Astronomy
&
Astrophysics

Comparing HARPS and *Kepler* surveys

The alignment of multiple-planet systems

P. Figueira[1], M. Marmier[2], G. Boué[1], C. Lovis[2], N. C. Santos[1,3], M. Montalto[1], S. Udry[2], F. Pepe[2], and M. Mayor[2]

[1] Centro de Astrofísica, Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
  e-mail: `pedro.figueira@astro.up.pt`
[2] Observatoire Astronomique de l'Université de Genève, 51 Ch. des Maillettes, Sauverny 1290, Versoix, Suisse
[3] Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Portugal

ABSTRACT

Context. The recent results of the HARPS and *Kepler* surveys provided us with a bounty of extrasolar systems. While the two teams extensively analyzed each of their data-sets, little work has been done comparing the two.
Aims. We study a subset of the planetary population whose characterization is simultaneously within reach of both instruments. We compare the statistical properties of planets in systems with $m \sin i > 5-10 \, M_\oplus$ and $R > 2 \, R_\oplus$, as inferred from the HARPS and *Kepler* surveys, respectively. If we assume that the underlying population has the same characteristics, the different detection sensitivity to the orbital inclination relative to the line of sight allows us to probe the planets' mutual inclination.
Methods. We considered the frequency of systems with one, two, and three planets as dictated by HARPS data. We used *Kepler*'s planetary period and host mass and radius distributions (corrected from detection bias) to model planetary systems in a simple, yet physically plausible way. We then varied the mutual inclination between planets in a system according to different prescriptions (completely aligned, Rayleigh distributions, and isotropic) and compared the transit frequencies with one, two, or three planets with those measured by *Kepler*.
Results. The results show that the two datasets are compatible, a remarkable result especially because there are no tunable knobs other than the assumed inclination distribution. For $m \sin i$ cutoffs of 7–10 $M_\oplus$, which are those expected to correspond to the radius cutoff of 2 $R_\oplus$, we conclude that the results are better described by a Rayleigh distribution with a mode of 1° or smaller. We show that the best-fit scenario only becomes a Rayleigh distribution with a mode of 5° if we assume a quite extreme mass-radius relationship for the planetary population.
Conclusions. These results have important consequences for our understanding of the role of several proposed formation and evolution mechanisms. They confirm that planets are likely to have been formed in a disk and show that most planetary systems evolve quietly without strong angular momentum exchanges such as those produced by Kozai mechanism or planet scattering.

Key words. planetary systems – techniques: radial velocities – techniques: photometric – surveys – methods: numerical – methods: statistical

1. Introduction

We live in very exciting times for extrasolar planet science. Since the first discovery of an extrasolar planet by Mayor & Queloz (1995), and thanks to the high efficiency of detection mechanisms, the planetary detection rate has been increasing rapidly. As of today, more than 16 years past, we count ~700 detected planets and many candidates to confirm. We have now both the data and the tools to explore the statistical properties of the planetary underlying population (e.g. Udry & Santos 2007).

From the technical perspective, during the last couple of years we have witnessed a spectacular increase in the precision (and consequently sensitivity) of the most efficient detection mechanisms. Using the radial velocity (RV) technique, the most efficient to date, the HARPS spectrograph detected 153 planets around G, K, and M stars, amongst which the lightest planets ever found, with a mass of less than 2 $M_\oplus$ (Mayor et al. 2009; Lovis et al. 2011) The only spectrograph with demonstrated sub-m/s precision, HARPS allowed a detailed study of the underlying planetary population, yielding that at least 50% of the stars of the solar neighborhood have a planet orbiting around them with a period shorter than 100 days (Mayor et al. 2011).

As the *Kepler* mission became operational and the data were reduced, it became clear that the mission would set a new standard on photometric precision (Borucki et al. 2010). The candidate extrasolar planets revealed through the transit technique outnumbered those that had been found up to now, from ground or space. Out of the ~1250 candidates found (Borucki et al. 2011)[1], one can highlight the large number of extrasolar planet systems and the first secure confirmation of planets through transit-timing variations (Lissauer et al. 2011a).

As information from the two surveys becomes available, a comparison between the results is warranted, and some authors started to tackle this problem (e.g. Wolfgang & Laughlin 2011). However, this task is complicated by several factors. The most obvious limitation is that these surveys rely on different techniques that probe different parameter's space. And, naturally, these techniques lead to different detection biases, providing two

---

[1] Batalha et al. (2012), submitted during the refereeing process of this paper, presented an updated planet count. However, the lack of information on the photometric limits as discussed in first by Howard et al. (2011) and then by Youdin (2011) for the Borucki et al. (2011) release prevents us from repeating the analysis for this new dataset.

incomplete snapshots of the underlying population which only overlap partially, giving us a fragmented picture of the mass-radius diagram.

Particularly interesting is the dependence of both methods on orbital inclination relative to our line-of-sight. Planets transit only when their orbital movement makes them cover the disk of the star, as seen from our vantage point; this only happens close to our line-of-sight, restraining detections to small angles relative to it. On the other hand, RV techniques probe a much wider range of inclinations: in theory one can detect a system as long as its plane does not coincide with the plane of the sky; in practice the amplitude of the signal depends in a complex way on several orbital parameters, and its detectability is an even more complicated matter.

This has a particularly interesting consequence for planetary systems: while transit technique detects planetary systems that are not only aligned with our line of sight but close to coplanar, RV detects systems with a much higher inclination and consequently allows the detection of systems with a high mutual inclination between planets. In this work we propose a first comparison between planets from HARPS and *Kepler* surveys as a way of studying the mutual inclination between planets. We will restrict our analysis to planets with masses and radii within reach of the detection limits of both surveys to allow a simpler and more meaningful comparison.

In Sect. 2 we present an overview of the HARPS results of interest for our study and in Sect. 3 we describe our selection and analysis of the *Kepler* candidates. In Sect. 4 we describe the methodology that allows us to compare the planets from both surveys and present the results of this comparison. In Sect. 5 we discuss these results and conclude in Sect. 6.

2. The HARPS survey results

The HARPS spectrograph (Mayor et al. 2003) has set a new benchmark in high-precision RV measurements, delivering several high-visibility results. Recently, Mayor et al. (2011) conducted a detailed statistical analysis of the volume-limited survey of HARPS+CORALIE and calculated the frequency of exoplanets with periods of up to 100 days, by calculating the detection limits for each star and correcting the measured planetary frequency from the detection bias. This analysis was performed on a sample of 822 FGK stars, 376 of which are followed at a level allowing the detection of small-mass planets.

Of interest to us is a subset of the survey, for which comparison with *Kepler* is possible. We restrain our analysis to exoplanets defined by a mass cutoff whose RV signal makes the detection possible up to the longest period covered by *Kepler* analysis. This cutoff should be taken with care, since one cannot define a priori how large K, the semi-major amplitude of the signal, has to be relative to the instrumental precision for the planet to be detected in all cases. Note that this ratio is non-constant and is the reason why to evaluate the presence of planets on datasets one must resort to time-intensive Monte-Carlo analysis as those presented in Mayor et al. (2011).

To provide some latitude to our analysis and test the results for different mass-radius relationships, we repeated the analysis of Mayor et al. (2011) for different cutoffs: 5, 7, 9, and 10 $M_\oplus$ and for the periods between 0.68 and 50 days, the range of periods explored by *Kepler*. It is important to note that the detection probability of a 5 $M_\oplus$ planet in a 50 day orbit is already of 10–20% (Fig. 7 of the paper); as a consequence the error bars on the corrected detection rates are quite large, making the results less constraining and less insightful. These mass-cutoff values will

turn out to be a very convenient choice for a different reason, as discussed in the next section.

The results for the true frequency of single planets and systems with two, three, and four planets are presented in Table 1. Note that these results are not corrected for sin i selection effects. This selection effect will bias the detections preferentially toward planets whose system's plane is closer to the line of sight. This might lead to an underestimation of the true frequency of planets, but in a way that is not a function of the number of planets in a system, but rather depends on the detectability of each planet's signal.

3. The *Kepler* mission results

The *Kepler* photometric mission measured high-precision photometric variations of stellar flux on 160 000 stars (down to ~10 ppm) and allowed (up to now) the detection of 1235 planet transit candidates (Borucki et al. 2011). Howard et al. (2011) extensively analyzed the published datasets and established the properties of the planetary population probed by the mission. Restricting the study to the subset of stars with high planet detectability, the authors quantified the planetary occurrence (as they named the true frequency of planets per star) as a function of planetary radius and period, from 0.68 to 50 days.

Of particular interest to us are planets with radii large enough for the planetary census to be considered complete and devoid of systematics, the threshold of which was defined as 2 $R_\oplus$ by the authors. Youdin (2011) repeated the analysis with a different methodology and, among many interesting results, confirmed this claim of completeness for at least $R > 3\,R_\oplus$. Setting 2 $R_\oplus$ as the lower limit for data analysis facilitates the comparison with HARPS data, because according to Sotin et al. (2007) a planet with 10 $M_\oplus$ has a radius between 1.74 and 2.37 $R_\oplus$.

Howard et al. (2011) presented the planet occurrence as the frequency of planets corrected from both geometrical probability and from the insufficient photometric precision. We repeated the same analysis, considering the same thresholds for detection but now separating planetary candidates into single planets, double transiting systems, triple transiting systems, and quadruple transiting systems. All planetary and stellar parameters were extracted from the *Released Planet Kepler Candidates* list[2]. We considered the same candidates by imposing the same thresholds on *Kepler* star magnitude, T_{eff} and log g, and setting the same signal-to-noise ratio threshold on transit detection. When a system consists of more than one planet we considered the photometric precision to be dictated by the most stringent of the two planets, in line with Howard et al. (2011); to do so we considered the number of stars that still allow us to reach the required precision to detect both planets. We corrected our planetary frequencies for the insufficient photometric precision but not for the transit probability. We did this because it is our objective to reproduce through simulations the frequency of planets that transit; these simulations include in a natural way the geometrical probability impact but cannot reproduce the effect of deficient photometric precision.

The frequency of single-transit planets, double transiting systems, triple transiting systems, and quadruple transiting systems is presented in Table 2.

---

[2] http://archive.stsci.edu/kepler/planet_candidates.html

Table 1. True frequency of single planets and systems with planetary periods between 0.68 and 50 days and $m \sin i$ larger than 5, 7, 9, and 10 $M_{\oplus}$, as calculated from HARPS data.

$m. \sin(i)$	Single planets	Syst. w/ 2. plan.	Syst. w/ 3. plan.	Syst. w/ 4. plan.
>5 $M_{\oplus}$	11.15 ± 3.69%	4.63 ± 2.27%	0.68 ± 0.68%	n.d.
>7 $M_{\oplus}$	10.51 ± 2.60%	2.19 ± 1.26%	n.d.	n.d.
>9 $M_{\oplus}$	9.82 ± 2.22%	1.21 ± 0.85%	n.d.	n.d.
>10 $M_{\oplus}$	8.75 ± 1.92%	0.57 ± 0.57%	n.d.	n.d.

Notes. n.d. stands for "not detected"; see Sect. 2 for more details.

Table 2. Transit frequency for planets and systems of up to four planets with period between 0.68 and 50 days, as calculated from the *Kepler* candidates using the methodology from Howard et al. (2011).

Radii	Single transits	Double transits	Triple transits.	Quadruple transits
>2 $R_{\oplus}$	5.23e-1%	5.97e-2%	9.26e-3%	5.51e-3%

Notes. n.d. stands for "not detected"; see Sect. 3 for more details.

4. Comparing the surveys

4.1. Methodology

Comparing the results from the two methods is not straightforward. It is particularly difficult to match a cutoff in radius with one in mass. Here, by using different cutoff values, we considered different mass-radius relationships. This relation is not only unknown, but hardly unique, owing to the different possible composition of exoplanets for the mass range considered (e.g. Valencia et al. 2007).

Starting form the observed HARPS frequencies stated above, we simulated systems whose inclination between the planets and the system's (fixed) plane was controlled by tunable knobs. It is important to note that we considered planet frequencies as fixed (here used in plural because one must distinguish the frequency of stars with 1, 2, or 3 planets[3]). These were dictated by HARPS results and only depend on the mass cutoff (i.e. the mass-radius relationship assumed). The methodology is depicted in the flowchart presented in Fig. 1 and can be summarized in the following way:

- we created 10^8 planetary systems; the number of planets in each system was dictated by the frequencies of stars with one, two, or three planets set by HARPS results.
- for each system the stellar radius was drawn from *Kepler*'s radius distribution of host stars. Each planet period P_k was drawn from the measured *Kepler* period distribution of planetary candidates.
  The observed stellar radius distribution and planetary period distribution were corrected both from photometric detectability bias and from geometry bias. The geometry bias is corrected simply by applying the formula of the transit probability (e.g. Youdin 2011)

$$P_{\text{transit}} = 0.051 \left(\frac{10 \, \text{days}}{P} \right)^{2/3} \left(\frac{\rho_{\odot}}{\rho_*} \right)^{1/3}, \qquad (1)$$

in which ρ is the mean stellar density and is valid for the cases of $R_p \ll R_*$ and e = 0 [4]. The photometric detectability

bias, as corrected by Howard et al. (2011), reduces the accuracy of our measurements because it is presented for binned data in the ($\log P$, $\log R_p$) space. To correct for this last effect we used instead the power-law formulas (4), (5), and (6) of Youdin (2011). The interested reader is referred to this work for the details on the advantages of the approach. We corrected for these two effects and only then binned the data into 20 equally sized bins and assigned a probability proportional to the (corrected) frequency of planets inside the bin; these distributions are presented in Fig. 2.

- For each drawn stellar radius, we drew a stellar mass with a flat probability between the extreme masses detected for the radii included in the bin. This allows us to translate orbital periods P_k into orbital semi-major axis a_k using the generalized Kepler third law:

$$a_i = \sqrt[3]{\frac{M_*}{M_{\odot}} \frac{P}{1 \, \text{yr}}^2} \, [\text{AU}]. \qquad (2)$$

- If more than one planet is drawn for the system in question, the orbits have to respect the Hill stability criterion, as defined by Eq. (24) of Gladman (1993),

$$\frac{a_2 - a_1}{a_1} > 2.40 \, (\mu_1 + \mu_2)^{1/3}, \qquad (3)$$

in which $\mu_1 = m_1/m_*$, $\mu_2 = m_2/m_*$. This equation is valid for circular orbits and when $m_* \gg m_1, m_2$. The planetary periods are redrawn until all pairs of planets respect this property. We chose for m_1 and m_2 the $m \sin i$ cutoff used for the considered HARPS frequency determination.
- The orientation of the system's plane relative to the line-of-sight plane, I, is drawn from a linear distribution in *sin* (note that angles closer to the line of sight have a higher probability of being detected).
- For each planet the angle i_k with respect to the reference plane of the system is drawn from the assumed distribution (see next subsection), and the longitude of the ascending node Ω_k is drawn from a linear distribution between 0 and 360°.
- Finally, each angle θ_k relative to the line of sight, as seen from our vantage point, is given by

$$\theta_k = \text{asin} \, | \cos I \cos \Omega \sin i + \sin I \cos i |, \qquad (4)$$
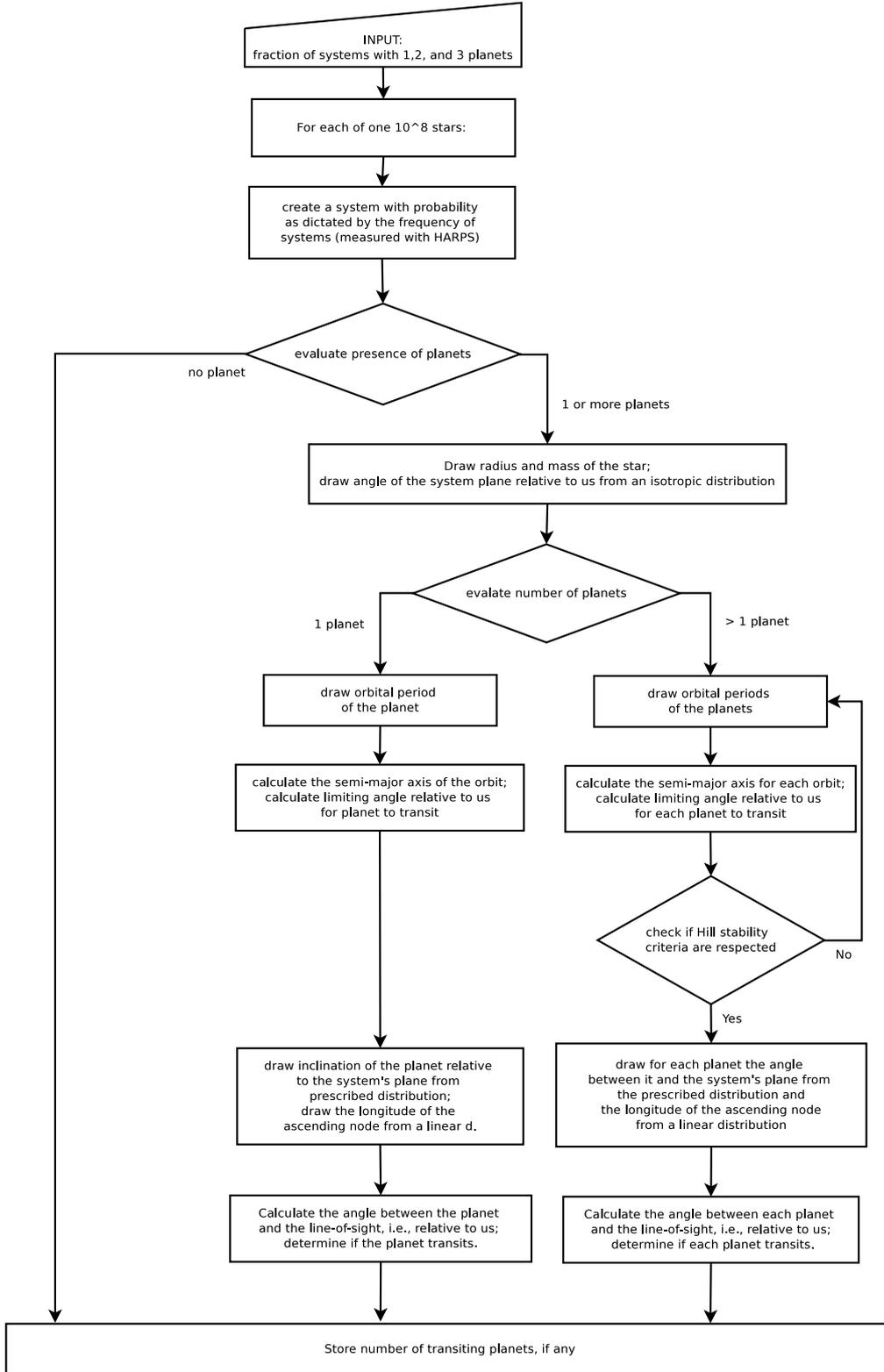
---

[3] In principle, there is no reason not to consider the frequency of stars with more planets around them; in practice these systems were not present in the sample.

[4] We considered that planetary orbits are circular, a hypothesis strengthened by the work of Moorhead et al. (2011), who showed that *Kepler* transit durations imply a mean eccentricity ≤0.2.

Fig. 1. Flowchart for the simulations described in Sect. 4.1.

and the planet is considered to transit if $\theta_k < \theta_{\lim} = \mathrm{asin}(R_*/a_k)$. The different planes and angles mentioned are depicted in Fig. 3.

It is important to note that we followed a methodology similar to that of Lissauer et al. (2011b), in which one starts from different primordial populations and tries to recover the frequency of observed *Kepler* transit candidates. However, in our case, and on top of several different minor aspects, we had a previously determined frequency of planets in systems for single planets, double-planet systems, and triple-planet systems. To estimate the error bars on our transit frequencies we used the 1-σ uncertainties in HARPS frequencies to calculate extreme-case scenarios for the transit frequencies.
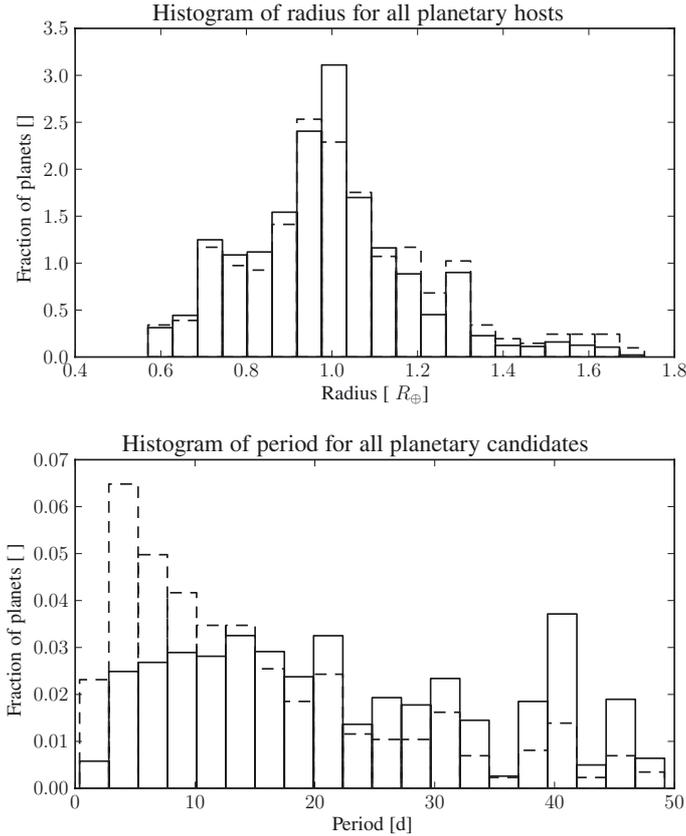
Fig. 2. Distribution of radii for all *Kepler* planetary host (*top*) and period for all *Kepler* planetary candidates considered for the $2\,R_\oplus$ radii cutoff (*bottom*); dashed histograms depict the distributions prior to the detection bias correction and solid line after.

4.2. Mutual inclination distributions and results

For the distribution of the angles of each planet relative to the plane of the system *i* we considered 1) planets aligned with their systems plane; 2) different Rayleigh distributions $R(\sigma)$ relative to the system's plane, where σ, the mode of the distribution, is the only governing parameter. For the latter we considered $\sigma \in [1, 5, 10, 20, 30, 45]°$; and 3) an isotropic distribution (i.e. a linear distribution in $\cos i$). The results are presented in Table 3.

The obtained frequencies for single and double planetary transits are plotted and compared to those recovered from *Kepler* data in Fig. 4. In this figure we plot in a two-dimensional space the transit frequencies (and associated uncertainties) for single-transiting planets and double-transiting planets, as calculated from our simulations, for the different assumed inclinations (different panels) and different $m\sin i$ cutoff (different markers and colors). In the same plot we present the values measured from *Kepler* data, so that both can be directly compared.

Note that for most cutoffs considered HARPS did not detect triple or quadruple systems for the given mass and period range. This was due to 1) the low number of stars surveyed when compared with *Kepler*'s, for which the detection of a system with frequency of 10^{-3} is quite low, and 2) to the short period range considered in the analysis, of 50 days. As a matter of fact, many of the systems detected by HARPS extend to more than 50 days, up to 100 days, and were naturally not included in the analysis.
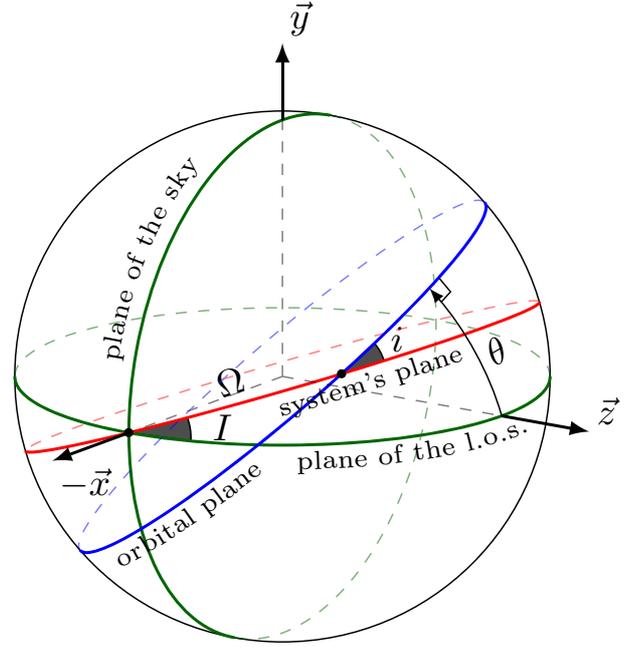

Fig. 3. Angles used in this paper. The observer is in the *z* direction.

5. Discussion

5.1. The likeliness of different i distributions

The first point to note is that the simulations based on HARPS frequencies agree remarkably well with the *Kepler* measurements. This is worth noting, especially because our simulation procedure has no tunable knobs; no effort was made to match the calculated with the observed data. The only free parameter is the inclination distribution, as described.

From Fig. 4 one can conclude that while all considered distributions reproduce the fraction of stars with one transiting planet inside or close to 1-σ error bars, most underestimate the frequency of stars with two transiting systems. It is interesting to quantify the difference between the observed quantities and those reproduced through simulations. To do so we calculated for each case the absolute deviation between the observed and calculated values, taking as units the 1-σ uncertainties (our proxy for the transit frequencies' σ). The results are presented in Table 4, and in Fig. 5 we plot the quadratic sum of the two absolute deviations $\sigma_{\text{tot}} = \sqrt{\sigma_{f_1}^2 + \sigma_{f_2}^2}$, as a function of the considered *i* distribution for the different $m\sin i$ cutoffs.

An inspection of the Table 4 and Fig. 5 shows that, in general, the more misaligned the planets (the larger the mode of the *i* distribution considered), the stronger the deviation relative to the measured *Kepler* values. For cutoffs of 9 and $10\,M_\oplus$ this is exactly the case, but for cutoffs of 5 and $7\,M_\oplus$ the most probable distributions are not the aligned one but Rayleigh with modes of 1 and 5°, respectively. This brings back the question of the impact of the $m\sin i$ cutoff on our results. We saw in Sect. 3 that a $2\,R_\oplus$ radius corresponds to a mass cutoff of around $10\,M_\oplus$, of which precise value depends on the composition of the planets. We also know that by using $m\sin i$ as a proxy for the real mass, we are underestimating the mass by a factor of 0.73, the average value of the $\sin i$ for an isotropic distribution. Accordingly, for the $m\sin i$ cutoffs of 5, 7, 9, and 10, we are considering (on average) mass cutoff values of 6.3, 8.9, 11.4, and $12.7\,M_\oplus$, respectively. At face value, this means that the results for $m\sin i$ of

Table 3. Frequency of simulated transiting systems with 1, 2, and 3 planets, for the different planetary frequencies of Table 2.

$m \sin i$ cutoff	distribution	Single planets [%]	Syst. w/ 2. plan. [%]	Syst. w/ 3. plan. [%]
5 $M_\oplus$	aligned	5.97e-01$^{+2.27e-01}_{-2.27e-01}$	1.51e-01$^{+7.80e-02}_{-7.79e-02}$	1.93e-02$^{+1.88e-02}_{-1.88e-02}$
5 $M_\oplus$	$R(1.0°)$	6.39e-01$^{+2.52e-01}_{-2.50e-01}$	1.38e-01$^{+7.30e-02}_{-7.47e-02}$	1.42e-02$^{+1.41e-02}_{-1.37e-02}$
5 $M_\oplus$	$R(5.0°)$	8.18e-01$^{+3.52e-01}_{-3.53e-01}$	6.60e-02$^{+4.10e-02}_{-4.02e-02}$	2.31e-03$^{+2.40e-03}_{-1.81e-03}$
5 $M_\oplus$	$R(10.0°)$	8.77e-01$^{+3.93e-01}_{-3.90e-01}$	3.81e-02$^{+2.37e-02}_{-2.40e-02}$	7.17e-04$^{+6.53e-04}_{-2.17e-04}$
5 $M_\oplus$	$R(20.0°)$	9.14e-01$^{+4.16e-01}_{-4.12e-01}$	2.12e-02$^{+1.33e-02}_{-1.35e-02}$	2.05e-04$^{+2.04e-04}_{-2.05e-04*}$
5 $M_\oplus$	$R(30.0°)$	9.25e-01$^{+4.25e-01}_{-4.20e-01}$	1.56e-02$^{+9.80e-03}_{-1.01e-02}$	1.07e-04$^{+1.11e-04}_{-1.07e-04*}$
5 $M_\oplus$	$R(45.0°)$	9.29e-01$^{+4.21e-01}_{-4.23e-01}$	1.30e-02$^{+8.10e-03}_{-8.54e-03}$	6.30e-05$^{+5.70e-05}_{-6.30e-05*}$
5 $M_\oplus$	isotropic	9.30e-01$^{+4.30e-01}_{-4.23e-01}$	1.27e-02$^{+8.30e-03}_{-8.18e-03}$	6.90e-05$^{+4.40e-05}_{-6.90e-05*}$
7 $M_\oplus$	aligned	4.94e-01$^{+1.39e-01}_{-1.40e-01}$	6.75e-02$^{+4.05e-02}_{-3.86e-02}$	n. d.
7 $M_\oplus$	$R(1.0°)$	5.13e-01$^{+1.49e-01}_{-1.51e-01}$	5.88e-02$^{+3.39e-02}_{-3.38e-02}$	n. d.
7 $M_\oplus$	$R(5.0°)$	5.82e-01$^{+1.91e-01}_{-1.91e-01}$	2.38e-02$^{+1.38e-02}_{-1.37e-02}$	n. d.
7 $M_\oplus$	$R(10.0°)$	6.03e-01$^{+2.04e-01}_{-2.02e-01}$	1.31e-02$^{+7.40e-03}_{-7.64e-03}$	n. d.
7 $M_\oplus$	$R(20.0°)$	6.15e-01$^{+2.11e-01}_{-2.09e-01}$	6.94e-03$^{+3.96e-03}_{-3.92e-03}$	n. d.
7 $M_\oplus$	$R(30.0°)$	6.20e-01$^{+2.13e-01}_{-2.13e-01}$	5.15e-03$^{+2.83e-03}_{-3.00e-03}$	n. d.
7 $M_\oplus$	$R(45.0°)$	6.22e-01$^{+2.15e-01}_{-2.13e-01}$	4.28e-03$^{+2.35e-03}_{-2.45e-03}$	n. d.
7 $M_\oplus$	isotropic	6.22e-01$^{+2.12e-01}_{-2.14e-01}$	4.09e-03$^{+2.50e-03}_{-2.32e-03}$	n. d.
9 $M_\oplus$	aligned	4.41e-01$^{+1.14e-01}_{-1.14e-01}$	3.77e-02$^{+2.66e-02}_{-2.66e-02}$	n. d.
9 $M_\oplus$	$R(1.0°)$	4.52e-01$^{+1.20e-01}_{-1.21e-01}$	3.26e-02$^{+2.28e-02}_{-2.29e-02}$	n. d.
9 $M_\oplus$	$R(5.0°)$	4.91e-01$^{+1.46e-01}_{-1.50e-01}$	1.34e-02$^{+8.90e-03}_{-9.50e-03}$	n. d.
9 $M_\oplus$	$R(10.0°)$	5.02e-01$^{+1.56e-01}_{-1.57e-01}$	7.31e-03$^{+5.09e-03}_{-5.12e-03}$	n. d.
9 $M_\oplus$	$R(20.0°)$	5.09e-01$^{+1.60e-01}_{-1.62e-01}$	3.87e-03$^{+2.79e-03}_{-2.72e-03}$	n. d.
9 $M_\oplus$	$R(30.0°)$	5.10e-01$^{+1.63e-01}_{-1.61e-01}$	2.83e-03$^{+1.90e-03}_{-2.00e-03}$	n. d.
9 $M_\oplus$	$R(45.0°)$	5.13e-01$^{+1.62e-01}_{-1.65e-01}$	2.37e-03$^{+1.61e-03}_{-1.68e-03}$	n. d.
9 $M_\oplus$	isotropic	5.11e-01$^{+1.63e-01}_{-1.62e-01}$	2.34e-03$^{+1.70e-03}_{-1.69e-03}$	n. d.
10 $M_\oplus$	aligned	3.81e-01$^{+9.60e-02}_{-9.40e-02}$	1.76e-02$^{+1.79e-02}_{-1.71e-02}$	n. d.
10 $M_\oplus$	$R(1.0°)$	3.86e-01$^{+9.90e-02}_{-1.01e-01}$	1.55e-02$^{+1.55e-02}_{-1.50e-02}$	n. d.
10 $M_\oplus$	$R(5.0°)$	4.03e-01$^{+1.19e-01}_{-1.16e-01}$	6.04e-03$^{+6.26e-03}_{-5.54e-03}$	n. d.
10 $M_\oplus$	$R(10.0°)$	4.08e-01$^{+1.24e-01}_{-1.22e-01}$	3.43e-03$^{+3.50e-03}_{-2.93e-03}$	n. d.
10 $M_\oplus$	$R(20.0°)$	4.13e-01$^{+1.26e-01}_{-1.27e-01}$	1.85e-03$^{+1.87e-03}_{-1.35e-03}$	n. d.
10 $M_\oplus$	$R(30.0°)$	4.13e-01$^{+1.28e-01}_{-1.27e-01}$	1.36e-03$^{+1.25e-03}_{-8.60e-04}$	n. d.
10 $M_\oplus$	$R(45.0°)$	4.14e-01$^{+1.26e-01}_{-1.27e-01}$	1.12e-03$^{+1.14e-03}_{-6.20e-04}$	n. d.
10 $M_\oplus$	isotropic	4.14e-01$^{+1.29e-01}_{-1.27e-01}$	1.13e-03$^{+1.07e-03}_{-6.30e-04}$	n. d.

Notes. The error bars were drawn from the lower and upper limit of HARPS frequencies uncertainties. n.d. stands for "not detected"; see Sect. 4 for more details. Note that for very low frequency values (marked with a *), lower than 5e-04%, the number of events is so low that the error bars are ill defined through Monte-Carlo analysis and cannot be considered meaningful.

Table 4. Absolute deviation between the *Kepler* transit frequencies for single and double transit systems for the 2 $R_\oplus$ cutoff and those obtained through simulation in this paper (presented in Table 3) for the different assumed inclination distributions.

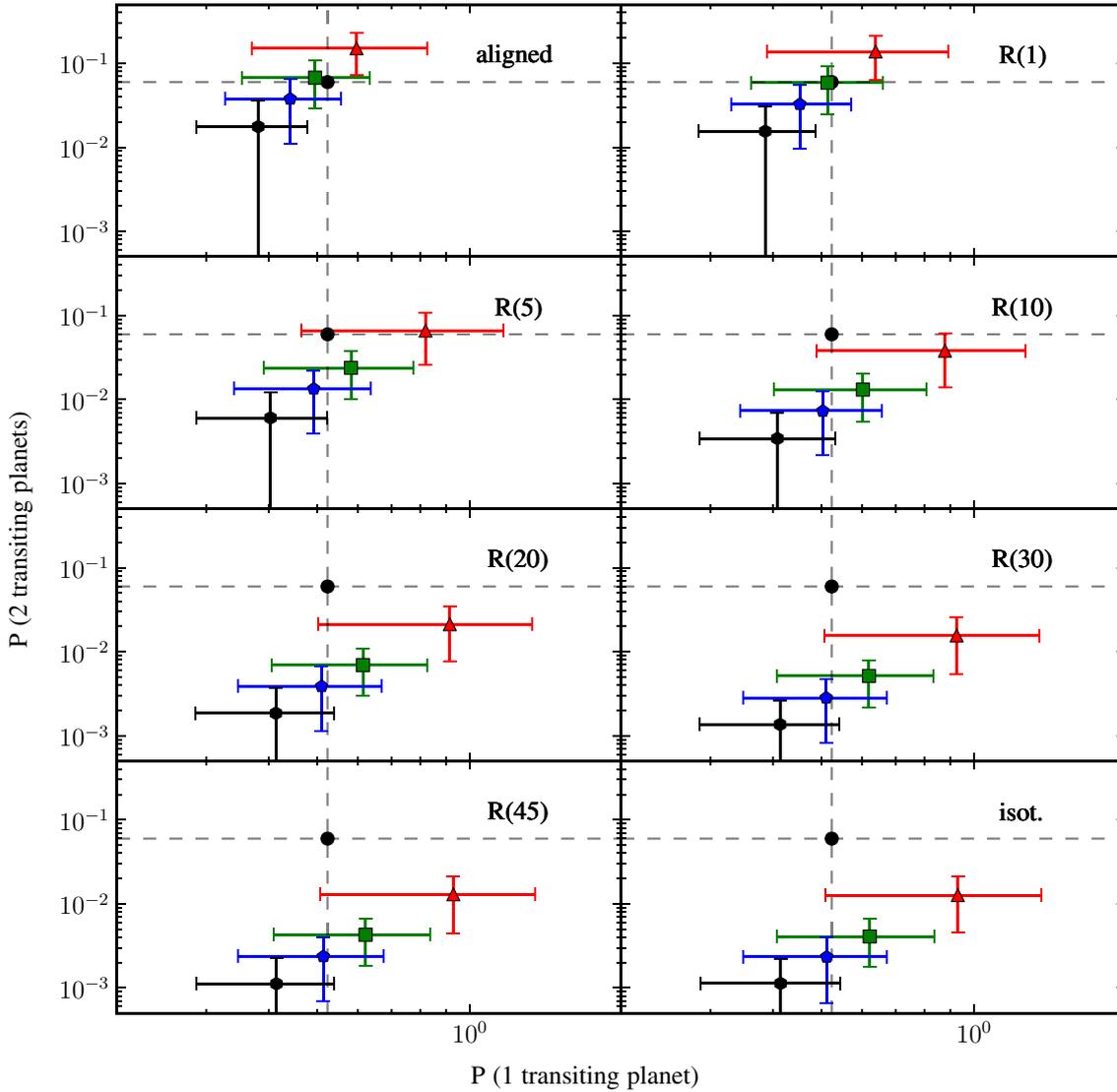| i dist.	O–C(5 $M_\oplus$)		O–C(7 $M_\oplus$)		O–C(9 $M_\oplus$)		O–C(10 $M_\oplus$)	
	σ_{f_1}	σ_{f_2}	σ_{f_1}	σ_{f_2}	σ_{f_1}	σ_{f_2}	σ_{f_1}	σ_{f_2}
aligned	0.33	1.17	0.21	0.20	0.72	0.83	1.48	2.35
$R(1.0°)$	0.46	1.05	0.07	0.03	0.59	1.19	1.38	2.85
$R(5.0°)$	0.84	0.16	0.31	2.60	0.22	5.20	1.01	8.57
$R(10.0°)$	0.91	0.91	0.40	6.30	0.13	10.29	0.93	16.08
$R(20.0°)$	0.95	2.89	0.44	13.32	0.09	20.01	0.87	30.94
$R(30.0°)$	0.96	4.50	0.46	19.28	0.08	29.93	0.86	46.67
$R(45.0°)$	0.96	5.77	0.46	23.58	0.06	35.61	0.87	51.39
isotropic	0.96	5.66	0.46	22.24	0.07	33.74	0.84	54.74

Fig. 4. Result of the simulations with the methodology described in 4.1 for the different i distributions considered. The figure depicts the match between the frequency of single and double transiting systems obtained from our simulations and *Kepler* results. In each plot the black circle represents *Kepler* results and the triangle, square, pentagon, and hexagon (red, green, blue, and black in electronic version only) represent the results for the 5, 7, 9, and 10 $M_\oplus$ cutoffs in $m \sin i$; the error bars result from the application of the 1-σ uncertainties. The different subpanels represent the application of different i distributions, from left to right and top to bottom: aligned, $R(1°)$, $R(5°)$, $R(10°)$, $R(20°)$, $R(30°)$, $R(45°)$, and isotropic, respectively.

7 and 9 $M_\oplus$ are those that better match the cutoff in mass; interestingly, the result for the $m \sin i$ cutoff of 7 $M_\oplus$ indeed provides the best match, and is twice as close to *Kepler*'s measured values than the second best. However, the question is more involved than that: the span in composition means that a planet with mass lower by a couple of % than the average mass cutoff has a non-zero probability of having a radius above the cutoff. This sends a clear warning about matching a sharp cutoff in mass with one in radius.

It is therefore important to note that only for an extreme choice of mass cutoff we have the (relatively) high Rayleigh mode of 5° as the best match. As we saw before, one would have to assume a mass cutoff significantly lower than 8.9 $M_\oplus$ and closer to 6.3 $M_\oplus$, which is at odds with the structure model's results. According to Sotin et al. (2007), a planet with mass of 6.3 $M_\oplus$ has a radius between 1.66 and 2.09 $R_\oplus$, for a water content of <0.1% and ~50%, respectively. This shows that planets with such mass are indeed too small to be included, or at least to

contribute significantly for the population we recover by imposing as limiting radius 2 $R_\oplus$. The only exception are strongly irradiated planets, as discussed by Rogers et al. (2011), for which, under very strong irradiation (500 < $T_{\rm eff}$ < 1000 K) and other assumptions, a planet with mass as low as 4 $M_\oplus$ can have a radius large enough to be included in our cutoff. However, once again, it is unlikely that these planets dominate our population.

It is important to note that Eq. (3) is only valid for coplanar systems. Multi-planetary systems with relative inclinations higher than zero have more degrees of freedom and are thus more easily unstable. Then, stability should require larger planet separations. Nevertheless, in this study we chose to treat all systems in the same way to avoid any uncontrolled bias. Since planets on wider orbits have lower transit probabilities, our choice slightly increases the transit probability of the outer planets of non-coplanar systems. The same argument can be used to show that if larger masses are used in the calculation of the Hill stability criterion, the spacing between the planets is increased and the
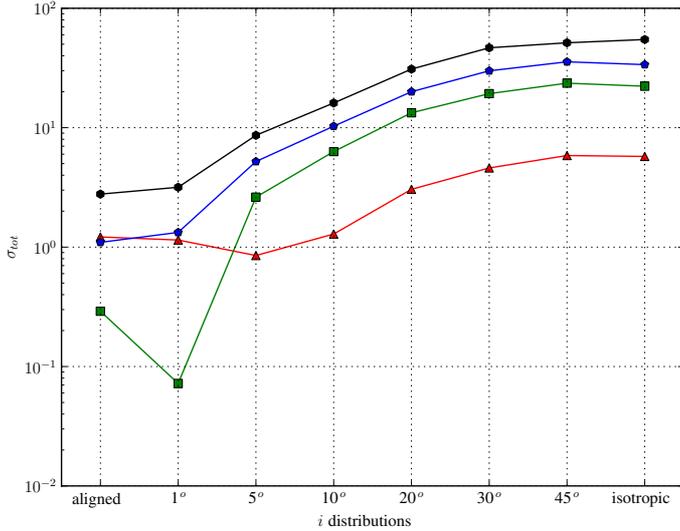
Fig. 5. Distance between *Kepler* point and the result from each simulation, measured in σ's; the triangle, square, pentagon, and hexagon (red, green, blue, and black in electronic version only) represent the results for the 5, 7, 9, and 10 $M_\oplus$ cutoffs in $m \sin i$, respectively.

transit probability decreased. Unfortunately, owing to the lack of appropriate stability studies for inclined systems, one cannot do better than we do here.

We can then conclude that the inclination distribution that better reproduces *Kepler* single and double-transit frequencies follow a Rayleigh distribution with mode of ~1° or even smaller, but we caution that the limitations of the approach presented here lead more to an order-of-magnitude result than to a clearly defined value. What can be asserted from our analysis is that a Rayleigh distribution with a mode of 5° can only be accommodated if we consider a population characterized by an extreme mass-radius relationship that would lead to very light planets (~6 $M_\oplus$) with a typical radius larger than 2 $R_\oplus$.

5.2. The impact of different stellar hosts: spectral types and metallicity

When comparing the outcome of different surveys, one must take into account that the population of stars surveyed even though similar, has different characteristics. The first point to evaluate is the spectral type of the hosts themselves, since planetary formation is expected to be a function of stellar mass (e.g. Alibert et al. 2011), among other parameters.

In Fig. 6 we overplot the two $T_{\rm eff}$ distributions obtained from HARPS and *Kepler* hosts considered in this study. We notice that *Kepler* host temperatures extend up to 400 K cooler than the coolest HARPS hosts, and conversely HARPS hosts extend to 400 K hotter than the hottest *Kepler* hosts. Both distributions have their peak around 5700 K, but the *Kepler* distribution is more peaked while HARPS is wider.

The $T_{\rm eff}$ for HARPS survey was determined as described by the calibration of Sousa et al. (2008) and its internal accuracy is of 50 K. The accuracy of the effective temperatures given by the *Kepler* Input Catalog (KIC) was estimated to be within 200 K for stars in the range between 4500 K and 6500 K (Brown et al. 2011). Recently, systematic differences between the griz magnitudes and those of the SDSS (DR8) have been noticed (Pinsonneault et al. 2012). Comparisons of $(J - Ks)$-based temperatures from the infrared-flux method (IRFM) and SDSS filter
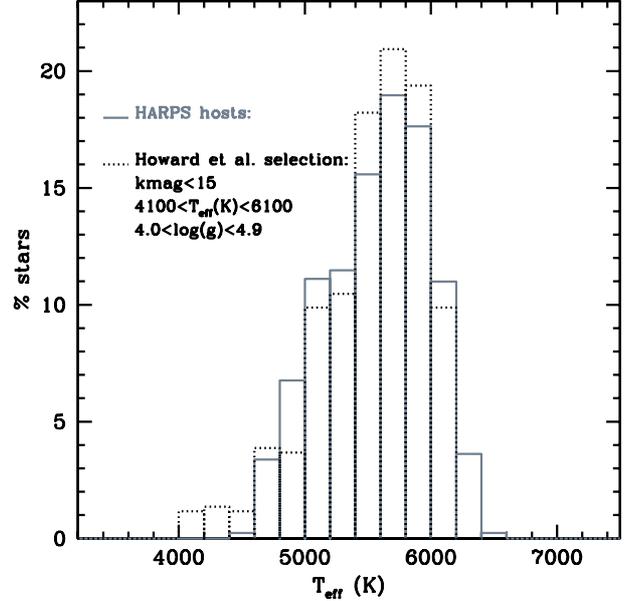

Fig. 6. Histogram of *Kepler* stars selected by effective temperature, from 3000 to 7500 K in bins of 200 K. In gray and with a solid line we present HARPS host and in black with a dotted line we present stars selected from the *Kepler* candidate stars according to the Howard et al. (2011) selection criteria (which are thus eligible for planetary detection).

indicated a mean shift toward hotter temperatures with respect to the KIC of the order of 215 K in between 4000 K and 6500 K. This result seems to be confirmed also by high-precision spectroscopic measurements. Metcalfe et al. (2010) determined the effective temperature of the sun-like star KIC11026764 by means of several spectroscopic approaches. They determined that this object is typically 100 to 300 K hotter than the value reported in the KIC, which is in line with the possible offset between our two distributions.

Another possible bias in our comparison may come from the different metallicity distribution of the HARPS and *Kepler* samples. On the one side, all HARPS stars are in the solar neighborhood (closer than 50 pc), and the vast majority belong to the Galactic thin disk. On the other hand, *Kepler* targets are all much farther away (no longer solar neighborhood objects). Adding to this, *Kepler* targets are slightly above the Galactic plane (Howard et al. 2011), suggesting that they may be more metal-poor (on average) than solar neighborhood stars because of a higher proportion of thick disk objects[5]. This may have important implications regarding the type of planets found in the two samples. For instance, it has been shown that giant planets are preferentially found around metal-rich stars (Santos et al. 2001, 2004; Fischer & Valenti 2005), while this same trend is not seen for neptunes and super-Earths (Udry & Santos 2007; Mayor et al. 2011; Sousa et al. 2011). The stellar metal content may also affect the stellar radius, since [Fe/H] correlates with stellar radius (Guillot et al. 2006). The impact of these effects is difficult to quantify; the two quantities are expected to be correlated. However, using the Besançon model, Howard et al. (2011) derived a ~0.1 dex effect in [Fe/H] as a function of spectral type, a value that is not very pronounced.

The implications of these differences are far from obvious and their study requires a detailed knowledge of the impact of stellar host properties on planet formation – and eventually a

---

[5] Thick-disk stars are on average more deficient than thin-disk counterparts (e.g. Adibekyan et al. 2011).

more refined analysis of *Kepler* hosts parameters. This is far beyond the reach of this paper, and we will content ourselves with noting that the two populations do not seem to differ in a significant way.

5.3. Comparison with other works and consequences for formation and evolution mechanisms

Lissauer et al. (2011b) analyzed the architecture of *Kepler* planetary system candidates (i.e. the observed multiplicity frequencies) and concluded that a single population of planetary systems that matched the higher multiplicities underpredicted the number of single-transiting systems. The authors also provided constraints on the frequency of systems with 1 to 6+ planets and mutual inclination, even though these two parameters are naturally correlated. However, their results were obtained using a methodology that required normalizing the number of simulated transits to reproduce the total number of transits detected by *Kepler*, making it vulnerable to the presence of false planet positives (FPP). Morton & Johnson (2011) calculated a FPP rate <10% for 90% of all *Kepler* candidates, the average being closer to 5% and with tails extending up to 30%. The presence of FPP also has an impact on our study, since it will modify the planetary period distribution and stellar radii and mass distributions. However, and unlike Lissauer et al. (2011b), here the distinction between the different models is made essentially through the frequency of double transiting extrasolar planets, for which the FPP is much lower than on average or for single-planet systems (Lissauer et al. 2012), which makes our results probably more robust.

Tremaine & Dong (2012) described a very general formalism to analyze the multiplicity function in transit and RV surveys based on the approximation of separability, that the probability of distribution of planetary parameters in a system is the product of identical one-planet distributions. The authors applied it both to the *Kepler* survey and to RV+*Kepler* data (in which they used the planets listed in *exoplanet.org* as of August 2010 for RV data) and reached several interesting conclusions. They found out that a wide range of inclinations were compatible with the *Kepler* data alone, as long as some of the stars contained many planets per system (curiously different from the conclusions of Lissauer et al. 2011b, but still in agreement with their upper limit on an rms inclination of 10°). Their joint analysis of RV and transit data led to an rms inclination between 0 and 5°, a result compatible with ours. Their wider range of inclinations can be explained by their different working hypothesis, different RV data, and especially by a more general and less detailed treatment of detection biases for both surveys.

The very recent work of Batalha et al. (2012) presents a new list containing 1091 new *Kepler* exoplanet candidates based on the analysis of an extended data-set covering 16 months. However, and as already discussed before, the lack of information on the photometric limits as discussed first by Howard et al. (2011) and then by Youdin (2011) for the Borucki et al. (2011) release prevents us from repeating our analysis for the new dataset. Still, some educated guesses can be made on its impact on our results. The first point to note is that the highest fractional increases are obtained for small-radius candidates (197% for candidates smaller than 2 $R_\oplus$ compared to 52% for candidates larger than it) and located at longer orbital periods (123% for planets outside of 50 day orbits versus 85% for candidates inside of 50 day orbits) and thus not for the population discussed here. If we apply to this dataset the same selection criteria we applied to Borucki et al. (2011), the number of stars with one

transiting candidate increase from 274 to 384 (a 40% increase) and the stars with two transiting candidates increase from 30 to 41 (a 37% increase). Since the match of observations to the models in our study is governed mostly by the frequency of stars with two transiting planets, we expect that a higher frequency of stars with two transiting planets will favor more aligned models. As discussed in Batalha et al. (2012), these frequencies will increase, owing to the higher number of events and the usage of a more precise pipeline. However, without applying the detectability correction coming from photometric limits we cannot translate the absolute number of events into the required frequencies and draw firm conclusions. We leave this analysis to a future paper.

The recently submitted paper of Fabrycky et al. (2012) revisits the architecture of *Kepler* multi-transit systems and discussed several of their most important properties. Among the addressed points is the distribution of inclinations for the candidates, which the authors estimate to be between 1 and 2.3°. The authors conclude then that the new release of *Kepler* data alone allows one to establish that planetary systems are strongly coplanar and declare that studies such as Tremaine & Dong (2012) and our own, which draw information from RV studies too, are unnecessary complicated by the RV surveys' different characteristics. However, as demonstrated by Tremaine & Dong (2012), the transit data alone cannot exclude, for instance, the presence of two population of planetary systems: one whose relative inclination is low and other population with different properties, and potentially higher planet relative inclinations and planetary frequencies. In other words, it is important to keep in mind that Fabrycky et al. (2012) evaluate the inclination of *planetary systems with multi-transiting planets*, while the studies that rely on RV and *Kepler* evaluate the inclination of *planetary systems as whole*. The two values are not derived for the same population of planets and cannot be compared directly.

Our work shows that planetary systems are likely to host planets with a very low inclination relative to the plane of the system. This is already the case for the solar system, which has an average inclination <2°, and favors the standard model for planet formation in a disk. More importantly, it suggests that most planets in systems do not have their orbital elements influenced by violent angular momentum exchanges such as planet-planet scattering (e.g. Nagasawa et al. 2008), Kozai oscillation (e.g. Wu & Murray 2003) or perturbation by a stellar encounter (e.g. Malmberg et al. 2011), mechanisms that are only expected to create short-period single planets and not systems. As already pointed out by Greg Laughlin[6], the distribution of orbital parameters for *Kepler* multi-transit systems is very similar to that of the same (scaled) parameters for Solar System giant planet satellites. This is well in line with the fact the planets should be formed in a disk, with relative low orbital inclinations, as we show here.

6. Conclusions

We attempted at a first comparison between the planetary population properties as characterized by the HARPS and *Kepler* surveys. We simulated the population of planets, with planetary frequencies dictated by HARPS survey results, and considered that these planets followed different inclination distributions relative to the systems' plane. We considered distributions from aligned to Rayleigh distributions with different modes, to finally isotropic (completely independent).

---

[6] Oral communication at the "First *Kepler* Science Conference", December 2011.

The first remarkable point is the compatibility of the results from the two surveys. This is made even more so by the very little freedom we have on influencing our simulations' outcome. There are no tunable knobs other than the inclination distribution. Concerning these, we showed that the results point to a strong alignment of the systems, with an inclination between the planet and the plane of the system better described by a Rayleigh distribution with mode of ~1° or even lower. This is a feature that depends on the assumed mass-radius relationship, but not in a strong way, an important point since the mass-radius relationship for low-mass planets is ill-defined and hardly unique because of the span of compositions and mass that can generate the same radius. We have shown that only in extreme and consequently highly unlikely cases the best fit for the alignment between planets is expected to follow an $R(5°)$.

These results have important consequences for our understanding of the role of several proposed formation and evolution mechanisms. They confirm that planets are likely to have been formed in a disk and show that most planetary systems evolve quietly without strong angular momentum exchanges such as those produced by the Kozai mechanism or planet scattering.

References

Adibekyan, V. Z., Santos, N. C., Sousa, S. G., & Israelian, G. 2011, A&A, 535, L11

Alibert, Y., Mordasini, C., & Benz, W. 2011, A&A, 526, A63

Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2012, ApJS, submitted [arXiv:1202.5852]

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

Borucki, W. J., Koch, D. G., Basri, G., et al. 2011, ApJ, 736, 19

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

Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2012, ApJ, submitted [arXiv:1202.6328]

Fischer, D. A., & Valenti, J. 2005, ApJ, 622, 1102

Gladman, B. 1993, Icarus, 106, 247

Guillot, T., Santos, N. C., Pont, F., et al. 2006, A&A, 453, L21

Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2011, ApJ, submitted [arXiv:1103.2541]

Lissauer, J. J., Fabrycky, D. C., Ford, E. B., et al. 2011a, Nature, 470, 53

Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011b, ApJS, 197, 8

Lissauer, J. J., Marcy, G. W., Rowe, J. F., et al. 2012, ApJ, 750, 112

Lovis, C., Ségransan, D., Mayor, M., et al. 2011, A&A, 528, A112

Malmberg, D., Davies, M. B., & Heggie, D. C. 2011, MNRAS, 411, 859

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

Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20

Mayor, M., Udry, S., Lovis, C., et al. 2009, A&A, 493, 639

Mayor, M., Marmier, M., Lovis, C., et al. 2011, A&A, submitted [arXiv:1109.2497]

Metcalfe, T. S., Monteiro, M. J. P. F. G., Thompson, M. J., et al. 2010, ApJ, 723, 1583

Moorhead, A. V., Ford, E. B., Morehead, R. C., et al. 2011, ApJS, 197, 1

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

Nagasawa, M., Ida, S., & Bessho, T. 2008, ApJ, 678, 498

Pinsonneault, M. H., An, D., Molenda-Żakowicz, J., et al. 2012, ApJS, 199, 30

Rogers, L. A., Bodenheimer, P., Lissauer, J. J., & Seager, S. 2011, ApJ, 738, 59

Santos, N. C., Israelian, G., & Mayor, M. 2001, A&A, 373, 1019

Santos, N. C., Israelian, G., & Mayor, M. 2004, A&A, 415, 1153

Sotin, C., Grasset, O., & Mocquet, A. 2007, Icarus, 191, 337

Sousa, S. G., Santos, N. C., Mayor, M., et al. 2008, A&A, 487, 373

Sousa, S. G., Santos, N. C., Israelian, G., Mayor, M., & Udry, S. 2011, A&A, 533, A141

Tremaine, S., & Dong, S. 2012, AJ, 143, 94

Udry, S., & Santos, N. C. 2007, ARA&A, 45, 397

Valencia, D., Sasselov, D. D., & O'Connell, R. J. 2007, ApJ, 665, 1413

Wolfgang, A., & Laughlin, G. 2011, ApJS, submitted [arXiv:1108.5842]

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

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