Draft version December 7, 2020
Typeset using LaTeX **twocolumn** style in AASTeX63

A Distinct Population of Small Planets: Sub-Earths

Yansong Qian (钱岩松)[1] and Yanqin Wu (武延庆)[1]

[1] *Department of Astronomy and Astrophysics, University of Toronto, Toronto, ON M5S 3H4, Canada*

ABSTRACT

There is a well-known gap in the sizes of small planets, between super-Earths and mini-Neptunes. This is explained by the envelope stripping of mini-Neptunes at short orbits. Here, we report the presence of another gap at a smaller size (around one Earth radius). By focussing on planets orbiting around GK-dwarfs inward of 16 days, and correcting for observational completeness, we find that the number of super-Earths maximize around 1.4 Earth radii and disappear shortly below this size. Instead, a new population of planets (sub-Earths) appear to dominate at sizes below ~ 1 Earth radius. This pattern is also observed in ultra-short-period planets.

The end of super-Earths supports earlier claims that super-Earths and mini-Neptunes, planets that likely form in gaseous proto-planetary disks, have a narrow mass distribution. The sub-Earths, in contrast, can be described by a power-law mass distribution and may be explained by the theory of terrestrial planet formation. We therefore speculate that they are formed well after the gaseous disks have dissipated. The extension of these sub-Earths towards longer orbital periods, currently invisible, may be the true terrestrial analogues. This strongly motivates new searches.

Keywords: planets

1. INTRODUCTION

Most extra-solar planets known to date are likely formed in gaseous protoplanetary disks. These include the Jovian planets, the super-Earths and their siblings, the mini-Neptunes (the latter two are sometimes jointly called 'Kepler planets' for short). However, there is little evidence for the existence of terrestrial planets outside our own solar system. Terrestrial planets are thought to have formed from collisional debris, well after the dispersion of proto-planetary disks. We term these **Generation II planets**, to distinguish them conceptually from the first group (Generation-I planets).

Currently, evidences for Gen-II planets are best seeked using data from the *Kepler* mission. The final *Kepler* mission transit planet search (DR25, Twicken et al. 2016) includes nearly 200,000 stellar targets and yields ~ 4700 planet candidates. Such a large sample allows us to examine the exoplanet populations in detail. A majority of these candidates are super-Earths and mini-Neptunes. It has been convincingly shown that these two populations are separated by a radius gap (Fulton et al. 2017; Van Eylen et al. 2018), readily explained by the theory of photo-evaporation (Lopez et al. 2012; Owen & Wu 2013, 2017; Jin et al. 2014), or theory of core-powered mass-loss (Ginzburg et al. 2018). As such, the two populations have the same origin, with super-Earths being former mini-Neptunes that happen to lie too close to their host stars. Moreover, the sizes of these planets appear to be narrowly distributed[1], and rise with the masses of the host stars (Fulton & Petigura 2018; Wu 2019; Cloutier & Menou 2020).

On the other hand, the initial aim of *Kepler*, Earth-like planets, remain elusive. While there are a smattering of small planets at short orbital periods that are Earth-sized, *Kepler* has a very low survey completeness for such planets at AU-distances. It is also unclear if the close-in ones are low-mass extension of the super-Earths (and are therefore Gen-I planets), or are a distinct population. This is the very question we set out to answer in this work.

yansong.qian@mail.utoronto.ca

wu@astro.utoronto.ca

[1] This is already clear for the mini-Neptunes, and becomes clear for the super-Earths as well in this work.

Our effort is made possible thanks to two recent advances. The planet detection efficiency of the DR25 pipeline has now been characterized for every star by injection tests at the flux-level (Burke & Catanzarite 2017), and characterized for the group as a whole by injection tests at the pixel-level (Christiansen et al. 2020). Meanwhile, uncertainties in the radii of KOIs have now been dramatically reduced using GAIA DR2 stellar radii (Berger et al. 2020a).

In this work, we focus on planet candidates that orbit around GK dwarfs, with periods shortwards of 16 days. We are able to show that the super-Earths, like the mini-Neptunes, have a narrow size distribution. Their occurrences fall off rapidly below a peak. This feature enables us to detect the rise of a new population at smaller sizes, the sub-Earths. There is a clear gap between the super-Earths and sub-Earths.

Results similar to ours, that there is a new population of small planets, has been obtained or alluded to previously. Neil & Rogers (2020) found that *Kepler* data is best explained by three populations of exoplanets: gaseous (mini-Neptunes), evaporated cores (super-Earths) and intrinsically-naked. However, their joint distribution does not show a clear gap as we find here. In a different approach, Hsu et al. (2019) studied the planet occurrence rates across a wide range of orbital periods and planet radii. Their radius bins are broader than ours here, so any possible gap feature is somewhat smeared. They do, however, report an excess of small planets, consistent with our results here.

Both these works aim to determine the more general landscape of small planets, while ours drills down to one specific question. This allows us to focus on the most revealing parameter space and to slim down the model. We are also able to test the robustness of our results by experimenting with different sample selections and completeness functions.

2. PLANET AND STAR SAMPLES

Starting from the Kepler DR25 Supplemental catalogue of 200,038 host stars and 4717 planet candidates (Mathur & Huber 2016) (excluding false positives, NASA Exoplanet Archive), we introduce a number of cuts to obtain a suitable sample for this study. We give our justifications for these cuts and list the resulting sample sizes in Table 1. In §5.1, we explore how results may be impacted by different sample cuts.

1. **GAIA-able:** We only consider stars that are in the Gaia-Kepler Stellar (GKS) Catalog (Berger et al. 2020b). In addition, we follow Berger et al. (2020a) to exclude stars with Gaia DR2 renormalized unit-weight error (RUWE) larger than 1.2 and with isochrone-derived goodness-of-fit parameters lower than 0.99. These cuts can remove binary stars and ensure the reliability of parameters. Compared to the much larger radius uncertainty in the KIC catalogue (typically 25%, Fulton et al. 2017), the GKS catalogue employs GAIA parallaxes to achieve a much lower uncertainty of $\sim 4\%$ in stellar radius (Berger et al. 2020b). This allows us to separate planets into smaller radius bins and to observe the fine details in planet occurrence.

2. **GK-dwarf:** It has been noticed before (Fulton et al. 2017; Wu 2019; Berger et al. 2020a; Cloutier & Menou 2020) that the sizes of both super-Earths and mini-Neptunes rise systematically with stellar mass. Such a trend is as yet unexplained. For our purpose, this tends to smear any features in the planet size distribution. This leads us to focus only on GK-dwarfs, within a stellar radius range of $R_* \in [0.75R_\odot, 1R_\odot]$ (and a similar stellar mass range). It is beneficial to study smaller stars, as small planets are more readily detectable around them. This choice of GK-dwarfs also ensures a sufficiently large sample. Finally, we choose our sample based on stellar radius, rather than stellar mass as provided by Berger et al. (2020b). The former has a median uncertainty of 4%, while the latter is larger at 7%.

3. **Brightness:** We discard stars with K_p magnitudes above 15. Very few Earth-sized planets are discovered around dim stars, and both the false-positive probabilities and radius errors rise for planet candidates around these stars (Petigura et al. 2013).

4. **NoisyTargetList** is a list of Kepler stars with too high noise to apply the the detection efficiency model. Thus, we exclude stars on this list.

5. **False Positives:** we remove planet candidates with a false-positive-probability $FPP > 0.1$, based on the False Positive Probabilities(FPP) Table produced by Morton et al. (2016). This alleviates confusion by background eclipsing binaries. For the remaining ones, we assume a reliability of 100% (more below on this).

6. **Impact Parameters:** candidates with high impact parameters typically suffer larger radius uncertainties. For instance, Berger et al. (2020a) shows that removing these candidates will clarify the radius gap between super-Earths and mini-

Neptunes. We agree with their assessment and remove all candidates with impact parameter greater than 0.9.

7. **Planet Radius**: we remove large planets with radii above $4R_\oplus$ and small planets below $0.6R_\oplus$. Below $0.6R_\oplus$, the detection efficiency drops too low to allow any secure statement.

8. **Period range:** a major distinction in this study is our focus on short period planets, inward of 16 days. By zooming in to the region where small planets can potentially be discovered (detection efficiency above $\sim 10\%$), we contain the uncertainties inflated by rare detections. While this reduces our planet sample, it actually helps increasing our sensitivity to small planets.

Eventually, we are left with 280 planet candidates and 13,297 qualified stars. In the following, we consider the pipeline completeness for this sample.

Cut	n_p	N_*
Exoplanet Archive	4717	200,038
GAIA-able	3898	162,657
$0.75 < R_*/R_\odot < 1$	1263	34,399
$K_p \leq 15$	670	13,498
NoisyTargetList	651	13,297
FPP < 0.1	516	-
$b \leq 0.9$	481	-
$0.6 \leq R_p \leq 4R_\oplus$	444	-
$1 \leq P_{\rm orb} \leq 16$	280	-

Table 1. Star and Planet Sample Cuts.

3. COMPLETENESS

For every planet in the *Kepler* field, its detection depends on a number of factors: geometric probability of transit; pipeline efficiency in identifying the transit as a threshold-crossing-event (TCE); and loss during the vetting process. In the following, we discuss the last two factors, together with the issue of reliability.

3.1. *Pipeline Detection Efficiency*

By injecting simulated transit signals into real light curves and then recovering them using the same search pipeline, one can determine the so-called 'pipeline efficiency' as a function of planet size, orbital period and stellar properties. The *Kepler* team have published multiple studies of pipeline completeness (e.g. Christiansen et al. 2016; Thompson et al. 2018). The most recent ones are those by Burke & Catanzarite (2017); Christiansen et al. (2020) for the DR25 pipeline. The Christiansen et al. (2020) work is based on the so-called pixel-level transit injection tests. Limited by the computation capacities, each Kepler star only receives one planet injection and the resulting completeness is an average over the entire star sample. In contrast, the much simpler flux-level tests (Burke & Catanzarite 2017) can inject many different signals into a given star, and one can survey a range of planet parameters (radius and period) on a per-star basis, to obtain the detection efficiency, $DE(p, r, \theta_*)$, as a function of orbital period, planet radius and host star properties. Such a model captures the target-to-target variations and is recommended for studies involving a small subset of *Kepler* targets, as is the case here. We will adopt this model and use the KeplerPorts python code (Burke & Catanzarite 2017) to generate the pipeline efficiency.

In detail, KeplerPorts calculate the detection efficiency, $DE(P, R_p, \theta_*)$, based on the so-called multiple event statistics, $MES(P, R_p, \bar{\theta})$ (Jenkins 2002; Christiansen et al. 2012), with $\bar{\theta}$ being stellar parameters (R_*, $\log g$ and $T_{\rm eff}$). This statistics is closely related to the transit signal-to-noise ratio. The detection efficiency is also empirically found to correlate strongly with the so-called CDPP slope, the temporal behaviour of the stellar noise. This dependency is taken into account by the Burke & Catanzarite (2017) model, based on injection experiments. Furthermore, the method of 'MES Smearing' is applied to transits with non-zero impact parameters.

The resultant detection efficiencies are then averaged over the $N_* = 13,297$ GK-dwarfs in our sample. For the radius and period ranges of interests, these efficiencies are plotted in the left-hand panel of Fig. 1. One observes that, inward of 16 days, transiting planets larger than $1.3R_\oplus$ can be detected fairly completely ($DE \geq 90\%$), and the efficiency drops below $\sim 50\%$ for planets smaller than $0.9R_\oplus$. The right-hand panel presents the overall efficiency, $\eta(p, r)$, one that also includes the geometric transit probability and impact parameter correction. This is calculated as 1.

$$\eta(P, R_p) = \frac{1}{N_*} \sum_{i=1}^{N_*} DE_i(P, R_p) \times P_{\rm geo,i}(P, R_p) \times 0.9\,, \tag{1}$$

where $P_{\rm geo,i} = 2R_{*,i}/a$ and we adopt an estimate for the star mass from Berger et al. (2020b). The factor 0.9 reflects our range for the impact parameter.

3.2. *Vetting Completeness*

Binary stars and some astrophysical noise can masquerade as transit signals. So every TCE needs to be vetted before qualifying as a planet candidate. This can introduce a reduction in completeness as some genuine

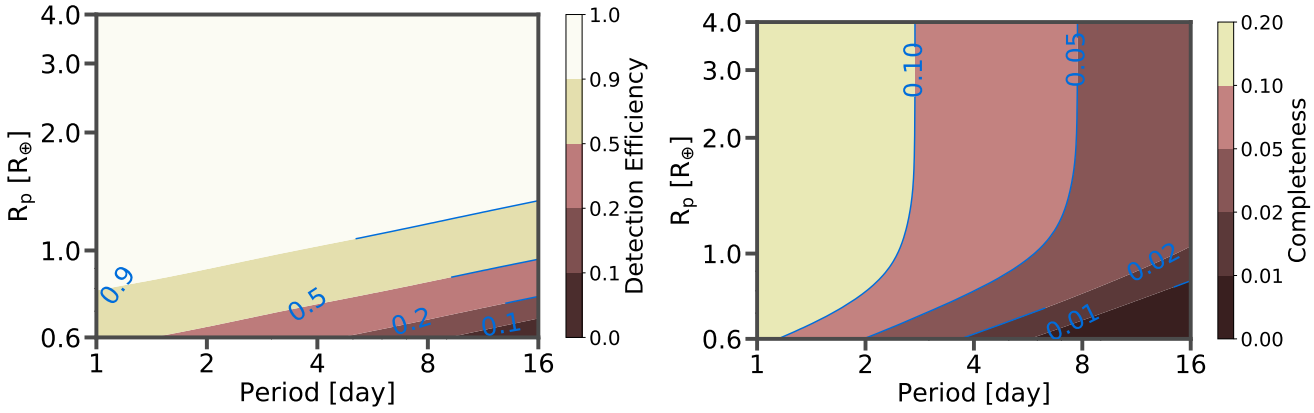

Figure 1. Detection completeness. The left panel shows the averaged pipeline detection efficiency for the GK-dwarfs in our sample, as a function of planet radius and period, obtained using the Burke & Catanzarite (2017) model. Within 16 days, nearly all transiting planets with radii greater than $1.3R_\oplus$ are recovered. Moreover, at least half of the planets with radii greater than $0.9R_\oplus$ are recovered. The right panel shows the overall efficiency (eq. 1) when transit probability is also considered.

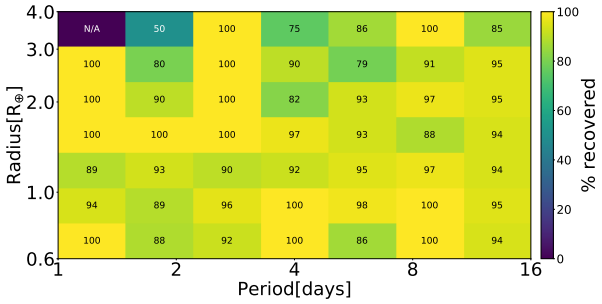

Figure 2. Vetting efficiency for our ranges of planet radius and orbital period, obtained using Robovetter (Coughlin 2017). On average, 94% simulated planets within this range pass the vetting tests.

planets may be vetted out as false positives. This vetting efficiency is presented in Fig 2 for our ranges of interest, produced using the robotic vetting code Robovetter (Thompson et al. 2018; Coughlin 2017). For the short period planets we are interested in, the vetting efficiency typically lies above 90%, with no obvious dependency on planet radius or period. So in this study, we simply assume this to be 100%.

3.3. *Reliability*

Morton et al. (2016) assigned a false positive score (FPP) for each planet candidate, using an automated validation procedure. This uses a full range of observational constraints to calculate the probabilities of various hypothesis scenario (genuine planets, brown dwarfs, eclipsing binaries, etc). We have discarded all False Positives, but retained candidates in both the 'confirmed' and the 'candidate' categories, with the proviso that FPP < 0.1. We follow convention and define a reliability index $= (1 - \mathrm{FPP})$. These are shown in Fig. 3 by the colors of the dots. Since reliability is high in our sample, we simply ignore this distinction and assume all candidates are genuine.

4. SIZE DISTRIBUTION OF SMALL PLANETS – A NEW GAP

Equipped with the planet sample and their detection efficiencies (Fig. 3), we proceed to recover the underlying distribution. We first present our analysis using the straight-forward IDEM method, before elaborating on a Bayesian approach. The two are shown to produce compatible results.

In the following, we divide our planet sample on a logarithmic grid of period and radius. To ensure a sufficient radius resolution, we choose to use 20 radius bins from 0.6 to $4R_\oplus$. To ensure a sufficient number of candidates in each bin, we choose a relatively crude period resolution, 5 bins from 1 to 16 days. Even with this choice, there are some bins that contain zero or only one candidate and their information contents have to be carefully evaluated (see below).

4.1. *IDEM: inverse detection efficiency method*

One straight-forward, non-parametric way of obtaining the occurrence rate is called 'IDEM' (Howard et al. 2012; Petigura et al. 2013). Assuming the observed population is drawn from an underlying population weighted by the detection efficiency, one can retrieve the differential occurrence rate by dividing the number of planets detected in the bin $\Delta \ln P \Delta \ln R_P$, $n(\mathrm{P}, \mathrm{R_P})$, by the detection efficiency,

$$\frac{d^2 f}{d \ln P d \ln R_P} = \frac{1}{\Delta \ln P \Delta \ln R_P} \times \frac{1}{N_*} \times \frac{n(\mathrm{P}, \mathrm{R_P})}{\eta(P, R_P)}, \quad (2)$$

where $\eta(P, R)$ is the detection efficiency, suitably averaged over the relevant grid, and N_* is the stellar sample

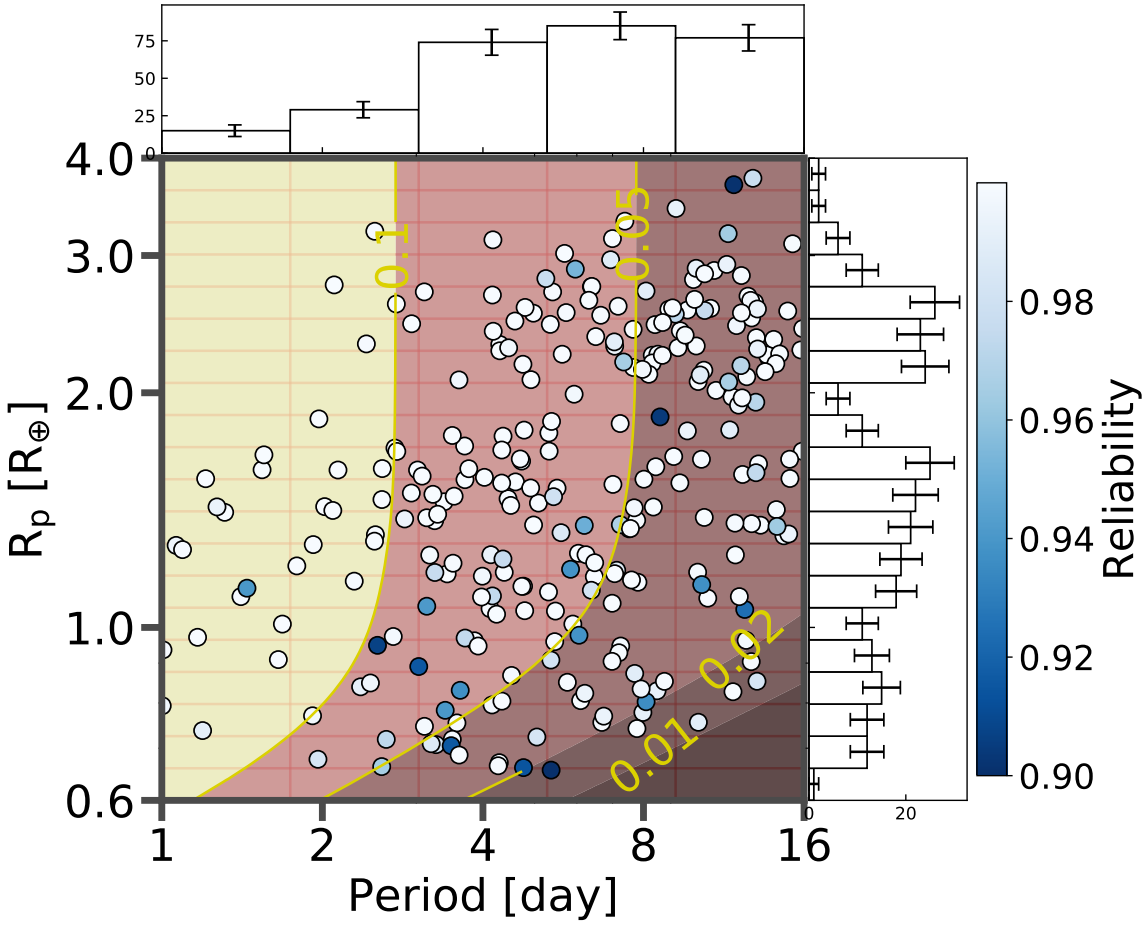

Figure 3. The sample of 280 planet candidates in our analysis, plotted as their observed radii vs. periods. The contours indicate the total detection efficiency (including transit probability), and the color of each dot indicates the reliability index (1-FPP). The size histogram and period histogram are shown to the left and the top of the main figure. Even without correcting for completeness, one can already see the two gaps, one between super-Earths and mini-Neptunes at around $2R_\oplus$, and one between sub-Earths and super-Earths at around $1R_\oplus$.

size. We then sum over the period bins to obtain a size distribution

$$\frac{df}{d\ln R_P} = \sum \frac{d^2 f}{d\ln P d\ln R_P} \Delta \ln P \,. \tag{3}$$

The dominant source of error is Poisson error, as opposed to measurement errors. To obtain an estimate on the uncertainty, we note that since we are dealing with small candidate numbers in each bin (some bins even contain zero detection), the confidence interval cannot be estimated using the usual $1/\sqrt{n}$ approach. Instead, we adopt Table VI of Feldman & Cousins (1998) to obtain the 95% confidence limits when presented with a detection $n(P, R_P) \geq 0$. When summing over different period bins, we assume that the confidence limits from each bin are independent and adds the errors quadratically. The uncertainty range will then rise with bin number. This, we believe, yields a maximal uncertainty range.

This approach departs from previous IDEM studies (e.g. Howard et al. 2012; Petigura et al. 2013; Kunimoto & Matthews 2020) in providing a more conservative error estimate. To obtain a more realistic and smaller uncertainties, one needs to adopt a model-based approach (e.g., Bayesian, see below) that accounts for correlated information among bins, while paying a price for extra assumptions.

The result of this analysis is presented in tandem with the Bayesian model in Fig. 5.

4.2. *Bayesian Modelling*

4.2.1. *Model Ansatz*

To conduct a Bayesian analysis, we need to make an array of assumptions for the underlying distribution.

First, we will depart from tradition and consider the distribution in planet mass and period. We use planet core-radius (R_c) as a proxy for its mass (roughly, $M_p \propto R_c^4$, Fortney et al. 2007). Since mini-Neptunes

and super-Earths are of the same origin and therefore likely share the same mass distribution, it seems sensible to combine their signals together by working on planet mass. Moreover, this enables us to separate the distribution in period-radius space (see below).

While $R_c \approx R_P$ for super-Earths, mini-Neptunes need some explanation. According to the theory of photo-evaporation, the radii of the two populations are spaced by a ratio that is nearly 2. We adopt a number 1.7 taken from Owen & Wu (2017); Wu (2019). This ratio is the result of fitting the observed distributions with evaporation theory. It is also the ratio between the observed super-Earth peak ($\sim 1.4R_\oplus$) and mini-Neptune peak ($\sim 2.4R_\oplus$) in our sample. So in our study, for any planets lying above $2R_\oplus$ (the so-called 'Fulton'-gap), we simply divide their observed radii by 1.7 to obtain their core radii. When correcting for detection completeness, we note that while super-Earths are smaller and are in principle more incomplete compared to mini-Neptunes, the difference is minor and can be ignored (Fig. 3).

One naturally wonders if this procedure produces artificial features. We show that folding Neptunes into Super-Earths, while enhancing the super-Earth peak, is not critical to our conclusions on the existence and characters of the sub-Earths.

Next, we assume that the planet distribution in the $P - R_c$ plane is separable,

$$\frac{d^2 f}{d\ln P\, d\ln R_c} = \frac{df_P}{d\ln P} \times \frac{df_R}{d\ln R_c}\,. \tag{4}$$

Such a separability assumption, often adopted in literature (e.g. Silburt et al. 2015; Mulders et al. 2018; Neil & Rogers 2020), is problematic when describing the distribution of super-Earths and/or mini-Neptunes individually, since their relative proportion depends on distances from the star. It is, however, reasonable when describing the joint distribution. Studies of the planet mass distribution (e.g. Wu 2019) have shown that it is relatively distance dependent. This is another reason behind our adoption of R_c as the relevant parameter.

In the following, we will experiment with different forms of f_R, while using the same broken-power-law for f_P,

$$\frac{df_P}{d\ln P} = \begin{cases} \left(\frac{P}{1\mathrm{day}}\right)^{\alpha}, & \text{if } P \le P_b \\ \left(\frac{P_b}{1\mathrm{day}}\right)^{\alpha-\beta}\left(\frac{P}{1\mathrm{day}}\right)^{\beta}, & \text{otherwise} \end{cases} \tag{5}$$

where P_b, α, β are free parameters to be determined from data. This form is motivated by previous studies of super-Earths/mini-Neptunes (e.g. Silburt et al. 2015; Mulders et al. 2018; Neil & Rogers 2020). We assume that any new population of planets also satisfy the same rules. We relax this assumption in Model 4.

For the planet mass (or core-radius) distributions, we adopt the following three models, each with a different physical significance.

Model 1: one Gaussian. Here, we assume that nature produces only one single population of planets, for core-radius from $0.6\oplus$ to $2\oplus$ (or from ~ 0.2 to $16M_\oplus$, if these planets all share the terrestrial composition). We describe this using a log-Normal distribution in core-radius,

$$\begin{aligned}\frac{df_R}{d\ln R_c} &= A \times N(\ln R_c, \mu_0, \sigma_0) \\ &= A \times \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\ln R_c - \ln\mu_0)^2}{2\sigma_0^2}\right]\,. \end{aligned} \tag{6}$$

We adopt broad linear priors for for both μ_0 and σ_0: $\mu_0 \in [0.6, 2]R_\oplus$ and $\sigma_0 \in [0, 1]$.

Model 2: two Gaussians. To allow for a separate population of planets, we assume the core-radius distribution is a superposition of two Gaussians,

$$\frac{df_R}{d\ln R_C} = B \times N(\ln R_c, \mu_1, \sigma_1) + C \times N(\ln R_c, \mu_2, \sigma_2)\,, \tag{7}$$

where the first term describes the Super-Earth/Neptune population, and we adopt linear priors of $\mu_1 \in [1.3, 1.6]R_\oplus$ and $\sigma_1 \in [0, 0.5]$. These narrower priors are intuited from both the above IDEM analysis and previous studies (e.g. Hsu et al. 2019; Neil & Rogers 2020). For the second term, employed to describe a new lower-mass population , we assign broad linear priors of $\mu_2 \in [0.6, 1.2]R_\oplus$ and $\sigma_2 \in [0, 1]$.

Model 3: power-law+Gaussian. While super-Earth/Neptunes appear to be well described by (narrow) Gaussians, this does not appear so for the new population of planets we are after. The above 2-Gaussian results, as well our IDEM analysis, suggest that they may have a broad radius range, compatible with a power-law distribution. We attempt the following power-law/Gaussian combination,

$$\frac{df_R}{d\ln R_C} = B \times N(\ln R_c, \mu_1, \sigma_1) + D \times \left(\frac{R_C}{R_\oplus}\right)^{\gamma}\,. \tag{8}$$

We select a linear prior of $\gamma \in [-3, 1]$. This is not restrictive, as it allows for both ascending and descending trends towards small sizes.

Model 4: 2-periods. The previous three models adopt the same period distribution for any planet population. Here we relax this assumption by allowing for a separate period distribution for each population (based

on model 3),

$$\frac{d^2f}{d\ln P d\ln R_C} = B \times N(\ln R_c, \mu_1, \sigma_1) \times \left(\frac{df_P}{d\ln P}\right)_I + D \times \left(\frac{R_c}{R_\oplus}\right)^\gamma \times \left(\frac{df_P}{d\ln P}\right) . \quad (9)$$

Here both period distributions have the same form as eq. (5) but they can have different parameters. Moreover, since the super-Earths/Neptunes extend to well beyond 16 days and their period-distribution has been well determined (see, e.g., Neil & Rogers 2020), we can adopt narrower priors on their parameters based on previous studies (see Table 2).

4.2.2. *Bayesian Operations*

Here, we apply Bayesian inference to determine the parameters of maximum posterior probability.

Given the observed properties of a planet sample ('PCs'), the posterior probability of a hypothesis θ (with its attendant parameters) can be calculated using the Bayes' theorem as

$$P(\boldsymbol{\theta} \mid \{PCs\}) = P(\{PCs\} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta}) , \quad (10)$$

where $P(\{PCs\} \mid \boldsymbol{\theta})$ is the likelihood and $\pi(\boldsymbol{\theta})$ the prior. We use exclusively flat priors.

As the planets are discreet points in the period-radius space, we treat them as Poisson point process, following earlier works (e.g., Youdin 2011; Burke et al. 2015; Bryson et al. 2020). Defining the occurrence rate within a bin $\Delta\ln P\,\Delta\ln R_C$ as

$$\Delta f(\boldsymbol{\theta}) = \frac{d^2f}{d\ln P d\ln R_C}\,\Delta\ln P\,\Delta\ln R_C \quad (11)$$

where $\boldsymbol{\theta}$ stands for an ansatz and its attendant parameters, we can write the total expected number of planet detection as

$$N_{\rm exp} = N_* \int\int \frac{d^2f}{d\ln P d\ln R_C} \times \eta(P, R_C)\, d\ln P d\ln R_C = N_* \sum_{\rm bins} \Delta f(\boldsymbol{\theta}) \times \eta(P, R_C) , \quad (12)$$

where $\eta(P, R_c)$ is the averaged detection completeness within the bin (§3). We choose a fine grid, with 20 period bins (1-16 days) and 30 core-radius bins ($0.6 - 2R_\oplus$).

The likelihood function is then

$$P(\{PCs\} \mid \boldsymbol{\theta}) = \exp(-N_{\rm exp}) \prod_{i=1}^{n_p} \Delta f(\boldsymbol{\theta}) , \quad (13)$$

where the product is carried over each detected candidate.

In detail, we first use the package *scipy.optimize* to obtain values that maximize the likelihood function. We then sample the posteriors using the MCMC algorithm provided by the *emcee* package (Foreman-Mackey et al. 2013). We use 50 walkers, 1000 "burn-in" steps and 5000 MCMC steps to find convergence, where the walkers of "burn-in" are initiated in a narrow Gaussian whose mean is the maximum likelihood solution. The parameters corresponding to the median of posterior distribution are shown in Table 2, with Fig. 4 illustrating the corresponding size distributions for our best models.

To compare the relative merits of various models, we employ the Akaike information criterion(AIC), a measure of information loss. This takes account of both the quality of the fit and the simplicity of the model to avoid both under-fitting and over-fitting. Let the latter be quantified by the number of free parameter,$n_{\boldsymbol{\theta}}$, we write

$$AIC = 2n_{\boldsymbol{\theta}} - 2\ln\left[P(\{PCs\} \mid \boldsymbol{\theta})\right] . \quad (14)$$

The best model has the lowest AIC score, and the relative merit (likelihood) between models is quantified as $e^{(-\Delta AIC/2)}$. So any model that has an AIC score higher by ~ 10 relative to the best model is clearly rejected.

4.3. *Results*

Here, we describe the results of our Bayesian inferences, and compare them against those from IDEM.

First, among the models we tested, the Bayesian analysis strongly rejects Model 1 (single population) and requires the presence of at least one other population of planets, in addition to the super-Earths. Model 1 is extremely unlikely, with a relative likelihood of $\exp(-20/2) \sim 10^{-5}$ compared to the 2-population models.

Second, the 2-period model (Model 4), while allowing for more flexibility in the period distributions, performs slightly worse than the models that insist on a single period distribution. In other words, there is no statistical evidence suggesting that the two populations (super-Earths and sub-Earths) have different period distributions, within the limited range that we probe (1 to 16 days). In fact, the obtained period distributions all feature a power-law rise to a few days, and a roughly logarithmically flat distribution afterwards, analogous to earlier studies using more extended samples (Silburt et al. 2015; Mulders et al. 2018; Neil & Rogers 2020).

Third, Models 2 and 3 are statistically indistinguishable (Model 3 is slightly preferred). Their size and period distributions also look similar (Fig. 5). This is because the Gaussian distribution of the sub-Earths, in the 2-Gaussian model, is very broad and is compatible

Parameter	1 Gaussian	2 Gaussian	Power-Gaussian	2 Period	Prior
A	$0.010^{+0.004}_{-0.003}$	-	-	-	$[0, 0.05]$
μ_0	$1.25^{+0.04}_{-0.04}$	-	-	-	$[0.6, 2]$
σ_0	$0.39^{+0.03}_{-0.03}$	-	-	-	$[0, 1]$
B	-	$0.0034^{+0.0018}_{-0.0012}$	$0.0032^{+0.0012}_{-0.0009}$	$0.0034^{+0.0015}_{-0.0013}$	$[0, 0.01]$
μ_1	-	$1.44^{+0.03}_{-0.03}$	$1.43^{+0.02}_{-0.02}$	$1.43^{+0.02}_{-0.02}$	$[1.3, 1.6]$
σ_1	-	$0.15^{+0.02}_{-0.02}$	$0.13^{+0.02}_{-0.02}$	$0.13^{+0.02}_{-0.02}$	$[0, 0.5]$
C	-	$0.00058^{+0.0027}_{-0.0025}$	-	-	$[0, 0.01]$
μ_2	-	$0.74^{+0.07}_{-0.08}$	-	-	$[0.6, 1.2]$
σ_2	-	$0.4^{+0.2}_{-0.2}$	-	-	$[0, 1]$
D	-	-	$0.0036^{+0.0013}_{-0.0010}$	$0.039^{+0.0018}_{-0.0014}$	$[0, 0.01]$
γ	-	-	$-1.2^{+0.5}_{-0.7}$	$-1.0^{+0.5}_{-0.6}$	$[-3, 1]$
α	$1.3^{+0.3}_{-0.3}$	$2.3^{+0.3}_{-0.3}$	$2.3^{+0.3}_{-0.3}$	$2.4^{+0.4}_{-0.4}$	$[1, 3]$
β	$-0.4^{+0.2}_{-0.2}$	$0.6^{+0.2}_{-0.2}$	$0.7^{+0.2}_{-0.2}$	$-0.0^{+0.4}_{-0.5}$	$[-1, 1]$
P_b	$4.4^{+0.7}_{-0.6}$	$4.4^{+0.7}_{-0.6}$	$4.4^{+0.7}_{-0.7}$	$4.5^{+0.9}_{-0.6}$	$[2, 10]$
α_I	-	-	-	$2.0^{+0.3}_{-0.2}$	$[1.5, 3]$
β_I	-	-	-	$0.2^{+0.2}_{-0.4}$	$[-0.5, 0.5]$
$P_{b,I}$	-	-	-	$8.4^{+1.1}_{-1.6}$	$[5, 10]$
AIC	2646	2606	2604	2607	

Table 2. The best fit value of parameters for each occurrence rate density model with $1-\sigma$ uncertainty and the corresponding AICs. Best models are Power-Gaussian as it has lowest AIC

with a power-law. From now, on, we choose to focus on Model 3, the so-called power-Gaussian model.

In Model 3, the super-Earths are centered around $1.4R_\oplus$ with a narrow width of $\sigma_1 \sim 0.13$. In planet mass, these correspond to a centroid of $\sim 4M_\oplus$ and a FWHM of a factor of 3.2. Given that we are focussing on host stars of lower masses (median mass $0.89M_\odot$), this centroid falls below that obtained by Wu (2019), $M_p \approx 8M_\oplus(M_*/M_\odot)$. The reason of the discrepancy is unclear for now. The FWHM does agrees with that in Wu (2019). The super-Earth population appear to peak narrowly in core mass, a mystery of the formation.

What is of most interest to us is the size distribution of the sub-Earths. The power-Gaussian model yields a size index $\gamma = -1.2^{+0.5}_{-0.7}$. Or, the number of sub-Earths most likely rises towards small sizes. This is in drastic contrast to the super-Earths who fall off steeply towards small sizes. In addition, not all objects bigger than Earth belongs to the super-Earth group. Even at the super-Earth peak ($\sim 1.4R_\oplus$), sub-Earths make a non-negligible contribution ($\sim 24\%$). Lastly, when we integrate all planets inward of 16 days and with core-sizes from 0.6 to $2R_\oplus$, we find that super-Earths and sub-Earths are comparable in numbers ($21 \pm 4\%$ vs. $28 \pm 4\%$).

To compare the parameter-free IDEM results against the Bayesian results, we first convert the planet observed radii to core radii, as detailed above. We then interpolate the IDEM results between bins, using a third-degree spline smoothing, to obtain the desired core-size distribution, $df/d\ln R_C$. The comparison is shown in Figure 5. They generally agrees with each other. And as expected, our treatment of uncertainties in the IDEM ap-

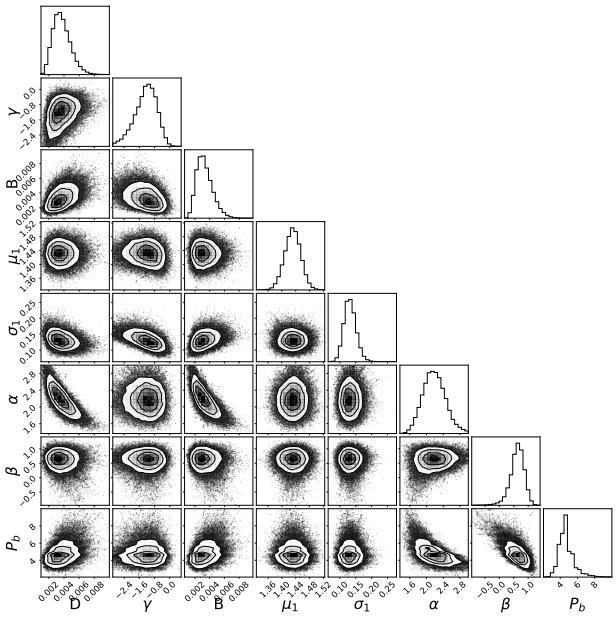

Figure 4. Corner-plot for the Power-Gaussian model, showing the co-variance between different parameters and their posterior distributions.

proach yield larger error-bars compared to those from forward models.

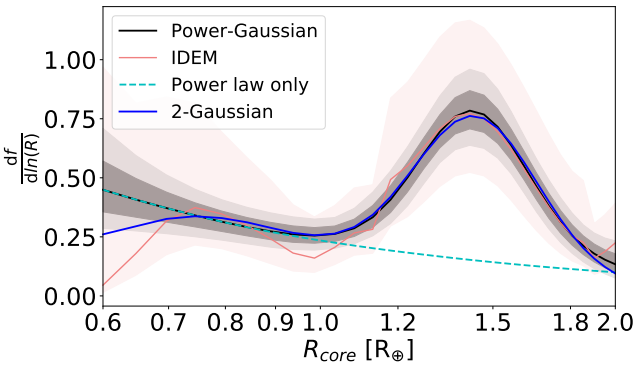

Figure 5. Comparison of results from both the Bayesian analysis (Power-Gaussian model in black line, dark-grey and light-grey zones indicating 1 and $2-\sigma$ uncertainties, respectively) and the parameter-free IDEM analysis (red curve, pink shadow zone being 95% confidence interval). The blue curve shows the best-fit 2-Gaussian model. All models agree on the fall of the super-Earths and the emergence of a separate population at smaller sizes. In addition, the broke cyan line shows the contribution of the power-law component (sub-Earths) in the Power-Gaussian model.

5. DISCUSSIONS

The decisive evidence for the new sub-Earth population is the radius gap located around $1R_\oplus$ (Fig. 5). From the peak of the super-Earths ($1.4R_\oplus$) to this gap, the occurrence rate of super-Earths drops markedly by a factor of a few; while the sub-Earth population gradually take over in dominance. The latter is only revealed thanks to the narrow mass distributions of the former. The mass distribution of the the sub-Earths appear much wider and can be described by a rising power-law towards small sizes.

Here, we discuss the robustness of these results and comparing them against previous works.

5.1. *Sample Cuts and Other variations*

Our analysis is performed on a small set of 'clean' planet sample (280 in total). Here, we vary the sample criteria and observe how the results for the power-gaussian model change. The priors remain the same as before. For each experiment, we remove (or add) one constraint on the property of the stars or PCS, keeping other cuts unchanged.

We also experiment with alternative priors and completeness treatment. All experiments yield similar results compared to our default case. We comment on each briefly below.

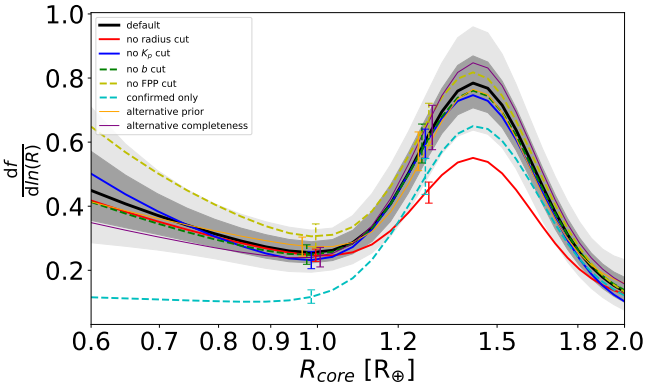

Figure 6. Comparing the core-radius distribution across models. The black curve and the shaded belts are as those in Fig. 5, while the colored curves from different experiments, and the error-bars marking their respective $1-\sigma$ uncertainties (only at two select radii for clarity). Regardless of sample selection or other variations, super-Earths are found to have a narrow size distribution. With the exception of the 'confirmed' cut, a new component (sub-Earth) that rises towards small sizes is required in all models.

- No stellar radius cut: we employ all stars, regardless of radii. There are then 90427 stars and 893 PCs. The stellar sample has expanded by a factor of 6, while the planet sample only a factor of 3. This explains the overall lower occurrence in Fig. 3. Moreover, the super-Earth peak is broadened by the correlation between planet mass and stellar mass.

- No K_p cut: while dimmer stars return fewer super-Earths and sub-Earths, the completeness correction largely accommodates this correctly. There is no perceptible change to the occurrence rates.

- No b cut: retaining candidates with any impact-parameters, not just those with $b < 0.9$ as in our default case, produces no perceptible changes.

- No FPP cut: retaining all objects, regardless of their FPP values, adds 36 new candidates to our list, but produces no change to the results. The newly added candidates do not make the gap features more convincing, presumably because they are more likely to be potential false positives.

- Confirmed only: rejecting all planets that are currently un-confirmed removes a majority of the sub-Earths (who typically have lower signal-to-noise ratios) from our sample. This cut is the only one that produces a substantial change to the sub-Earth distribution – the small planet population are much reduced and flattened. However, we believe that unconfirmed PCs with a low FPP value

are reliable enough and the sub-Earth distribution in our default case is likely genuine.

- Alternative Priors: We have chosen broad and flat priors previously. Here, we follow the insights from Wu (2019) and use a narrower prior on the super-Earths: μ_1 is now normally distributed with a mean at $1.43R_\oplus$ and a variance of $0.05R_\oplus$, and σ_1 is normally distributed with a mean of 0.1 and a variance of 0.1. The results remain the same.

- Alternative completeness: while we have adopted the Burke & Catanzarite (2017) to to obtain the detection completeness, there exists an alternative approach (Christiansen et al. 2020). In this method, the completeness of a detection is expressed as a function of its MES (multiple-event-statistics) value, as well as some parameters a, b, c that are empirically obtained from pixel-level injection experiments,

$$\eta(MES, a, b, c) = \frac{c}{b^a \Gamma(a)} \int_0^{MES} t^{a-1} e^{-\frac{t}{b}} dt\,. \tag{15}$$

Because pixel-level injections are computationally expensive, typically one injection per star is possible. These empirical parameters are therefore best fits for the entire Kepler stellar sample. This type of correction is used by e.g., Neil & Rogers (2020); Hsu et al. (2019). Here, we employ KeplerPORTS to map radius and period into MES for every planet host in our sample, and then adopt $a = 29.41$, $b = 0.284$, $c = 0.891$ from Neil & Rogers (2020). Running the Bayesian analysis return results that do not differ significantly from our original ones (see Fig. 6).

5.2. *Comparison with Other works*

There are two recent works that employ the updated completeness corrections and have included sub-Earths into their modelling efforts (Hsu et al. 2019; Neil & Rogers 2020). We compare our results with theirs here, for the period range from 1 to 16 days.. The Neil & Rogers (2020) work infers distributions in planet masses. We use simple scaling relations of $M/M_\oplus = (R_C/R_\oplus)^4$ for $R_C > 1.2R_\oplus$ and $M/M_\oplus = (R_C/R_\oplus)^3$ for $R_C <= 1.2R_\oplus$ to translate their mass distribution into core radius distribution. The other study (Hsu et al. 2019) infers the distribution in planet transit radius. We convert from these to core-radius as laid out in §4.2 and interpolate between bins where necessary. The comparisons are displayed in Fig. 7 and we briefly describe any similarities or differences here.

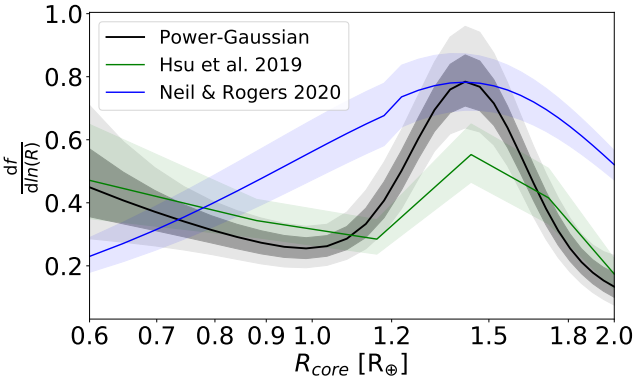

Figure 7. Comparison of our core-radius distribution (same model as that in Fig. 5) and those from Hsu et al. (2019) (green) and Neil & Rogers (2020) (blue), for planets with periods from 1 to 16 days. The color shaded regions indicate their respective 1-σ uncertainties. Three studies differ in sample selection, completeness correction and inversion method.

Similar to our study here, both works show a droping-off of super-Earths towards small sizes, and a new population below about $1R_\oplus$.

In the parameter-free, grid-based inversion model of Hsu et al. (2019), one observes a radius gap between super-Earths (including Neptunes) and sub-Earths. However, because their stellar sample covers a wide range of spectral types (mostly FGK dwarfs with no cuts in stellar radius), their super-Earth peak is broader and lower in amplitude, resembling closely our 'no-radius-cut' case in Fig. 6. By focussing on a narrow range of spectral type, we end up with a smaller planet sample, but can detect more prominent features.

The forward-modelling approach of Neil & Rogers (2020) is very similar in spirit to our Bayesian models. By parametrizing the size distributions as a number of distinct Gaussians, they find that a new population of small planets (which they term 'intrinsically rocky') is strongly demanded by data. However, their overall size distribution does not show a clear 'gap'. We can compare their results with our 2-Gaussian models in detail. Their best-fit super-Earth peak lies at $R_C = 1.53R_\oplus$ (compared to our $1.44R_\oplus$) and has a much broader dispersion of $\sigma = 0.63$ (compared to our 0.15). We attribute both these differences to their wide range of stellar properties (they include almost all dwarfs). Since super-Earth properties are known to correlate with stellar properties, this tends to broaden the super-Earth peak. Moreover, their sub-Earth can be described by a log-normal distribution with a peak at size $R_C \approx 0.96R_\oplus$ (compare to our 0.74), and a radius dispersion of $\sigma \sim 0.4$ (compare to our 0.4). The fact that their sub-Earths are more massive than ours gives rise

to a gap-free transition from the super-Earths to the sub-Earths (Fig. 7), as well as a dropping occurrence towards very small sizes. These are qualitatively different from our results and those in Hsu et al. (2019). A recent investigation, using their original model but with more relaxed assumptions, leads to better agreements with our results (Neil, private communication).

5.3. *Insights from Ultra-Short-Period Planets*

An exotic class of planets, called ultra-short-period planets (USP for short), have orbital periods short-ward of one day (e.g. Rappaport et al. 2013; Sanchis-Ojeda et al. 2014; Winn et al. 2018). At these short periods, *Kepler* can be sensitive to much smaller planets than otherwise possible. The origin for these boiling worlds is still under debate (e.g. Lee & Chiang 2017; Petrovich et al. 2019; Pu & Lai 2019; Millholland & Spalding 2020), though scenario involving dynamical migration by companion planets have received some observational supports (Sanchis-Ojeda et al. 2014; Adams et al. 2020). If so, their mass distribution should be similar to those further-away. Here, we exploit this advantage to shed light on the mass model.[2]

Fig. 8 compares the observed size distribution of USPs (< 1 day), against that of our default sample ($1 - 16$ days), and against our preferred model. These 127 USPs are discovered by the *Kepler* main mission, are smaller than $2R_\oplus$, include both "confirmed" and "candidate" dispositions. Their host magnitudes are mostly between 14 and 16 mags. We have also updated their radii using GAIA stellar parameters (Berger et al. 2020b).

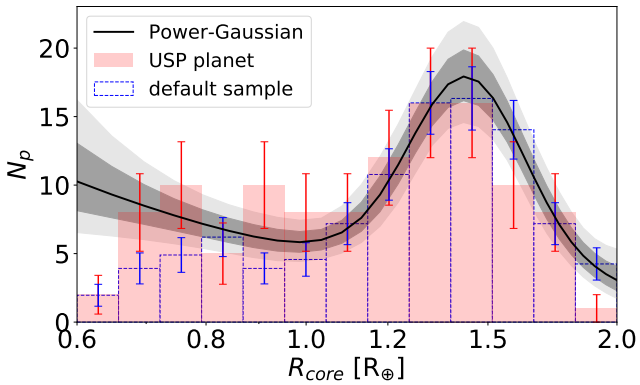

Figure 8. Comparing the observed size distributions of ultra-short-period planets (< 1 day) (red histogram) with our default sample (blue histogram, $1 - 16$ days), and our power-Gaussian model (black curve). The latter two are scaled to have the same heights as USPs at $1.3R_\oplus$. The USPs show an excess of sub-Earths than the longer period sample, due to better completeness. They therefore provide support to our model.

[2] There is another advantage to using USPs. Ultra-short-period candidates are less likely to be false positives due to background eclipsing binaries. At such short periods, ellipsoidal variations would have been clearly visible.

Fig. 8 shows that USPs show an excess of sub-Earths, over our default sample, due to their better completeness in these small sizes. More importantly, one can see near perfect agreements between the un-corrected USPs and our power-Gaussian model. Completeness correction may enhance the rise of sub-Earths, possibly even over that of our current model.

Other than supporting our model, this result also suggests that USPs are indeed formed similarly as planets further away.

5.4. *Siblings of the sub-Earths*

In our sample, there are 71 PCs with radii smaller than $1.1R_\oplus$. These most likely belong to the sub-Earth population, according to our Power-Gaussian model. Among them, 33 sub-Earths have at least one (though frequently more than one) transiting sibling planet. These multiple-planet systems are shown in Figure 9. We see that sub-Earths appear to happily co-exist with larger planets (super-Earths, mini-Neptunes, even Jovians), with no substantial differences in both their period distribution and their stellar mass distribution.

The fact that about half of all sub-Earths have transiting siblings may indicate that the majority of them prefer to reside in systems with other larger planets.

6. CONCLUSION

The primary goal of the *Kepler* mission is to find Earth analogs in the habitable zones (Koch et al. 2010). It has had some successes in finding rocky planets around low mass stars where the habitable zones are close-in (see e.g., Borucki et al. 2013; Kane et al. 2016), but there are few solid such detections around Sun-like stars (see Fig. 3 in Borucki et al. 2019). While *Kepler* did find an abundance of super-Earths, these are likely stripped-down version of the mini-Neptunes and do not exist beyond a few tens of days.

In this study, by choosing a mostly complete planet sample, by assuming simple models with broad priors, we are able to establish that the core-radii of super-Earths peak narrowly around $1.4R_\oplus$ (corresponding to $\sim 4M_\oplus$ for terrestrial composition). At smaller sizes, we find that these planets disappear rapidly in number, with the data showing a clear gap between super-Earths and a new population of small planets (sub-Earths). The existence of a gap is robust, regardless of sample selection, inversion method, or completeness correction. The size distribution of ultra-short-period planets provides additional supports to our model.

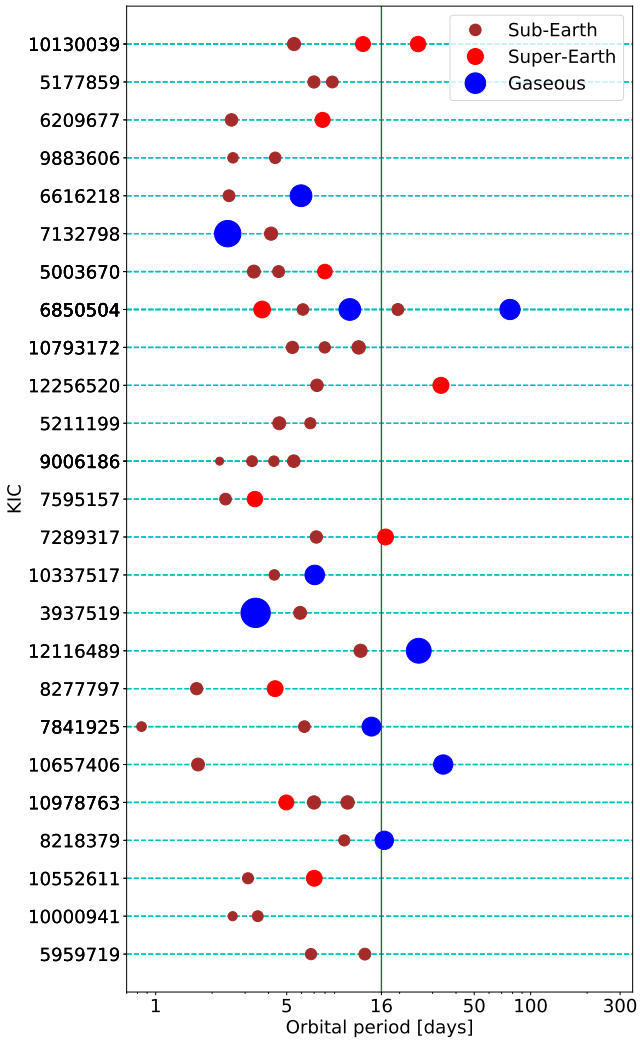

Figure 9. About half sub-Earths in our default sample have transiting planet siblings. They are plotted here according to their sizes. The systems are arranged from top to bottom by decreasing host star masses.

We find that the sub-Earths, dominant in number below $\sim 1R_\oplus$, have a similar period distribution as the super-Earths, co-exist with larger planets in the same planetary systems, and, most interestingly, can be described by a rising power-law towards small sizes. This is in contrast to the nearly singular distribution of the super-Earths.

This implies that sub-Earths and super-Earths are formed differently. While super-Earths (and their cousins that still retain the primordial H/He envelopes, mini-Neptunes) are most likely formed in the gaseous proto-planetary disks (and therefore may be called Generation-I planets), the sub-Earths, we speculate, may form later in a collisional debris disk, much like the terrestrial planets are theorized to do (e.g. Chambers & Wetherill 1998; Kokubo & Ida 1998; Morishima et al. 2008). In such a debris disk, the final planetary outcome depends on the initial mass budget. One therefore expects the planetary masses to be broadly distributed. We argue that these planets may be thought of as Generation-II planets.

If our speculation, that these planets are true terrestrial analogs, is correct, we expect them to extend to at least the AU region. They are then likely the only population of rocky planets in the habitable zones, and they are the most likely habourer of life. While *Kepler* just fell short of discovering them in AU distances, more sensitive new missions should be organized to go after these targets.

To provide a rough guidance for these future searches, we boldly estimate the occurrence rate of sub-Earths in the habitable zone (Table 3), extrapolating from models that apply only within 16 days. The estimates show substantial uncertainties, largely because of our short lever arm in constraining the period distribution. Even an extension by a factor of 2 in period would have helped firming down the $\eta_\oplus$ value substantially.

Models	f (1-16 days)	f (200-400 days, $\eta_\oplus$)
Power-Gaussian	$28^{+4}_{-4}\%$	$124^{+128}_{-69}\%$
2-Gaussian	$22^{+5}_{-5}\%$	$90^{+95}_{-50}\%$
2-period	$25^{+5}_{-4}\%$	$9^{+40}_{-8}\%$

Table 3. Estimated occurrences of sub-Earths ($0.6 - 2R_\oplus$) around GK dwarfs.

ACKNOWLEDGMENTS

This work makes use of the NASA Exoplanet archive, and results from the NASA Kepler mission, the ESA Gaia mission and the California-Kepler survey. Danley Hsu, Andrew Neil and Eric Ford provided helpful feedbacks. We acknowledge NSERC for a research grant, including support for QYS as a summer undergraduate researcher.

REFERENCES

Adams, E. R., Jackson, B., Johnson, S., et al. 2020, arXiv e-prints, arXiv:2011.11698. https://arxiv.org/abs/2011.11698

Berger, T. A., Huber, D., Gaidos, E., van Saders, J. L., & Weiss, L. M. 2020a, AJ, 160, 108, doi: 10.3847/1538-3881/aba18a

Berger, T. A., Huber, D., van Saders, J. L., et al. 2020b, AJ, 159, 280, doi: 10.3847/1538-3881/159/6/280
Borucki, W., Thompson, S. E., Agol, E., & Hedges, C. 2019, arXiv e-prints, arXiv:1905.05719. https://arxiv.org/abs/1905.05719
Borucki, W. J., Agol, E., Fressin, F., et al. 2013, Science, 340, 587, doi: 10.1126/science.1234702
Bryson, S., Coughlin, J., Batalha, N. M., et al. 2020, AJ, 159, 279, doi: 10.3847/1538-3881/ab8a30
Burke, C. J., & Catanzarite, J. 2017, Planet Detection Metrics: Per-Target Flux-Level Transit Injection Tests of TPS for Data Release 25, Kepler Science Document KSCI-19109-002
Burke, C. J., Christiansen, J. L., Mullally, F., et al. 2015, ApJ, 809, 8, doi: 10.1088/0004-637X/809/1/8
Chambers, J. E., & Wetherill, G. W. 1998, Icarus, 136, 304, doi: 10.1006/icar.1998.6007
Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. 2012, PASP, 124, 1279, doi: 10.1086/668847
Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2016, ApJ, 828, 99, doi: 10.3847/0004-637X/828/2/99
—. 2020, AJ, 160, 159, doi: 10.3847/1538-3881/abab0b
Cloutier, R., & Menou, K. 2020, AJ, 159, 211, doi: 10.3847/1538-3881/ab8237
Coughlin, J. L. 2017, Planet Detection Metrics: Robovetter Completeness and Effectiveness for Data Release 25, Kepler Science Document KSCI-19114-002
Feldman, G. J., & Cousins, R. D. 1998, PhRvD, 57, 3873, doi: 10.1103/PhysRevD.57.3873
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067
Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, ApJ, 659, 1661, doi: 10.1086/512120
Fulton, B. J., & Petigura, E. A. 2018, AJ, 156, 264, doi: 10.3847/1538-3881/aae828
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109, doi: 10.3847/1538-3881/aa80eb
Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, MNRAS, 476, 759, doi: 10.1093/mnras/sty290
Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, ApJS, 201, 15, doi: 10.1088/0067-0049/201/2/15
Hsu, D. C., Ford, E. B., Ragozzine, D., & Ashby, K. 2019, AJ, 158, 109, doi: 10.3847/1538-3881/ab31ab
Jenkins, J. M. 2002, ApJ, 575, 493, doi: 10.1086/341136
Jin, S., Mordasini, C., Parmentier, V., et al. 2014, ApJ, 795, 65, doi: 10.1088/0004-637X/795/1/65
Kane, S. R., Hill, M. L., Kasting, J. F., et al. 2016, ApJ, 830, 1, doi: 10.3847/0004-637X/830/1/1
Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, ApJL, 713, L79, doi: 10.1088/2041-8205/713/2/L79
Kokubo, E., & Ida, S. 1998, Icarus, 131, 171, doi: 10.1006/icar.1997.5840
Kunimoto, M., & Matthews, J. M. 2020, AJ, 159, 248, doi: 10.3847/1538-3881/ab88b0
Lee, E. J., & Chiang, E. 2017, ApJ, 842, 40, doi: 10.3847/1538-4357/aa6fb3
Lopez, E. D., Fortney, J. J., & Miller, N. 2012, ApJ, 761, 59, doi: 10.1088/0004-637X/761/1/59
Mathur, S., & Huber, D. 2016, Kepler Stellar Properties Catalog Update for Q1-Q17 DR25 Transit Search, Kepler Science Document KSCI-19097-004
Millholland, S., & Spalding, C. 2020, arXiv e-prints, arXiv:2010.15087. https://arxiv.org/abs/2010.15087
Morishima, R., Schmidt, M. W., Stadel, J., & Moore, B. 2008, ApJ, 685, 1247, doi: 10.1086/590948
Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, ApJ, 822, 86, doi: 10.3847/0004-637X/822/2/86
Mulders, G. D., Pascucci, I., Apai, D., & Ciesla, F. J. 2018, AJ, 156, 24, doi: 10.3847/1538-3881/aac5ea
Neil, A. R., & Rogers, L. A. 2020, ApJ, 891, 12, doi: 10.3847/1538-4357/ab6a92
Owen, J. E., & Wu, Y. 2013, ApJ, 775, 105, doi: 10.1088/0004-637X/775/2/105
—. 2017, ApJ, 847, 29, doi: 10.3847/1538-4357/aa890a
Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, Proceedings of the National Academy of Science, 110, 19273, doi: 10.1073/pnas.1319909110
Petrovich, C., Deibert, E., & Wu, Y. 2019, AJ, 157, 180, doi: 10.3847/1538-3881/ab0e0a
Pu, B., & Lai, D. 2019, MNRAS, 488, 3568, doi: 10.1093/mnras/stz1817
Rappaport, S., Sanchis-Ojeda, R., Rogers, L. A., Levine, A., & Winn, J. N. 2013, ApJL, 773, L15, doi: 10.1088/2041-8205/773/1/L15
Sanchis-Ojeda, R., Rappaport, S., Winn, J. N., et al. 2014, ApJ, 787, 47, doi: 10.1088/0004-637X/787/1/47
Silburt, A., Gaidos, E., & Wu, Y. 2015, ApJ, 799, 180, doi: 10.1088/0004-637X/799/2/180
Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, ApJS, 235, 38, doi: 10.3847/1538-4365/aab4f9
Twicken, J. D., Jenkins, J. M., Seader, S. E., et al. 2016, AJ, 152, 158, doi: 10.3847/0004-6256/152/6/158
Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, MNRAS, 479, 4786, doi: 10.1093/mnras/sty1783
Winn, J. N., Sanchis-Ojeda, R., & Rappaport, S. 2018, NewAR, 83, 37, doi: 10.1016/j.newar.2019.03.006
Wu, Y. 2019, ApJ, 874, 91, doi: 10.3847/1538-4357/ab06f8
Youdin, A. N. 2011, ApJ, 742, 38, doi: 10.1088/0004-637X/742/1/38