# The Architecture of Multiplanet Systems as a Tracer of Their Formation Mechanisms 

Udit Arora ${ }^{1}$ © and Yasuhiro Hasegawa ${ }^{2}$<br>${ }^{1}$ Indraprastha Institute of Information Technology, Okhla Industrial Estate, Phase III, New Delhi, 110020, India; udit18417@iiitd.ac.in<br>${ }^{2}$ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA; yasuhiro.hasegawa@jpl.nasa.gov<br>Received 2021 April 8; revised 2021 June 8; accepted 2021 June 10; published 2021 July 5


#### Abstract

Exoplanets observed by the Kepler telescope exhibit a bimodal, radius distribution, which is known as the radius gap. We explore an origin of the radius gap, focusing on multiplanet systems. Our simple theoretical argument predicts that type I planetary migration produces different configurations of protoplanets with different masses and such different configurations can result in two distinguishable populations of small-sized multiplanet systems. We then perform an observational analysis to verify this prediction. In the analysis, multiple Kolmogorov-Smirnov tests are applied to the observed systems, using the statistical measures that are devised to systematically characterize the properties of multiplanet systems. We find with $99.5 \%$ confidence that the observed, small-sized multiplanet systems are divided into two distinct populations. The distinction likely originates from different spatial distributions of protoplanets, which are determined by type I migration and subsequently trigger giant impact. We also show that these distinct populations are separated around the radius gap when the gas surface density of protoplanetary disks is $\sim 10^{2} \mathrm{~g} \mathrm{~cm}^{-2}$ in the vicinity of the host stars. This work therefore emphasizes the importance of planetary migration and the inner disk properties.


Unified Astronomy Thesaurus concepts: Planet formation (1241); Planetary migration (2206); Super Earths (1655); Mini Neptunes (1063); Protoplanetary disks (1300); Exoplanet dynamics (490)

## 1. Introduction

The rapid increase of observed exoplanets revolutionizes our understanding of planet formation (e.g., Winn \& Fabrycky 2015 for a review), which has been accelerated thanks to the Kepler mission (Borucki et al. 2010). One famous, astonishing result is that small-sized planets whose radii are larger than $1 R_{\oplus}$ and smaller than $4 R_{\oplus}$ are dominant in the observed population (e.g., Howard et al. 2010). This is consistent with the pioneering discovery of radial velocity observations (Mayor et al. 2011).
The ubiquity of observed small-sized planets has challenged the canonical picture of planet formation (e.g., Ida \& Lin 2004). This is because there is no such analog in the solar system, and hence both the observational characterization of these planets and the theoretical development of possible formation mechanisms have been required. For the former, the follow-up observations of exoplanet-host stars reveal that the radius range of small-sized planets is divided into two regimes (Fulton et al. 2017): $1 R_{\oplus} \lesssim R_{p} \lesssim 1.8 R_{\oplus}$ and $1.8 R_{\oplus} \lesssim R_{p} \lesssim 4 R_{\oplus}$. Based on the bulk density (e.g., Weiss \& Marcy 2014; Rogers 2015), planets in the first regime can literally be called a "super-Earth," and those in the second regime may be named a "sub-Neptune." The presence of these bimodal populations has been popularized as the "radius gap (or valley)," and triggered a number of theoretical studies. Recent studies suggest that both the formation and evolution processes can generate the gap (e.g., Owen \& Wu 2013; Ginzburg et al. 2018).

The multiplicity is another important feature of observed small-sized planets (e.g., Lissauer et al. 2011), which also sheds light on their formation mechanisms (e.g., Hansen \& Murray 2013; Dawson et al. 2016). Intriguingly, Weiss et al. (2018) confirm the presence of the radius gap, focusing only on exoplanets in multiplanet systems. This finding may suggest that physical processes inevitable for shaping multiplanet systems would contribute to the generation of the radius gap.

Here, we show through a simple theoretical argument and an observational analysis that planetary migration produces different configurations of protoplanets with different masses and that this difference and the subsequent giant impact lead to two distinguishable populations of small-sized multiplanet systems. It is prominent that the dividing radius of these two populations broadly corresponds to the radius gap when the gas surface density of the natal protoplanetary disk has a low value $\left(\sim 10^{2} \mathrm{~g} \mathrm{~cm}^{-2}\right)$ in the vicinity of the central star. Thus, this work points out the importance of planetary migration and the inner disk properties to better understand the observed properties of super-Earths and sub-Neptunes.

## 2. Theoretical and Observational Analyses

### 2.1. Exoplanet Data

We first introduce the exoplanetary data that will be used in the following analyses; in these analyses, the mass $\left(M_{p}\right)$, radius $\left(R_{p}\right)$, and semimajor axis $\left(a_{p}\right)$ of planets and the mass $\left(M_{s}\right)$ of the host stars are needed.

We obtain the Kepler data from the NASA Exoplanet Archive (NASA Exoplanet Science Institute 2019). ${ }^{3}$ We focus only on multiplanetary systems orbiting around single stars, which are comprised of both confirmed and candidate planets. We filter out the systems that do not have certain quantities and hence our analyses cannot be applied to them. These are the surface gravity, effective temperature, and radius of the host stars, and the radius and orbital period of planets. Note that the stellar parameters are needed both to identify whether the host stars are the main-sequence ones, and to compute their masses. For the former, we follow the approach of Mulders et al. (2018), where dwarf stars are identified, based on the relationship

[^0]

Figure 1. The mass-radius relation for exoplanets in our data set. The cyan dots with the error bars represent planets in the Terran world (i.e., $M \lesssim 2 M_{\oplus}$ ), and the light green dots with the error bars represent planets in the Neptunian worlds (i.e., $2 M_{\oplus} \lesssim M \lesssim 100 M_{\oplus}$ ). The corresponding fitting profiles are denoted by the blue dashed and the green solid lines, respectively.
between the surface gravity and effective temperature (Huber et al. 2016). For the latter, we use the mass-luminosity and the mass-radius relations (Eker et al. 2018). ${ }^{4}$

We select planets whose radius is in the range of $0.1 R_{\oplus} \leqslant R_{p} \leqslant 30 R_{\oplus}$. Following Thompson et al. (2018), we adopt the disposition score of $>0.6$, where the trade-off between reliability and completeness is achieved properly (see the top panel of their Figure 13). We have confirmed that our results do not change very much even if planets with the disposition score of $>0.5$ are chosen. The semimajor axis of planets is computed, using Kepler's third law.

In summary, we have a total of 348 planetary systems with 870 exoplanets observed by Kepler. ${ }^{5}$

### 2.2. Mass-Radius Relation

The planet mass is one fundamental quantity for our analyses and is computed from the mass-radius relation.

We utilize the pretrained model developed by Chen \& Kipping (2017), where an unbiased forecasting model was built upon for a probabilistic mass-radius relation with the Bayesian framework. They found that the mass-radius relation is divided into two regimes for small-sized planets: $M_{p} \lesssim 2 M_{\oplus}$ and $2 M_{\oplus} \lesssim M_{p} \lesssim 100 M_{\oplus}$, and the corresponding regimes are referred to as the Terran worlds and the Neptunian worlds, respectively.

Figure 1 shows the computed masses for given radii. We subsequently fit these data and confirm that the resulting powerlaw indices for the two types of planets are similar to those of Chen \& Kipping (2017). It is interesting that the Neptunian worlds reside in the radius range of $1.2 R_{\oplus} \lesssim R_{p} \lesssim 10 R_{\oplus}$, and therefore they cover both (massive) super-Earths and the entire sub-Neptunes.

[^1]
### 2.3. Theoretical Prediction: The Importance of Planetary Migration

The formation of small-sized, multiplanet systems is currently the target of active research (e.g., Ida \& Lin 2010; Chiang \& Laughlin 2013; Hansen \& Murray 2013; Izidoro et al. 2017), and it is still far from being completely understood.

Our theoretical argument assumes that type I migration and the subsequent giant impact play an important role in forming small-sized, multiplanet systems; type I migration takes place due to the disk-planet interaction and becomes effective for low-mass planets ( $M_{p} \gtrsim 1 M_{\oplus}$; e.g., Kley \& Nelson 2012). Giant impact is the collision between protoplanets and serves as the final mass assembly for terrestrial planets in the solar system (e.g., Chambers 2001, hereafter C01).

Under this assumption, the importance of planetary migration to trigger the subsequent giant impact between neighboring protoplanets may be determined by the difference between the mutual spacing ( $\Delta_{\text {mig }}$ ) achieved by type I migration and that ( $\Delta_{\text {esc }}$ ) by the escape velocity; $\Delta_{\text {esc }}$ is the result of the pure gravitational interaction between the neighboring protoplanets and is the largest separation between them before they undergo giant impact (e.g., Schlichting 2014):

$$
\begin{equation*}
\Delta_{\mathrm{esc}} \simeq 2 a_{p} \frac{v_{\mathrm{esc}}}{v_{\mathrm{Kep}}} \tag{1}
\end{equation*}
$$

where $v_{\text {Kep }}$ is the Keplerian velocity, and $v_{\text {esc }}=\sqrt{2 G M_{p} / R_{p}}$ is the escape velocity for the neighboring protoplanets with comparable masses and radii.

Planetary migration affects the onset of the subsequent giant impact considerably if the following condition is met:

$$
\begin{equation*}
\Delta_{\mathrm{mig}}<\Delta_{\mathrm{esc}} \tag{2}
\end{equation*}
$$

Mathematically, $\Delta_{\text {mig }}$ can be computed from the consideration that the speed ( $v_{\text {mig }}$ ) of (differential) type I migration, which reduces the mutual spacing $\left(b=\left|a_{p, i}-a_{p, j}\right|\right)$ between the neighboring protoplanets $(i, j)$, is compensated for by the gravitational repulsion between them (Ida \& Lin 2010). For $v_{\text {mig }}$, it is written as (e.g., Hasegawa \& Pudritz 2011)

$$
\begin{equation*}
v_{\mathrm{mig}} \simeq 2 K_{\mathrm{mig}} \frac{\Sigma_{\mathrm{g}} a_{p}^{2} M_{p}}{M_{s}^{2}}\left(\frac{H}{a_{p}}\right)^{-2} v_{\mathrm{Kep}} \tag{3}
\end{equation*}
$$

where $\Sigma_{\mathrm{g}}$ and $H$ are the gas surface density and the pressure scale height of the natal protoplanetary disk, respectively, and $K_{\text {mig }}$ is a coefficient that accounts for the differential speed and the detail of disk properties. For simplicity, we adopt here that $H / a_{p}=0.02, K_{\mathrm{mig}}=1$, and $M_{s}=1 M_{\odot}$. For the gravitational repulsion, the expansion $(\delta b)$ of the mutual spacing is given by linear theory as (Hasegawa \& Nakazawa 1990)

$$
\begin{equation*}
\delta b \simeq 30\left(\frac{b}{r_{\mathrm{H}}}\right)^{-5} r_{\mathrm{H}} \tag{4}
\end{equation*}
$$

where $r_{\mathrm{H}}=a\left(2 M_{p} / 3 M_{s}\right)^{1 / 3}$ is the mutual Hill radius for the neighboring protoplanets with comparable masses. Since the expansion is driven by encounters that occur at every synodic period $\left(T_{\text {syn }}\right)$, the expansion rate becomes

$$
\begin{equation*}
\frac{d b}{d t} \simeq \frac{\delta b}{T_{\mathrm{syn}}} \tag{5}
\end{equation*}
$$



Figure 2. The mutual spacing as a function of planet mass for the case where $\Sigma_{\mathrm{g}}=5 \times 10^{2} \mathrm{~g} \mathrm{~cm}^{-2}$ at $a_{p}=0.1 \mathrm{au}$. Planetary migration achieves narrower mutual spacings for the mass range of $1 M_{\oplus} \lesssim M_{p} \lesssim 10 M_{\oplus}$ (see the red shaded region). The mass-radius relation is used to plot $\Delta_{\text {esc }}$ (see the black dashed line). For $\Delta_{\text {mig }}$, the solid line denotes the mass regime where type I migration becomes effective (see the magenta line). The background cyan and light green regions represent the Terran and Neptunian worlds, respectively.
where $T_{\text {syn }} \simeq 4 \pi a_{p}^{2} / 3 b v_{\text {Kep }}$. In summary, $\Delta_{\text {mig }}$ can be computed, by equating $v_{\text {mig }}$ with $\delta b / T_{\text {syn }}$.

Figure 2 shows the resulting behaviors of $\Delta_{\text {esc }}$ and $\Delta_{\text {mig }}$ as a function of planet mass. In this plot, the mass-radius relation obtained in Figure 1 is used to compute $\Delta_{\text {esc }}$. As an example, we adopt the value of $\Sigma_{\mathrm{g}}=5 \times 10^{2} \mathrm{~g} \mathrm{~cm}^{-2}$ at $a_{p}=0.1$ au to calculate $\Delta_{\text {mig }}$. In addition, the mass regime where type I migration becomes effective is denoted by the solid line of $\Delta_{\text {mig }}$, following Hasegawa \& Pudritz (2012).

We find that $\Delta_{\text {mig }}>\Delta_{\text {esc }}$ for the planet mass of $M_{p} \lesssim 1 M_{\oplus}$, $\Delta_{\text {mig }}<\Delta_{\text {esc }}$ for $1 M_{\oplus} \lesssim M_{p} \lesssim 10 M_{\oplus}$, and $\Delta_{\text {mig }} \simeq \Delta_{\text {esc }}$ for $M_{p} \gtrsim 10 M_{\oplus}$ in this particular setup. Note that the mutual spacing of $20 r_{\mathrm{H}}$ corresponds to the period ratio of $\sim 1.5$, which is consistent with the peak value of the observed exoplanets (e.g., Choksi \& Chiang 2020). Our value is, however, larger than what the typical numerical simulations predict (i.e., $\sim$ a few $r_{H}$; e.g., Izidoro et al. 2017). This difference comes from a lower value of $\Sigma_{\mathrm{g}}$ (see Section 3 for more discussion).

Given that the mutual spacing is one key parameter for giant impact, it is natural to consider that protoplanets with $\Delta_{\text {mig }} \lesssim \Delta_{\text {esc }}$ undergo giant impact more efficiently/rapidly than those in $\Delta_{\text {mig }}>\Delta_{\text {esc }}$. This consideration leads to the expectation that (proto)planets in the red shaded region may surely experience giant impact and obtain higher masses. Note that the size of the region depends on the power-law index of the mass-radius relation; assuming that $R_{p} \propto M_{p}^{\beta}, \Delta_{\mathrm{mig}} \propto M^{-1 / 12}$ and $\Delta_{\text {esc }} \propto M^{(1 / 3-\beta) / 2}$ in the unit of $r_{\mathrm{H}}$. The region expands if $\beta>1 / 3$ and shrinks if $\beta<1 / 3$ in the Neptunian worlds, compared with Figure 2. However, the region surely exists if $\Sigma_{\mathrm{g}}$ is larger than a threshold value (see Section 3).

In the following, we conduct the observational analysis to verify our theoretical prediction.

### 2.4. Statistical Measures

The direct observables (e.g., the period ratio of neighboring planets) can be used to characterize multiplanet systems (e.g., Weiss et al. 2018; Zhu et al. 2018). However, as the number of planets in the systems increases, the complexity in systematically characterizing multiplanet systems as a whole increase (e.g., Gilbert \& Fabrycky 2020).

In order to resolve this issue, we adopt the so-called "statistical measures" that are devised by C01. We focus here on two quantities that are computed directly from $M_{p}, a_{p}$, and $M_{s}$.

The first quantity is the mass concentration $\left(S_{c}\right)$ that measures the degree to which the planet mass is concentrated in a certain location of the system and is calculated as

$$
\begin{equation*}
S_{c}=\max \left(\frac{\sum_{j} M_{p, j}}{\sum_{j} M_{p, j}\left[\log _{10}\left(a / a_{p, j}\right)\right]^{2}}\right) \tag{6}
\end{equation*}
$$

where the summation is done for the index $j$. By changing $a$, the maximum value is searched. Based on $N$-body simulations, the value of $S_{c}$ reflects the initial spatial distribution of protoplanets that undergo giant impact eventually (Hansen 2009, hereafter H09; Hansen \& Murray 2012, hereafter HM12).

The second quantity is the orbital spacing $\left(S_{s}\right)$ that is somewhat similar to the averaged mutual spacing normalized by the mutual Hill radius. The main difference is that $S_{s}$ is normalized by $M_{p}^{1 / 4}$ (not $M_{p}^{1 / 3}$ ) and is motivated by the results of $N$-body simulations that explore the stability of multiplanet systems (Chambers et al. 1996):

$$
\begin{equation*}
S_{s}=\frac{6}{N-1}\left(\frac{a_{p, \max }-a_{p, \min }}{a_{p, \max }+a_{p, \min }}\right)\left(\frac{3 M_{s}}{2 \bar{M}_{p}}\right)^{1 / 4} \tag{7}
\end{equation*}
$$

where $N, a_{p, \max }, a_{p, \min }$, and $\bar{M}_{p}$ are the number, the maximum and the minimum values of the semimajor axis, and the mean mass of the planets in a system, respectively.

### 2.5. Observational Analysis: KS Tests

We now perform the Kolmogorov-Smirnov (KS) test for the cumulative distributions of $S_{c}$ and $S_{s}$ to determine whether the population of small-sized, multiplanet systems can be divided into subgroups.

The KS test is a nonparametric test and is used to discover whether two samples are drawn from the same distribution, based on the null hypothesis. The significance level (also referred to as $\alpha$ ) is the probability of rejecting the null hypothesis when it is actually true. The KS test for the two samples returns a KS statistic $(\mathcal{K})$ and a $p$-value $(\mathcal{P})$. Based on the value of $\alpha$, we calculate a critical value $\left(D_{\text {crit }}\right)$, which is given as

$$
\begin{equation*}
D_{\mathrm{crit}}=c(\alpha) \cdot \sqrt{\left(\frac{1}{n_{1}}+\frac{1}{n_{2}}\right)} \tag{8}
\end{equation*}
$$

where $c(\alpha)=\sqrt{-\ln (\alpha / 2) / 2}$, and $n_{1}$ and $n_{2}$ are the two sample sizes we employ. If $\mathcal{K}>D_{\text {crit }}$ or $\mathcal{P}<\alpha$, then we can reject the null hypothesis at a level of $\alpha$.

We choose the average mass of planets in the systems ( $M_{p, \text { ave }}$ ) as a parameter to divide the full samples into two groups; the current Kepler data suggest that at least the planet radius difference in the systems may be insignificant (Weiss et al. 2018). This does not necessarily mean that the planet mass difference is small as well (see Figure 1). However, the mass-radius relation is the best tool currently available in the literature to estimate the planet mass, and hence we use the planet mass for our KS tests. Practically, we bin the mass range of $1 M_{\oplus} \leqslant M_{p, \text { th }} \leqslant 10 M_{\oplus}$ with a bin size of $0.5 M_{\oplus}$, where $M_{p, \text { th }}$ is a parameterized threshold mass to divide the full samples into two groups, and compare these two samples:


Figure 3. The results of our KS tests and the best case of the cumulative distributions of $S_{c}$ and $S_{s}$ (i.e., $M_{p, \text { th }}=3.5 M_{\oplus}$ ). The left panel shows that $S_{c}$ can be used to divide the full sample of small-sized, multiplanet systems into two populations in the mass range of $2.5 M_{\oplus} \lesssim M_{p, \text { th }} \lesssim 6.3 M_{\oplus}$ (see the red shaded region). In this region, the KS statistic (the red dashed line) is greater than the critical value (the black solid line), and hence the null hypothesis is rejected with $99.5 \%$ confidence; equivalently, low-mass ( $M_{p, \text { ave }} \leqslant M_{p, \text { th }}$ ) and high-mass ( $M_{p, \text { ave }}>M_{p, \text { th }}$ ) systems are not drawn from the same distribution. On the contrary, $S_{s}$ does not show the similar feature. The central and right panels show the best case, where low-mass and high-mass systems are denoted by the blue dashed and the green solid lines, respectively.
planetary systems with $M_{p, \text { ave }} \leqslant M_{p, \text { th }}$ and those with $M_{p, \text { ave }}>M_{p, \text { th }}$.

Figure 3 summarizes our results; the left panel shows the results of our KS tests, and the central and right ones depict the cumulative distributions of $S_{c}$ and $S_{s}$ for the best case (i.e., $M_{p, \text { th }}=3.5 M_{\oplus}$ ). We find that for the mass concentration, the null hypothesis can be rejected for $2.48 M_{\oplus}<M_{p, \text { th }}<6.3 M_{\oplus}$ with the $99.5 \%$ confidence (i.e., $\alpha=0.005$ ); this is evident from the KS statistic, which is greater than the value of $D_{\text {crit }}$ in this range. Our analysis therefore suggests that the full samples can be divided into two groups at $M_{p, \text { th }}=3.5 M_{\oplus}$. On the other hand, a similar distinction is not possible for the orbital spacing.

We discuss below how our observational analysis is consistent with our theoretical prediction, and why the mass concentration and orbital spacing behave differently.

## 3. Discussion

Our theoretical argument has suggested that planetary migration should have played a more important role for planets with $M_{p} \gtrsim 2 M_{\oplus}$ (Figure 2), and our observational analysis has shown that planetary systems with $M_{p \text {,ave }} \lesssim 3 M_{\oplus}$ have a higher value of $S_{c}$ than those with $M_{p, \text { ave }} \gtrsim 3 M_{\oplus}$ (Figure 3). Based on these results, we here discuss how small-sized, multiplanet systems form.

First, we point out that both of our analyses lead to a consistent picture that super-Earth (i.e., $M_{p, \text { ave }} \lesssim 3 M_{\oplus}$ ) systems form out of narrow ring-like distributions of protoplanets while sub-Neptune (i.e., $M_{p \text {,ave }} \gtrsim 3 M_{\oplus}$ ) systems form from radially wide distributions of protoplanets. The difference in the spatial distribution of protoplanets is produced by planetary migration; a chain of protoplanets tends to form readily with a wider radial extent for massive protoplanets due to faster migration. Importantly, the existing $N$-body simulations already show that different (narrow versus wide) spatial distributions of protoplanets leads to different (high versus low) values of $S_{c}$. Figure 4 shows the simulation results obtained by H 09 and HM12. The former targets the formation of terrestrial planets in the solar system and hence starts



Figure 4. The cumulative distributions of $S_{c}$ and $S_{s}$ for the simulation results made by H 09 and HM12. The cumulative distribution of $S_{c}$ shows a clearer difference, which comes from different spatial distributions of protoplanets that eventually undergo giant impact.
from narrow rings whose radial extent is comparable to the resulting value of $S_{s}$. On the other hand, the latter targets the formation of small-sized planets observed by Kepler and begins with wide distributions whose radial extent is a few times larger than the resulting value of $S_{s}$.

It is obvious that quantitative comparison between the observations (Figure 3) and the simulations (Figure 4) cannot be made due to the observational bias and idealization adopted in simulations. For instance, the observational bias may partially wash out the effect of planetary migration; this may be a reason of why there is no clear difference in the cumulative distribution of $S_{s}$ between low-mass and high-mass systems (Figure 3). However, qualitative comparison may be useful; it may be reasonable to consider that the difference in the cumulative distribution of $S_{c}$ is significant enough to infer the spatial distribution of protoplanets.


Figure 5. The optimal value of $\Sigma_{\mathrm{g}}$ as a function of the distance from the central star. The maximum and minimum values of $\Sigma_{g}$ are computed (see the black solid and dashed lines, respectively), using the results of Figure 3 (see the vertical dashed lines in the left panel). For comparison, the MMSN model is denoted by the blue dotted line.

Second, we consider the value of $\Sigma_{\mathrm{g}}$ in the vicinity of the central star. As discussed in Section 2.3, a lower value of $\Sigma_{\mathrm{g}}$ may be preferred, in order to limit the effect of planetary migration (Figure 2). Under the assumption that the population of low-mass and high-mass systems is divided at $M_{p, \mathrm{th}}$, one can compute the optimal value of $\Sigma_{\mathrm{g}}$. Figure 5 shows the results. It is interesting that the resulting values are much lower than the minimum-mass solar nebula (MMSN) model (Hayashi 1981). This difference is readily understood by the fact that there is no close-in planet in the solar system, and hence the MMSN model is not sensitive to such a region. Furthermore, the recent studies support the lower $\Sigma_{\mathrm{g}}$ in the vicinity of the central star (e.g., Choksi \& Chiang 2020), which may be caused by stellar magnetic fields (Hasegawa et al. 2019) and/or disk winds (Ogihara et al. 2015).

Finally, we comment on caveats involved in this work. The most important one would be the observational bias as discussed above. Since it would be complicated to reliably determine the observational bias for multiplanet systems, it may be straightforward to run simulations and impose the bias on the simulation results as done by Mulders et al. (2020) and He et al. (2020). This approach allows one to compare simulations and observations directly and quantitatively. Detailed simulations are desired to verify the results of this work, which is the target of our future work.

Thus, planetary migration and the properties of protoplanetary disks in the vicinity of the central stars are the key to better understanding the origin and properties of observed superEarths and sub-Neptunes.

The authors thank an anonymous referee for useful comments on our manuscript. This research was carried out in part at JPL/Caltech, under a contract with NASA. Y.H. is supported by JPL/Caltech.

## ORCID iDs

Udit Arora (1) https:// orcid.org/0000-0001-7250-0862

## References

Berger, T. A., Huber, D., van Saders, J. L., et al. 2020, AJ, 159, 280
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Sci, 327, 977
Chambers, J. E. 2001, Icar, 152, 205
Chambers, J. E., Wetherill, G. W., \& Boss, A. P. 1996, Icar, 119, 261
Chen, J., \& Kipping, D. 2017, ApJ, 834, 17
Chiang, E., \& Laughlin, G. 2013, MNRAS, 431, 3444
Choksi, N., \& Chiang, E. 2020, MNRAS, 495, 4192
Dawson, R. I., Lee, E. J., \& Chiang, E. 2016, ApJ, 822, 54
Eker, Z., Bakış, V., Bilir, S., et al. 2018, MNRAS, 479, 5491
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Gilbert, G. J., \& Fabrycky, D. C. 2020, AJ, 159, 281
Ginzburg, S., Schlichting, H. E., \& Sari, R. 2018, MNRAS, 476, 759
Hansen, B. M. S. 2009, ApJ, 703, 1131
Hansen, B. M. S., \& Murray, N. 2012, ApJ, 751, 158
Hansen, B. M. S., \& Murray, N. 2013, ApJ, 775, 53
Hasegawa, M., \& Nakazawa, K. 1990, A\&A, 227, 619
Hasegawa, Y., \& Pudritz, R. E. 2011, MNRAS, 417, 1236
Hasegawa, Y., \& Pudritz, R. E. 2012, ApJ, 760, 117
Hasegawa, Y., Yu, T. Y. M., \& Hansen, B. M. S. 2019, A\&A, 629, L1
Hayashi, C. 1981, PThPS, 70, 35
He, M. Y., Ford, E. B., Ragozzine, D., \& Carrera, D. 2020, AJ, 160, 276
Howard, A. W., Marcy, G. W., Johnson, J. A., et al. 2010, Sci, 330, 653
Huber, D., Bryson, S. T., Haas, M. R., et al. 2016, ApJS, 224, 2
Ida, S., \& Lin, D. N. C. 2004, ApJ, 604, 388
Ida, S., \& Lin, D. N. C. 2010, ApJ, 719, 810
Izidoro, A., Ogihara, M., Raymond, S. N., et al. 2017, MNRAS, 470, 1750
Kley, W., \& Nelson, R. P. 2012, ARA\&A, 50, 211
Lissauer, J. J., Fabrycky, D. C., Ford, E. B., et al. 2011, Natur, 470, 53
Mayor, M., Marmier, M., Lovis, C., et al. 2011, arXiv:1109.2497
Mulders, G. D., O’Brien, D. P., Ciesla, F. J., Apai, D., \& Pascucci, I. 2020, ApJ, 897, 72
Mulders, G. D., Pascucci, I., Apai, D., \& Ciesla, F. J. 2018, AJ, 156, 24
NASA Exoplanet Science Institute 2019, Kepler Objects of Interest Cumulative Table, NASA Exoplanet Archive, doi:10.26133/NEA4
Ogihara, M., Morbidelli, A., \& Guillot, T. 2015, A\&A, 584, L1
Owen, J. E., \& Wu, Y. 2013, ApJ, 775, 105
Rogers, L. A. 2015, ApJ, 801, 41
Schlichting, H. E. 2014, ApJL, 795, L15
Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, ApJS, 235, 38
Weiss, L. M., \& Marcy, G. W. 2014, ApJL, 783, L6
Weiss, L. M., Marcy, G. W., Petigura, E. A., et al. 2018, AJ, 155, 48
Winn, J. N., \& Fabrycky, D. C. 2015, ARA\&A, 53, 409
Zhu, W., Petrovich, C., Wu, Y., Dong, S., \& Xie, J. 2018, ApJ, 860, 101


[^0]:    3 We accessed https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/ nph-tblView?app=ExoTbls\&config=cumulative on 2021 January 21 at 16:17, and obtained the data with the size of $9564 \times 49$ columns.

[^1]:    4 Note that the stellar masses estimated by our approach are comparable to those obtained using the Gaia data (Berger et al. 2020); we find that the difference is small ( $<20 \%$ ) enough for most cases that our results do not change very much.
    5 Note that we have removed one planetary system, Kepler-1659, due to the unphysically high mass of one planet.

