On the Orbital Spacing Pattern of Kepler Multiple-planet Systems

Chao-Feng Jiang, Ji-Wei Xie ©, and Ji-Lin Zhou ©
School of Astronomy and Space Science & Key Laboratory of Modern Astronomy and Astrophysics in Ministry of Education, Nanjing University, 210093, People’s Republic of China; jwxie@nju.edu.cn

Received 2020 January 25; revised 2020 August 2; accepted 2020 August 16; published 2020 September 24

Abstract

The Kepler space mission has detected a large number of exoplanets in multiple transiting planet systems. Previous studies found that these Kepler multiple-planet systems exhibit an intra-system uniformity, namely planets in the same system have similar sizes and correlated orbital spacings. However, it is important to consider the possible role of selection effects due to observational biases. In this paper, we revisit the orbital spacing aspect of the pattern after taking observational biases into account using a forward modeling method. We find that orbital spacings, in terms of period ratios, of Kepler multiple-planet systems are significantly correlated only for those tightly packed systems, and the transition from correlation to noncorrelation is abrupt with a boundary at mean period ratio $PR \sim 1.5–1.7$. In this regard, the pattern of orbital spacing is more like a dichotomy rather than a global correlation. Furthermore, we find that such an apparent orbital spacing dichotomy could be essentially a projection of a dichotomy that related to mean motion resonance (MMR), which we dub as MMR dichotomy, and itself could be a natural result of planet migration and dynamical evolution.

Unified Astronomy Thesaurus concepts: Exoplanet astronomy (486); Astrostatistics (1882)

1. Introduction

Hitherto, the number of detected exoplanets has been boosted to over 4000 thanks to various ground-based and space-based surveys, among which the Kepler mission (Borucki et al. 2010) has played a major role in contributing over two-thirds of these discoveries.1 The bulk of exoplanets detected by the Kepler mission are so-called super-Earth or sub-Neptunes with radii between Earth and Neptune and orbital periods less than several hundred days (Thompson et al. 2018). Although super-Earths are found to be common (Dong & Zhu 2013; Howard 2013; Mullally et al. 2015; Zhu et al. 2018), they do not exist in our solar system, and how they were formed remains an open question (Lissauer et al. 2014; Morbidelli & Raymond 2016).

Among the Kepler discoveries, one of the most valuable parts is the large sample of multiple transiting planet systems (Ragozzine & Holman 2010), which has greatly advanced our knowledge on exoplanets in many aspects, including planetary masses and thus physical compositions (Carter et al. 2012; Wu & Lithwick 2013; Hadden & Lithwick 2014), orbital eccentricities, and inclinations (Fang & Margot 2012; Fabrycky et al. 2014; Xie et al. 2016; Van Eylen et al. 2019), and etc., shedding light on their formation and evolution history (Mills et al. 2016; Owen & Campos Estrada 2020).

In this paper, we focus on the aspect of orbital spacing, which has attracted numerous studies. Bovaird & Lineweaver (2013) and Huang & Bakos (2014) investigated the orbital spacings of Kepler’s multiple systems in a context of extended Titus–Bode law of our solar system. Pu & Wu (2015) found that the orbital spacings of Kepler planets are clustered around the theoretical stability threshold. Some studies investigated the spacings of Kepler planets in terms of orbital period ratio (PR; Lissauer et al. 2011; Steffen 2013; Steffen & Hwang 2015). From the PR distribution, the majority of Kepler planets were found to be not in mean motion resonance (MMR). Nevertheless, the PR distribution has shown overabundances just wide of first-order MMRs and deficits short of them (Fabrycky et al. 2014), which may have implications to planet formation and evolution (Lithwick & Wu 2012; Batygin & Morbidelli 2013; Delisle & Laskar 2014; Xie 2014; Chatterjee & Ford 2015; Millholland & Laughlin 2019).

Recently, Weiss et al. (2018) found that planets orbiting the same host tend to be similar in size (see also in Millholland et al. 2017; Wang 2017) and have regular orbital spacings (i.e., PR correlation), a pattern which they dubbed as “peas in a pod.” However, whether such a pattern is astrophysical or a selection effect due to observational biases is still currently in debate (Gilbert & Fabrycky 2020; Murchikova & Tremaine 2020; Weiss & Petigura 2020; Zhu 2020).

Here, we revisit one aspect of the pattern, i.e., the PR correlation, in detail by taking observational biases into account. This paper is organized as follows. In Section 2, we select different planet samples by applying different criteria. Then, for each planet sample, we evaluate the significance of PR correlation and the effects of observational biases (Section 3.1). We find evidences, in Section 3.2, which show that the orbital spacing pattern is more like a dichotomy rather than a global correlation. In Section 4, we discuss the implications of such an orbital spacing dichotomy. Section 5 is the summary of the paper.

2. Sample

Our study is based on the multiple transiting planet systems detected by the Kepler mission. We use the Q1–Q17 table of Kepler Objects of Interest (hereafter KOIs) from the NASA Exoplanet Archive.2 First, we exclude all the KOIs which are identified as false positives. Second, we adopt three filters as follows to the remaining planetary systems.

1. The multiplicity of planetary systems $N_p \geq 4$.
2. The maximum radius of planets in the systems $R_{\text{max}} \leq 6R_\oplus$, where $R_\oplus$ is the Earth radius.

---

1. http://exoplanet.eu
2. https://exoplanetarchive.ipac.caltech.edu
3. Results

3.1. Revisit the PR Correlation

First, we revisit the PR correlation (Weiss et al. 2018) in different samples in Section 2.

3.1.1. Correlation Evaluation

In the work of Weiss et al. (2018) the authors measured the correlation of the orbital PR of each pair of neighboring planets $p_{i+1}/p_i$ and that of the outer pair of neighboring planets $p_{i+2}/p_{i+1}$. They found a Pearson-R correlation coefficient of 0.46 with a significance of $P$ value $<10^{-3}$, leading to a conclusion that there is a strong correlation among orbital PRs of planets in the same systems. The Pearson correlation coefficient, however, is not appropriate for searching for correlations in a relatively small sample, because it assumes the linear correlation and Gaussian scatter. For this reason, besides using the Pearson correlation coefficient (mainly for comparison with Weiss et al. 2018), we further repeat all the analyses using Kendall’s tau correlation coefficient, which is nonparametric without making either assumptions, and is thus more robust. The detailed procedure is as follows.

Step 1: we calculate Kendall’s tau nonparametric correlation coefficient $\tau_{obs}$ (or Pearson’s correlation coefficient $R_{obs}$) for each sample in Section 2.

Step 2: we randomly scramble PRs of neighboring planets among planetary systems then recalculate the correlation coefficient for each simulated realization $\tau_{sim}$ (or $R_{sim}$).

Step 3: we repeat Step 2 for 10,000 times and calculate the fraction of times with $\tau_{sim} > \tau_{obs}$ (or $R_{sim} > R_{obs}$). This fraction gives the $P$ value $P_{Kendall}$ (or $P_{Pearson}$) of the Kendall correlation test and $1 - P_{Kendall}$ (or $1 - P_{Pearson}$) is the confidence level of the observed PR correlation.

Figure 1 shows the PR correlation evaluation for the four samples as defined in Table 1. For each of the samples described in Table 1, the PR correlation is significant with a confidence of larger than 99.99% in the Kendall correlation test, which is consistent with the result in Weiss et al. (2018) although we use different samples and correlation tests. However, as for the bottom left panel for the result of sample 3, the Pearson test returns a much larger $P$ value of 0.274. This is probably because planet pairs with larger PRs, i.e., $PR > 4$, are included in sample 3. In fact, each panel also shows an apparent trend that the points with larger PRs become more dispersed with respect to the 1:1 ($y = x$) line. In Section 3.2, we will investigate this trend in more detail.

### Table 1

Summary of the Samples 1, 2, 3, and 4 Mentioned in Section 2

| Sample ID | 1 | 2 | 3 | 4 (Weiss + 2018) |
|-----------|---|---|---|-----------------|
| $R_{max} \leq 6R_J$ | Yes | No | Yes | No |
| $PR_{max} \leq 4$ | Yes | Yes | No | Yes |
| $N_{sys}$ | 56 | 60 | 65 | 95 |
| $N_T$ | 0 | 0 | 0 | 53 |
| $N_S$ | 39 | 41 | 46 | 31 |
| $N_E$ | 15 | 17 | 17 | 10 |
| $N_R$ | 2 | 2 | 2 | 1 |

Note. Different cut conditions (i.e., $R_{max} \leq 6R_J$ and $PR_{max} \leq 4$) are applied to some of the samples. $N_{sys}$ is the total number of systems and $N_T$, $N_E$, $N_S$, and $N_R$ are the specific numbers of systems with three, four, five, and six transiting planets respectively.

3. The maximum of PRs of adjacent planets in the systems

$PR_{max} \leq 4.0$.

We adopt the first filter for the reason that systems of lower multiplicities tend to be not dynamically packed and thus have a higher likelihood of missing nontransiting planets in between the transiting planets (see more discussions in Section 4.1.3), causing a systematic overestimation of the PRs of neighboring planets. Through the second filter, we exclude giant planets, allowing us to focus on smaller planets, i.e., super-Earths and sub-Neptunes. We adopt the third filter according to Weiss et al. (2018) for comparison with their results. After all these three filters, we have 56 multiple-planet systems in our nominal sample (sample 1 in Table 1).

For comparison with sample 1, we adjust the above three filters to construct our sample 2 and sample 3. In sample 2, we release the radius cutoff to include those systems which host giant planets with $R_{max} > 6R_J$. In sample 3, we release the spacing cutoff of $PR_{max} \leq 4.0$. Besides, we adopt the same sample of Weiss et al. (2018) as our sample 4. The descriptions of the samples are summarized in Table 1.

#### Figure 1

The period ratio correlation evaluation for the four observed samples (see Table 1 and Section 2). The x-axis and the y-axis in each panel denote the period ratio of the inner pair of neighboring planets ($p_{i+1}/p_i$) and that of the outer pair ($p_{i+2}/p_{i+1}$). On the upper right of each panel, we printed the $P$ value of the Kendall correlation test and Pearson correlation test (Section 3.1.1). The gray dashed line shows the perfect correlation, i.e., $y = x$. We can see that all the samples show strong PR correlation in the Kendall correlation test. However, in the Pearson correlation test, all the samples except sample 3 show strong PR correlation. We note that the relatively weaker PR correlation in sample 3 is probably attributed to the inclusion of planet pairs with larger period ratios, i.e., $PR > 4$. (The axes scale in the bottom left panel is different from the other panels.) In fact, the trend that the period ratio correlation becomes weaker with increasing period ratio can be indeed seen in all the samples.
Simulated Systems

Figure 2. Similar to Figure 1, but for a set of typical Monte Carlo realizations of simulated samples with the assumption that planets are intrinsically randomly paired (see Section 3.1.2). Compared to Figure 1, the period ratio correlations vanish, with much larger $P$ values of the Kendall (Pearson) correlation tests, $P_{\text{Kendall}}(P_{\text{Pearson}})$.

Note, although $P$ values are reported to high precision here, one should not over interpret the numbers in high precision (Boos & Stefanski 2011; Lazzeroni et al. 2014). For example, $P_{\text{Kendall}} = 0.279$ and $P_{\text{Kendall}} = 0.378$ are essentially the same; both indicate no correlation at all. What really matters is the order of magnitude of the $P$ value.

3.1.2. Effect of Observational Biases

Before reaching any conclusion, one should address the issue of observational bias. How do the transit selection effect and detection efficiency affect the observed orbital spacing pattern? Could the observed pattern (Figure 1) be reproduced by the observational bias (Zhu 2020)?

Here, we address this issue by forward modeling the transit detection and selection process with a Monte Carlo method (see the Appendix for the detailed procedure). With this method, we create 1000 corresponding simulated samples of equal size as each observed sample. We then perform the same PR correlation evaluation (Section 3.1.1) to the simulated samples. Figure 2 shows the typical result of each set of simulated samples. As can be seen, all the Pearson test $P$ values for the Monte Carlo realizations are larger than 0.1, and all the Kendall test $P$ values are larger than 0.05, indicating almost no correlation at all.

In Figure 3, we plot the distributions of the $P$ values for the four simulated sample sets, and calculate the fractions of simulations whose $P$ values are not smaller than the observed ones. As can be seen, in most cases (except the Pearson test in sample 3) the fraction numbers are close to 100%, implying a high confidence level that the PR correlations observed in these samples are likely to be physical rather than the results of observational biases. As for the low fraction number (68.6%) for the Pearson test in sample 3, this is because the inclusion of larger PRs largely reduces the PR correlation as mentioned in Figure 1. In the following section, we will investigate how the PR correlation changes with PR itself.

3.2. Evidence of PR Dichotomy

In this subsection, we further perform a moving sample analysis, which reveals that the orbital spacing pattern as a whole is more like a dichotomy rather than a correlation. For the sake of clarity, hereafter, we only present the results of analyzing the nominal sample (sample 1 in Table 1), since other samples generally give similar results.

The procedure of such an analysis is described as follows.

1. First, we sort all the systems in the sample according to the average PR of neighboring planets $\text{PR}$ of each system.
2. Second, we select the first 15 systems as a subsample and perform the Kendall (and Pearson) correlation evaluation (Section 3.1.1) to the subsample, obtaining the $P$ value $P_{\text{Kendall}}$ and $P_{\text{Pearson}}$.
3. Third, we repeat the above correlation evaluation to a series of continuously moving subsamples until the entire sample goes through. Specifically, for each time, we move the subsample one step toward larger PR. For example, we select 15 systems from the second and the sixteenth in the sorted sample next time.

In Figure 4, we plot the result of the above moving sample analysis, which is the $P$ value of the correlation test $P_{\text{Kendall}}$ and $P_{\text{Pearson}}$ as a function of the median of $\text{PR}$ in each moving subsample. As can be seen, the $P$ value ($P_{\text{Kendall}}$ (blue solid curve) and $P_{\text{Pearson}}$ (red solid curve)) increases from $\sim10^{-3}$.
regardless of Table 1 and red shaded region on the top of the plot. For comparison, we also describe the proximity of a PR to \( j + 1 \) MMR,

\[
\Delta = \frac{j}{j + 1} \text{PR} - 1,
\]

where PR is the PR of adjacent planets. Note that \( \Delta \) is calculated with respect to the nearest first-order MMR (making the absolute value of \( \Delta \) a minimum). Figure 6 shows the \( \Delta \) distribution of neighboring planet pairs in the Kepler multiple transit systems. Similar to Lissauer et al. (2011) and Fabrycky et al. (2014), we also see an overabundance just outside the MMR center (i.e., \( \Delta = 0 \)). As the overabundance is mainly within \(|\Delta| = 0.03\), therefore, we set it as the boundary to select those near-MMR PRs.

We plot in Figure 7 an overview of the orbital architecture of the planetary systems in our nominal sample. Each dot denotes a planet or planet candidate, and each line of dots represents a planetary system with its name on the right edge of the figure. The orbital periods of the planets are normalized by the orbital periods of the innermost planets in the same systems. Between each pair of adjacent planets, there is a number indicating the orbital PR. All the systems are sorted bottom-up according to the average PRs, \( \overline{\text{PR}} \). We have an intuitive impression that near-MMR PRs are clustered in compact systems rather than randomly and evenly distributed among all systems. In order to see how the distribution of near-MMR PRs deviates from random distribution, we perform the following statistical test. Note, in the following analysis, we consider only the first-order MMRs for simplicity as we found that the result would be similar if the second-order MMRs were included.

We classify all the planetary systems into three groups according to the number of near-MMR pairs in each system: MMR-poor (zero near-MMR pair), MMR-middle (one or two near-MMR pairs), and MMR-rich (three or more near-MMR pairs) systems. For our nominal sample (sample 1), the numbers of MMR-poor, MMR-middle, and MMR-rich systems are 22, 26, and 8, respectively. We then apply the same classification to those 10,000 randomly simulated systems where near-MMR PRs are randomly distributed. The average

![Figure 4](image-url)
than 0.001 for curves show a similar trend that overabundance. MMR-middle, and MMR-rich systems, respectively. These as in Lissauer et al. systems. We can see the overabundance of planet pairs just outside exact MMRs observed numbers (number of systems of these three groups. We compare the numbers (expectation) are 15.2, 37.6, and 3.2 in MMR-poor, MMR-middle, and MMR-rich systems, respectively. These results are plotted in Figure 8. In the top panel, we count the number of systems of these three groups. We compare the observed numbers (red histogram) with what we would expect (gray histogram) if all near-MMR pairs are randomly distributed. The chi-square test gives a \( \chi^2 = 13.286 \) for the deviation of the observed sample from the expectation, and there are only 7 in 10,000 times of random realizations resulting larger \( \chi^2 \). This gives a \( P \) value of \( 7 \times 10^{-4} \), indicating that the distribution of near-MMR PRs significantly deviates from a random distribution. As can be seen from the bottom panel, with respect to random distribution, the distribution of near-MMR PRs is polarized into the two ends: MMR rich and MMR poor. There is a deficit in MMR-middle class systems.

Note, MMR is loosely defined here, namely, it generally refers to planets pairs whose PRs are close to MMR, regardless of whether they are dynamically in an MMR state with librating resonant angles. Previous studies (Lissauer et al. 2011; Fabrycky et al. 2014) have shown that the global PR distribution deviates somewhat from random distribution in the sense that there is an overabundance of near-MMR ones. Here, we further show that the local PR distribution also deviates from random distribution, namely, those near-MMR PRs are not evenly distributed among individual systems. Some systems are MMR rich, while some are MMR poor, forming an MMR dichotomy (Figure 8).

4.1.2. PR Dichotomy or MMR Dichotomy?

So far, we have revealed two dichotomous features on the orbital spacing, i.e., the PR dichotomy and the MMR dichotomy. In fact, the two dichotomies are largely equivalent to each other. On one hand, MMR dichotomy could be nothing more than a restatement of the PR dichotomy (the small PR correlation) given the fact that MMRs are denser for smaller PRs. On the other hand, the apparently small PR correlation (i.e., PR dichotomy) could also be just a projection of the MMR dichotomy. As shown in Figure 7, most of the first-order and second-order MMRs (except for the 2:1 MMR and 3:1 MMR) have PRs in a relatively small range (PR \( \leq 5 \): 3 \( \sim 1.7 \)). Thus, PRs of an MMR-rich system are more likely to be correlated to each other, while such a correlation is not expected in an MMR-poor system, causing the apparent PR correlation dichotomy. These are clearly shown in Figure 9. As can be seen, the two broken dashed lines generally match the envelopes of the data in Figure 9. The envelopes of MMR-rich systems generally follow the parts that are parallel to the 1:1 line, and thus resulting in strong PR correlation with a \( P \) value of the Kendall correlation test of \( P_{\text{Kendall}} = 0.003. \) In contrast, the envelopes of other systems generally follow the parts that are parallel to the \( x \) and \( y \) axes respectively, resulting in weak PR correlations in MMR-middle (\( P_{\text{Kendall}} = 0.08 \)) and MMR-poor (\( P_{\text{Kendall}} = 0.35 \)) systems. The break points of the dashed lines are at PR = 1.65, which are consistent with the transition zone (PR = 1.5–1.7) as shown in Figure 4.

That being so, then which one is more essential to reflect the orbital spacing pattern? PR dichotomy or MMR dichotomy?
Figure 7. Overview of the orbital architectures of planetary systems in the nominal sample. Each dot denotes a planet or planet candidate and each line of dots represents a planetary system with its name on the right edge of the figure. The orbital periods of the planets are normalized by the orbital periods of the innermost planets in the same systems. Between each pair of adjacent planets, there is a number indicating the orbital period ratio. The red color denotes the proximity to first-order MMRs and the blue to second-order MMRs.
Here, we prefer the MMR dichotomy rather than the PR dichotomy for the following reasons.

First, PR dichotomy or small PR correlation is just a mathematical correlation whose boundary (PR ~ 1.5–1.7; Figure 4) itself needs an additional explanation, while the MMR dichotomy is more physically based and naturally explains the PR correlation boundary (as discussed above and shown in Figure 9).

Second, perhaps more importantly, the MMR dichotomy could be a natural result of planet migration and dynamical evolution. One of the leading models on the formation of close-in super-Earths is the inward migration model, namely planets formed at larger distances (e.g., snowline) from the star followed by inward migration driven by a gas disk (Terquem & Papaloizou 2007; Ida & Lin 2008; Hellary & Nelson 2012; Cossou et al. 2014). At the beginning when the gas disk was present, planets grew and migrated inward to form an MMR chain. Afterwards, when the gas disk dissipated, these MMR chains generally evolved to the following two branches (Izidoro et al. 2017). On one hand, some of the MMR chains could become dynamically unstable, which underwent a phase of giant impact that erased the footprint of MMR. On the other hand, some MMR chains could remain relatively stable. Although most of these MMRs could still be broken afterwards due to various mechanisms, e.g., tides damping (Lithwick & Wu 2012; Batygin & Morbidelli 2013; Delisle & Laskar 2014), planetesimal interaction (Chatterjee & Ford 2015), and etc., many of these effects are gentle and planets are able to stay near-MMR with approximately commensurable PRs. These two branches of dynamical evolution naturally lead to the MMR dichotomy.

As a conclusion of above discussions, we therefore consider the orbital spacing pattern dichotomy shown in Figure 4 as a consequence of the MMR dichotomy (Figure 8).

4.1.3. Effect of Missing Planets

Planets which intrinsically exist between the detected transiting planets could be missed by the transit survey, due to either weak signals (low signal-to-noise ratio; S/N) or nontransiting geometry. In our forward modeling simulations, we found ~2–3% of planets in the simulated transiting multi-planet systems were missed due to low S/N, and ~1%–14% (depending on the intrinsic inclination dispersion, $\sigma_{i5}$) of them were missed because of nontransiting geometry. These missing planets cause the observed PRs larger than the intrinsic ones, which randomizes the PR distribution to some degree. If adopting a typical minimum intrinsic PR of 1.2, this effect can affect PRs larger than 1.2$^\pm$ 1.4. Therefore, one might concern that the observed tendency of weaker PR correlation at larger PRs could be caused by the effect of missing planets. In the follows, we quantify this effect.

First, we investigate a one-population scenario with a toy model, in which PRs are intrinsically correlated (along the diagonal line of Figure 9), and those observed uncorrelated PRs (outliers away from the diagonal line) are caused by the missing planets. Specifically, to generate an intrinsic system, we randomly draw the first PR from the debiased PR distribution (Appendix), then draw other PRs with a random deviation within 10% from the first one. A typical result of the one-population scenario is shown in the right panel of Figure 9. Similar to the upper left panel in Figure 1, but here we divide the nominal sample into three subsamples, MMR poor (blue), middle (green), and rich (red) (see Section 4.1.1). For each subsample, we repeat the Kendall correlation test and print the corresponding $P$ value, $P_{\text{Kendall}}$. As can be seen, the period ratio correlation is significant ($P_{\text{Kendall}} = 0.003$) in MMR-rich systems, but weak in MMR-middle ($P_{\text{Kendall}} = 0.08$) and MMR-poor ($P_{\text{Kendall}} = 0.35$) systems. The two broken dashed lines generally match the envelopes of the data. The break points are at PR = 1.65, which are consistent with the transition zone (PR = 1.5–1.7) in Figure 4 (see Section 4.1.2 for more discussion).
Figure 10. Ratio distributions of period ratios for the one-population scenario (right panel, all planets are generated from a period ratio correlated population, i.e., \( f_{\text{correlated}} = 100\% \)), two-population scenario \( f_{\text{correlated}} = 35\% \) (left panel), and the real sample 1 (middle panel, the same data as Figure 9). In each panel, the black solid line shows perfect correlation, i.e., \( y = x \). The two black dashed lines \( y = \frac{11}{10}x \) and \( y = \frac{2}{1}x \) represent the 10% deviation from the perfect period ratio correlation. The two blue dashed lines \( y = x^2 \) and \( y = \sqrt{x} \) denote the expected locations of outliers caused by missing the intermediate planets. In each panel, we print the fraction \( f_{\text{outliers}} \) of outliers, namely the data points further away from the perfect correlation line, i.e., \( y = x \), than the two black dashed lines. See more details in Section 4.1.3.

Figure 10 with \( \sigma_{i,5} = 1^\circ \) and \( f_{\text{correlated}} = 100\% \) (i.e., 100% systems are PR-correlated). As compared to the result of real sample shown in the middle panel, the one-population scenario fails to reproduce the observation in the following two aspects.

1. It produces too few outliers (points away from the diagonal line further than the two dashed lines; \( 11y = 9x \) and \( 9y = 11x \) in a \( x-y \) plane, see the caption of Figure 10). The outliers fraction is 5.7% versus the observed 32.8% in this case. Although increasing the intrinsic orbital inclination dispersion \( \sigma_{i,5} \) generally increases the numbers of nontransiting planets and thus the fraction of outliers, it is still significantly lower than the observed one even if assuming an unrealistically large \( \sigma_{i,5} = 10^\circ \) (as shown in the bottom right part of Figure 11).

2. Its envelopes (set by the outliers), as expected, follow the blue dashed lines in Figure 10 \( (y = x^2 \text{ and } y = x^{0.5} \text{ in a } x-y \text{ plane}) \), which is significantly different from the observed one (red dashed lines in Figures 9 and 10).

Second, we then further consider a two-population scenario with a toy model, in which only a fraction \( f_{\text{correlated}} < 100\% \) of systems are assumed as PR correlated as in the above one-population scenario. For the other the \( 1-f_{\text{correlated}} \) fraction of systems, the PRs are randomly drawn from the debiased PR distribution but with a lower limit truncated at 1.35 (motivated by the apparent envelopes). As shown in Figure 11, by adding more uncorrelated population systems (i.e., decreasing \( f_{\text{correlated}} \)), the outlier fraction generally increases, and it meets the observed value if \( f_{\text{correlated}} \sim 35\% \) for \( \sigma_{i,5} = 1^\circ \). In the left panel of Figure 10, we plot the ratio distribution of PRs for this specific case. As can be seen, the two-population toy model largely reproduces the result of the real observed sample, especially in terms of both the outlier fraction (31.7% versus 32.8%) and the distribution envelopes.

As a summary of this subsection, we conclude that the effect of missing planets (either low S/N planets or nontransiting planets) alone is too small to reproduce the observed ratio distribution of PRs (Figure 10). In addition, we find that the observed results could be largely reproduced with a two-population toy model, which further demonstrates the dichotomy nature of the orbital spacing pattern.

4.1.4. Effect of Ultra-short-period Planets

Systems with ultra-short-periods (USPs; period < 1 day) are found to have relative larger PRs (Winn & Fabrycky 2015) and larger orbital inclinations (Dai et al. 2018), and they could have undergone some different formation history (Petrovich et al. 2019; Pu & Lai 2019). Thus, one might be concerned whether USP planets are related to the observed trend of weaker PR correlation in systems with larger PRs. However, the occurrence rate of USP planets is in fact very low (<5%) around Sun-like stars (Sanchis-Ojeda et al. 2014). In our
nominal sample, only 2 out of 56 systems host USP planets. After removing these two systems, we repeat the moving sample analysis and find that the result is nearly unchanged as compared to Figure 4. We therefore conclude that our results are not affected by USP planets.

4.2. Predictions

Based on the above discussions on the dynamical origin of the MMR dichotomy, we may further make some predictions for future studies.

First, we predict that the planets in MMR-poor systems (with relatively larger and thus uncorrelated PRs) may have larger masses, densities, and orbital eccentricities/inclinations than those in MMR-rich systems (with relatively smaller and thus correlated PRs). This is simply because the giant impact process which erased the footprint of MMRs also increased the masses and the orbital eccentricities/inclinations of planets. The prediction on mass and density is consistent with the recent finding that the masses and densities of transit-timing variation planets (most are near-MMR) are systematically lower than those of the radial velocity planets (most are not near-MMR; Steffen 2016). The confirmation of the prediction on orbital eccentricity/inclination is not trivial, because the increase in eccentricity/inclination is moderate, which requires future dedicated studies on orbital characterization.

Second, we may predict that MMR-poor systems (with relatively larger and thus uncorrelated PRs) are relatively older than those MMR-rich systems (with relatively smaller and thus correlated PRs). This is simply based on the consideration that the longer time of dynamical evolution (e.g., giant impact, tidal damping, and planet–planetesimal interaction), the larger probability to erase the footprint of MMR. The prediction on age is qualitatively consistent with the result of a previous study (Koriski & Zucker 2011) based on the radial velocity planet sample. Future studies with large and diverse samples are needed to fully establish this point.

5. Summary

In this paper, we studied the pattern of orbital spacings (in terms of PRs) of Kepler multiple-planet systems. We confirm that PRs are indeed somewhat correlated (Figure 1), and such a correlation is unlikely to be caused by observational biases (Figures 2 and 3). Furthermore, we reveal that the above orbital spacing pattern is dichotomous, namely, PRs are strongly correlated to each other in the tightly packed systems, but uncorrelated at all in the loosely packed systems. The transition from correlation to noncorrelation is abrupt with the boundary at Median(PR) ~ 1.5–1.7 (Section 3.2 and Figure 4).

Then, we relate such a PR dichotomy to another dichotomy that reflects the near-MMR PRs trend to be clustered rather than evenly distributed (dubbed as MMR dichotomy for short, see Section 4.1 and Figures 7 and 8). The MMR dichotomy naturally leads to a transition from PR correlation to noncorrelation around PR ~ 1.5–1.7 (Figure 9), and it could be also a natural result of planet migration and dynamical evolution (Section 4.1.2). The transition from PR correlation to noncorrelation cannot be explained by the missing intermediate planets (due to either low S/N or nontransiting geometry, Section 4.1.3) or by USP planets (Section 4.1.4). Nevertheless, it can be largely reproduced with a two-population toy model, further demonstrating the dichotomy nature of the orbital spacing pattern.

Finally, based on the formation of the MMR dichotomy, we predict that planets in MMR-poor systems are more massive, denser, and dynamically hotter (larger orbital eccentricities and inclinations) than those in MMR-rich ones (Section 4.2).

We thank W. Zhu and J.-Y. Yang for helpful comments and suggestions. This work is supported by the National Key RD Program of China (No. 2019YFA0405100) and the National Natural Science Foundation of China (NSFC) (grant No. 11933001). J.-W.X. also acknowledges the support from the National Youth Talent Support Program and the Distinguish Youth Foundation of Jiangsu Scientific Committee (BK20190005).

Appendix

Monte Carlo Simulations of Transit Systems

In order to quantify the probability of reproducing the observed PR correlation by observational bias, we perform Monte Carlo simulations of transit systems by the following forward modeling of transit observations. Specifically, first (Appendix A.1), we generate intrinsic planetary systems based on some reasonable assumptions that are studied and justified by previous studies. Then (Appendix A.2), we apply some criteria to simulate transit detection from the above generated systems. Finally (Appendix A.3), we evaluate the PR correlation as in Section 3.1.1 for the simulated transit sample. By repeating the above simulation and evaluation 1000 times, we access the probability of reproducing the observed PR correlation by observational bias (Figure 3).

A.1. Generating Intrinsic Planet Systems

In the following, we describe the procedure to generate an intrinsic planet system.

1. We randomly select a star from the Kepler Input Catalog, whose stellar properties have been revised by Gaia data (Berger et al. 2018).
2. We assign K planets to the star, where K = 1–6 is drawn from the multiplicity function obtained by Zhu et al. (2018; their Figure 8).
3. We draw the orbital period of the innermost planet randomly from the distribution of orbital periods of innermost transiting planets in the observed sample after correcting the transit geometric bias. To determine the period of other planets in the system, we multiply the period of the inner planet by a PR, which is randomly drawn from a distribution debiased from observation using the CORBITS algorithm (Brakensiek & Ragozzine 2016; Wu et al. 2019). Specifically, we calculate the probability of detecting an outer planet given that the inner planet is detected. The inverse of the probability is adopted as the weight of the PR of the planet pair.
4. The radius of each planet is drawn from a debiased radius distribution that is constructed as in Fang & Margot (2012). Specifically, for a planet with radius R and period P in the observed sample, we calculate η as the fraction of stars that can detect the transit of such a planet. Since η is the ratio of the number of detectable events to the number of actual planets, the inverse of η is an estimate of the actual number of planets represented by each detection.
Therefore, we set $\frac{1}{\varepsilon}$ as the weight of each observed specific radius $R$ to obtain the debiased radial distribution.

5. To avoid the cases where two planets are too close to each other and become dynamically unstable, we also adopt the stability criterion as in Fang & Margot (2012), i.e.,

$$\Delta = \frac{a_2 - a_1}{R_{\text{Hill}}} \geq 3.46$$  \hspace{1cm} (A1)

where $a_1$ and $a_2$ are the semimajor axis of the inner and outer planet respectively and $R_{\text{Hill}}$ is their mutual Hill radius,

$$R_{\text{Hill}} = \left(\frac{M_1 + M_2}{3M_*}\right)^{1/3} \frac{a_2 + a_1}{2},$$  \hspace{1cm} (A2)

with $M_1$ and $M_2$ being the mass of the inner and outer planet and $M_*$ being the mass of the host star. Masses of planets are estimated using a nominal mass–radius relation (Lissauer et al. 2012), i.e.,

$$\frac{M}{M_{\oplus}} = \left(\frac{R}{R_{\oplus}}\right)^{2.06}$$  \hspace{1cm} (A3)

where $M$ and $R$ are the mass and radius of the planet, and $M_{\oplus}$ and $R_{\oplus}$ is the mass and radius of Earth, respectively.

6. For each system that passed the orbital stability check, we assign $I_p$, the orbital inclination relative to the observer to the planets. Following Zhu et al. (2018), in practice, we calculate

$$\cos I_p = \cos I \cos i - \sin I \sin i \cos \phi,$$  \hspace{1cm} (A4)

where $I$ is the inclination of the system invariant plane, $i$ the planet inclination with respect to this invariant planet, and $\phi$ is the phase angle. The distribution of $I$ is isotropic (i.e., $\cos I$ is uniform for $0^\circ < I < 180^\circ$) and $\phi$ is a random between $0^\circ$ and $360^\circ$. For single planet systems, $i = 0^\circ$ and $I_p = I$. For multiple-planet systems, following Zhu et al. (2018), $i$ is modeled as a Fisher distribution,

$$P(i|\kappa) = \frac{\kappa \sin i}{2 \sinh \kappa} e^{\kappa \sin i \cos \phi}.$$  \hspace{1cm} (A5)

The $\kappa$ parameter is related to the inclination dispersion as

$$\sigma_{i,k}^2 = \langle \sin^2 i \rangle = \frac{\kappa}{\kappa_k} \left(\coth \kappa - \frac{1}{\kappa_k}\right).$$  \hspace{1cm} (A6)

Here, also following Zhu et al. (2018), the inclination dispersion is a power-law function of the planet multiplicity, $k$,

$$\sigma_{i,k} = \sqrt{\langle \sin^2 i \rangle} = \sigma_{i,5} \left(\frac{k}{5}\right)^{-\alpha}.$$  \hspace{1cm} (A7)

Here, we adopt the typical results from Zhu et al. (2018), i.e., $\sigma_{i,5} = 0.58$ and $\alpha = -4$.

A.2. Simulating Transit Observation

We first consider the transit geometric effect. A transit is defined as the impact parameter less than 1, i.e., $|\cos (I_p)|/e < 1$, where $e = R/a$ is the transit parameter. As in Zhu et al. (2018), we ignore the minor impact of the planet size and eccentricity.

We then consider the effect of detection efficiency. Specifically, we remove the nondetectable transiting planets with transit $S/N$ lower than 7.1 according to Mullally et al. (2015). Following Narang et al. (2018), the transit $S/N$ is calculated as

$$S/N = \left(\frac{R}{R_*}\right)^2 \frac{\sqrt{N}}{\sigma_{\text{CDPP}}}$$  \hspace{1cm} (A8)

where $R$ and $R_*$ are the radii of planet and star, respectively, and $N$ is the effective transiting times. $\sigma_{\text{CDPP}}$ represents the combined differential photometric precision of the star.

A.3. Evaluation of PR Correlation

We repeat above procedure until obtaining the same number of simulated transit systems after the same filters as the observed ones (Table 1 in Section 2). For the four simulated samples, we perform the same PR correlation evaluation as for the observed ones (Section 3.1). The typical results are illustrated in Figure 2. As can be seen, the $P$ value, $P_{\text{Kendall}}$ and $P_{\text{Pearson}}$, of all the four simulated samples are of the magnitude of $10^{-5}$, which are consistent with no PR correlation. The $P_{\text{Kendall}}$($P_{\text{Pearson}}$) distributions of 1000 Monte Carlo realizations are plotted in Figure 3. For samples 1, 2, and 4, the simulations lead to $P_{\text{Kendall}}$ ($P_{\text{Pearson}}$) larger than that of the observed one in most cases. Therefore, the PR correlations observed in the samples 1, 2, and 4 are likely to be physical rather than the results of observational biases.

Note, although our model is relatively simple and suffers some uncertainties, for example, the intrinsic multiplicity is actually not well constrained and the transit detection efficiency is considered as a simple $S/N$ cut, it catches the bases of transit simulation. A more sophisticated state-of-the-art model may improve the estimate the planet occurrence rates, but it is unlikely to change the conclusion, namely, the process that generates transit systems cannot produce significant PR correlation.

ORCID iDs

Ji-Wei Xie @ https://orcid.org/0000-0002-6472-5348
Ji-Lin Zhou @ https://orcid.org/0000-0003-1680-2940

References

Batygin, K., & Morbidelli, A. 2013, AJ, 145, 1
Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, ApJ, 866, 99
Boos, D. D., & Stefanski, L. A. 2011, The American Statistician, 65, 213
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Sci, 327, 977
Bovaird, T., & Lineweaver, C. H. 2013, MNRAS, 435, 1126
Brakensiek, J., & Ragozzine, D. 2016, ApJ, 821, 47
Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, Sci, 337, 556
Chatterjee, S., & Ford, E. B. 2020, AJ, 159, 281
Cossou, C., Raymond, S. N., Hersant, F., & Pierens, A. 2014, A&A, 569, A56
Dai, F., Masuda, K., & Wunn, J. N. 2018, ApJL, 864, L38
Delisle, J. B., & Laskar, J. 2014, A&A, 570, L7
Dong, S., & Zhu, Z. 2013, ApJ, 778, 53
Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, ApJ, 790, 146
Fang, J., & Margot, J.-L. 2012, ApJ, 761, 92
Gilbert, G. J., & Fabrycky, D. C. 2020, AJ, 159, 281
Hadden, S., & Lithwick, Y. 2014, ApJ, 787, 80
He, M. Y., Ford, E. B., & Ragozzine, D. 2019, MNRAS, 490, 4575
Hellary, P., & Nelson, R. P. 2012, MNRAS, 419, 2737
Hebb, A. W. 2013, Sci, 340, 572
Huang, C. X., & Bakos, G. A. 2014, MNRAS, 442, 674
Iida, S., & Lin, D. N. C. 2008, ApJ, 673, 487
Izidoro, A., Ogihara, M., Raymond, S. N., et al. 2017, MNRAS, 470, 1750
