$	$0.9^{+0.1}_{-0.3}$	$0.6^{+0.3}_{-0.3}$
7449.01	$132.62^{+0.03}_{-0.14}$	$1.3245^{+0.0029}_{-0.0006}$	$1.2^{+0.8}_{-0.3}$	$1.4^{+0.5}_{-0.2}$	$0.4^{+0.6}_{-0.3}$	$0.03^{+0.15}_{-0.08}$	$-0.0^{+0.3}_{-0.2}$	$0.3^{+0.6}_{-0.3}$	$0.71^{+0.09}_{-0.44}$	$0.6^{+0.1}_{-0.2}$

2.2.2.3 *Compact multi-planet systems*

Finally, we consider compact multi-planet systems, which are not expected to be dynamically stable unless their constituent planets are on low-eccentricity orbits. Van Eylen & Albrecht (2015) quantify this expectation by examining 28 *Kepler* multi-planet host stars with asteroseismic ρ_* measurements. They use the discrepancy between the asteroseismic and transit-derived ρ_* measurement to measure the eccentricity of each of the 74 KOIs in their sample. They find that the resulting eccentricities are well described by a Rayleigh distribution, $P(e, \sigma) = \frac{e}{\sigma^2} \exp \frac{-e^2}{2\sigma^2}$, with $\sigma = 0.049 \pm 0.013$.

We cannot properly use this Rayleigh distribution as a prior e distribution to measure ρ_* of the host stars in the Van Eylen & Albrecht (2015) sample itself, because the ρ_* information contained in those KOIs' transits was used to define the prior in the first place. Rather, we must identify an independent sample of KOIs which resembles the sample of Van Eylen & Albrecht (2015).

To assemble this sample, we compare the distribution of period ratios of the Van Eylen & Albrecht (2015) sample to that of the remaining *Kepler* multi-planet systems. For each multi-planet system, we calculate the ratio of the orbital period of each outer planet to its nearest inner neighbor. The distribution of period ratios in the Van Eylen & Albrecht (2015) sample serves as a reference distribution; we identify a sample of 1340 KOIs which is consistent with the Van Eylen & Albrecht (2015) sample by an Anderson-Darling test ($p = 0.4$; Anderson & Darling 1952). We further subject the two samples to a Kolmogorov-Smirnov test and find them consistent at the $p = 0.15$ level (Kolmogorov, 1933; Smirnov, 1948).

We impose a signal-to-noise cutoff on this KOI sample, discarding planetary systems where the *Kepler*-reported transit model SNR < 50 for one or more of the KOIs. After the cutoff, 27 systems remain; of these 27 systems, 12 exhibit TTVs (Holczer et al., 2016) and are removed from the target list. We remove a further 2 systems, comprising 5 KOIs, because because their MCMC analysis was prohibitively slow (see Section 2.2.4) as a result of their

unusually high number of data points ($\sim 10^5 - 10^6$, compared to $\sim 10^3 - 10^5$ for successfully analyzed multi-planet targets), and 4 further systems for having very few remaining posterior samples after we perform some quality checks (see Section 2.3 for details). The remaining 9 systems, comprising 18 KOIs, are listed in Table 2.3.

Table 2.3: Compact multi-planet systems ("multis") selected to be statistically consistent with the sample of Van Eylen & Albrecht (2015), with measured transit parameters, stellar densities, and limb darkening coefficients. A machine-readable version of this table with more significant figures and the complete target list is available online.

KOI	t_0 [BKJD]	P [days]	b	$\log_{10}\rho_*$ [kg m^{-3}]	R_p/R_*	$\sqrt{e}\cos\omega$	$\sqrt{e}\sin\omega$	α_r	α_h	α_θ
124.01	$137.136^{+0.004}_{-0.011}$	$12.69085^{+0.00025}_{-5e-05}$	$0.68^{+0.31}_{-0.05}$	$2.8^{+0.3}_{-1.5}$	$0.0154^{+0.0066}_{-0.0005}$	$0.25^{+0.04}_{-0.01}$	$0.1^{+0.06}_{-0.03}$	$0.87^{+0.02}_{-0.02}$	$0.58^{+0.04}_{-0.03}$	$0.3^{+0.01}_{-0.17}$
124.02	$142.833^{+0.008}_{-0.004}$	$31.71962^{+0.00025}_{-5e-05}$	$0.83^{+0.09}_{-0.03}$		$0.019^{+0.0886}_{-0.0004}$	$0.14^{+0.1}_{-0.06}$	$-0.05^{+0.08}_{-0.29}$			
153.01	$139.7138^{+0.0005}_{-0.0011}$	$8.925083^{+2e-06}_{-9e-06}$	$0.98^{+0.04}_{-0.01}$	$1.95^{+0.05}_{-0.02}$	$0.11^{+0.02}_{-0.03}$	$0.03^{+0.02}_{-0.02}$	$-0.246^{+0.009}_{-0.009}$	$0.65^{+0.06}_{-0.03}$	$0.51^{+0.09}_{-0.04}$	$0.991^{+0.007}_{-0.011}$
153.02	$128.72^{+0.03}_{-0.03}$	$4.754008^{+2e-06}_{-9e-06}$	$1.06^{+0.04}_{-0.06}$		$0.08^{+0.04}_{-0.05}$	$-0.16^{+0.04}_{-0.03}$	$-0.513^{+0.006}_{-0.009}$			
678.01	$172.61^{+0.3}_{-0.01}$	$6.0634^{+0.0001}_{-0.0201}$	$0.02^{+0.84}_{-0.02}$	$0.7^{+1.9}_{-0.2}$	$0.023^{+0.024}_{-0.001}$	$0.14^{+0.02}_{-0.01}$	$0.4^{+0.2}_{-0.3}$	$0.18^{+0.13}_{-0.02}$	$0.76^{+0.01}_{-0.2}$	$0.64^{+0.08}_{-0.44}$
678.02	$132.04^{+0.41}_{-0.08}$	$4.1615^{+0.0001}_{-0.0201}$	$1.0^{+0.01}_{-0.23}$		$0.02^{+0.03}_{-0.02}$	$-0.37^{+0.03}_{-0.04}$	$-0.24^{+0.02}_{-0.02}$			
708.01	$171.12^{+0.01}_{-0.13}$	$17.404^{+0.003}_{-0.015}$	$0.88^{+0.08}_{-0.15}$	$3.9^{+0.3}_{-0.2}$	$0.02^{+0.06}_{-0.02}$	$0.01^{+0.02}_{-0.11}$	$0.0^{+0.4}_{-0.2}$	$0.68^{+0.06}_{-0.05}$	$0.45^{+0.05}_{-0.13}$	$0.5^{+0.3}_{-0.1}$
708.02	$176.315^{+0.15}_{-0.008}$	$7.691^{+0.003}_{-0.015}$	$0.79^{+0.21}_{-0.03}$		$0.003^{+0.057}_{-0.003}$	$-0.4^{+0.5}_{-0.1}$	$-0.24^{+0.09}_{-0.09}$			
1089.01	$175.5974^{+0.0004}_{-0.0004}$	$86.67856^{+1e-05}_{-2e-05}$	$0.53^{+0.08}_{-0.09}$	$2.63^{+0.08}_{-0.1}$	$0.0876^{+0.0011}_{-0.0007}$	$0.3^{+0.1}_{-0.6}$	$0.31^{+0.08}_{-0.13}$	$0.4^{+0.5}_{-0.3}$	$0.3^{+0.1}_{-0.2}$	$0.66^{+0.14}_{-0.04}$
1089.02	$140.323^{+0.003}_{-0.002}$	$12.21829^{+1e-05}_{-2e-05}$	$1.1^{+0.2}_{-0.4}$		$0.6^{+0.3}_{-0.5}$	$0.2^{+0.8}_{-0.1}$	$-0.5^{+0.3}_{-0.1}$			
1359.01	$370.48^{+0.03}_{-0.02}$	$37.1018^{+0.0011}_{-0.0009}$	$1.1^{+0.3}_{-0.7}$	$3.5^{+2.2}_{-0.8}$	$0.4^{+0.6}_{-0.4}$	$0.2^{+0.3}_{-0.2}$	$-0.6^{+0.5}_{-0.4}$	$0.5^{+0.4}_{-0.4}$	$0.4^{+0.3}_{-0.4}$	$0.3^{+0.2}_{-0.2}$
1359.02	$374.79^{+0.02}_{-0.05}$	$104.8213^{+0.0011}_{-0.0009}$	$1.0^{+0.2}_{-0.5}$		$0.09^{+0.24}_{-0.02}$	$0.3^{+0.5}_{-0.3}$	$-0.1^{+0.3}_{-0.2}$			
1779.01	$134.17798^{+5e-05}_{-5.2e-04}$	$4.685735^{+3e-06}_{-3e-06}$	0.3138^{+4e-04}_{-7e-04}	$5.0355^{+7e-04}_{-3.6e-03}$	$0.06695^{+7e-05}_{-1.3e-04}$	0.6894^{+4e-04}_{-9e-04}	$0.00706^{+9e-05}_{-7.1e-04}$	0.2489^{+1e-03}_{-3e-04}	0.9863^{+2e-04}_{-7e-04}	$0.9629^{+2e-04}_{-1.3e-03}$
1779.02	$133.36^{+0.0003}_{-0.0001}$	$11.838017^{+3e-06}_{-3e-06}$	$0.2759^{+0.0016}_{-0.0003}$		$0.0723^{+0.0002}_{-0.0003}$	$0.8987^{+0.0006}_{-0.0007}$	$0.3516^{+0.0004}_{-0.0006}$			
1809.01	$144.523^{+0.001}_{-0.0006}$	$13.0939^{+8e-06}_{-7e-06}$	$0.97^{+0.04}_{-0.04}$	$2.14^{+0.05}_{-0.04}$	$0.1^{+0.04}_{-0.04}$	$-0.27^{+0.02}_{-0.02}$	$-0.26^{+0.01}_{-0.02}$	$0.57^{+0.08}_{-0.06}$	$0.76^{+0.03}_{-0.04}$	$0.9^{+0.05}_{-0.04}$
1809.02	$132.695^{+0.002}_{-0.001}$	$4.915382^{+8e-06}_{-7e-06}$	$0.84^{+0.03}_{-0.02}$		$0.023^{+0.0007}_{-0.0016}$	$-0.04^{+0.03}_{-0.06}$	$-0.33^{+0.08}_{-0.04}$			
2687.01	$131.96^{+0.07}_{-0.22}$	$1.717^{+0.016}_{-0.003}$	$1.0^{+0.5}_{-0.1}$	$2.9^{+1.2}_{-2.0}$	$0.1^{+0.4}_{-0.1}$	$0.0^{+0.1}_{-0.1}$	$-0.34^{+0.18}_{-0.08}$	$0.42^{+0.3}_{-0.08}$	$0.6^{+0.1}_{-0.3}$	$0.7^{+0.2}_{-0.4}$
2687.02	$136.2^{+0.4}_{-0.4}$	$8.167^{+0.016}_{-0.003}$	$1.0^{+0.4}_{-0.3}$		$0.3^{+0.2}_{-0.3}$	$-0.2^{+0.3}_{-0.1}$	$-0.2^{+0.4}_{-0.4}$			

2.2.3 Detrending

Here, we describe our procedure for detrending the *Kepler* light curves of our 75 target KOIs (orbiting 66 target stars) in preparation for transit modeling. The trends in question are due to stellar activity or instrumental effects and are superimposed on the planetary transits in the light curve.

2.2.3.1 Outlier removal

We begin by splitting each target KOI's full *Kepler* simple aperture photometry light curve into individual transits, each bookended by sufficient out-of-transit observation time to capture out-of-transit trends in the light curve. For targets with available short-cadence observations (58.86 seconds per exposure), we apply the below procedure to both short- and long-cadence data; otherwise we use long-cadence data (29.4 minutes per exposure).

To slice the light curve, we use the NEA-reported transit ephemeris t_0, orbital period P, and transit duration T_{14} (Akeson, 2015). We divide each light curve into segments centered at t_0 plus successive integer values of P. For each segment, we keep out-of-transit data spanning an interval $t_{1/2} + t_{OOT}$ on either side of t_0, where we define $t_{1/2}$ as slightly more than half a transit duration, and t_{OOT} as an "out-of-transit window:"

$$t_{1/2} = 1.1 \left(\frac{T_{14}}{2} + t_{LC} \right) \tag{2.12}$$

$$t_{OOT} = \sqrt{10}\, t_{1/2}, \tag{2.13}$$

or roughly $3\, t_{1/2}$. Here, t_{LC} is the integration time of a long-cadence *Kepler* exposure, equal to 29.4 minutes. We discard data points where $|t - t_0| > (t_{1/2} + t_{OOT})$.

Once each KOI's light curve is divided into individual transit segments, we remove flux outliers and discard transit segments with insufficient data. To remove outlying data points

within each transit segment, we perform a moving median smoothing of the out-of-transit (i.e. $|t - t_0| > t_{1/2}$) flux data points, with a kernel size of 21 data points. We then reject any data points more than 3σ away from the moving median-smoothed light curve, where σ is defined as the *Kepler*-reported uncertainty of each flux measurement. A small number of transit segments also exhibit clear outliers within $t_{1/2}$ of a segment midpoint, identifiable as data points with anomalously high flux. We remove any data point that lies more than 3σ *above* the within-transit light curve.

Finally, after removing individual outlying data points, we reject any full transit segment where one of the following conditions is met:

1. There are fewer than 3 out-of-transit data points on one side of t_0;

2. There are more than 3 out-of-transit data points, but they span a very short time interval (i.e., less than $2\,t_{LC}$); or

3. The out-of-transit data points immediately adjacent to the transit are missing. Such missing data could lead to poor constraints on the transit depth or duration.

After outlier removal, each target KOI's light curve is reduced to a series of individual transit segment light curves. Each transit observed at *Kepler's* long cadence contains ~ 30 data points, and each transit observed in short cadence contains ~ 750 data points.

2.2.3.2 Evaluation of out-of-transit trends

The transit segments isolated by the above procedure are individually afflicted by out-of-transit trends due to stellar activity and instrumental variation. To fit a precise transit model to each KOI, we must first account for these trends (e.g. Aigrain et al. 2016; Luger et al. 2016). One approach to detrending would be to add additional free parameters to our transit model to describe each transit's trends individually; fit them all; and then marginalize over them to recover the physically interesting parameters describing the planet-star

system. However, this would add prohibitive computational cost, all for the sake of nuisance parameters.

Instead, we elect to detrend each transit segment using linear least squares regression (see e.g. Kundurthy et al. 2011). We assume that the out-of-transit trend for each transit segment is well-fit by a low-order polynomial of predetermined order, then divide out the best-fitting polynomial trend at each MCMC step before calculating the likelihood of the transit model parameters. Polynomial detrending is a common approach to analyzing *Kepler* data (see e.g. O'Leary & Burkart 2014; Fabrycky et al. 2012; Orosz et al. 2012; Lissauer et al. 2011b).

To choose the appropriate polynomial order for each transit segment, we use the Bayesian Information Criterion (BIC), a model selection statistic which balances goodness-of-fit against the number of free parameters in the model, i.e. the polynomial order:

$$BIC = \chi^2 + k \ln n \tag{2.14}$$

Here, k is the number of free parameters in the model, n is the number of data points, and χ^2 is the squared error of the model, scaled by the measurement uncertainties.

For each transit segment, with the in-transit data masked, we test polynomials of orders ranging from 0 to 3 and select the polynomial model with the lowest BIC. At each MCMC step (see section 2.2.4, below), before evaluating the likelihood of the transit model calculated from the sampled parameters, we (i) calculate, analytically, the best-fitting polynomial of this pre-selected order for each transit segment using linear least squares regression and (ii) impose this best-fitting polynomial trend upon the transit model. We are then evaluating the likelihood of the transit model given the data, both subject to the same out-of-transit trends.

2.2.4 Transit modeling

With this polynomial-fitting procedure in place, we explore the parameter space of our transit model to identify the region that describes each planet's transit light curve best. For the occultation and tidal targets, this space is ten-dimensional. For the multi-planet targets, it is $(4 + 6N)$-dimensional, where N is the number of planets in the system; 4 parameters describe the star and are the same for every KOI in the system (ρ_* and the three LDCs, α_r, α_h, and α_θ), and 6 describe each KOI (the transit epoch t_0, the period P, the impact parameter b, the ratio-of-radii R_p/R_*, and the reparametrized eccentricity and argument of periastron, $\sqrt{e}\cos\omega$ and $\sqrt{e}\sin\omega$).

We evaluate the likelihood of any given set of $4 + 6N$ transit parameters by using the transit modeling package `BATMAN` (Kreidberg, 2015) to calculate a light curve directly from the parameters. To this calculated light curve, we calculate and apply the best-fitting out-of-transit polynomial trend of pre-determined order (see Section 2.2.3) for each observed transit of the target KOI to enable a direct comparison of the model to the data. We then calculate the likelihood of the data given the transit model parameters. We adopt a Gaussian likelihood function.

We explore the $4+6N$-dimensional parameter space of the transit model with the affine-invariant MCMC ensemble sampler package `emcee`. `emcee` initializes an ensemble of MCMC walkers in this parameter space and calculates the posterior probability of the sampled set of transit parameters at every step in their random walk, given a choice of prior distributions and our Gaussian likelihood function.

We adopt the following priors for the transit parameters:

1. Intrinsic priors:

   (a) t_0: A uniform prior from $t_{0,\,reported}-0.5$ days to $t_{0,\,reported}+0.5$ days , where $t_{0,\,reported}$ is the transit epoch reported in the *Kepler* catalog.

(b) P: A uniform prior from $0.9\,P_{reported}$ to $1.1\,P_{reported}$, where $P_{reported}$ is the orbital period reported in the *Kepler* catalog.

(c) b: A uniform prior from 0 to 2, allowing for grazing transits.

(d) ρ_*, reparametrized as $\log_{10}(\rho_*[\mathrm{kg/m^3}])$: A uniform prior in $\log_{10}(\rho_*[\mathrm{kg/m^3}])$ from 0 to 6.

(e) R_p/R_*: A uniform prior from 0 to 1.

(f) e and ω: Uniform priors from -1 to 1 in $\sqrt{e}\cos\omega$ and $\sqrt{e}\sin\omega$, with additional uniform priors restricting e to the range $(0, 1)$ and ω to the range $(-\pi, \pi)$.

(g) Nonlinear limb-darkening coefficients $\alpha_r, \alpha_h, \alpha_\theta$: Uniform priors from 0 to 1 (Kipping, 2016).

(h) A prior insisting that b be less than (a/R_*), calculated from Kepler's third law, in order to prevent unphysical inclinations.

(i) A prior insisting that b be less than $(1 + R_p/R_*)$, in order to prevent unphysical transit durations.

2. Target selection-motivated e and ω priors:

(a) For the occultation targets, which have secondary eclipse-measured constraints on e and ω, we adopt Gaussian priors in $e\cos\omega$ and $e\sin\omega$, where the means are given by the measured values of $e\cos\omega$ and $e\sin\omega$ from Shabram et al. (2016) or Coughlin & López-Morales (2012) and the standard deviations by their measurement uncertainties.

(b) For the tidal targets, we adopt an exponential prior on e, with scale parameter $\lambda = 0.00796$, consistent with the findings of Wang & Ford (2011).

(c) For the multi-planet targets, we adopt a Rayleigh prior in e, with scale parameter $\sigma = 0.049$, consistent with the findings of Van Eylen & Albrecht (2015).

With `emcee`, we initialize 100 MCMC walkers per KOI and run them for 10^5 steps each, generating 10^7 posterior samples per KOI. We initialize the walkers in the P-dimension by drawing from a Gaussian distribution centered at the *Kepler* catalog-reported P, with standard deviation 0.01. We initialize the walkers in the other 9 dimensions of parameter space by sampling randomly in a 9-dimensional box spanning the range in each parameter that is allowed by its intrinsic prior.

We discard the first 20,000 steps per walker chain as "burn-in," based on a conservative by-eye judgment of when the walkers "forget" their initial conditions and begin to explore the parameter space freely. We also discard walker chains which fail to converge to the same value of P as the majority of the ensemble of walkers. More specifically, we calculate the median and median absolute deviation (MAD) of P over all the walker chains. We use $\hat{\sigma} = 1.4826 \times \mathrm{MAD}$ as an estimator for the standard deviation of the P distribution and discard any chain whose median P differs from the overall median by more than $5\hat{\sigma}$ (Rousseeuw & Croux, 1993).

Finally, for KOIs which are confirmed per their NEA disposition (two-thirds of our targets, or 50 KOIs), we discard all posterior samples with $R_p/R_* > 0.15$ on the grounds that they represent unphysically large planets. We note that all of our target planets which are dispositioned as "confirmed" in the NEA are validated by Morton et al. (2016).

2.3 Results

We obtain successful transit fits (i.e., MCMC convergence) for 66 target stars (hosting 75 individual KOIs). For four multi-planet targets (those hosting KOIs 156.01, 156.02, and 156.03; 723.01, 723.02, and 723.03; KOIs 1805.01, 1805.02, and 1805.03; and KOIs 1824.01 and 1824.02), less than a few hundred posterior samples for each system remain after we discard chains that fail to converge in P and samples with unphysically large R_p/R_*; we count these as failed fits. The best-fit transit parameters for the 66 successes are listed in

Tables 2.1-2.3; we present the median of the posterior distribution for each parameter, with uncertainty bounds describing the 16th and 84th percentiles.

Of order $10^6 - 10^7$ samples from the posterior distributions of the transit parameters remain for each KOI after we discard the burn-in phase of the MCMC chain, as well as chains which fail to converge in P and samples with unphysically large R_p/R_*. The files containing all of the posterior samples are prohibitively large to be made available for online download, so we downsample the posteriors by a factor of 10^2 and publish the resulting $10^4 - 10^5$ posterior samples for each KOI at https://doi.org/10.5281/zenodo.1028515.

As an example, in Figure 2.3, we present a well-converged transit fit, for occultation target KOI-929.01 (a confirmed planet, per the NEA). We plot the corresponding posterior distributions for the 10 fitted transit parameters in Figure 2.4. From each posterior sample, we calculate nine other parameters describing the system (the transit duration T_{14}, the flat-bottomed transit duration T_{23}, the normalized semi-major axis a/R_*, the inclination i, the eccentricity e, the argument of periastron ω, and the three traditional nonlinear limb-darkening coefficients c_2, c_3, and c_4). We plot the distributions of these derived parameters in Figure 2.5.

KOI-929.01's posterior distributions typify our broader results: t_0 and P are by far the most precisely constrained parameters, ρ_* is constrained to within 5% of its median value, and $\sqrt{e}\cos\omega$ and $\sqrt{e}\sin\omega$ are centered at zero, in agreement with the prior constraints on this planet's orbit from secondary eclipse observations. When we derive the distributions of e and ω themselves (see Figure 2.5), we find that e is strongly peaked at $e = 0$, and ω is very poorly constrained, which is sensible for a nearly-circular orbit. Also typical are the constraints on the three limb-darkening coefficients: α_θ is constrained to within $\sim 15\%$ of its median value, and α_r and α_h only to within $\sim 60\%$.

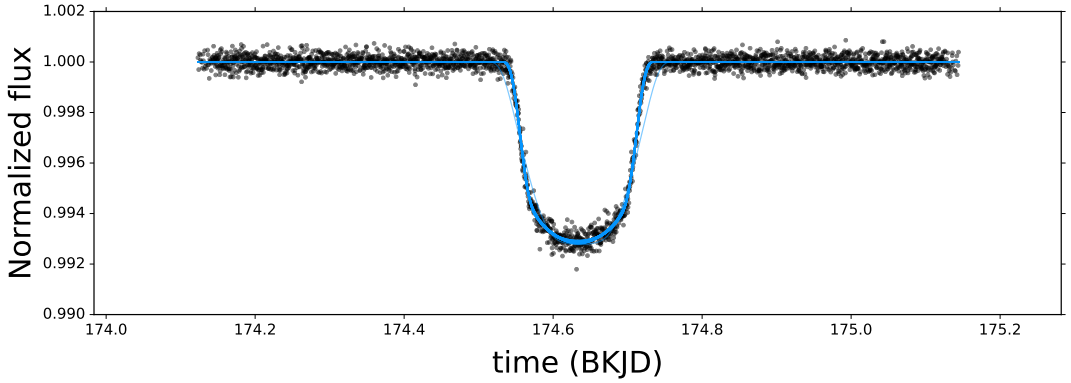

Figure 2.3: An example transit fit, for occultation target KOI-929.01, a confirmed planet per its NEA disposition. The black points are the 183 transits observed for this KOI, detrended and stacked; it is not observed in short cadence, so all of these data points are long-cadence observations. The blue lines are light curve models computed by `BATMAN` (Kreidberg, 2015) from 500 random draws from our 10-dimensional transit parameter posterior distributions. This KOI has an orbital period of 6.491683 ± 0.000002 days, and it orbits a star of *Kepler*-band magnitude 15.649.

2.3.1 Covariances

In this section, we investigate covariances between the transit parameters, which indicate degeneracies in the transit model. In other words, if independently adjusting two or more of the parameters can create the same effect in the shape of the model light curve, these parameters will correlate with each other, or co-vary.

A well-known effect in transit modeling (see e.g. Carter et al. 2008) is the covariance between stellar density ρ_*, impact parameter b, and ratio-of-radii R_p/R_*, which results from the mixed influence of these three parameters on the transit duration. For example, a larger R_p/R_*, a smaller b, and a lower ρ_* all lead to a longer transit duration. This covariance manifests itself in the posterior distributions of several of our less-well-constrained targets, especially those for which no short-cadence observations are available and those which are not confirmed per the NEA (for which we cannot discard posterior samples with $R_p/R_* > 0.15$). For such targets, our posterior plots show an elongated positive correlation between the b

and R_p/R_* distributions, as well as a tail of low b values which are negatively correlated with $\log_{10}\rho_*$. Figure 2.4, although it is a confirmed planet per the NEA, exhibits these trends.

The ultimate overall effect of this covariance is a distinct bimodality in each of the b, $\log_{10}\rho_*$, and R_p/R_* posterior distributions. For physical intuition, this bimodality signifies that two transit models are likely given the observations: one in which a relatively small planet undergoes a non-grazing transit across a compact star, and one in which a relatively large planet undergoes a grazing transit across a large, low-density star.

A-priori, the high-b, low-$\log_{10}\rho_*$, high-R_p/R_* peak is physically implausible, on the grounds that we are much more likely to observe a small planet transiting across the mid-point of its host star than we are to observe an enormous planet transiting across the limb (Kipping & Sandford, 2016). For independently confirmed KOIs, we exclude all posterior samples with $R_p/R_* > 0.15$ on these physical grounds, because anything larger than this approximate limit would be an eclipsing binary, not a transiting planet. We cannot, however, exclude the large R_p/R_* samples for KOIs which are not formally confirmed per the NEA, in the case that they turn out to be eclipsing binaries.

A strong intrinsic covariant prior on R_p/R_* and $\log_{10}\rho_*$, i.e., a way to formally encode our skepticism of grazing, large-planet fits, would address this problem of bimodality, as discussed in Kipping & Sandford 2016. However, the exoplanet population data are not yet robust enough to define such a prior.

Also evident in the posterior distributions of the transit parameters is a covariance between the limb darkening coefficients α_r and α_h. For the vast majority of our target stars, these parameters are not tightly constrained—the posteriors displayed in Figure 2.4 are typical. Although the peaks of the distributions of α_r and α_h are broad, however, there is a clear negative correlation between the two, with high α_r corresponding to low α_h and vice versa. This covariance explains the trends we discuss in Section 2.3.2.1, where we compare our observed αs to theoretical predictions from stellar atmosphere modeling.

Finally, we note a strong covariance between c_2, c_3, and c_4, evident in the rightmost panels

of Figure 2.5. This covariance, which motivated the transformation to α-space originally (Kipping, 2016), exists because only a relatively small region of the three-dimensional c-space describes physically realistic limb-darkening behavior.

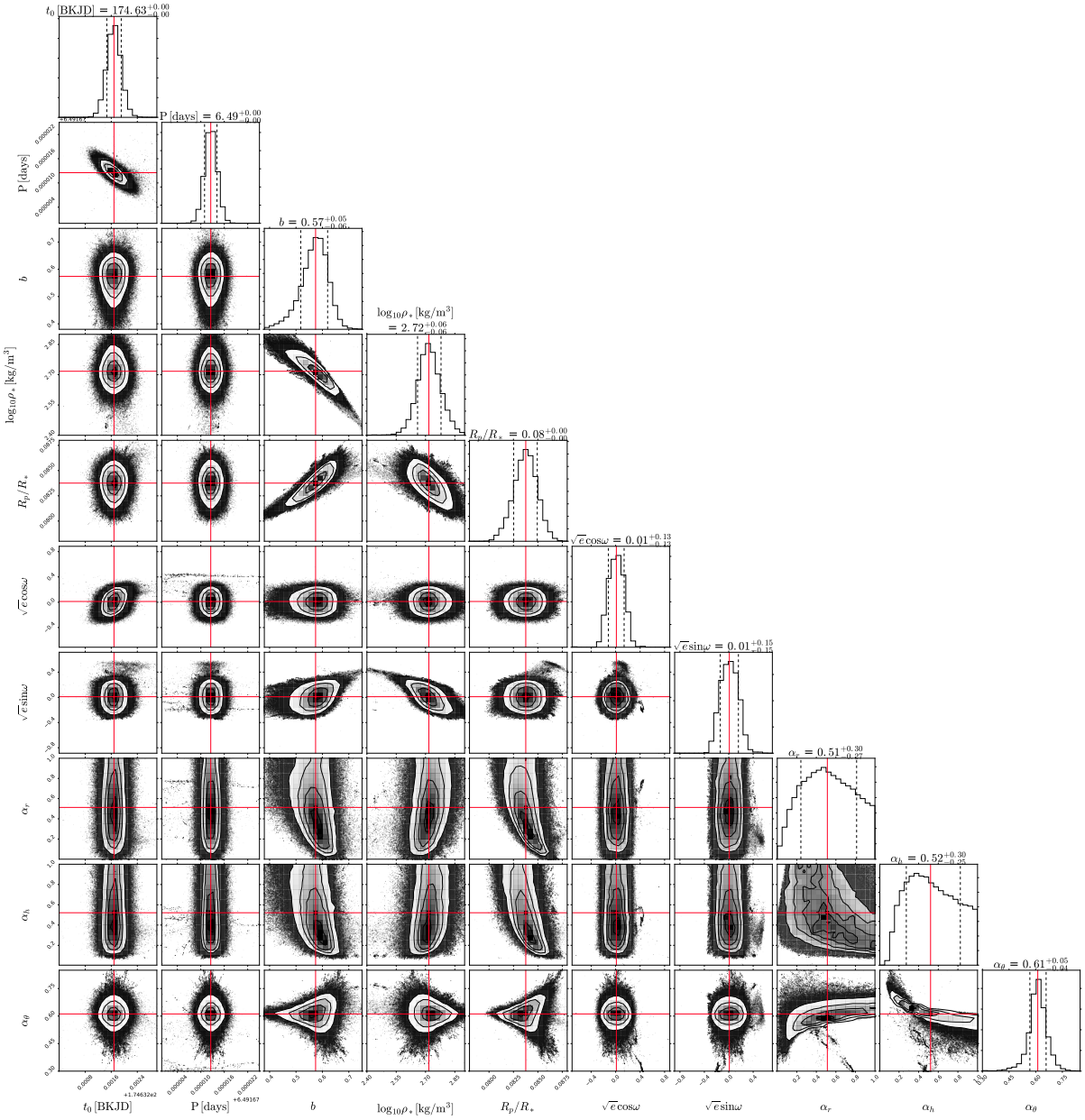

Figure 2.4: Posterior distributions for the ten fitted transit parameters of occultation target KOI-929.01, a confirmed planet per the NEA. The red lines mark the median of each distribution; the black dotted lines mark the 16th and 84th percentiles.

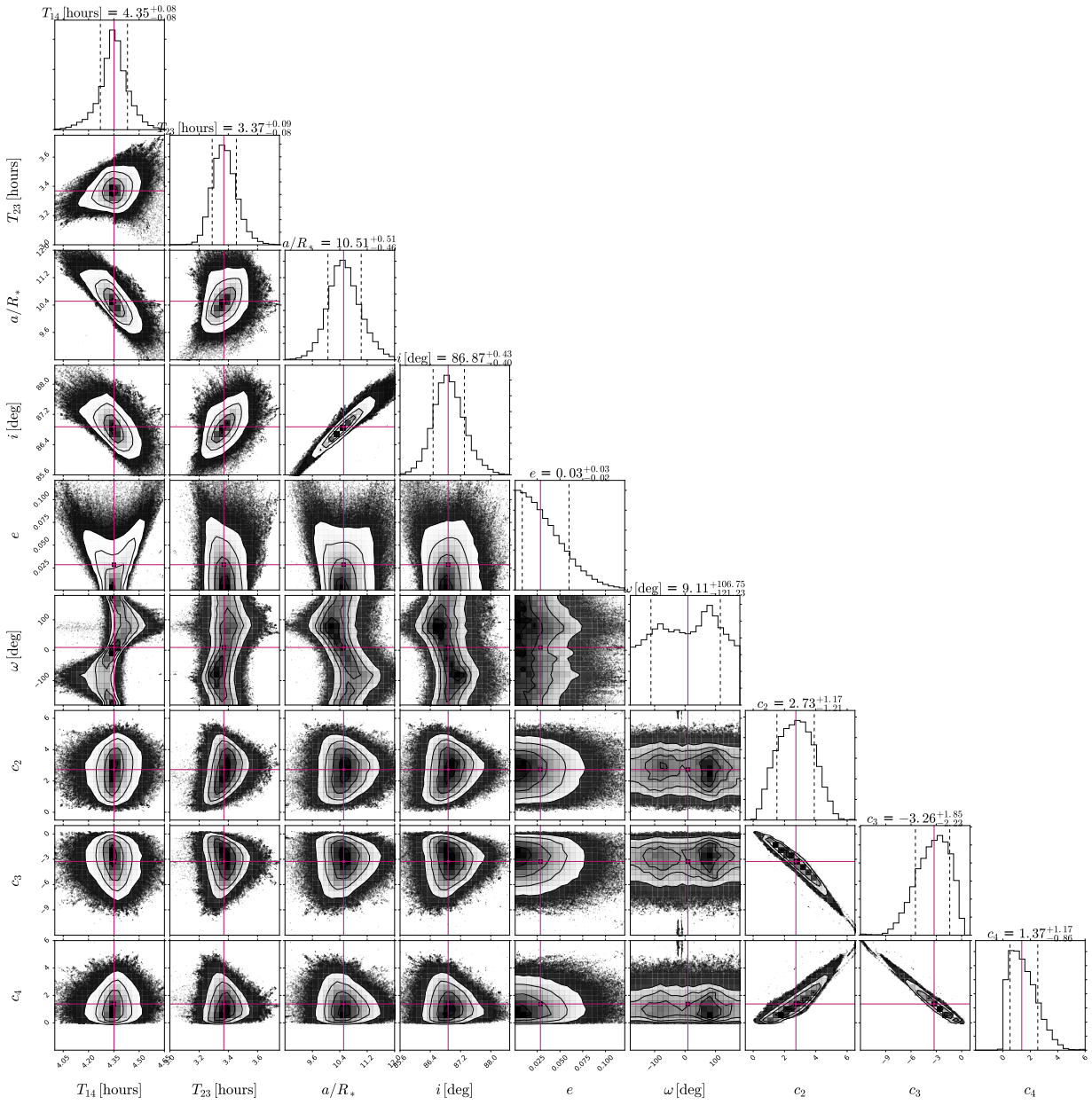

Figure 2.5: Distributions for nine derived parameters of occultation target KOI-929.01, a confirmed planet per the NEA. These are parameters which we did not fit for directly but can compute from the posterior samples plotted in Figure 2.4. Here, c_2, c_3, and c_4 are the traditional coefficients of a modified nonlinear limb-darkening law, computed from our reparametrized αs. The purple lines mark the median of each distribution; the black dotted lines mark the 16th and 84th percentiles.

2.3.2 Stellar densities

The ensemble results of our stellar density measurements are presented in Figure 2.6. In this figure, for each target star, we compare the posterior distributions of $\log_{10} \rho_*$ derived from our transit fitting with the constraints on $\log_{10} \rho_*$ from the *Kepler* Data Release 25 (DR25) Stellar Properties Catalog by Mathur et al. (2017). The posteriors from the Mathur et al. (2017) catalog are derived by performing Dartmouth Stellar Evolution Database isochrone modeling on input values of $T_{\rm eff}$, $\log g$, and [Fe/H] obtained from earlier studies relying on a variety of experimental methods, including spectroscopy, flicker, asteroseismology, and previous transit modeling.

Our transit modeling-derived $\log_{10} \rho_*$ is in 1σ or better agreement with the isochrone modeling-derived DR25 constraint for 55% of our target stars, and in 3σ agreement for 95%. The only three target stars for which we disagree with the DR25 stellar density constraint at the 3σ level are tidal targets KOI-5157.01 and 7430.01 and the multi-planet system consisting of KOIs 153.01 and 153.02. We note that all four of these KOIs have unusually poorly constrained transit epochs, the parameter that is generally best constrained by our modeling. Correspondingly, we recommend against adopting our modeled transit parameters for these KOIs and their host stars.

Our transit modeling-derived $\log_{10} \rho_*$ is more precise than the isochrone modeling-derived DR25 constraint for 50% of our targets. The median improvement to fractional uncertainty for these 50% is a factor of 2.3, meaning that our fractional uncertainty is less than half that of the literature value for a typical target star. We achieve comparable precision to asteroseismology (i.e., fractional uncertainty in $\rho_* \leq 5\%$) for 62% of our targets. For some others, e.g. KOI-1793.01, KOI-4351.01, and KOI-3913.01, the bimodality discussed in section 2.3.1 is apparent, and we derive a poor constraint on $\log_{10} \rho_*$.

We are able to extend sub-5%-fractional-uncertainty stellar density measurements to *Kepler* stars three magnitudes fainter than asteroseismology can, across a broad range in

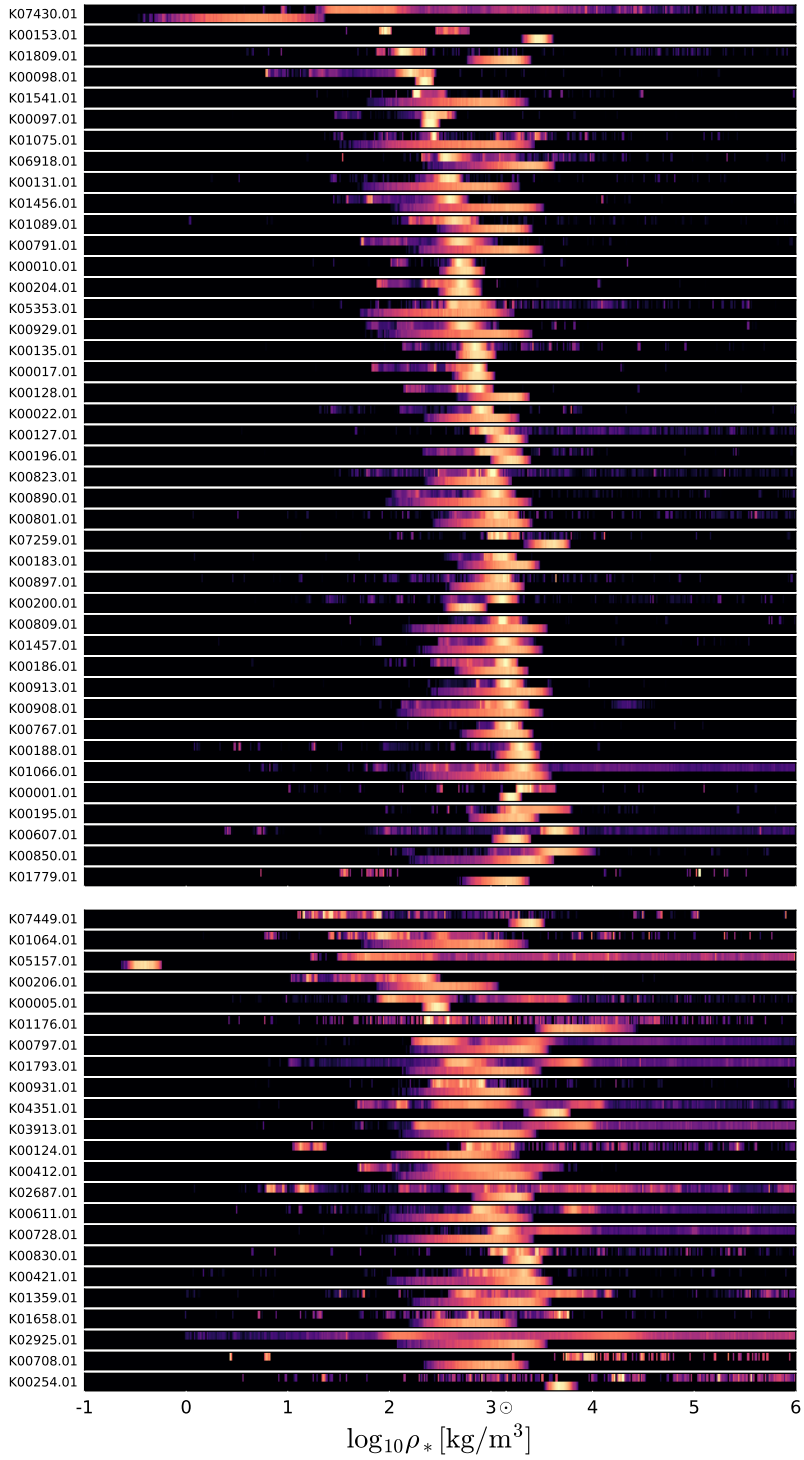

Figure 2.6: A comparison of our ρ_* posteriors to previously published constraints. Top block: KOIs for which we achieve $\leq 5\%$ fractional uncertainty on $\log_{10}\rho_*$ (62% of targets); bottom block: KOIs for which we do not. For each KOI, the upper row shows the posterior distribution of $\log_{10}\rho_*$ derived in this work, and the lower row shows the *Kepler* Data Release 25 constraint (Mathur et al., 2017) on $\log_{10}\rho_*$ derived from Dartmouth Stellar Evolution Database isochrone modeling. Within each block, the KOIs are sorted from top to bottom in order of increasing median $\log_{10}\rho_*$ from our results.

$T_{\rm eff}$, as shown in Figure 2.2. Opaque circles in this figure represent stars for which we achieve fractional $\log_{10}\rho_*$ uncertainty of less than 5%, and transparent squares represent stars for which we do not.

2.3.2.1 Comparisons with asteroseismology

Four of our occultation targets (KOIs 1.01, 5.01, 97.01, and 98.01) have previously been targeted for asteroseismic density measurement by Huber et al. (2013). In Figure 2.7, we compare our ρ_* posteriors directly to the asteroseismically measured ρ_* for each of these four targets.[2]

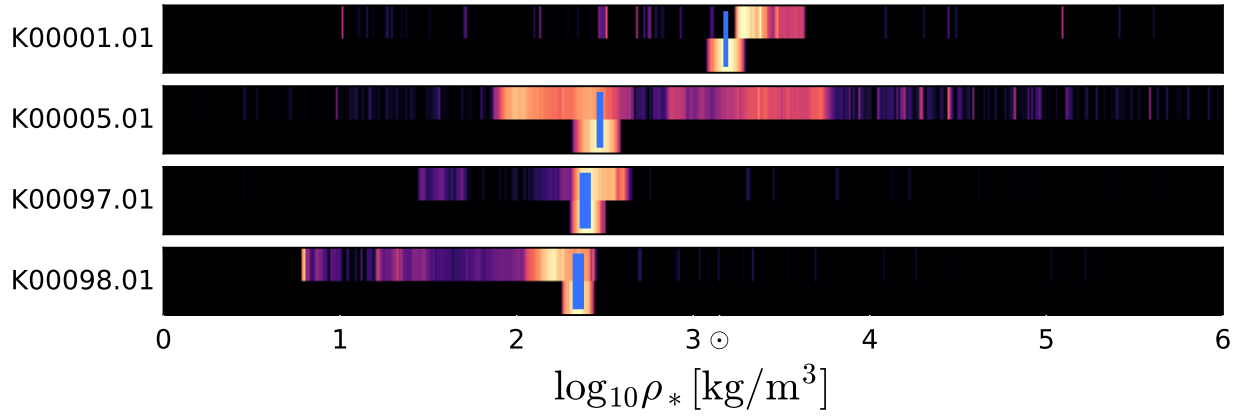

Figure 2.7: A comparison of our ρ_* posteriors (top rows) to the Mathur et al. (2017) ρ_* posteriors (bottom rows) and asteroseismic constraints on ρ_* by Huber et al. (2013) (blue boxes). We discuss the evident discrepancy between our results and asteroseismology for KOI-1.01 (TrES-2b) in Section 2.3.2.1.

Our results are in good agreement for KOIs 5.01, 97.01, and 98.01; KOIs 5.01 and 97.01 are in 1σ agreement with asteroseismology, and KOI-98.01 in 2σ agreement. We achieve comparable precision to asteroseismology except in the case of KOI 5.01, which undergoes grazing transits at low signal to noise and is subject to the parameter covariances discussed in section 2.3.1.

[2]The Mathur et al. 2017 posteriors are derived by feeding these asteroseismic constraints through isochrone modeling, so they are in excellent agreement with the Huber et al. 2013 results.

For the final target with available asteroseismic data, KOI-1.01 (TrES-2b), we derive a higher stellar density than previously published constraints. Although our transit model for this planet is well-converged and a good match to the *Kepler* data, we note that our best-fitting parameters conflict with earlier results from very reliable analyses—in particular, we derive an eccentricity $e = 0.2^{+0.16}_{-0.08}$, while radial velocity observations agree that this planet's orbit is consistent with being circular (O'Donovan et al., 2006, 2010; Kipping & Bakos, 2011; Coughlin & López-Morales, 2012). e and ρ_* are covariant, and the sense of the covariance is such that a too-high e would indeed cause us to overestimate ρ_*. We must then explain why we derive such a high e, especially given that we impose a strong e prior which should favor near-zero values of e.

We attribute our implausibly high e to a failure of our transit model to accurately capture the limb-darkening behavior of KOI-1.01's host star. Upon closer inspection of KOI-1.01's posterior distributions, we observe that the posterior distributions of the limb-darkening coefficients are strange, particularly that of α_r. While the vast majority of our target KOIs exhibit well-behaved α_r distributions like those of KOI-929.01 (see Figure 2.4: the α_r distribution is broad and peaked in the middle of the allowable α_r range), KOI-1.01's is instead narrowly peaked at $\alpha_r = 0.01^{+0.07}_{-0.01}$. In other words, it abuts the lower boundary of our uniform prior on α_r, which indicates that the limb-darkening behavior of our highest-likelihood model is unphysical.

Our transit model fails to capture the limb darkening behavior of KOI-1.01's host star because KOI-1.01 undergoes a grazing transit (we derive $b = 0.79 \pm 0.01$; the NEA reports $b = 0.818 \pm 0.001$). In other words, the planet transits across the limb of the star, so the transit data contains no information about the star's limb darkening behavior near the center of the sky-projected star ($\mu = 1$). Our three-parameter limb darkening law is actually somewhat of a liability in this situation: a very flexible model, subject to minimal constraining data, is free to adopt physically implausible (though still technically permitted within the bounds of the priors) combinations of the αs in pursuit of the highest-likelihood

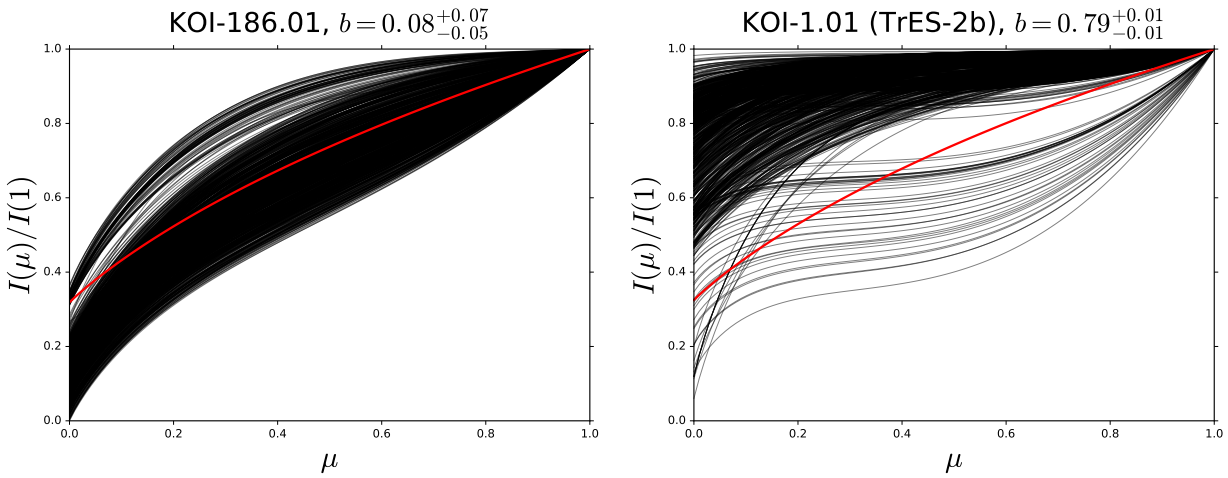

Figure 2.8: The limb darkening profile of KOI-186.01, which transits close to the midplane of its host star, compared to that of KOI-1.01 (TrES-2b), which undergoes a grazing transit. In each panel, we plot profiles $I(\mu)/I(1) = 1 - c_2(1-\mu) - c_3(1-\mu^{3/2}) - c_4(1-\mu^2)$ computed from 1000 random draws from the posterior distributions of the LDCs (black lines) and theoretical predictions for the limb darkening profile from Sing (2010) (red lines).

solution, where a less flexible model, with fewer free parameters, would be fixed by fewer constraints.

To illustrate the undesirable effects of the flexibility in the limb darkening model in the case of grazing transits, in Figure 2.8, we compare our measured limb darkening profile of KOI-1.01 (TrES-2b) to that of KOI-186.01, which transits very close to the midplane of its host star. KOI-186.01's limb darkening profile is well-constrained from $\mu = 0$ (the stellar limb) to $\mu = 1$, while KOI-1.01's is poorly constrained, with a wide range of plausible α behavior.

Visual inspection of the posterior distributions for KOI-1.01 confirms that the posterior samples with near-zero α_r correspond to unrealistically high values of e and ρ_*.

We examine the remainder of our target list for other stars which exhibit similarly suspicious α posteriors, and we also compare our results to the theoretical predictions of Sing (2010), based on stellar atmosphere modeling, in figure 2.9. This figure shows a comparison of our observed α-values to the Sing (2010) predictions, which we calculate by linearly inter-

polating their Table 2 results and evaluating the interpolation at the NEA-provided stellar spectroscopic parameters for our target stars. To check for general consistency between our results and the Sing (2010) predictions, we plot 3σ uncertainty bounds on our α-values. We highlight the results for KOI-1.01 in bright blue; note in particular its anomalously low α_r value.

We identify a handful of suspect stars which exhibit similarly anomalous values of any of the three αs, abutting either the lower ($\alpha = 0$) or upper ($\alpha = 1$) boundaries of our prior, and which also have derived values of e which are inconsistent with their strong eccentricity prior. The KOIs meeting these criteria are, of the occultation targets, KOIs-1.01 (as discussed already) and 823.01; of the tidal targets, KOIs-1075.01 and 1658.01; and of the multi-planet targets, the KOI-153 and 1779 systems. All of these lone KOIs, and at least one KOI orbiting each of the suspect multi-planet targets, undergo grazing transits, so their behavior is overall consistent with the case of KOI-1.01, discussed above. (We note that occultation target KOI-1541.01 also has an anomalously high α_h, but that its correspondingly high eccentricity is consistent with the priors in $e\cos\omega$ and $e\sin\omega$ adopted from Coughlin & López-Morales (2012), and also that it transits at very low impact parameter, so its transits contain information about its entire limb darkening profile.)

The case of KOI-1.01 demonstrates that strangely behaved LDC posterior distributions strongly indicate that other transit model parameters—especially ρ_*—may not be reliable. We therefore advise extreme caution in adopting our transit model parameter posteriors for these stars.

Aside from these isolated cases, which comprise 12% of our target list, Figure 2.9 establishes that our results are generally in good agreement with the predictions of Sing (2010). 79% of our target stars are consistent with Sing (2010) at the 3σ level in all three α-dimensions, and 94% in at least two of the three α-dimensions. We note that, although our results statistically agree, there are systematic offsets between our α_r and α_h values and those of Sing (2010); in particular, we overpredict α_r and underpredict α_h. These two pa-

rameters, however, as we discuss in Section 2.3.1, are not independent—rather, they co-vary in exactly the sense observed in this figure, with high α_r corresponding to low α_h. We find that the fractional uncertainty in the αs is positively correlated with impact parameter b, consistent with the results for KOI-1.01.

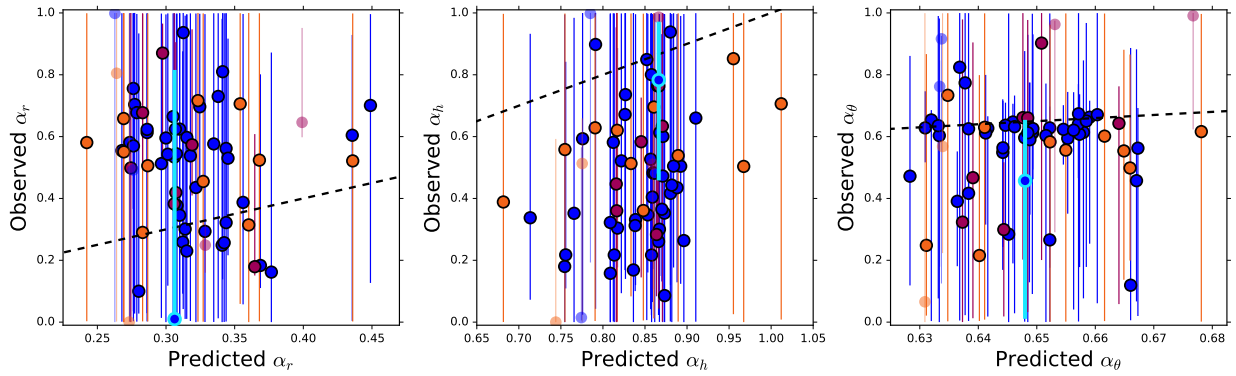

Figure 2.9: A comparison of our α_r, α_h, and α_θ to the theoretical predictions of Sing (2010), based on stellar atmosphere modeling. We plot 3σ uncertainty bands on our values to check for broad consistency. Blue points are occultation targets; orange points are tidal targets; purple points are multi-planet targets. The dotted lines indicate one-to-one correspondence, and KOI-1.01 (TrES-2b) is highlighted in bright blue. Transparent points are those with at least one anomalous α distribution, abutting either the upper or lower boundary of the prior.

2.3.3 Limb darkening coefficients

We also investigate the relationship of our measured nonlinear limb-darkening coefficients, α_r, α_h, and α_θ, to various stellar properties from the *Kepler* input catalog. In Figure 2.10, we plot various projections of this high-dimensional stellar parameter space to look for correlations. We find that the three αs are totally uncorrelated with *Kepler*-band magnitude, $T_{\rm eff}$, $\log g$, stellar radius, stellar mass, and each other. The only pattern of note in this parameter space is the sharp peak of α_θ about a value of approximately 0.6; that so many disparate target stars share this value indicates that α_θ is especially uninformative with respect to stellar properties.

When we transform the three αs into the more traditional nonlinear LDCs c_2, c_3, and c_4 (see e.g. Sing 2010; Claret 2000), the LDCs remain essentially uncorrelated with any stellar properties. The strong peak in the α_θ distribution evident in Figure 2.10 translates to a strong peak in c_4, indicating that c_4 is the least informative LDC with respect to stellar properties.

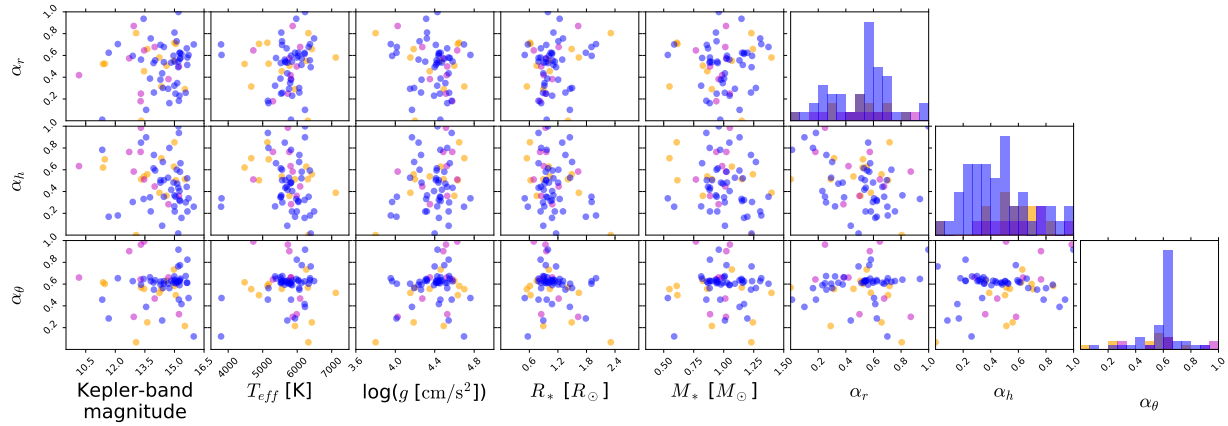

Figure 2.10: Reparametrized limb darkening coefficients (α_r, α_h, and α_θ) vs. various properties of our target stars. Blue points are occultation targets; orange points are tidal targets; purple points are multi-planet targets. There are no significant correlations between the LDCs and stellar properties. We note that the α_θ distribution is strongly peaked at 0.6 across a broad range of stellar properties, suggesting that this coefficient contains almost no information about the properties of the star.

2.3.3.1 Which stars are the best transit-modeling targets?

Having demonstrated the capability of transit modeling to yield high-precision measurements of stellar density, we now ask: Are some stars better suited to measurement by this method, and if so, can we identify those stars ahead of time? In other words, are there any properties of a star or its associated KOIs that predict a successful, precise transit-based ρ_* measurement, or disqualify a star from such a measurement?

In Figure 2.11, we plot the fractional uncertainty of each of our transit-based ρ_* measurements against various stellar and KOI data properties. Stellar properties include the *Kepler*-band magnitude, $T_{\rm eff}$, $\log g$, stellar radius, and stellar mass; KOI properties include

whether short-cadence observations were available, whether the KOI is "confirmed" per its NEA disposition, and the NEA threshold-crossing event signal-to-noise ratio (SNR).

We find that the achieved precision on ρ_* does not depend on any stellar properties, meaning that faint and bright, hot and cool, large and small stars are equally appropriate targets, *a priori*, for transit-based stellar density measurements. This lack of any dependence of the success of our method on the properties of our target stars is evident in the distribution of opaque circles (target stars measured to high ρ_* precision) across Figure 2.2.

The fractional uncertainty in our ρ_* measurements does, however, correlate strongly with the NEA-reported SNR, which is sensible given that the precision of our derived transit parameters, including ρ_*, depends on our ability to successfully model transits. We furthermore find that planets observed in short cadence are more likely to have precise ρ_* measurements, but that short-cadence data is not necessary to achieve this level of precision. Planets which are confirmed per the NEA are also more likely to have high-precision ρ_* measurements, which is partly due to our ability to discard posterior samples with unphysical R_p/R_* (and corresponding ρ_*) for these KOIs. Another contributing factor is that planets which are easy to confirm by other exoplanet detection methods (large, close to their host stars) are also likely to have high transit SNR.

2.4 Conclusions

In this work, we demonstrate the promise of exoplanetary transits to characterize planet-hosting stars with high precision. We select 66 target planet-star systems with strong prior constraints on planetary eccentricity, either directly measured from secondary eclipses, or strongly implied by a short tidal circularization timescale or compact multiplicity. We fit transit models to these targets and derive posterior distributions of ten transit parameters: transit epoch, period, impact parameter, stellar density, ratio-of-radii, reparametrized eccentricity and argument of periastron, and three reparametrized coefficients of a modified

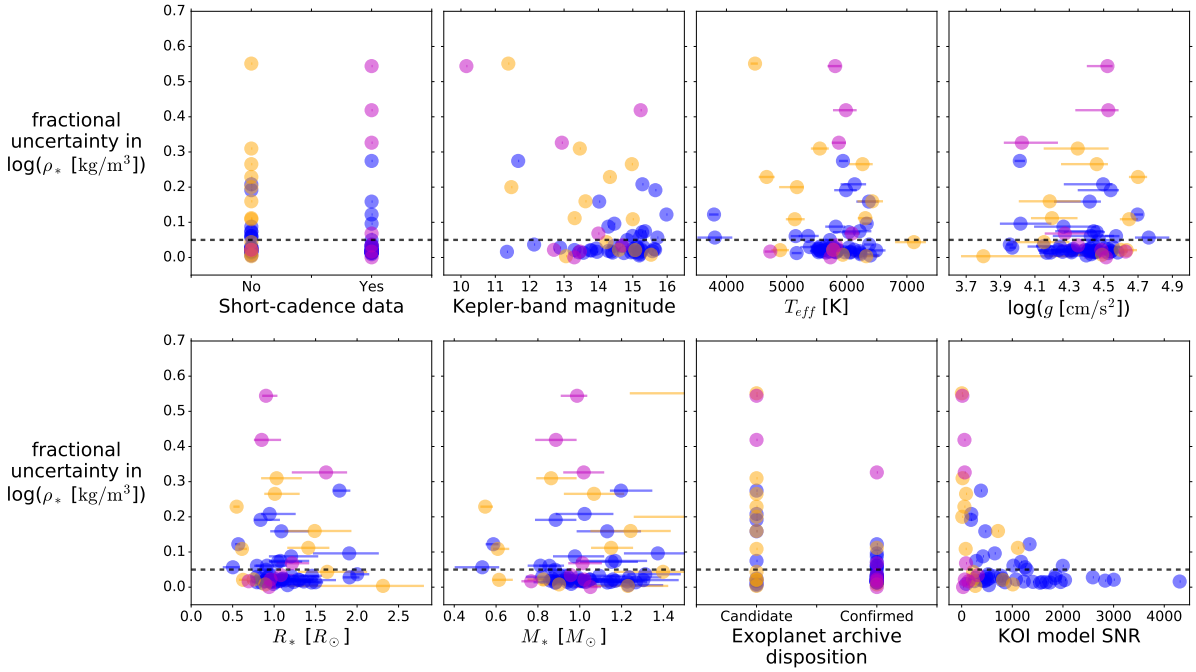

Figure 2.11: The fractional uncertainty in ρ_* vs. stellar and KOI data properties for each of our 66 targets. Blue points are occultation targets; orange points are tidal targets; purple points are multi-planet targets; the dotted lines in each panel mark 5% fractional uncertainty in ρ_*.

nonlinear limb darkening law. We make downsampled posterior distributions for the transit parameters of the 75 KOIs orbiting our 66 target stars available at https://doi.org/10.5281/zenodo.1028515.

For 95% of our targets, our measured stellar densities are in agreement with previously published constraints at the 3σ level (55% at the 1σ level). Furthermore, for 50% of our targets, we improve upon the best available published constraint on stellar density; the median improvement is slightly greater than a factor of two, meaning that we achieve a fractional uncertainty less than half of the literature value. For 62% of our targets, we achieve comparable precision to asteroseismology (typical fractional uncertainty $\leq$ 5%), generally considered the gold-standard method of stellar density measurement. We demonstrate that the success of our method for any given target planet-star system does not depend on any of the star's properties, but instead depends only on the signal-to-noise ratio of the planetary

transits.

Correspondingly, we successfully use this method to extend asteroseismic-level-precision stellar density measurements to stars three magnitudes fainter than the *Kepler* asteroseismic limit, across a broad range of effective temperatures. We note that, although TESS will observe brighter stars than *Kepler*, its asteroseismic limit will be several magnitudes brighter ($\sim 8^{\text{th}}$ magnitude) due to its small aperture (Campante et al., 2016; Ricker et al., 2014), and therefore that this transit-based method will be invaluable in characterizing TESS stars which are inaccessible to asteroseismology.

We emphasize that this method requires no data beyond a transiting exoplanet light curve, and it therefore promises to aid greatly in exoplanet host star characterization in the era of TESS and LSST, when we expect to discover many more transiting planets than we can hope to quickly follow up spectroscopically.

Finally, in advance of TESS, we note the potential of this precise transit fitting technique to characterize not just stars, but also singly-transiting planets, as shown by Yee & Gaudi 2008. We demonstrate in this work that transiting planets with strong prior eccentricity constraints may be used to strongly constrain their host stars, yielding very precise ($\leq 5\%$ uncertainty) measurements of their hosts' properties, including ρ_*. Once a host star is "anchored" by a transiting planet (a "*stellar anchor*" planet) in this way, the properties of any *other* transiting planets in the system can be derived to higher precision using the transit-measured stellar parameters than would be possible without this information.

TESS' observational baseline will be only 27.4 days over $\sim 30,000\,\text{deg}^2$ on the sky; in these regions, TESS will be unable to directly measure the period of any planet with a period greater than 27.4 days, because it will observe at most one transit of such a planet (Ricker et al., 2014). As shown in Figure 2.12, this short baseline precludes direct measurement of the periods of planets in a large region of "habitable zone" parameter space.

However, if any of these long-period, singly-transiting planets orbits the same star as a short-period stellar anchor planet, we will be able to use the anchor's transits to precisely

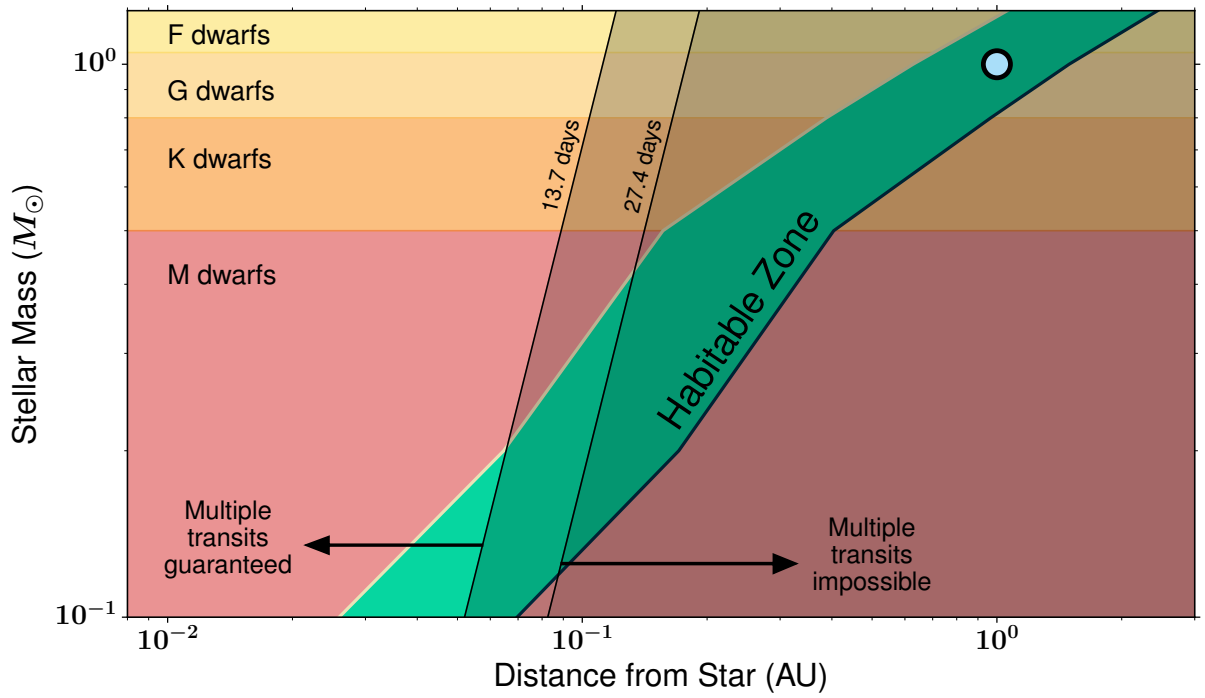

Figure 2.12: The approximate location of the habitable zone about late-type stars (Kopparapu et al., 2013). Planets in the dark gray shaded region will transit at most once during TESS' 27.4-day single-visit observational baseline. Earth is plotted as a blue dot.

measure ρ_* with the method developed in this work, then place better constraints on the period of the single-transiter using the stellar density constraint.

In the simplest case of an outer single transiter on a circular orbit, we expect the fractional uncertainty of the period P to equal

$$\frac{\sigma_P}{P} = \frac{1}{2}\sqrt{\left(\frac{\sigma_{\rho_*}}{\rho_*}\right)^2 - \left(3\frac{\sigma_{(a/R_*)}}{a/R_*}\right)^2} \tag{2.15}$$

by the propagation of uncertainty through equation 2.3. Here, P and a/R_* are parameters of the outer single transiter, and ρ_* is the density of the star that the outer single transiter and the inner stellar anchor both orbit.

We assume that the fractional uncertainty in the stellar density, $\frac{\sigma_{\rho_*}}{\rho_*}$, will dominate over the fractional uncertainty in the normalized semimajor axis, $\frac{\sigma_{(a/R_*)}}{a/R_*}$, because a/R_* can be

measured directly from the shape of the single observed transit of the outer planet, while ρ_* requires detailed modeling of the inner stellar anchor's transits subject to a prior constraint on eccentricity, as described in this work. Under this assumption, a 5% fractional uncertainty in ρ_* corresponds to a 2.5% fractional uncertainty in the period P of the single transiter.

While considerably less precise than a direct period measurement, this degree of fractional uncertainly could establish whether a given planet orbits in the habitable zone of its host star or not, and hence whether it merits follow-up observations. Obtaining a similar constraint on the period using radial velocity measurements of the host star would be time-consuming, generally requiring an observational baseline comparable to the orbital period of the planet (see e.g. Ford 2005). Our method therefore promises to aid greatly in the characterization of long-period TESS planets.

Acknowledgements

The authors thank members of the Cool Worlds Lab for useful discussions. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

3. Estimation of singly-transiting K2 planet periods with Gaia parallaxes

3.1 Introduction

Choosing an observational baseline for a transit survey places fundamental limits on which planets will be observed to transit. For a baseline B, a planet with orbital period $P > B$ will transit *at most* once during the baseline, with the transit probability falling off as P^{-1} for longer periods (Yee & Gaudi, 2008).

As a result, transit observations of relatively long-period exoplanets are rare, even as long-period exoplanets themselves merit intense study. Planets with long periods relative to the *Kepler* baseline ($\sim$ 1500 days), for example, are interesting as analogs to the outer planets of the Solar System, and as examples of planets at or beyond the snow line (e.g. Kipping et al. 2016). The shortest TESS baseline, 27.4 days for most of the sky (Ricker et al., 2015), relegates even habitable-zone planets around FGKM stars to the "long-period" regime.

However, even when a long-period transiting exoplanet is observed, constraining that planet's orbital period (and hence its distance from its host star) is difficult unless we see at least two successive transits. The observational baseline of the *Kepler* survey was so long that few single-transit candidates were observed. Wang et al. (2015) identify 17 single-transit candidates; Uehara et al. (2016) enumerate a further 23 (of which 14 are new discoveries, and 9 rediscoveries of single transits identified by the *Kepler* transit search pipeline); Foreman-

Mackey et al. (2016) identify a further 6. Altogether, this yields a catalogue of 46 single-transit events from the initial *Kepler* mission.

In the repurposed K2 mission, which has a shorter observational baseline over each observed field ($\sim$ 75 days), Osborn et al. (2016) identify 7 single-transit candidates. LaCourse & Jacobs (2018) catalogue 164 single-transit events, although they caution that most are likely eclipsing binaries.

Villanueva et al. (2019), in contrast, estimate that TESS will observe more than 200 single transits among its postage-stamp targets (observed at 2-minute cadence), and a further $\sim$ 1000 in full-frame images (observed at 30-minute cadence). They further estimate that, if they are confirmed as planets, these single-transits will double the postage-stamp targets' yield of planets with $P > 25$ days, and increase the yield of $P > 250$ day planets tenfold. Huang et al. (2018), meanwhile, predict a more modest 75 single transits among the postage-stamp targets, and $\sim$ 700 in the full-frame images, but add that the single transiters will increase TESS's yield of temperate planets around FGK stars roughly threefold.

Figuring out how to accurately constrain the period distribution of single-transiters is therefore critical to enhancing the long-period planet yield of the forthcoming TESS data, as knowledge of that distribution will be important for follow-up observations aiming at confirming the planetary nature of any signals detected. There is a simple approach, first suggested by Yee & Gaudi (2008), to constrain the period with observations of a single transit. From Kepler's third law, we can relate the period of a planet to the density of its host star (ρ_*) and its normalised semi-major axis (a/R_*):

$$P^2 = \frac{3\pi}{G}\left(\frac{a}{R_*}\right)^3 \rho_*^{-1} \tag{3.1}$$

We may measure the single-transiter's a/R_* directly by modeling its transit shape; ρ_*, however, must come from an independent observation of the star.

Propagating uncertainty through that equation, we may derive:

$$\frac{\sigma_P}{P} = \frac{1}{2}\sqrt{\left(\frac{\sigma_{\rho_*}}{\rho_*}\right)^2 + \left(3\frac{\sigma_{(a/R_*)}}{a/R_*}\right)^2} \tag{3.2}$$

If we assume that the uncertainty in ρ_* will dominate over the uncertainty in a/R_*, the above equation means that a $\sim 5\%$ uncertainty in ρ_* translates to a $\sim 2.5\%$ uncertainty in P for the single transiter ($\sigma_P \simeq 9$ days for an Earth analog). In summary, a precise constraint on ρ_*, plus a single transit observation, may yield a precise constraint on the single transiter's period P.

Of course, this simple order-of-magnitude calculation omits some important details of the transit modeling, with which one actually obtains a/R_*. First, the retrieved value of this parameter, which comes from knowledge of the overall shape of the light curve in addition to the transit duration, is dominated by other factors in the transit modeling, such as limb-darkening and whether or not the planet's orbit is assumed to be circular. These two effects can lead to large biases in the retrieval of transit parameters if not accounted for (see, e.g., Espinoza & Jordán, 2015, on the impact of limb-darkening on the retrieval of a/R_*). In addition, the long cadence of the data that missions like *Kepler* and TESS provide for most stars puts an even stricter limit on the accuracy with which a/R_* can be retrieved (Kipping, 2010b; Price & Rogers, 2014). (If we were to assume that the uncertainty in a/R_* dominated in Equation 3.2, a $\sim 5\%$ uncertainty in a/R_* would translate to a $\sim 7.5\%$ uncertainty in P for a single transiter, or ~ 1 month for an Earth analog.)

Nevertheless, the above approach, of constraining a single-transiter's P by constraining its host star's density, has been adopted by nearly all of the single-transit-catalogue works listed above. However, in most cases, because ρ_* is not especially well-constrained by many types of stellar observations, they derive large P uncertainties for their single-transiters. Wang et al. (2015), for example, estimate P for their single-transiters using host star densities interpolated from isochrones, achieving a typical fractional P uncertainty of $\sim 100\%$. Uehara et al. (2016) adopt stellar density constraints from the *Kepler* Community Follow-Up

Observing Program, and furthermore assume circular orbits for their single-transiters, which limits the achievable P precision severely. Foreman-Mackey et al. (2016) fit transit models, including inferred P, using priors on stellar mass and radius from the *Kepler* DR25 Stellar Properties Catalog (Mathur et al., 2017) and derive similarly large uncertainties on P (the typical fractional uncertainty for their single transiters is also $\sim 100\%$). Osborn et al. (2016) estimate stellar parameters by deriving stellar temperatures from broad-band colours, then calculating stellar mass and radius from those temperatures under the assumption that their stars were on the main sequence, and derive typical fractional uncertainty $\sim 50\%$; however, they still assume that these single transiters are on circular orbits.

Other methods of stellar characterisation, however, can yield significantly more precise constraints on ρ_*, which may translate to correspondingly narrow bounds on P for single-transiters. Asteroseismology, for example, yields a typical ρ_* precision of $\sim 5\%$ (see e.g. Huber et al. 2013, Silva Aguirre et al. 2017). Sandford & Kipping (2017) demonstrate that the method of "stellar anchors," in which ρ_* is measured by modeling the transits of a planet with independently well-constrained eccentricity, can also yield ρ_* uncertainties of order 5%.

In this paper, we investigate in detail the performance of a third method of constraining ρ_*: combining stellar radius measurements derived from Gaia DR2 parallaxes (Gaia Collaboration et al., 2018) with stellar mass measurements from isochrone fitting. We apply such stellar density constraints to long-period planets observed by K2. First, we use these stellar density constraints to model individual transits of known-period K2 planets *as if they were single-transiters*, infer their periods, and investigate the precision and accuracy of the inferences. We then apply the method to 12 true single-transiters observed by K2. Throughout, we treat eccentricity as a free parameter in the transit model fits.

3.2 Method

3.2.1 Stellar density estimation

In order to determine the stellar physical parameters, we follow a procedure similar to that presented in Brahm et al. (2019, 2018). For a given star, we first estimate the stellar radius by combining its publicly available photometry with its GAIA DR2 parallax measurement. For this step we also require an estimate of the stellar atmospheric parameters in order to select a spectral energy distribution (SED) model to represent the star being analysed. For all the systems analysed in the present study we adopt the `BT-Settl-CIFIST` (Baraffe et al., 2015) SED models. We consider the following sets of photometric surveys/bands in our analysis: APASS (V,B,g,r,i; Henden et al., 2009), 2MASS (J,H,Ks; Skrutskie et al., 2006), WISE (W1,W2,W3; Wright et al., 2010). For each star we construct the observed reddened emitted flux density at the surface of the star:

$$\vec{F_o} = 4\pi d_{gaia}^2 \cdot \vec{f}, \tag{3.3}$$

where d_{gaia} is the distance to the star computed from the Gaia parallax, $\vec{f}$ are the flux densities computed from the observed magnitudes, and the vectors have a length equal to the number of passband filters that are being considered. We compute the uncertainties in $\vec{F_o}$ by propagating the uncertainties on the observed magnitudes and parallax. The distances are estimated using the estimation procedures presented in Bailer-Jones et al. (2018).

Meanwhile, our model for the reddened emitted flux density at the surface of the star takes the form of a vector $\vec{F}$, where the m^{th} component is given by:

$$F_m = 4\pi R_*^2 f_m \cdot e^{-A_\lambda}, \tag{3.4}$$

where R_* is the stellar radius, f_m is the synthetic flux density in bandpass m generated from

the `BT-Settl-CIFIST` SED, and A_λ is the reddening or extinction factor in that bandpass.

We assume that reddening follows the Cardelli et al. (1989) law, and therefore we consider just a single reddening parameter A_V in our model, which combined with the reddening law generates extinction factors for each passband filter. With the model and observed flux vectors in hand, we explore the posterior distributions for R_* and A_V using the `emcee` package (Foreman-Mackey et al., 2013) and a log-likelihood given by $\log \mathcal{L} = \sum_i (F_{o,i} - F_i^m)^2 / \sigma^2_{F_{o,i}}$. We adopt uniform priors in R_* ($[0.1 R_\odot, 100 R_\odot]$)and A_V ($[0, 1]$) .

Once the stellar radius is estimated, we proceed to estimate the stellar mass and age by comparing the radius and effective temperature given by stellar evolutionary models to the observed values for these parameters. Specifically, we use the Yonsei-Yale isochrones (Yi et al., 2001) as our model, where we fix the metallicity to the reported value, and we use the interpolating code provided with the isochrones for generating a set of modeled $\mathrm{R}_\star$ and $T_{\rm eff}$ from an arbitrary stellar mass and age. The distributions for $\mathrm{M}_\star$ and $\mathrm{AGE}_\star$ are explored using the `emcee` package. Again we use uniform priors for $\mathrm{M}_\star$ ($[0.4\ \mathrm{M}_\odot, 4.5\ \mathrm{M}_\odot]$) and $\mathrm{AGE}_\star$ ([0.05 Gyr, 4.5 Gyr]). Finally, we use the obtained distributions of the of stellar masses and radii to determine the distribution of the stellar bulk density (assuming sphericity):

$$\rho_\star = \frac{M_\star}{\frac{4\pi}{3} R_\star^3}. \tag{3.5}$$

The use of stellar evolutionary models for the estimation of stellar densities can in principle produce systematic biases if the models are not well calibrated. In order to test the accuracy of our derived densities, we compare the results obtained with our method with those obtained with an independent and more precise technique. Specifically, we use a sample of stars that have asteroseismic density determinations. From the study presented in Silva Aguirre et al. (2015, 2017) of *Kepler* stars with densities derived through asteroseismology, we select stars that have Gaia DR2 parallaxes and no reported companions closer than 4". We also select two giant stars, K2-97 and K2-132, that host close-in planets, which have

density estimations from K2 photometry (Grunblatt et al., 2016, 2017; Jones et al., 2018).

We use the methods described above to compute the densities for these stars, adopting for them the stellar atmospheric parameters reported in the literature. Figure 3.1 compares our density estimates to those obtained through asteroseismology. The densities range from 0.02 g cm^{-3} to 2.5 g cm^{-3}, which is equivalent to a red giant and a K-type dwarf, respectively. The densities computed with our method are consistent to those obtained through asteroseismology and the residuals show no significant biases or trends between the two methods. We also find that the residuals present a root mean square of 0.048 g cm^{-3}, which is slightly larger than our mean uncertainty in density (0.033 g cm^{-3}), and could signify that our method and/or the asteroseismic method underestimates the uncertainty in density (by $\approx 40\%$ at maximum).

To confirm that the two methods are consistent, we compute the Bayesian evidence for seven possible relationships between our stellar densities and those obtained via asteroseismology:

1. A one-to-one relationship with an additional noise term, $\rho_*^{gaia+YY} = \rho_*^{aste} + \sigma_{extra}^2$;
2. A relationship with a constant offset, $\rho_*^{gaia+YY} = \rho_*^{aste} + b$, with and without σ_{extra}^2;
3. A relationship with a stellar-density dependent offset, $\rho_*^{gaia+YY} = a\rho_*^{aste}$, with and without σ_{extra}^2; and
4. A relationship with both a constant and stellar-density dependent offset, $\rho_*^{gaia+YY} = a\rho_*^{aste} + b$, with and without σ_{extra}^2.

Of these, model (i) has the highest Bayesian evidence, with measured $\sigma_{extra}^2 = 0.0313 \pm 0.0044$g/cm^3, which is roughly equal to our mean uncertainty in density. We therefore conclude that we have underestimated our error bars by roughly a factor of two.

We conclude from this exercise that the stellar densities determined using Gaia parallaxes and the Yonsei-Yale isochrones are reliable estimates to characterise the host stars of single transiters in order to predict their orbital periods.

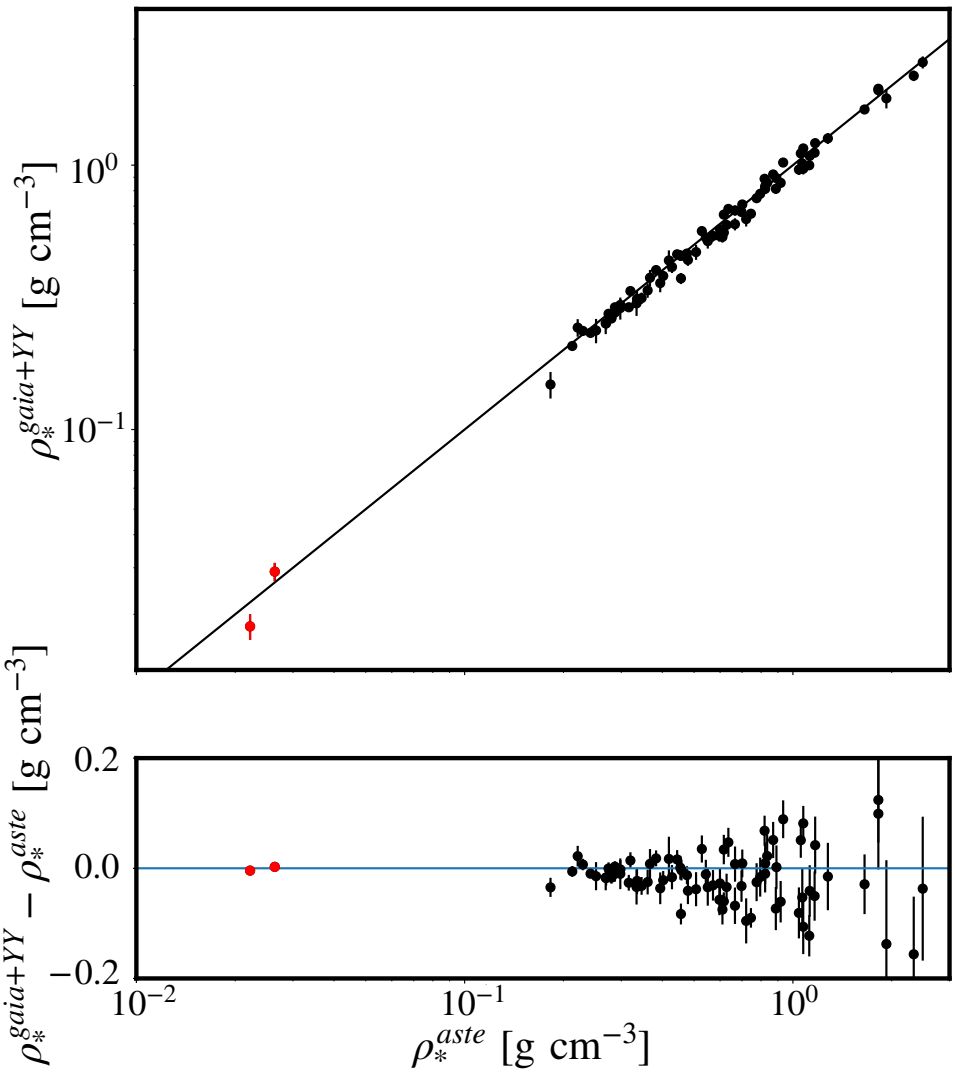

Figure 3.1: The top panel shows a comparison between stellar densities estimated with asteroseismology (x-axis) and those computed using Gaia parallaxes and the Yonsei-Yale isochrones (y-axis). The black points correspond to the sample of Kepler host stars presented in Silva Aguirre et al. (2015, 2017), while the red points correspond to the two giant stars that have been found to have transiting giant planets using K2 data, K2-97 and K2-132. The bottom panel presents the density difference between the two methods as a function of the asteroseismic density.

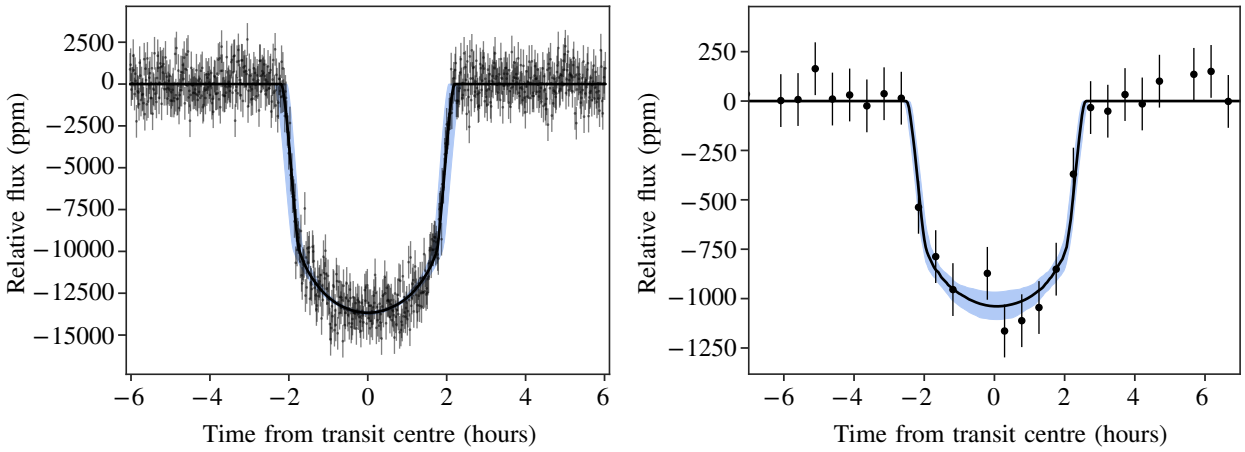

Figure 3.2: Illustrative plots of our single-transit fits to known exoplanets HATS-11b (short cadence, left) and K2-96c (long cadence, right) from K2 photometry (black dots with error bars). Solid black lines present our best-fit models; blue bands the 1-sigma credibility band given our posterior parameters. The out-of-transit trend, which we fit by Gaussian process regression, has been subtracted off of the K2 data.

3.2.2 Light curve analysis

Fitting single-transit light curves is a complex problem on its own, as it entails sampling parameters that are strongly correlated with one another due to the fact that we are extracting information of several parameters from the same portions of the light curve (Seager & Mallén-Ornelas, 2003; Winn, 2010). For example, the transit duration simultaneously gives information on every parameter that defines the transit light curve except for the limb-darkening parameters, whereas the ingress and egress times (which are slightly different if we assume eccentric orbits) constrain all the parameters, including the limb-darkening coefficients. All this implies that complex, possibly multi-modal solutions are possible for a given transit light curve.

Because of these possible complexities, we fit these transits using `MultiNest` (Feroz et al., 2009) through the `PyMultiNest` package (Buchner et al., 2014), which allows us to efficiently explore the posterior distribution of the parameters given the data in this potentially multi-modal and degenerate case, and thus marginalise over those possibilities when estimating the posterior distribution of the periods of our single transiters given the data.

In our transit fits, we assume as free parameters $\vec{\theta}$ all the physical and orbital parameters that define the transit light curve: the planet-to-star radius ratio R_p/R_* and the transit impact parameter b, reparametrised as described in Espinoza (2018) for improved sampling efficiency; the stellar density ρ_*; the period P; the time of transit centre t_0; the argument of periastron passage ω; and the eccentricity e. We describe the limb-darkening effect with a linear law through a parameter q, so the stellar intensity profile has the form $I(\mu) = 1 - q(1-\mu)$ with $q \in (0,1)$ in order to sample physically plausible intensity profiles. The reason for selecting a linear law is the small number of informative, in-transit data points in any given single long-cadence transit: in such a situation, the variance generated by other laws is greater than the bias generated by the linear law (i.e., the achievable precision on the limb darkening parameters of a more complex law would be severely hampered by the small number of data points in our light curves; see e.g. Espinoza & Jordán 2016). To allow a like-for-like comparison between all of the transits in our validation sample, we apply the same linear law to the short-cadence transits as well. However, in general, single transiters should be assessed individually in order to find the best limb-darkening law for a given candidate following the prescriptions of Espinoza & Jordán (2016).

To account for out-of-transit trends in the light curve, we simultaneously fit (i) the eight-parameter transit model described above and (ii) a Gaussian process model with an exponential squared kernel in time, implemented in `george` (Foreman-Mackey et al., 2014a). This model demands an additional four parameters: a constant flux offset f_0; a "jitter" term j to model the white noise in the light curve (j equals the natural logarithm of the white noise variance added to the diagonal of the covariance matrix); and the two hyperparameters of the exponential squared kernel (a multiplicative constant and a scale parameter).

We keep 30 and 100 out-of-transit points at each side of our long- and short-cadence transits, respectively, and fit the combined transit and GP model to this subset of the full time series. Our likelihood, which we assume to be Gaussian, is constructed from the light curve information, packed in a vector $\vec{y}_{tr}$ which contains times, fluxes and errors from the

photometry. To compute the transit model, we use the `batman` package (Kreidberg, 2015), which is re-sampled for the case of K2 30-minute cadence light curves following Kipping (2010b).

For both the validation transits and the true singles, we fit the EVEREST CBV-corrected, detrended K2 light curves (Luger et al., 2016). We perform no outlier rejection for the validation transits; for the singles, we exclude a single 60σ discrepant out-of-transit data point from the light curve of EPIC 211311380d, a single 40σ discrepant in-transit data point from the transit of EPIC 211311380f, and a series of five 5σ discrepant in-transit data points (apparently the result of stellar activity) from the transit of EPIC 203311200b. These excluded points are plotted with the rest of the transit in Figure 3.11.

We impose a split-normal prior on ρ_*,

$$\begin{aligned} \Pr(\rho_*) &= A \exp\left(-\frac{(\rho_* - \mu)^2}{2\sigma_{\text{left}}^2}\right), \ \rho_* < \mu \\ &= A \exp\left(-\frac{(\rho_* - \mu)^2}{2\sigma_{\text{right}}^2}\right) \text{ otherwise} \end{aligned} \tag{3.6}$$

with mode μ and lower and upper standard deviations σ_{left} and σ_{right}derived from the procedure described in 3.2.1. (The constant A is set to normalize the integral of the prior to 1.) We assume wide uniform priors for the reparametrised b and R_p/R_*; t_0; e; ω; q; and f_0, and wide log-uniform priors for P, j, and the two Gaussian process kernel hyperparameters. Specifically, for our validation K2 planets, we choose the P prior to be a log-uniform distribution bounded between 1.0 and 1000 days.

For both our validation transits and the true single transits, we also tested the single-transit period prior suggested by Kipping (2018a) (hereafter K18), which explicitly accounts for the fact that the planet was only observed to transit once over the observational baseline, as well as the "phase" of the single transit relative to the beginning and end of the observation. As in that work, we set α (the power-law index on the underlying intrinsic period prior) equal to $-2/3$, which implies uniformity over semi-major axis; the end result is a prior which is

$\propto P^{-7/3}$ over the interval $P_{\rm min}$ and $P_{\rm max}$ and zero elsewhere.

For the true single transiters, we have a meaningful, observed $P_{\rm min}$, based on when the transit is observed (t_0) relative to the beginning ($t_{\rm start}$) and end ($t_{\rm end}$) of its K2 campaign,

$$P_{\rm min} = \max[|t_0 - t_{\rm start}|, |t_0 - t_{\rm end}|]. \tag{3.7}$$

However, for the validation sample, in which our observational baseline is in fact longer than the planet's period, we must choose $P_{\rm min}$ arbitrarily. When we adopt a K18 prior with arbitrary $P_{\rm min}$ (specifically, $P_{\rm min} = 1.0$ days to enable comparison to the log-uniform prior described above), we find that the validation fits universally converge to solutions with $P = P_{\rm min} = 1.0$ days, regardless of their true periods. In these fits, the eccentricity converges to implausibly high values in order to maintain the observed transit duration in the face of such a short P—this is true regardless of whether the prior on e is uniform, or instead a Beta distribution with parameters adopted from Kipping (2013). We conclude from this exercise that, while the K18 prior is appropriate and philosophically motivated for true single transiters, it is inappropriate for our artificial validation "single" transits with arbitrary minimum period. We therefore use the log-uniform prior described above for these validation fits.

Additionally, we impose the physical constraints that $b \leq (1 + R_p/R_*)$ (this is already imposed by the b and R_p/R_* reparametrisation of Espinoza 2018); $b \leq (a/R_*)$; and $0° \leq i \leq 90°$.

Our single-transit fitter code, `single`, is available on GitHub[1].Typical fits for our targets are presented in Figure 3.2.

[1] http://www.github.com/nespinoza/single

Table 3.1: The K2 validation planets. The 21 planets above the horizontal line were observed at long (30 minute) cadence, and the 6 planets below the line were observed at short (1 minute) cadence. To the left of the vertical line are known or assumed parameters; to the right are parameters we fit to each individual transit with MultiNest. Because the t_0 posterior distributions are generally quite symmetrical, the reported fit t_0 is the 50th percentile of the posterior distribution, with uncertainties given by the 16^{th} and 84^{th} percentiles. The posterior distributions of the other parameters are asymmetrical, so we fit each with a split-normal distribution and report $\mu^{+\sigma_{\rm right}}_{-\sigma_{\rm left}}$. (Note that the parameters reported in this table are summary statistics over multiple single-transit fits for each of these planets, and therefore are not representative of any of our individual transit fits. For posterior distributions of each individual single-transit fit, please contact the authors.)

K2 name	EPIC ID	$P_{\rm known}$ [d]	ρ_* [kg/m^3]	$P_{\rm fit}$ [d]	t_0 [BJD-2454833]	R_p/R_*	b	ω [°]	e	q
K2-03b	201367065	10.054	$5609.8^{+153.8}_{-156.2}$	$8.3^{+30.0}_{-2.0}$	$1980.419^{+0.003}_{-0.002}$	$0.035^{+0.002}_{-0.002}$	$0.4^{+0.2}_{-0.2}$	$39.9^{+122.2}_{-9.9}$	$0.0^{+0.3}_{-0.0}$	$0.7^{+0.2}_{-0.4}$
K2-03c	201367065	24.649	$5609.8^{+153.8}_{-156.2}$	$4.2^{+80.0}_{-0.6}$	$1979.270^{+0.006}_{-0.003}$	$0.027^{+0.002}_{-0.002}$	$0.4^{+0.3}_{-0.2}$	$157.1^{+106.8}_{-105.6}$	$0.0^{+0.4}_{-0.0}$	$0.6^{+0.2}_{-0.3}$
K2-10b	201577035	19.304	$1698.0^{+62.1}_{-65.2}$	$8.3^{+20.0}_{-1.0}$	$1986.583^{+0.004}_{-0.001}$	$0.038^{+0.0008}_{-0.001}$	$0.4^{+0.2}_{-0.2}$	$44.9^{+96.8}_{-14.3}$	$0.0^{+0.2}_{-0.0}$	$0.3^{+0.2}_{-0.1}$
K2-19b	201505350	7.919	$2327.0^{+62.8}_{-64.4}$	$7.2^{+10.0}_{-0.6}$	$1980.384^{+0.003}_{-0.003}$	$0.078^{+0.001}_{-0.003}$	$0.1^{+0.4}_{-0.0}$	$85.0^{+61.5}_{-29.1}$	$0.0^{+0.2}_{-0.0}$	$0.3^{+0.3}_{-0.1}$
K2-19c	201505350	11.907	$2327.0^{+62.8}_{-64.4}$	$7.7^{+30.0}_{-0.9}$	$1984.274^{+0.005}_{-0.003}$	$0.046^{+0.0009}_{-0.002}$	$0.4^{+0.2}_{-0.2}$	$32.0^{+118.6}_{-5.8}$	$0.0^{+0.2}_{-0.0}$	$0.4^{+0.4}_{-0.2}$
K2-32b	205071984	8.992	$2094.0^{+85.2}_{-78.8}$	$4.3^{+20.0}_{-0.7}$	$2067.927^{+0.001}_{-0.001}$	$0.057^{+0.002}_{-0.002}$	$0.2^{+0.3}_{-0.1}$	$33.6^{+119.4}_{-5.9}$	$0.0^{+0.3}_{-0.0}$	$0.3^{+0.3}_{-0.1}$
K2-32d	205071984	31.719	$2094.0^{+85.2}_{-78.8}$	$9.1^{+60.0}_{-0.7}$	$2070.787^{+0.005}_{-0.006}$	$0.038^{+0.001}_{-0.003}$	$0.5^{+0.2}_{-0.3}$	$47.6^{+153.6}_{-15.8}$	$0.0^{+0.3}_{-0.0}$	$0.3^{+0.4}_{-0.3}$
K2-56b	210848071	41.686	$1029.1^{+41.8}_{-38.6}$	$5.4^{+50.0}_{-0.6}$	$2235.531^{+0.002}_{-0.003}$	$0.022^{+0.001}_{-0.0006}$	$0.2^{+0.4}_{-0.1}$	$39.4^{+136.2}_{-9.8}$	$0.0^{+0.3}_{-0.0}$	$0.4^{+0.3}_{-0.1}$
K2-96c	220383386	29.845	$1913.8^{+60.1}_{-55.6}$	$10.0^{+70.0}_{-1.0}$	$2561.979^{+0.003}_{-0.002}$	$0.030^{+0.002}_{-0.0007}$	$0.4^{+0.2}_{-0.2}$	$35.1^{+140.8}_{-6.5}$	$0.0^{+0.3}_{-0.0}$	$0.6^{+0.1}_{-0.2}$
K2-98b	211391664	10.137	$417.5^{+21.7}_{-14.7}$	$2.9^{+20.0}_{-0.3}$	$2312.980^{+0.007}_{-0.004}$	$0.030^{+0.002}_{-0.001}$	$0.5^{+0.2}_{-0.3}$	$42.7^{+135.0}_{-11.4}$	$0.0^{+0.3}_{-0.0}$	$0.4^{+0.3}_{-0.2}$
K2-99b	212803289	18.249	$123.3^{+54.8}_{-12.2}$	$15.2^{+60.0}_{-2.0}$	$2400.826^{+0.002}_{-0.002}$	$0.042^{+0.001}_{-0.0008}$	$0.7^{+0.1}_{-0.3}$	$59.4^{+95.2}_{-18.5}$	$0.0^{+0.3}_{-0.0}$	$0.4^{+0.1}_{-0.1}$
K2-110b	212521166	13.864	$2879.6^{+84.1}_{-61.0}$	$5.6^{+20.0}_{-1.0}$	$2400.738^{+0.001}_{-0.001}$	$0.033^{+0.001}_{-0.0009}$	$0.3^{+0.2}_{-0.1}$	$36.9^{+110.9}_{-5.2}$	$0.0^{+0.2}_{-0.0}$	$0.7^{+0.1}_{-0.3}$
K2-113b	220504338	5.818	$838.0^{+43.3}_{-45.1}$	$1.3^{+6.0}_{-0.1}$	$2565.702^{+0.002}_{-0.002}$	$0.086^{+0.006}_{-0.006}$	$0.6^{+0.2}_{-0.3}$	$56.0^{+88.3}_{-9.8}$	$0.0^{+0.2}_{-0.0}$	$0.8^{+0.0}_{-0.3}$
K2-114b	211418729	11.391	$2196.4^{+97.0}_{-107.7}$	$7.3^{+10.0}_{-2.0}$	$2307.325^{+0.001}_{-0.001}$	$0.113^{+0.005}_{-0.001}$	$0.1^{+0.3}_{-0.1}$	$186.4^{+98.8}_{-111.2}$	$0.0^{+0.2}_{-0.0}$	$0.6^{+0.1}_{-0.1}$
K2-115b	211442297	20.273	$1671.1^{+61.0}_{-50.7}$	$10.9^{+10.0}_{-3.0}$	$2324.157^{+0.000}_{-0.002}$	$0.128^{+0.001}_{-0.004}$	$0.7^{+0.0}_{-0.1}$	$178.2^{+61.4}_{-113.5}$	$0.0^{+0.2}_{-0.0}$	$0.5^{+0.3}_{-0.0}$
K2-139b	218916923	28.382	$2129.9^{+57.5}_{-80.2}$	$24.2^{+30.0}_{-0.5}$	$2492.817^{+0.001}_{-0.001}$	$0.096^{+0.004}_{-4e-05}$	$0.0^{+0.4}_{-0.0}$	$86.9^{+19.9}_{-21.3}$	$0.0^{+0.2}_{-0.0}$	$0.6^{+0.0}_{-0.1}$
K2-140b	228735255	6.569	$1368.1^{+51.6}_{-55.6}$	$5.3^{+5.0}_{-0.5}$	$2755.286^{+0.001}_{-0.001}$	$0.114^{+0.001}_{-0.0009}$	$0.2^{+0.1}_{-0.1}$	$64.8^{+54.6}_{-9.5}$	$0.0^{+0.1}_{-0.0}$	$0.5^{+0.1}_{-0.0}$
K2-232b	247098361	11.168	$837.9^{+27.7}_{-32.1}$	$10.7^{+6.0}_{-1.0}$	$2992.353^{+0.001}_{-0.001}$	$0.090^{+0.0005}_{-0.0008}$	$0.4^{+0.0}_{-0.2}$	$67.2^{+51.3}_{-1.7}$	$0.0^{+0.1}_{-0.0}$	$0.5^{+0.0}_{-0.0}$
K2-77b	210363145	8.2	$2599.0^{+49.9}_{-60.6}$	$2.6^{+20.0}_{-2.0}$	$2237.805^{+0.003}_{-0.003}$	$0.028^{+0.002}_{-0.001}$	$0.6^{+0.2}_{-0.4}$	$69.0^{+171.7}_{-40.5}$	$0.0^{+0.4}_{-0.0}$	$0.0^{+0.5}_{-0.0}$
EPIC 248777106b	248777106	11.814	$304.3^{+48.6}_{-7.2}$	$2.8^{+10.0}_{-0.4}$	$3080.805^{+0.004}_{-0.002}$	$0.037^{+0.002}_{-0.0005}$	$0.3^{+0.3}_{-0.2}$	$34.4^{+143.9}_{-5.0}$	$0.0^{+0.2}_{-0.0}$	$0.5^{+0.2}_{-0.1}$
HATS-9b	217671466	1.915311	$375.0^{+30.2}_{-21.6}$	$1.4^{+2.0}_{-0.1}$	$2547.702^{+0.000}_{-0.000}$	$0.084^{+0.002}_{-0.0007}$	$0.1^{+0.2}_{-0.1}$	$61.6^{+61.5}_{-11.8}$	$0.0^{+0.2}_{-0.0}$	$0.6^{+0.0}_{-0.1}$
HATS-11b	216414930	3.6191634	$452.9^{+36.6}_{-35.2}$	$2.0^{+5.0}_{-0.2}$	$2545.419^{+0.000}_{-0.000}$	$0.109^{+0.001}_{-0.001}$	$0.1^{+0.2}_{-0.1}$	$39.4^{+105.6}_{-5.0}$	$0.0^{+0.2}_{-0.0}$	$0.5^{+0.0}_{-0.1}$
WASP-47b	206103150	4.1591399	$942.7^{+76.3}_{-54.1}$	$2.6^{+3.0}_{-0.2}$	$931.349^{+0.001}_{-0.000}$	$0.102^{+0.001}_{-0.0006}$	$0.3^{+0.0}_{-0.2}$	$171.6^{+84.4}_{-105.2}$	$0.0^{+0.2}_{-0.0}$	$0.6^{+0.0}_{-0.0}$
WASP-55b	212300977	4.4656291	$1050.0^{+82.2}_{-62.4}$	$3.2^{+3.0}_{-0.5}$	$1583.716^{+0.000}_{-0.000}$	$0.125^{+0.001}_{-0.002}$	$0.2^{+0.0}_{-0.1}$	$50.5^{+84.6}_{-2.7}$	$0.0^{+0.1}_{-0.0}$	$0.5^{+0.0}_{-0.0}$
WASP-75b	206154641	2.484193	$797.1^{+42.8}_{-41.5}$	$1.7^{+8.0}_{-0.1}$	$1183.269^{+0.001}_{-0.001}$	$0.098^{+0.07}_{-0.001}$	$0.9^{+0.1}_{-0.1}$	$75.3^{+56.5}_{-25.6}$	$0.0^{+0.3}_{-0.0}$	$0.0^{+0.4}_{-0.0}$
WASP-118b	220303276	4.0460407	$365.4^{+30.8}_{-23.5}$	$4.4^{+4.0}_{-0.2}$	$2590.045^{+0.000}_{-0.000}$	$0.082^{+0.0008}_{-0.0007}$	$0.3^{+0.1}_{-0.1}$	$92.4^{+23.6}_{-29.7}$	$0.0^{+0.2}_{-0.0}$	$0.5^{+0.0}_{-0.1}$

3.3 Validation Fits

In order to test the accuracy of our method, we applied it to known K2 multi-transit planets, with known periods. For a given planet, we treated each transit independently as a single transit, and we compared our estimate of the orbital period to the known one. Our validation sample of 27 planets consists of the confirmed systems from K2 that have available Gaia parallaxes and stellar atmospheric parameters (which allow us in turn to obtain precise stellar masses and radii using the method outlined in Section 3.2.1), as well as at least one transit with SNR > 7, the threshold at which we would expect a transit to be detectable individually. 21 of the validation planets were observed at long cadence (30 minutes), but we also considered a set of 6 planets observed at 1 minute cadence.

Among our validation sample were K2-19b and 19c, which underwent three overlapping transits during K2 Campaign 1. We remove these overlapping transits from our sample. We also remove one transit of WASP-47b, one transit of WASP-55b, and two transits of WASP-118b because they have in-transit light curve discontinuities due to stellar activity; the Gaussian process detrending was able to remove these discontinuities but caused the resulting de-trended transits to be unphysically much shallower than the other transits of the same planets.

The systems that we analysed are listed in Table 3.1, along with their assumed and determined parameters. We adopted the atmospheric parameters reported in the discovery publications as input for the estimation of the stellar radii and masses. In the table, we report both the known period (which is measured from the interval between successive transits, and is not used in our inference in any way) and the period we measure by fitting the planet's transits individually. (Note that the parameters reported in this table are summary statistics over multiple single-transit fits for each of these planets, and therefore are not representative of any of our individual transit fits. For posterior distributions of each individual single-transit fit, please contact the authors.)

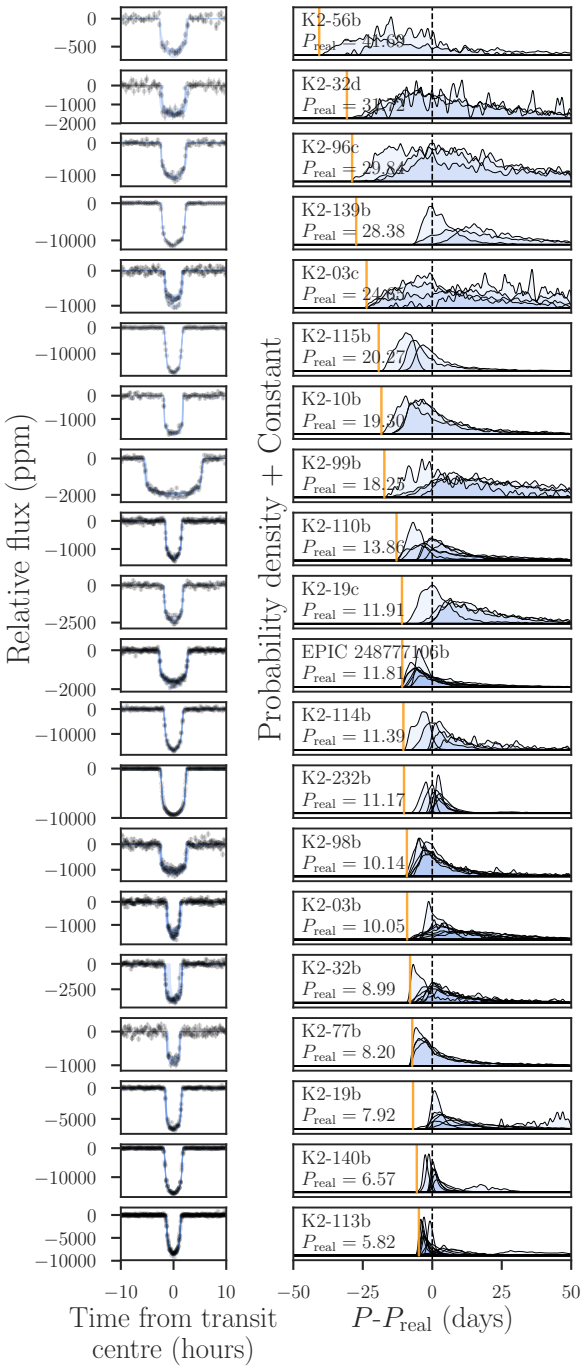

Figure 3.3: The posterior distributions for the 21 validation planets observed at long cadence. Planets are arranged in order of increasing period from bottom to top. Each histogram represents the posterior P distribution from the fit to each individual transit of the planet, with the planet's true period subtracted. The vertical yellow lines represent the lower prior bound on P, equal to 1.0 days. Transits with in-transit outliers and/or missing in-transit data points are susceptible to inaccurate measurement of P; see e.g. K2-32b and K2-140b, which both have transits with missing data points during ingress or egress. The posterior distributions appear to dip below the lower prior bound in some cases because of the smoothing we applied to the histograms in plotting.

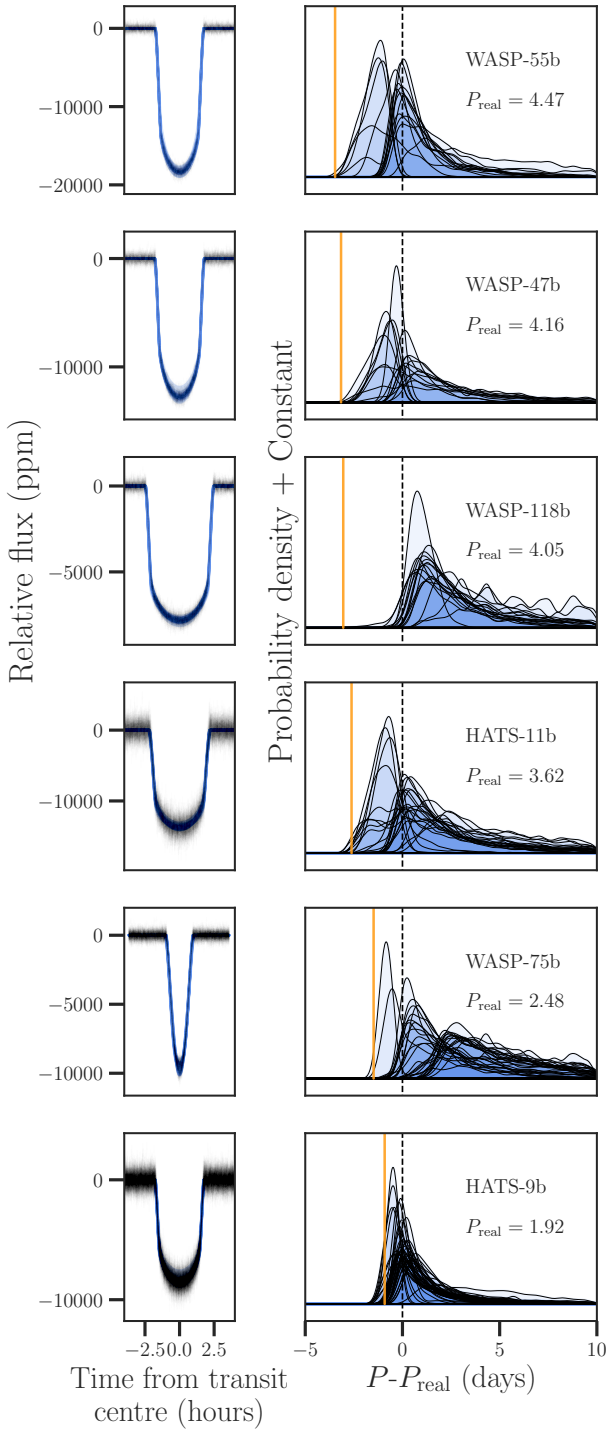

Figure 3.4: The posterior distributions for the 6 validation planets observed at short cadence. Planets are arranged in order of increasing period from bottom to top. Each histogram represents the posterior P distribution from the fit to each individual transit of the planet, with the planet's true period subtracted. The vertical yellow lines represent the lower prior bound on P, equal to 1.0 days. The posterior distributions appear to dip below the lower prior bound in some cases because of the smoothing we applied to the histograms in plotting.

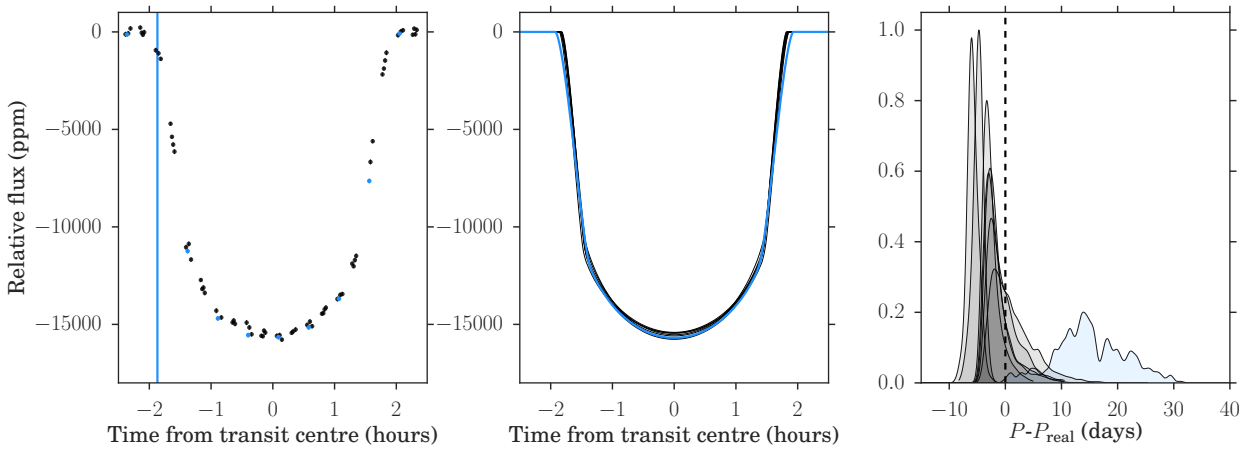

Figure 3.5: A demonstration of the extreme effect a slightly different transit shape measurement can have on the recovered P posterior, for validation planet K2-140b (P_{real} = 6.569 days). Eight transits are observed for this planet (left panel: transit data; middle panel: MultiNest-fit transit models; right panel: corresponding P posteriors). Seven have very similar transit shapes (black data points and transit models), corresponding to very similar P posteriors (black histograms). One transit, plotted in blue, has a missing data point at the position indicated by the vertical blue line in the left panel. As a result, MultiNest converges to a visibly wider transit shape, and the P posterior is offset by $\sim +12.5$ days from the true P.

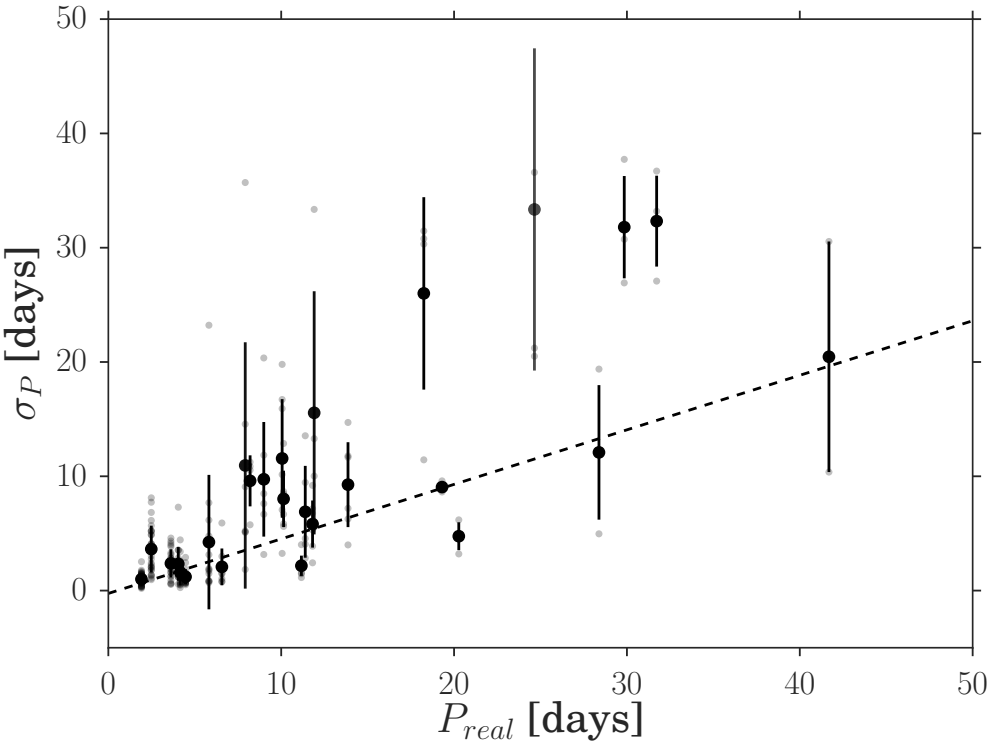

Figure 3.6: The posterior uncertainty on P as a function of P_{real}. Semi-transparent small points represent individual transits of a given planet; larger opaque black points with error bars summarise the results over all individual transits of each planet. The dotted line is the best weighted least-squares fit. We achieve smaller uncertainties for planets with smaller orbital periods.

Below, we discuss in turn the accuracy and precision of our period inferences for these planets.

3.3.1 Accuracy of Period Inferences

In Figures 3.3 and 3.4, we present the posterior P distributions from our single-transit fits to the long- and short-cadence validation planets, respectively, as well as plots of the fits themselves. Because of the physical requirement that $P > 0$ (i.e., $P - P_{\rm real} > -P_{\rm real}$), these posterior distributions have a characteristic asymmetric shape, skewed toward long P. To summarise these posteriors, we therefore choose to fit each one with a split-normal distribution and report the mean and left/right standard deviations (Penoyre, prep). (In Table 3.1, the given summary statistics are from split-normal fits to the concatenated posterior samples of all individual single transit fits for a given planet.)

Generally, these posterior P distributions are in good agreement with $P_{\rm real}$. This is particularly true for the six validation planets observed at short cadence, where the transit shape is much better constrained.

Occasionally, a particular transit will yield a P posterior that is significantly offset from P_{real}; this is never because of a poor fit to the available light curve data, as Figures 3.3 and 3.4 show. Rather, these are cases where our ability to accurately measure the transit shape is compromised by missing data, outliers, or both.

In Figure 3.5, we illustrate the surprisingly dramatic effect that even a small discrepancy in transit shape can have on the accuracy of the P posterior, using K2-140b as an example. This planet is observed to transit eight times during K2 Campaign 10. Of our fits to the eight transits individually, seven have very similar transit shapes and return very similar P posteriors, in reasonably good agreement with $P_{\rm real} = 6.569$ days.

The seventh of eight, however, is missing a data point during ingress, and the resulting MultiNest fit has a noticeably longer transit duration than the others. The resulting P posterior is centred at $\mu = 19.0^{+7.9}_{-4.0}$ days, more than $3\sigma_{\rm left}$ discrepant with $P_{\rm real}$. This is

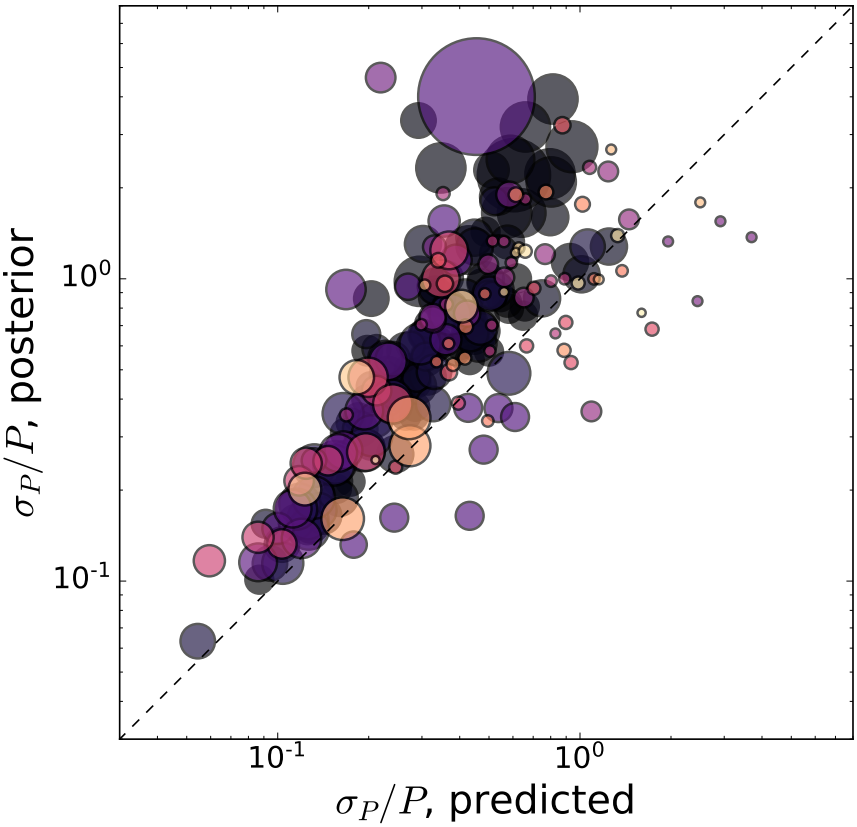

Figure 3.7: The posterior fractional P uncertainty for each single-transit fit of the validation planets, as a function of the fractional P uncertainty predicted by Equation (3.2). The black dotted line is a one-to-one line (the predicted relationship of Equation 3.2). Each data point represents one transit of one validation planet; points are colour-coded by $P_{\rm real}$ (dark for short $P_{\rm real}$ to light for large $P_{\rm real}$) and sized by the best-fit posterior R_p/R_*.

sensible because P scales as $(a/R_*)^{3/2}$, so an overestimate in a/R_* gets amplified in P.

We note also that missing data during ingress and egress can also cause MultiNest to converge to an erroneously short transit duration, and a correspondingly short period—this is true for two of K2-32b's seven observed transits.

We therefore advise caution in fitting single transits with missing data or obvious in-transit outliers, particularly during ingress and egress, as the fits for such transits are not reliable. The twelve true single transits discussed in section 3.4 are well-sampled during ingress and egress, so we have confidence in our ability to measure their transit shapes (and corresponding periods) accurately (see Figures 3.10 and 3.11, left panels).

3.3.2 Precision of Period Inferences

We also consider the precision of the period inferences we can make about our validation planets. In particular: are there specific characteristics of a single transit that enable a precise period measurement?

We note that we calculate the posterior period uncertainty, σ_P, by fitting a split-normal distribution to the P posterior, then taking the average of the resulting σ_{left} and σ_{right}. We make this choice because our posteriors are quite asymmetric, as explained in 3.3.1, and the traditional percentile summary statistics do not characterise them well. (We adopt the same procedure for calculating posterior uncertainties σ_{a/R_*} and σ_e, below.)

In Figure 3.6, we examine the trend between the posterior period uncertainty σ_P and the true period P_{real} for our 27 validation planets. These quantities are positively correlated: in other words, we derive smaller σ_P for smaller P_{real} and vice versa. (This trend is also visible in Figure 3.3, in the sense that the average width of the P posteriors increases with increasing P_{real}.)

Given that σ_P increases with P, we next investigate trends in the fractional period uncertainty, σ_P/P, with other quantities. First, in Figure 3.7, we compare the σ_P/P from our posterior P distributions to the σ_P/P predicted by Equation (3.2). Broadly, the data follow the expected relationship, but we find that Equation (3.2) tends to underestimate the observed fractional uncertainty, particularly when σ_P/P is high.

In Figure 3.8, we investigate σ_P/P more carefully, examining which sources of uncertainty contribute most strongly with it. In particular, we examine the correlation of σ_P/P with the fractional uncertainty of the input Gaia ρ_* measurement; with the fractional uncertainty of a/R_* and e from our posterior distributions; and with the estimated contribution to σ_P/P from K2's photometric uncertainty in observing each planet's transit. (This last quantity is calculated according to Equation 13 of Yee & Gaudi 2008.)

The top panel of Figure 3.8 plots the relationship between the fractional uncertainty in

ρ_*, which was a prior input to our MultiNest fits, and the resulting σ_P/P. Since σ_{ρ_*}/ρ_* is the same for all planets in a particular system, the planets in this panel bunch up in vertical lines. Surprisingly, there is no apparent correlation between the prior fractional stellar density uncertainty and the posterior fractional uncertainty in P: in other words, the stellar density uncertainty does not appear to be the dominant contribution to the ultimate P uncertainty. Our order-of-magnitude calculation for an Earth analog, following Equation (3.2), is clearly too simplistic.

In contrast, the second panel shows that σ_P/P is very strongly predicted by $\sigma_{a/R_*}/a/R_*$. a/R_* is measured directly from the transit shape; this panel shows that our ability to constrain the transit shape is the single most important predictor of how well we can constrain P for a single transiter.

Since we also expect our constraint on the linear limb darkening coefficient q to depend on how well we can constrain the transit shape, we plot in the third panel σ_P/P vs. σ_q/q, and find the expected positive correlation; in other words, transits with worse-constrained shapes have both worse-constrained limb darkening coefficients and worse-constrained periods.

Our ability to constrain e is less important, as illustrated in the fourth panel. There is a weak positive correlation between σ_P/P and e, indicating that more eccentric planets (larger e) have worse period constraints, and vice versa.

Finally, there is essentially no correlation between σ_P/P and the K2 photometric uncertainty over the planet transit. Visible in the colours and sizes of points on this plot, however, are the strong negative correlation between photometric uncertainty and planet size ($(\sigma_P/P)_{\text{photometric}} \propto R_p^{-5/2}$) and the weaker negative correlation between photometric uncertainty and orbital period ($(\sigma_P/P)_{\text{photometric}} \propto P^{-1/6}$).

In summary, a strong constraint on a/R_*, which is measured from the transit shape, is the best predictor of a strong posterior constraint on P. To improve the constraint on the transit shape, one could place stronger prior constraints on the limb-darkening profile of the star, for example by jointly fitting transits of multiple planets in the same system.

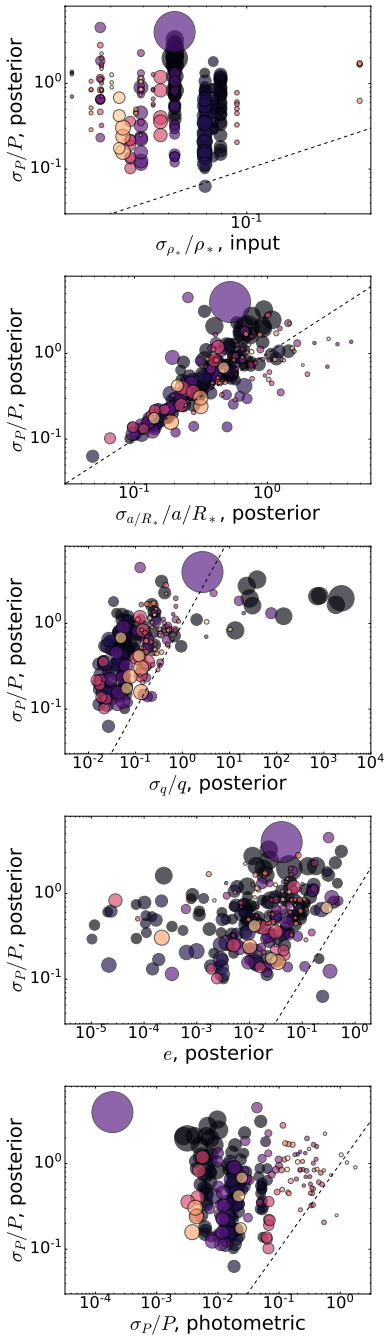

Figure 3.8: An exploration of which terms contribute most significantly to σ_P/P. Top: σ_P/P as a function of the fractional uncertainty on the Gaia-derived ρ_* measurement input to MultiNest. Second row: σ_P/P as a function of the fractional posterior uncertainty on normalised semi-major axis a/R_*. Third row: σ_P/P as a function of the fractional posterior uncertainty on linear limb darkening coefficient q. Fourth row: σ_P/P as a function of the posterior modal value of eccentricity e. Bottom: σ_P/P as a function of the K2 photometric uncertainty for each transit (calculated from the formula given by Yee & Gaudi 2008). Each point represents one transit of one validation planet; points are colour-coded by P_{real} (dark for short P_{real} to light for large P_{real} and sized by the best-fit posterior R_p/R_*. A one-to-one line is plotted in each panel.

Table 3.2: Summary statistics for fits to the true K2 single transits. To the left of the vertical line are known or assumed parameters; to the right are parameters fit with MultiNest. Because the t_0 posterior distributions are generally quite symmetrical, the reported fit t_0 is the 50^{th} percentile of the posterior distribution, with uncertainties given by the 16^{th} and 84^{th} percentiles. The posterior distributions of the other parameters are asymmetrical, so we fit each with a split-normal distribution and report $\mu^{+\sigma_{\text{right}}}_{-\sigma_{\text{left}}}$.

Name (EPIC)	Reference	P_{min} [d]	ρ_* [kg/m^3]	P_{inferred} [d]	t_0 [BJD-2454833]	R_p/R_*	b	ω [°]	e	q
EPIC 201635132b	Osborn16	63.3	$4992.3^{+222.5}_{-224.7}$	$125.9^{+42.6}_{-22.6}$	$1993.985^{+0.001}_{-0.001}$	$0.227^{+0.02}_{-0.02}$	$0.9^{+0.1}_{-0.0}$	$115.5^{+106.5}_{-58.6}$	$0.03^{+0.09}_{-0.02}$	$0.84^{+0.05}_{-0.32}$
EPIC 201892470b	Osborn16	78.7	$199.8^{+15.9}_{-7.5}$	$78.8^{+488.1}_{-0.1}$	$2056.159^{+0.020}_{-0.030}$	$0.026^{+0.0008}_{-0.009}$	$1.0^{+0.0}_{-0.3}$	$69.3^{+84.5}_{-29.5}$	$0.88^{+0.05}_{-0.35}$	$0.97^{+0.02}_{-0.42}$
EPIC 203311200b	Osborn16	59.7	$325.1^{+30.2}_{-27.2}$	$251.0^{+145.9}_{-33.9}$	$2121.016^{+0.001}_{-0.001}$	$0.060^{+0.0006}_{-0.0005}$	$0.8^{+0.0}_{-0.0}$	$88.6^{+30.4}_{-17.5}$	$0.03^{+0.14}_{-0.02}$	$0.44^{+0.09}_{-0.05}$
EPIC 204634789b	Osborn16	50.8	$848.0^{+48.3}_{-53.9}$	$64.5^{+110.4}_{-5.4}$	$2088.016^{+0.001}_{-0.001}$	$0.090^{+0.003}_{-0.003}$	$0.8^{+0.0}_{-0.1}$	$63.9^{+22.3}_{-6.4}$	$0.92^{+0.02}_{-0.03}$	$0.73^{+0.07}_{-0.25}$
EPIC 211311380d	Vanderburg16	48.2	$658.0^{+26.0}_{-26.0}$	$49.3^{+328.5}_{-0.1}$	$2333.270^{+0.001}_{-0.001}$	$0.026^{+0.001}_{-0.0008}$	$0.1^{+0.3}_{-0.1}$	$60.4^{+143.4}_{-29.8}$	$0.02^{+0.36}_{-0.01}$	$0.61^{+0.07}_{-0.14}$
EPIC 211311380e	Vanderburg16	72.4	$658.0^{+26.0}_{-26.0}$	$131.7^{+245.9}_{-12.4}$	$2309.016^{+0.001}_{-0.001}$	$0.036^{+0.001}_{-0.0005}$	$0.5^{+0.1}_{-0.2}$	$79.2^{+93.4}_{-25.8}$	$0.02^{+0.22}_{-0.01}$	$0.46^{+0.08}_{-0.02}$
EPIC 211311380f	Vanderburg16	47.3	$658.0^{+26.0}_{-26.0}$	$599.2^{+525.3}_{-159.5}$	$2353.915^{+0.001}_{-0.001}$	$0.066^{+0.001}_{-0.0005}$	$0.4^{+0.1}_{-0.2}$	$88.3^{+30.7}_{-36.6}$	$0.01^{+0.27}_{-0.01}$	$0.44^{+0.05}_{-0.04}$
EPIC 212813907b	Crossfield18	32.8	$2309.8^{+188.0}_{-164.8}$	$2014.9^{+1117.2}_{-331.6}$	$3380.823^{+0.001}_{-0.001}$	$0.119^{+0.001}_{-0.0003}$	$0.2^{+0.1}_{-0.1}$	$103.5^{+20.2}_{-3.7}$	$0.41^{+0.08}_{-0.12}$	$0.63^{+0.02}_{-0.02}$
EPIC 228801451d	Santerne18	25.0	$2702.2^{+84.8}_{-104.9}$	$25.0^{+49.8}_{-0.1}$	$2790.062^{+0.003}_{-0.003}$	$0.030^{+0.0003}_{-0.002}$	$0.5^{+0.2}_{-0.2}$	$85.0^{+37.3}_{-32.6}$	$0.69^{+0.11}_{-0.13}$	$0.18^{+0.25}_{-0.09}$
EPIC 248045685b	Vanderburg18	72.7	$3082.8^{+267.4}_{-186.7}$	$118.0^{+119.3}_{-41.4}$	$2995.387^{+0.001}_{-0.001}$	$0.027^{+0.0004}_{-0.0008}$	$0.5^{+0.2}_{-0.1}$	$95.2^{+30.3}_{-23.1}$	$0.44^{+0.11}_{-0.16}$	$0.26^{+0.17}_{-0.10}$
EPIC 248847494b	Giles18	59.8	$86.6^{+8.1}_{-7.3}$	$1658.2^{+4191.1}_{-445.8}$	$3134.174^{+0.020}_{-0.020}$	$0.044^{+0.002}_{-0.003}$	$0.9^{+0.0}_{-0.1}$	$308.2^{+19.5}_{-105.2}$	$0.08^{+0.29}_{-0.05}$	$0.75^{+0.08}_{-0.34}$
EPIC 246445793b	Vanderburg15	3.9	$2750.5^{+158.8}_{-252.8}$	$4.7^{+11.6}_{-0.2}$	$1865.133^{+0.002}_{-0.001}$	$0.030^{+0.002}_{-0.002}$	$0.2^{+0.3}_{-0.1}$	$97.1^{+37.6}_{-45.9}$	$0.10^{+0.26}_{-0.06}$	$0.86^{+0.09}_{-0.13}$

3.4 Period Predictions for K2 Single Transiters

We next proceed to apply our fitting code to twelve single transiters observed by K2. These single transiters were first reported by Osborn et al. (2016); Crossfield et al. (2018); Santerne et al. (2018); Vanderburg et al. (2018); Giles et al. (2018); Vanderburg et al. (2015); and Vanderburg et al. (2016), as detailed in Table 3.2.

In Figures 3.10 and 3.11, we present our transit fits and posterior distributions for these twelve single transiters. For all twelve, MultiNest converges to a good fit to the transit, with a stellar density posterior distribution in complete agreement with the Gaia prior. (In other words, there were no cases for which MultiNest needed to wander far from the input Gaia stellar density to fit the transit data.)

Over the twelve single transits, our posterior fractional uncertainty σ_P/P is $94^{+87}_{-58}\%$, which is comparable to that achieved in previous work. We emphasise that we treat eccentricity e as a free parameter in these fits.

For a more direct comparison to previous work, we re-fit the single transits with eccentricity e fixed to zero, and find that we achieve posterior period fractional uncertainty σ_P/P of $15^{+30}_{-6}\%$, a roughly threefold improvement over typical uncertainties of previous studies.

3.4.1 A note on priors

As discussed in Section 3.2.2, the log-uniform prior adopted for our validation fits does not formally account for the fact that a single-transiter is only observed to transit once, and as such does not exploit all of the available information about a single transit. The prior defined by K18 does account for this fact, and it strongly enforces the expectation that, if we observe a single transit of a planet over a given baseline, the planet is a-priori less likely to have a long period than a short one ($\Pr(P) \propto P^{\alpha-5/3}$ between $P_{\rm min}$ and $P_{\rm max}$, where we choose $\alpha = -2/3$, the power-law index of the underlying intrinsic period prior, for uniformity

[1]This single transit was observed during a 9-day test of K2 in February 2014, hence its short $P_{\rm min}$.

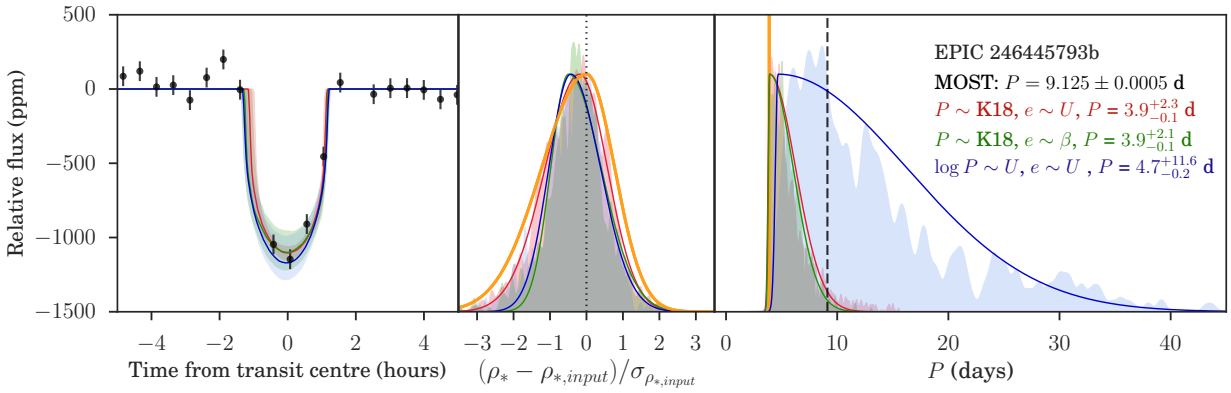

Figure 3.9: A comparison of the period posteriors derived for single transiter EPIC 246445793b with three different choices of prior: in red, a K18 period prior with $\alpha = -2/3$ and a uniform eccentricity prior between 0 and 1; in green, a K18 period prior with $\alpha = -2/3$ and a Beta distribution eccentricity prior with $a = 0.867, b = 3.03$ (adopted from Kipping 2013); and in blue, a log-uniform period prior between 1 and 10000 days with a uniform eccentricity prior between 0 and 1. Left panel: The best-fit transit model given each prior. Middle panel: The stellar density posterior given each prior. Right panel: The period posterior given each prior. The dashed black line marks the MOST photometry-measured period for this planet (Vanderburg et al., 2015).

in semi-major axis).

We test both period priors on our twelve true single transiters. For both, we set $P_{\rm min}$ according to Equation 3.2.2 and set $P_{\rm max} = 10000$ days. With the log-uniform period prior, we keep the eccentricity prior uniform between 0 and 1. With the K18 prior, we test two eccentricity distributions: (1) e uniform between 0 and 1 and (2) e Beta-distributed, with parameters $a = 0.867, b = 3.03$ adopted from Kipping (2013).

The results of these three sets of P and e prior choices for single transiter EPIC 246445793b (HIP 116454b) are plotted in Figure 3.9. The K18 prior, as expected, yields a P posterior which drops off much more steeply with increasing period than the log-uniform P prior, regardless of the choice of e prior.

In the case of EPIC 246445793b, the resulting K18 posteriors drop off so sharply with P as to be 2σ inconsistent with the period measured from MOST photometry of 9.125 ± 0.0005 days (Vanderburg et al., 2015). We find that the K18 period posteriors for EPIC 211311380f

(HIP 41378f) are similarly inconsistent with previously published period constraints (Becker et al., 2019).

These inconsistencies suggest that the K2 long-cadence transit data are insufficient to constrain our 12-parameter transit model in the face of such a strong period prior. It is possible that a lower-dimensional model—e.g., one that relied on a less flexible detrending algorithm than Gaussian process regression—would be less easily overwhelmed by a strong prior. It is also possible that the choice of intrinsic period prior power-law index $\alpha = -2/3$ is too steep to describe the long-period exoplanet population, but determining the intrinsic period prior for long-period exoplanets is well beyond the scope of this work.

As a result of these considerations and its success in the validation fits, we choose to fit the twelve single-transiters in our sample with a log-uniform prior in P between $P_{\rm min}$ (determined by Equation 3.2.2) and $P_{\rm max} = 10000$ days.

3.4.2 Comparison to previous work

For some of the twelve single transiters, period constraints have been published before. Vanderburg et al. (2015), for example, obtained MOST photometry and radial velocity measurements of EPIC 246445793b and measured its period at $P = 9.1205 \pm 0.0005$ days. Our measurement, though obviously considerably less precise, agrees well with this one.

Santerne et al. (2018) analyze the K2 photometry of the EPIC 228801451 system and conclude that there are two possible orbital solutions for EPIC 228801451d: one with $P = 31.0 \pm 1.1$ days, and one with $P > 50$ days. Our period measurement, $P = 25.0^{+49.8}_{-0.1}$ days, is compatible with both.

Vanderburg et al. (2018) estimate the period of EPIC 248045685b by a very similar method to ours, also exploiting information from the host star's Gaia parallax and public broadband photometry. They conclude that $P = 106^{+74}_{-26}$ days, and we agree, with $P = 118^{+119}_{-41}$ days.

Giles et al. (2018) use high-resolution spectra combined with Gaia parallax to measure the stellar parameters of EPIC 248847494, and then model the transit of EPIC 248847494b to conclude that $P = 3650^{+1280}_{-1130}$ days, which agrees with our $P = 1700^{+4200}_{-400}$ days.

Crossfield et al. (2018) estimate that EPIC 212813907b has P of order 1000 days, while we estimate that it has $P = 2000^{+1100}_{-300}$ days.

Vanderburg et al. (2016), who discovered the EPIC 211311380 (HIP 41378) system in K2 Campaign 5 observations, estimate from fits to the respective single transits that planet d has period $P = 156^{+163}_{-78}$ days, planet e has period $P = 131^{+61}_{-36}$ days, and planet f has period $P = 324^{+121}_{-127}$ days. All three are consistent with our respective estimates of $49.3^{+328.5}_{-0.1}$ days, 131^{+250}_{-12} days, and 600^{+530}_{-160} days. Becker et al. (2019) observed subsequent transits of planets d and f in K2 Campaign 18 and refine the possible periods of all three planets based on the additional observations and simulations of dynamical stability. Their maximum-probability period for planet d is consistent with our 1σ credibility band, but their maximum-probability period for planet f (361 days) is shorter than we estimate. Our modal period for planet e is comfortably allowed by their dynamical simulations.

Finally, Osborn et al. (2016) fit their single transiters by an analogous method to ours; they derive effective temperatures for the planet host stars using broad-band photometry, then estimate stellar masses and radii from these temperatures using stellar models, assuming they are on the main sequence. Four single transits in our sample—EPIC 201635132b, EPIC 201892470b, EPIC 203311200b, and EPIC 204634789b—are drawn from their work.

However, for these four single transits, we do not agree with period determinations of Osborn et al. (2016). This is because our stellar properties do not agree; in particular, the bulk densities we derive from Gaia parallaxes plus photometry do not agree with the bulk stellar densities calculable from their stellar mass and radius estimates. Only in the case of EPIC 203311200b are the period measurements even in 1σ agreement; for the other three, they are very different. Obtaining high-resolution spectra for these host stars would enable a precise cross-check on their stellar parameters and hopefully clear up this disagreement.

(In the case of EPIC 204634789b, our $P_{\rm min}$ prior excludes their best-fit P, so consistent prior choices are also important.)

Finally, we note that four of our twelve single transit fits—EPIC 201892470b, EPIC 204634789b, EPIC 228801451d (K2-229d), and EPIC 248045685b—have high modal posterior eccentricity, albeit with large uncertainty (see Table 3.2). We are unable to achieve good fits to these transits with e fixed to zero. These candidates, particularly EPIC 228801451d, which has two inner planetary companions, merit further study to determine their orbital properties more precisely.

3.5 Conclusions

In this work, we have presented new transit fits to twelve K2 single transiters, based on stellar density priors derived from GAIA parallaxes and publicly available broadband photometry. We achieve good precision in our period posteriors—when we let e vary, the fractional P uncertainty over the twelve single transiters is $94^{+87}_{-58}\%$, and when we fix $e = 0$, it is $15^{+30}_{-6}\%$ (a roughly threefold improvement over typical period uncertainties of previous studies). In future, the best way to handle the question of eccentricity is likely to perform `single` transit fits with both fixed and free e, then use Bayesian model averaging to combine them and obtain a posterior P estimate, but we leave this for future work.

Our fit period values also agree well with previously published period constraints. (Where we do not agree with earlier period constraints, it is because our stellar properties disagree with those used in previous work, e.g. Osborn et al. 2016; further study of these host stars, ideally with high-resolution spectroscopy, will be necessary to resolve the disagreement).

Additionally, we test this fitting method on 27 validation planets observed by K2. These planets have been observed to transit more than once, so their periods are known precisely; however, we model each transit individually, to evaluate the accuracy and precision of our derived P posteriors given this limited information. We conclude that our method is robust

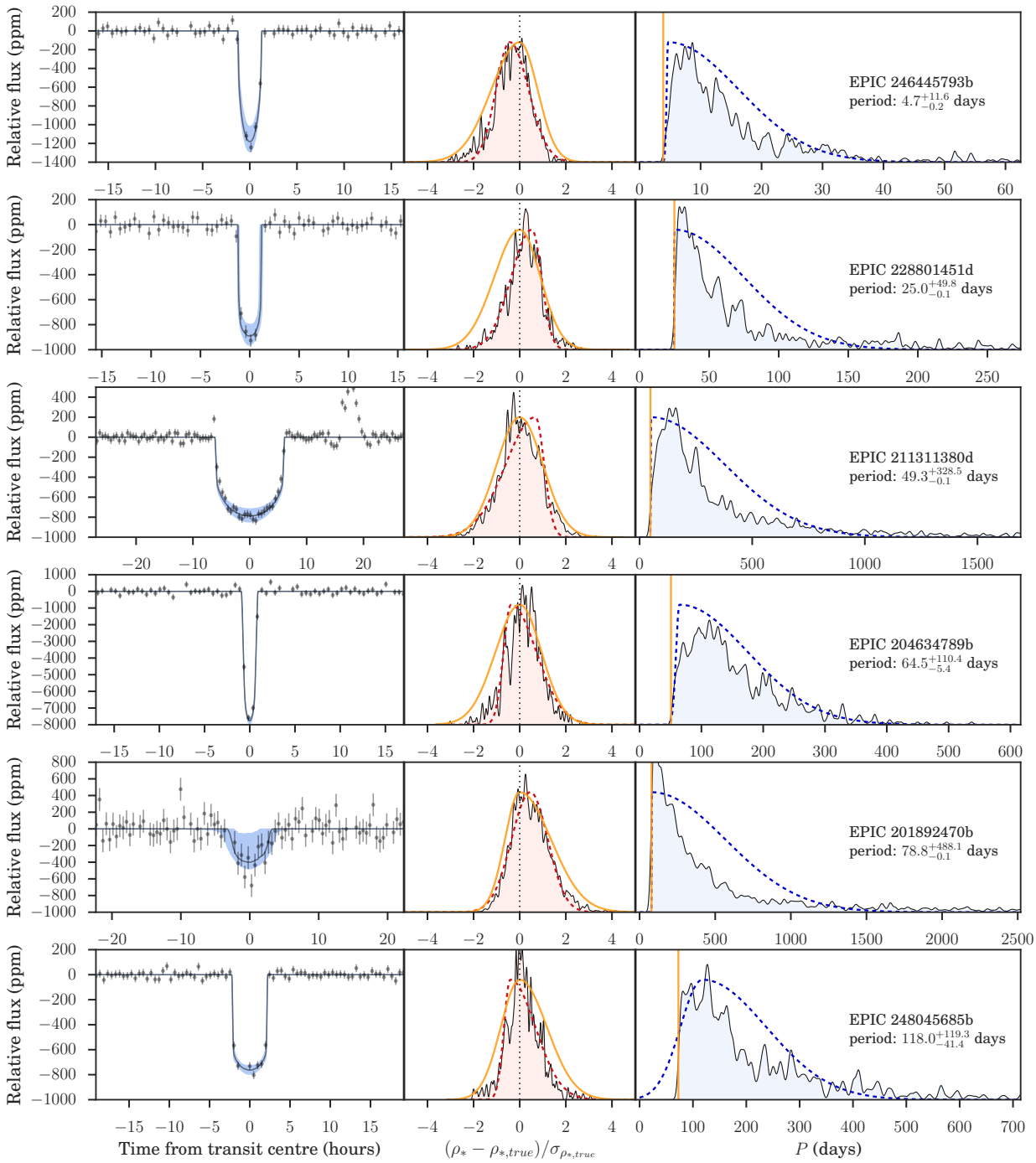

Figure 3.10: The results of fits to the twelve true K2 single transits, compiled from sources detailed in Table 3.2. Each row represents one planet. Left panel: The observed transit (black data points), best-fit model (solid black line), and 1-sigma credibility band given our posterior parameters (blue band). Middle panel: The Gaussian stellar density prior (yellow line) derived from Gaia distances plus available photometry for each host, and the corresponding posterior ρ_* distribution (red histogram). Right panel: The posterior P distribution, bounded on the left by the K2 baseline-deduced $P_{\rm min}$ (yellow line). We fit each P posterior with a split normal distribution (blue dotted line) to allow us to write down summary statistics (right panel text).

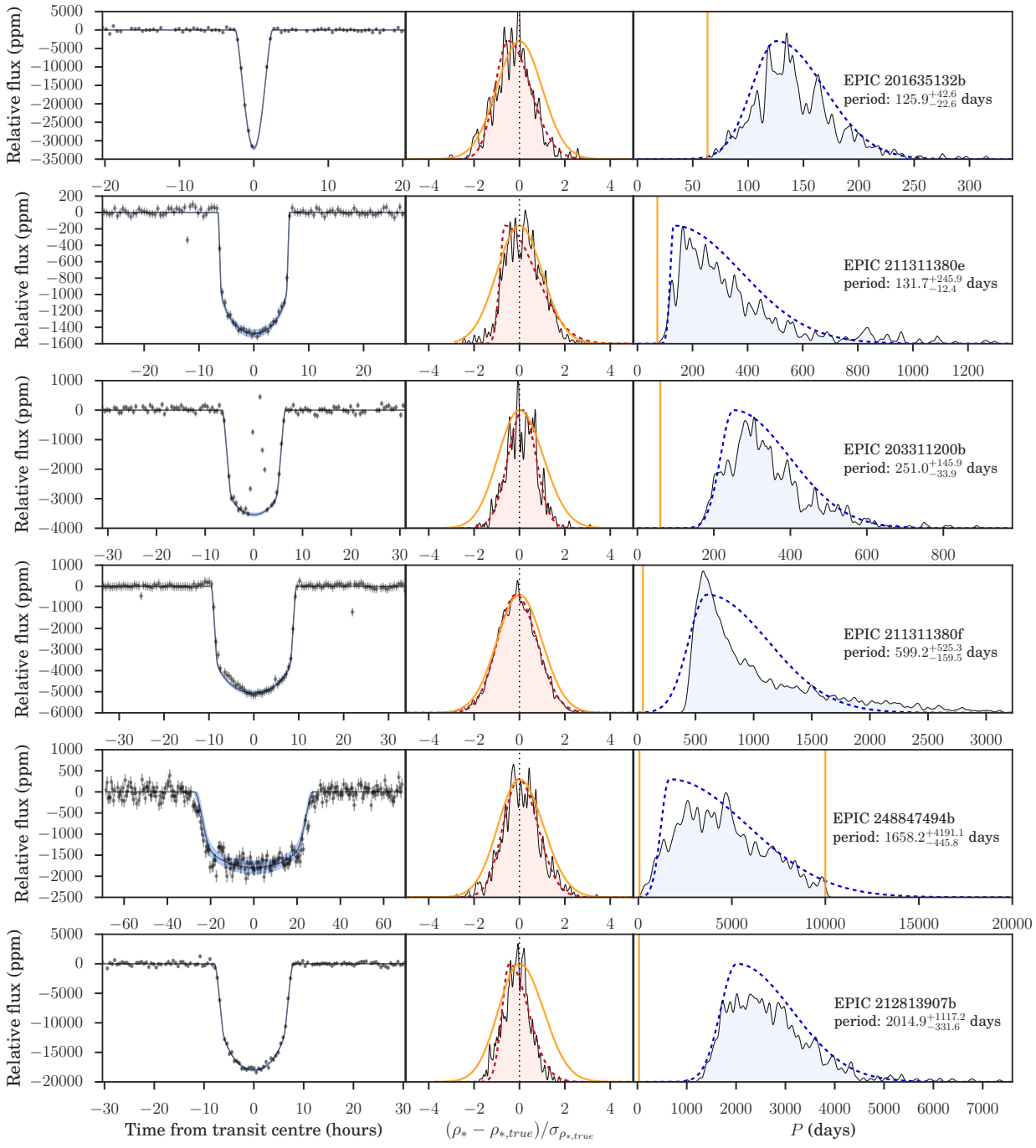

Figure 3.11: A continuation of Figure 3.10. The P posterior for the second-to-last planet, EPIC 248847494b, abuts the upper end of the P prior, $P_{\max} = 10000$ days.

as long as the individual transits we fit are well-sampled during ingress and egress, because our ability to measure a/R_* from the transit shape is the most important predictive factor in the success of the method, both in terms of accuracy and precision.

TESS, with its relatively short 27.4 day observational baseline over much of the sky, is predicted to reveal tens to hundreds of new single transiters, and it will be necessary to estimate the periods of these planets to constrain their orbits, temperatures, surface and atmospheric properties, and potential habitability. Combining information from multiple surveys, including Gaia, promises to aid greatly in their characterisation.

Acknowledgements

E.S. thanks David Kipping, Marcel Agüeros, and the Columbia University President's Global Innovation Fund for their support, and Zephyr Penoyre for useful discussions and sharing his split normal fitting code. N.E. would like to thank the Gruber Foundation for its generous support. R.B. acknowledges support from FONDECYT Post-doctoral Fellowship Project No. 3180246., and from the Millennium Institute of Astrophysics (MAS). A.J. acknowledges support from FONDECYT project 1171208 and by the Ministry for the Economy, Development, and Tourism's Programa Iniciativa Científica Milenio through grant IC 120009, awarded to the Millennium Institute of Astrophysics (MAS).

4. Shadow imaging of transiting objects

4.1 Introduction

Transit light curves are rich in information. If we assume a physical model for a transiting object—usually, a spherical body in a Keplerian orbit about a host star—we may then infer the parameters of this model, including physical properties of the transiter, its orbit, and the host star, from the light curve.

However, anomalous transit-like events, such as those observed in star KIC 8462852 (Boyajian et al., 2016), resist this type of analysis, because their physical cause, and consequently the appropriate model, is not apparent. In this paper, we consider the general problem of inferring the transiting shape, or shadow image, that generated a particular light curve. We wish to infer this image from the light curve alone, with as few additional assumptions as possible.

A number of problems related to shadow imaging have been studied before. The inverse problem, of how to calculate the light curve of an arbitrary transiting shape, has been tangentially addressed by several numerical transit-light-curve-calculating codes. Generally, however, these assume some parametric model for the transiting object—in the case of `BATMAN` (Kreidberg, 2015), the transiting object must be a spherical planet; `LUNA` (Kipping, 2011), a spherical planet accompanied by a spherical moon; `PyTranSpot` (Juvan et al., 2018), circular starspots projected on a stellar surface; and the Universal Transit Simulator (Deeg,

2009), a planet with moons or rings.

Meanwhile, significant advances have been made in another problem closely related to shadow imaging: eclipse mapping, which attempts to reconstruct the two-dimensional surface features of an exoplanet undergoing secondary eclipse from the light reflected off the planet's surface as it disappears and reappears from behind the star. Majeau et al. (2012) and de Wit et al. (2012) were the first to demonstrate this method, on hot Jupiter HD189733b. Kawahara & Fujii (2011) extended this theory to surface mapping of exoplanets in face-on orbits using scattered light, and Farr et al. (2018) recently released the `exocartographer` code to carry out surface mapping in a fully Bayesian framework with robust uncertainty estimation. Berdyugina & Kuhn (2019) showed that next-generation coronagraphic telescopes will be able, using these techniques, to map the surface of a handful of nearby planets, including Proxima b.

Analogous two-dimensional mapping methods have been successfully applied to the problem of starspot inversion, or deducing the pattern of starspots responsible for time variations in the spectrum or light curve of a star. Goncharskii et al. (1982) were among the first to attempt starspot inversion, aiming to explain spectral variations in Ap stars by inferring the pattern of chemical inhomogeneities on the surface that would generate them. Vogt & Penrod (1983) introduced Doppler imaging to infer maps of starspots on rapidly rotating stars from time series spectra, and Vogt et al. (1987) refined the technique by introducing maximum entropy regularization as a means of choosing from a set of degenerate solutions to the same observations. Piskunov et al. (1990) compared the maximum entropy method, which prefers a solution with the minimum spatial correlation between points on the star's surface, to an alternative constraint, Tikhonov regularization, which prefers the smoothest possible pattern of starspots that matches the observations. Similar techniques, with varying choices of regularization, have been applied to stellar light curves by e.g. Lanza et al. (1998).

In this work, we build upon these techniques to develop a mathematical and numerical treatment of shadow imaging, which has a similar geometric setup to and is subject to

similar degeneracies as eclipse mapping and starspot inversion. In Section 4.2, we investigate, analytically, the degeneracies inherent to the light curve imaging problem. We explain how discretizing the problem—modeling the transiting object as a grid of pixels of fixed opacity, rather than as a smooth, continuous image—allows us to make progress on the problem despite these degeneracies. In Section 4.3, we define the pixel-grid model which can be used to represent any transiting object and explain how to calculate its light curve. In Section 4.4, we consider how, starting from a transit light curve, we may infer the pixel grid image which generated it, and we discuss the results of this inference on a number of test cases. In Section 4.5, we consider the results of light curve inversion on the real cases of the TRAPPIST-1c,e,f triple transit and the anomalous transits observed in Boyajian's Star. We conclude in Section 4.6.

4.2 Transit Degeneracies

Calculating the light curve of a transiting object is an act of projection. It begins with a three-dimensional object in space, projected against the sky to make a two-dimensional image. At a few discrete points in time, as this image crosses a star, the starlight that the image does not block is summed up, and the sums strung together to make a light curve: a one-dimensional time series.

Deducing the image that generated a particular light curve, therefore, is a problem of inferring two-dimensional data from one-dimensional. As such, we do not expect to find a unique solution to match each light curve. Vogt et al. (1987); Piskunov et al. (1990); Majeau et al. (2012), and de Wit et al. (2012) note similar degeneracies in starspot inversion and eclipse mapping, respectively.

We begin by examining mathematically the degeneracies inherent to the problem of inferring the shape that generated a particular light curve. We operate under the assumptions, discussed further in 4.3, that the occulting shape is unchanging in time and moving at a

constant velocity across the star; that the star is spherical and of uniform brightness; and that the observed light curve is well sampled in time.

4.2.1 The Flip Degeneracy

The first important degeneracy in the shadow imaging problem results from the reflection symmetry of the star about its horizontal midplane. An opaque shape that transits at an impact parameter b above the midplane produces the same light curve as a "flipped" shape that transits below the midplane.

In planetary transit modeling, this degeneracy can be ignored, because the sign of the impact parameter $b = \cos i \left(\frac{a}{R_*}\right) \left(\frac{1-e^2}{1+e\sin\omega}\right)$ is a function of the inclination angle i of the planet's orbital plane and does not describe any inherent property of the transiting planet. However, if we wish to model more general transiting shapes, we must consider the full space of flip-degenerate solutions.

To express the degree of flip degeneracy in a given shadow imaging problem, we consider an image made up by a grid of opaque and transparent pixels, N rows by M columns. (See Figure 4.1 for an example.) Although there are 2^{NM} unique permutations of opaque and transparent pixels arranged in this grid, each of these permutations does not yield a unique light curve, and in general, a light curve cannot be inverted to produce a unique pixel grid shadow image.

We can express the degree of degeneracy by calculating the number of unique light curves, U_{LC}, possible for this N-row by M-column grid,

$$U_{LC} = \begin{cases} \left(2 \times 3^{\frac{(N-1)}{2}}\right)^M, & N \text{ odd} \\ \left(3^{\frac{N}{2}}\right)^M, & N \text{ even.} \end{cases} \tag{4.1}$$

For intuition, consider first the even-N case. In each of the M columns, there are $\frac{N}{2}$ pixels above the midplane; each of these has a counterpart below the midplane with the

same impact parameter. There are four possible opacity states of this pair of pixels: 00, 01, 10, or 11. However, because of the flip degeneracy, the 01 and 10 cases produce the same light curve, so only three arrangements are unique. Hence, three unique opacity values for each degenerate pixel pair, raised to the power of the total number of pixels above the midplane.

The arithmetic in the odd-N case is the same, except that the middle pixel row has no across-midplane counterpart. Each pixel in that row may only take on opacity 0 or 1.

In the case of a square, 3×3 pixel grid, then, there are $2^9 = 512$ unique permutations of opaque and transparent pixels, but only 216 unique light curves.

As a result, the binary-opacity pixel grid solution to any given light curve inversion is not unique, *unless* the light curve was, in truth, generated by a grid of binary-opacity pixels ($\tau = 0$ or 1) that is symmetrical about its horizontal mid-plane. Physically, we would only expect such a situation for the case of a perfectly spherical planet, or perhaps a planet-moon system or ringed planet, transiting a star at an impact parameter of 0.

In general, therefore, the inverted pixel grid which generates a light curve is not unique. Figure 4.1 shows four transiting pixel images which generate identical light curves. Starting with the pixel image at the top and flipping any pixel about the horizontal mid-plane leaves the light curve unchanged. The pixel image in the bottom panel is the average of the full set of flip-degenerate solutions.

We hope, therefore, to recover shadow images analogous to this bottom panel, which represent a kind of "superposition" of the full set of flip-degenerate solutions to a particular light curve.

4.2.2 The Arc Degeneracy

There is, however, another degeneracy inherent to the shadow imaging problem by which the set of physically allowable images matching any particular light curve becomes infinitely large. This degeneracy allows a transiting pair of semicircular arcs to generate the same

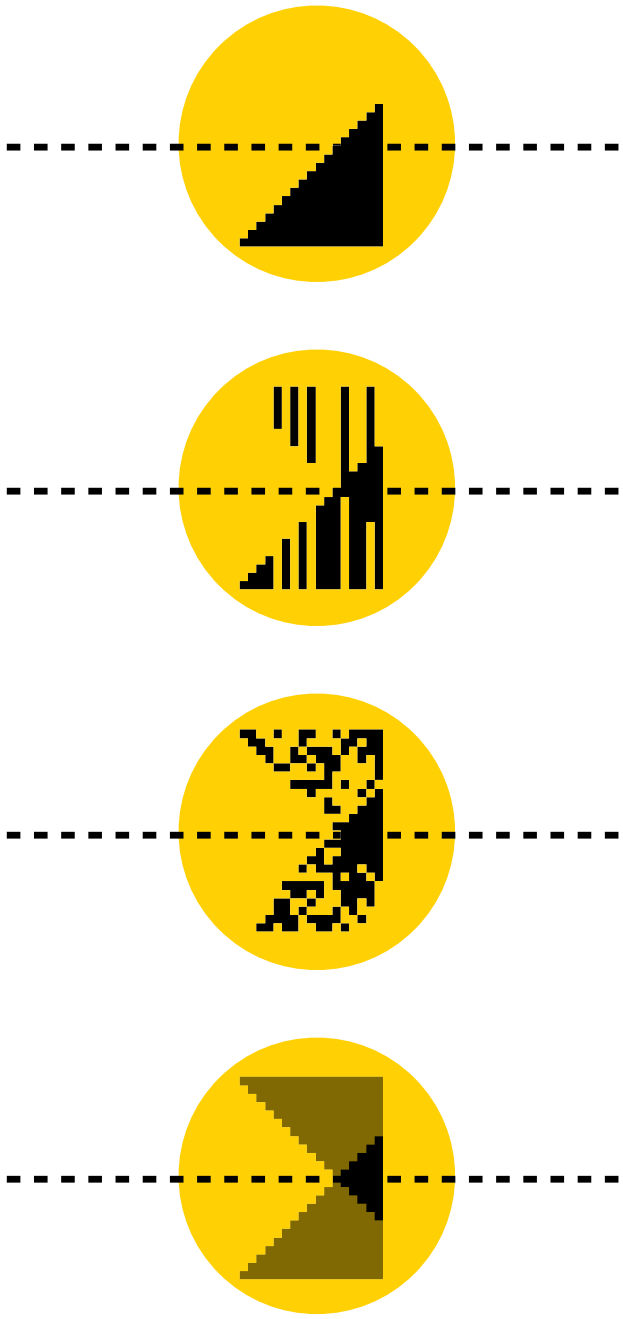

Figure 4.1: Four transiting binary-opacity pixel images which generate the same light curve. The bottom pixel image (opaque black pixels have $\tau = 1$; semi-transparent gray pixels have $\tau = 0.5$) is the average of the full set of flip-degenerate solutions which match this image's light curve.

light curve as a single opaque point, and we term it the "arc" degeneracy.

Figure 4.2 illustrates the geometry of the pair of arcs which generates the same light curve as an infinitesimally small opaque point transiting exactly along the horizontal midplane of the star. Consider this shape to transit from left to right across the star: because the right-hand arc traces the shape of the stellar limb, the entire right-hand arc will ingress upon the star at the same moment, yielding the same vertical ingress feature in the light curve that we would expect from an infinitesimally small transiting planet. (A correspondingly sharp egress feature in the light curve happens when the left-hand arc egresses all at once some time later.)

After the moment of ingress, the top- and bottom-most edges of the right-hand arc immediately egress again. However, this egress is balanced by the ingress of the middle of the left-hand arc. If opacity is appropriately distributed along each arc, then the ingress of the left-hand arc and egress of the right-hand arc may balance exactly. Here, we derive the functional form of the opacity distribution along the arc to allow this exact balance.

Let α (see Figure 4.2) denote the angle between the horizontal midplane of the star and the point of intersection between the stellar limb and the right-hand arc (which ingresses first). At the moment of ingress, $\alpha = \frac{\pi}{2}$; at the moment of egress, $\alpha = 0$. Let β denote the corresponding angle to the point of intersection on the left-hand arc, and let β range from 0 at ingress to $\frac{\pi}{2}$ at egress.

Let θ represent an angle measured from the horizontal midplane of either arc to its outermost point, and let $\lambda(\theta)$ represent the opacity along the arc as a function of this angle. Figure 4.3 illustrates this setup. Note that $\lambda(\theta)$ cannot be constant, because, for example, during some small time interval dt immediately after the moment of ingress, the length of the right-hand arc which egresses is greater than the length of left-hand arc which ingresses.

Let T be the duration of the transit of the pair of arcs (in other words, the interval between the moment of ingress and the moment of egress). Let the moment of ingress happen at $t = 0$, and let us define a dimensionless time coordinate $\kappa = \frac{t}{T}$ to parametrize the

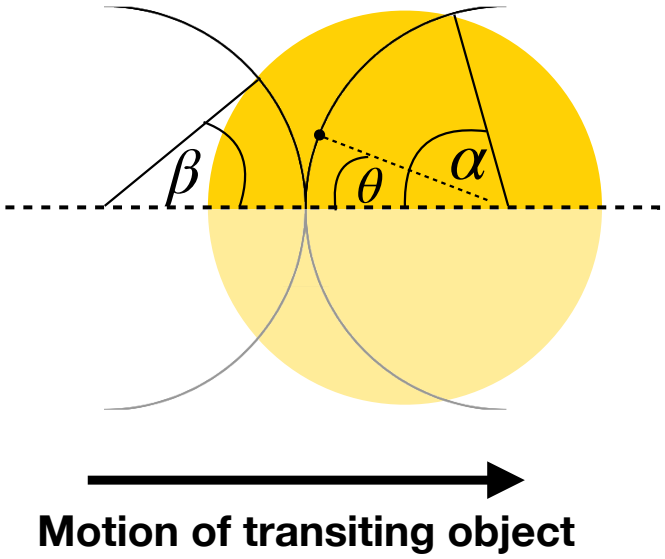

Figure 4.2: A pair of arcs which generates the same light curve as a single opaque point transiting along the horizontal midplane of the star. For this shape to generate a perfect box-like transit, the arcs must be infinitely thin and cannot be of uniform opacity; rather, opacity must be distributed symmetrically along them as a function of θ.

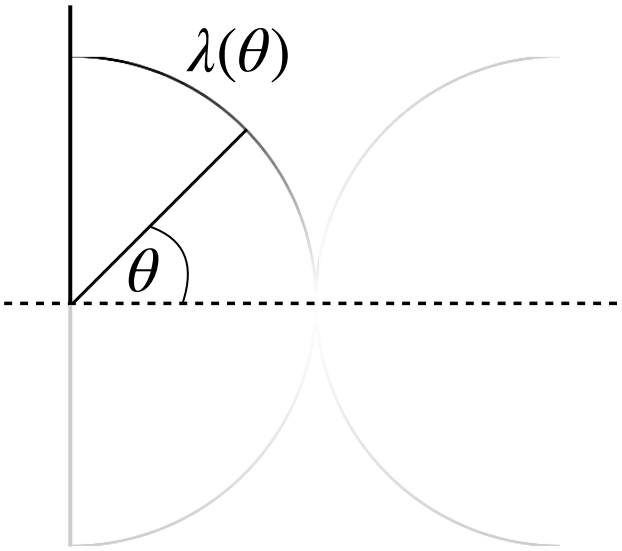

Figure 4.3: θ represents the angle from the horizontal midplane of either arc to any point along it. $\lambda(\theta)$ represents the opacity of the arc at θ. We wish to solve for $\lambda(\theta)$ such that the arc pair can produce a flat-bottomed transit.

progress of the transit. At ingress, then, $\kappa = 0$, and at egress, $\kappa = 1$.

Following these definitions, we may write

$$\cos\alpha = \kappa, \tag{4.2}$$

$$\cos\beta = 1 - \kappa. \tag{4.3}$$

The total opacity $L(\kappa)$ transiting the star at a particular moment κ is equal to

$$L(\kappa) = \int_0^{\alpha(\kappa)} \lambda(\theta)d\theta + \int_0^{\beta(\kappa)} \lambda(\theta)d\theta. \tag{4.4}$$

For the transit to be flat-bottomed, we require that $L(\kappa)$ be constant, or that $\frac{dL}{d\kappa} = 0$. Differentiating both sides of Equation (4.4) by κ, we obtain

$$\frac{dL}{d\kappa} = \lambda(\alpha)\frac{d\alpha}{d\kappa} + \lambda(\beta)\frac{d\beta}{d\kappa}, \tag{4.5}$$

because λ is time-independent and therefore independent of κ.

Setting this expression equal to 0 and substituting, we obtain

$$0 = \lambda(\alpha)\frac{1}{\sqrt{1-\kappa^2}} + \lambda(\beta)\frac{1}{\sqrt{1-(1-\kappa)^2}}, \tag{4.6}$$

or

$$\frac{\lambda(\alpha)}{\lambda(\beta)} = -\frac{\sqrt{1-\kappa^2}}{\sqrt{1-(1-\kappa)^2}}. \tag{4.7}$$

By the definitions of α and β, we may write

$$\frac{\lambda(\alpha)}{\lambda(\beta)} = -\frac{\sqrt{1-\cos^2\alpha}}{\sqrt{1-\cos^2\beta}} = \frac{\sin\alpha}{\sin\beta}. \tag{4.8}$$

By inspection, then,

$$\lambda(\theta) \propto \sin\theta, \tag{4.9}$$

where we choose the sign to be positive because physically meaningful opacities are between 0 and 1.

The overall normalization of $\lambda(\theta)$ sets the transit depth of the arcs' light curve. Figure 4.4 shows a transiting arc pair with $\lambda(\theta) = \sin\theta$.

We note that there are two other solutions to $\lambda(\theta)$ that satisfy the condition that $\frac{dL}{d\kappa} = 0$. The first is the trivial solution, $\lambda(\theta) = 0$. The second is a Dirac delta function at $\theta = 0$,

$$\lambda(\theta) \propto \delta(\theta = 0), \tag{4.10}$$

where again the overall normalization sets the transit depth.

For intuition, the two non-trivial solutions to $\lambda(\theta)$ given by Equations (4.9) and (4.10) represent two extremes: the least and most compact arrangements of opacity, respectively, that produce the same flat-bottomed, box-like transit. Any linear combination of these solutions also satisfies $\frac{dL}{d\kappa} = 0$ and generates a box-like transit.

The above derivation has demonstrated that a pair of arcs of variable opacity can match the transit shape of an infinitesimal point of opacity transiting along the horizontal midplane of the star, at impact parameter $b = 0$. The same logic applies to an infinitesimal point at arbitrary impact parameter b. Figure 4.5 illustrates the geometry of this situation.

Mathematically, a change in the impact parameter b means that the limits of integration in Equation (4.4) change,

$$L(\kappa) = \int_{\arcsin b}^{\alpha(\kappa)} \lambda(\theta) d\theta + \int_{\arcsin b}^{\beta(\kappa)} \lambda(\theta) d\theta. \tag{4.11}$$

Since b is constant, the subsequent steps and resulting solutions for $\lambda(\theta)$ do not change, except that the delta function solution is localized at $\theta = \arcsin b$.

We note finally that the arc degeneracy technically only operates for an occulter transiting

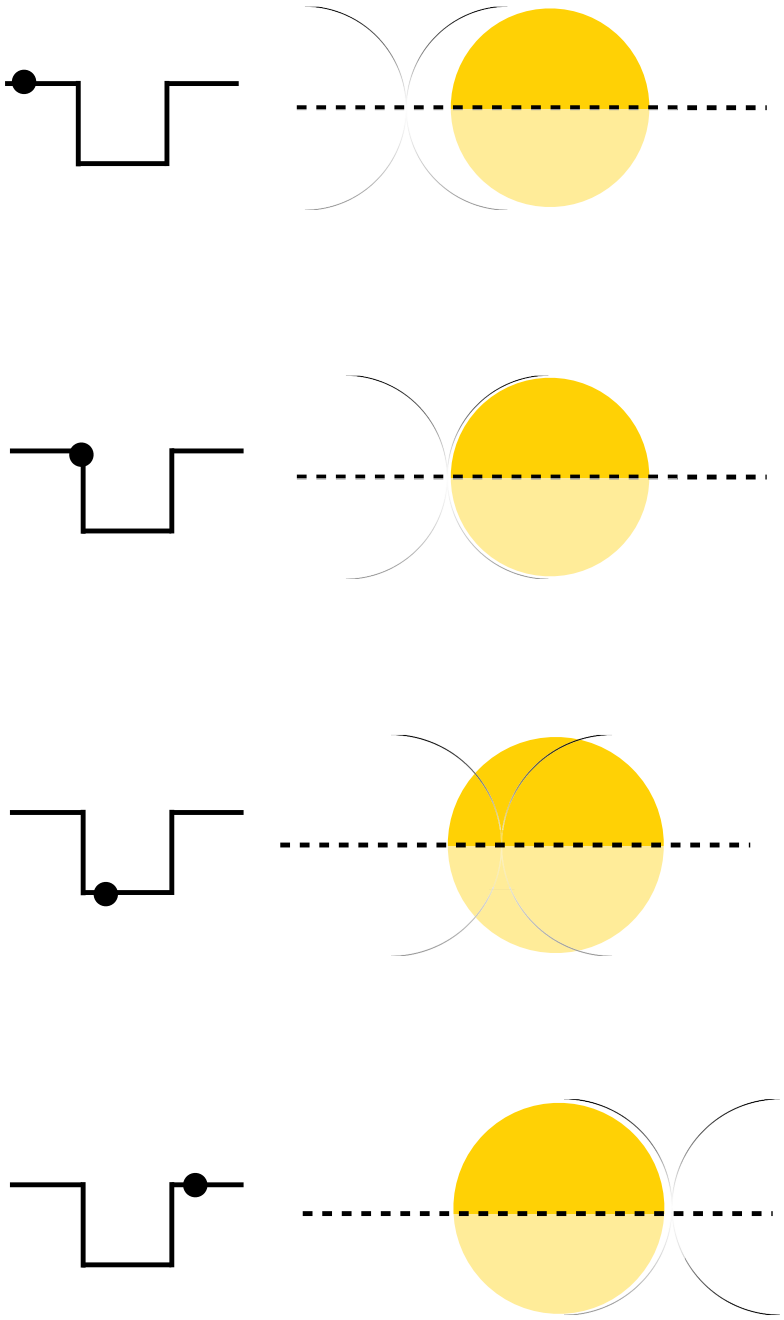

Figure 4.4: A transiting arc pair with opacity distributed as $\lambda(\theta) = \sin\theta$. This shape generates a box-like transit light curve. The circles in the left-hand panels mark the time along the transit at which the right-hand panels occur.

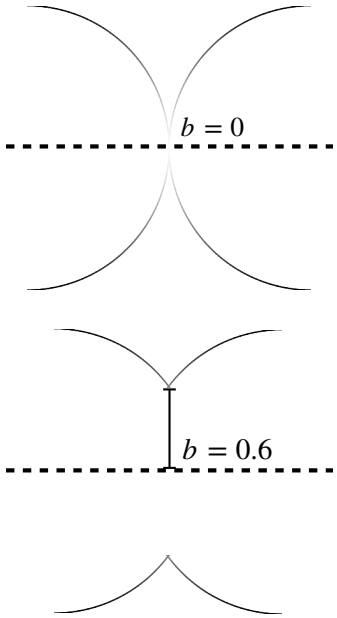

Figure 4.5: A pair of truncated arcs, as illustrated in the lower panel, can match the transit shape of an infinitesimal opaque point transiting at arbitrary impact parameter.

a uniformly bright star: if the star is limb-darkened, then there is no (unchanging) arc arrangement which can maintain the perfect opacity ingress-egress balance described by Equation (4.6). However, in practice, the limited time resolution of light curve observations leaves room for significant arc-degenerate behavior in shadow images recovered from real transit data (see 4.5 below).

4.2.3 The Stretch Degeneracy

A third degeneracy inherent to light curve imaging results from the "scale-free" nature of the problem, and allows a wide image moving at high velocity to generate the same light curve, within an arbitrarily small measurement uncertainty, as a narrower image moving at lower velocity. We term this degeneracy the "stretch" degeneracy.

The stretch degeneracy is mathematically simpler than the arc or flip degeneracies. Two

occulters with the same transit duration T both obey

$$T = \frac{W}{v}, \tag{4.12}$$

where W is the width of the occulter, and v is its velocity. The right-hand side of this equation can be multiplied by the same constant in the numerator and denominator without consequence to T. In other words, a "stretched" image traveling fast can generate a transit event of the same duration as a narrow image traveling slowly.

Figure 4.6 illustrates the stretch degeneracy for a simple, low-resolution circular occulter. Note in particular two features of the "stretched" image: first, that it is semi-opaque rather than fully opaque like the un-stretched image, and second, that its edges are less opaque than its middle. The semi-opacity of the stretched image is necessary in order to match the transit depth of the un-stretched image: because the stretched image is wider, it occults more of the stellar surface, so it must let some light through, or it will produce a much deeper transit than the un-stretched image. Meanwhile, the lightened edges of the stretched image are necessary to better match the ingress and egress shape of the un-stretched image's light curve; with arbitrarily high image resolution, it is possible to match the un-stretched image's light curve to arbitrary accuracy.

In practice, the stretch degeneracy is the least important of the three non-trivial degeneracies we explore in this section, because a fast-transiting, stretch-degenerate image can only match a narrow, slower image's light curve if the image resolution is high enough, as suggested by the example in Figure 4.6. For real data, image resolution is constrained by the number of observed data points, which causes us to prefer the narrowest, slowest possible image which can match an observed light curve (see 4.4.2 for further discussion).

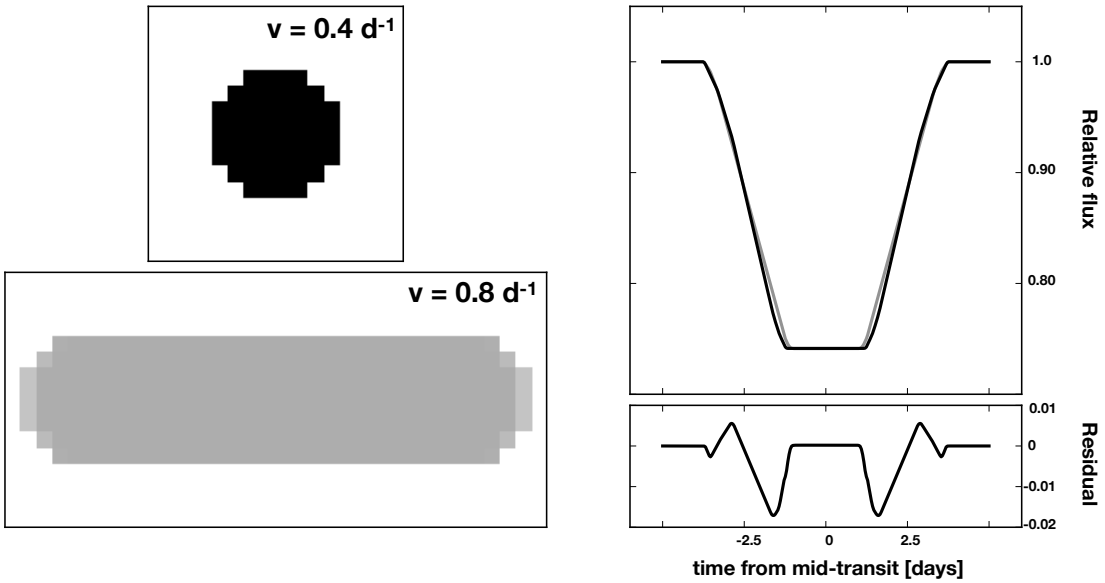

Figure 4.6: Two transiting images with light curves that differ by $\mathcal{O}(1\%)$. The lower image transits at a velocity twice that of the upper image. Note that the left- and rightmost edges of the lower image are slightly less opaque than the middle, an adjustment made to better match the ingress and egress shape of the upper image's light curve. At higher resolution for the lower image, an even better match to the upper image's light curve could be found.

4.2.4 Trivial Degeneracies

Finally, we note two trivial degeneracies which do not affect the inference of a shadow image. The first relates to the arbitrary sign of the velocity of the transiter; an image which transits left-to-right across the star generates the same light curve as the same image, horizontally mirrored, transiting right-to-left across the star at the same velocity. We choose positive v to indicate that the image transits left-to-right (see 4.3.2, below).

The second trivial degeneracy relates to a time translation of the entire transit event. As we discuss in 4.3.2, we must choose a "reference time," analogous to a transit midpoint time, along a light curve in order to recover a shadow image; shifting this reference time forward or backward along the light curve results in a shadow image which is shifted right or left, respectively (given our choice of v direction, above).

4.3 Model: Generating a Light Curve from a Discretized Image

By the arguments of Section 4.2, a given light curve may be generated by infinitely many images. To constrain the solution set, we therefore conclude that it is necessary to impose further constraints on the shadow image. (Starspot inversion requires an analogous constraint—popular choices include the maximum entropy principle, which chooses the solution with minimum spatial correlation between points on the stellar surface, and Tikhonov regularization, which chooses the smoothest solution, or the solution with minimum spatial derivative.)

In this section, we define a forward model for generating a light curve, sampled at discrete time intervals, from a pixelated image. This simulated light curve can be compared to observations of a real transit event. After we establish this forward model, we investigate the inverse problem, of how to infer a pixelated image from an observed light curve, in the next section. We return to the question of degeneracies in 4.4.3.

4.3.1 Discretizing the Image

Pixelating, or discretizing, the shadow image is motivated by recognizing that real light curves are themselves discrete time series. A light curve is not infinitely resolved in time, and therefore we should not attempt to recover a shadow image that is infinitely resolved spatially. Similarly, each flux measurement in a light curve has an associated uncertainty; we should not attempt to recover a shadow image with pixel elements too small to be definitively detected within that uncertainty (see Section 4.4.2 below for further discussion).

Discretizing the pixel image, furthermore, enables us to investigate two physical variants of the shadow imaging problem:

1. What if the transiting object which generated the light curve is a solid body, and

therefore our shadow image should only admit of completely transparent (opacity $\tau = 0$) or completely opaque ($\tau = 1$) pixels?

2. What if the transiting object is dusty or translucent, or is a solid body smaller than the pixel scale, and our shadow image can contain pixels of intermediate opacity ($0 \leq \tau \leq 1$)?

These two variations of the shadow imaging problem have different constraints on the pixel opacities, and require different mathematical approaches to inversion. In case (1), discretizing the pixel image is necessary to divide it up into opaque and transparent elements. In case (2), discretizing the pixel image enables us to set up the light curve inversion problem as a single matrix equation (Equation (4.35), below), and to explore both analytic and numerical approaches to solving this equation. (Similar mathematical formulations exist for both starspot inversion (Vogt et al., 1987) and eclipse mapping, e.g. Berdyugina & Kuhn 2019.)

The same forward model, or procedure for generating a light curve from a pixelated image, can be used in both cases, so we begin there. How do we calculate the light curve of a pixelated image grid transiting a star?

4.3.2 Grid Definitions and Positions

We consider a pixel grid of N rows and M columns transiting a star. We normalize the physical scale of the problem such that the radius of the star is unity.

The grid lives in the X-Y sky-projected plane, with the observer at $Z = +\infty$. The grid moves laterally along the X axis, with $\mathrm{d}X/\mathrm{d}t > 0$, and does not translate up or down (i.e. $\mathrm{d}Y/\mathrm{d}t = 0$). We illustrate this setup in Figure 4.7.

We treat the grid as moving at a constant lateral velocity $v \equiv \mathrm{d}X/\mathrm{d}t$, where $\mathrm{d}v/\mathrm{d}t \equiv 0$. This is a reasonable approximation over the timescale of a transit, unless the object resides on a very tight orbit, or the object is near pericenter on a highly eccentric orbit. We define

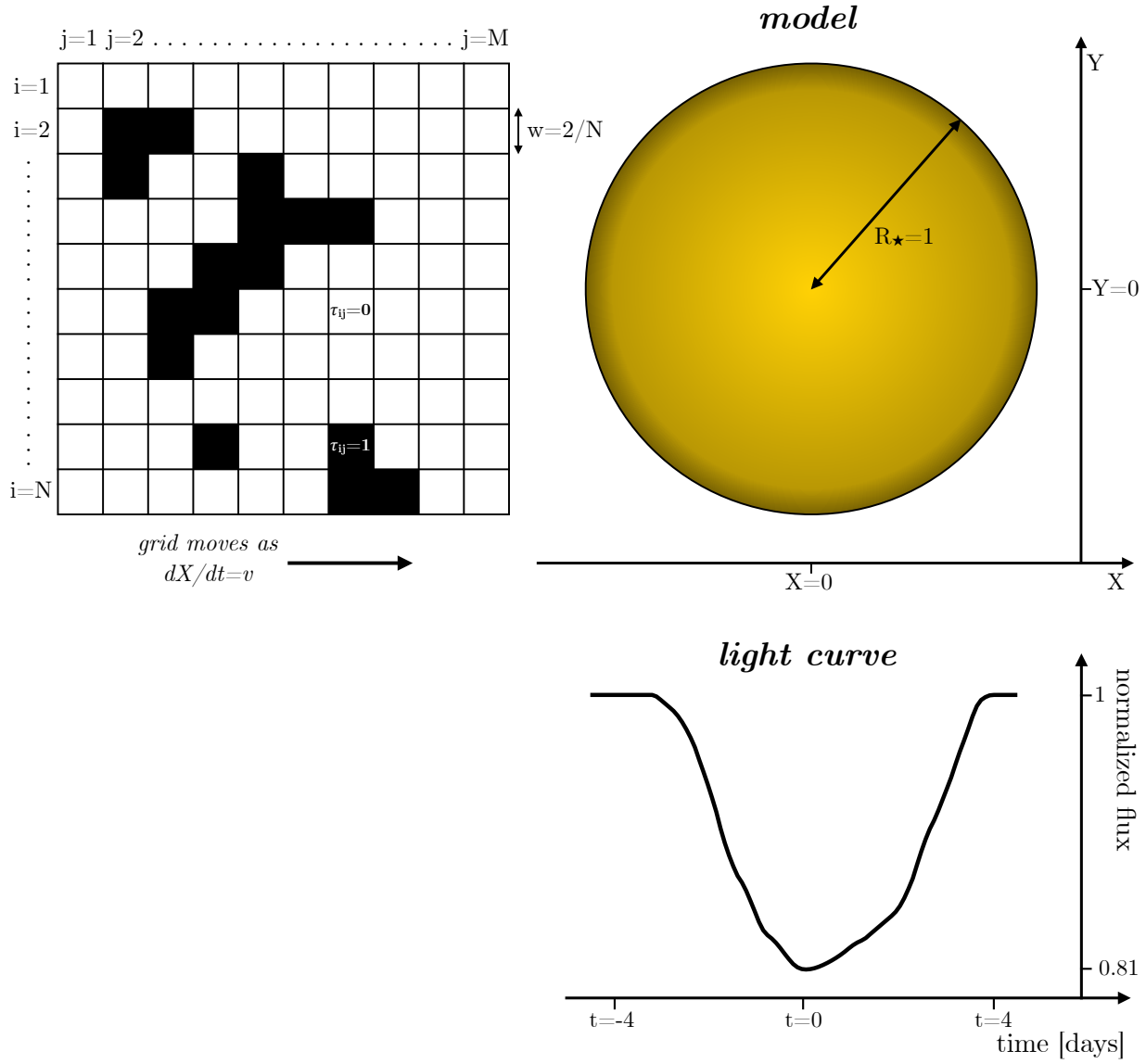

Figure 4.7: (Top panels) Illustration of a $N = 10$ by $M = 10$ binary-opacity grid model with 16 opaque pixels. The star itself is not pixelated; rather, the pixelated grid transits across the star and the exact area of overlap of each square pixel and the star is evaluated at each discrete time step in order to generate a light curve. (Bottom panel) The light curve generated when this grid transits across a uniformly bright star at $v = 0.4$ days^{-1}, $t_{\rm ref} = 0$ days.

positive v to mean that the grid transits from left to right across the star, such that the rightmost column of pixels ingresses first.

We define the vertical position of the grid such that the top of the highest row of pixels falls at $Y = 1$ and the bottom of the lowest row of pixels falls at $Y = -1$. In this way, the grid perfectly overlaps with the star in the vertical dimension.

This definition sets the size of each pixel to have a width, w, of

$$w = 2/N. \tag{4.13}$$

We emphasize that every pixel has the same square shape with this dimension. For $N = 1$, then, $w = 2$ and is thus equal to the diameter of the star.

To refer to individual pixels, we adopt the index notation $i \in [1, N]$ to denote the row and $j \in [1, M]$ to denote the column. To calculate the amount of stellar flux the grid blocks at each discrete time step t_k of the transit observation, we must first calculate the X and Y positions of each grid pixel i, j at each time step.

We may write the Y-position of the center of pixel i, j as

$$Y_{i,j} = 1 - (w/2) - (i - 1)w, \tag{4.14}$$

where setting $i = 1$ recovers $Y_{1,j} = 1 - (w/2)$, and setting $i = N$ recovers $Y_{N,j} = -1 + (w/2)$. The Y positions of the grid pixels are constant.

For the X positions of the pixels, which evolve in time, we first define a reference X position for each pixel at a reference time $t = t_{\rm ref}$ as

$$X^{\rm ref}_{i,j} = X_{i,j}[t = t_{\rm ref}]. \tag{4.15}$$

We choose the reference time such that the grid is centered on the star at $t = t_{\rm ref}$.

Therefore:

$$X_{i,j}^{\mathrm{ref}} = (j - j_{\mathrm{mid}})w, \tag{4.16}$$

where

$$j_{\mathrm{mid}} = 1 + (M-1)/2. \tag{4.17}$$

We may now write the time-evolving X-position of the center of any pixel as

$$X_{i,j}(t_k) = X_{i,j}^{\mathrm{ref}} + (t_k - t_{\mathrm{ref}})v, \tag{4.18}$$

where t_k is the k^{th} time index, and $k \in [1, K]$. Practically speaking, t_{ref} is analogous to the transit mid-time fitted in conventional transit models.

We may use the above equation for the time-evolving $X_{i,j}$ to solve for the time at which the grid makes first and last contact with the star, t_{enter} and t_{exit}. The grid moves from left to right across the star, so at first contact, the M^{th} column of pixels has an X position equal to $-1 - (w/2)$, and for the last contact the 1^{st} column of pixels has an X position equal to $1 + (w/2)$, giving

$$t_{\mathrm{enter}} = t_{\mathrm{ref}} - \frac{1 + Mw/2}{v}, \tag{4.19}$$

$$t_{\mathrm{exit}} = t_{\mathrm{ref}} + \frac{1 + Mw/2}{v}. \tag{4.20}$$

We assign a time-independent opacity $\tau_{i,j}$ to each pixel. $\tau_{i,j}$ is a binary value equal to zero or unity—in other words, we construct our grid of perfectly transparent pixels ($\tau_{i,j} = 0$) and opaque pixels ($\tau_{i,j} = 1$).

In total then, our model has MN opacity parameters, which are binary-valued (case 1) or real numbers between 0 and 1, inclusive (case 2), and two auxiliary parameters, t_{ref} and

v, which are real-valued.

4.3.3 Computing the Light Curve of a Pixel

As pixel i, j transits the star, it occludes a fractional area $A_{i,j}(t_k)$ of the stellar disk at time t_k; $A = 0$ for pixels which do not overlap the star, and $A = \frac{w^2}{\pi}$ for pixels which overlap completely (since we choose $R = 1$, the area of the entire stellar disk is equal to π).

If we assume that the stellar disk is uniformly bright (i.e., there is no limb darkening), we may then compute the light curve $\boldsymbol{F}(t)$ of the transiting grid by recognizing that the fractional flux *blocked* by the grid at each time step t_k is equal to the fractional area of the star occulted by non-transparent pixels ($\tau_{i,j} > 0$), in proportion to their opacity. Therefore, the *unocculted* flux at time t_k is given by:

$$F(t_k) = 1 - \sum_{i=1}^{N} \sum_{j=1}^{M} \tau_{i,j} A_{i,j}(t_k). \tag{4.21}$$

This is the equation for the transit light curve, normalized such that $F = 1$ out-of-transit.

We emphasize that, while opacities $\tau < 0$ and $\tau > 1$ are mathematically permissible in this equation, they are unphysical: a $\tau < 0$ would represent a transiting pixel brighter than the stellar surface it occulted, and a $\tau > 1$ would describe a pixel that blocked more than its proper area's worth of starlight.

We compute the area of overlap $A_{i,j}(t_k)$ of pixel i, j at time step t_k from the (X, Y) position of the pixel's center at t_k, given by Equations (4.14) and (4.18), and the pixel's width, given by Equation (4.13). When a pixel partially overlaps the star, we approximate its overlap area as either a triangle, a trapezoid, or a square missing a triangular corner. We choose the appropriate overlap-shape by computing the number of intersection points between the edge of the star and the sides of the pixel, and also noting whether the center of the pixel falls inside or outside the star. We then correct this approximation by using the length of the chord between intersection points to calculate the area of the sliver of occluded

star yet unaccounted for by this approximation.

For a limb-darkened star, we must also account for the position of each opaque pixel relative to the stellar limb at each time step in order to determine how much flux it occludes. We adopt the small-planet approximation of Mandel & Agol (2002), in which it is assumed that the star's surface brightness is constant across a pixel. In other words, we treat the pixel as occulting a thin, uniform-surface-brightness, annular slice of the stellar disk, where the radius of the annulus is the distance from the center of the stellar disk to the center of the pixel, and the annulus is just wide enough to encompass the pixel.

We denote the area of this annulus as A_{annulus}, and its emitted flux as F_{annulus}. As a rule of thumb, this small-planet approximation is only appropriate for $w \lesssim 0.2$ (i.e. $N > 10$), which corresponds roughly to an occulter-to-star ratio-of-radii of 0.1, for a circular occulter of the same area as the pixel. The exact ratio-of-radii at which the small-planet approximation becomes inappropriate depends on the impact parameter of the pixel, the size of the pixel, the limb-darkening profile of the star, and the bandpass of the observations, so there is no general exact cutoff.

To calculate the light curve in the limb-darkened case, we must re-normalize Equation (4.21): the fractional flux occulted by an opaque pixel is no longer equal to the fractional area of the stellar disk occluded by the pixel, but rather to:

$$\bar{A}_{i,j}(t_k) = \frac{A_{i,j}(t_k)}{A_{\text{annulus}}} \frac{F_{\text{annulus}}}{F_\star}, \tag{4.22}$$

where $F_\star$ is equal to the flux of the entire unocculted, limb-darkened star, relative to the non-limb-darkened star (which must be calculated given a choice of limb-darkening coefficients). We note that this equation reduces to $\bar{A}_{i,j}(t_k) = A_{i,j}(t_k)$ in the case of uniform limb-darkening.

The value of the light curve at t_k is then given by:

$$F(t_k) = 1 - \sum_{i=1}^{N} \sum_{j=1}^{M} \tau_{i,j} \bar{A}_{i,j}(t_k). \tag{4.23}$$

We provide `Python` code to calculate the transit light curve of any grid in the case of uniform, linear, quadratic, or 4-parameter nonlinear (Claret, 2000) limb-darkening in the software package accompanying this paper, `EightBitTransit`.

4.4 Fitting: Shadow Imaging a Pixel Grid from a Light Curve

In this section, we describe how we use the forward model described above to solve the inverse problem, "shadow imaging." We observe a light curve $\boldsymbol{F}$, made up of discrete flux measurements $F_k \equiv F(t_k)$ over K points in time: what pixelated image generated that light curve?

To illustrate the complexity of this problem, we begin with an order-of-magnitude estimation of the number of arrangements of pixels in a binary-valued shadow image (case (1)). There are 2^{NM} unique permutations of transparent and opaque pixels for an N by M grid, and $\mathcal{O}[3^{NM/2}]$ unique light curves that can result (by the flip degeneracy, discussed in 4.2.1). For a 10 by 10 grid, then, there are $\mathcal{O}[10^{30}]$ unique permutations of the binary pixel opacities; accounting for the flip degeneracy, if one wished to find the binary pixel opacity arrangement of the just top half of a 10 by 10 grid to best match a particular light curve, one would have to evaluate $\mathcal{O}[10^{24}]$ possibilities.

A full parameter search is therefore not practically feasible. The largest square grid which could be reasonably fully searched is 5 by 5, for which there are 33.6 million full-grid permutations and 1.9 million half-grid permutations (by Equation (4.1)). We must therefore infer the pixel opacities from the light curve, not attempt to guess them.

To infer a pixel grid from a light curve $\boldsymbol{F}$, we must first select the grid parameters: the dimensions N and M, the velocity v, and the reference time $t_{\rm ref}$. Given these choices, we may calculate the areas of overlap of each grid pixel at each light curve time step, and the corresponding $\bar{A}_{i,j}(t_k)$ for any choice of limb-darkening law. All that remains is to solve Equation (4.23) for the opacities of the grid pixels, $\tau_{i,j}$, subject to the constraints of either case (1) ($\tau_{i,j} = 0$ or 1) or case (2) ($0 \leq \tau_{i,j} \leq 1$).

4.4.1 Mathematical Setup

To be exact, we note that $\boldsymbol{F}$ is a column vector of length K, of which each scalar entry $F_k \equiv F(t_k)$ is given by Equation (4.23). Let us "unravel" the double sum in Equation (4.23) by defining a new index l, such that

$$l[i,j] = j + (i-1)M. \tag{4.24}$$

Since i ranges from 1 to N, and j from 1 to M, l ranges from 1 to MN.

We may then rewrite Equation (4.23) as

$$F_k = 1 - \sum_{l=1}^{L} \tau_l \bar{A}_l(t_k), \tag{4.25}$$

where we define $L \equiv MN$. If we further define $\bar{A}_{k,l} \equiv \bar{A}_l(t_k)$, then

$$F_k = 1 - \sum_{l=1}^{L} \tau_l \bar{A}_{k,l}. \tag{4.26}$$

Let us now rewrite $\bar{A}_{k,l}$ in matrix form:

$$\bar{\mathbf{A}} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,L} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ a_{K,1} & a_{K,2} & \cdots & a_{K,L} \end{pmatrix}. \tag{4.27}$$

$\bar{\mathbf{A}}$ is a matrix of shape K by L, where the k^{th} row encodes the state of overlap of the entire pixel grid at time step k, and the l^{th} column encodes the overlap state of pixel l across all time steps.

Similarly, we may "unravel" the opacity matrix τ into a column vector $\boldsymbol{\tau}$ of length L:

$$\boldsymbol{\tau} = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_L \end{pmatrix}. \tag{4.28}$$

We may now re-express Equation (4.23) in matrix form:

$$\boldsymbol{F} = \boldsymbol{1} - \bar{\mathbf{A}}\boldsymbol{\tau}, \tag{4.29}$$

where

$$\boldsymbol{F} = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_k \end{pmatrix} \tag{4.30}$$

and $\boldsymbol{1}$ is a column vector of ones, equal in length to $\boldsymbol{F}$.

If we define a vector $\boldsymbol{R} = \boldsymbol{1} - \boldsymbol{F}$, we may rearrange this equation to read

$$\overline{\mathbf{A}}\boldsymbol{\tau} = \boldsymbol{R}. \tag{4.31}$$

If $\overline{\mathbf{A}}$ were invertible, then our work would be done: we could solve Equation (4.31) directly for the vector of pixel opacities $\boldsymbol{\tau}$. However, because of the flip degeneracy, pixel i, j has the same area-of-overlap at every time step as pixel $(N+1-i), j$, and as a result, $\overline{\mathbf{A}}$ always has repeated columns. By the invertible matrix theorem, a matrix with repeated columns is not invertible.

We may proceed by recognizing that $\overline{\mathbf{A}}$ and $\boldsymbol{\tau}$, since they describe the entire pixel grid, contain redundant information. We need only solve for the opacities of one half of the pixels (we choose the top half, for convenience). We define a new index

$$L_{\text{half}} \begin{cases} \frac{(N-1)M}{2} + M, & N \text{ odd} \\ \frac{NM}{2}, & N \text{ even.} \end{cases} \tag{4.32}$$

We define a new area-of-overlap matrix $\overline{\mathbf{A}}_{\text{half}}$, which represents the left half (columns 1 through L_{half}, inclusive) of $\overline{\mathbf{A}}$, and a new opacity vector $\boldsymbol{\tau}_{\text{half}}$, which represents the corresponding top half of $\boldsymbol{\tau}$. We may then write:

$$\overline{\mathbf{A}}_{\text{half}}\boldsymbol{\tau}_{\text{half}} = \boldsymbol{R}. \tag{4.33}$$

Since, in general, $K \neq L_{\text{half}}$, we may multiply both sides of this equation by $\overline{\mathbf{A}}_{\text{half}}^{\text{T}}$ to yield

$$\overline{\mathbf{A}}_{\text{half}}^{\text{T}}\overline{\mathbf{A}}_{\text{half}}\boldsymbol{\tau}_{\text{half}} = \overline{\mathbf{A}}_{\text{half}}^{\text{T}}\boldsymbol{R} \tag{4.34}$$

so that both sides of the equation are column vectors of length L_{half}, and $\overline{\mathbf{A}}_{\text{half}}^{\text{T}}\overline{\mathbf{A}}_{\text{half}}$ is a square matrix.

For notational simplicity, let $\mathbf{B} \equiv \bar{\mathbf{A}}_{\rm half}^{\rm T} \bar{\mathbf{A}}_{\rm half}$, and let $\boldsymbol{C} \equiv \bar{\mathbf{A}}_{\rm half}^{\rm T} \boldsymbol{R}$, such that

$$\mathbf{B}\boldsymbol{\tau}_{\rm half} = \boldsymbol{C}. \tag{4.35}$$

We have therefore reduced our shadow imaging problem to the problem of solving a system of linear equations for the entries of the column vector $\boldsymbol{\tau}_{\rm half}$. These entries, reshaped into the matrix τ, correspond to the opacities of the pixels making up the top half of the grid, which define the image.

In the sections below, we elaborate upon the two steps of shadow imaging: first, selecting the grid parameters, and second, solving Equation (4.35) for the pixel opacities subject to our chosen physical constraints.

4.4.2 Constraining the Grid Parameters

In general, the auxiliary parameters $t_{\rm ref}$ and v can be set to reasonable approximations of their "true" values, and the pixel image will slightly shift or stretch, respectively, relative to the "truth," without disturbance to its principal morphology. This means we can proceed by fixing these terms and optimizing the opacities $\boldsymbol{\tau}$ only. We may then, depending on the success of the solution $\boldsymbol{\tau}$, perform further iterations, varying the grid parameters each time, to reach an optimal grid with optimal auxiliary parameters. We discuss here some constraints of the grid parameters which allow us to estimate their values initially.

The first constraint we consider is that the number of pixel elements should not exceed the number of data points obtained during the transit event of interest. For regularly sampled data, such as that of *Kepler*, we may write the sampling constraint as

$$NM \leq \frac{t_{\rm event}}{t_{\rm cadence}}, \tag{4.36}$$

where $t_{\rm event}$ is the timescale of the event we wish to image and $t_{\rm cadence}$ is the cadence of

the time series, i.e. the interval between successive observations.

The second constraint we consider is that a pixel should not be too small to detect individually. In other words, the transit depth of a single opaque pixel should not be smaller than the uncertainties on the flux measurements. In principle, smaller pixels could be resolved over repeated transit observations, but this approximation again aids in selecting a unique initial grid size from which to begin optimizing the grid opacities.

Mathematically, we can express the precision constraint as:

$$\frac{w^2}{\pi} \gtrsim \sigma \tag{4.37}$$

where σ is the typical photometric uncertainty. Combining Equation (4.13) with this constraint gives

$$N \lesssim \sqrt{\frac{4}{\pi\sigma}}. \tag{4.38}$$

For reference, using a 60 ppm uncertainty, this yields $N \lesssim 146$. (In practice, we are usually limited to much smaller values of N by the number of data points in the observed light curve.)

The third constraint we consider is the size of M. Our grid must be wide enough to create a total duration sufficient to explain the event timescale, t_{event}. We require that $t_{\text{exit}} - t_{\text{enter}} \geq t_{\text{event}}$, or

$$\frac{2 + 2(M/N)}{v} \geq t_{\text{event}}. \tag{4.39}$$

Similarly, we consider that a single pixel needs to be able to traverse the entire disk of the star within the event timescale. The actual duration of a single pixel's transit will depend on

the pixel's latitude Y, but to simplify things, we consider an equatorial pixel of infinitesimal size and use an approximate symbol for the inequality, to give

$$v \gtrsim 2/t_{\text{event}}. \tag{4.40}$$

Together, these expressions constrain the velocity to the range

$$2/t_{\text{event}} \lesssim v \leq 4/t_{\text{event}}. \tag{4.41}$$

As a general strategy, then, we choose a grid velocity v equal to $2/t_{\text{event}}$, and t_{ref} to correspond to the minimum of the observed light curve. To choose N and M, we recognize that, for a chosen N, we may solve for M such that the grid continuously overlaps the star, by rearrangement of Equations (4.20). We can then adjust N to accommodate the constraint that NM be less than the number of observed data points. Once the grid dimensions have been chosen, we re-execute the inversion for different velocities, until the fit ceases to improve.

Because of the resolution constraint, we prefer the slowest grid velocity v which returns a reasonable fit to the observed light curve, because this slow velocity corresponds to the highest image resolution N. This is, in a sense, an image prior which prefers narrow, slow images to their fast, stretch-degenerate counterparts.

4.4.3 Solving for the Pixel Opacities

Once we have reasonable first estimates for t_{ref}, v, N, and M, and have chosen a limb-darkening law to describe the stellar disk, we may use Equation (4.22) to solve for $\bar{A}_{i,j}(t_k)$ for each grid pixel at each light curve time step. At this stage of shadow imaging, it is helpful to think of the grid pixels as containers for as-yet-to-be-determined opacity: each

transits the star in a definite way according to the grid parameters, so $\bar{A}_{i,j}(t_k)$ and hence $\bar{\mathbf{A}}$ are well-defined, but its opacity is not yet known.

For a chosen t_{ref}, v, N, and M and , we restrict our attention to the observed light curve data points that satisfy $t_{\text{enter}} < t < t_{\text{exit}}$. In other words, we consider only the points in time during which the grid partially overlaps the star, because the transiting grid could not influence points outside this range.

To determine the opacities, we must solve Equation (4.35) for the entries of the opacity vector $\boldsymbol{\tau}$. Since this matrix equation is linear, in principle it can be directly, analytically solved.

However, direct solution of Equation (4.35) cannot accommodate constraints on the pixel opacities. Namely, there is no way to restrict the entries of $\boldsymbol{\tau}$ to the physically meaningful range $[0, 1]$ (case (1)), let alone to the binary values 0 or 1 (case (2)). Mathematically, introducing these constraints transforms the problem into a nonlinear optimization problem, which is not susceptible to solution by a linear matrix equation. We furthermore find that transforming the opacity variables through a logistic function, which maps the real numbers to the range $[0, 1]$, results in numerical instabilities in our attempts to solve Equation (4.35) both directly and iteratively (e.g. with SART; see 4.4.3.2 below).

Furthermore, we find that choosing grid parameters t_{ref}, v, N, and M that deviate even slightly from the true values leads to completely nonsensical recovered $\boldsymbol{\tau}$. Direct analytic solution is therefore not robust enough to apply to a light curve of unknown origin, where our initial guesses for the grid parameters are unlikely to be so accurate.

We therefore explore less exact, but significantly more robust, algorithmic approaches to solving for $\boldsymbol{\tau}$. Below, we discuss each of these algorithms in turn. The first two address case (1), where pixels may take on intermediate opacities, and the latter three address the more restrictive case (2), where pixels are constrained to be binary-valued.

In Figures 4.8 and 4.9, we compare their performances in recovering a number of known test grids from noiseless light curves. In these recovery tests, the parameters N, M, v, and

t_{ref} were assumed to be known. Eight of the test grids are binary-valued, and three (the low-resolution planet-moon, planet-ring, and comet) include intermediate-opacity pixels.

In Figures 4.8 and 4.9, we have chosen to generate our test light curves with a uniformly bright star, i.e., without limb darkening. We make this choice because non-limb-darkened light curves are sharper and less rounded than limb-darkened light curves, and the inversions result in correspondingly sharper image grids, among which the differences between the images generated by our four recovery algorithms stand out most clearly.

We find that introducing realistic limb darkening results in very similar recovered images to those shown in Figures 4.8 and 4.9, with two notable qualitative differences: first, for the limb-darkened case, opacity tends to be pushed farther out towards the top and bottom edges of the recovered image. This effect is most obvious in the arc-combinatoric images. Second, the recovered images appear blurrier, which makes intuitive sense given the more rounded features of a limb-darkened transit event compared to a non-limb-darkened transit.

4.4.3.1 Arc-Averaging

The first algorithmic approach we explore relies on the time derivative of Equation (4.31). At each time step dt, the overlap state of the grid changes; we can express the change in overlap area as the matrix $d\overline{\mathbf{A}}/dt$, calculated at each time step. Most of the entries of this matrix will be equal to 0, because only the pixels overlapping the stellar limb at that time step will have nonzero change in overlap area.

Meanwhile, at each dt, we can calculate the net change in the observed light curve, $d\boldsymbol{R}/dt$. Two effects can contribute to nonzero $d\boldsymbol{R}/dt$ at a particular time step: (i) one or more pixels with nonzero opacity overlapping the stellar limb at that time step, and (ii) in the case of non-uniform limb darkening, one or more pixels with nonzero opacity overlapping any part of the star. For the low-resolution grid inversions possible given the time resolution of currently available transit data (see e.g. 4.5.1 and 4.5.2), effect (i) is much larger than effect (ii). Additionally, the stellar intensity profile changes most steeply near the limb, so

effect (ii) is most prominent for limb pixels anyway.

Therefore, for the "arc-averaged" algorithm, we take the naive approach of calculating the average $d\boldsymbol{R}/dt$ per pixel which overlaps the limb at that time step. Then, we endow each limb pixel with that average opacity, weighted by $\frac{1}{\sin\theta_{\text{pixel}}} = \frac{R_*}{b_{\text{pixel}}} = \frac{1}{b_{\text{pixel}}}$ to mitigate the effects of the arc degeneracy.

We do the above arc-averaging independently for each time step dt, then average the results over all time steps to compute the final grid. Finally, we re-normalize the pixel grid to match the transit depth of the observed light curve. (Renormalization is necessary because the arc-averaging algorithm only exploits information from the derivative of the light curve, not from the light curve itself.)

Arc-averaged pixel solutions, because they exploit the arc degeneracy, exhibit semicircular arc-like features. They are also horizontally symmetrical as a result of the flip degeneracy. Overall, they are smoother and more dispersed than their true pixel grid counterparts, with smoother light curves, because the averaging step precludes sharp, isolated islands of opacity. The impact parameter weighting causes opacity to be concentrated at the midplane of the grid.

As shown in Figures 4.8 and 4.9, the light curves of arc-averaged solutions match observed light curves well, particularly for large, centrally-concentrated test shapes. The worst matches are for grazing shapes (see e.g. the 16 by 16-pixel grazing circle), because the $\frac{1}{b_{\text{pixel}}}$ weighting pushes opacity strongly toward the grid midplane and away from the top and bottom of the grid. The arc-averaged light curves also tend to be more rounded than the observed light curve, meaning that the arc-averaging approach struggles to reproduce sharp light curve features. This is sensible because, by design, it produces solutions where opacity is distributed continuously along overlapping arcs rather than confined to discrete islands.

We also note that, because arc-averaging can easily accommodate pixel opacities between 0 and 1, it can be applied to semi-opaque pixel grids, like the low-resolution planet-moon, planet-ring, and comet test grids.

4.4.3.2 Simultaneous Algebraic Reconstruction Technique (SART)

The next algorithm we test is called the Simultaneous Algebraic Reconstruction Technique, or SART (Andersen & Kak, 1984). SART was originally developed for medical computed tomographic imaging. Specifically, SART reconstructs a 2D image from the projections of X-rays through the body—this is directly analogous to our shadow imaging problem, where the "projections" of the pixelated image against the stellar disk are the individual data points in the light curve.

SART operates iteratively upon an initial guess for the opacity vector $\boldsymbol{\tau}_{\text{half}}$, which encodes the opacities of the pixels of the top half of the image grid. Beginning from this initial guess, it computes subsequent corrective updates to the individual entries of $\boldsymbol{\tau}_{\text{half}}$.

The $(q+1)^{\text{th}}$ iteration of τ_l, the opacity of pixel l, is given by

$$\tau_l^{q+1} = \tau_l^q + \frac{\sum_{k=1}^{L_{\text{half}}} \left(B_{kl} \frac{C_k - \sum_{\lambda=1}^{L_{\text{half}}} (B_{k\lambda}\, \tau_\lambda^q)}{\sum_{\lambda=1}^{L_{\text{half}}} B_{k\lambda}} \right)}{\sum_{k=1}^{L_{\text{half}}} B_{kl}}, \tag{4.42}$$

The scalar τ_l^q is the l^{th} entry of $\boldsymbol{\tau}_{\text{half}}$ at iteration q, representing the opacity of pixel l; the scalar B_{kl} is k^{th}-row, l^{th}-column entry of $\mathbf{B}$; and the scalar C_k is the k^{th} entry of $\boldsymbol{C}$. $\mathbf{B}$ and $\boldsymbol{C}$ are defined in Equation (4.35).

For intuition, the update term in Equation (4.42) is equal to the average correction to pixel opacity τ_l over all rows and all columns of $\mathbf{B}$. (Hence, the sum in the denominator is over all rows of column vector $\boldsymbol{B}_l$, and the sum in the numerator term's denominator is over all columns of row vector $\boldsymbol{B}_k$.) The numerator, specifically, is the average value over all pixels in the grid of a sort of "residual" between the observed light curve and the model. This residual is equal to C_k minus the scalar projection of $\boldsymbol{\tau}_{\text{half}}^q$ along $\boldsymbol{B}_k$. In effect, these two averages allow for a correction to the opacities which is averaged over all time steps of the light curve and all pixels in the grid.

By running the SART algorithm for a large number of iterations (usually, $\sim 10^4$ for a 16 by 16 pixel grid), we achieve good convergence to the observed light curve for a number of test cases. The RMS error between the light curve of the input image and the light curve of the SART solution declines monotonically over the SART iterations, indicating that SART achieves a progressively better fit to the light curve as it proceeds.

We find that starting from an initial guess of all $\tau_l = 0.5$ works well, because the step-by-step updates to $\boldsymbol{\tau}$ are generally small, so the algorithm does not wander far into unphysical parameter space (i.e., $\tau_l < 0$ or $\tau_l > 1$). In the event that the resulting SART solution does have slightly unphysical opacities, we redistribute the excess positive or negative opacity uniformly among the pixels whose centers fall within a distance of $w/2$ of the arc pair that intersects at the unphysical pixel. This redistribution renders the SART solution fully physical without drastically changing its light curve. Because SART exploits information in the light curve, not just its derivative, it is not necessary to re-normalize the SART solution pixel opacities.

SART solutions exhibit horizontal symmetry as a result of the flip degeneracy, and semi-circular arc-like features as a result of the arc degeneracy. Like the arc-averaging algorithm, SART tends to smear out sharp features in the true input image along arcs, resulting in pixel grid solutions which are smoother, with more dispersed opacity than the true image. (SART solutions are even smoother than the corresponding arc-averaged solutions.) As as a result, SART fails, for example, to match the sharply flat-bottomed transits of the 16 by 16-pixel circle and square test grid light curves (Figure 4.9), producing slightly rounded light curve shapes instead. On the other hand, because SART allows the pixel opacities to take any continuous value between 0 and 1, it can accurately reproduce the light curves of non-binary test grids, like the planet-moon, planet-ring, and comet.

4.4.3.3 Brute Force Search

The next three algorithms we explore attempt to invert light curves subject to the constraint of binary pixel opacities: in other words, we attempt to recover grids with pixel opacities of 0 (completely transparent) or 1 (completely opaque). We begin with the simplest, a brute-force search of every possible arrangement of binary pixel opacities.

As discussed in 4.2.1, by the flip degeneracy, a grid of N by M opaque and transparent pixels can generate $\mathcal{O}[3^{NM/2}]$ unique light curves. Correspondingly, one would have to evaluate $\mathcal{O}[3^{NM/2}]$ permutations of transparent and opaque pixels to find the grid that matches a given light curve best. The largest square grid for which such a full search is feasible is 5 by 5 pixels, which has 1.9 million associated pixel arrangements with unique light curves (for comparison, a 6 by 6 grid has $\sim$ 390 million).

In Figure 4.8, we illustrate the results of a brute force full-grid search for noiseless test light curves generated by number of 5 by 5 known input grids. The brute force algorithm returns the pixel arrangement which, when transiting the star, generates a light curve with the lowest RMS error compared to the truth.

Unsurprisingly, when the input grid is truly binary, i.e. made up of completely opaque and completely transparent pixels, the full search converges to the best possible solution every time. However, when the input grid includes semi-opaque pixels, as in the low-resolution planet-moon, planet-ring, and comet test cases, the brute force search struggles; the lowest-RMS solution does not necessarily bear any resemblance to the input grid, even though its light curve matches the true light curve well. This is a testament to the complex and multimodal likelihood landscape of the pixel opacities, and also an illustration of why conventional nonlinear optimization methods cannot solve the light curve inversion problem. (We note here that we also investigated both a genetic algorithm and a downhill simplex algorithm (Nelder & Mead, 1965), without success—both methods tended to reach local optima and stall, and as illustrated here, locally optimal grids are not necessarily morphologically similar to the true grid.)

Brute-force search solutions are not presented in Figure 4.9, because these grids are far too large to be exhaustively permuted.

4.4.3.4 Parsimonious Opacity Assignment

The next two algorithms we test rely, like arc-averaging, on the time derivative of Equation (4.31). However, instead of averaging the ingress or egress opacity over all of the limb pixels at each time step, we attempt to parcel it out in units of 0.5 opacity (to accommodate the flip degeneracy). We note that consequently, these two algorithms do not work well for inverting shallow transits observed with low time sampling (i.e., few light curve data points), because in such cases, the grid will be low-resolution, and the transit depth of a single pixel's worth of opacity can be greater than the observed transit depth. There will then be no good match to the light curve.

First, we explore the "parsimonious" approach, which assigns opacity to as few pixels as possible in order to accommodate the change in the light curve. This algorithm is motivated by compactness–is it possible to match the light curve with as few "on" pixels as possible?

The parsimonious approach assigns opacity first to the pixel with the largest change in overlap area $d\overline{\mathbf{A}}/dt$, then steps through successively "less influential" pixels until the entire change in the light curve has been accounted for. As with the arc-averaging approach, it is necessary to average the results over all time steps dt, then renormalize the resulting pixel grid to match the observed transit depth; the pixel grid solutions presented in Figures 4.8 and 4.9 therefore have some pixel opacities between 0 and 1.

In practice, this algorithm generates pixel grids which are strongly concentrated at the stellar midplane, because these middle pixels undergo the greatest change in overlap area at fixed dt during their ingress and egress. Correspondingly, it fails to reproduce high-impact-parameter features in the input grids, and is especially poor at matching the light curves of grazing shapes, like the grazing circle and grazing triangle (Figure 4.9). Overall, it is the least successful of the four algorithms.

4.4.3.5 Arc Combinatorics

Finally, we consider an algorithm which attempts to assign units of 0.5 opacity to the best *combination* of limb pixels at every time step in order to match the observed light curve. At every dt, the algorithm calculates the number of "spaces" on the stellar limb, s, that could accommodate a unit of 0.5 opacity. This is equal to twice the number of limb pixels of the appropriate "sign:" if the light curve is decreasing at dt, we need only consider the limb pixels which are undergoing ingress, and vice versa.

Next, it calculates the number of 0.5-opacity units n that need to be accommodated. This is equal to the change in the light curve, $d\boldsymbol{R}/dt$, divided by the mean overlap area of the limb pixels at that time step, multiplied by 2 (because we wish to distribute opacity in units of 0.5, not 1).

The number of ways to arrange n opacity units over s spaces is then $\binom{s}{n}$. The algorithm explores each combination and chooses the one which matches the vector $d\boldsymbol{R}/dt$ best. Finally, as with arc-averaging and the parsimonious approach, the resulting grid is averaged over all time steps and renormalized to match the observed transit depth (so once again, the pixel grid solutions presented in Figures 4.8 and 4.9 therefore have some pixel opacities between 0 and 1).

The arc combinatorics approach is able to match certain vertically-sharp features in the input images, such as the 16 by 16-pixel annulus and column test cases (Figure 4.9). It can also accommodate narrow features at high impact parameter; to see this, compare the parsimonious and arc combinatorics solutions to the 16 by 16-pixel four-squares test case.

Because of the arc degeneracy, however, the arc combinatorics algorithm tends to prefer solutions where opacity is pushed too far toward the top and bottom edges of the grid (e.g. the 16 by 16-pixel circle and square test grids, Figure 4.9). (This is the opposite problem of the parsimonious algorithm.) It also struggles to capture the nuances of semi-opaque test grids, like the planet-moon, planet-ring, and comet. Finally, we note that the computational cost of this algorithm scales poorly with increasing grid resolution (i.e., increasing s), because

the algorithm needs to evaluate $\binom{s}{n}$ opacity arrangement possibilities.

4.5 Real Data

In this section, we discuss the performance of shadow imaging on two real test cases: first, the light curve of the triple transit of TRAPPIST-1c, e, and f (Gillon et al., 2017), and second, two unexplained transit-like events observed in KIC 8462852, or Boyajian's Star (Boyajian et al., 2016).

4.5.1 TRAPPIST-1c,e,f triple transit

We begin with the TRAPPIST-1c,e,f triple transit, for which the expected shadow image is known. We hope to recover an image of three transiting planets, analogous to the diagram presented in Gillon et al. (2017) Extended Data Figure 1.

In attempting to invert this light curve, we have useful prior information beyond the expected image. First, because of the repeated transit observations and N-body dynamical simulations presented in (Gillon et al., 2017), the periods and eccentricities, respectively, of planets c, e, and f are well-constrained. This enables us to calculate the Keplerian orbital velocities of c, e, and f, which we can use as v of our transiting grid. (We note that since these three orbital velocities are different, the pixel image we are attempting to recover changes during the transit, so we will only be able to recover an approximate image for any single choice of v.)

Second, because the physical behavior of this system is so well-understood and the other properties of these planets (R_p/R_*, b, a/R_*) are so well-constrained by transit modeling, we can generate an extremely finely time-sampled model light curve, based on a BATMAN model (Kreidberg, 2015), of this triple transit, which matches the observed light curve. We can use this high-resolution light curve to test the effects of grid resolution on the success of shadow imaging: when the light curve is finely sampled, we can recover a much higher-

resolution grid than when the light curve is sparsely sampled. Finally, we can adopt the same approach to determining the quadratic limb-darkening coefficients for TRAPPIST-1 as Gillon et al. (2017) did in their analysis, interpolating for TRAPPIST-1's stellar properties from the tables of Claret et al. (2012).

In Figure 4.10, we present three inversions of the BATMAN-modeled high-resolution TRAPPIST-1c,e,f, triple transit light curve, conducted with grid v equal to the Keplerian velocity of planets c, e, and f, respectively. We choose $N = 16$ because it is a high-enough resolution that pixel width $w \lessapprox R_p/R_*$ for planet e, the smallest planet ($(R_p/R_*)^2 = 0.52$, according to the transit modeling of Gillon et al. 2017). We show the results of the arc-averaging algorithm here, because it produces the cleanest and most interpretable shadow images, although results from the other three algorithms are qualitatively similar.

In the shadow images, which transit the star moving left-to-right (i.e., the pixels at the right-hand edge of the image transit first), clear ingress and egress arcs for each planet are visible, in the expected order: first, planets c and f ingress together; then, e ingresses; c egresses; f egresses; and e egresses.

The three planets move at three different velocities to produce the light curve, but the grid moves as a unit, so none of the three inversions perfectly matches the light curve model. When the velocity is correct for a particular planet, that planet's image is a pair of arcs whose points of intersection fall at the planet's impact parameter as measured by Gillon et al. (2017), demonstrating that shadow imaging of that planet is successful within the constraints of the arc and flip degeneracies.

When the grid v is slower than the planet's velocity, the planet's ingress and egress arcs are spaced too closely together; this effect is most visible for planet c in the top panel, where the grid moved at planet f's velocity. In the light curve, the overlapping arcs manifest themselves in a too-early dip, caused by c's *egress* arc entering too quickly, and in a too-deep transit depth between the egresses of c and f, caused by c's *ingress* arc remaining in front of the star for too long.

Conversely, when the grid velocity is faster than the planet's, the planet's arcs are too widely separated; this effect is most visible for planet f in the bottom panel. This time, the light curve is too shallow between the egresses of planets c and f, because f's *ingress* arc egresses too soon.

We next investigate what happens if we attempt to invert the observed Gillon et al. (2017) light curve of this triple transit, which is noisy and much more coarsely time-sampled, rather than a high-resolution BATMAN model light curve. Additionally, we ask what happens if we attempt to recover a shadow image without knowing the true velocity of the transiting object: what happens if we use the guidelines presented in 4.4.2 instead?

We invert the observed TRAPPIST-1 triple transit light curve at a range of velocities: the slowest is 31.9 d^{-1}, corresponding to 2 divided by the entire triple-transit event duration (in accordance with the guidelines presented in 4.4.2), and the fastest is 135.9 d^{-1}, corresponding to 4 divided by the duration of planet c's transit by itself. At each velocity, we choose the maximum grid resolution N that, when combined with v to solve for M, allows the transiting grid to partially overlap the star at all time steps of the light curve, while still maintaining NM less than the number of observed data points. Accordingly, the resolution N decreases as v increases, because M increases with v to maintain full light curve coverage.

In Figure 4.11, we present the results of these inversions. There are a number of interesting features about these results. We note, first of all, that SART is consistently the most successful inversion algorithm—this is true across the range of tested grid velocities. Furthermore, the SART shadow image consistently resembles the expected shadow image illustrated in Figure 4.10, even at low image resolutions. Arc combinatorics is somewhat successful at matching the observed light curve at the slowest tested velocity (corresponding to the highest grid resolution), but fails otherwise.

The other algorithms fail consistently across the range of tested velocities. For arc parsimony and arc combinatorics, this results because these algorithms assign binary opacities (0 or 1) to individual pixels, rather than assigning continuous opacities. (We note that the

shadow images presented in Figure 4.11 do not have binary opacities because the final step of both the arc parsimony and arc combinatorics algorithms is to average the binary shadow images produced at each time step dt and re-normalize the average to match the observed transit depth.)

When the pixel resolution of the grid is too low, a single pixel's transit depth can exceed the transit depth of a shallow event like the TRAPPIST-1c,e,f triple transit (maximum transit depth $\sim 2\%$). As a result, the smallest unit of opacity that the arc parsimony or arc combinatorics algorithms can assign is too deep, and these algorithms cannot reproduce the observed light curve. Instead, they tend to assign opacity to pixels along the top and bottom of the image grid, which have the smallest impact on the light curve. This is especially visible in the high-v arc combinatorics panels in Figure 4.11.

Meanwhile, the arc averaging algorithm also fails to match the observed light curve, regardless of velocity. This is because the arc-averaging algorithm, unlike SART, is not robust to noise in the light curve; noise is tantamount to light-curve fluctuations at much higher frequency than can be accommodated by the grid velocity. While SART is able to average out high-frequency noise over many corrective iterations, arc averaging calculates only one arc arrangement per time step dt; if these arrangements are wildly different for neighboring time steps, as they will be for noisy light curves, arc averaging fails.

From these investigations, we conclude that SART is the most robust light curve imaging algorithm. In particular, light curves with large measurement uncertainties and/or shallow transit depths should only be inverted with SART.

4.5.2 Boyajian's Star

Next, we proceed to a light curve with an unknown generative shadow image: that of KIC 8462852, Boyajian's Star (Boyajian et al., 2016). This star exhibits aperiodic, deep transit events of unknown origin; hypotheses to explain these events include a family of transiting comets (Boyajian et al., 2016; Bodman & Quillen, 2016); circumstellar rings (Katz,

2017); an intervening occulter not orbiting Boyajian's Star directly, such as structure in the interstellar medium or an object with a dusty disk (Wright & Sigurdsson, 2016); or circumstellar debris following the star's earlier engulfment of a planet (Metzger et al., 2017); or alien megastructures orbiting the star (Wright et al., 2016).

We focus on the two deepest dimming events observed in the Boyajian's Star light curve during the *Kepler* mission, which Boyajian et al. (2016) label Dip 5 and Dip 8, respectively. Since these events are aperiodic, the appropriate grid velocity v is not obvious; furthermore, in both dips, the light curve smoothly tapers to a sharp point, so the "beginning" and "ending" points of the event are not obvious. Correspondingly, we start from a velocity $v = 2/t_{\text{event, max}}$, where $t_{\text{event, max}}$ is a wide window around the deepest part of the transit, outside of which the flux of the star has essentially returned to 1 again. (These time ranges are plotted in Figures 4.12 and 4.13). We then test several other velocities doubled from this starting point. We interpolate quadratic limb-darkening coefficients for Boyajian's Star from Sing (2010).

The inverted images for Dips 5 and 8 are presented in Figures 4.12 and 4.13, respectively. There are a number of interesting features in these images.

First, there is a circular ring-like feature, of the same radius as the star, that appears generally in images from all four algorithms when the grid velocity v is too slow. (See Dip 5, $v \leq 1.6$ days, and Dip 8, $v \leq 1.2$ days, for examples.) This happens because, when the grid velocity is too slow, the grid struggles to produce narrow features in the light curve; the rate of change of the state of the grid overlap is simply too slow. Under this constraint, the circular ring is the grid pattern that matches a narrow light curve feature best, in the sense that it generates the narrowest possible V-shaped light curve feature.

For intuition, consider a copy of the circular ring with the addition of some opaque interior pixels: at some time steps, these interior pixels will be entirely contained within the stellar disk, and their effect on the light curve during these time steps will be constant. In other words, their transit will be flat-bottomed. This is not the case for the ring, whose overlap

state changes at every time step of the transit; the ring is the "opposite" of a flat-bottomed semicircular arc pair in this sense.

Velocities which produce a ringed shadow image are therefore too slow. This is also obvious from the light curve of the shadow image, which is wider than the observed transit event.

When v is fast enough, we find that all four algorithms produce qualitatively similar shadow images for both Dip 5 and Dip 8, and furthermore that the shadow images of the two dips are similar to each other. As in the case of the TRAPPIST-1 triple transit, the arc parsimony and arc combinatorics algorithms generate "noisy" shadow images where the light curve is shallow, because they cannot assign opacity in units smaller than 1 fully opaque pixel. For Dip 5, SART is clearly the best match to the shadow image; arc averaging produces a light curve which is too narrow, likely because the near-transparent pixels farther from the center which would have produced the "wings" of the transit event have been averaged away in the last combination-and-normalization step of the algorithm. Meanwhile, for Dip 8, the shadow images from both arc averaging and SART match the light curve well.

We strongly caution that there is no straightforward way to interpret these images, for two reasons. First, these images are subject to both the flip and arc degeneracies; second, the grid resolution is low (N = 5 for Dip 5; N = 6 for Dip 8) because we are limited by the 30-minute cadence of *Kepler* observations, and technically this resolution is so low that the small-planet limb-darkening approximation used to calculate the light curves of these shadow images is inappropriate. Nevertheless, these limitations should only affect the distribution of opacity among the semi-opaque pixels in the shadow images of Figures 4.12 and 4.13; pixels which are fully transparent in the shadow images should remain so, even if we were to obtain a much higher-resolution time series of these events.

We therefore note that the "gaps" of near-zero opacity (i.e., nearly transparent regions) which symmetrically frame the opaque transiting blob at the center of the shadow images in Figures 4.12 and 4.13 suggest that structured occulters are responsible for Dips 5 and 8 of

the light curve of Boyajian's Star. The gap structure appears to be necessary to produce a shadow image which matches the ingress and egress shape of both Dip 5 and Dip 8; we note, for example, that this gap structure is missing in the arc-averaged shadow image of Dip 5 at $v = 3.2\mathrm{d}^{-1}$, and the ingress and egress features of the corresponding light curve are too sharp to match the observed light curve.

4.6 Conclusions

Here, we have developed a mathematical and computational framework to address the problem of shadow imaging, or inferring the shape of a transiting object from its light curve alone. We find that this problem, which amounts to reconstructing a two-dimensional map from a one-dimensional time series, is degenerate, like the analogous problems of eclipse mapping and starspot inversion. In particular, by the flip degeneracy, shadow images are horizontally symmetrical; by the arc degeneracy, any infinitesimal opaque point in a shadow image can be replaced by a pair of intersecting semicircular arcs without consequence to the light curve; and by the stretch degeneracy, a wide image transiting at high velocity can produce the same light curve as a narrow image transiting slowly, given high enough pixel resolution.

In spite of these degeneracies, we are able to recover informative shadow images by adopting additional assumptions in algorithmic approaches to inverting light curves. We investigate four algorithms with different underlying assumptions. The first is arc averaging, which assumes that opacity should be distributed along arcs in inverse proportion to the $\sin\theta$ opacity distribution characteristic of the arc degeneracy. The second is the Simultaneous Algebraic Reconstruction Technique, an iterative approach which assumes that opacity should be distributed so as to minimize the RMS averaged over all time steps of the light curve and all pixels in the grid. The third is arc parsimony, which assumes that opacity should be distributed to as few individual opaque pixels as possible. The fourth is

arc combinatorics, which assumes that opacity should be assigned to the best combination of individual opaque pixels to match the light curve. More broadly, the first two algorithms require only that the grid opacities be physical (i.e., restricted to the range $[0, 1]$), while the latter two algorithms operate under the more restrictive assumption that the grid pixel opacities ought to be binary-valued. The less-restrictive case can accommodate pixel images of dusty, translucent, or solid objects smaller than the pixel scale, while the more-restrictive case is in principle more appropriate for recovering an image of a solid body which is larger than the pixel scale..

Overall, we conclude that SART is the approach which is most robust to our choices of grid resolution and velocity, most robust to noise in the observed light curve, and best able to accommodate shallow transit events. The only downside of SART is that, because it is an iterative optimization method, it is not parallelizable. For grids of the size investigated here ($N \leq 16$), it is of perfectly manageable computational cost.

We evaluate the performance of the four algorithms on a number of test cases, and find that we can recover informative shadow images for both binary- and continuous-valued opacity grids. We also apply them to real transit events—first, the triple transit of TRAPPIST-1 c, e, and f, for which the true shadow image is known. We recover a shadow image of TRAPPIST-1 c, e, and f which matches our expectations, subject to the constraint that our model grid transits the star at a single velocity, while the real TRAPPIST-1 planets all move individually.

We also apply our techniques to two of the dips observed in Boyajian's Star, for which the true shadow image is unknown. We recover images which are self-consistent in the sense that the results from all four algorithms are qualitatively similar; also, the shadow images of Dip 5 and Dip 7 resemble each other. Transparent gaps in the shadow images of both events suggest that both dips were caused by structured occulters. However, we caution that these shadow images are difficult to interpret: they are subject to both the flip and arc degeneracies, and they are limited in resolution by the cadence of the original *Kepler*

observations. In the future, for successful shadow imaging of events like these, high time sampling of the light curve is key.

An important next step in shadow imaging will be to expand the framework presented here to encompass a true *inference* of shadow images: in other words, to recover, given a transit light curve, a distribution of images which could have generated it, complete with uncertainties on the pixel opacities. Such a distribution would meaningfully represent the full set of degenerate solutions that could generate a particular observed set of uncertain flux measurements in a way that a single image cannot.

Accounting for measurement uncertainties is certainly possible within the work presented here; one could, for example, draw repeated "realizations" of a particular light curve given the uncertainties on the individual flux measurements, then invert each realization to recover a single shadow image. The deeper question is how to take what is currently a deterministic retrieval procedure—one light curve, inverted with any of our algorithms, yields exactly one reproducible shadow image—and build in a way to account for the physical degeneracies of the problem, particularly the arc degeneracy, such that one light curve can generate an ensemble of possible shadow images.

In principle, one could also attempt to engineer such an ensemble from a single shadow image by perturbing opacity along arcs. We find that in practice, because of the complex overlapping pattern of the ingress and egress arcs, it is very difficult to perturb opacities and maintain a good fit to the observed light curve. In other words, the arc structure renders the pixel opacities strongly and non-trivially correlated. It remains nevertheless an interesting avenue for future work.

To accompany this work, we present the software package `EightBitTransit`, implemented in `Python`, which is able to calculate the light curves of arbitrary pixel arrangements and to recover shadow images from an input light curve, given the user's choice of grid parameters and inversion algorithm. This software package is available at https://github.com/esandford/EightBitTransit.

Acknowledgements

The authors thank the referee for a thorough and thoughtful review, and members of the Cool Worlds Lab for useful discussions. ES thanks Zephyr Penoyre for help building the mathematical formalism of the arc degeneracy, and for many conversations throughout the project. ES thanks Moiya McTier for test-driving the `EightBitTransit` installation instructions.

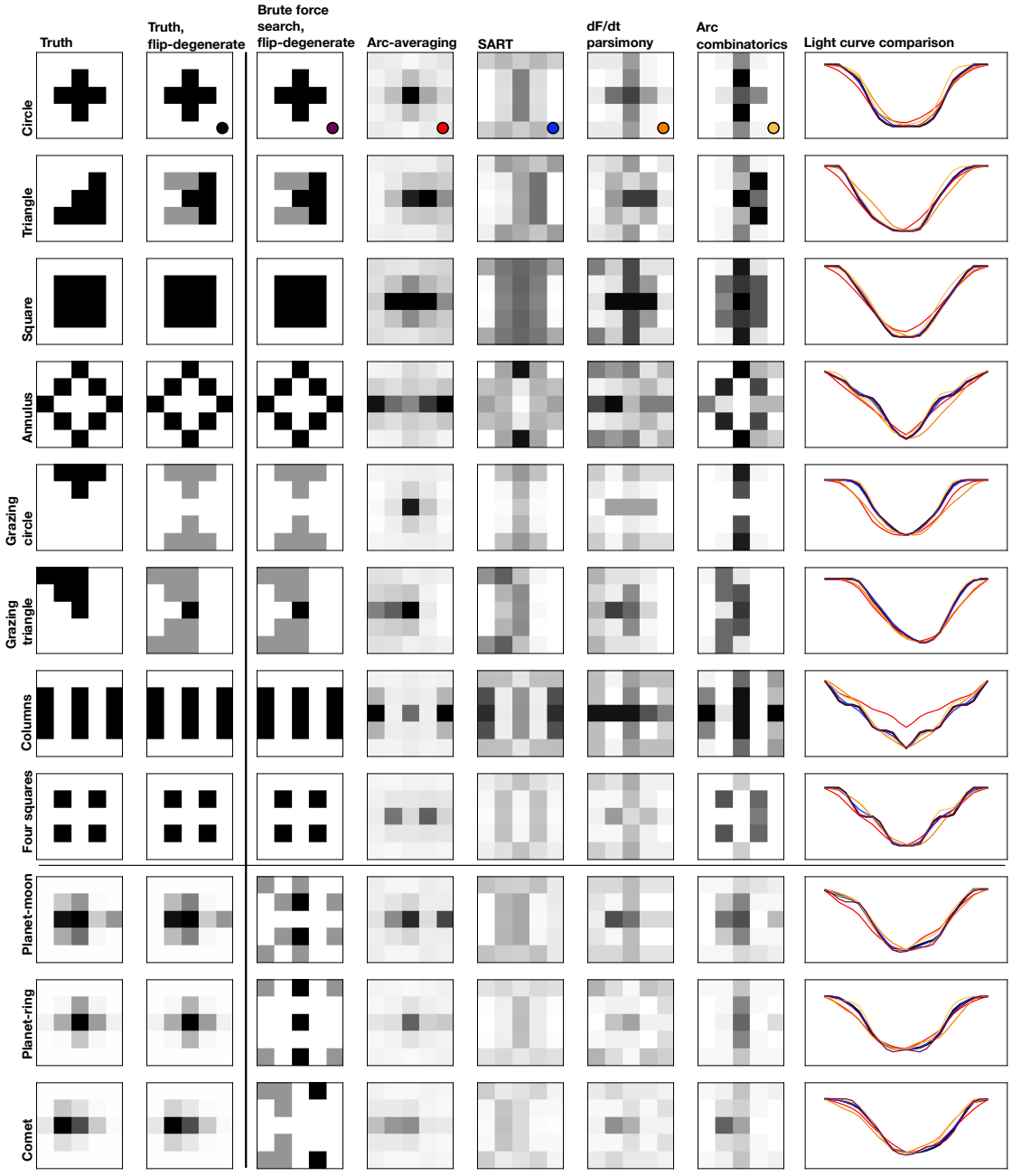

Figure 4.8: The performance of several light-curve inversion algorithms on eleven known 5 by 5-pixel test grids. The leftmost two columns represent the true input grid; the subsequent columns represent the grid recovered by each inversion algorithm given only the (noiseless) true light curve as input. The eight test grids above the horizontal line are pure binary grids (i.e., pixel opacities are either 0 or 1); the three below have intermediate, semi-opaque pixels. Each algorithm was initialized with correct grid parameters N, M, $t_{\rm ref}$, and v, and the light curves were generated with a uniform limb darkening law. The brute-force search algorithm performs the best, i.e. returns the light curve with lowest RMS error compared to the true image's light curve, in every pure binary test case, but SART performs best on the semi-opaque test cases.

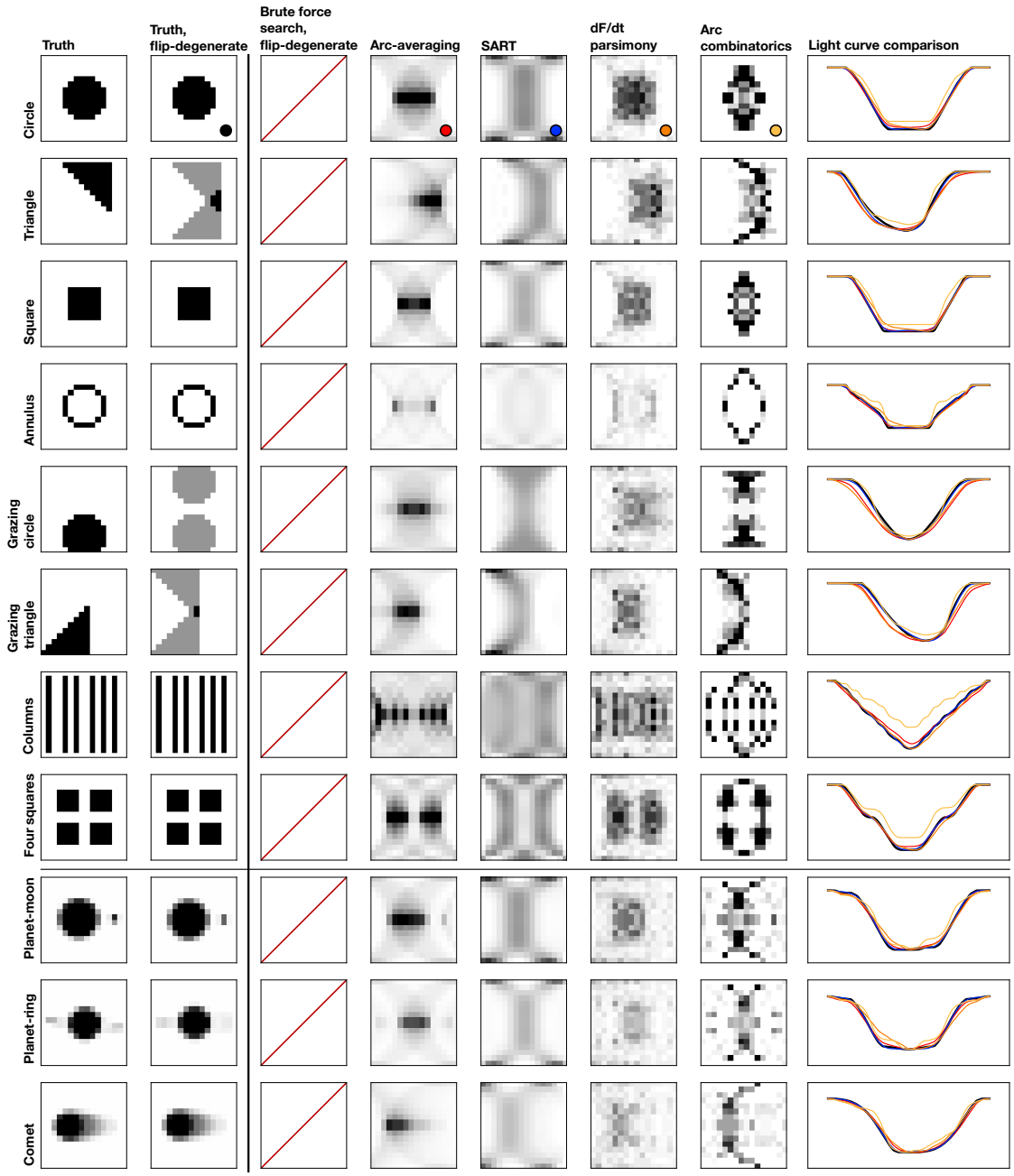

Figure 4.9: The performance of several light-curve inversion algorithms on eleven known 16 by 16-pixel test grids, which are too large to allow for a brute-force permutation search. The leftmost two columns represent the true input grid; the subsequent columns represent the grid recovered by each inversion algorithm given only the (noiseless) true light curve as input. The eight test grids above the horizontal line are pure binary grids (i.e., pixel opacities are either 0 or 1); the three below have intermediate, semi-opaque pixels. Each algorithm was initialized with correct grid parameters N, M, $t_{\rm ref}$, and v, and the light curves were generated with a uniform limb darkening law. SART performs the best, i.e. returns the light curve with lowest RMS error compared to the true image's light curve, in every test case; arc-averaging is second-best in every case except the offset circle, for which arc combinatorics does better.

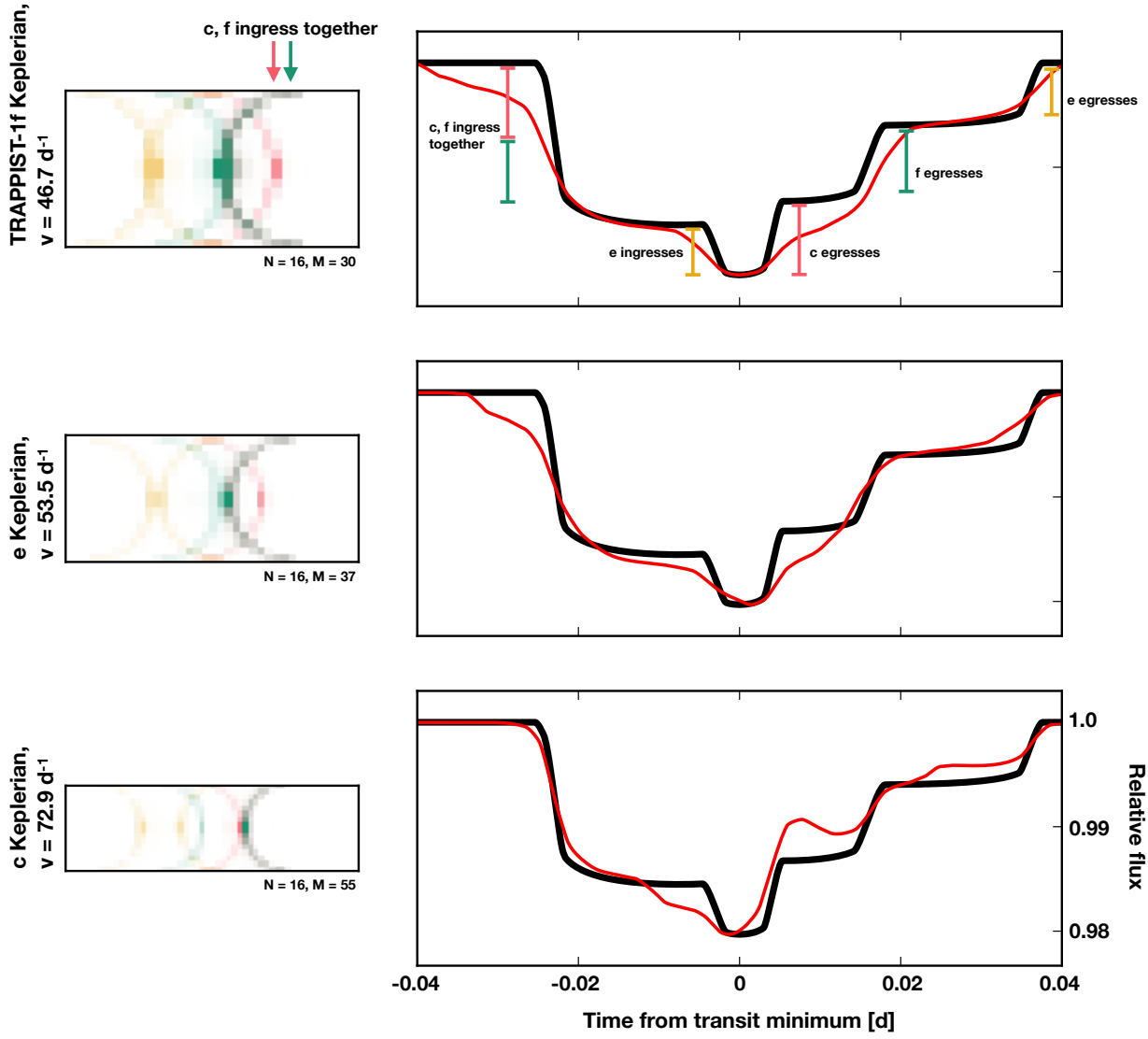

Figure 4.10: Three inversions of a BATMAN-modeled, high-resolution TRAPPIST-1c,e,f, triple transit light curve, conducted with the arc-averaging algorithm, with grid v equal to the Keplerian velocity of planets c (bottom), e (middle), and f (top). These images transit the star moving left to right, so the features at the right-hand side of the image influence the light curve first. The BATMAN model light curve (black) and arc-averaged shadow image light curve (blue) are compared in the right-hand panels. We have added color to the shadow images to indicate the positions of planets c (pink), e (yellow), and f (green). (Note that c and f ingress together, so their ingress arc is green + pink = gray.)

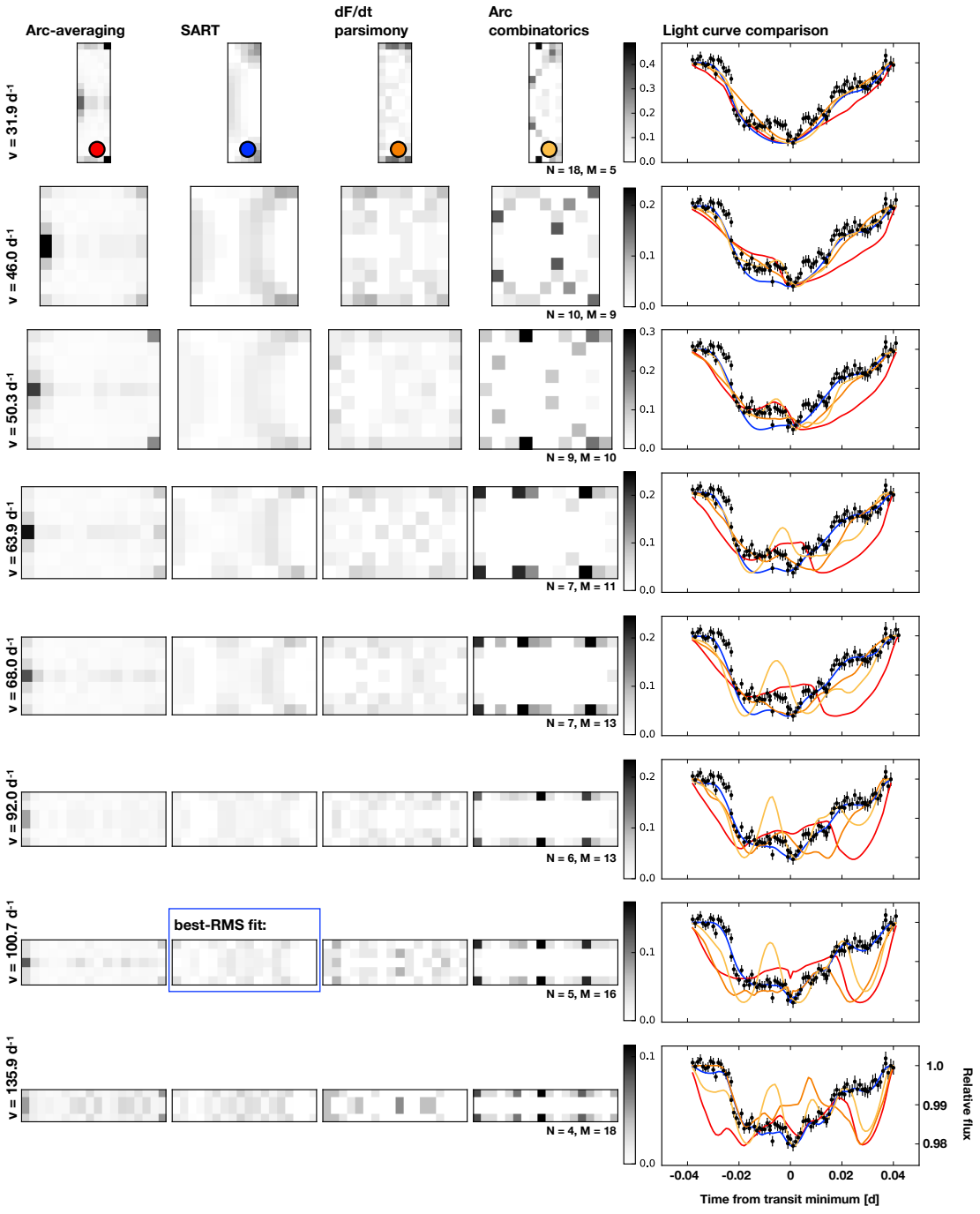

Figure 4.11: The performance of several light-curve inversion algorithms on the observed TRAPPIST-1c,e,f triple transit light curve. The test velocities and corresponding grid resolutions were chosen according to the guidelines set out in 4.4.2. The shadow image whose light curve has the lowest RMS error compared to the observed light curve is the SART inversion at $v = 100.7\mathrm{d}^{-1}$, marked by the blue box. Arc combinatorics performs best, by RMS, at the two slowest tested velocities, but SART performs best at all the others.

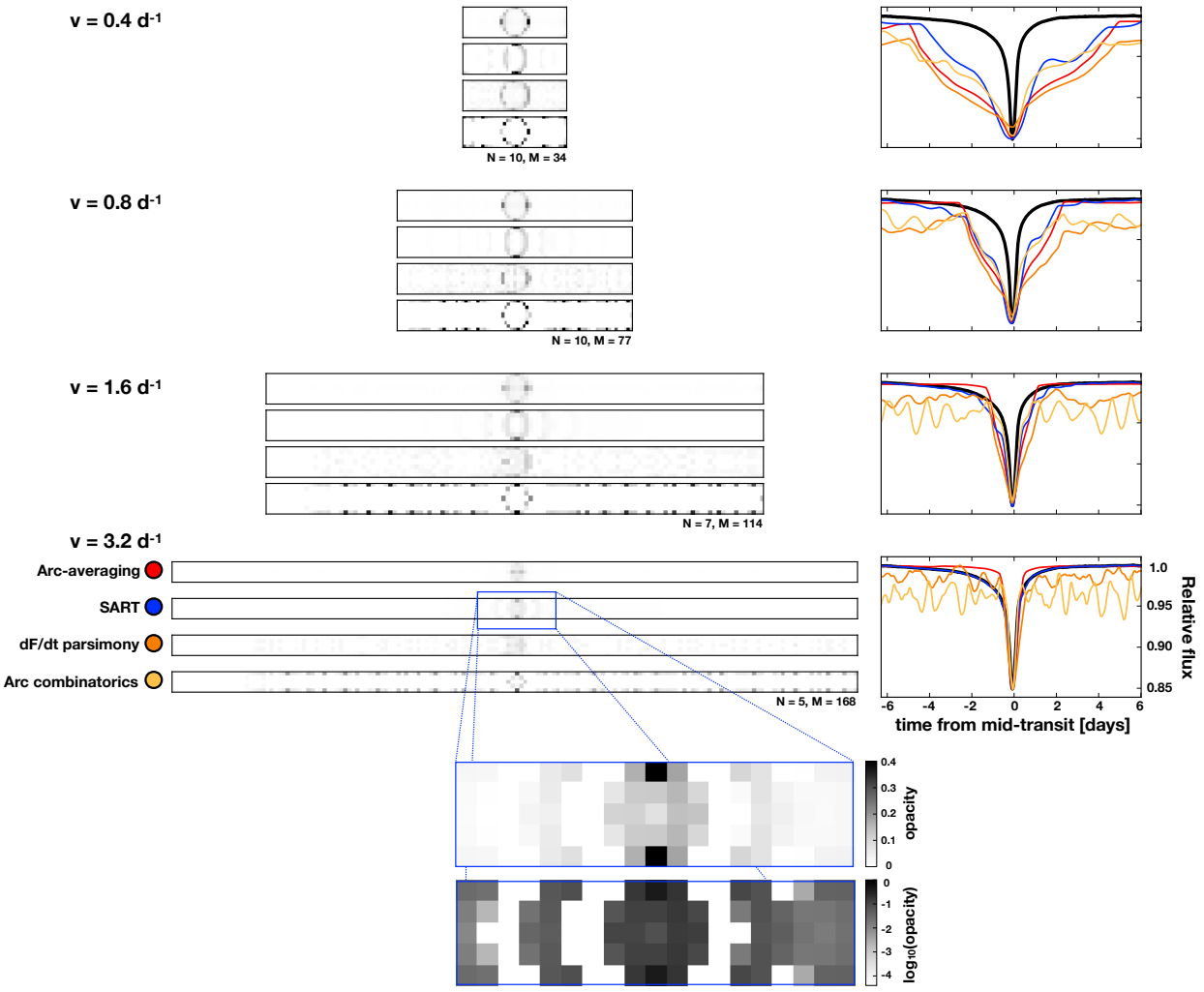

Figure 4.12: The performance of several light-curve inversion algorithms on Dip 5 of Boyajian's Star. Inset: a zoomed-in view of the central SART shadow image, with both linear and logarithmic color scaling to represent opacity. SART performs best, by RMS, at all four choices of v.

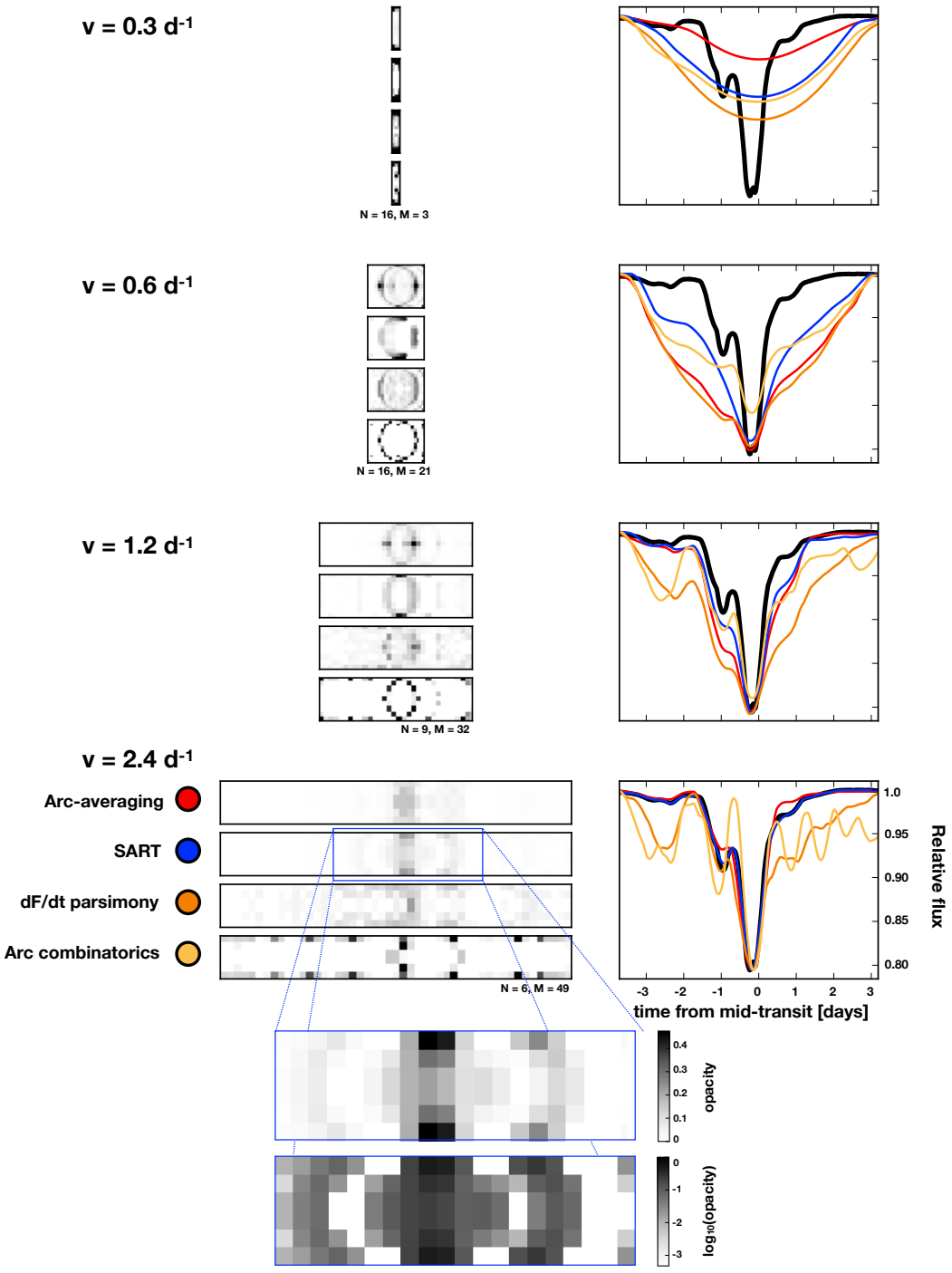

Figure 4.13: The performance of several light-curve inversion algorithms on Dip 8 of Boyajian's Star. Inset: a zoomed-in view of the central SART shadow image, with both linear and logarithmic color scaling to represent opacity. SART performs best, by RMS, at all four choices of v.

5. The multiplicity distribution of *Kepler*'s exoplanets

5.1 Introduction

From our vantage point within the Solar System, it seems natural to expect that stars should be accompanied by multiple planets. Around stars similar to the Sun, planetary systems are common, with numerous studies of the *Kepler* sample converging on an occurrence rate of at minimum one planet per star (Petigura et al., 2013; Foreman-Mackey et al., 2014b; Burke et al., 2015; Hsu et al., 2018). In the majority of *Kepler* systems, there is just a single transiting planet detection (Thompson et al., 2018) but the strong observational biases plaguing transit surveys (Kipping & Sandford, 2016) mean that one might reasonably expect many of these to in fact be multi-planet systems ("multis") yet to be revealed.

Measuring the multiplicity distribution is crucial, as it is a vital clue to the origins and evolution of detected systems. For example, hot Jupiters rarely reside in multi-planet systems (Wright et al., 2009; Steffen & Agol, 2005; Gibson et al., 2009; Latham et al., 2011; Steffen et al., 2012) barring exceptional cases like WASP-47b (Becker et al., 2015; Weiss et al., 2017). This is often interpreted as evidence for late inward migration from beyond the snow line, leading to scattering of interior planetesimals (Beaugé & Nesvorný, 2012; Spalding & Batygin, 2017; Heller, 2019; Dawson & Johnson, 2018). At the other extreme, systems like Kepler-11 (Lissauer et al., 2011b) and TRAPPIST-1 (Gillon et al., 2017) pack half a dozen planets within the orbit of Mercury, which suggests that disk migration and resonant

trapping may guide the evolution of such systems (Quillen, 2006; Mustill & Wyatt, 2011; Ormel et al., 2017; Tamayo et al., 2017; Papaloizou et al., 2018).

Unfortunately, the multiplicity distribution is not directly observable from transit surveys like *Kepler*, because of the extreme biases inherent to the technique. Nevertheless, the unparalleled volume and homogeneous detection biases of the *Kepler* planets still make it arguably the best resource for the task. If we imagine a system of multiple planets around a star, it is likely that only a subset of the planets (if any) will be detected by a transit survey such as *Kepler*, because it is guaranteed neither that all of the planets will be aligned with our line of sight, nor that all will transit at high enough signal-to-noise to be detected. Furthermore, Zink et al. (2019) investigate the detection efficiency of the *Kepler* pipeline and find that it drops in multi-planet systems—specifically, the detection efficiency is higher for the first transiting planet discovered around a star than for subsequent transiting planets in the same system.

Unveiling the true multiplicity distribution from the observed one is therefore a challenging task that needs to account for both geometric and detection biases. Specifically, the mutual inclination distribution between planets should be taken into account in any exploration, since it combines with the underlying multiplicity distribution to produce the observed catalog (Tremaine & Dong, 2012; Brakensiek & Ragozzine, 2016).

One of the first attempts to model the *Kepler* multiplicity distribution is presented by Lissauer et al. (2011a), who test a Poisson multiplicity model and find that the *Kepler* catalog is best fit when the mean of this model is equal to 5.5 planets, and the mutual inclinations of the planets are low. However, they find that this best-fitting model significantly underpredicts the number of single-planet systems observed by *Kepler*. The case for low mutual inclinations in particular has been reproduced in numerous studies (Fang & Margot, 2012; Figueira et al., 2012; Weissbein et al., 2012; Fabrycky et al., 2014). Ballard & Johnson (2016), who study the *Kepler* M-dwarfs, suggest that the under-prediction of single-planet systems can be resolved by introducing a dichotomous population, with one component being a dynamically

cold set of multis and the second being a population of singles or highly-mutually-inclined multis.

This *Kepler* dichotomy, if it extended to FGK stars, may also explain the under-prediction of singles observed by Lissauer et al. (2011a). This has motivated follow-up efforts to determine if there are fundamental differences in the stellar hosts between singles and multi-planet systems. In particular, Munoz Romero & Kempton (2018) search for, but ultimately find no evidence of, metallicity differences among the hosts of the two types of systems, which might be expected if giant planets were responsible for the dynamically hot population.

The lack of any significant metallicity difference has led some to question the dichotomous hypothesis, even as it explains the observed M-dwarf multiplicities better than a single-population model. The drop in detection efficiency for subsequent planets detected in multi-planet systems noted by Zink et al. (2019) hints at another explanation for the overabundance of observed singles, that the *Kepler* pipeline simply fails to detect subsequent planets around some percentage of "singles." Indeed, after accounting for this effect, Zink et al. (2019) find that a modified Poisson model fits the observed multiplicities of *Kepler* GK stars well.

Another alternative explanation to a dichotomous model is a flatter inclination distribution in the inner parts of multi-planet systems than previously assumed—because this inclination distribution works together with the underlying multiplicity distribution to create the observed multiplicities, it is important to model both (Tremaine & Dong, 2012). Bovaird & Lineweaver (2017) show that by adopting a flat-disk model, rather than a "flared" disk, they can match the observed *Kepler* multiplicities without invoking a dichotomous model.

In this work, we aim to address the question of multiplicity by presenting a comparison of several plausible multiplicity distribution models, including both single and dichotomous populations.

We structure this paper as follows. In Section 5.2, we introduce the methods, models and inference approach of this work. In Section 5.3, we present the results, visualizations and analysis of the fits. Finally, we place our work in a broader context in Section 5.4.

Table 5.1: Observed multiplicities in the final subset of 1966 KOIs considered in this work. Taking the sum of each multiplicity by its count yields 1966, as expected.

Multiplicity, m	Counts, $n_{\mathrm{obs},m}$
1	1225
2	218
3	76
4	15
5	1
6	2
7	0
8	0
9	0
10	0

5.2 Methods

5.2.1 Input catalog

We downloaded the *Kepler* DR25 Kepler Objects of Interest (KOIs) catalog via the NASA Exoplanet Archive (NEA; Akeson et al. 2013), with several filters applied. First, we selected only KOIs for which the "Disposition using Kepler Data" was reported as "CANDIDATE". Second, we required that the NEA-reported surface gravity of the star satisfied $\log g > 4$ and that the stellar mass was $0.8 < (M_\star/M_\odot) < 1.2$, in order to focus on FGK dwarfs. Finally, we filtered for planetary candidates which satisfied $6.25 < (P/\mathrm{days}) < 400$ (the same range considered by Petigura et al. 2013 and Foreman-Mackey et al. 2014b) and $0.5 < (R_P/R_\oplus) < 32$. This led to a population of 1966 KOIs, of which the majority were dispositioned as "CONFIRMED".

These 1966 KOIs define our observed data set, $\mathcal{D}_{\mathrm{obs}}$, which comprises three key pieces of information. First, the observed multiplicity distribution, which is simply the occurrence tally of multiplicities from 1 to 10 and is reported in Table 5.1. Second, the list of maximum a-posteriori probability orbital periods, of which there are 1966 elements. Third, the list of maximum a-posteriori probability planetary radii, of which again we have 1966 entries.

After compiling $\mathcal{D}_{\rm obs}$, we also queried all *Kepler* target stars for stars which match the filters imposed above. For this, we took the Mathur et al. (2017) DR25 catalog of stellar parameters, which listed 197,096 stars, and cross-matched these with the CDPP_6 values (combined differential photometry on a 6-hour timescale; see Christiansen et al. 2012) as obtained from the Mikulski Archive for Space Telescopes (MAST). The mean CDPP_6 across all quarters of a given star was saved as the representative CDPP_6. In some rare instances, these values were not available on MAST and thus these stars were dropped, leaving us with 196,792 stars. We then applied the same cuts for $\log g$ and $M_\star$ as described in the previous paragraph, leaving us with 108,429 FGK dwarfs. This catalog of stars will be used later in Section 5.2.3.

5.2.2 Tackling completeness

Our objective is to infer the multiplicity distribution from the *Kepler* catalog. At a very basic level, this objective is challenged by the incompleteness of the *Kepler* catalog itself - just because a star has a planet doesn't mean *Kepler* is guaranteed (or even likely) to see it. A great deal of attention has been paid to this issue in connection to estimating the underlying planet occurrence rate from *Kepler*, and so although our objective is distinct, it is useful to briefly review the approaches used in such studies a source of guidance.

The simplest form of incompleteness to deal with is the geometric transit probability, which decreases with increasing planet orbital radius. In estimates of *Kepler* planet occurrence rates, this can be most easily accounted for by simply dividing apparent occurrence rates by $R_\star/a$ (the geometric transit probability), under the assumption of close-to-circular orbits (e.g. see Howard et al. 2012).

The second, and more challenging, component to completeness is detection efficiency. The simplest solution is to limit one's analysis to a parameter subset where one assumes that the completeness is approximately unity (e.g. Howard et al. 2012; Fang & Margot 2012). This naturally comes at the expense of a smaller sample size. In order to expand the

sample to lower signal-to-noise ratio (SNR) events, it is necessary to estimate the detection efficiency in more detail. A typical approach is the so-called "inverse detection efficiency method" (IDEM)[1], where each planet is assigned a detection efficiency score and ultimately the true occurrence rate is inferred by dividing by both the transit probability and the detection efficiency. Detection efficiencies are typically estimated by injection and recovery exercises (e.g. see Petigura et al. 2013; Dressing & Charbonneau 2015). As an example, for a given choice of orbital period and planetary radius, the associated detection efficiency for a particular star may be computed using the `KeplerPORTs` software (Burke et al., 2015; Burke & Catanzarite, 2017).

In the case of planet occurrence rate estimation, the simplest strategy to account for detection efficiency is the IDEM approach. However, recently Hsu et al. (2018) argue that this approach leads to systemic biases in the inferences since the efficiencies are drawn from estimated planet properties, which are themselves uncertain. Instead, they use a forward-model to inject a population, filter it through a realistic detection efficiency model, and then compare the surviving population to the observed population with some distance metric and Approximate Bayesian Computation (ABC). This approach is shown to more faithfully infer the occurrence rate of the injected population and so a similar approach is adopted here. Rather than comparing occurrences, our work ultimately is interested in the frequency of various multiplicities, but the same approach can be employed (and is discussed further in Section 5.2.4).

In principle, it should be possible to define a detection efficiency model unique to each star using `KeplerPORTs`. However, when conducting Bayesian inference, detailed calculations of these efficiencies for every star and at every step in period and radius comes at high computational cost. A simpler yet still accurate approach is to use a global *Kepler* detection efficiency model, for which one inputs the so-called multiple event statistic (MES) of a planetary candidate and the model returns a detection probability. This is appropriate since

[1] As dubbed by Foreman-Mackey et al. (2014b).

we primarily care about the ensemble rather than individual systems. One example of a global detection efficiency model comes from Christiansen et al. (2016), who use the actual *Kepler* detection pipeline to inject and recover and planets and find that, for FGK stars, the average detection efficiency is well-approximated by either a cumulative gamma function or a logistic function. We use the latter in this work since it is faster to compute. It is given by

$$\Pr(\text{detection}|\text{MES}) = d_l - \frac{d_l}{1 + (\text{MES}/c_l)^{b_l}}, \tag{5.1}$$

where b_l, c_l and d_l are coefficients defined in Christiansen et al. (2016) using *Kepler* DR24. An update to the cumulative detection fraction versus MES is presented in Thompson et al. (2018) for DR25, who find a similar distribution which we use in this work.

The MES (see Jenkins 2002) is a statistic that measures the combined significance of all the observed transits in the detrended, whitened *Kepler* light curves, assuming a linear ephemeris. In practice, it is not feasible to generate very large populations of synthetic planets (required for Bayesian Monte Carlo work), inject their transits, detrend, whiten, and thus compute MES in the same way as the real *Kepler* pipeline.

Instead, we use the transit SNR as a proxy for the MES in what follows. We note that the two are not equal—specifically, the MES depends on the goodness-of-fit between a transit search template and a transit signal mediated by stellar noise, while the SNR does not depend on the template—but they are, to first order, proportional to each other (Burke & Catanzarite, 2017). By combining the SNR with Equation (5.1), we are able to estimate the detectability of any synthetic KOI.

5.2.3 The forward model

Our model works by first choosing a random star from the filtered stellar catalog described in Section 5.2.1. We then inject a planetary system around it composed of m planets, where

m is always less than or equal to $m_{\max} = 10$ and is drawn from a chosen multiplicity distribution as described later in Section 5.2.5. Each of the planets is assigned a random period drawn from a log-uniform distribution from 6.25 to 400 days. A log-uniform distribution was chosen since it both provides a reasonably close match to the observed marginalized period distribution reported by Foreman-Mackey et al. (2014b), and is the same assumption used in previous multiplicity studies, such as Ballard & Johnson (2016).

Next, the innermost planet in the system is assigned a random radius drawn from a double-sided power law (DSPL) distribution, described by

$$\Pr(R) \propto \begin{cases} (\log R - \log R_{\min})^{-\alpha_{\rm small}} & \text{if } R_{\min} < R \leq R_{\rm crit}, \\ (\log R - \log R_{\rm crit})^{-\alpha_{\rm big}} & \text{if } R_{\rm crit} < R < R_{\max}. \end{cases} \tag{5.2}$$

We normalize the DSPL distribution such that the two sides meet at $R_{\rm crit}$ and integrate to unity over the interval $R_{\min} < R < R_{\max}$. The terms $R_{\min}$ and $R_{\max}$ are fixed to $0.5\,R_\oplus$ and $32\,R_\oplus$ respectively, but the parameters $R_{\rm crit}$, $\alpha_{\rm small}$ and $\alpha_{\rm big}$ are treated as unknown shape parameters to be inferred.

To reflect the observed "peas-in-a-pod" covariance of planet radii (Weiss et al., 2018), the radii of subsequent planets in the system are drawn from a Gaussian distribution centred at the innermost planet's radius. The scale parameter of this distribution, σ_R, is treated as another free parameter to be inferred. This parameter is able to extend out to very large values, thereby accounting for the possibility of no correlation (Zhu, 2020).

The simplified double-sided power law radius distribution is designed to capture the turn-over in planet occurrence seen at around mini-Neptune radii, reported in numerous studies (Fressin et al., 2013; Petigura et al., 2013; Foreman-Mackey et al., 2014b). It does not, however, describe the radius valley reported by Fulton et al. (2017). This effect was only revealed by substantial improvements to the precision of measured stellar radii, and we argue that it is not influential enough to significantly affect our study which focusses on multiplicity.

The radius distribution used here essentially represents a set of nuisance parameters which is marginalized over in the final results.

Having generated m proposal planets around the star, with periods and radii drawn from the distributions described above, we next check whether the system is dynamically stable. Following the same approach as Ballard & Johnson (2016), we test for Hill stability using Equation (3) of Fabrycky et al. (2014). Specifically, we define the mutual Hill radius between planets "1" and "2" as

$$R_H = \left(\frac{M_1 + M_2}{3M_\star} \right)^{1/3} \frac{a_1 + a_2}{2}, \tag{5.3}$$

where the Hill stability criterion is satisfied if

$$\frac{a_2 - a_1}{R_H} > \Delta_{\rm crit}, \tag{5.4}$$

where M and a refer to the masses and semi-major axes of the planets. To estimate M for each planet, we use the maximum a-posteriori probability `forecaster` mass-radius relation derived by Chen & Kipping (2017). We compute semi-major axes are from periods using the stellar mass and Kepler's Third Law.

The critical separation is $\Delta_{\rm crit} = 2\sqrt{3}$ for neighboring planets, and for three-or-more planets, Fabrycky et al. (2014) require $\Delta_{\rm inner} + \Delta_{\rm outer} > 18$ for neighboring inner and outer pairs of planets.

If the proposed planetary system violates Hill stability, we use the same star and same multiplicity but make a new realization of the periods and radii for the planetary system. We allow this process to repeat up to 1000 times, after which we abandon the star and draw a new star from the KIC catalog.

After this point we have generated a stable multi-planet system. Next, we need to

calculate how many of these planets actually transit. The innermost planet, labelled with subscript "1", has a transit impact parameter, b_1, of $(a_1/R_\star)\cos I_1$, where I_1 is the orbital inclination angle[2]. Inclination is isotropically distributed and thus we adopt a uniform distribution for $\cos I_1$, which in practice means that we draw a random real number for $\cos I_1$ from $\mathcal{U}[0,1]$, where $\mathcal{U}$ denotes a uniform distribution. The other planets in the system are assumed to have inclinations perturbed away from this angle by an angle ΔI, representing their mutual inclinations within a flared disk (see Bovaird & Lineweaver 2017 for a flat-disk model). ΔI is drawn from a Rayleigh distribution characterized by a scale parameter σ_I:

$$\Pr(\Delta I) = \frac{\Delta I}{\sigma_I^2} \exp\left(\frac{-\Delta I^2}{2\sigma_I^2}\right) \tag{5.5}$$

For each planet in the system, we calculate the impact parameter $b_j = (a_j/R_\star)\cos(I_1 + \Delta I_j)$. Any planet for which $b_j < 1 + (R_j/R_\star)$ is treated as a transiting planet and is saved. Systems with zero transiting planets need not be considered further and are discarded, leading us to draw a new star from the KIC catalog.

At this point, we now have a simulated system of at least one transiting planet orbiting a chosen KIC star. The final component of our forward model is to simulate what fraction of transiting planets in the system would actually be detectable. To do this, we first assign each planet a random transit epoch. Next, we query which quarters that particular KIC star was observed by *Kepler* for, since many stars were not observed in every quarter due to spacecraft rotation and loss of CCDs during the mission.

Using our simulated ephemeris for each planet, we can now calculate how many transits of each planet would have been observed by *Kepler*. We estimate the SNR of each planet using Equation (10) of Kipping & Sandford (2016), multiplied by the square-root of the number of observed transits. Finally, the detection probability is computed using Equation (5.1) from Christiansen et al. (2016). To decide if the transiting planet is detectable or not, we make a

[2]Note that a near-circular orbit is assumed here and throughout.

random Bernoulli draw, with probability equal to this computed detection probability.

This process culminates in a set of m simulated detected transiting planets around a particular star. At this point, we loop back to the beginning of the forward model and keep going until 1966 detected planets have been generated (since this represents the size of the observed sample, $\mathcal{D}_{\rm obs}$, that we will ultimately compare to).

The forward model therefore ultimately yields a simulated data set, $\mathcal{D}_{\rm sim}$, with the same elements and form as $\mathcal{D}_{\rm obs}$. Further,$\mathcal{D}_{\rm sim}$ is clearly dependent upon the simulation's choice of multiplicity, radius and inclination distributions - which are characterized by model parameters $\boldsymbol{\theta} = \{\beta, \alpha_{\rm small}, \alpha_{\rm big}, R_{\rm crit}, \sigma_R, \sigma_I\}$ (where β is a stand-in term(s) describing the multiplicity distribution, described below in Section 5.2.5, and the other terms have been previously defined).

We note that since non-detections are discarded and not counted, our approach does not enable an estimate of the underlying planet occurrence rate.

5.2.4 Comparison to observations

A single run of the forward model described in Section 5.2.3 generates a population of simulated detected transiting planets described by $\mathcal{D}_{\rm sim}$. Our task is now to infer the parameters of the forward model, $\boldsymbol{\theta}$, which would have generated the observed *Kepler* systems, $\mathcal{D}_{\rm obs}$, by comparing $\mathcal{D}_{\rm sim}$ to $\mathcal{D}_{\rm obs}$. In particular, we are interested in inferring the parameters of the multiplicity distribution, β.

Conventional Bayesian inference might proceed using hierarchical Bayesian modeling (HBM), where the multiplicity distribution is described by some parameterized form and then each system's true multiplicity is treated as a free parameter drawn from this overall distribution - giving rise to a large number of unknown variables to solve for (see Hogg et al. 2010 for an astronomer's introduction to HBMs). In this case, the likelihood function used for inference would be well-defined as the product of the likelihoods for each individual system.

Hierarchical models allow for rigorous inference but typically come at great computational expense. Instead, we seek to learn the multiplicity distribution by comparing some distance metric which quantifies how closely the simulated population matches the observed population - thereby ignoring the individual systems and treating the population as an ensemble. By using one or more distance metrics to quantify goodness-of-fit, we are thus conducting what is typically referred to as approximate Bayesian computation, or ABC (see Ishida et al. 2015; Hahn et al. 2017; Hsu et al. 2018; Witzel et al. 2018 for recent applications in astronomy).

The three key ingredients for ABC inference are a forward model which generates $\mathcal{D}_{\rm sim}$, prior distributions for the model parameters, $\Pr(\boldsymbol{\theta})$, and a distance function $\rho(\mathcal{D}_{\rm obs}, \mathcal{D}_{\rm sim})$ which quantifies how well the simulated distribution resembles the observed sample.

Although our primary goal is to learn the multiplicity distribution, we elect to define a distance metric which considers the agreement between the simulated and observed multiplicities but also the agreement between the simulated and observed radius distribution. This is because the two cannot be assumed to be independent: planetary radii determine planetary masses, which in turn determine their stability and whether they could reside in a high-multiplicity system. We therefore infer not only the multiplicity model parameters, but also the radius distribution parameters. Further, the mutual inclination distribution strongly influences the fraction of planets observed to transit and thus is also a parameter we should expect to constrain and be covariant with the other model terms.

Our goal here, of course, is not to infer the true radius distribution of the *Kepler* catalog, nor the distribution of mutual inclinations among its multi-planet systems. However, the known inter-relationships between these terms necessitates that we have some reasonable description of them and that we freely explore them in conjunction with the multiplicity distribution. At the end, we can simply marginalize over these "nuisance" terms in our final calculation of the multiplicity distribution.

Having established that we require a distance metric which incorporates both the mul-

tiplicities and radii, let us consider the multiplicity component first. Previous works have most commonly invoked a Poisson likelihood function in comparing a simulated multiplicity to the observed value (Weissbein et al., 2012; Ballard & Johnson, 2016). This essentially asserts that probability distribution for the observed number of m-planet systems, n_m, is a Poisson distribution with a mean rate given by $n_{\mathrm{sim},m}$. The Poisson model is well-motivated for inference based on counting statistics, such as this, and thus is adopted in this work too. Accordingly, the probability of observing a particular number of m-planet systems, $n_{\mathrm{obs},m}$, is given by

$$\Pr(n_m = n_{\mathrm{obs},m}) = \frac{e^{-n_{\mathrm{sim},m}} n_{\mathrm{sim},m}^{n_{\mathrm{obs},m}}}{n_{\mathrm{obs},m}!}. \tag{5.6}$$

The Poisson likelihood function is defined by a product of the above over all m (=multiplicities), and this function certainly describes how close a simulated set of multiplicities, $\mathbf{n}_{\mathrm{sim}}$, resembles the observed set, $\mathbf{n}_{\mathrm{obs}}$ - thereby providing a suitable distance metric:

$$\rho(\mathbf{n}_{\mathrm{obs}}, \mathbf{n}_{\mathrm{sim}}) = \prod_{m=1}^{M_{\mathrm{max}}} \frac{e^{-n_{\mathrm{sim},m}} n_{\mathrm{sim},m}^{n_{\mathrm{obs},m}}}{n_{\mathrm{obs},m}!}, \tag{5.7}$$

where $\mathbf{n}_{\mathrm{obs}}$ and $\mathbf{n}_{\mathrm{sim}}$ represent vectors containing all the m-indexed observed and simulated population multiplicities, respectively. The $\mathbf{n}_{\mathrm{obs}}$ vector is fixed and given by Table 5.1. Meanwhile, the $\mathbf{n}_{\mathrm{sim}}$ vector, which needs to be counted up after each forward model call, is directly controlled by the choice of forward model parameters $\boldsymbol{\theta}$, which we ultimately wish to infer. It should be noted that our choice of distance metric here, shown in Equation (5.7), does not decrease as the distributions approach one another, but rather increases. Accordingly, it could perhaps be better thought of as an inverse distance metric although we'll continue to refer to it as distance metric in what follows, with the only important consequence being that our task is to maximize ρ, rather than minimize ρ.

We now turn our attention to the component of the distance metric which characterizes the planetary radius distribution. The objective here is not to fit each and every planetary radius - which would be more in line with an HBM. Instead, we wish to simulate a population whose statistical properties broadly match those of the observations. A straightforward approach for accomplishing this is to the use the Kolmogorov-Smirnov (K-S) test, following on from the approach adopted by Fang & Margot (2012).

We therefore compute the K-S p-value between the observed radii and the simulated set as our radius distance metric, since this follows the behaviour of the multiplicity component in terms of being a term we seek to maximize. We multiply this by the multiplicity distance metric given by Equation (5.7) to define an overall distance metric, $\rho(\mathcal{D}_{\rm obs}, \mathcal{D}_{\rm sim})$. The two components are equally weighted under this definition.

A variety of sampling techniques are suitable for ABC inference (Beaumont, 2019), and in this work we elect to use the Markov Chain Monte Carlo (MCMC) approach (Marjoram et al., 2003; Marin et al., 2012). We sample the model parameter space of $\boldsymbol{\theta}$ with Gaussian proposals where the acceptance criterion is chosen such that the probability of accepting a proposal is $\rho_{\rm proposal}/\rho_i$, where i denotes the current index in the chain. In this way, improvements in the distance metric (which recall equates to an increase in the "distance" under our definition) are always accepted. This means that samples near the beginning of the chain, prior to convergence, can often have poor distance scores. We remove these burn-in samples by only including samples in the chain past the first instance exceeding the median distance metric. We demand that 50,000 accepted samples are achieved for each model, with the final chains inspected to verify convergence and mixing.

Because of the somewhat subjective nature of choosing an appropriate distance function in ABC, there is no formal guarantee the model posterior will converge to the true posterior distribution. For this reason, it is important to test the accuracy of our ABC inference framework through fake data generation and recovery simulations, which we present in Section 5.3.2.

5.2.5 Proposed multiplicity models

In Section 5.2.3, we described how exoplanetary radii could be described using a DSPL distribution, mutual inclinations with a Rayleigh distribution, and orbital periods with a log-uniform distribution. However, we did not propose a specific form for the multiplicity distribution itself - which we turn our attention to here.

Specifically, we here describe ten choices of multiplicity distribution: five single-population models, each parameterized by a single free parameter β, and five corresponding "dichotomous" models. The dichotomous models are mixture models of two populations: (1) a fraction f of single-planet systems and (2) a fraction $1 - f$ of multi-planet systems, distributed according to one of the models parametrized by β.

5.2.5.1 Constant model

One of the first multiplicity models proposed in the literature is presented by Ballard & Johnson (2016), who initially adopt a simple approach where every system has the same multiplicity, $\beta_{\rm const}$, which is treated as a free parameter (we refer to this as the constant model):

$$\Pr(m|\beta_{\rm const}) = \begin{cases} 1 & \text{if } m = \beta_{\rm const}, \\ 0 & \text{otherwise }, \end{cases} \tag{5.8}$$

Ballard & Johnson (2016) conclude that this model is unable to provide a good fit to the observed multiplicities of *Kepler* M-dwarf systems and thus expand upon the constant distribution in that same work to include a second component of single-planet systems ($m =$

1), which represent a fraction f of all systems:

$$\Pr(m|\beta_{\rm const}, f) = \begin{cases} f & \text{if } m = 1, \\ 1 - f & \text{if } m = \beta_{\rm const}, \\ 0 & \text{otherwise ,} \end{cases} \tag{5.9}$$

Following the terminology used by the authors, we refer to this as the "dichotomous" constant model.

Simple models are often attractive since one might plausibly purport that the laws that govern planetary architectures (or whatever other phenomenon one is considering) are fundamentally simple themselves. However, we suggest here that the constant model is almost certainly too simple a model to be a realistic description of the exoplanet multiplicity distribution. It is rather implausible to suppose that every system should have an identical number of planets, and even after including a second population of singletons, this still yields a highly unnatural distribution composed of two distinct peaks at $m = 1$ and $m = \beta_{\rm const}$, with zero probability that systems have multiplicities other than this. For this reason, we felt motivated to consider other models in addition to the constant formalism.

5.2.5.2 Uniform model

A simple improvement to consider would be to adopt a discrete uniform distribution, where every multiplicity is just as likely as any other, above some minimum multiplicity, $\beta_{\rm uniform}$:

$$\Pr(m|\beta_{\rm uniform}) = \begin{cases} \frac{1}{(m_{\rm max}+1)-\beta_{\rm uniform}} & \text{if } \beta_{\rm uniform} \leq m \leq m_{\rm max}, \\ 0 & \text{otherwise .} \end{cases} \tag{5.10}$$

This can be similarly be extended to a dichotomous model by assuming that some fraction of planetary systems, f, belong to a separate population of singles.

5.2.5.3 *Truncated Poisson model*

Arguably, a more natural model is a Poisson multiplicity distribution, which appears to be the most commonly adopted law (e.g. see Lissauer et al. 2011a; Fang & Margot 2012; Gaidos et al. 2016; Bovaird & Lineweaver 2017). This might be expected if the multiplicity of an exoplanetary system were the result of a constant rate of generating planets within a fixed interval of time or space. One may write that the probability of forming an m-planet system would thus be

$$\Pr(m|\beta_{\rm poisson}) \propto \begin{cases} \frac{\beta_{\rm poisson}^m}{m!} & \text{if } 1 \leq m \leq m_{\rm max}, \\ 0 & \text{otherwise }, \end{cases} \tag{5.11}$$

where we drop normalization terms which do not depend on m. For $\beta_{\rm poisson} \geq 2$, this implies a peaked, non-monotonic distribution at a specific multiplicity, unlike the uniform case. In practice, we reject any trial m equal to zero or exceeding $m_{\rm max}$, i.e. we truncate the distribution. The Poisson can again be extended to a dichotomous model as was done before.

5.2.5.4 *Exponential model*

Another previously adopted law is that of a discrete exponential distribution (e.g. Bovaird & Lineweaver 2017) which imposes that the multiplicity, m, follows

$$\Pr(m|\beta_{\rm exp}) \propto \begin{cases} \beta_{\rm exp}^m & \text{if } 1 \leq m \leq m_{\rm max}, \\ 0 & \text{otherwise}, \end{cases} \tag{5.12}$$

which can again be extended to a dichotomous model as above.

5.2.5.5 *Zipfian model*

Finally, we consider a Zipfian distribution, which represents a discrete power-law given by

$$\Pr(m|\beta_{\rm zipf}) \propto \begin{cases} m^{-1-\beta_{\rm zipf}} & \text{if } 1 \leq m \leq m_{\rm max}, \\ 0 & \text{otherwise,} \end{cases} \tag{5.13}$$

and a corresponding dichotomous model constructed as above.

Zipf's Law (Zipf, 1935) is known to be an excellent approximation for word frequency versus rank for human languages, and even some animal communications (Doyle et al., 2011a). On this basis, it might seem like a peculiar choice to use when modeling the exoplanet multiplicity distribution, but Zipf's Law also appears in much wider array of problems, such as the population frequency of cities (Auerbach, 1913). Zipf's Law has been argued to be a natural by-product of models with many underlying latent variables, somewhat analogous to the arguments behind the Central Limit Theorem (Belevitch, 1959; Aitchison et al., 2016), and thus on this basis would seem a very reasonable model to propose for exoplanets too - despite the fact it has seemingly not been used in the past for this purpose.

5.3 Analysis

We run the forward model described in Section 5.2.3 ten times, once for each choice of multiplicity model described above, and fit for the free parameters $\boldsymbol{\theta} = \{\beta, \alpha_{\rm small}, \alpha_{\rm big}, R_{\rm crit}, \sigma_R, \sigma_I\}$ (and f, for the dichotomous models) via ABC, as described in Section 5.2.4.

In these fits, we adopt a uniform prior on $R_{\rm crit}$ between $R_{\rm min} = 0.5R_\oplus$ and $R_{\rm max} = 32R_\oplus$, and a uniform prior on f between 0 and 1. We also adopt (improper) priors insisting that $\alpha_{\rm small}$ and $\alpha_{\rm big}$ be greater than -1 and that σ_R and σ_I be positive.

For the constant model, we adopt a prior that restricts $\beta_{\rm const}$ to be an integer between 1 and $m_{\rm max}$, inclusive. For the uniform model, we adopt a prior that restricts $\beta_{\rm uniform}$ to be

Table 5.2: One-sigma credibility intervals of the model parameters for each of the ten multiplicity models. Recall that β is defined differently for each model, as described in Section 5.2.5.

Model	α_{small}	α_{big}	$R_{\text{crit}}\,[R_\oplus]$	$\sigma_R\,[R_\oplus]$	β	$\sigma_I\,[^\circ]$	f
Constant	$0.13^{+0.37}_{-0.26}$	$4.47^{+1.82}_{-1.39}$	$2.91^{+0.88}_{-0.55}$	$0.27^{+0.30}_{-0.18}$	$4.54^{+0.64}_{-0.87}$	$3.95^{+0.76}_{-0.57}$	-
Uniform	$2.27^{+0.42}_{-0.47}$	$3.17^{+2.16}_{-1.39}$	$3.09^{+0.58}_{-0.37}$	$0.19^{+0.21}_{-0.13}$	$1.54^{+0.73}_{-0.36}$	$6.16^{+0.45}_{-0.50}$	-
Poisson	$-0.32^{+0.18}_{-0.22}$	$3.97^{+0.80}_{-0.76}$	$2.77^{+0.28}_{-0.29}$	$0.15^{+0.18}_{-0.10}$	$5.10^{+0.60}_{-0.62}$	$2.18^{+0.16}_{-0.17}$	-
Exponential	$0.54^{+0.48}_{-0.37}$	$4.55^{+1.06}_{-0.78}$	$2.59^{+0.31}_{-0.27}$	$0.14^{+0.18}_{-0.10}$	$1.66^{+0.34}_{-0.26}$	$2.71^{+0.96}_{-0.79}$	-
Zipfian	$0.63^{+0.45}_{-0.37}$	$4.69^{+1.02}_{-0.69}$	$2.62^{+0.30}_{-0.24}$	$0.13^{+0.16}_{-0.09}$	$0.80^{+0.28}_{-0.33}$	$2.28^{+0.87}_{-0.71}$	-
Di-Constant	$-0.22^{+0.23}_{-0.22}$	$3.18^{+0.88}_{-0.63}$	$2.96^{+0.38}_{-0.33}$	$0.11^{+0.18}_{-0.07}$	$5.00^{+0.32}_{-0.37}$	$2.75^{+0.31}_{-0.32}$	$0.72^{+0.04}_{-0.05}$
Di-Uniform	$1.80^{+0.42}_{-0.44}$	$4.24^{+1.03}_{-0.93}$	$2.84^{+0.25}_{-0.25}$	$0.13^{+0.17}_{-0.09}$	$1.77^{+0.63}_{-0.49}$	$4.28^{+0.52}_{-0.48}$	$0.65^{+0.06}_{-0.08}$
Di-Poisson	$-0.36^{+0.19}_{-0.19}$	$4.24^{+1.01}_{-0.81}$	$2.74^{+0.26}_{-0.27}$	$0.12^{+0.16}_{-0.08}$	$4.94^{+0.66}_{-0.62}$	$0.72^{+0.29}_{-0.30}$	$0.53^{+0.04}_{-0.05}$
Di-Exponential	$0.52^{+0.63}_{-0.40}$	$4.74^{+1.06}_{-0.77}$	$2.66^{+0.29}_{-0.28}$	$0.15^{+0.16}_{-0.10}$	$1.55^{+0.42}_{-0.35}$	$2.09^{+1.24}_{-0.95}$	$0.44^{+0.12}_{-0.12}$
Di-Zipfian	$0.82^{+0.50}_{-0.42}$	$4.59^{+0.97}_{-0.79}$	$2.66^{+0.28}_{-0.24}$	$0.15^{+0.16}_{-0.11}$	$0.40^{+0.40}_{-0.52}$	$2.70^{+0.76}_{-0.75}$	$0.23^{+0.22}_{-0.15}$

an integer between 1 and $m_{\text{max}} - 1$, inclusive. For the exponential and Poisson models, we adopt a prior that insists β_{exp} or β_{poisson} be positive. We place no prior constraints on β_{zipf}.

One-sigma credibility intervals for these parameters in each of our ten model fits are presented in Table 5.2, and an example posterior distribution, for the single-population Zipfian model, is presented in Figure 5.1.

5.3.1 Model comparison

To compare models, we use the Akaike Information Criterion, AIC (Akaike, 1974). There are two major reasons behind this choice. First, the AIC does not require that one of the models being tested is the correct model, it merely asks which of the models is the closest approximation to the truth (unlike the BIC; Schwarz 1978). Second, the AIC does not functionally depend on the sample size, which is somewhat ill-defined in our problem since our inference employed a likelihood approximation. We therefore calculate the AIC for each model by first finding the most probable realization from the 50,000 posterior samples, as defined by the distance function, and then using

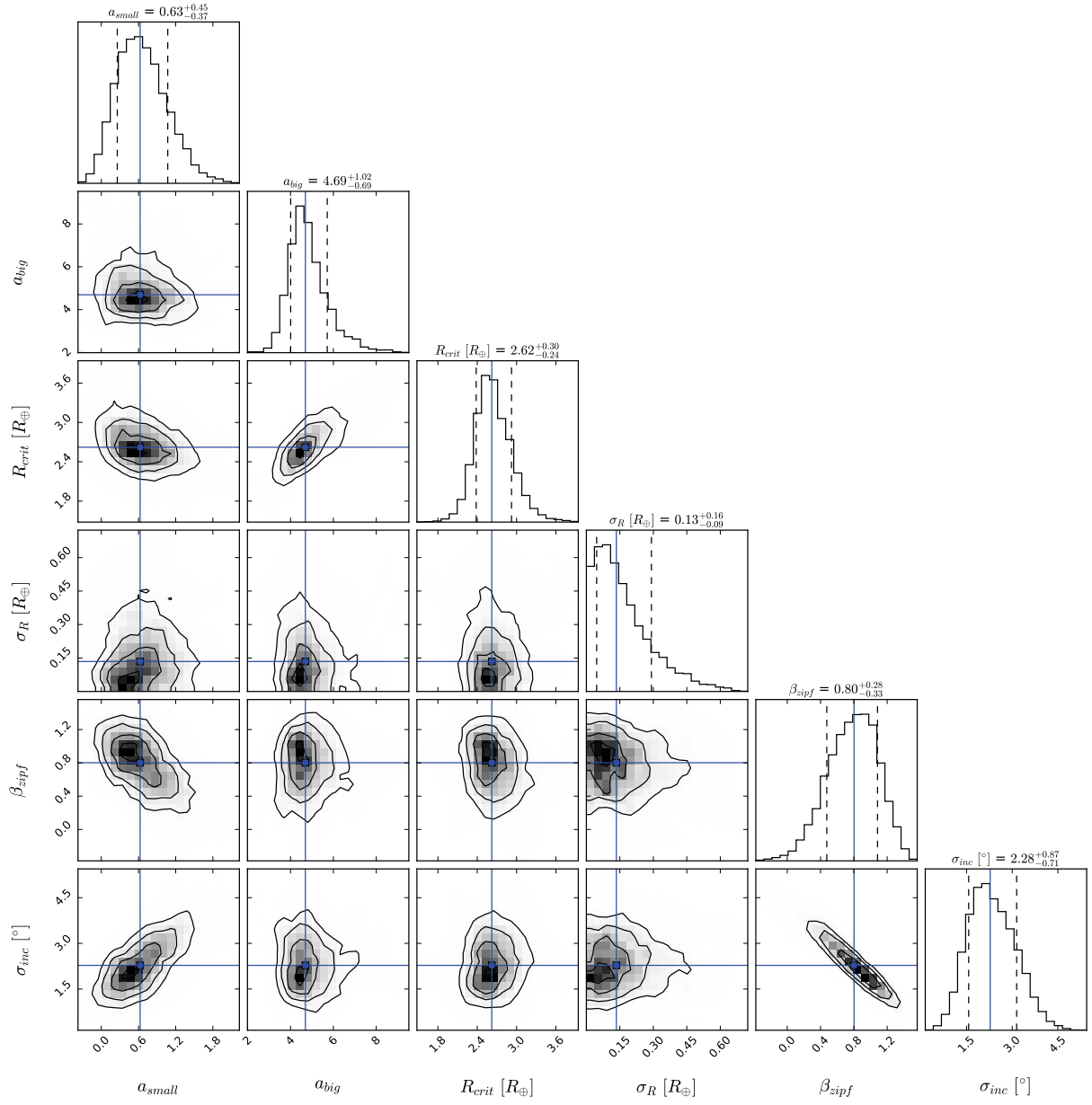

Figure 5.1: Joint posterior probability distribution for the non-dichotomous Zipfian multiplicity model - the favored model deduced in this work. We find that the parameters converge to unique and physically plausible values.

Table 5.3: AIC scores and estimated uncertainties for the ten different models used to describe the *Kepler* exoplanet multiplicity distribution.

Multiplicity Model	AIC
Constant	72.58 ± 1.75
Uniform	64.72 ± 0.86
Poisson	72.15 ± 1.15
Exponential	56.92 ± 0.88
Zipfian	53.79 ± 0.54
Di-Constant	53.99 ± 0.47
Di-Uniform	53.30 ± 0.60
Di-Poisson	54.10 ± 0.61
Di-Exponential	53.85 ± 0.67
Di-Zipfian	54.14 ± 0.74

$$\mathrm{AIC} = -2 \log \hat{\rho} + 2k, \tag{5.14}$$

where k is the number of free parameters used by each model. Since we constructed our distance metric ρ as a product of two likelihood-like terms, ρ approximates the likelihood in the AIC calculation above. For the non-mixture models, $k = 6$ since we have free parameters β, σ_I, R_{crit}, σ_R, α_{small} and α_{big}. The dichotomous models add one extra free parameter, the fraction of single-planet systems f.

We assign uncertainties to our AIC scores through a bootstrapping procedure. First, we split the chain up into S segments. For each segment, we compute the AIC, and estimate its standard deviation as 1.4826 multiplied by the median absolute deviation. We repeat this procedure, varying S from 2 to 50 in unity steps, and for the non-mixture models use the median score across all experiments. For the dichotomous models, we find that the scatter tends to decrease as we approach small S, and thus we fit a simple quadratic model of scatter versus S to estimate the scatter at $S = 1$. These uncertainty estimates, along with the overall AIC scores, are compiled in Table 5.3.

Amongst the non-mixture models, Table 5.3 shows that the Zipfian distribution is pre-

ferred, favored over the next-best model (exponential) with an odds ratio of $e^{(56.92-53.79)/2} = 4.78$. Since all of these models have the same number of parameters, this preference is purely driven by the much improved distance metrics. The constant model is found to be the worst description of the multiplicity distribution, disfavored versus a Zipfian model by a factor of 12,000.

Amongst the dichotomous models, the field is much more level, with all five models roughly equally favorable to each other and also to the single-population Zipfian distribution model. As a check on this, we also tried computing the Savage-Dickey ratio by evaluating the posterior density at $f_{\rm single} = 0$. Only the di-exponential and di-Zipfian models had enough samples around this region to reliably estimate the single-population model:dichotomous model odds ratio, yielding ratios of 0.040 and 1.607, respectively. These are broadly consistent with the AIC results of approximately equal weights ($= e^{-\Delta {\rm AIC}/2}$), demonstrating that the AIC approach is a suitable approximation for this model selection problem.

On this basis, we conclude that the simpler hypothesis of a single population model is not significantly rejected by the current data. A single Zipfian distribution appears quite capable of describing the *Kepler* exoplanet multiplicity distribution for FGK hosts.

In Figures 5.2 and 5.3, we plot the underlying simulated population of planetary systems (inset figures) and the "detected" subset of these systems, for all ten multiplicity models, to investigate the effect of detection biases on this subset and to compare it to the real *Kepler* detections. Despite very different underlying multiplicity models, the detected sample is qualitatively similar in all ten cases: strongly peaked at $m = 1$, and falling off at higher multiplicities. It is not surprising in this light that the Zipfian model performs best of the five single-population models, as it has this general shape already, and that the dichotomous models (which by definition include a peak at $m = 1$) perform equally well.

Consequently, we conclude overall that the current *Kepler* data prefer multiplicity models which peak at $m = 1$, but have little distinguishing power between such models. Choosing among them thus becomes a question of prior beliefs about the underlying planet

distribution—e.g., is there theoretical support for two planetary system formation pathways?

In what follows, we investigate more fully the single-population Zipfian multiplicity model, on the basis of its simplicity.

5.3.2 Testing the inference framework

The use of ABC and also AIC model selection are both approximate tools and thus one might reasonably question how robust they really are. To test this, we decided to generate a total of 20 fake data sets where the true multiplicity distribution is known and test how well we can recover that true distribution using the same machinery used thus far.

We generate the first ten mock populations assuming a Poisson multiplicity distribution. Every population has the same input parameters, chosen to be close to the inferred solutions in our earlier fits, specifically $\alpha_{\rm small} = 0.33$, $\alpha_{\rm big} = 5.0$, $R_{\rm crit} = 2.5\,R_\oplus$, $\sigma_R = 0.05\,R_\oplus$, $\beta_{\rm Poisson} = 1.0$, $\sigma_I = 2.0°$. However, these ten fake "observed" data sets, $\mathcal{D}'_{\rm obs}$, are slightly different to each other due to the stochastic nature of the forward simulation. The second ten are generated in the same way except we switch to a Zipfian distribution (replacing $\beta_{\rm Poisson}$ with $\beta_{\rm Zipf} = 1.0$).

We fit *each* of these twenty data sets with two models: a Poisson and a Zipfian. Thus, we should be able to test whether AIC scoring is able to pick out the correct model in each case - a basic assumption upon which the previous subsection rests. Second, we can test whether the inferred parameters (in cases where the correct model is regressed) are compatible with the input values. In this way, we can provide a detailed assessment of the validity of our inference framework. To save computational time, we use 10,000 post-burn-in steps for each MCMC fit.

The AIC results, summarized in Table 5.4, show that the correct model is identified in 18 out of 20 cases. In general, the Zipfian model appears to be more flexible and gets closer to describing the Poisson model than vice versa, likely as a result of the very harsh selection

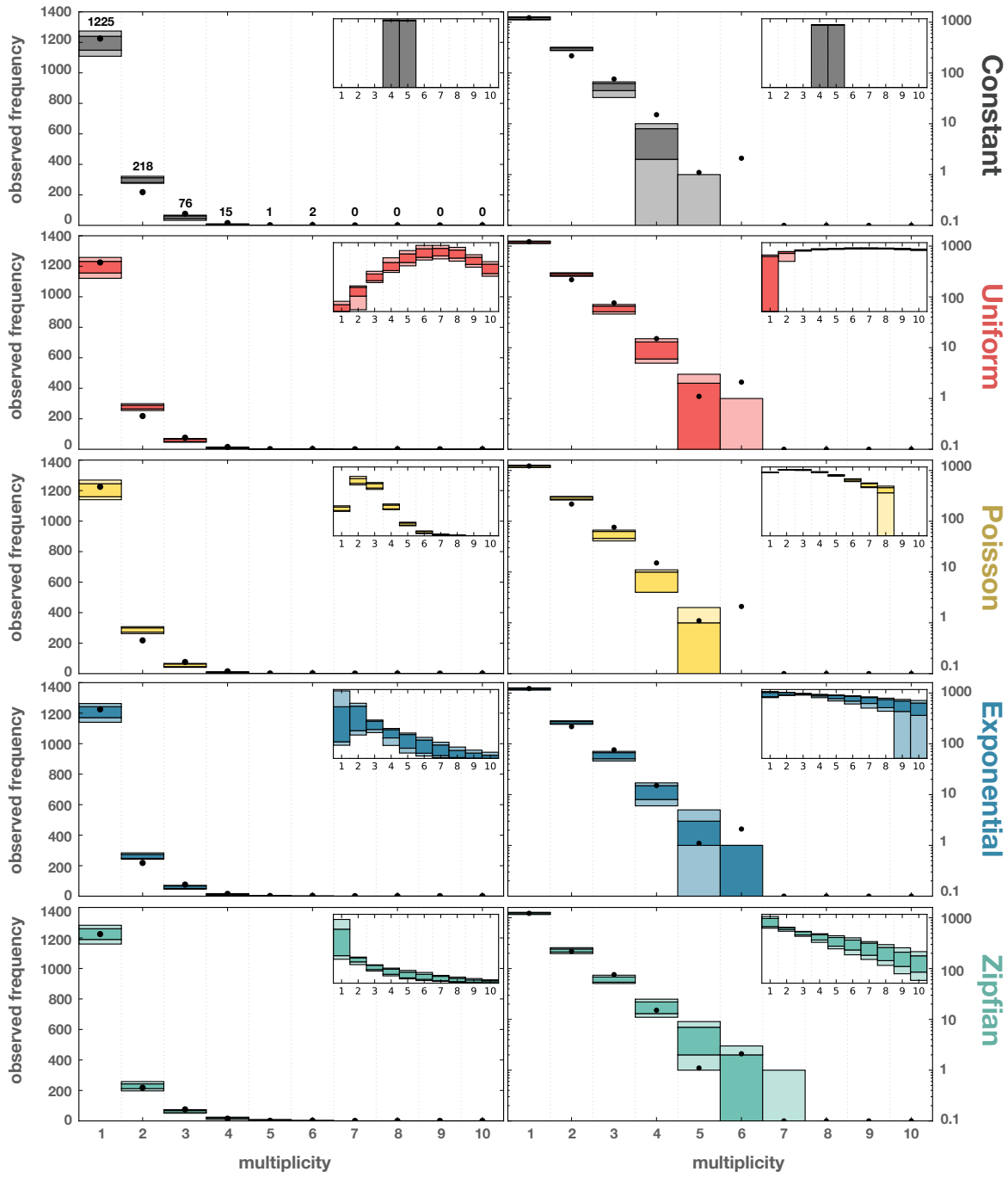

Figure 5.2: Left: Linear-scale histogram of the multiplicities of "detected" simulated planetary systems for *Kepler* FGK stars for the five single-population models. We inset the underlying simulated multiplicity distribution in each panel. The dark regions signify the 1-σ credible interval, and light regions give 2-σ. Black circles represent the real observed *Kepler* sample. Right: Same as left except log-scaled.

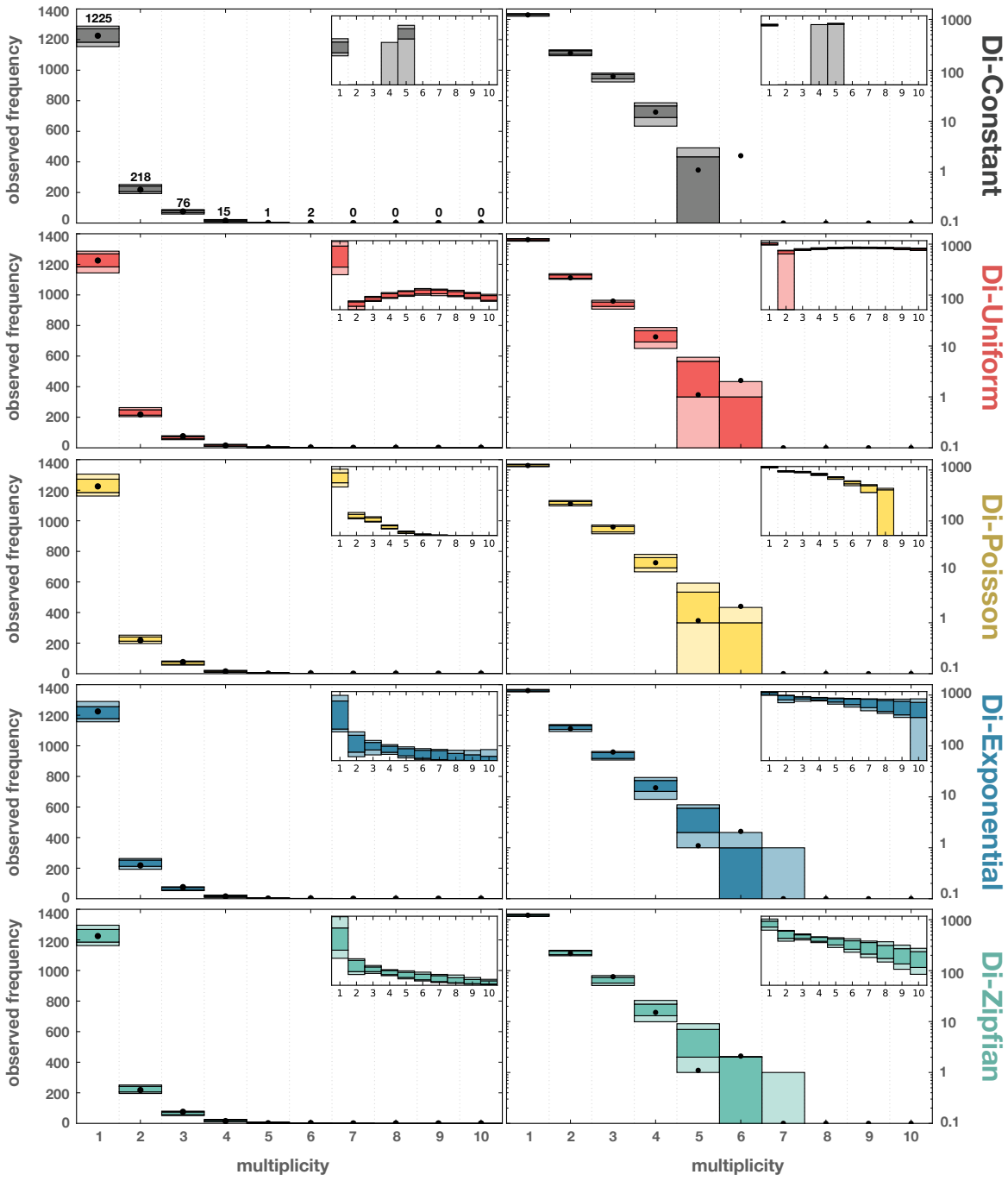

Figure 5.3: Left: Linear-scale histogram of the multiplicities of "detected" simulated planetary systems for *Kepler* FGK stars for the five dichotomous models. We inset the underlying simulated multiplicity distribution in each panel. The dark regions signify the 1-σ credible interval, and light regions give 2-σ. Black circles represent the real observed *Kepler* sample. Right: Same as left except log-scaled.

Table 5.4: AIC scores for twenty fake data sets fitted using two models. Boldened numbers indicate the favored model, which equals the true model in 18/20 cases.

Experiment	AIC (Poisson)	AIC (Zipfian)
Truth = Poisson		
1	**38.73**	40.59
2	**39.22**	39.69
3	**39.78**	40.62
4	39.70	**39.66**
5	**38.78**	39.34
6	**38.45**	42.42
7	38.16	**37.95**
8	**38.93**	40.61
9	**38.50**	38.55
10	**39.17**	40.96
Truth = Zipfian		
1	76.27	**48.02**
2	71.06	**50.32**
3	64.30	**45.90**
4	64.84	**47.91**
5	53.59	**45.97**
6	62.84	**47.51**
7	74.26	**52.27**
8	75.54	**48.42**
9	58.27	**47.00**
10	63.34	**47.01**

functions which push the distributions towards an ostensibly monotonic form. Nevertheless, the AIC scoring system appears to be a reliable tool for identifying the best model.

Comparing the actual parameters which result, as shown in Figure 5.4, we find good agreement between the results and the injected truths: over the 10 fits to the 6 parameters, the recovered parameter is in 1σ (2σ) agreement with the injected parameter in 37 (50) of 60 cases for the Poisson trial and 49 (60) of 60 cases for the Zipf trial. We also inspected the distances metrics versus parameter samples and verified that the distance metrics approach their maximum around the true injected values, as expected. This establishes that the ABC inference framework is able to accurately recover the correct parameters, as well as being suitable for model selection via the AIC.

5.3.3 Properties of the preferred Zipfian distribution

Given that the non-dichotomous Zipfian multiplicity model is the favored model of this work, it is worthwhile to consider the parameters inferred from this model. We show a corner plot of the joint posteriors in Figure 5.1, where the converged, unique nature of the inferred solution is evident.

It is also instructive to compare the multiplicity of the generated systems in our fits, versus the apparent multiplicity of these same systems after being filtered through our mock *Kepler* pipeline. This is is shown in Figures 5.2 and 5.3, where one can see how the true sample (inset figures) is considerably diminished as a result of detection bias.

5.3.3.1 Interpreting the Zipfian slope

When interpreting the inferred value of $\beta_{\rm Zipf}$, it is worth highlighting that $\beta_{\rm Zipf} = -1$ leads to a precisely uniform distribution, $\beta_{\rm Zipf} < -1$ leads to distributions whose probability density monotonically increases with increasing multiplicity, and vice versa $\beta_{\rm Zipf} > -1$ leads to distributions which decrease with increasing multiplicity.

With a shape parameter of $\beta_{\rm zipf} = (0.80^{+0.28}_{-0.33})$, our fit strongly favors a distribution which

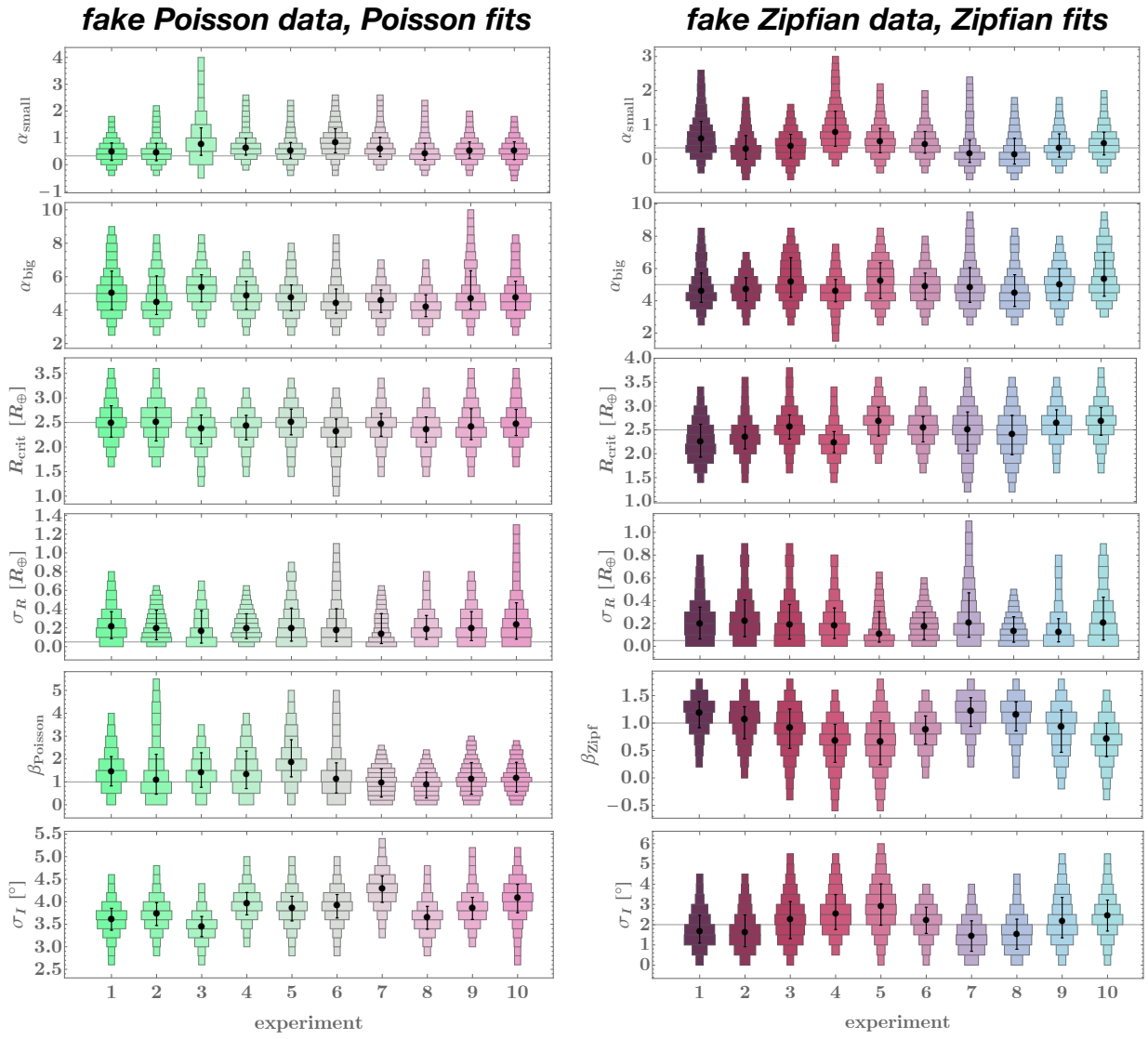

Figure 5.4: Violin plots comparing the retrieved a-posteriori distributions of the five free parameters in our Poisson (left) and Zipfian (right) model, using the described ABC inference framework. Each panel shows a different parameter, with each labeled experiment representing an independently generated fake data set. The injected truth is given by the horizontal lines. and the median and 1-sigma credibility band of the recovered parameter in each trial are plotted in black.

Table 5.5: One-sigma credible intervals of the underlying planet multiplicity for *Kepler* FGK stars with periods $6.25 < (P/\text{days}) < 400$ and sizes $0.5 < (R/R_\oplus) < 32$, as computed from the marginal posterior of the favored non-dichotomous Zipfian model. Quoted scores are defined as the percentage of FGK planetary systems with at least one planet in the quoted period and radius range.

Underlying multiplicity	Credible interval
1	$31.08^{+7.18}_{-8.94}\%$
2	$19.72^{+1.44}_{-2.07}\%$
3	$13.93^{+1.13}_{-1.25}\%$
4	$10.48^{+1.66}_{-1.61}\%$
5	$7.85^{+1.95}_{-1.61}\%$
6	$5.83^{+1.98}_{-1.43}\%$
7	$4.22^{+1.76}_{-1.20}\%$
8	$2.97^{+1.52}_{-0.95}\%$
9	$2.03^{+1.16}_{-0.70}\%$
10	$1.40^{+0.88}_{-0.54}\%$

decreases with increasing multiplicity. This is evident from Table 5.5, which presents the relative frequency of each multiplicity as determined from the Zipfian fit. Since Mercury is less than half an Earth radius and Mars is beyond the period threshold used in this work, the Solar System is a 2-planet system in our framework - a configuration found in $19.7^{+1.4}_{-2.1}\%$ of our simulated FGK planetary systems. Packed, compact systems are rare, with 7-planets or more constituting 10.6% of our simulated FGK planetary systems.

5.3.3.2 Probability of additional planets in known systems

One can also see that although the fraction of observed one-planet systems represents 80% of all systems (see Table 5.1), in reality only 31% of system are truly single (see Table 5.5). Another way to think about this is that the Zipfian model finds that 31% of the simulated systems are genuinely single and all of these must yield a planet with the correct geometry and detectability to have been "detected" by the simulated *Kepler* survey (else they would not have been included in the final simulated catalog since our code would have not saved the realization). Since 100% of the 31% truly single planet systems appear as singletons in the final catalog, $80 - 31 = 49\%$ of the detected singles are multiple planet systems for

which only one planet was detected to transit. Thus, of the 80% of ostensibly single planet systems, a fraction $31/80 = 0.39$ are indeed genuinely single and the other $49/80 = 0.61$ are yet-to-be-revealed multi-planet systems.

Accordingly, radial velocity follow-up of single transiting FGK *Kepler* systems has an a-priori $60.8^{+11.2}_{-8.8}$% chance of detecting new planets (after correctly propagating the uncertainties) with periods and radii in our specified range. Given that there are 1225 single-planet systems in our sample, that equates to 745^{+137}_{-108} hidden planets in the single-planet systems.

5.3.3.3 Total number of missing planets

By calculating the total number of planets generated in the simulated systems, we find that the Zipfian model predicts a total of 4843^{+915}_{-586} planets residing around the 1537 FGK systems with known detections. Since only 1966 known planets reside around these stars, that means that there are 2877^{+915}_{-586} hidden planets - which are expected to be dynamically stable. Discovering these planets, perhaps through radial velocity follow-up, could increase the planet count around these stars by a factor of $2.46^{+0.47}_{-0.30}$.

5.4 Discussion

The principal finding of this work is that the observed multiplicities of the *Kepler* FGK transiting systems can be well-explained without invoking a dichotomous population model. Specifically, we find that a Rayleigh mutual inclination distribution with a Zipfian multiplicity distribution (the latter of which appears to have never been tried before) is able to well-reproduce the observed catalog. This is not to say that dichotomous models are disfavored—indeed, the single-population Zipfian and the five dichotomous models perform equally well—only that invoking a dichotomous population is not necessary to explain the detected *Kepler* multiplicities. Furthermore, we find that the *Kepler* data do decisively prefer multiplicity models peaked at $m = 1$ over those peaked at higher multiplicities.

Bovaird & Lineweaver (2017) also suggest that the dichotomous model may not be necessary by considering an alternative inclination distribution. Since inclination and multiplicity both affect the final catalog (Tremaine & Dong, 2012), then it is certainly plausible then either (or both) of these effects are able to explain the observed multiplicities without invoking dichotomy. Zink et al. (2019) note, furthermore, that the *Kepler* pipeline's decreased detection efficiency for multi-planet systems could also explain the overabundance of *Kepler* singles. Finally, although our work centers on FGK stars, we highlight that Gaidos et al. (2016) also find that a dichotomous distribution may not be necessary by changing the underlying models in the case of M-dwarfs.

It is curious that Zipf's Law (Zipf, 1935), most commonly associated with linguistics, works well for exoplanet multiplicities. Zipfian laws are argued by Aitchison et al. (2016) to be natural outcomes of systems involving a large number of latent variables, and this may represent another example. Extending our analysis to M-dwarfs, particularly from TESS, will provide a good test as to whether the Zipfian model can persist in the face of new data.

Using our preferred model, we are able to make predictions about the numbers of missing planets. For example, we predict that 7 or more planet systems are rare, with just 10.6% of detected systems being so packed. This is in sharp contrast to Mulders et al. (2018), who recently estimated that 42% of Sun-like stars have nearly coplanar planetary systems with 7 or more exoplanets. Although our numbers are not measuring precisely the same quantity, it would be difficult to reconcile the Mulders et al. (2018) value with our estimates given the stark paucity of such systems in our observed sample.

Our model does predict a large number of missing planets, $\simeq 2900$ around the 1537 host stars considered, of which some $\simeq 750$ reside in ostensibly single-transiting-planet systems. It may therefore be possible to test the predictions of these models by conducting radial velocity follow-up of the *Kepler* field in the future to measure the true multiplicities.

Acknowledgements

ES, DK, & MC acknowledge support from the Columbia University Data Science Institute "Seed Funds Program". Thanks to members of the Cool Worlds Lab for useful discussions in preparing this manuscript.

6. On planetary systems as ordered sequences

6.1 Introduction

The planets belonging to any particular system are not random, and neither are they independent. We broadly expect them to have formed from the same protoplanetary disk, collapsed initially from the same cloud, around the same star (Williams & Cieza, 2011). We expect early-forming planets to determine which planets form later. Within our own Solar System, for example, Jupiter and Saturn are thought to have formed early, migrated inward, and truncated the protoplanetary disk at ~ 1 AU, which left relatively little mass to form the terrestrial planets over the subsequent ~ 50 Myr (Walsh et al., 2011; Batygin & Laughlin, 2015). Within exoplanetary systems, we observe correlations between the properties of sibling planets, suggesting similar interdependence (Ciardi et al., 2013; Millholland et al., 2017; Weiss et al., 2018).

After the disk dissipates, we expect the planets to continue interacting dynamically, through orbital migrations, resonances, and scatterings; the present-day configuration of the system, from the multiplicity (Nesvorný, 2011; Sandford et al., 2019b), to their spacing and ordering (Lissauer et al. 2011a; Fabrycky et al. 2014; Weiss et al. 2018; Kipping 2018b; see Winn & Fabrycky 2015 for a review), to their mutual inclinations and consequent co-transiting probability (Tremaine & Dong, 2012; Fang & Margot, 2012; Figueira et al., 2012; Weissbein et al., 2012; Ballard & Johnson, 2016), will naturally depend on this dynamical

history.

The planetary system—the star and its planets, in their specific configuration—encodes its formation and dynamical history; there is information in the individual components, but also in their *arrangement*. It is therefore valuable, especially now that we know of so many exoplanetary systems, to, in the words of Gilbert & Fabrycky (2020), "treat the *system* as the fundamental unit [of exoplanet science]" and investigate planets within the context of their siblings and host stars.

Historically, the relationship of planets to their context has been difficult to investigate because of the combinatoric proliferation of relationships between planets as system multiplicity grows, and because of the lack of an obvious way to compare systems of different multiplicity. Gilbert & Fabrycky (2020) address these problems by defining seven scalar statistics, each of which captures some aspect of the planetary system overall: three relevant examples are the system multiplicity; the monotonicity, which describes how strictly size-ordered the planets are; and the characteristic spacing of the planets in mutual Hill radii.

Here, we take a different approach to studying planetary systems, inspired originally by the study of linguistics. Linguistics concerns itself not just with vocabularies of individual words, but also with their arrangement into meaningful sequences; a particular arrangement of words conveys information which the same words, randomized, do not. Furthermore, any particular language has grammatical rules and conventions that govern these arrangements, which is interesting for two reasons: first, because grammatical logic allows you to make inferences about missing words, and second, because studying the set of allowable arrangements as a whole can allow you to infer the underlying rules.

By analogy, we concern ourselves here with planetary systems as ordered sequences, the structure of which is governed, presumably self-consistently, by the rules of physics and probability. We adapt models from computational linguistics to explore two questions. First, can we infer the properties of a missing or unobserved planet based on its observed context?

If yes, these predictions would be useful in planet searches, to narrow the range of parameter space in which we might expect to find the planet. Second, what can the patterns of planets in systems tell us? Can we infer anything about the underlying "grammatical structure" of planetary systems, if there is such a structure?

In Section 6.2, we describe our input catalog of KOIs arranged in systems. In Section 6.3, we investigate the regression question: how well can we predict the period and radius of an exoplanet based on its context, i.e. the properties of its host star and the periods and radii of its neighbour planets? Furthermore, which contextual information is most important in making those predictions? In Section 6.4, we turn to classification, and apply a model based on linguistic part-of-speech tagging to explore the categories of exoplanet and the "grammatical roles" they play in their systems. In Section 6.5, we conclude and highlight some open questions.

6.2 Input catalog

The goal of this work is to model the relationships between planets and their surrounding system context. For this investigation, we limit ourselves to planets discovered by the *Kepler* mission. We do not expect *Kepler* planetary systems to be complete, because of the inherent biases of the transit method and the detection efficiency of the *Kepler* pipeline (see e.g. Zink et al. 2019), but we nevertheless expect the arrangement of these systems to contain some interesting information.

To construct our input catalog, we downloaded the list of *Kepler* DR25 Objects of Interest (KOIs) from the NASA Exoplanet Archive (NEA; Akeson et al. 2013).[1] We selected planetary systems containing only KOIs with "Disposition Using Kepler Data" of "CANDIDATE" and cross-matched these with the catalog of Chen & Kipping (2018), who predicted masses for the DR25 KOIs using `forecaster` (Chen & Kipping, 2017). We discarded any system which included a KOI without a mass prediction.

[1] Accessed 24 January 2019.

We also discarded any system which contained a KOI with $R_p > 500\ R_\oplus$ (to eliminate unphysically big entries in the catalog), `forecaster`-predicted $M_p > 0.08\ M_\odot$ (the Jovian-star boundary identified by Chen & Kipping 2017), or $P > 3000$ days. After these cuts, our catalog comprised 3690 KOIs, grouped into 2804 systems. These systems range in multiplicity from 1 to 7 KOIs each, where the count of each multiplicity is given in Table 6.1.

We note here that we do not make any effort to account for the selection bias of the transit method in our sample: it is certain that there exist additional planets in these systems which either do not transit or are too small to be detected in transit. For our purposes, this incompleteness does not matter—if we find that we are able to predict the properties of a planet based on its surrounding context, that will remain true even if additional planets are later discovered in the same system. Likewise, any grammatical structure we find in these incomplete systems may be incomplete, but will not be rendered incorrect by further planet discoveries.

We decided to limit our investigation to four salient features of each planetary system: the effective temperature T_{eff} and surface gravity $\log g$ of its host star, and the radius R_p (in units of $R_\oplus$) and period P (in days) of each planet. We choose R_p and P because they are simply interpretable, physically meaningful dimensions. Period (which has the additional advantage of being directly measurable from transit light curves) encodes the ordering of planets outward from their star, as well as some information about their insolation. Planetary radius (although it is not directly observable from transits, and depends on isochrone modeling of the host star) contains information about both composition and equilibrium temperature, and is an axis along which we know there to be meaningful boundaries between categories of planets, such as the small rocky planets and super-Earths on opposite sides of the radius gap at $1.5 - 2.0 R_\oplus$ (Fulton et al., 2017). Likewise, we choose T_{eff} and $\log g$ because they are simple to interpret, capture a lot of information about the host star, and are directly spectroscopically observable.

Because R_p and P span 2 and 3 orders of magnitude, respectively, we take $\log_{10}$ of both.

Table 6.1: Multiplicities in the final subset of 3690 KOIs, grouped into 2804 systems, considered in this work. Taking the sum of each multiplicity by its count yields 3690, as expected.

Multiplicity, m	Counts
1	2211
2	393
3	132
4	47
5	18
6	2
7	1

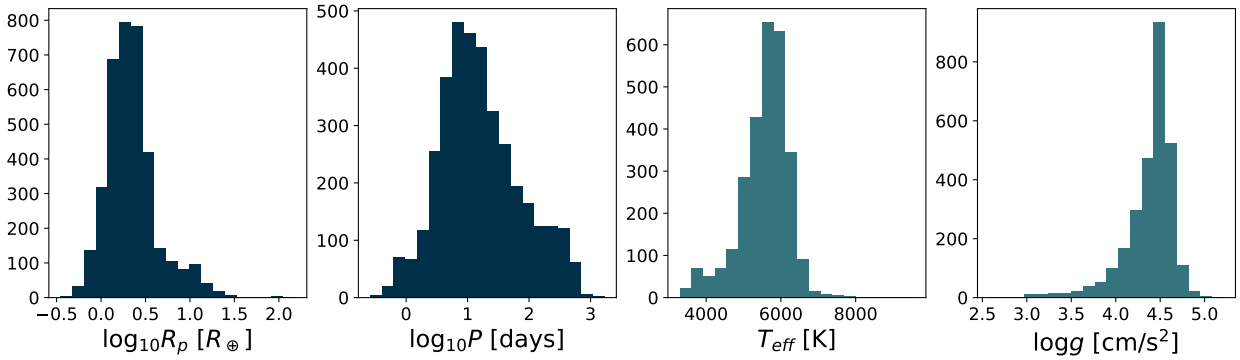

Figure 6.1: Histograms of the two planetary (left) and two stellar (right) features of interest across our catalog of KOIs. The two planetary-feature histograms are over the 3690 planets, and the stellar-feature histograms are over the 2804 systems.

Histograms of these four features are presented in Figure 6.1.

6.3 Regression: Prediction of planet properties from system context

The first question we ask of our KOI catalog is: what can we say about the properties of a "target" planet, based only on its "context"? We define "context," as above, as consisting of the host star's $T_{\rm eff}$ and $\log g$, as well as the radii and periods of the other planets in the system. We seek to constrain the target planet's R_p and P.

6.3.1 Analytic considerations

We begin by considering the analytic constraints we can place on the period and radius, respectively, of a target planet, based on the assumption that the planetary system is dynamically stable. Following Fabrycky et al. (2014), this means that the separation Δ between neighbouring planets in units of mutual Hill radii satisfies:

$$\Delta > 2\sqrt{3}, \tag{6.1}$$

where Δ is defined as

$$\Delta \equiv \frac{a_{\rm outer} - a_{\rm inner}}{R_H} \tag{6.2}$$

and R_H by

$$R_H \equiv \left(\frac{M_{\rm inner} + M_{\rm outer}}{3M_*}\right)^{1/3} \frac{a_{\rm inner} + a_{\rm outer}}{2}. \tag{6.3}$$

Fabrycky et al. (2014) further suggest a conservative heuristic stability criterion for neighbouring pairs of planets, but we do not consider it here.

The goal of this section is to adopt the most conservative possible constraints (in the sense of adopting the fewest further assumptions about the planetary system), yielding the widest physically allowable intervals of $R_{p,\rm target}$ and $P_{\rm target}$. As we will see, the analytic constraints we derive this way are so broad as to be practically irrelevant, so the further assumptions that we do adopt in this section are not important to subsequent sections.

6.3.1.1 Period

Assuming circular orbits, we can bracket the period of any unobserved target planet by the periods of its inner and outer neighbour planets: $P_{\rm inner} < P_{\rm target} < P_{\rm outer}$. (Circularity is a conservative assumption, given that orbital eccentricity of either neighbour planet would

impose stricter bounds on the allowed period range of the middle one.) If a target planet is the innermost in its system, we can similarly bracket its period by

$$\sqrt{\frac{4\pi^2 R_*^3}{G(M_* + M_p)}} < P_{\rm target} < P_{\rm outer}, \tag{6.4}$$

where the inner limit comes from calculating the period of an orbit at a semi-major axis of R_* using Kepler's Third Law.

If the planet is the outermost in its system, we can bracket its period by $P_{\rm inner} < P_{\rm target} \leq P_{\rm max\,in\,data\,set}$, where the latter is 1694 days.

6.3.1.2 Radius

The radius constraint is slightly more involved, and requires us to adopt a mass-radius relationship for the KOIs in our sample. In brief, the procedure for constraining $R_{p,\rm target}$ is:

1. Using the mass-radius relationship, calculate $M_{p,\rm inner}$ and $M_{p,\rm outer}$ from $R_{p,\rm inner}$ and $R_{p,\rm outer}$;
2. Use these to calculate the maximum allowable stable mass of the target planet, $M_{p,\rm target,max}$;
3. Again use the mass-radius relationship to translate this maximum mass into an upper limit on radius, $R_{p,\rm target,max}$.

For steps 1 and 3, we adopt the probabilistic broken power law mass-radius relationship of Chen & Kipping (2017), implemented in the `forecaster` package. We use `forecaster` to translate the NEA-reported 1σ uncertainty range of R_p into a corresponding range of M_p for the inner and outer planets (see Figure 6.2 for examples from the Kepler-1073 (or KOI-2055) system).

Step 2 involves a rearrangement of the stability equations of Fabrycky et al. (2014). For stability, we require that $\Delta > 2\sqrt{3}$ for both the inner-target and target-outer pairs, i.e. that the middle target planet is sufficiently far from both its inner and outer neighbours. In

equation form, it must hold both that:

$$M_{p,\mathrm{target}} < \frac{M_*}{\sqrt{3}} \left(\frac{a_{\mathrm{target}} - a_{\mathrm{inner}}}{a_{\mathrm{target}} + a_{\mathrm{inner}}} \right)^{-3} - M_{p,\mathrm{inner}} \tag{6.5}$$

and

$$M_{p,\mathrm{target}} < \frac{M_*}{\sqrt{3}} \left(\frac{a_{\mathrm{outer}} - a_{\mathrm{target}}}{a_{\mathrm{outer}} + a_{\mathrm{target}}} \right)^{-3} - M_{p,\mathrm{outer}}. \tag{6.6}$$

Since M_* is known, and a and M_p are known for both neighbours, these constraints represent upper limits on $M_{p,\mathrm{target}}$ as a function of a_{target}, which (in accordance with our circularity assumption, above) could take any value between a_{inner} and a_{outer}.

The maximum allowable mass for the target planet *overall*, i.e. anywhere in the allowable range of semi-major axis, is at the intersection of the two constraints, where a_{target} satisfies

$$\left(\frac{a_{\mathrm{target}} - a_{\mathrm{inner}}}{a_{\mathrm{target}} + a_{\mathrm{inner}}} \right)^{3} - \frac{M_{p,\mathrm{inner}}\sqrt{3}}{M_*} = \left(\frac{a_{\mathrm{outer}} - a_{\mathrm{target}}}{a_{\mathrm{outer}} + a_{\mathrm{target}}} \right)^{3} - \frac{M_{p,\mathrm{outer}}\sqrt{3}}{M_*}. \tag{6.7}$$

We adopt this maximum allowable mass, $M_{p,\mathrm{target,max}}$, as our upper mass constraint, and translate it to an $R_{p,\mathrm{target,max}}$ with `forecaster`.

For the innermost planet in any system, we may similarly derive an $M_{p,\mathrm{target,max}}$ based on the outer-planet mass constraint, Equation 6.6. We must also make sure that the planet is small enough not to be tidally destroyed by the star, so we use `forecaster` to translate the upper limit on $M_{p,\mathrm{target}}$ as a function of a_{target} into an upper limit on $R_{p,\mathrm{target}}$ as a function of a_{target}, then calculate the star's Roche limit d as a function of that $R_{p,\mathrm{target}}$:

$$d = R_{p,\mathrm{target}} \left(2 \frac{M_*}{M_{p,\mathrm{target}}} \right)^{1/3}. \tag{6.8}$$

We take $M_{p,\mathrm{target,max}}$ to be the maximum allowed $M_{p,\mathrm{target}}$ at which a_{target} exceeds d.

For the outermost planet, we can adopt the largest allowable mass by the inner-planet mass constraint (Equation 6.5) at the semi-major axis of the maximum period in the data

set,

$$a_{\max} = \left(\frac{G M_* P^2_{\text{max in data set}}}{4\pi^2} \right)^{1/3} . \tag{6.9}$$

In Figure 6.2, we take as an example the first three planets of the Kepler-1073 system, and plot the mass upper limits we derive for the middle planet (Kepler-1073c) based on its inner neighbour (KOI 2055.03) and its outer neighbour (KOI 2055.04). The mass constraints from the inner and outer neighbour as a function of a_{target} are plotted as blue and red curves, respectively, with the corresponding excluded regions shaded in blue and red. The true mass and semi-major axis of the middle planet are plotted as a black point.

This system is representative of our results for the KOI data set overall: in practice, the upper limit on $M_{p,\text{target}}$ that we derive by this procedure—and the corresponding upper limit on $R_{p,\text{target}}$—is orders of magnitude larger than the true value, and indeed orders of magnitude too large to be practically relevant for our prediction problem.

Consequently, having explored this analytic avenue, we drop it for the remainder of this work.

6.3.2 Machine learning approach

We next explore another approach to predicting the properties of a planet based on its system context: training a neural network model. A diagram of this approach is presented in Figure 6.3.

We begin by dividing our data set of 2804 *Kepler* systems into a training set of 1962 systems (70% of the total) and a test set of 842 systems (the remaining 30%).

For each *target planet* in the training set, the network's goal is to take as input the planet's context, and provide as output a prediction of the log planetary radius $\log_{10} R_p$ and log period $\log_{10} P$, which can then be compared to the true values.

Specifically, the network's contextual input consists of:

1. the stellar T_{eff} and $\log g$;

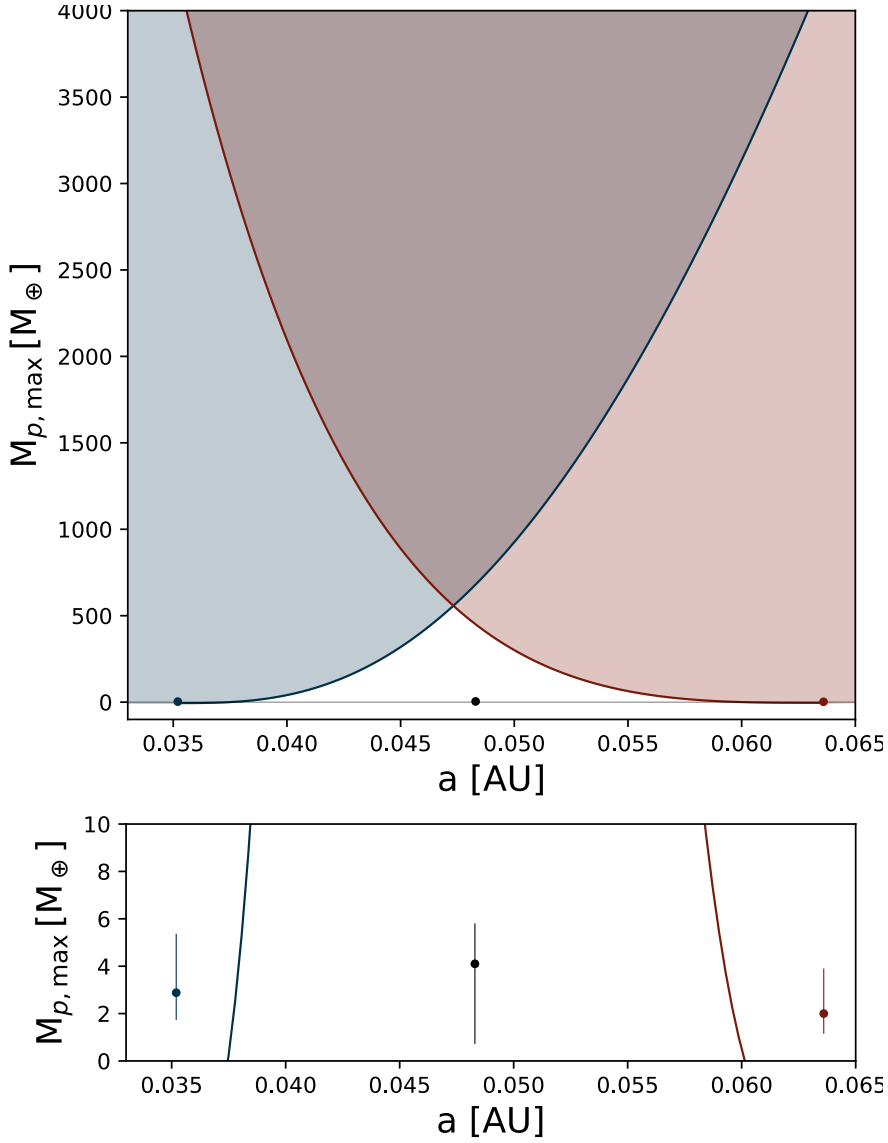

Figure 6.2: An investigation into which regions of parameter space are excluded by simple analytic stability criteria for representative planet Kepler-1073c, which is bounded by inner planet KOI 2055.03 and outer planet KOI 2055.04. Top panel: Masses above the blue line are excluded by the inner planet; masses above the red line are excluded by the outer planet. Lower panel: A zoomed-in view, showing the Forecaster-predicted 1σ mass ranges for all three planets based on their NEA-reported 1σ radius ranges. In practice, the analytic upper mass limits are generally much too large to meaningfully constrain the middle planet's mass.

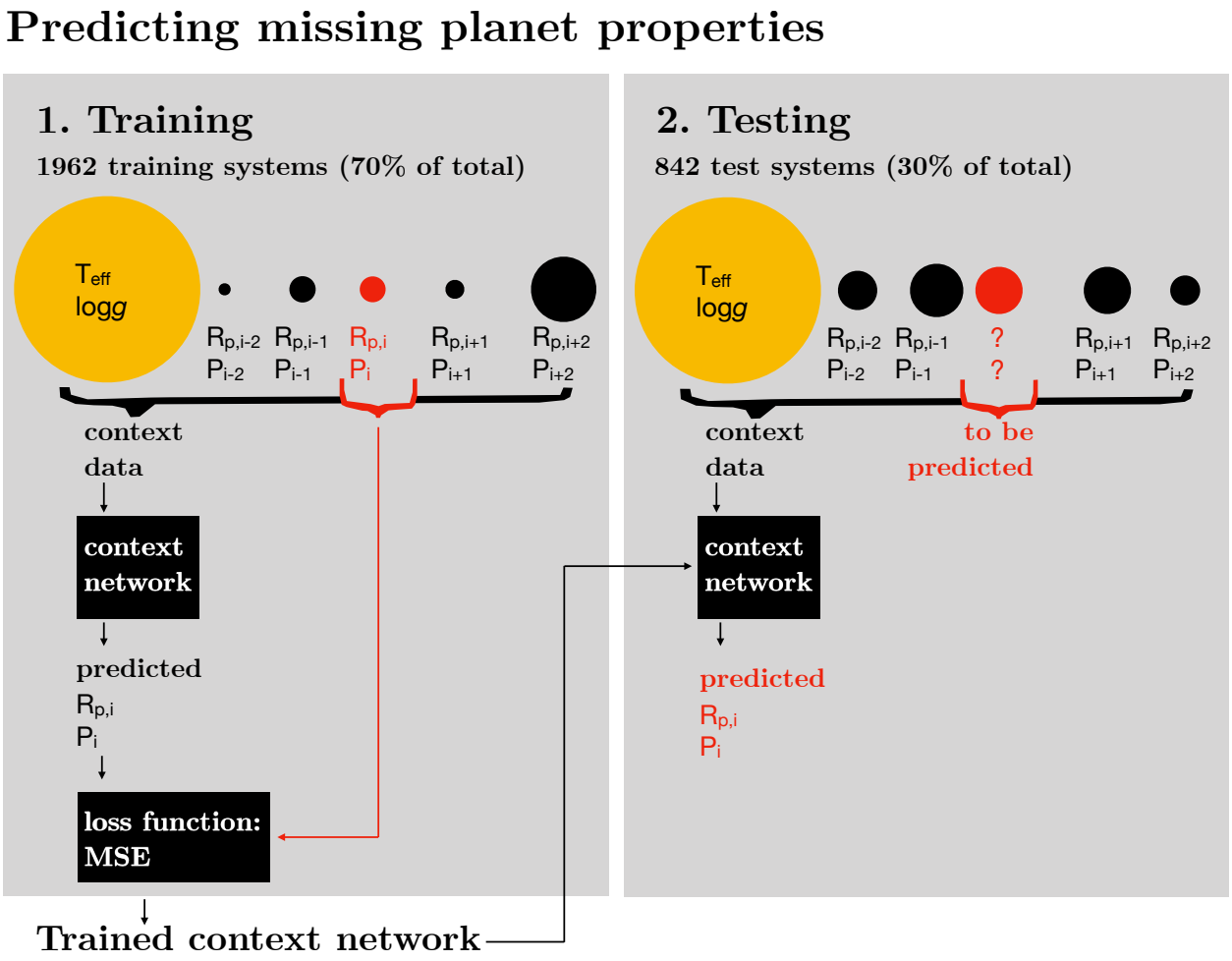

Figure 6.3: A diagram of the machine learning approach to predicting the properties of an unobserved middle planet bounded by four observed planets.

2. a vector of $\log_{10} P$ of the w planets immediately inner to the target planet and the w planets immediately outer to the target planet, where w is a network hyperparameter called the context width;

3. a corresponding vector of $\log_{10} R_p$ of these $2w$ neighbour planets.

In other words, for each target planet, the network sees the planet's host star and the neighbouring planets within a window of width $2w$ centered on the target planet. Of course, because the maximum multiplicity of systems in our data set is $m = 7$ and these systems are heavily skewed toward low multiplicity, most planets will not have w inner and w outer neighbours, even for $w = 1$. In this case, the period and radius vectors are zero-padded such that they still have length $2w$. Based on testing and considering the tradeoff between capturing more information about high-multiplicity systems (high w) and training speed (low w), we choose $w = 2$.

Our network is implemented in `pytorch`. It has two fully-connected hidden layers of width 10 and 5, respectively. For the hidden layers, we use a rectified linear unit (ReLU) activation function; for the output layer, which needs to generate real-valued predictions of arbitrary sign, we use a linear activation function.

The cost function the network minimizes over training is the mean squared error:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} |\boldsymbol{X}_{\rm true} - \boldsymbol{X}_{\rm pred}|^2, \tag{6.10}$$

where $\boldsymbol{X}_{\rm true}$ is the vector $[\log_{10} R_{p,\rm true},\ \log_{10} P_{\rm true}]$ for the target planet, $\boldsymbol{X}_{\rm pred}$ is the corresponding network prediction, and N is the number of planets in the training set (or, in practice, training batch). We note that this cost function does not account for the observational uncertainty in $\log_{10} R_{p,\rm true}$ and $\log_{10} P_{\rm true}$: in other words, we are training the network to predict the maximum-likelihood values of $\log_{10} R_{p,\rm true}$ and $\log_{10} P_{\rm true}$, not their full credible intervals.

We train the network with an `Adam` optimizer (Kingma & Ba, 2014) with learning rate

0.002, although this hyperparameter has minimal effect on the training efficiency. We train the network in batches of 100 training planets over 500 epochs, which (by visual inspection of cost vs. epoch number) is more than sufficient for the network cost to decline to a roughly constant value. To all neurons, we impose a dropout probability of 1% at every training epoch to discourage overfitting.

We train the network 100 times with different random initializations of the neuron weights and take the version which achieves the minimal MSE on our training set as our result. This optimal network's predictions for the 842 *test* systems, grouped by system multiplicity, are shown as yellow points in Figure 6.4.

For comparison, we also plot the "predictions" of an extremely naive model: for each target planet, we "predict" the radius and period by drawing 100 random training set planets which satisfy the basic stability criterion $P_{\rm inner} < P_{\rm draw} < P_{\rm outer}$. If the target planet is the innermost (outermost) of its system, we adopt as $P_{\rm inner}$ ($P_{\rm outer}$) the minimum (maximum) period of any planet in our training set. We plot the 1 and 2σ confidence intervals over the 100 draws as gray bars in Figure 6.4. We note that in the period-prediction scatterplots, you can clearly pick out the planets which are innermost (outermost) in their systems, as these are the planets for which the gray bars extend all the way to the lower (upper) limit of the plotted period range. The gray bars for all of the one-planet systems, for example, fill the entire period range (with some statistical noise that results from our sample size of 100 draws).

As Figure 6.4 shows, the network predictions are no better than random for the vast majority of the 1-planet systems, but improve dramatically with system multiplicity. This indicates that the network is learning primarily from the other planets in the system, not from the stellar features.

For period, especially, it makes sense that the network is learning mainly from the context of neighbouring planets—after all, the network is directly given the bracketing periods of the neighbour planets. However, the network does better than the naive model, which relies only

on period bracketing—so the network is learning something deeper from the arrangement of planets in the training systems. The network does better than the naive model for middle planets as well as inner/outermost planets, meaning that the network's improvement over the naive model does not result purely from our choice to bracket the inner/outermost planet periods by P_{minimum} and P_{maximum}.

The network's radius predictions, while less tightly scattered around the true values than the period predictions, are also markedly better than those of the naive model. Interestingly, unlike with period, there is a small subset of high-radius single-planet systems for which the network is able to make a meaningful prediction: there is some stellar information which correlates with radius for this subset. Again, the radius predictions generally improve with system multiplicity.

A summary of the network vs. naive model performance is presented in Figure 6.5, which shows the mean absolute error (MAE) in both radius and period as a function of system multiplicity, transformed into linear R_p and P space, for the trained network and the naive model. The mean absolute error is defined:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |x_{\text{true}} - x_{\text{pred}}|, \tag{6.11}$$

where N is the number of planets, x_{true} is the true value of R_p (or P), and x_{pred} is the predicted value of R_p (or P). We choose to plot MAE in linear space to make it easier to judge the accuracy of the two models' predictions in meaningful units.

Because the mean absolute error is already normalized by the number of planets, we can compare the different multiplicities in this plot directly. Both the network and the naive model's R_p predictions become more accurate for multiplicities up to $m = 5$, then worsen for $m = 6$ to 7; this could be the result of the very small number of $m \geq 5$ systems in the data set (and one of the two 6-planet systems, as well as the lone 7-planet system, were by chance assigned to the test set, not the training set, so the network did not train on either).

In period space, there is an equally dramatic uptick in MAE beginning at $m = 5$.

Despite this uptick in MAE for high multiplicities, the network consistently outperforms the naive model: its radius predictions are better on (weighted) average by a factor of 2.4, and its period predictions by a factor of 1.8.

6.4 Classification: The grammar of planets and planetary systems

We next turn to the question of planet grammar. In this section, we adopt a model used for part-of-speech tagging in computational linguistics to ask: do planets fall into natural categories with different roles in their systems, as words can be categorized into parts of speech? If these categories exist, can we deduce them from the set of observed planetary systems? Finally, what can we infer about the underlying rules governing the arrangement of planetary systems?

We begin by exploring the analogous linguistic question that inspired this work. What is part-of-speech tagging, why is it useful, and how is it done?

6.4.1 A model for our model: part-of-speech tagging

Part-of-speech tagging is the problem of categorizing the words in a body of text by their grammatical function or behavior (see e.g. Greene & Rubin 1971; Brown et al. 1992). In the preceding sentence, for example, "is" is a verb; "tagging", "problem", "categorizing", "words", "body", "text" "function", and "behavior" are nouns, etc. This task is deceptively simple: often, the part of speech of a word changes depending on its context, so it has no single true classification. For example, in other sentences, "words", "text", and "function" can all be verbs, as in, "She words the sentence carefully," "I will text you later," or "The computer cannot function."

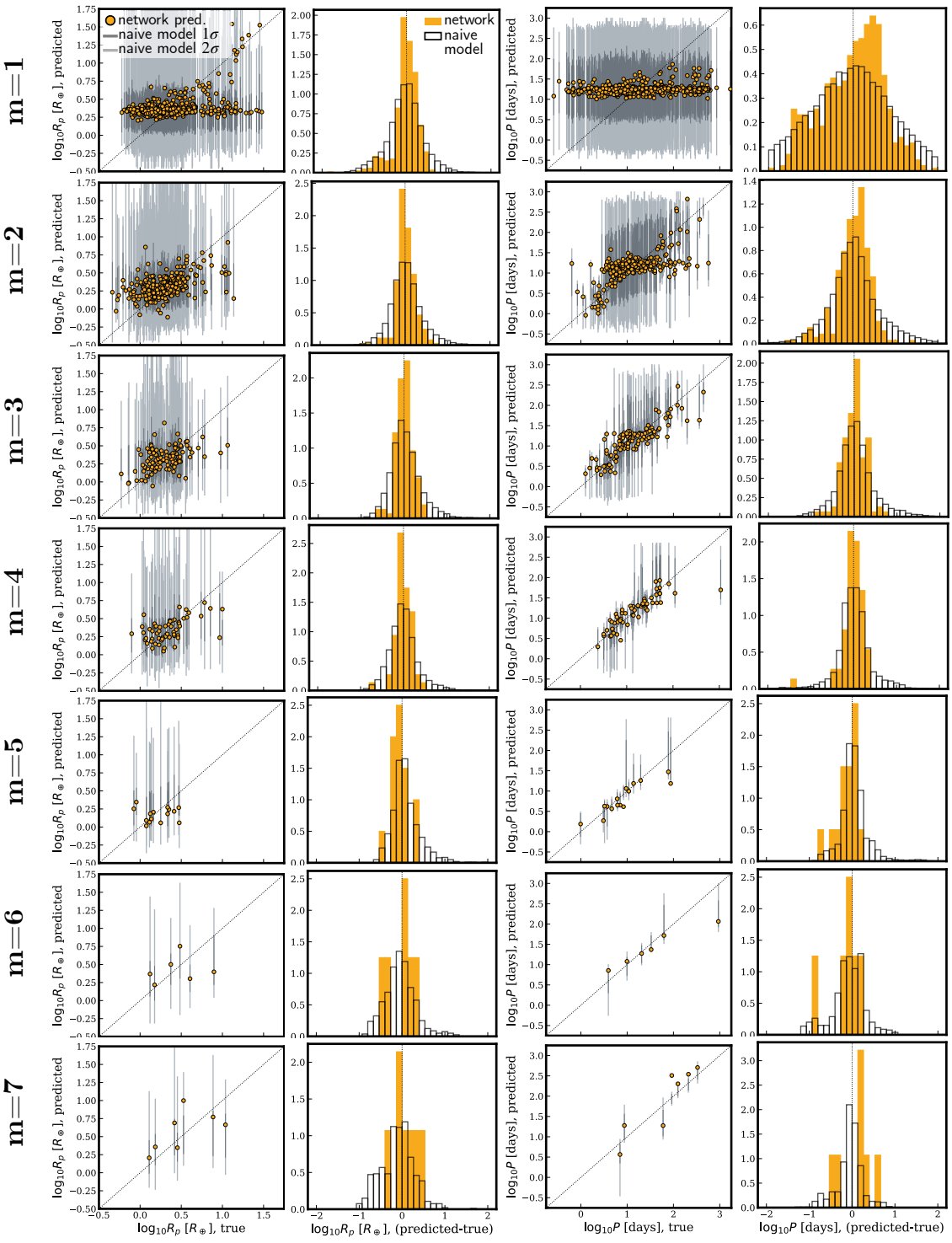

Figure 6.4: The performance of the trained regression network on the 842 test systems. Scatterplots: Network-predicted planet properties vs. true values. In yellow are the network predictions; in dark (light) gray are the 1σ (2σ) confidence intervals of the "naive" model (see text). The corresponding histograms show distributions of (predicted value - true value) for the network (yellow) and naive model (gray).

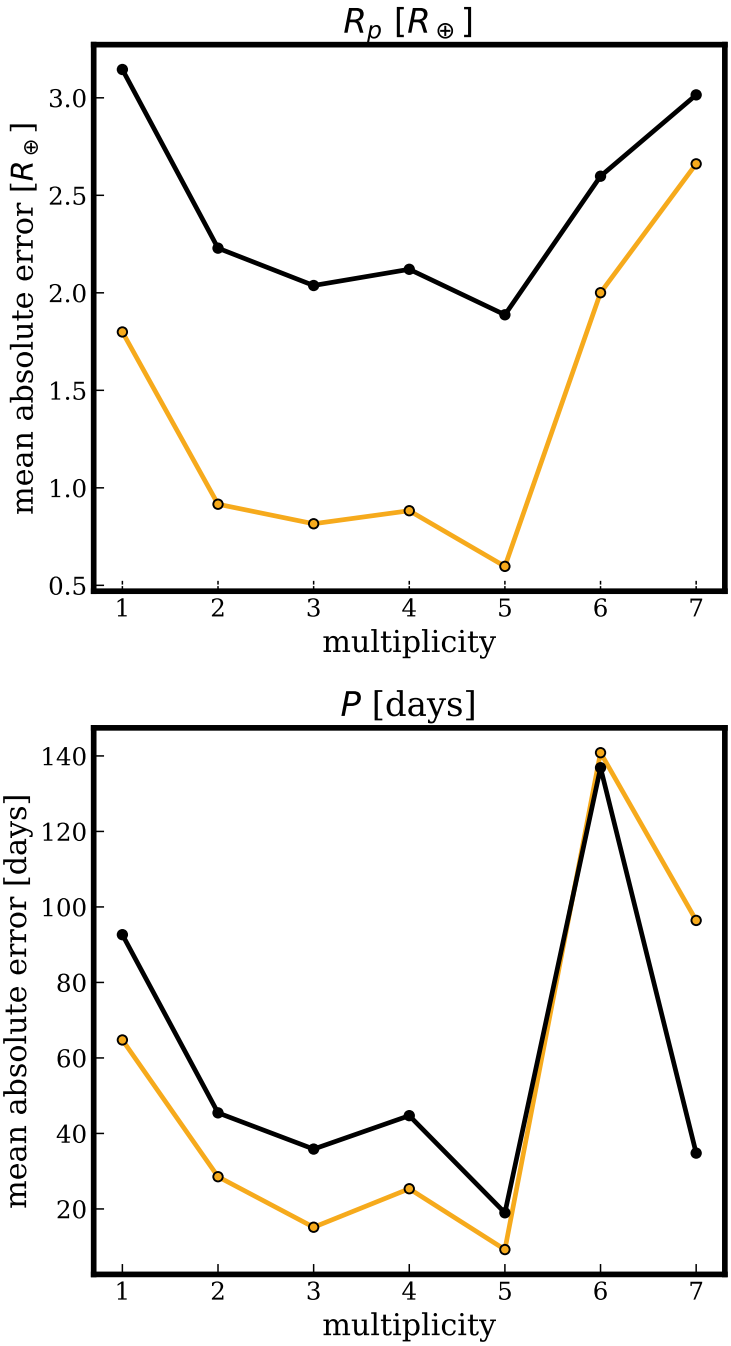

Figure 6.5: The mean absolute error per planet of the regression network (yellow) and naive model (black) predictions, as a function of multiplicity, transformed into linear R_p and P space. The network consistently outperforms the naive model, by (on average, weighted by the number of planets per multiplicity) a factor of 2.4 in radius and 1.8 in period.

Part-of-speech tagging is interesting and useful because it represents sentences abstractly. It captures the patterns of words organized according to a grammar, independent of the choice of words in any particular sentence. For example, "The bird sings" and "The plot thickens" both follow the same part-of-speech pattern, "article noun verb," despite very different meaning. The abstraction to "article noun verb" reveals the common structure.

Tagging an entire corpus of English text in this way would reveal several common patterns: "noun verb" if the sentence ends in a period, but "verb noun" if it ends in a question mark; verbs and adjectives allowed to occur in sequence, but not prepositions. These patterns in turn reveal the underlying grammatical rules of the language. Grammar induction has many further linguistic applications, including translation of dead languages, machine translation of extant languages, and machine reading comprehension.

Over the past two decades, computational linguists have made great strides in *unsupervised* part-of-speech tagging (see Christodoulopoulos et al. 2010 for a review), where a model learns to identify the correct parts of speech without training on a data set labeled with the correct answers. Because grammatical patterns in sentences exist regardless of the content of those sentences, it is not especially necessary to train part-of-speech tagging models on a labeled training set; part-of-speech tagging can be treated more like a clustering problem (e.g. Brown et al. 1992). For our purposes, where we do not have "correct" planet labels to train on, unsupervised learning is ideal.

Here, we investigate a particular unsupervised machine learning technique introduced by Stratos (2019), based on theoretical work by McAllester (2018), and adapt it to classify planets in systems. This technique, called mutual information maximization, simultaneously trains *two* models to predict the part-of-speech (class membership) of any word (planet). The first model, called the *target network*, takes as input the target word (planet); the second model, called the *context network*, takes as input the surrounding context words (star and neighbouring planets). Both networks output a prediction of the target's class membership, and the cost function rewards agreement between these predictions. In other words, the

goal is to maximize the mutual information—the amount of information learned about one variable by measuring another—between the target network's representation of the target (based on itself) and the context network's representation of the target (based on its context).

6.4.1.1 The mutual information cost function

Here, we explore this cost function in slightly more detail. This will be a simplified treatment; for the full derivation and information-theoretic justification, see Stratos (2019).

Following the derivation of Stratos (2019) section 3.2 (and adopting the same notation), let us label our target planet as $x \in X$, where X is the space of all possible planets, and our planet context as $y \in Y$, where Y is the space of all possible contexts. If we aim to classify the planets into N_{classes} classes, our goal is to train the target network to assign a class label $z \in 1, ..., N_{\text{classes}}$ to the target, to train the context network to assign a class label $z \in 1, ..., N_{\text{classes}}$ to the target, and for those assigned labels to agree as far as possible.

Let us call the output of the target network $p(z|x)$. This is a vector of probabilities of length N_{classes}, where e.g. $p(1|x)$ is the probability that planet x belongs to class 1, etc. Similarly, let us call the output of the context network $q(z|y)$.

Let us assume first that we know $q(z|y)$ and we wish to train the target network such that $p(z|x)$ matches it as well as possible. We can reward that agreement by minimizing the cross-entropy $H(q, p)$:

$$H(q, p) = \mathrm{E}[-\sum_{z} q(z|y) \log p(z|x)], \tag{6.12}$$

where the expectation value is over all target-context pairs (x, y) in the training set. However, we do not actually know $q(z|y)$, and rewarding agreement alone would allow the two networks to trivially agree by e.g. assigning all planets to the same class. So we must simultaneously penalize that behavior, i.e. reward assigning examples to different classes.

This can be achieved by maximizing the entropy of $q(z|y)$ over the different choices of label z. First we marginalize $q(z|y)$ over all y, so we are left with the distribution of labels $q(z)$. Let Z represent a random draw from this distribution. We wish to maximize the

entropy of Z:

$$H(Z) = -\sum_z q(z) \log q(z). \tag{6.13}$$

The overall cost function to be *maximized* is then -1 times the cross-entropy term, to reward agreement, plus the label entropy term, to penalize trivial agreement. Mathematically, we wish to maximize:

$$J = H(Z) - H(q, p). \tag{6.14}$$

6.4.1.2 Network design and performance on simulated data set

To apply this mutual information maximization framework to our planets, we first design the target and context networks, then test them on a (highly artificial) data set of simulated planetary systems. A diagram of our classification model is presented in Figure 6.6.

The target and context networks (again implemented in `pytorch`) have the same internal architecture, but with different inputs. The target network takes as input the vector $[\log_{10} R_{p,\text{target}}, \log_{10} P_{\text{target}}]$. The context network (as in Section 6.3.2) takes as input the stellar features T_{eff} and $\log g$, as well as a length-$2w$ vector of the $\log_{10} R_p$ of the w inner and w outer context planets, and a corresponding length-$2w$ vector of the $\log_{10} P$ of these planets. As in the regression problem, we adopt a context width $w = 2$ and zero-pad the context vectors where there are fewer than $2w$ context planets.

Both networks have two fully-connected hidden layers of width 100 and 10, respectively, with ReLU activations. The output layer is of size N_{classes}, and we apply a softmax function to its output to transform it into a vector of class membership probabilities (i.e. the first entry is the probability that the particular training example belongs to class 1, etc.). We again train the networks with an `Adam` optimizer with learning rate 0.002 in batches of 100 training examples over 500 epochs, from 100 different random initializations of the neuron weights. To all neurons, we again impose a dropout probability of 1% at every training epoch to discourage overfitting.

To test this setup, we simulate a data set of planets that follow an invented toy planetary

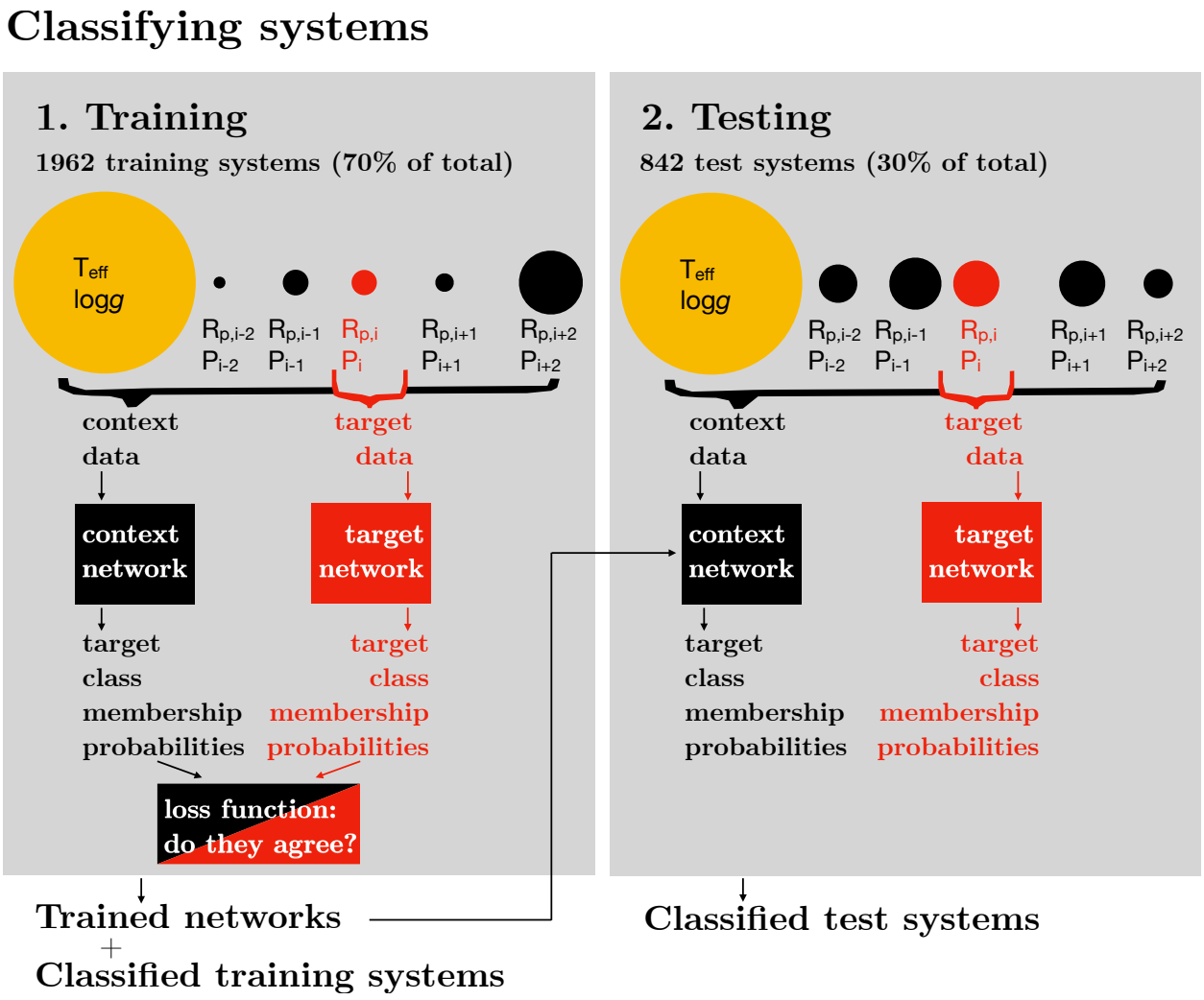

Figure 6.6: A diagram of the machine learning approach to classifying planetary systems, which is adapted from the part-of-speech tagger of Stratos (2019).

grammar. A diagram of allowed planetary systems under this grammar is shown in the top panel of Figure 6.7.

In the simulated data set, planets belong to three classes: small, or class A ($R_p = 1R_\oplus$); medium, or class B ($R_p = 5R_\oplus$); and large ($R_p = 10R_\oplus$), or class C. The radius distributions are extremely narrow, and there is no overlap in radius between the classes (in other words, a planet's radius determines its class exactly). In our invented grammar, small planets act like articles (e.g. "the" or "a"), medium planets act like compoundable nouns (a single medium planet could represent the word "planet," a sequence of two medium planets could represent the phrase "planety planet," and so on), and large planets act like verbs (e.g. "is").

Allowable planetary systems under this grammar include $[A, B, C]$ ("the planet is"), $[A, B, B, C]$ ("the planety planet is"), $[A, B, B, B, C]$ ("the planety planety planet is"), and $[A, B, B, C, A, B, B]$ ("the planety planet is a planety planet"), but *not* $[A, A, A]$ ("the the the") or $[C, B, A, C]$ ("is planet the is").

To generate a simulated system, we first draw a random star from the catalog of *Kepler* planet hosts satisfying $3 < \log g[\mathrm{cm/s^2}] < 5.3$, $2400K < T_{\mathrm{eff}} < 9600K$, and $R_* < 10R_\odot$. (These cuts are somewhat arbitrary, because our toy grammar does not depend on stellar properties.) We then draw a system multiplicity m from a Zipfian distribution with maximum multiplicity $m_{\mathrm{max}} = 10$ and index $\beta = 0.80$, as found to be the best-fit model to the multiplicity distribution of *Kepler* multis in Sandford et al. (2019b):

$$\Pr(m) \propto \begin{cases} m^{-1-\beta_{\mathrm{zipf}}} & \text{if } 1 \leq m \leq m_{\mathrm{max}}, \\ 0 & \text{otherwise,} \end{cases} \tag{6.15}$$

We then draw periods for the m planets in the system from a uniform distribution between 6.25 and 400 days. (Because class A planets often begin systems, and class C planets often end them, a planet's period contains weak information about its class membership.) We then populate the radii according to our grammar rules, and check for dynamical stability

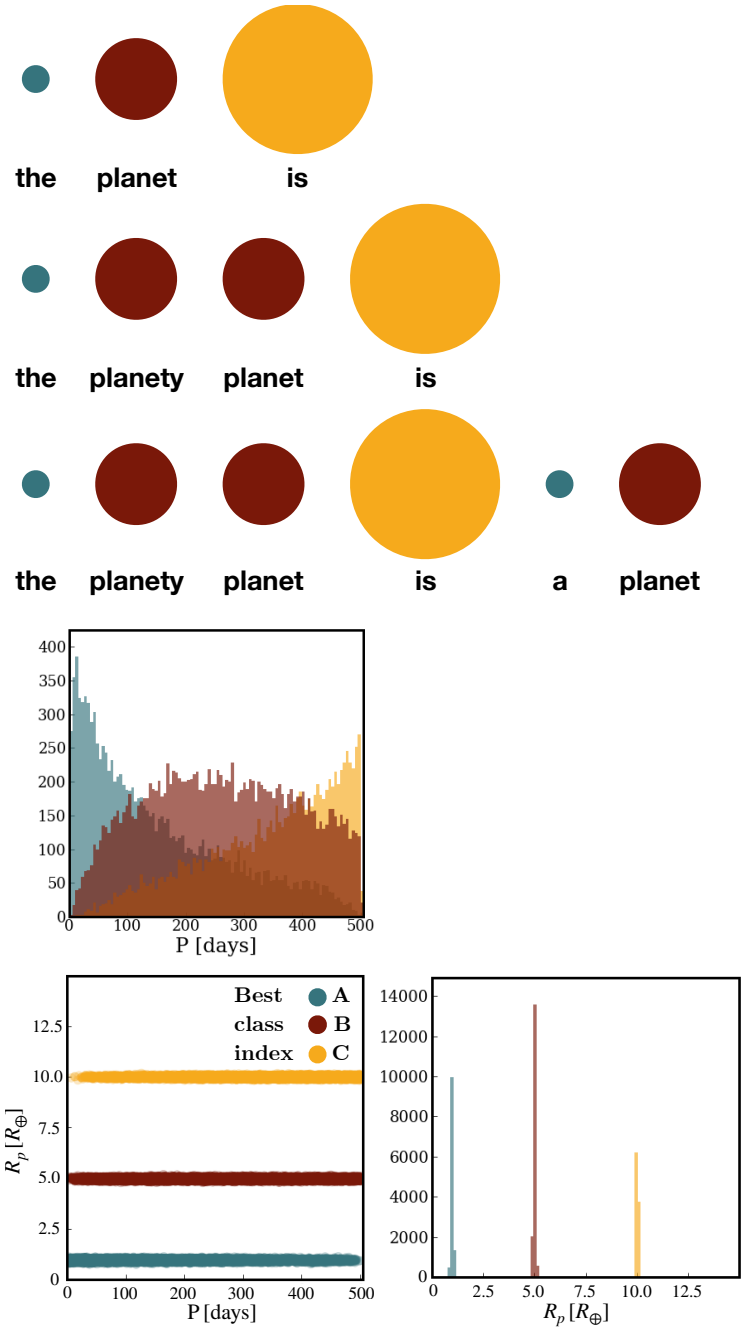

Figure 6.7: Top: A toy planetary "grammar." The three sizes of planet serve different grammatical functions in their systems, analogous to articles, adjectivey-nouns, and verbs. Bottom: The joint period-radius distribution of the planets in 10,000 test planetary systems generated according to these grammar rules. The planets are colormapped by the context network's decision about their class membership—for this very artificial population, both the context network and the target network recover the truth exactly.

per the equations of Fabrycky et al. (2014). If the system is stable, we keep it; otherwise, we discard it and draw new periods; if it is still unstable after 1000 period draws, we draw a new m and start over.

We generate a population of 10,000 training and 10,000 test systems according to this procedure, train the context and target networks on the training set, and then run the test set through the trained models. For both the training and test phases, we must also choose the number of categories $N_{classes}$ into which the networks attempt to sort the planets; when we choose the truth, $N_{classes} = 3$, both the target and context networks classify the simulated planets perfectly. When we choose $N_{classes} > 3$, the networks begin to subdivide the true classes arbitrarily.

The lower panel of Figure 6.7 shows the radius and period distributions of the planets in the test set, color-mapped according to the (perfectly accurate) class membership assigned by our *context* network. The target network's classifications are identical, so we do not plot them.

Having demonstrated that the network is capable of classifying planets belonging to systems of this arbitrary grammar, we now turn to the real KOI data set to see what we can learn.

6.4.2 Kepler systems, classified

We next apply the model shown in Figure 6.6, with the same architectures and hyper-parameters given above, to the set of real KOI systems described in Section 6.2. For these real systems, we do not know in advance whether or if there is a sensible choice of $N_{classes}$, so we run the network for $N_{classes} = 2$ to 10, inclusive. For each choice of $N_{classes}$, we again run 100 different training runs with different random initializations of the network weights, then take the one with the optimal cost function on the training set as our result. Figure 6.8 shows the cost function on the training set as a function of $N_{classes}$ for all these trials.

We next investigate what class boundaries the networks drew, to see if it picked up on any

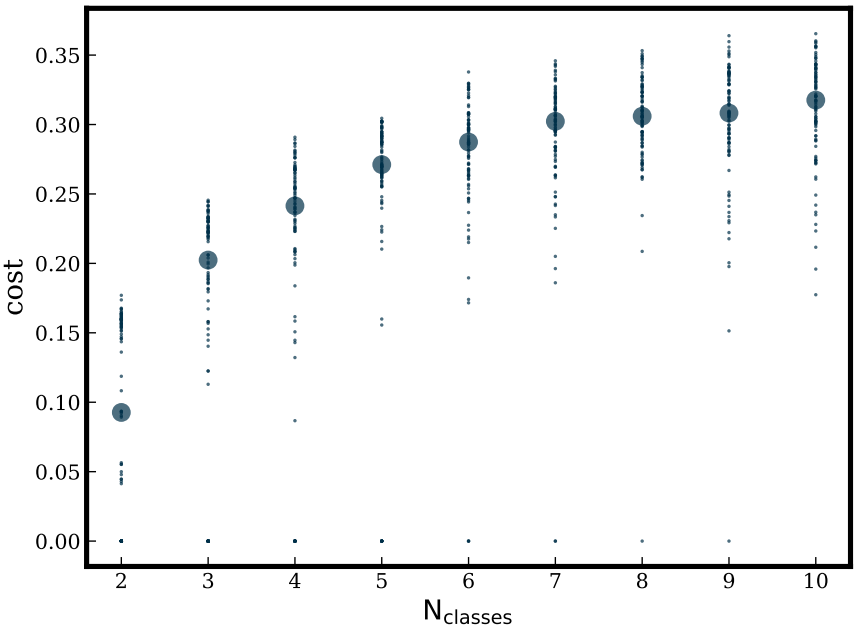

Figure 6.8: The best-achieved value of the cost function on the training set as a function of $N_{classes}$. For each choice of $N_{classes}$, we ran the network with 100 different random initializations of the network weights (small circles); the median value over the 100 trials is plotted as a large circle.

physically meaningful categories. The *target* network's classifications for the varying choices of $N_{classes}$ are shown in Figure 6.9. The *context* network's classifications are considerably noisier, with much more overlap between the classes; this makes sense because the context network is attempting to predict the target planet's class membership based on contextual, indirect information, rather than based on the period and radius of the target planet itself.

Interestingly, although the context network sees the stellar properties $T_{\rm eff}$ and $\log g$, both networks draw their boundaries exclusively in the period-radius plane; the stellar properties seem irrelevant, in the sense that there is no information about planet categories in those dimensions. Given that the *Kepler* planet hosts in our sample occupy a relatively narrow range of $T_{\rm eff}$ and $\log g$ (see Figure 6.1), we might expect a fairly homogeneous planet population around these stars, so this null result is sensible.

What can we make of the classification boundaries plotted in Figure 6.9? Some features persist despite the choice of $N_{classes}$; for $N_{classes} \geq 6$, for example, the short-period ($\sim 0.5 - 8$

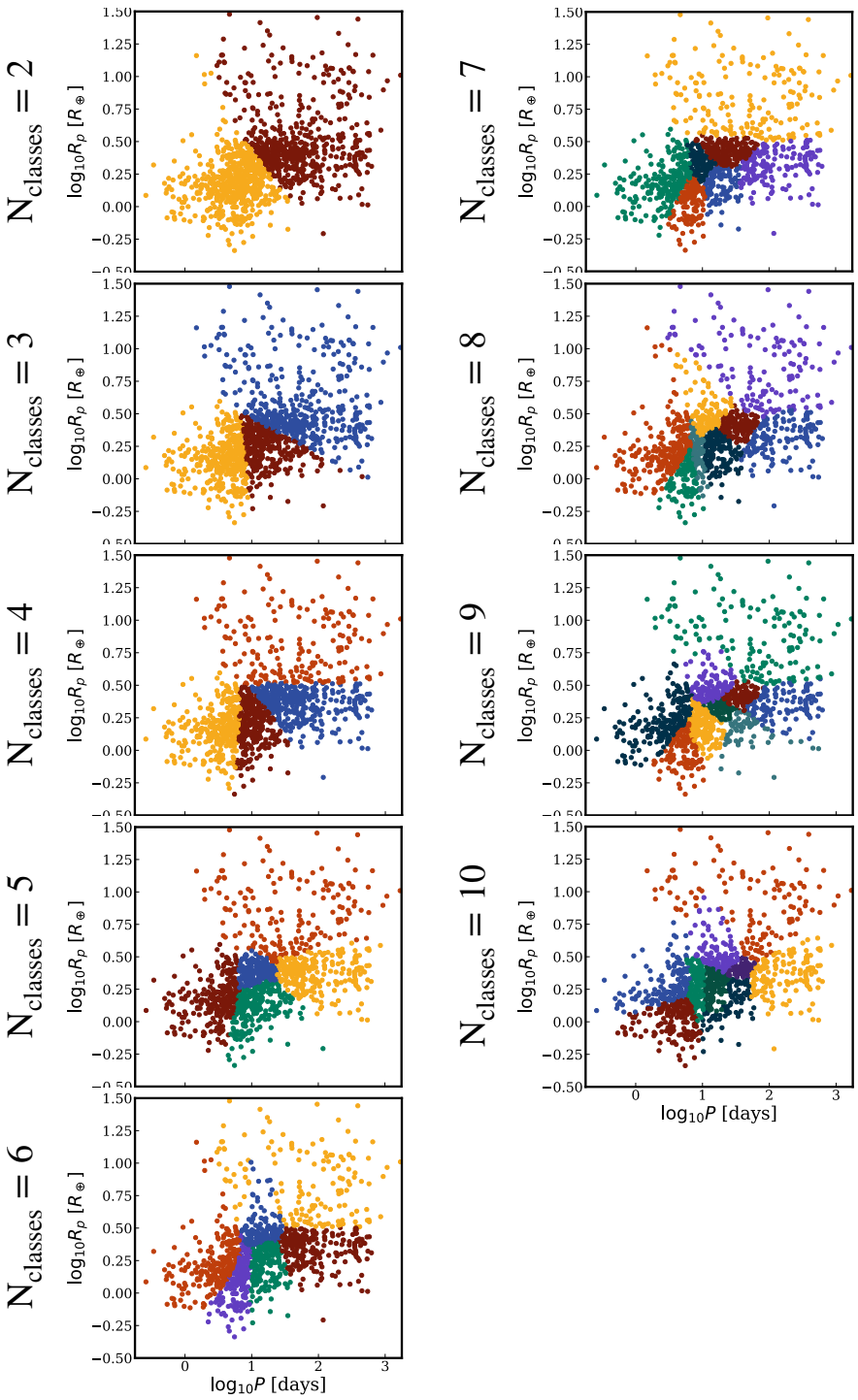

Figure 6.9: The class boundaries drawn by the *target* network in the period-radius plane, as a function of $N_{classes}$. The *context* network's class boundaries are considerably noisier, with much more overlap between classes; see the lower-right inset panels in Figure 6.10 for an example. We show only the period-radius plane because the networks do not draw boundaries in the stellar dimensions.

day) planets at the leftmost edge of the plot (plotted in orange for $N_{classes} = 6$, green for $N_{classes} = 7$) are assigned to a class of roughly the same shape and size, as are the slightly longer-period ($\sim 3 - 10$ day) small ($\sim 0.4 - 2R_\oplus$) planets (plotted in purple for $N_{classes} = 6$, orange for $N_{classes} = 7$).

Similarly, the intermediate-to-long-period ($\sim 30 - 2000$ day), large ($\sim 3 - 15R_\oplus$) planets (plotted in yellow for $N_{classes} = 6$) are assigned to an essentially stable class from $N_{classes} = 6$ upwards. So are the intermediate-to-long-period ($\sim 30 - 2000$ day), super-Earth-sized ($\sim 1 - 3R_\oplus$) planets just below them in the plane (plotted in red for $N_{classes} = 6$).

Class assignments for the middle section of the plane are less stable.

It is not obvious how to decide which choice of $N_{classes}$ is best: the achieved cost on the training set increases monotonically with $N_{classes}$, but we would expect this trend to continue indefinitely (albeit slowly), because a higher $N_{classes}$ means a model with more parameters, more easily moldable to our data set. Ideally, we would wish to define some analog of the Bayesian evidence, which could account for both the goodness-of-fit (i.e., cost) and the model complexity and allow us to compare the varying choices of $N_{classes}$ more directly, but it is not obvious how to do this either.

In Figure 6.8, we can see that the optimal network cost on the training set jumps noticeably from $N_{classes} = 5$ to 6 in Figure 6.8, then grows more slowly for higher $N_{classes}$. On this basis, we select $N_{classes} = 6$ to investigate further.

We next ask: what do the class boundaries the network draws when $N_{classes} = 6$ really mean? How are planets of these 6 types arranged in systems, and can we deduce any "grammatical rules" from these arrangements? In Figure 6.10, we plot the systems from our *test* set, with each planet size-scaled according to its radius and color-coded by the *target* network's class assignment. For reference, we show the class boundaries drawn by the target and context networks as insets in the lower-right corner.

We can see immediately that certain trivial patterns arise from the period ranking of systems in the panels: "orange" planets, which come from the shortest-period group, occur

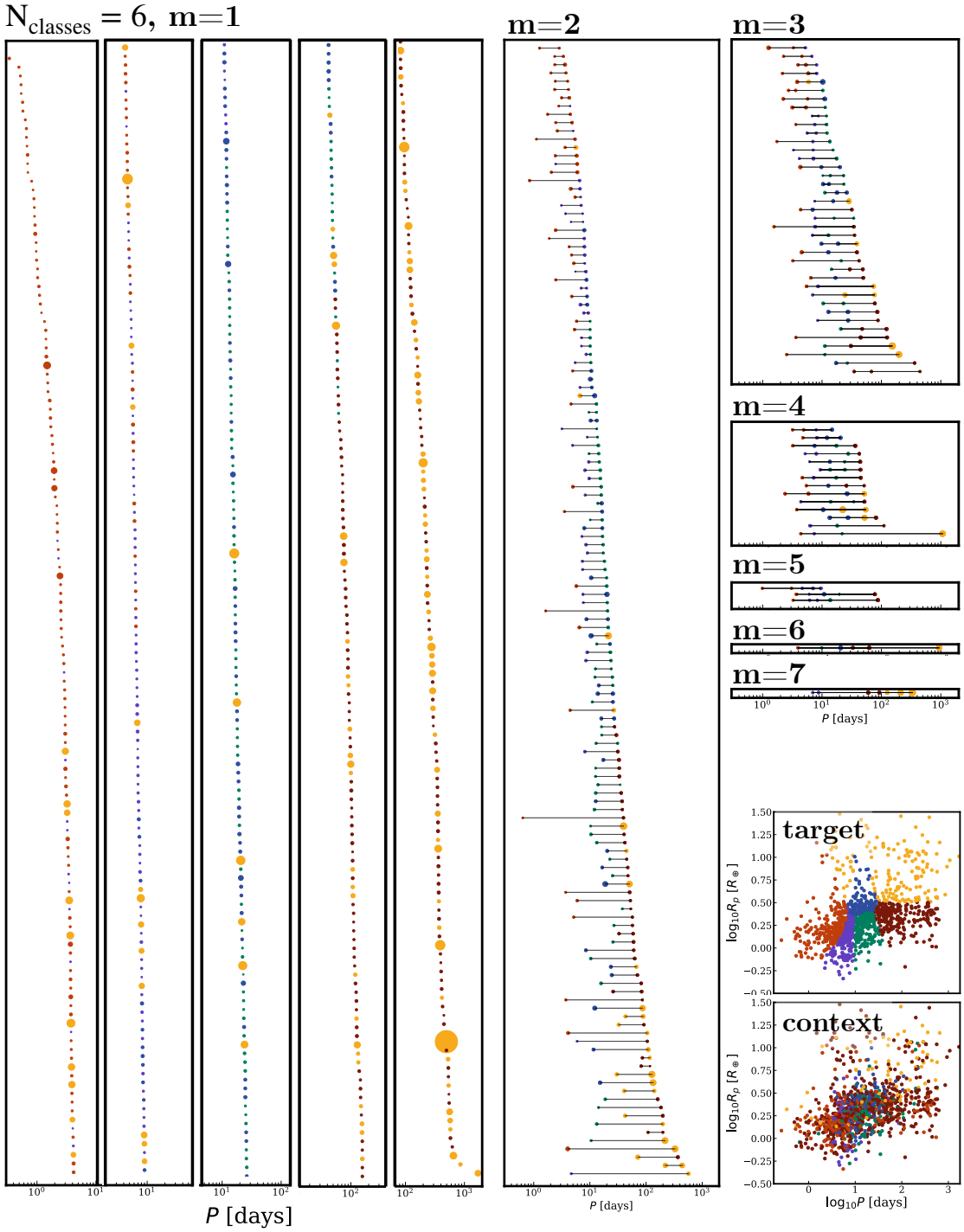

Figure 6.10: *Target* network classification of planets in test-set systems. Each row represents one system; each circle represents one planet, with area proportional to R_p. Systems are sorted by outermost planet period (note the change in x-axis limits across the 1-planet systems). Planets are color-coded by their target network class assignment. Bottom right: Maps of the class boundaries by the target and context networks, respectively.

more frequently toward the top of each multiplicity panel, and red and yellow planets, from the long-period groups, occur more frequently toward the bottom. Within systems, which are individually sorted by period, certain sequences are impossible given the period distributions of the groups: a green planet cannot precede an orange, for example.

Also evident are the well-known patterns which result from the observational biases of the transit method. In general, planet size and period are correlated, i.e. planets at longer periods are larger; this pattern is sculpted by *Kepler*'s increasing incompleteness toward the lower-right quadrant of the period-radius plane (see e.g. Petigura et al. 2013, figure 1).

There are several interesting patterns beyond these, however. First, the largest planets, which generally belong to the yellow group, are overrepresented among the single-planet ($m = 1$) systems. In contrast, planets from the small-radius, intermediate-period green group are underrepresented among the single-planet systems.

Within the multi-planet systems ($m \geq 2$), an interesting pattern arises overall, which perhaps sheds some light on why the network decided on these particular class boundaries: it is rare for neighbouring planets in a system to come from non-contiguous classes in the map (i.e., classes that do not share a border). Only 16% of multi-planet systems (31/194) have neighbouring pairs from non-contiguous classes (e.g., have a green planet immediately following an orange, or a yellow immediately following a green). Vastly more common are either (1) successive planets from the same class, e.g. "orange orange" or (2) neighbouring planets from neighbouring classes—the patterns "orange purple", "purple green", "green red", and "blue yellow" occur frequently, which can be seen clearly in the 2-planet systems. Planet pairs from non-contiguous classes which are farther apart in the map are even rarer: "orange red" happens only about half as often as "orange green".

The commonness of same-group pairs and contiguous pairs is overall superficially consistent with a "peas in a pod" picture of planetary systems, where planets are correlated in size with their neighbours and periods are regularly spaced (Weiss et al., 2018).

6.5 Discussion & open questions

Here, we have explored two avenues to understanding planetary systems as ordered sequences, in which the arrangement of individual planets contains information beyond that of the planets themselves. In other words, we have explored ways to understand planets in the context of their systems (i.e., their host star and their sibling planets). We specifically investigate a data set consisting of 3690 *Kepler* objects of interest, grouped into 2804 planetary systems of multiplicity ranging from 1 to 7.

We first explore a regression problem: is it possible to accurately predict the radius and period of a planet, if the radii and periods of its surrounding planets are known? We find that our trained model indeed can. We compare our network's predictions to those of a naive model, which accounts only for basic orbital stability, and find that the network can predict both planetary radius to a factor of 2.4 better, and period to a factor of 1.8 better, than the naive model.

We find furthermore that the network's predictions improve with increasing system multiplicity, and that neither the network nor the naive model are able to make meaningful predictions about most single-planet systems, indicating that planet context information, not stellar context information, is driving the accuracy of the network's predictions. The only exception to this is a small subset of large (i.e., high-radius) single-planet systems, for which the network seems to be exploiting contextual stellar information to accurately predict high radii. This work is still in progress, but we will identify which planets these are.

We next explore a planetary classification model based on recent advances in linguistic part-of-speech tagging. Our model consists of two networks; one which sees the "target" planet whose properties we are trying to predict, and the other which sees the surrounding stellar and planetary context. The network trains by maximizing the mutual information between the class assignments of the target network and those of the context network, a principle which has proved very successful in linguistic applications but has not, to our

knowledge, previously been used in the noisier realm of exoplanet science.

We find, most importantly, that the context and target networks draw their classification boundaries in the period-radius plane only, despite the context network also having access to the stellar properties $T_{\rm eff}$ and $\log g$. This indicates that the network has found no useful correlations to exploit in the stellar dimensions, which is not surprising given the relative homogeneity of the stars in our sample. In future, it would be very interesting to expand our sample to include e.g. TESS systems, where cool stars are much better represented than among the *Kepler* target stars. (It would be extremely interesting also to apply these methods to a data set of transiting planets containing more hot Jupiters, which are overall larger, closer-in, and lonelier than the planets investigated here, and form a cluster in the period-radius plane that is quite distinct from the planets in our sample.)

For $\mathrm{N}_{\rm classes} \geq 6$, we find that the target network consistently identifies four stable classes. The first consists of short-period ($\sim 0.5 - 8$ day) planets of all sizes. The second consists of slightly longer-period ($\sim 3 - 10$ day) small ($\sim 0.4 - 2R_\oplus$) planets. The third consists of intermediate-to-long-period ($\sim 30 - 2000$ day), super-Earth-sized ($\sim 1 - 3R_\oplus$) planets, and the fourth of intermediate-to-long-period ($\sim 30 - 2000$ day), large ($\sim 3 - 15R_\oplus$) planets. The classification boundaries drawn by the context network are unsurprisingly much noisier.

When we look at the "grammatical patterns" in the classifications assigned by the network, we find that all neighbour-planet pairs in 84% of multi-planet systems come from either the same class or classes which are contiguous in the period-radius plane. In other words, pairs where the two members come from non-contiguous classes are rare. This pattern is superficially consistent with a peas-in-a-pod picture, where sequences of similar planets are common.

Overall, these methods show promise for picking up on subtle patterns in the arrangement of planetary systems. We caution that our data set is small and quite noisy, so the conclusions offered here are tentative, and certainly fall short of the ultimate goal of inducting planetary grammar—but prospects can only improve as we gather more data.

Acknowledgements

ES, DK, & MC acknowledge support from the Columbia University Data Science Institute "Seed Funds Program". Thanks to members of the Cool Worlds Lab for useful discussions in preparing this manuscript.

7. Conclusion

There is much to be learned from the shadows of exoplanets. By carefully modeling the shapes of transit light curves, we can deduce the properties of individual exoplanets, their orbits, and their host stars, as well as begin to understand the population of exoplanets and the arrangement of planets in systems. Here, we summarize the results presented in the preceding chapters and discuss avenues for future work.

7.1 Summary of results

In Chapter 2, we measured the densities of 66 *Kepler* planet host stars from the light curves of their transiting planets, following the method originally suggested by Seager & Mallén-Ornelas (2003). We selected these planets because they have strong prior eccentricity constraints, either because their secondary eclipses have been measured, because their theoretical tidal circularization timescales are short, or because they belong to compact multi-planet systems; because both stellar density and eccentricity influence the transit duration, these strong eccentricity constraints are necessary to derive a correspondingly tight constraint on stellar density. We demonstrated that density constraints derived from the light curve can be as precise as density constraints derived from asteroseismology, for stars up to three magnitudes fainter than the *Kepler* asteroseismic limit, for mid-F to mid-K dwarfs. Finally, we show that the major limiting factor in the precision of stellar density measured from a transit light curve is the signal-to-noise ratio of the transits. In other words, our ability to precisely measure the density depends mainly on our ability to precisely measure the transit shape.

We concluded Chapter 2 with a brief discussion of the potential of this method to help constrain the periods of singly-transiting planets: if a star's density can be "anchored" by precisely modeling the light curve of another planet in the same system, the stellar density can then be used, in conjunction with the shape of the single transit, to constrain the single transiter's period (an application originally suggested by Seager & Mallén-Ornelas (2003) and Yee & Gaudi 2008).

Of course, as we explored in Chapter 3, a stellar density constraint need not come from transits to be useful in constraining the period of a single transiter. In this chapter, we derived stellar density constraints for the host stars of 12 single transiters observed by K2. We combined their Gaia-measured parallaxes with published broad-band photometry, then ran these data through Yonsai-Yale isochrone modeling to obtain bulk stellar densities; we demonstrated that these densities are consistent with densities derived by asteroseismology. We then modeled the transits of these 12 planets assuming these stellar density values *a priori* and achieved good precision in our period posteriors: when we treated eccentricity as a free parameter, the fractional period uncertainty over the 12 single transiters was $94^{+87}_{-58}\%$, and when we fixed $e = 0$, it was $15^{+30}_{-6}\%$, which represents a roughly threefold improvement over typical period uncertainties of previous studies. We also explored the accuracy of period posteriors derived this way by applying our method to a sample of 27 validation planets of known period. We demonstrated that the method's accuracy depends strongly on our ability to measure a/R_* from the transit shape, which means that the transit must be well-sampled during ingress and egress for the period estimation to be reliable.

In Chapter 4, we considered the general problem of "shadow imaging," or inferring the shape of a transiting object from its light curve without assuming a physical model for the object itself. We identified several interesting degeneracies inherent to this problem, including the "flip" degeneracy, by which shadow images are horizontally symmetrical; the "arc" degeneracy, by which an infinitesimal opaque point produces the same transit light curve as a pair of intersecting semicircular arcs; and the "stretch" degeneracy, by which

a wide image transiting at high velocity produces the same light curve as a narrow image transiting slowly. Despite these degeneracies, if we adopted additional assumptions in our algorithmic "inversion" of the light curves, we could recover informative shadow images of arbitrary transit light curves, including the TRAPPIST-1 c, e, and f triple transit and Dips 5 and 8 of Boyajian's Star. The conclusions we could draw about the Boyajian's Star occulter were limited by the time-sampling of their *Kepler* light curves, but our shadow images suggested structure, with alternating bands of transparency and opacity.

In Chapter 5, we shifted from considering the shape of individual transits to the "shape" or arrangement of planets in systems. First we considered the underlying multiplicity distribution of *Kepler* systems orbiting FGK dwarfs. We generated ten catalogs of "observed" simulated planetary systems, each from a different intrinsic multiplicity distribution, accounting for both geometric and detection biases. We then used approximate Bayesian computation to compare the "observed" simulated catalogs to the real *Kepler* multis, and concluded that, in contrast to previous work which invokes a dichotomous planet population to explain the high number of observed single-planet systems, a single-population Zipfian model matches the real data as well as any of the dichotomous models. According to the best-fit Zipfian model, we predicted that ~ 750 additional planets should exist in ostensibly single-planet *Kepler* systems, and that RV follow-up of these single-planet systems could test this prediction.

Finally, in Chapter 6, we investigated the relationship between planets and their surrounding "context," i.e. their host star and sibling planets. We built a neural network model to predict the radius and period of a planet based on its context, and found that this model predicts both of these properties to a factor of 2 better than a "naive" model which considers only basic dynamical stability constraints. We also adapted a part-of-speech tagging model from computational linguistics to explore the "grammar" of planetary systems and concluded that (1) within our relatively homogeneous sample of FGK host stars, the model does not seem to learn anything from the stellar features; (2) the network consis-

tently identifies certain clusters in the period-radius plane, including short-period ($\sim 0.5-8$ day) planets of all sizes; slightly longer-period ($\sim 3-10$ day) small ($\sim 0.4-2R_\oplus$) planets; intermediate-to-long-period ($\sim 30-2000$ day), super-Earth-sized ($\sim 1-3R_\oplus$) planets; and intermediate-to-long-period ($\sim 30-2000$ day) large ($\sim 3-15R_\oplus$) planets; and (3) that it is rare for a multi-planet system to include a pair of neighboring planets from non-contiguous classes in the period-radius plane (only 16% of multis contain such a pair). This last pattern is superficially consistent with a peas-in-a-pod picture, where we expect sequences of planets belonging to the same class to be common.

7.2 Future work

A natural extension of Chapters 2, 3, 5, and 6 would be to apply these methods to systems discovered around TESS stars, which are overall closer, brighter, and cooler than *Kepler* host stars. As we note in the conclusion to Chapter 2, the TESS asteroseismic limit is brighter than that of *Kepler* because of its smaller aperture (Campante et al., 2016; Ricker et al., 2014), so transit light curves could be used to measure the densities of TESS stars too faint for asteroseismology. Since the publication of Chapter 2, the asteroseismic capabilities of TESS have been more carefully modeled: Schofield et al. (2019) constructed a list of TESS stars to target for asteroseismic measurements and predicted the yield of asteroseismic detections from this list, finding that the distribution of V-band magnitudes for the stars with detections peaks at $V \simeq 8.5$, on average 4 to 5 magnitudes brighter than *Kepler* (see their Figure 11).

Similarly, since the publication of Chapter 3, the first of the hundreds of TESS single-transit detections predicted by Huang et al. (2018) and Villanueva et al. (2019) have begun rolling in. Several TESS single transits have been followed up with RV observations or ground-based photometry and confirmed as planets, including HD 21749 b (also known as TOI-186.01; Dragomir et al. 2019), GJ 143 b (Trifonov et al., 2019), HD 332231 b (Dalba

et al., 2020), HD 332231 b (Dalba et al., 2020), NGTS-11 b (also known as TOI-1847.01; Gill et al. 2020a; TIC 238855958 b (Gill et al., 2020b), and TIC 231005575 b (Gill et al., 2020c), but so far there exists no catalog of TESS single transits overall.

Meanwhile, there have been further injection-recovery simulation studies done about the TESS single transiters as a population. Cooke et al. (2019) investigate how the planned re-observation of the southern ecliptic hemisphere during the TESS extended mission will further illuminate the population of single transiters detected there during the primary mission: they calculate that approximately 80% of their predicted single transiters will transit again during the second year of observations, but that 75% these will only transit again once (i.e., TESS will have seen these planets transit exactly twice: once during the primary mission, and once during the extended mission). Because ~ 3 years will have elapsed between the two observations, these second transits will not suffice to identify unique periods for this population, so period constraints from transit modeling based on prior stellar density information could be very useful.

Yao et al. (2019), meanwhile, consider how TESS single transiters could be identified in ground-based photometric surveys, namely the Kilodegree Extremely Little Telescope (KELT) survey: they inject TESS single transits as periodic signals into KELT light curves, then attempt to detect these signals using a box-least squares search tuned to the TESS-observed transit midpoint time, depth, and duration over a range of possible periods. The methods of Chapter 3 could constrain this range, making such searches more efficient.

An extension to TESS systems is even more important for the work described in Chapters 5 and 6, because these chapters attempt to infer certain aspects of the underlying structure of planetary systems from the *Kepler* sample, which is limited to a relatively narrow range of stellar types similar to the Sun (FGK dwarfs). Ballard (2019) predicts that TESS, in contrast, will detect multi-planet systems around ~ 200 M-dwarfs (a factor of ~ 4 higher than *Kepler*'s yield of M-dwarf multis). Because Ballard & Johnson (2016) initially suggested a dichotomous planet population to explained the observed *Kepler* M-dwarf sys-

tem multiplicities, this new TESS sample will be a critical test of the underlying multiplicity distribution.

It will also be extremely interesting to apply the part-of-speech tagging-inspired model from Chapter 6 to M-dwarf systems. We found that, within our relatively homogeneous sample of FGK host stars, our neural network did not seem to learn anything from the stellar features we included as part of the "context;" this may or may not hold true for a sample expanded to include M-dwarf systems. Furthermore, this expanded sample would allow us to see if there are meaningful differences between the "grammar" of FGK and M systems, offering another window onto the *Kepler* dichotomy.

The next steps for shadow imaging (Chapter 4) are much more theoretical than applied: as it stands now, despite our good understanding of the degeneracies of the problem, shadow imaging is a deterministic algorithmic procedure to recover exactly one image from exactly one light curve, with no attempt made to estimate the uncertainty of this image, i.e. the uncertainties of the opacities of the constituent pixels. Because of the degeneracies (particularly the arc degeneracy), these opacities are strongly and non-trivially correlated, so even small perturbations to the best-fit shadow image generally substantially worsen the fit to the light curve. I believe there is analytic progress to be made on this front, beginning by further exploring the mapping between a maximally compact image and its maximally arc-dispersed counterpart.

Bibliography

Aigrain S., Parviainen H., Pope B. J. S., 2016, MNRAS, 459, 2408

Aitchison L., Corradi N., Latham P. E., 2016, PLoS Computational Biology, 12, e1005110

Akaike H., 1974, IEEE Transactions on Automatic Control, 19, 716

Akeson R. L., 2015, in AAS/Division for Planetary Sciences Meeting Abstracts. p. 417.01

Akeson R. L., et al., 2013, Publications of the Astronomical Society of the Pacific, 125, 989

Andersen A., Kak A., 1984, Ultrasonic Imaging, 6, 81

Anderson T. W., Darling D. A., 1952, Ann. Math. Statist., 23, 193

Auerbach F., 1913, Petermanns Geographische Mitteilungen, 59, 74

Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, AJ, 156, 58

Ballard S., 2019, AJ, 157, 113

Ballard S., Johnson J. A., 2016, ApJ, 816, 66

Baraffe I., Homeier D., Allard F., Chabrier G., 2015, A&A, 577, A42

Barnes J. W., Fortney J. J., 2003, ApJ, 588, 545

Barnes J. W., Fortney J. J., 2004, ApJ, 616, 1193

Batalha N. M., et al., 2013, ApJS, 204, 24

Batygin K., Laughlin G., 2015, Proceedings of the National Academy of Science, 112, 4214

Beaugé C., Nesvorný D., 2012, ApJ, 751, 119

Beaumont M. A., 2019, Annual Review of Statistics and Its Application, 6, 379

Becker J. C., Vanderburg A., Adams F. C., Rappaport S. A., Schwengeler H. M., 2015, ApJ, 812, L18

Becker J. C., et al., 2019, AJ, 157, 19

Belevitch V., 1959, Annales de la Société Scientifique de Bruxelles, 73, 310

Berdyugina S. V., Kuhn J. R., 2019, AJ, 158, 246

Bodman E. H. L., Quillen A., 2016, ApJ, 819, L34

Borucki W. J., Summers A. L., 1984, Icarus, 58, 121

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

Borucki W. J., et al., 2011, ApJ, 736, 19

Bovaird T., Lineweaver C. H., 2017, MNRAS, 468, 1493

Boyajian T. S., et al., 2016, MNRAS, 457, 3988

Brahm R., et al., 2018, MNRAS, 477, 2572

Brahm R., et al., 2019, MNRAS, 483, 1970

Brakensiek J., Ragozzine D., 2016, ApJ, 821, 47

Brown P. F., Della Pietra V. J., deSouza P. V., Lai J. C., Mercer R. L., 1992, Computational Linguistics, 18, 467

Buchner J., et al., 2014, A&A, 564, A125

Budding E., 1977, Ap&SS, 48, 207

Burke C. J., Catanzarite J., 2017, Technical report, Planet Detection Metrics: Per-Target Detection Contours for Data Release 25

Burke C. J., et al., 2015, ApJ, 809, 8

Campante T. L., et al., 2016, ApJ, 830, 138

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

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

Chapman A., 1990, QJRAS, 31, 333

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

Charbonneau D., et al., 2009, Nature, 462, 891

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

Chen J., Kipping D. M., 2018, MNRAS, 473, 2753

Christiansen J. L., et al., 2012, Publications of the Astronomical Society of the Pacific, 124, 1279

Christiansen J. L., et al., 2016, ApJ, 828, 99

Christodoulopoulos C., Goldwater S., Steedman M., 2010, in Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Cambridge, MA, pp 575–584, https://www.aclweb.org/anthology/D10-1056

Ciardi D. R., Fabrycky D. C., Ford E. B., Gautier T. N. I., Howell S. B., Lissauer J. J., Ragozzine D., Rowe J. F., 2013, ApJ, 763, 41

Claret A., 2000, A&A, 363, 1081

Claret A., Hauschildt P. H., Witte S., 2012, A&A, 546, A14

Cooke B. F., Pollacco D., Bayliss D., 2019, A&A, 631, A83

Coughlin J. L., López-Morales M., 2012, AJ, 143, 39

Cowan N. B., Agol E., 2011, ApJ, 729, 54

Crossfield I. J. M., et al., 2018, The Astrophysical Journal Supplement Series, 239, 5

Dalba P. A., et al., 2020, AJ, 159, 241

Dawson R. I., Johnson J. A., 2018, Annual Review of Astronomy and Astrophysics, 56, 175

Deeg H., 2009, in Transiting Planets. pp 388–391 (arXiv:0807.3915), doi:10.1017/S1743921308026720

Désert J.-M., et al., 2011, ApJS, 197, 14

Doyle L. R., McCowan B., Johnston S., Hanser S. F., 2011a, Acta Astronautica, 68, 406

Doyle L. R., et al., 2011b, Science, 333, 1602

Dragomir D., et al., 2019, ApJ, 875, L7

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

Dressing C. D., Charbonneau D., 2015, ApJ, 807, 45

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

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

Espinoza N., Jordán A., 2016, MNRAS, 457, 3573

Fabrycky D. C., et al., 2012, ApJ, 750, 114

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

Faigler S., Mazeh T., 2011, MNRAS, 415, 3921

Fang J., Margot J.-L., 2012, ApJ, 761, 92

Farr B., Farr W. M., Cowan N. B., Haggard H. M., Robinson T., 2018, AJ, 156, 146

Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601

Figueira P., et al., 2012, A&A, 541, A139

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

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

Foreman-Mackey D., Hoyer S., Bernhard J., Angus R., 2014a, George: George (V0.2.0), doi:10.5281/zenodo.11989

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

Foreman-Mackey D., Morton T. D., Hogg D. W., Agol E., Schölkopf B., 2016, AJ, 152, 206

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

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

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

Gaidos E., Mann A. W., Kraus A. L., Ireland M., 2016, MNRAS, 457, 2877

Galindo Trejo J., Allen C., 2005, in Kurtz D. W., ed., IAU Colloq. 196: Transits of Venus: New Views of the Solar System and Galaxy. pp 124–137, doi:10.1017/S1743921305001328

Gassendi P., 1632, Mercurius in sole visus et Venus invisa Parisiis a 1631. Pro voto & admonitione Kepleri. Epistolae duae cum observatis quibusdam aliis. https://books.google.co.uk/books?id=6_v0PgAACAAJ

Gibson N. P., et al., 2009, ApJ, 700, 1078

Gilbert G. J., Fabrycky D. C., 2020, AJ, 159, 281

Giles H. A. C., et al., 2018, A&A, 615, L13

Gill S., et al., 2020a, arXiv e-prints, p. arXiv:2005.00006

Gill S., et al., 2020b, MNRAS, 491, 1548

Gill S., et al., 2020c, MNRAS, 495, 2713

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

Giménez A., 2006, A&A, 450, 1231

Goncharskii A. V., Stepanov V. V., Khokhlova V. L., Yagola A. G., 1982, Soviet Ast., 26, 690

Goodricke J., 1783, Philosophical Transactions of the Royal Society of London Series I, 73, 474

Greene B., Rubin G., 1971, Automatic Grammatical Tagging of English. Department of Linguistics, Brown University, https://books.google.co.uk/books?id=VznTygAACAAJ

Grunblatt S. K., et al., 2016, AJ, 152, 185

Grunblatt S. K., et al., 2017, AJ, 154, 254

Hahn C., Vakili M., Walsh K., Hearin A. P., Hogg D. W., Campbell D., 2017, MNRAS, 469, 2791

Haswell C. A., 2010, Transiting Exoplanets

Heller R., 2019, A&A, 628, A42

Henden A. A., Welch D. L., Terrell D., Levine S. E., 2009, in American Astronomical Society Meeting Abstracts #214. p. 669

Henry G. W., Marcy G. W., Butler R. P., Vogt S. S., 2000, ApJ, 529, L41

Hill G., Hutchings J. B., 1970, ApJ, 162, 265

Hogg D. W., Myers A. D., Bovy J., 2010, ApJ, 725, 2166

Holczer T., et al., 2016, ApJS, 225, 9

Howard A. W., et al., 2012, The Astrophysical Journal Supplement Series, 201, 15

Hsu D. C., Ford E. B., Ragozzine D., Morehead R. C., 2018, AJ, 155, 205

Huang C. X., et al., 2018, preprint, (arXiv:1807.11129)

Hubbard W. B., Fortney J. J., Lunine J. I., Burrows A., Sudarsky D., Pinto P., 2001, ApJ, 560, 413

Huber D., et al., 2013, ApJ, 767, 127

Hutchings J. B., 1968, MNRAS, 141, 329

Ishida E. E. O., et al., 2015, Astronomy and Computing, 13, 1

Ivezić Ž., et al., 2019, ApJ, 873, 111

Jenkins J. M., 2002, ApJ, 575, 493

Jones M. I., et al., 2018, A&A, 613, A76

Juvan I. G., Lendl M., Cubillos P. E., Fossati L., Tregloan-Reed J., Lammer H., Guenther E. W., Hanslmeier A., 2018, A&A, 610, A15

Katz J. I., 2017, MNRAS, 471, 3680

Kawahara H., Fujii Y., 2011, ApJ, 739, L62

Kepler J., 1630, Joannis Keppleri...admonitio ad astronomos, rerumque coelestium studiosos, de raris mirisque anni 1631 phænomenis, Veneris puta et Mercurii in Solem incursa; excerpta ex ephemeride anni 1631, & certo authoris consilio huic præmissa, iterumque edita a Jacobo Bartschio. Tambach, Gottfried, fl. 1607-1632, https://books.google.co.uk/books?id=xTDcwQEACAAJ

Kingma D., Ba J., 2014, International Conference on Learning Representations

Kipping D. M., 2008, MNRAS, 389, 1383

Kipping D. M., 2010a, MNRAS, 407, 301

Kipping D. M., 2010b, MNRAS, 408, 1758

Kipping D. M., 2011, MNRAS, 416, 689

Kipping D. M., 2013, MNRAS, 434, L51

Kipping D. M., 2016, MNRAS, 455, 1680

Kipping D., 2018a, Research Notes of the American Astronomical Society, 2, 223

Kipping D., 2018b, MNRAS, 473, 784

Kipping D., Bakos G., 2011, ApJ, 733, 36

Kipping D. M., Sandford E., 2016, MNRAS, 463, 1323

Kipping D. M., Tinetti G., 2010, MNRAS, 407, 2589

Kipping D. M., et al., 2016, ApJ, 820, 112

Knutson H. A., et al., 2007a, Nature, 447, 183

Knutson H. A., Charbonneau D., Noyes R. W., Brown T. M., Gilliland R. L., 2007b, ApJ, 655, 564

Kollerstrom N., 2005, in Kurtz D. W., ed., IAU Colloq. 196: Transits of Venus: New Views of the Solar System and Galaxy. pp 34–40, doi:10.1017/S1743921305001249

Kolmogorov A. N., 1933, Giornale dell'Istituto Italiano degli Attuari, 4, 83

Kopal Z., 1959, Close binary systems

Kopal Z., 1979, Language of the stars. A discourse on the theory of the light changes of eclipsing variables, doi:10.1007/978-94-009-9466-9.

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

Kreidberg L., 2015, PASP, 127, 1161

Kundurthy P., Agol E., Becker A. C., Barnes R., Williams B., Mukadam A., 2011, ApJ, 731, 123

LaCourse D. M., Jacobs T. L., 2018, Research Notes of the American Astronomical Society, 2, 28

Lanza A. F., Catalano S., Cutispoto G., Pagano I., Rodono M., 1998, A&A, 332, 541

Latham D. W., et al., 2011, ApJ, 732, L24

Lissauer J. J., et al., 2011a, ApJS, 197, 8

Lissauer J. J., et al., 2011b, Nature, 470, 53

Love A., 1934, A Treatise on the Mathematical Theory of Elasticity. University Press, https://books.google.com/books?id=AD79MAAACAAJ

Lucy L. B., 1968, ApJ, 153, 877

Luger R., Agol E., Kruse E., Barnes R., Becker A., Foreman-Mackey D., Deming D., 2016, AJ, 152, 100

Majeau C., Agol E., Cowan N. B., 2012, ApJ, 747, L20

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

Marin J.-M., Pudlo P., Robert C. P., Ryder R. J., 2012, Statistics and Computing, 22, 1167

Marjoram P., Molitor J., Plagnol V., Tavaré S., 2003, Proceedings of the National Academy of Sciences, 100, 15324

Mathur S., et al., 2017, The Astrophysical Journal Supplement Series, 229, 30

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

McAllester D., 2018, CoRR, abs/1802.07572

McTier M. A. S., Kipping D. M., 2018, MNRAS, 475, 4978

Metzger B. D., Shen K. J., Stone N., 2017, MNRAS, 468, 4399

Millholland S., Wang S., Laughlin G., 2017, ApJ, 849, L33

Mochnacki S. W., Doughty N. A., 1972, MNRAS, 156, 51

Morton T. D., Bryson S. T., Coughlin J. L., Rowe J. F., Ravichandran G., Petigura E. A., Haas M. R., Batalha N. M., 2016, ApJ, 822, 86

Mulders G. D., Pascucci I., Apai D., Ciesla F. J., 2018, AJ, 156, 24

Munoz Romero C. E., Kempton E. M. R., 2018, AJ, 155, 134

Mustill A. J., Wyatt M. C., 2011, MNRAS, 413, 554

Nelder J. A., Mead R., 1965, The Computer Journal, 7, 308

Nelson B., Davis W. D., 1972, ApJ, 174, 617

Nesvorný D., 2011, ApJ, 742, L22

O'Donovan F. T., et al., 2006, ApJ, 651, L61

O'Donovan F. T., Charbonneau D., Harrington J., Madhusudhan N., Seager S., Deming D., Knutson H. A., 2010, ApJ, 710, 1551

O'Leary R. M., Burkart J., 2014, MNRAS, 440, 3036

Ogilvie G. I., Lin D. N. C., 2004, ApJ, 610, 477

Ormel C. W., Liu B., Schoonenberg D., 2017, A&A, 604, A1

Orosz J. A., et al., 2012, Science, 337, 1511

Osborn H. P., et al., 2016, MNRAS, 457, 2273

Pál A., 2012, MNRAS, 420, 1630

Papaloizou J. C. B., Szuszkiewicz E., Terquem C., 2018, MNRAS, 476, 5032

Penoyre Z., in prep., Research Notes of the American Astronomical Society

Penoyre Z., Sandford E., 2019, MNRAS, 488, 4181

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

Piskunov N. E., Tuominen I., Vilhu O., 1990, A&A, 230, 363

Poulsen S. K., 2009, Master's thesis, Niels Bohr Institute, University of Copenhagen

Price E. M., Rogers L. A., 2014, ApJ, 794, 92

Quillen A. C., 2006, MNRAS, 365, 1367

Rappaport S., et al., 2016, MNRAS, 462, 1812

Ricker G. R., et al., 2014, in Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave. p. 914320 (arXiv:1406.0151), doi:10.1117/12.2063489

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

Rosenblatt F., 1971, Icarus, 14, 71

Rousseeuw P. J., Croux C., 1993, Journal of the American Statistical Association, 88, 1273

Russell H. N., 1912a, ApJ, 35, 315

Russell H. N., 1912b, ApJ, 36, 54

Russell H. N., 1912c, ApJ, 36, 133

Russell H. N., 1948, ApJ, 108, 388

Russell H. N., Merrill J. E., 1952, The determination of the elements of eclipsing binaries

Sackett P. D., 1999, in Mariotti J. M., Alloin D., eds, NATO Advanced Science Institutes (ASI) Series C Vol. 532, NATO Advanced Science Institutes (ASI) Series C. p. 189 (arXiv:astro-ph/9811269)

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

Sandford E., Kipping D., 2019, AJ, 157, 42

Sandford E., Espinoza N., Brahm R., Jordán A., 2019a, MNRAS, 489, 3149

Sandford E., Kipping D., Collins M., 2019b, MNRAS, 489, 3162

Santerne A., et al., 2018, Nature Astronomy, 2, 393

Schofield M., et al., 2019, ApJS, 241, 12

Schwarz G., 1978, Ann. Statist., 6, 461

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

Shabram M., Demory B.-O., Cisewski J., Ford E. B., Rogers L., 2016, ApJ, 820, 93

Shibahashi H., Kurtz D. W., 2013, in Shibahashi H., Lynas-Gray A. E., eds, Astronomical Society of the Pacific Conference Series Vol. 479, Progress in Physics of the Sun and Stars: A New Era in Helio- and Asteroseismology. p. 503

Silva Aguirre V., et al., 2015, MNRAS, 452, 2127

Silva Aguirre V., et al., 2017, ApJ, 835, 173

Sing D. K., 2010, A&A, 510, A21

Skrutskie M. F., et al., 2006, AJ, 131, 1163

Smirnov N., 1948, Ann. Math. Statist., 19, 279

Spalding C., Batygin K., 2017, AJ, 154, 93

Steffen J. H., Agol E., 2005, MNRAS, 364, L96

Steffen J. H., et al., 2012, Proceedings of the National Academy of Science, 109, 7982

Stratos K., 2019. pp 1095–1104, doi:10.18653/v1/N19-1113

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

Sullivan P. W., et al., 2015, ApJ, 809, 77

Tamayo D., Rein H., Petrovich C., Murray N., 2017, ApJ, 840, L19

Teachey A., Kipping D. M., Schmitt A. R., 2018, AJ, 155, 36

Terquem C., Papaloizou J. C. B., 2007, ApJ, 654, 1110

Thompson S. E., et al., 2018, The Astrophysical Journal Supplement Series, 235, 38

Tremaine S., Dong S., 2012, AJ, 143, 94

Trifonov T., Rybizki J., Kürster M., 2019, A&A, 622, L7

Uehara S., Kawahara H., Masuda K., Yamada S., Aizawa M., 2016, ApJ, 822, 2

Van Eylen V., Albrecht S., 2015, ApJ, 808, 126

Vanderburg A., et al., 2015, ApJ, 800, 59

Vanderburg A., et al., 2016, ApJ, 827, L10

Vanderburg A., et al., 2018, AJ, 156, 46

Villanueva Steven J., Dragomir D., Gaudi B. S., 2019, AJ, 157, 84

Vogt S. S., Penrod G. D., 1983, Publications of the Astronomical Society of the Pacific, 95, 565

Vogt S. S., Penrod G. D., Hatzes A. P., 1987, ApJ, 321, 496

Walsh K. J., Morbidelli A., Raymond S. N., O'Brien D. P., Mandell A. M., 2011, Nature, 475, 206

Wang J., Ford E. B., 2011, MNRAS, 418, 1822

Wang J., et al., 2015, ApJ, 815, 127

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

Weiss L. M., et al., 2017, AJ, 153, 265

Weiss L. M., et al., 2018, AJ, 155, 48

Weissbein A., Steinberg E., Sari R., 2012, arXiv e-prints, p. arXiv:1203.6072

Williams J. P., Cieza L. A., 2011, ARA&A, 49, 67

Wilson R. E., 1994, PASP, 106, 921

Wilson R. E., Devinney E. J., 1971, ApJ, 166, 605

Winn J. N., 2010, Exoplanet Transits and Occultations. pp 55–77

Winn J. N., Fabrycky D. C., 2015, ARA&A, 53, 409

Winn J. N., Fabrycky D., Albrecht S., Johnson J. A., 2010, ApJ, 718, L145

Witzel G., et al., 2018, ApJ, 863, 15

Wood D. B., 1971, AJ, 76, 701

Wright J. T., Sigurdsson S., 2016, ApJ, 829, L3

Wright J. T., Upadhyay S., Marcy G. W., Fischer D. A., Ford E. B., Johnson J. A., 2009, ApJ, 693, 1084

Wright E. L., et al., 2010, AJ, 140, 1868

Wright J. T., et al., 2011, PASP, 123, 412

Wright J. T., Cartier K. M. S., Zhao M., Jontof-Hutter D., Ford E. B., 2016, ApJ, 816, 17

Yao X., et al., 2019, AJ, 157, 37

Yee J. C., Gaudi B. S., 2008, ApJ, 688, 616

Yi S., Demarque P., Kim Y.-C., Lee Y.-W., Ree C. H., Lejeune T., Barnes S., 2001, ApJS, 136, 417

Zhu W., 2020, AJ, 159, 188

Zink J. K., Christiansen J. L., Hansen B. M. S., 2019, MNRAS, 483, 4479

Zipf G., 1935, The Psychobiology of Language: An Introduction to Dynamic Philology. M.I.T. Press, Cambridge, Mass.

de Wit J., Gillon M., Demory B. O., Seager S., 2012, A&A, 548, A128

van Helden A., 1976, Journal for the History of Astronomy, 7, 1