7-30-2018

Systematic Mischaracterization of Exoplanetary System Dynamical Histories from a Model Degeneracy near Mean-motion Resonance

John H. Boisvert
University of Nevada, Las Vegas, johnb@physics.unlv.edu

Benjamin E. Nelson
CIERA – Northwestern University

Jason H. Steffen
University of Nevada, Las Vegas, jason.steffen@unlv.edu

Repository Citation

Boisvert, J. H., Nelson, B. E., Steffen, J. H. (2018). Systematic Mischaracterization of Exoplanetary System Dynamical Histories from a Model Degeneracy near Mean-motion Resonance. *Monthly Notices of the Royal Astronomical Society, 480*(2), 2846-2852.
http://dx.doi.org/10.1093/mnras/sty2023

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **480,** 2846–2852 (2018)
Advance Access publication 2018 July 30

doi:10.1093/mnras/sty2023

Systematic mischaracterization of exoplanetary system dynamical histories from a model degeneracy near mean-motion resonance

John H. Boisvert,[1]★ Benjamin E. Nelson[2] and Jason H. Steffen[1]★

[1]*Department of Physics and Astronomy, University of Nevada, Las Vegas, 4505 S Maryland Parkway, Box 454002, Las Vegas, NV 89154-4002, USA*
[2]*CIERA – Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3112, USA*

Accepted 2018 July 20. Received 2018 July 20; in original form 2018 June 6

ABSTRACT

There is a degeneracy in the radial velocity exoplanet signal between a single planet on an eccentric orbit and a two-planet system with a period ratio of 2:1. This degeneracy could lead to misunderstandings of the dynamical histories of planetary systems as well as measurements of planetary abundances if the correct architecture is not established. We constrain the rate of mischaracterization by analysing a sample of 60 non-transiting, radial velocity systems orbiting main-sequence stars from the NASA Exoplanet Archive (NASA Archive) using a new Bayesian model comparison pipeline. We find that 15 systems (25 per cent of our sample) show compelling evidence for the two-planet case with a confidence level of 95 per cent.

Key words: methods: data analysis – techniques: radial velocities – planetary systems.

1 INTRODUCTION

The architectures of planetary systems give insight into their formation and dynamical histories. For example, interactions with the protoplanetary disc tend to drive adjacent planets into first-order, mean-motion resonances (MMRs; such as the 2:1), while simultaneously damping their eccentricities to values that are difficult to measure (Lee & Peale 2002; Tinney et al. 2006). On the other hand, planet–planet scattering (Chatterjee et al. 2008; Ford & Rasio 2008) or Kozai–Lidov oscillations (Kozai 1962; Lidov 1962; Fabrycky & Tremaine 2007) can produce single planets with eccentric orbits. While not all planetary systems must pass through these phases of disc migration or eccentricity growth, the system architectures that they produce rarely occur from *in situ* formation. Thus, reliable estimates of their frequencies will reveal the relative importance of these processes in planet formation and evolution in general.

For radial velocity (RV) observations in particular, the challenge in identifying the true system architecture is a degeneracy between two models – one with a single planet with eccentric orbits (single eccentrics) and one with two planets with circular orbits at the 2:1 (circular doubles) (Anglada-Escudé et al. 2010; Wittenmyer et al. 2013). Historically, the single-planet model has been favoured on the grounds of Occam's razor (Kürster et al. 2015), since a system with a single planet is simpler than a system with two. However, the circular double model has the same number of model parameters as single eccentrics (it is just as simple) and it is a consequence of dynamical processes known to occur. These facts motivate careful scrutiny of existing discoveries in order to properly characterize the systems. If circular doubles are more common than currently suggested, then disc migration may be more important than previously thought (Tinney et al. 2006).

The source of the degeneracy between these models is in a first-order expansion of the RV signal of a single eccentric planet

$$RV_{\rm single} \approx K\cos(M+\omega) + Ke\cos(2M+\omega) + \mathcal{O}(e^2), \quad (1)$$

where $RV_{\rm single}$ is the observed radial velocity, K is the velocity semi-amplitude, e is the eccentricity, ω is the longitude of periastron, and M is the mean anomaly, which is a function of time. By comparison, the signal of a circular double is

$$RV_{\rm double} = K_{\rm out}\cos(M_{\rm out}) + K_{\rm in}\cos(M_{\rm in}), \quad (2)$$

where $RV_{\rm double}$ is the observed radial velocity, $K_{\rm out}$ and $K_{\rm in}$ are the velocity semi-amplitudes, and $M_{\rm out}$ and $M_{\rm in}$ are the mean anomalies. At the 2:1 MMR, $M_{\rm in} = 2M_{\rm out}$ and the inner planet signal ($K_{\rm in}$) masquerades as the eccentricity signal (Ke) of the single planet.

This degeneracy is widely known though rarely addressed. Nevertheless, there is precedent for reconsidering certain systems. For example, Kürster et al. (2015) reanalysed RV data for HD 27894 and found that a circular double model was a better fit than the reported single eccentric model. Also, Anglada-Escudé et al. (2010) and Wittenmyer et al. (2013) found similar results for several RV systems. At the same time, new measurements from the *Kepler* mission show that planet pairs near 2:1 are quite common. For example, using the method of Steffen & Hwang (2015) on the *Kepler* DR25 catalogue (Thompson et al. 2017), we estimate that 20 per cent of *Kepler*'s transiting adjacent planet pairs with period ratios between 1 and 6 are within 10 per cent of 2 – including the most prominent peak

★ E-mail: johnb@physics.unlv.edu (JHB); jason.steffen@unlv.edu (JHS)

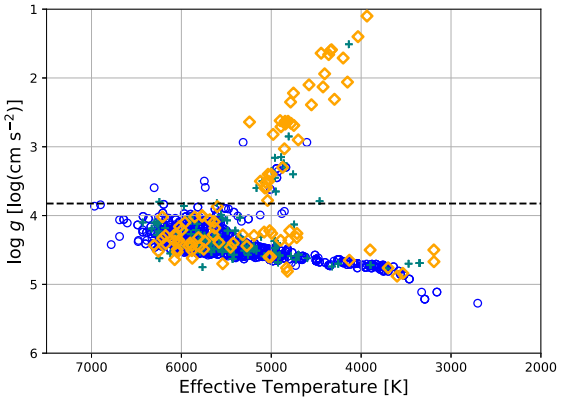

Figure 1. Stellar effective temperature versus stellar surface gravity for the RV multiplanet systems from the NASA Archive as grey crosses and the *Kepler* multiplanet systems as blue circles. Our sample of 95 stars, ignoring stellar type, are shown as orange diamonds. For our sample of main-sequence stars, we select those with $\log g \geq 3.825$; there are 60 main-sequence stars in the main sample and 95 stars in the entire sample.

of the period ratio distribution at 2.17 (Steffen & Hwang 2015). Motivated by these new facts and the results of previous studies, we reanalyse a sample of 60 single eccentric planetary systems using a new Bayesian analysis pipeline in an effort to discover their true architectures.

2 METHODS

Our sample contains 60 systems and comprises every non-transiting RV system from the NASA Exoplanet Archive (NASA Archive), as of 2016 November (Akeson et al. 2013), which is listed as having only a single planet, orbiting a main-sequence star, and whose system properties were derived from a single data set. Our pipeline did not have the capability to analyse multiple data sets when the analysis began. We did not limit our sample by eccentricity. Our main results will focus on the main-sequence stars, but we will also report on an extended sample which ignores stellar type. The extended sample contains 95 systems, which is nearly a quarter of all RV-discovered single-planet systems. Fig. 1 shows how we determine stellar types based on their reported surface gravity and how we select the main-sequence sample.

2.1 The pipeline

Our pipeline estimates the Bayes factor – the ratio of the probabilities of the RV data given the circular double model and the single eccentric model – to quantitatively compare the two models. For each system, we test four planetary system models: a single eccentric; two circular doubles [one with a period ratio fixed at 2 and the other fixed at 2.17 – where there are two large peaks in the period ratio distribution from *Kepler* (Steffen & Hwang 2015)]; and a 'floating' circular double with no period ratio constraint. This last model has an additional model parameter, but the Bayes factor calculation can account for different numbers of model parameters. Our primary results work with the two fixed models given the compelling theoretical and observational reasons to consider them, the fact that the numbers of model parameters are identical (and thus more directly comparable), and because a narrow-band signal at a fixed period ratio is less susceptible to a false positive detection from stellar RV jitter or statistical noise.

The parameters for the single eccentric model are: period (P), velocity semi-amplitude (K), eccentricity (e), longitude of periastron (ω), and mean anomaly at the time of the earliest RV measurement [$M_o \equiv M(t_o)$]. The parameters for the circular double models are: outer planet period (P_{out}), the velocity semi-amplitude for the outer/inner planets (K_{out}/K_{in}), and mean anomaly at the time of the earliest RV measurement for the outer/inner planets ($M_{o,\,out}/M_{o,\,in}$). The floating circular double model includes the inner planet orbital period (P_{in}). Each model also has a linear trend (A) and a velocity offset (C).

Our model fitting is a three-step process. First, we determine the starting values for our Markov Chain-Monte Carlo (MCMC) by maximizing the likelihood function:

$$\ln P(t, RV, \sigma_{RV}|\theta) = -\frac{1}{2}\sum_n \left[\frac{(RV_n - \mathcal{M}(\theta, t_n))^2}{\sigma^2_{RV,n}} + \ln 2\pi\sigma^2_{RV,n}\right], \quad (3)$$

where t, RV, and σ_{RV} are the observed times, RV measurements, and RV errors, respectively. $\mathcal{M}$ is the RV model and θ are its associated parameters. We draw our set of initial conditions for the maximum likelihood estimation (MLE) from the NASA Archive. The time of periastron passage is used to determine the initial M_o. We first fit for C, fixing the other parameters at their nominal values and setting $A = 0$. We next fit for A and C simultaneously. Some systems did not have K, ω, and/or the time of periastron passage reported on the NASA Archive. In those cases, an MLE was done with the missing quantities as the only free parameters. We initialized the fixed circular double models to their first-order, single eccentric equivalent values using equations (1) and (2). For the floating circular double, the inner planet orbital period is initialized to either the 2:1 or the 2.17:1, depending on which fixed model produced a larger Bayes factor.

The second step in our pipeline estimates the posterior distributions of the model parameters using an ensemble sampler MCMC (Foreman-Mackey et al. 2013). Each run has 30 Markov chains, thins the chains every hundred steps, and ignores the first 20 per cent of the chain as burn-in. We allow the chains to evolve until they yield a set of at least 10 000 independent samples per model per RV data set. We measure the autocorrelation length after each run to determine the number of independent samples. If the number of independent samples falls short of 10 000, then the autocorrelation length is used to determine how many additional steps are needed to yield 10 000 independent samples and the MCMC is rerun with the new number of steps. The different chains were initialized using the parameter values from the MLE, with each parameter scattered by a sufficiently small amount to allow the ensemble sampler to fill the posterior mode.

We impose a modified Jeffery's prior for the orbital period and velocity semi-amplitude: $p(X) = [(1 + X) \times \ln(1 + X_{max}/X_0)]^{-1}$ with bounds between 0–10 000 d and 0–2000m s^{-1}, respectively, and X_0 equal to 1 d and 1 m s^{-1}, respectively. We use this prior because it is normalizable, objective, and intended for scalable parameters that could have zero as a value. We use uniform priors for the remaining parameters, (e, ω, M_0), because they are also normalizable and objective. We sample the parameters for the single eccentric model in $\{P, K, \sqrt{e}\sin(\omega), \sqrt{e}\cos(\omega), \omega + M_0\}$-space in order to maintain uniform priors (Daniel Foreman-Mackey & Ben Nelson private communication).

The prior bounds for K, K_{out}, and K_{in} are between 0 and 2000m s^{-1}. The prior bounds for P, P_{out}, and P_{in} are between 0 and 10 000 d. The prior bounds for $\sqrt{e}\sin(\omega)$ and $\sqrt{e}\cos(\omega)$ are such

that $0 < (\sqrt{e}\sin(\omega))^2 + (\sqrt{e}\cos(\omega))^2 < 1$, i.e. $0 < e < 1$. The prior bounds for $(M_0 + \omega)$, $M_{0,\,\mathrm{out}}$, and $M_{0,\,\mathrm{in}}$ are between -2π and 4π. These limits allow the Markov chains to cross the 0 and 2π coordinate singularities while remaining well behaved. Furthermore, these values are modded by 2π before doing any calculations. The prior bounds for C are between $-100\,000$ and 100 000m s^{-1} to accommodate the wide range in offset values in the real sample.

Finally, we estimate the Bayes factors between the single eccentric model and the circular double models by taking the ratio of the fully marginalized likelihoods (FMLs, i.e. Bayesian evidence) for the two models. We approximate the FML using an importance sampling algorithm where the sampling distribution is informed by a set of posterior samples taken from the aforementioned MCMC (Nelson et al. 2016). For each system we take the larger of the Bayes factor for the two fixed circular double models.

In this context, importance sampling is essentially a general form of Monte Carlo integration to estimate the FML, $\mathcal{Z}$. The value of $\mathcal{Z}$ is the integral over the prior probability distribution $p(\theta)$ times the likelihood function $\mathcal{L}(\theta) \equiv p(t, RV, \sigma_{RV}|\theta)$, i.e.

$$\mathcal{Z} = \int p(\theta)\mathcal{L}(\theta)\mathrm{d}\theta. \tag{4}$$

We multiply the numerator and denominator of the integrand by $g(\theta)$, a distribution over the model parameters with a known normalization.

$$\mathcal{Z} = \int \frac{\mathcal{L}(\theta)p(\theta)}{g(\theta)} g(\theta)\mathrm{d}\theta. \tag{5}$$

Equation (5) is in a form such that $\mathcal{Z}$ can be estimated numerically by drawing N samples from $g(\vec{\theta})$,

$$\widehat{\mathcal{Z}} \approx \frac{1}{N} \sum_{\theta_i \sim g(\theta)} \frac{\mathcal{L}(\theta_i)p(\theta_i)}{g(\theta_i)}. \tag{6}$$

The key to an accurate and efficient estimate of $\widehat{\mathcal{Z}}$ lies in choosing an appropriate $g(\theta)$. Assuming our parameter space contains one dominant posterior mode, we choose a multivariate normal $\mathcal{N}(\vec{\mu}_g, \vec{\Sigma}_g)$, where $\vec{\mu}_g$ and $\vec{\Sigma}_g$ describe the mean vector and covariance matrix of the model parameters, respectively. After we perform an MCMC on a particular model/data set, we can estimate $\vec{\mu}_g$ and $\vec{\Sigma}_g$ using a set of posterior samples. That information is fed into our importance sampling algorithm to estimate $\widehat{\mathcal{Z}}$ for that model. Nelson et al. (2016), Guo (2012), and Weinberg, Yoon & Katz (2013) provide more detailed prescriptions and investigations of this method.

2.2 Pipeline characterization

We characterized the pipeline efficiency with an ensemble of 1000 synthetic RV time series whose system and data properties match the real systems. We use the Bayes factors of these synthetic systems to characterize our model comparison pipeline. This Monte Carlo simulation was initialized as follows:

The start time (t_0) is a uniform random draw between 1 and 1000 d. The number of observations is drawn from the real systems with a normally distributed adjustment with a standard deviation 10 per cent of the nominal value rounded to the nearest whole number. The observation time series is produced by selecting a set of observation differences ($t_i - t_{i-1}$) from the real distribution of observation differences with a similar, normally distributed 10 per cent variation added to each difference. The number of orbits is the number of orbits of a randomly chosen real system with a normally distributed 10 per cent variation.

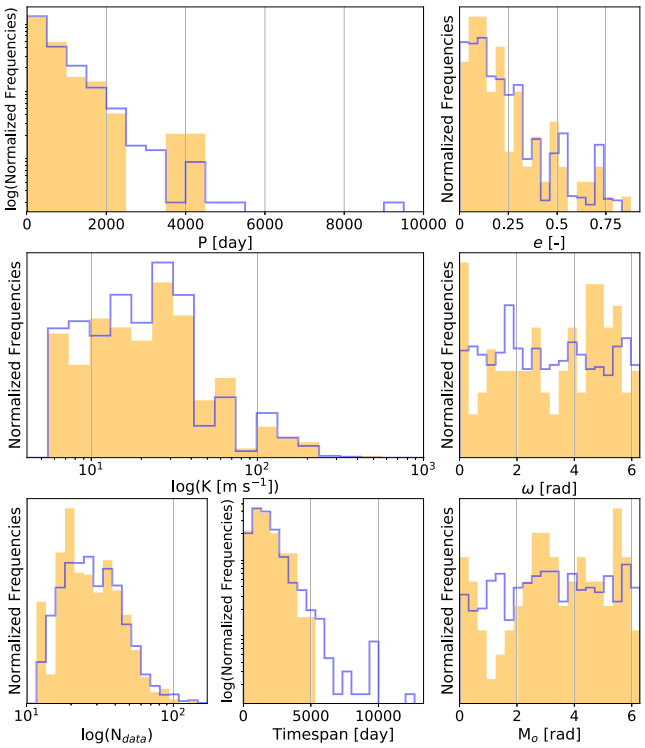

Figure 2. Property distributions for our sample of 95 real systems from the NASA Exoplanet Archive in orange and the 1000 synthetic time series in blue.

We determine the orbital period (P) using the selected number of orbits and the observation time series. The velocity semi-amplitude (K) and the eccentricity (e) are separate random draws from the real systems. The mean anomaly of the start time (M_0) and argument of periastron (ω) are randomly drawn between 0 and 2π. The linear trend (A) is a 10 per cent variation to a random draw from the real systems.

We assume that the RV errors are normally distributed with a standard deviation that is the quadrature sum of stellar jitter and instrumental and photon noise. The instrumental and photon noises (σ_{RV}) are drawn randomly from the RV errors of the real systems and our error bars are assigned to this value. Stellar jitter is selected from a log uniform distribution between 0.5 and 5 m s^{-1}. The observation errors are added to the synthetic RV measurement – not to the error in the RV measurement. Fig. 2 shows the parameter distributions for the 1000 synthetic time series and the real systems as reported in the NASA Archive.

The resulting Bayes factors from this characterization are shown as the blue distribution in Fig. 3. The vertical lines denote the 95th and 90th percentiles of the distribution. The shape of the distribution is not symmetric, and the vast majority of our synthetic data sets favour the single-planet case – as expected since the synthetic systems were constructed to be single eccentrics. Real systems with Bayes factors larger than those thresholds may be circular double systems mischaracterized as single eccentrics.

The approach outlined above is different from earlier studies. For example, Wittenmyer et al. (2013) used the reduced χ^2 to determine the preferred model and refined their results with stability tests using the N-body integrator *Mercury* (Chambers & Migliorini 1997). Anglada-Escudé et al. (2010) randomized individual sets of data to calculate the false positive rate per

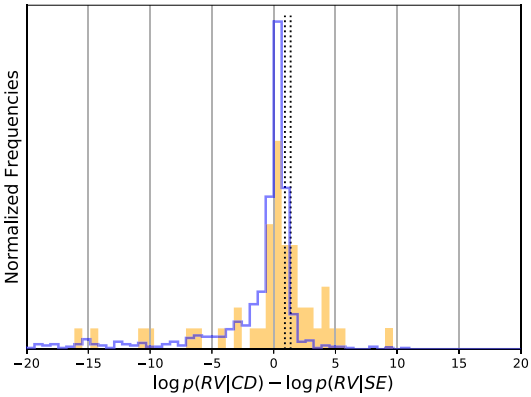

Figure 3. The log Bayes factor distribution for the 1000 synthetic single eccentric time series in blue and 60 real systems hosted by main-sequence stars in orange. Here, we compare only the single eccentric model to the fixed, circular double model with the largest Bayes factor. The 95th and 90th percentiles are indicated with the dotted lines near a Bayes factor of 24 and 8, respectively.

system. Their model selection was also based on the reduced χ^2 of least-squares fitting. In this work, we use an FML to calculate the Bayes factor for the model comparison and we estimate our false positive rate by analysing a large simulated data set with our pipeline.

3 RESULTS

After analysing our synthetic systems, we ran our sample of 60 real systems (95 systems for the extended sample) through the same pipeline. Fig. 3 shows that the Bayes factor distributions for the synthetic and real systems (in orange) are not similar. We find that 15 (25 per cent) of the systems have Bayes factors larger than the 95th percentile of the synthetic systems. (For the extended sample of 95 systems, the numbers are 30 and 31 per cent of the entire sample, respectively.) Nine of these systems prefer the 2.17:1 model (22 of the extended sample) while the remaining six (eight from the extended sample) prefer the 2:1 model. Assuming a false positive rate of 5 per cent from our 95 per cent confidence level, our estimate of the number of false positives is 0.75 ± 0.87 (1.5 ± 1.2 for the extended sample). The systems from the extended sample that prefer the fixed circular double model, the model parameters, and Bayes factors are shown in Table 1. A CSV file containing the model parameters with errors for all four models, Bayes factors between the circular double models and the single eccentric model, and percentile of the best fixed model for each system in the extended sample are available online as Table 2.

We examine the consequences of these potential discoveries on several distributions of planet properties. Fig. 4 shows the planet mass versus orbital period for known RV planets along with the new planets favoured by our analysis orbiting main-sequence stars. These potential new systems lie well within the range of values measured in known systems. We point out, however, that some systems may yet be false positives. For instance, there are a few candidate circular double systems that would be hot Jupiters (planet with $P \lesssim 10$ d) with interior companions. Presently, there is only a single known system [WASP-47 (Becker et al. 2015)] where a hot Jupiter has a known interior companion. And the period ratio in this case is over 5:1 – far from the degeneracy we consider here. However, the hot Jupiter has an outer companion with a period ratio near 2.17. Fig. 5 shows how the predicted mass ratios for the main-sequence systems that favour the two-planet model compare with the mass ratios for RV systems on the NASA Archive.

Our primary analysis does not include stellar jitter (even though our synthetic data have jitter added to the simulated data). We made this choice for a number of reasons. One is that since we are considering a fixed period ratio, only noise that occurs at that specific frequency could produce a spurious signal. Most sources of stellar noise occur on much different time-scales. The stellar rotation periods (typically ranging from 4 to 40 d; McQuillan, Mazeh & Aigrain 2014) are shorter than the inner planet periods for most of these systems. Stellar p-mode oscillations have typical time-scales of 5–15 min (Haywood 2015). And surface granulation variations last minutes to hours, with the largest granules remaining on the surface of stars for about a day (Del Moro et al. 2004; Haywood 2015). The time-scales of long-term stellar activity arising from the cyclical appearance of starspots are of the order of years to decades (Strassmeier 2009).

These facts support the interpretation that stellar noise is not the cause of the inner companion signal for the majority of our systems. Nevertheless, we did a separate analysis that included a white noise jitter term to all models and found that five of the 15 systems still remain in the 95th percentile of likely two-planet systems, four of which prefer the 2.17:1 architecture. Thus, even if we adopt the much more conservative approach – which assumes stellar jitter does indeed affect our data at precisely the relevant time-scales – we still see a number of systems that favour the two-planet models.

While our results are primarily from the fixed circular double models, we examined the results of a floating circular double model in order to estimate the likely distribution of orbital periods for the inner companion. We analysed the real and synthetic systems with the floating circular double model and find an even larger portion of the systems that have Bayes factors above the 95th percentile – 19 systems, 32 per cent of the main-sequence sample, (41 systems, 43 per cent of the extended sample) with an estimated false positive rate of 0.95 ± 0.97 (2.1 ± 1.4 for the extended sample). Of these 19 systems, nine prefer the floating circular double model, six prefer the fixed 2.17:1 model, and the remaining four prefer the 2:1 double circular model. Four systems remain in the 95th percentile when including stellar jitter in the model as a white noise term.

We show the period ratio posteriors that result from this analysis for these 19 systems and the synthetic systems in Fig. 6. These histograms show the combined, period ratio posterior distribution from fitting the circular double model without a constraint on orbital periods to the 19 systems and to the synthetic systems. The distribution for the synthetic systems clearly shows the degeneracy at the location of the 2:1 MMR. If the real systems (in orange) were single-planet systems, then the expected distribution should be the same as for the synthetics. However, the two distributions differ significantly. In fact, the distribution for the real systems mirrors the period ratio distribution from the *Kepler* data (Steffen & Hwang 2015). Most of the combined posteriors favour period ratios just wide of the 2:1 or between 2.15 and 2.2. Only a few systems preferred the circular double model near the 2:1 because the degeneracy is located at the 2:1 and the power to distinguish between the models diminishes. Thus, in that regime, more data with appropriate phase-sampling are essential to distinguish between the models.

Table 1. The preferred fixed circular double models of the extended sample that has Bayes factor larger than the 95th percentile of the synthetic systems. Under star type, MS and nMS refer to main-sequence and non-main-sequence stars, respectively. The order of the table is by Bayes factor.

Star name	Star type	Model	P_{out} (d)	K_{out} (m s^{-1})	M_{out} (rad)	P_{in} (d)	K_{in} (m s^{-1})	M_{in} (rad)	A (km s^{-1} day^{-1})	C (m s^{-1})	Bayes factor
HD 240237	nMS	2.17:1	770.45 ± 3.09	69.75 ± 1.59	5.45 ± 0.04	355.05 ± 3.09	57.52 ± 2.57	1.43 ± 0.08	9.96 ± 2.49	82.31 ± 1.54	1.09×10^{85}
HD 2952	nMS	2.17:1	318.83 ± 0.27	23.16 ± 0.94	4.13 ± 0.05	146.92 ± 0.27	15.90 ± 0.71	4.41 ± 0.10	6.37 ± 0.62	3.15 ± 0.74	6.88×10^{57}
α Arietis	nMS	2.17:1	372.60 ± 0.52	29.67 ± 0.52	4.31 ± 0.04	171.70 ± 0.52	10.25 ± 0.46	6.17 ± 0.10	−9.5 ± 0.64	0.91 ± 0.45	9.80×10^{49}
HD 96127	nMS	2.17:1	632.39 ± 2.19	96.66 ± 1.50	5.95 ± 0.04	291.42 ± 2.19	41.02 ± 1.26	1.41 ± 0.07	22.83 ± 3.49	−923.95 ± 0.93	9.63×10^{43}
HD 95089	nMS	2.17:1	496.57 ± 2.88	20.65 ± 0.55	2.40 ± 0.05	228.83 ± 2.88	6.73 ± 0.57	4.64 ± 0.11	−4.47 ± 1.39	0.23 ± 0.35	3.75×10^{16}
11 Ursae Minoris	nMS	2.17:1	513.22 ± 0.75	185.38 ± 1.15	4.30 ± 0.02	236.51 ± 0.75	19.43 ± 1.15	4.80 ± 0.08	−11.14 ± 2.18	−10.39 ± 0.84	1.06×10^{15}
HD 136418	nMS	2.17:1	474.99 ± 1.26	41.62 ± 0.39	5.63 ± 0.03	218.89 ± 1.26	10.41 ± 0.47	2.86 ± 0.08	−6.93 ± 1.1	−4.99 ± 0.42	8.03×10^{12}
HD 81688	nMS	2.17:1	184.08 ± 0.17	60.05 ± 0.92	5.47 ± 0.04	84.83 ± 0.17	6.96 ± 0.97	5.44 ± 0.15	4.91 ± 1.53	−1.21 ± 0.67	1.41×10^{11}
HIP 57050	MS	2.17:1	41.31 ± 0.01	29.43 ± 0.62	2.60 ± 0.11	19.04 ± 0.01	6.98 ± 0.64	3.86 ± 0.22	9.01 ± 0.53	−15.77 ± 0.83	4.27×10^{9}
HD 206610	nMS	2.17:1	628.84 ± 6.03	34.50 ± 0.55	1.67 ± 0.05	289.79 ± 6.03	6.92 ± 0.79	3.06 ± 0.09	−17.3 ± 1.7	19.30 ± 0.47	1.59×10^{7}
HD 216770	MS	2:1	121.77 ± 0.38	14.05 ± 2.97	4.30 ± 0.33	60.88 ± 0.38	33.86 ± 3.44	0.79 ± 0.18	36.91 ± 6.87	31 146.80 ± 1.92	1.61×10^{5}
HD 114386	MS	2.17:1	955.46 ± 15.81	33.79 ± 1.55	3.92 ± 0.11	440.31 ± 15.81	14.08 ± 1.96	2.23 ± 0.22	−5.24 ± 3.82	33 367.71 ± 1.28	8.80×10^{4}
o Coronae Borealis	nMS	2:1	187.61 ± 0.13	33.64 ± 0.67	6.00 ± 0.05	93.81 ± 0.13	8.64 ± 0.67	4.66 ± 0.10	1.92 ± 0.57	1.82 ± 0.49	8.66×10^{4}
HD 101930	MS	2.17:1	71.30 ± 0.17	17.93 ± 0.41	1.38 ± 0.04	32.86 ± 0.17	2.37 ± 0.50	1.49 ± 0.20	−10.72 ± 3.43	18 363.19 ± 0.35	3.00×10^{4}
14 Andomedae	nMS	2.17:1	185.89 ± 0.22	99.80 ± 1.41	4.90 ± 0.04	85.66 ± 0.22	6.56 ± 1.61	4.60 ± 0.22	−7.64 ± 2.96	1.68 ± 1.24	2.59×10^{4}
HD 180902	nMS	2.17:1	482.22 ± 2.92	28.68 ± 1.16	0.30 ± 0.04	222.22 ± 2.92	4.20 ± 0.88	1.05 ± 0.23	−3.74 ± 2.69	9.27 ± 0.54	1.83×10^{4}
HD 218566	MS	2:1	225.54 ± 0.14	7.60 ± 0.24	0.90 ± 0.07	112.77 ± 0.14	2.47 ± 0.26	0.17 ± 0.17	0.61 ± 0.11	0.82 ± 0.19	1.70×10^{4}
GJ 649	MS	2:1	601.38 ± 2.17	11.55 ± 0.31	3.41 ± 0.11	300.69 ± 2.17	2.88 ± 0.39	1.83 ± 0.32	0.86 ± 0.31	6.18 ± 0.46	1.10×10^{4}
75 Ceti	nMS	2.17:1	694.41 ± 1.40	37.12 ± 0.74	3.84 ± 0.04	320.01 ± 1.40	3.31 ± 0.70	4.32 ± 0.21	4.59 ± 0.48	0.40 ± 0.47	1.04×10^{4}
HD 27894	MS	2.17:1	17.97 ± 0.01	58.40 ± 0.49	5.04 ± 0.07	8.28 ± 0.01	3.91 ± 0.73	0.50 ± 0.17	−29.59 ± 10.58	82 907.65 ± 1.94	6118.83
HD 32518	nMS	2.17:1	157.45 ± 0.19	117.90 ± 2.19	6.00 ± 0.03	72.56 ± 0.19	9.90 ± 2.38	4.92 ± 0.17	13.84 ± 4.14	−11.90 ± 1.26	3259.36
HD 231701	MS	2.17:1	141.30 ± 0.35	41.56 ± 1.48	1.32 ± 0.10	65.11 ± 0.35	12.81 ± 3.37	2.98 ± 0.23	−8.44 ± 6.3	−0.02 ± 1.81	993.37
γ_1 Leonis	nMS	2:1	428.87 ± 0.17	206.77 ± 0.62	1.09 ± 0.01	214.43 ± 0.17	31.92 ± 0.70	4.99 ± 0.03	9.44 ± 0.86	178.66 ± 0.53	598.25
HD 2638	MS	2.17:1	3.45 ± 0.00	66.75 ± 0.45	5.26 ± 0.09	1.59 ± 0.00	1.36 ± 0.46	4.44 ± 0.35	49.46 ± 13.72	9619.32 ± 2.32	407.32
HD 31253	MS	2:1	464.44 ± 0.64	10.75 ± 0.34	0.27 ± 0.06	232.22 ± 0.64	3.58 ± 0.33	2.87 ± 0.13	0.59 ± 0.17	1.96 ± 0.25	295.22
HD 221287	MS	2.17:1	452.51 ± 1.00	73.50 ± 1.37	2.10 ± 0.04	208.53 ± 1.00	12.95 ± 2.46	2.34 ± 0.05	10.72 ± 2.36	−21 861.09 ± 1.29	124.62
HD 190647	MS	2:1	931.10 ± 76.66	30.25 ± 2.44	3.81 ± 0.44	465.55 ± 76.66	7.90 ± 1.04	2.94 ± 0.93	−20.43 ± 5.67	−40 266.66 ± 1.65	82.74
HD 220773	MS	2:1	2877.77 ± 87.74	14.60 ± 1.51	1.63 ± 0.12	1438.88 ± 87.74	10.15 ± 1.42	4.80 ± 0.25	2.89 ± 0.72	−4.96 ± 1.05	54.06
HD 330075	MS	2.17:1	3.39 ± 0.00	106.81 ± 0.73	5.87 ± 0.01	1.56 ± 0.00	0.49 ± 0.49	3.22 ± 1.21	−13.2 ± 4.57	61 278.58 ± 0.40	37.21
HIP 79431	MS	2.17:1	113.99 ± 0.40	155.87 ± 2.20	2.19 ± 0.02	52.53 ± 0.40	30.65 ± 1.82	4.00 ± 0.09	−385.58 ± 24.21	10.53 ± 1.90	26.99

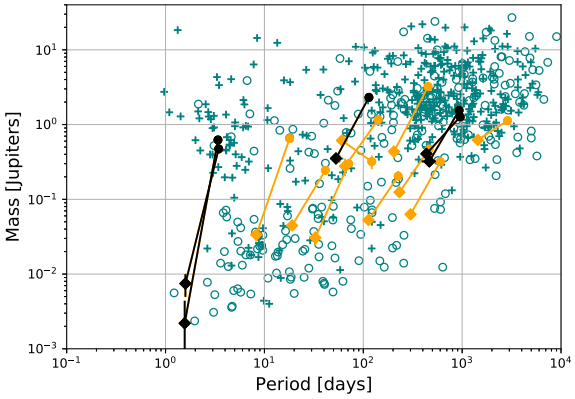

Figure 4. Orbital period versus planetary mass for all RV planets. Systems from the NASA Exoplanet Archive are in teal, with multi and single-planet systems as open circles and crosses, respectively. Each system orbiting a main-sequence star with a measured Bayes factor larger than 95th percentile of the synthetic systems is plotted in orange. Each putative system is represented by a line on the plot, with the diamonds as the inner planet and the circles are the outer companion. Systems that remain in the 95th percentile after including a white noise stellar jitter term are in black. We note that these results lie well within normal parameter values of known systems.

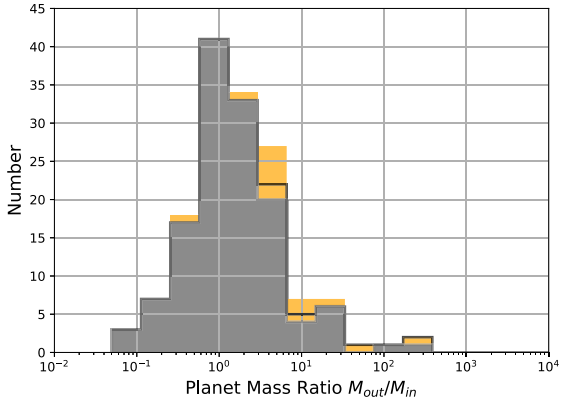

Figure 5. Mass ratio distribution for all RV adjacent planet pairs in grey. The stacked, orange distributions are our systems with Bayes factors larger than 24 (95th percentile) around main-sequence stars. The nature of the signal favours more massive outer planets. The black outline shows the systems that have stellar jitter included and still had Bayes factors larger than the 95th percentile.

4 DISCUSSION AND CONCLUSIONS

There are currently 395 confirmed solitary RV planets, and we re-analysed about 15 per cent of them. Our extended sample contains nearly a quarter of the confirmed RV planets. The distribution of eccentricities, periods, velocity semi-amplitudes, etc. of our extended sample is shown in Fig. 2. If the 15 systems in the main-sequence sample (30 systems in the extended sample) that we identify are indeed circular doubles, then they would increase the number of RV multiplanet systems by ∼12.5 per cent (∼25 per cent) since there are 120 confirmed systems reported with at least two planets discovered by RV. They would also significantly alter the estimated mixture of these two architectures – shifting the relative importance of their implied dynamical histories.

If the fraction of misidentified single eccentrics in the entire NASA Archive is similar to the misidentification fraction seen in our sample, then there could be as many as ∼100 planets missing, or ∼15 per cent of the overall confirmed RV planets (∼120 in the extended sample, or ∼18 per cent of the overall confirmed RV planets). Moreover, the apparent propensity for some systems to cluster around period ratios near 2.17 is a further indication that there is something fundamental, but still unknown, that attracts planet pairs into this period ratio. We encourage observers to consider planning follow-up observations of these systems and make additional measurements at phases where the degeneracy is at its weakest. New observations near these phases could confirm or refute the existence of these putative interior companions. The success of such a campaign opens the door to identifying the architectures of the systems where the preferred model is still ambiguous.

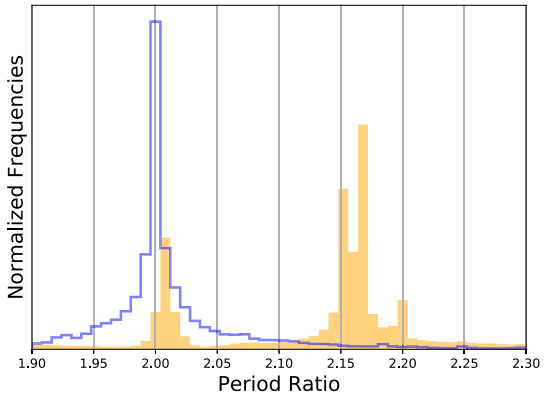

Figure 6. The posterior distributions for the period ratio when considering the inner planet period as a free parameter. The blue distribution is the 1000 synthetic single eccentric systems. This distribution peaks at 2:1 – the location of the degeneracy. The orange distribution is the 19 systems with Bayes factors larger than the 95th percentile thresholds that are hosted by main-sequence stars.

ACKNOWLEDGEMENTS

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. http://exoplanetarchive.ipac.caltech.edu. The authors declare no competing interests. JHB wrote the model fitting software, did the analysis, and wrote the paper. JHS assisted with doing the analysis and editing the paper. BEN wrote the analysis software. We would like to thank Dan Foreman-Mackey for valuable conversations. JHB and JHS acknowledge support from NASA under grant NNH13ZDA001N-OSS issued through the Origins of Solar Systems programme. BEN acknowledges support from the Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and the Data Science Initiative at Northwestern University.

REFERENCES

Akeson R. L. et al., 2013, PASP, 125, 989
Anglada-Escudé G., López-Morales M., Chambers J. E., 2010, ApJ, 709, 168
Becker J. C., Vanderburg A., Adams F. C., Rappaport S. A., Schwengeler H. M., 2015, ApJ, 812, L18
Chambers J. E., Migliorini F., 1997, in AAS/Division for Planetary Sciences Meeting Abstracts #29. p. 1024

Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., 2008, ApJ, 686, 580
Del Moro D., Berrilli F., Duvall T. L. Jr., Kosovichev A. G., 2004, Sol. Phys., 221, 23
Fabrycky D., Tremaine S., 2007, ApJ, 669, 1298
Ford E. B., Rasio F. A., 2008, ApJ, 686, 621
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Guo P. C., 2012, PhD thesis, University of Florida
Haywood R. D., 2015, PhD thesis, University of St Andrews
Kozai Y., 1962, AJ, 67, 591
Kürster M., Trifonov T., Reffert S., Kostogryz N. M., Rodler F., 2015, A&A, 577, A103
Lee M. H., Peale S. J., 2002, ApJ, 567, 596
Lidov M. L., 1962, Planet. Space Sci., 9, 719
McQuillan A., Mazeh T., Aigrain S., 2014, ApJS, 211, 24
Nelson B. E., Robertson P. M., Payne M. J., Pritchard S. M., Deck K. M., Ford E. B., Wright J. T., Isaacson H. T., 2016, MNRAS, 455, 2484
Steffen J. H., Hwang J. A., 2015, MNRAS, 448, 1956
Strassmeier K. G., 2009, A&AR, 17, 251
Thompson S. E. et al., 2017, ApJS, 235, 38
Tinney C. G., Butler R. P., Marcy G. W., Jones H. R. A., Laughlin G., Carter B. D., Bailey J. A., O'Toole S., 2006, ApJ, 647, 594
Weinberg M. D., Yoon I., Katz N., 2013, preprint (arXiv:1301.3156)
Wittenmyer R. A. et al., 2013, ApJS, 208, 2

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.