The Exoplanet Radius Valley from Gas-driven Planet Migration and Breaking of Resonant Chains

André Izidoro[1,2] , Hilke E. Schlichting[3] , Andrea Isella[1] , Rajdeep Dasgupta[2] , Christian Zimmermann[4], and Bertram Bitsch[4]

[1] Department of Physics and Astronomy 6100 MS 550, Rice University, Houston, TX 77005, USA
[2] Department of Earth, Environmental and Planetary Sciences, 6100 MS 126, Rice University, Houston, TX 77005, USA
[3] Department of Earth, Planetary, and Space Sciences, The University of California, Los Angeles, Charles E. Young Drive East, Los Angeles, CA, USA
[4] Max-Planck-Institut für Astronomie Königstuhl 17, D-69117 Heidelberg, Germany

Abstract

The size frequency distribution of exoplanet radii between 1 and $4R_\oplus$ is bimodal with peaks at ~ 1.4 $R_\oplus$ and ~ 2.4 $R_\oplus$, and a valley at ~ 1.8 $R_\oplus$. This radius valley separates two classes of planets—usually referred to as "super-Earths" and "mini-Neptunes"—and its origin remains debated. One model proposes that super-Earths are the outcome of photoevaporation or core-powered mass loss stripping the primordial atmospheres of the mini-Neptunes. A contrasting model interprets the radius valley as a dichotomy in the bulk compositions, where super-Earths are rocky planets and mini-Neptunes are water-ice-rich worlds. In this work, we test whether the migration model is consistent with the radius valley and how it distinguishes these views. In the migration model, planets migrate toward the disk's inner edge, forming a chain of planets locked in resonant configurations. After the gas disk dispersal, orbital instabilities "break the chains" and promote late collisions. This model broadly matches the period-ratio and planet-multiplicity distributions of Kepler planets and accounts for resonant chains such as TRAPPIST-1, Kepler-223, and TOI-178. Here, by combining the outcome of planet formation simulations with compositional mass–radius relationships and assuming the complete loss of primordial H-rich atmospheres in late giant impacts, we show that the migration model accounts for the exoplanet radius valley and the intrasystem uniformity ("peas in a pod") of Kepler planets. Our results suggest that planets with sizes of ~ 1.4 $R_\oplus$ are mostly rocky, whereas those with sizes of ~ 2.4 $R_\oplus$ are mostly water-ice-rich worlds. Our results do not support an exclusively rocky composition for the cores of mini-Neptunes.

Unified Astronomy Thesaurus concepts: Planet formation (1241); Exoplanets (498)

1. Introduction

Kepler transit observations have shown that planets with sizes between those of Earth (1 $R_\oplus$) and Neptune (~ 4 $R_\oplus$) are extremely common (Lissauer et al. 2011; Batalha et al. 2013; Fressin et al. 2013; Howard 2013; Fabrycky et al. 2014; Marcy et al. 2014). Demographics analysis suggests that at least 30%–55% of the Sun-like stars host one or more planets within this size range and with orbital periods shorter than 100 days (Mayor et al. 2011; Howard et al. 2012; Fressin et al. 2013; Petigura et al. 2013; Mulders 2018; Mulders et al. 2018; Zhu et al. 2018; He et al. 2019, 2021). Uncertainties in stellar radius estimates from photometric Kepler observations prevented a detailed assessment of the intrinsic planet size distribution (Fulton et al. 2017; Petigura et al. 2017). Specific trends in planet sizes only started to emerge from the data with more precise determination of stellar radii by follow-up surveys (California Kepler Survey, CKS) and the use of Gaia improved parallaxes (Johnson et al. 2017; Van Eylen et al. 2018; Petigura et al. 2022). These studies showed that the size frequency distribution of planets between ~ 1 and $\sim 4R_\oplus$ is bimodal with peaks at ~ 1.4 $R_\oplus$ and ~ 2.4 $R_\oplus$, and a valley at ~ 1.8 $R_\oplus$ (Fulton et al. 2017; Fulton & Petigura 2018; Petigura 2020). The best-characterized planets with sizes of about ~ 1.4 $R_\oplus$ are consistent with rocky composition, as constrained by their estimated bulk densities (Fortney et al. 2007; Adams et al. 2008;

Lopez & Fortney 2014; Weiss & Marcy 2014; Dorn et al. 2015; Wolfgang et al. 2016; Zeng et al. 2016, 2019; Bashi et al. 2017; Chen & Kipping 2017; Otegi et al. 2020). These planets are usually referred to as "super-Earths." Planets with sizes of about $\sim 2.4R_\oplus$ are consistent with the presence of volatiles—which could reflect either rocky cores with H-He-rich atmospheres or ice/water-rich planets (Kuchner 2003; Rogers & Seager 2010; Lopez & Fortney 2014; Zeng et al. 2019; Mousis et al. 2020; Otegi et al. 2020). These planets are commonly referred to as "mini-Neptunes." For a detailed discussion, see reviews by Bean et al. (2021) and Weiss et al. (2022), and references therein.

The planet-size distribution also reveals an intrasystem uniformity in planet radii, where planets in the same systems tend to have similar sizes ($R_{j+1}/R_j \approx 1$, where R_{j+1} and R_j are the radii of two adjacent planet pairs; Weiss et al. 2018). This is popularly known as the "peas-in-a-pod" feature of exoplanets.

The planet-size distribution strongly constrains planet formation and evolution models. Different mechanisms have been proposed to explain the Kepler planets' bimodal distribution, including atmospheric loss via photoevaporation (Lopez et al. 2012; Owen & Wu 2013; Lopez & Fortney 2013; Kurosaki et al. 2014; Luger et al. 2015; Mordasini 2020) and core-powered effects (Ginzburg et al. 2018; Gupta & Schlichting 2019; Rogers et al. 2021; Gupta et al. 2022). The photoevaporation model suggests that super-Earths are the photoevaporated rocky cores of mini-Neptunes (Owen & Wu 2013; Mordasini 2020; Owen & Campos Estrada 2020; Zhang et al. 2022). Atmospheric loss via core-powered effects also supports a predominantly rocky composition for super-Earths and the cores of mini-Neptunes (Ginzburg et al. 2018; Gupta & Schlichting 2019; Gupta et al. 2022). However, the

bimodal size distribution of planets smaller than $4R_\oplus$ has been also interpreted as reflecting distinct compositions between rocky super-Earths and the ice-/water-rich mini-Neptunes (Zeng et al. 2019; Venturini et al. 2020; Izidoro et al. 2021a).

The goal of this work is to investigate whether the observed radius distribution of super-Earths and mini-Neptunes is consistent with planet-formation models and, in particular, with the migration model (Izidoro et al. 2017, 2021a; see Raymond et al. 2020 and Bean et al. 2021 for a detailed discussion of other plausible formation models). Some of the key questions we will address in the following are: Can the size difference between super-Earths and mini-Neptunes be explained by planet formation models? What is the role and relevance of atmospheric loss processes in the context of the formation and early evolution of planetary systems?

The migration scenario proposes that super-Earths and mini-Neptunes formed during the gas disk phase and experienced gas-driven planet migration (e.g., Terquem & Papaloizou 2007; Ida & Lin 2008, 2010; McNeil & Nelson 2010; Hellary & Nelson 2012; Coleman & Nelson 2014; Cossou et al. 2014; Coleman & Nelson 2016; Izidoro et al. 2017; Carrera et al. 2018, 2019; Ogihara et al. 2018; Raymond et al. 2018). This model suggests that convergent migration promotes the formation of chains of planets anchored at the disk's inner edge and locked in first-order mean motion resonances with each other, the so-called resonant chains. In a first-order mean motion resonance, the ratio of orbital periods of two resonant planets is a ratio of integers with a difference of one (e.g., $P_2/P_1 = 2/1$, but it also requires libration of associated resonant angles). After gas disk dispersal, a large fraction of the resonant chains become dynamically unstable (Izidoro et al. 2017; Lambrechts et al. 2019; Esteves et al. 2020, 2022; Izidoro et al. 2021a). This instability phase leads to orbital crossing among planets and giant impacts, producing planets spaced by Hill radii in agreement with Kepler observations (Pu & Wu 2015; Izidoro et al. 2017). Giant impacts are expected to erode primordial atmospheres (Liu et al. 2015; Inamdar & Schlichting 2016; Biersteker & Schlichting 2019; Chance et al. 2022) and change planetary architectures. It remains elusive, however, if this scenario is consistent with the size distribution of exoplanets. In this paper, we revisit the migration model to test if it is consistent with the observed peaks in the size distribution of planets at $\sim 1.4R_\oplus$ and $\sim 2.4R_\oplus$, and a valley at $\sim 1.8R_\oplus$. We will also test how this model matches the so-called "peas-in-a-pod" feature of exoplanets (Weiss et al. 2018; Millholland & Winn 2021).

2. Methods

2.1. Planet-formation Simulations

Our analysis is based on the numerical simulations presented in Izidoro et al. (2021a) and a small sample of new simulations. These simulations model the formation of super-Earths and mini-Neptunes ($1 < R < 4R_\oplus$; $1 < M < 20M_\oplus$) by following the evolution of Moon-mass planetary seeds as they grow inside a circumstellar disk. Our model accounts for several physical processes such as gas-assisted pebble accretion (e.g., Ormel & Klahr 2010; Lambrechts & Johansen 2012; Johansen & Lambrechts 2017), gas-driven planet migration (e.g., Baruteau et al. 2014), gas tidal damping of orbital eccentricity and inclination (e.g., Cresswell & Nelson 2006; Cresswell 2008), and mutual gravitational interaction of planetary embryos.

From Izidoro et al. (2021a), we selected three models that differ in terms of the starting location of planetary seeds in the gaseous disk (w), the flux of pebbles available for the seeds to grow (S_{peb}), the starting time of the simulations relative to the disk age (t_{start}), and the size of pebbles inside the disk water snow line[5] (R_{peb}; see Table 1 in Izidoro et al. 2021a). More specifically, the three models have the following initial setup:

1. Model-III with $w = 0.2-2$ au, $t_{start} = 0.5$ Myr, $R_{peb} = 1$ cm, and $S_{peb} = 5$, hereafter referred to as model A;
2. Model-II with $w = 0.7-20$ au, $t_{start} = 3$ Myr, $R_{peb} = 1$ mm, $S_{peb} = 5$, hereafter referred to as Model B;
3. Model-III with $w = 0.2-2$ au, $t_{start} = 0.5$ Myr, $R_{peb} = 1$ mm, and $S_{peb} = 10$; hereafter referred to as model C.

The selected models produce (i) planetary systems dominated by planets with a rocky composition (model A); (ii) planetary systems dominated by water-rich worlds (model B); and (iii) planetary systems with mixed populations of water-rich and rocky worlds (model C).

The breaking-the-chains scenario broadly matches the observed period ratio distribution of exoplanets with radii smaller than $\sim 4R_\oplus$ if the great majority of the resonant chains become dynamically unstable after gas disk dispersal (Izidoro et al. 2017, 2021a). A good match to observations requires an instability rate of more than 90%–95%, but not 100%. The breaking-the-chains model suggests that the Kepler sample is well matched by mixing >90% of unstable systems with less than <10% of stable systems. This is because observations also show that resonant chains exist, so not all chains should become dynamically unstable. Iconic examples of resonant chains are TRAPPIST-1 (Gillon et al. 2017; Luger et al. 2017), Kepler-223 (Mills et al. 2016), and TOI-178 (Leleu et al. 2021) systems.

For each model of Izidoro et al. (2021a), we have ~ 50 simulations available that slightly differ in the initial distribution of planetary seeds and masses. For models A and B, about 90%, and 76% of the planetary systems become naturally unstable after the gas disk dispersal, respectively. For model C, the fraction of unstable systems is about 50%. It is possible that all of these fractions could be higher if we had integrated our simulations for longer timescales (e.g., >1 Gyr, instead of ~ 50 Myr; Izidoro et al. 2021a), but this is computationally too demanding. It is also possible that our simulations miss some important physics that may help trigger dynamical instabilities after gas disk dispersal. These include planetesimal scattering effects (Chatterjee & Ford 2015; Raymond et al. 2022), interactions of planet chains with distant external perturbers (Lai & Pu 2017; Bitsch et al. 2020), planet-star tidal interaction effects (Bolmont & Mathis 2016), and spin–orbit misalignment effects (Spalding & Batygin 2016).

Whereas models A and B provide 45 and 38 unstable systems, respectively, model C produces only 25 unstable systems (see Figure 20 of Izidoro et al. 2021a). To overcome potential problems caused by the smaller number of unstable systems in model C (e.g., misleading results due to small number statistics), we have increased their number by artificially triggering dynamical instabilities in the 25 systems that remain stable after 50 Myr.[6] This was done by changing the pericenter of the orbit of the two innermost planets by a random amount

---

[5] Location in the gaseous disk where water condenses as ice.
[6] We have verified that the main results presented in this paper would not change qualitatively if we had used the original sample of Izidoro et al. (2021a).

varying between -5% and $+5\%$ of their respective values. A similar approach was adopted in simulating dynamical instabilities within the solar system (e.g., Levison et al. 2011; Nesvorný & Morbidelli 2012). Artificial triggering of dynamical instabilities, this time via an artificial reduction of the planet mass, has also been used in studies of the dynamical architecture of super-Earth and mini-Neptune systems (Matsumoto & Ogihara 2020; Goldberg & Batygin 2022). After triggering instability, we extend the numerical integration of these 25 systems for another ~ 50 Myr.

Finally, we expanded our model C sample with 15 new simulations based on Model III of Izidoro et al. (2021a) that use slightly different pebble fluxes ($S_{peb} = 5$ and $S_{peb} = 2.5$; $t_{start} = 0.5$ Myr) and/or initial distribution of seeds (inside 1 au, instead of 2 au as in the nominal case). Out of these 15 new simulations, only one system represents a stable system. All these simulations were numerically integrated for 50 Myr.

In summary, models A, B, and C are composed of 47 (2 stable systems), 41 (3 stable systems), and 65 (1 stable system) simulations, respectively. The ratio between stable and unstable systems broadly satisfies the fact that the migration model typically requires less than 10% of stable systems in order to match observations in terms of period ratios and planet multiplicity distributions (Izidoro et al. 2017, 2021a).

2.2. Accretion of Water/Ices

In Izidoro et al. (2021a), planetary seeds grow via pebble accretion and mutual collisions. Pebbles beyond the snow line are assumed to have 50% of their masses in water ice. Pebbles that drift inwards and cross the water snow line are assumed to sublimate, losing their water component and releasing their silicate counterpart in the form of small silicate grains (Morbidelli et al. 2015). Planets growing by pebble accretion beyond the snow line tend to become water rich whereas those growing inside the snowline tend to be rocky (Bitsch et al. 2019b; Izidoro et al. 2021a). We model collisions as perfect merging events that conserve mass (including water) and linear momentum. Statistically speaking, collisional fragmentation has a negligible effect on the final dynamical architecture of planetary systems produced in the breaking-the-chains model (Esteves et al. 2022; see also Poon et al. 2020).

2.3. Converting Planetary Mass to Planetary Radius

Our planet formation simulations provide the mass and composition of planets but not their size/radius. In order to compare our model to the exoplanet radius valley and the peas-in-a-pod trend, we use compositional mass–radius relationships (Zeng et al. 2016, 2019). To this end, we use the mass–radius relationship (MRR) fits from Zeng et al. (2019),[7] who modeled planets with the following compositions:

1. Earth-like rocky composition: 32.5% Fe + 67.5% $MgSiO_3$. We refer to these planets simply as rocky planets.
2. Earth-like rocky composition with H_2 atmosphere: 99.7% Earth-like composition + 0.3% H_2 envelope by mass. We refer to these planets as rocky planets with H-rich/primordial atmospheres.

---

3. Water-rich composition: 50% Earth-like rocky core + 50% H_2O layer by mass. We refer to these planets as water-rich planets.
4. Water-rich composition with H_2 atmosphere: 49.85% Earth-like composition + 49.85% H_2O layer + 0.3% H_2 envelope by mass. We refer to these planets as water-rich planets with H-rich/primordial atmospheres.

The water mass fraction of our synthetic planets is an outcome of the simulations that track the composition of the material accreted by each planet. In general, these compositions vary between 0% and 50% of mass in water. However, because mass–radius relationships for intermediate compositions (e.g., water-mass fraction of 25%) are not publicly available, we assume that planets with water-mass fractions larger than 10% are water-rich worlds and those with lower water-mass fractions are rocky. We argue that this simplification does not degrade the quality of our study because most of our final planets have either 0% or >20% of water content. Nevertheless, in order to account for uncertainties coming from our model simplifications (e.g., we ignore water/mass loss via impacts; see Marcus et al. 2010; Leinhardt & Stewart 2012; Biersteker & Schlichting 2021; Esteves et al. 2022) and the few planets with intermediate water content, we calculated planet radii assuming 1σ uncertainties of 7%. This uncertainty is motivated by the typical difference in the size of planets (with no atmosphere) with a 25% water-mass fraction and those with 50% (see Zeng et al. 2016, 2019). In the Appendix, we also test our results against the empirical mass–radius relationship of Otegi et al. (2020), and we show that our main conclusions do not qualitatively change.

2.4. Atmospheric Loss via Giant Impacts

In simulations of Izidoro et al. (2021a), gas accretion onto growing planets was not taken into account (see papers by Bitsch et al 2019a, 2020, which focused on the formation of giant planets). This is a reasonable approximation because state-of-the-art 3D hydrodynamical simulations show that the atmosphere of planets with masses smaller than 10–$15 M_\oplus$ should correspond to only a few percent of the core mass. In this mass range and below, recycling between the planetary atmosphere and the circumstellar disk is an efficient process, limiting atmospheric mass (Cimerman et al. 2017; Lambrechts & Lega 2017; Béthune & Rafikov 2019; Moldenhauer et al. 2022).

In all simulations presented here, the masses of planets at the end of the gas disk dispersal are systematically smaller than $10 M_\oplus$. Both hydrodynamical simulations and analytic calculations show that their putative atmospheric masses would be limited to a few percent of their total masses (e.g., Lee & Chiang 2015; Ginzburg et al. 2016). Following previous studies (e.g., Owen & Wu 2017; Ginzburg et al. 2018; Gupta & Schlichting 2019), we assume that the atmosphere-to-core-mass ratio of our planets at the end of the gas disk phase is 0.3% in our nominal simulations, but we also test cases with 0.1%, 1% and 5% (these cases are presented in the Appendix).

We also assume that giant impacts ($M_p/M_t > 0.1$; where M_p, M_t are the projectile and target masses, respectively) that take place after the gas disk dispersal completely strip primordial atmospheres (Liu et al. 2015; Biersteker & Schlichting 2019), either leaving behind bare rocky or bare water-rich cores. Following this definition, $\gtrsim 80\%$–90% of the late impacts in our

simulations are flagged as giant impacts. We do not model the formation of secondary/outgassed atmospheres in this work.

2.5. Atmospheric Loss via Photoevaporation

The model of Izidoro et al. (2021a) also does not include photoevaporation or core-powered mass loss of planetary atmospheres. As discussed before, we do not model gas accretion, but the existence of gaseous atmospheres is assumed during our data analysis.

Recent studies discovered that photoevaporation and/or core-powered mass loss might explain the exoplanet radius valley (Owen & Wu 2017; Jin & Mordasini 2018; Gupta & Schlichting 2019, 2020). To investigate the robustness of our results to the possibility of subsequent atmospheric loss after the giant-impact phase has concluded, we analyze our simulations including atmospheric mass loss by photoevaporation after the giant-impact phase.

In order to test the impact of photoevaporation in our model, we follow a simple energy-limited escape prescription to estimate the planet's atmosphere stability when subject to stellar X-ray and ultraviolet (XUV) radiation (e.g., Owen & Wu 2017). We follow the criterion by Misener & Schlichting (2021), which compares the atmospheric binding energy to the energy the planet receives from 100 Myr to 1 Gyr. If the ratio between these two quantities Φ is $\lesssim 1$, then sufficient energy is received by the planet to have its atmosphere photoevaporated. Φ is given by

$$\Phi \simeq \frac{f}{3.3 \times 10^{-3}} \left(\frac{E_{\rm out}}{5.2 \times 10^{45} {\rm erg}} \right)^{-1} \left(\frac{\eta}{0.1} \right)^{-1}$$
$$\times \left(\frac{M_{\rm c}}{3M_{\oplus}} \right)^{5/4} \left(\frac{T_{\rm eq}}{1000 \ {\rm K}} \right)^{-4} \left(\frac{R_p}{R_c} \right)^{-2}, \tag{1}$$

where f is the atmosphere-to-core-mass ratio, which in our simulations varies from 0.1% to 5%, $M_{\rm c}$ is the planetary core mass, $T_{\rm eq}$ is the planet equilibrium temperature, $R_{\rm p}$ is the planet radius, and R_c is the core radius. The integrated stellar energy output is set to $E_{\rm out} = 5.2 \times 10^{45}$ erg, while the dimensionless efficiency parameter describing the amount of energy available for driving mass loss is set as $\eta = 0.1$ (Owen & Wu 2017).

For every planet in our simulations, M_c is provided directly by our planet formation simulation, while R_p and R_c are calculated from Zeng et al. (2019) assuming planets with a fixed and identical atmosphere-to-core-mass ratio (f) and bare planets, respectively. The planet equilibrium temperature is calculated as

$$T_{\rm eq} = 279 \left(\frac{a_{\rm p}}{1 \ {\rm au}} \right)^{-1/2} \ {\rm K}, \tag{2}$$

where a_p is the planet's orbital semimajor axis. For all the planets with an atmosphere, i.e., those that did not experience a late giant impact, we calculate Φ, and, if $\Phi < 1$, we remove the planet's atmosphere and assign to that planet a radius equal to the core radius R_c. Vice versa, if $\Phi > 1$, we assume that photoevaporation is inefficient and assign to that planet the original radius R_p. We verified that our simplified treatment of photoevaporation is qualitatively consistent with more sophisticated photoevaporation models of the literature (e.g., Owen & Wu 2017).

3. Results

Figure 1 shows the orbital period versus planet radius for every planet in our simulations at the end of the gas disk phase, i.e., before orbital dynamical instabilities take place. From top to bottom, we show the results of models A, B, and C. The panels on the right show planets color-coded by their water/ice mass fraction, while the panels on the left show planets color-coded by mass. At this stage, all the planets have an hydrogen atmosphere-to-core-mass ratio of 0.3% and most of the planets are locked in resonant chains. The simulated planets are typically less massive than 10 $M_{\oplus}$ (Izidoro et al. 2021a) and, due to the presence of primordial atmospheres, are larger than ∼1.5 $R_{\oplus}$. Models A, B, and C are dominated by rocky (black), water/ice-rich (yellow), and mixed-composition planets, respectively.

Figure 2 shows the final architecture of our planetary systems after dynamical instabilities have taken place (and without the effects of photoevaporation; see the Appendix). In addition to showing planet masses and water-mass fractions via color coding (as in Figure 1), we also show the number of late giant impacts that each planet experienced during the final stage of our simulations between 5 and 100 Myr. Panels in the left, central, and right columns show models A, B, and C, respectively. During the dynamical instability phase, planets experienced up to four giant impacts, with most of the planets experiencing one or two giant impacts. About 26% of the planets in model A, 36% of the planets in model B, and 43% of the planets in model C did not experience any giant impact after gas disk dispersal. Model C exhibits the lowest number of late giant impacts because its planetary systems are less crowded at the beginning of the dynamical instability phase as compared to models A and B.

The green-shaded regions in the bottom panels of Figure 2 show the density distribution of the simulated planets, with darker green indicating higher density. For comparison purposes, we also show with a black dashed line the location of the center of the exoplanet planet radius valley as derived by previous studies (Gupta & Schlichting 2019; Van Eylen et al. 2019). In model A, the planet density distribution has a single peak that covers a broad range of planet radii from roughly 1.5 to 3 $R_{\oplus}$. This distribution is not consistent with exoplanet observations. Model B also shows a single-peaked distribution centered at about 2.4 $R_{\oplus}$, but, in contrast to model A, the peak covers a narrow range of planet radii, from roughly 2 to 3 $R_{\oplus}$. However, exoplanets show a second peak in the radius distribution at about 1.4 $R_{\oplus}$, which is not accounted for by model B. Finally, model C shows two peaks at ∼1.4 $R_{\oplus}$ and 2.4 $R_{\oplus}$ and a deficit of planets at ∼1.8 $R_{\oplus}$, which coincides fairly well with the location of the exoplanet planet radius valley (see the black dashed line; Fulton & Petigura 2018; Van Eylen et al. 2019). Model C is therefore the model that best matches the demographics of exoplanetary systems.

In order to understand the origin of the radius valley in model C, it is important to recall that this model produces a dichotomy in composition, with planets larger than $2R_{\oplus}$ being water rich and planets under $1.5R_{\oplus}$ being mostly rocky. The top-right panel of Figure 2 shows that all planets smaller than $1.5R_{\oplus}$ experienced at least one late giant impact. To understand how the radius valley emerges, let us take as an example a rocky planet that before the instability phase has an orbital period of 23 days, a mass of about $3M_{\oplus}$, and an atmosphere-to-core-mass ratio of 0.3%. The radius of this planet, as calculated using the MRR of Zeng et al. (2019), is roughly 2.1$R_{\oplus}$. If this planet collides with an equal-mass planet of similar composition
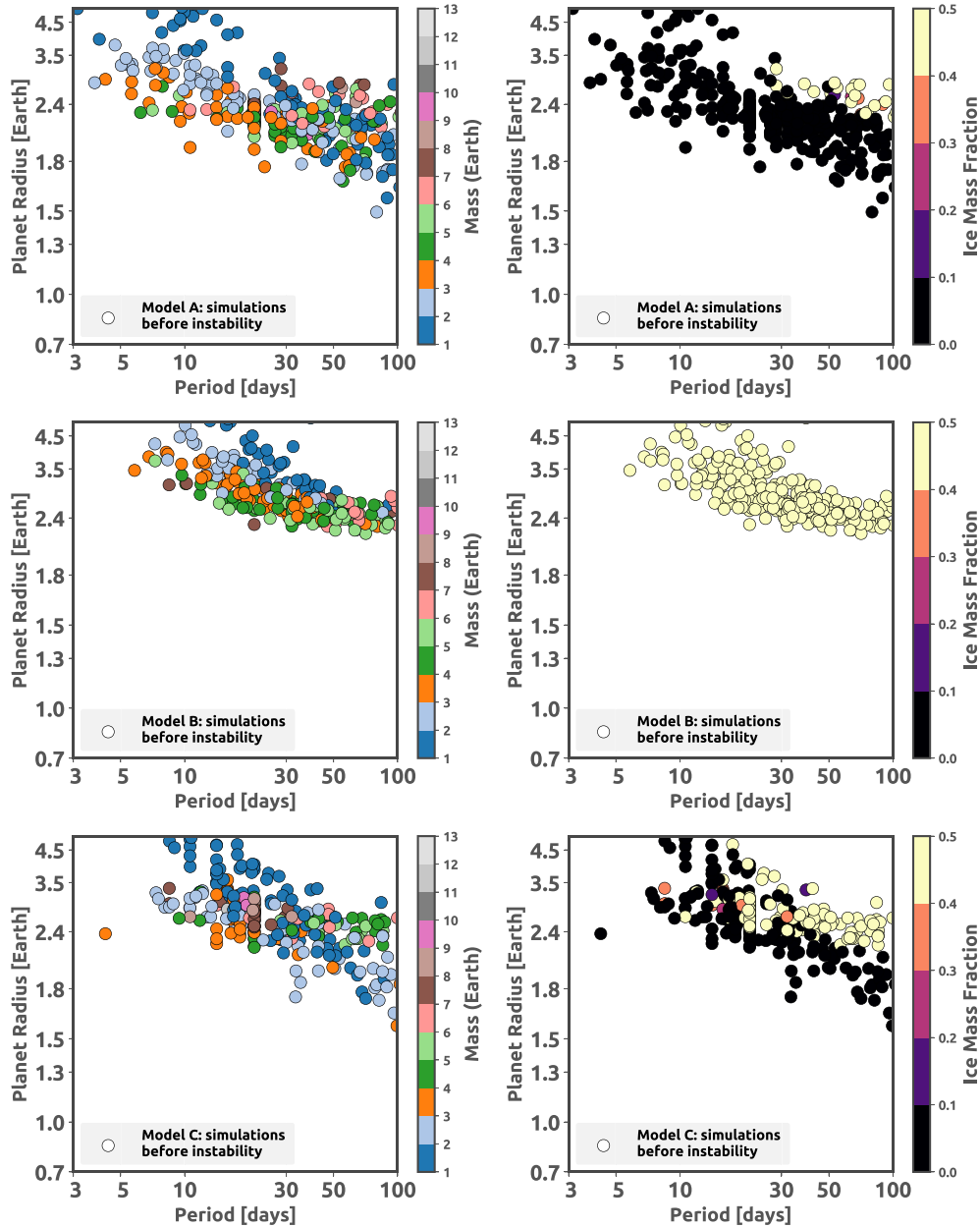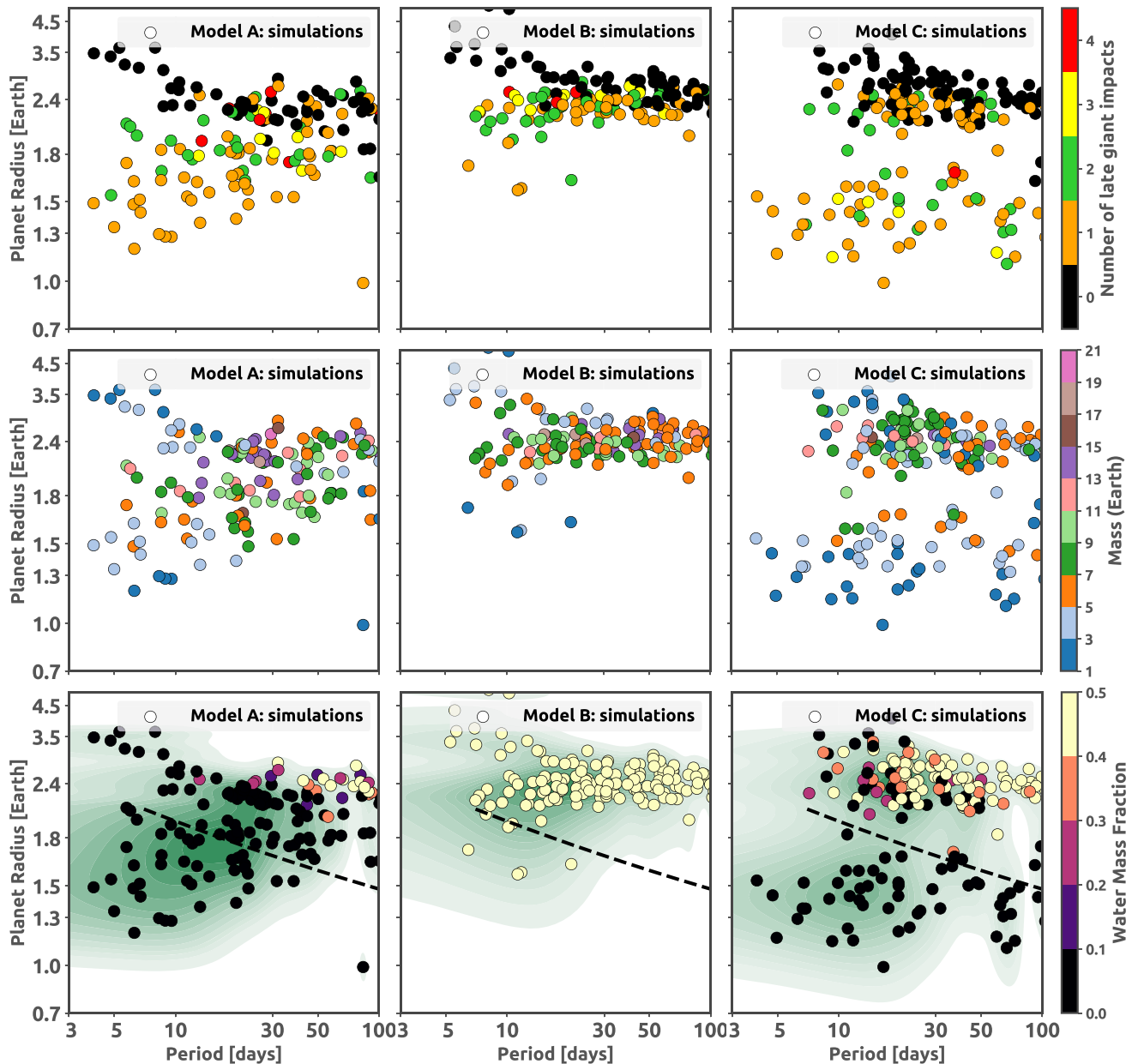
Figure 1. Architecture of planetary systems at the end of the gas disk dispersal phase, i.e., before the onset of dynamical instabilities and the breaking of resonant chains. Planet sizes are calculated assuming primordial atmosphere-to-core-mass ratios of 0.3%. The horizontal axis shows the orbital period and the y-axis shows the planetary radius as calculated from the MRR from Zeng et al. (2019). Planets are shown as individual dots. Panels show the results of all simulations and all planets ($P < 100$ days) produced in each model. From top to bottom, it shows models A, B, and C, respectively. Individual planets are color-coded according to their masses (left-column) and ice/water-mass fractions (right-column panels).

during the instability phase, it will lose its primordial atmosphere (Biersteker & Schlichting 2019), and its final mass and radius will be roughly $6M_{\oplus}$ and $1.6R_{\oplus}$, respectively. Collisions result in atmospheric loss, which reduces the radius of rocky planets significantly, moving them from above to below the radius valley.

Note, however, this is not the case for water-rich planets. To illustrate this contrasting scenario, let us take a second planet with the same orbital period of 23 days, the same mass and atmosphere-to-core-mass ratio, but with a water-rich composition (water-mass fraction of 50%) at the end of the gas disk phase. In this case, the planet's radius before the instability phase is about $2.9R_{\oplus}$. If the planet collides with an equal-mass planet of similar composition, its mass will double and its

radius will be roughly $2.1R_{\oplus}$, which is still above the radius valley. This shows that the origin of the radius valley in our model is associated with a dichotomy in (core) composition. In addition, water-rich planets/cores are typically more massive at the end of the gas disk phase than rocky ones, due to the higher efficiency of pebble accretion and larger pebble isolation mass beyond the snow line than in the inner disk (Lambrechts et al. 2014; Bitsch et al. 2018; Bitsch 2019). This reduces the likelihood that water-rich planets move below or fill the radius valley. If dynamical instabilities after the disk dispersal, and the resulting planetary collisions, are a common process of planet formation, our models indicate that the radius valley does not form if planets/cores have similar compositions and masses, as in models A and B (see also Section 4).

Figure 2. Architecture of planetary systems at the end of our simulations, i.e., after dynamical instabilities have taken place. The horizontal axis shows the orbital period and the *y*-axis the planetary radius as calculated from the MRRs of Zeng et al. (2019). Planets are shown as individual dots. We only show planets with orbital periods shorter than 100 days. In all these models, more than ∼95% of the resonant chains became unstable after gas disk dispersal. Models A, B, and C are shown from left to right. Points are color-coded according to the planets' number of late giant impacts (top row of panels), masses (middle row of panels), and ice-mass fractions (right column). We ignore the effects of photoevaporation and/or other subsequent atmospheric mass loss in all these cases (see the Appendix for simulations where the effects of photoevaporation are included). The dashed lines in the bottom panels show the exoplanet radius valley slope, calculated as $R = 10^{-0.11 \log_{10}(P) + 0.4}$ (Gupta & Schlichting 2019; Van Eylen et al. 2019). The green background contours show the kernel density distribution of the planets in our simulations (dots of the figure).

Supporting the results of Figure 2, Figure 3 shows that the only model showing a valley in the planet-radius distribution is model C. The peaks at 1.4–1.5$R_\oplus$ and 2.4$R_\oplus$ are also pronounced, matching fairly well the location of the peaks in the CKS data. Our simulations, however, produce a peak at 2.4$R_\oplus$, higher than that of exoplanets after completeness corrections (Fulton & Petigura 2018). Assuming that the relative sizes of the observed peaks are truly representative of reality—which may not be the case, for instance, due to observational bias—one could imagine several ways to reconcile our model with observations. Possible solutions could be achieved via an increase in the number of systems with rocky planets, an increase in the efficiency of

photoevaporation or the inclusion of additional atmospheric mass-loss mechanisms, or a reduction in the efficiency of formation of water-rich planets (see also the Appendix, where we test a different MRR and different atmosphere-to-core-mass ratios).

The panels on the right column of Figure 3 show the size-ratio distribution of adjacent planets. The exoplanet sample in this case comes from Weiss et al. (2018) but still from the CKS sample. All our models produce size-ratio distributions that are as narrow as the exoplanet sample and broadly match the peak at $R_{j+1}/R_j \approx 1$. However, model C provides the best match to the peak and shape of the frequency distribution. These results demonstrate that our model, which self-consistently accounts
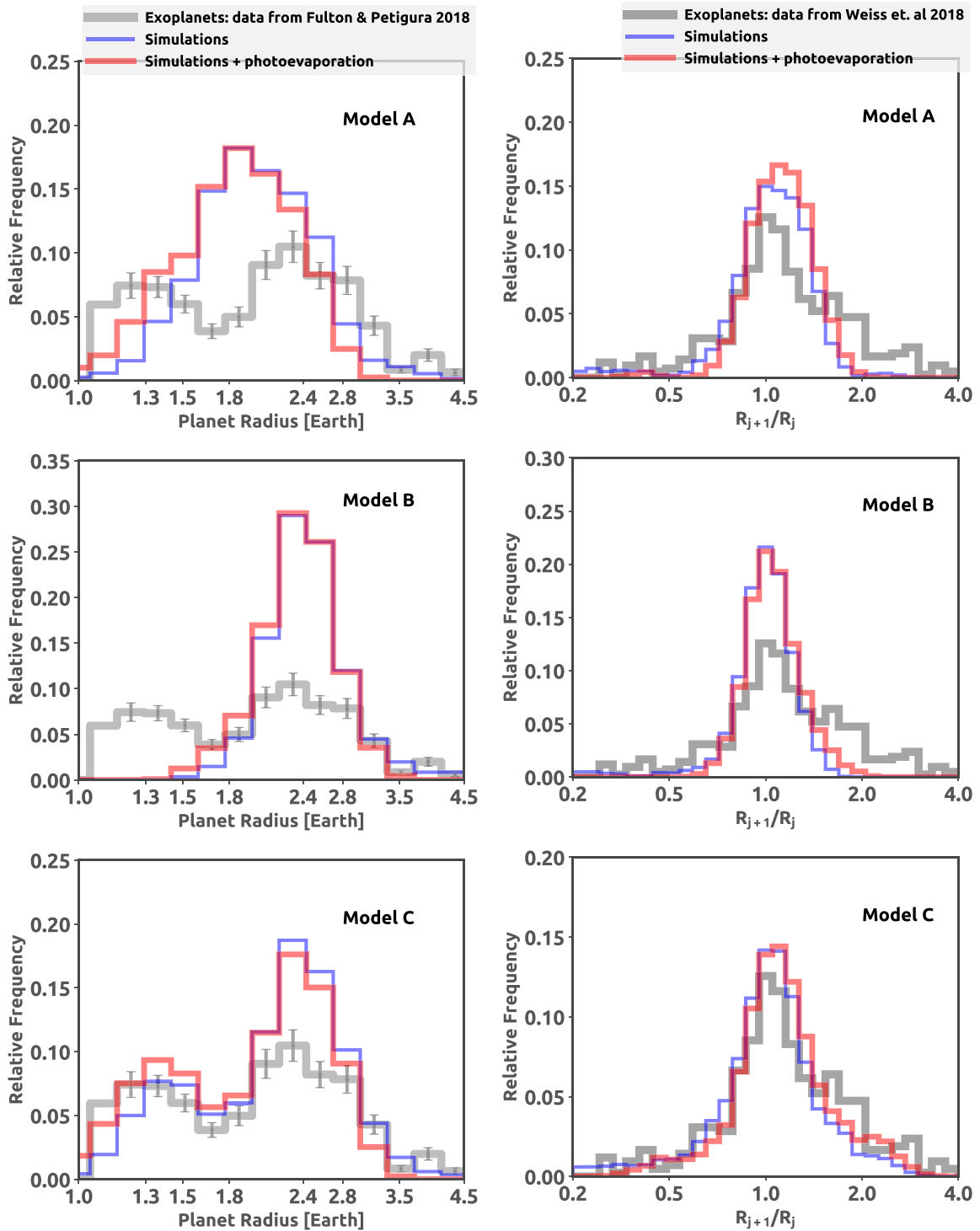
Figure 3. Left: planet-radius distribution. Right: planet size-ratio distribution of adjacent planet pairs. These distributions are computed after dynamical instabilities and the breaking of the resonant chains. Exoplanets are shown in gray. Blue shows the outcome of our planet-formation simulations. Red shows the outcome of our planet-formation simulations including the effects of photoevaporation. From top to bottom, the panel rows show models A, B, and C. For all models, we use the MRR of Zeng et al. (2019) assuming an uncertainty of 7% in size and an initial atmosphere-to-core-mass ratio of 0.3%. We build these histograms accounting for uncertainties in radius by generating the possible radius of each planet 100 times.

for the planets' compositions, dynamical evolution, and late giant impacts, provides a natural explanation for both the observed distribution of planetary radii (i.e., the radius valley feature) and size ratios (i.e., the peas-in-a-pod feature).

Additional results for cases including the effects of photoevaporation are shown in the Appendix. Overall, we find

that in our model photoevaporation has a negligible to small impact on the distribution of planets in the planet radius versus orbital period space. Planets populating the radius valley in model A do not have atmospheres because of giant impacts, and therefore, they cannot be stripped via photoevaporation. In conclusion, if late giant impacts are common, photoevaporation

(a) Before photoevaporation   (b) After photo-evaporation
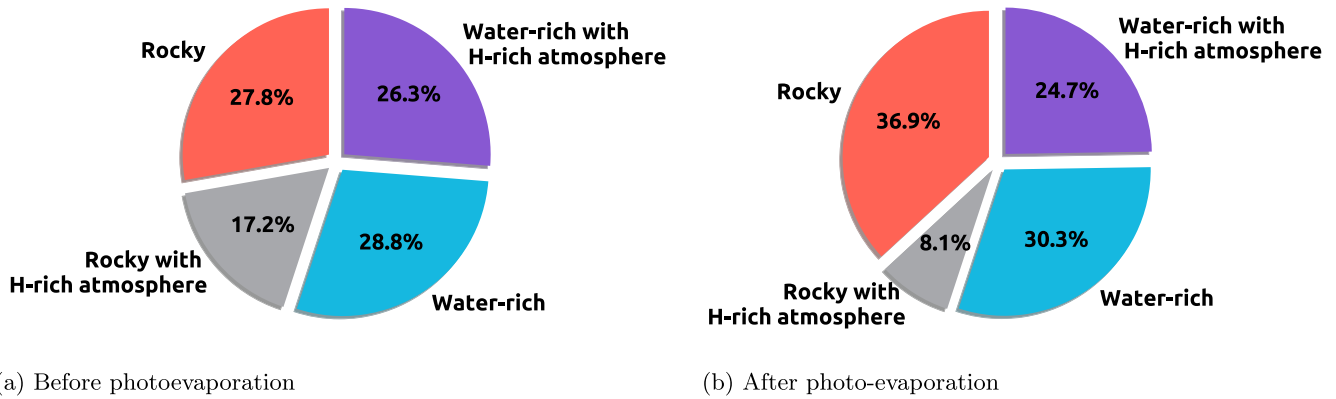
Figure 4. Fraction of planets of different types produced in our simulations of model C that match the radius valley and peas-in-a-pod feature. The left-hand-side pie chart shows the relative fractions when the effects of photoevaporation are ignored. The right-hand-side panel shows the relative fraction when we include the effects of photoevaporation.

does not seem to affect the general result that the radius valley requires a dichotomy in the composition of planetary cores. We have also included the effects of observational bias in our simulations by performing synthetic transit observations of our planetary systems, and our results also do not qualitatively change (see the Appendix).

In Figure 4 we show the fraction of planets of different types produced in our simulations of model C. If we ignore the effects of photoevaporation or any other sub-subsequent atmospheric mass loss (Figure 4(a)), 27.8% of our planets are rocky, 17.2% are rocky with primordial atmospheres, 29.8% are water rich, and 26.3% are water rich with primordial atmospheres. The number of planets with atmospheres—regardless of the composition—drops when we included the effects of photoevaporation, as expected. However, the effects of photoevaporation increases the fraction of bare rocky planets by only $\sim$10% and that of water-rich planets by only $\sim$2% (see Figure 4). The breaking-the-chains model and photoevaporation, or possibly other atmospheric mass-loss mechanisms, are not mutually exclusive processes but our results shows that giant impacts play the dominant role in sculpting the radius valley.

Figure 5 shows the final planetary systems produced in our simulations including the effects of photoevaporation. The innermost planets are typically rocky whereas water-rich mini-Neptunes tend to have relatively larger orbital periods (Izidoro et al. 2021a). This orbital arrangement is broadly consistent with observations (Millholland & Winn 2021). Figure 5 also shows that migration and dynamical instabilities promote a great diversity of planetary compositions. Some of the systems show, for instance, adjacent planets with distinct compositions as a rocky planet adjacent to a water-rich one or a rocky planet adjacent to a rocky planet with a primordial atmosphere. These systems are consistent with the orbital architecture of intriguing systems like Kepler-36 (Carter et al. 2012) and TOI-178 (Leleu et al. 2021; see example systems in Figure 5 and Raymond et al. 2018).

4. Discussion

Our model A, which was designed to produce planetary systems dominated by rocky planets/cores (Izidoro et al. 2021a), failed to reproduce the exoplanet radius valley. Model A was based on a specific set of initial conditions (see Section 2.1) but systems dominated by rocky planets—as Model A—can be achieved using different initial parameters in the model of Izidoro et al. (2021a). So, can we conclude that

the breaking-the-chains scenario is generally inconsistent with the exoplanet radius valley when planetary systems are dominated by planets/cores with rocky composition? The best way to address this question would be to perform an extensive exploration of the parameter space that defines our model. However, this is impossible due to the computational time required to perform the simulations. Nevertheless, we can try to solve this problem using a different approach, based on the following reasoning.

Let us first assume that as postulated by our model, all super-Earths are the result of collisions between mini-Neptunes. In this case, super-Earths should be more massive than mini-Neptunes, and this is in direct contrast with the observations (see Chen & Kipping 2017; Otegi et al. 2020). For instance, the "super-Earth" peak is observed at $\sim$1.4 $R_{\oplus}$ (Fulton & Petigura 2018), which corresponds to a mass of about $3.5M_{\oplus}$, assuming an Earth-like composition. The inner edge of the valley corresponds to a size of about $1.6R_{\oplus}$ and translates to a mass of $\sim 6M_{\oplus}$ (Zeng et al. 2019), also for an Earth-like composition. In the context of the breaking-the-chains model, rocky planets less massive than $\sim 6M_{\oplus}$ ($<1.6R_{\oplus}$) are envisioned to be the outcomes of one or more giant impacts that stripped primordial planetary atmospheres of mini-Neptunes. As one or two collisions are a common outcome of dynamical instabilities, it implies that mini-Neptunes should have masses in the range of $\lesssim$1–3$M_{\oplus}$ before collisions.

We have performed additional simulations to further investigate the plausibility of this scenario. We refer to this additional set of simulations as model D. These simulations were designed to produce exclusively rocky planets with masses relatively lower than those produced in model A. Before dynamical instabilities, the average mass of planets in this model is 1.9$M_{\oplus}$ (compare to 3.6$M_{\oplus}$ in model A). We adjusted some of the free parameters of our model in order to produce lower-mass planets in model D. In model D, Moon-mass planetary seeds were initially distributed between 0.2 and $\sim$1 au, $t_{\rm start}$ = 0.5 Myr, $R_{\rm peb}$ = 1 mm, and $S_{\rm peb}$ = 5. (compare with parameters of model A described in see Section 2.1).

Figure 6 shows the outcome of simulations of Model D after dynamical instabilities (as Figure 2). It shows that even when rocky planets are systematically less massive than those of model A, there is no clear valley in the planet size distribution at $\sim$1.8$R_{\oplus}$. Interestingly, the final planet-mass distribution of model D is broadly consistent with that predicted by photoevaporation models (Owen & Wu 2017). We have verified that—after
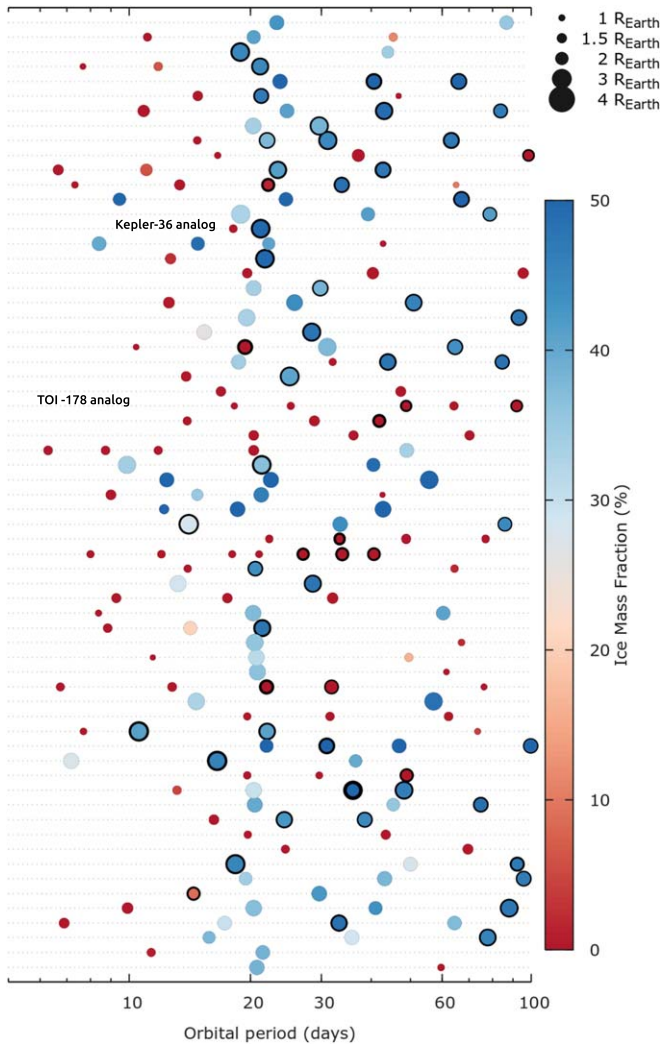
Figure 5. Final planetary systems produced in Model-C. Each line shows one planetary systems. The color coding indicates the planet's composition. A black ring around the dot indicates the presence of an atmosphere, and the full size of the dot scales with its radius. When determining the final planet sizes, we include atmospheric stripping by photoevaporation after the giant-impact phase. Planets' atmospheric masses, when present, are assumed to correspond to 0.3% of the core mass.

dynamical instabilities—model D's planet-mass distribution broadly corresponds to a Rayleigh distribution with a mode of $3M_\oplus$ (see Owen & Wu 2017). Yet, a valley in the radius distribution is not created in model D because giant impacts play the dominant role in stripping planetary atmospheres and filling/destroying the radius gap in our model. Note that this is different from standard photoevaporation and core-powered atmospheric mass-loss models (e.g., Owen & Wu 2017; Jin & Mordasini 2018; Gupta & Schlichting 2019). In these models, atmospheric accretion during planet formation is assumed to "fill" the radius valley before it can be sculpted by photoevaporation or cored-powered mass-loss effects. In model D (as in model A), planets populating the radius valley have no atmospheres to be stripped via photoevaporation or core-powered mass-loss effects because they were stripped via late giant impacts.

The last issue with model D is that most planets with radii larger than $\sim 2R_\oplus$ (dark-blue dots in the middle panel of Figure 6) have masses of ~ 1–$3M_\oplus$. This mass range conflicts with the exoplanet data, which suggest that mini-Neptunes typically have
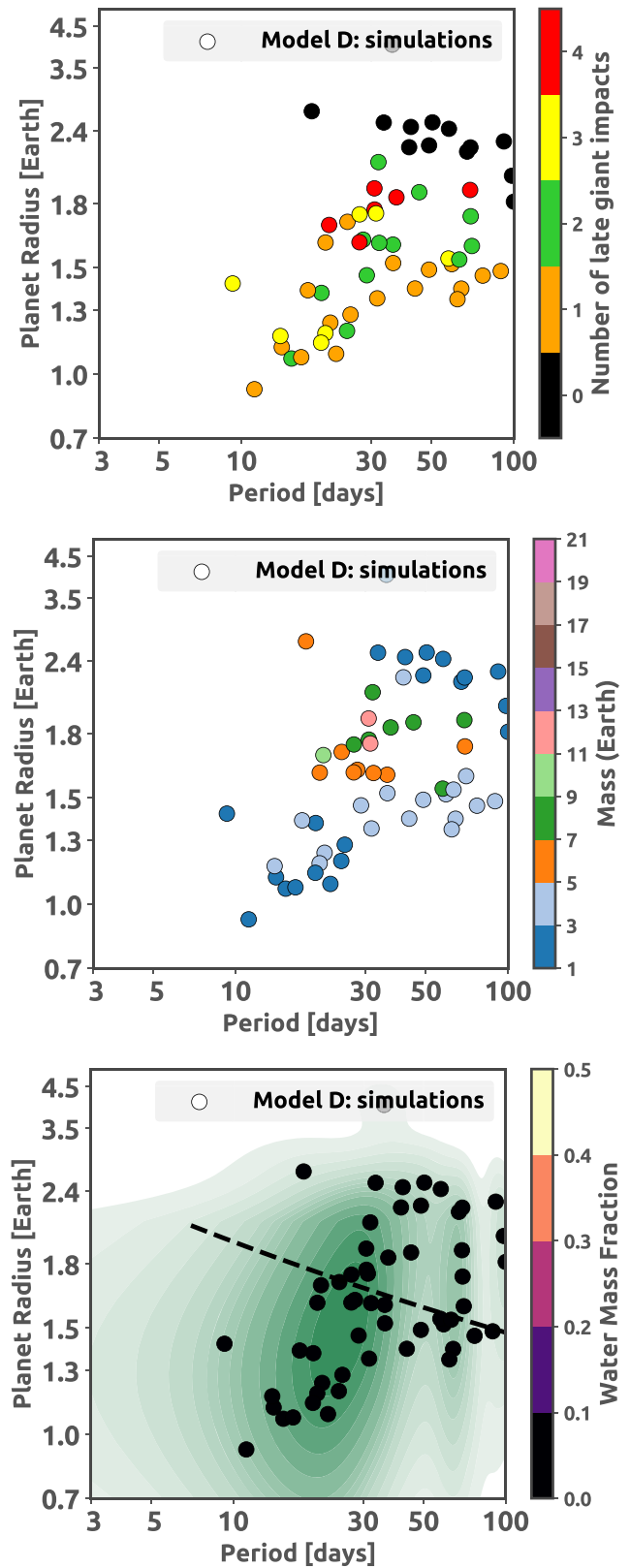
Figure 6. The same as Figure 2 but for model D.

masses larger than $5M_\oplus$ (e.g., Zeng et al. 2019; Otegi et al. 2020). Consequently, this scenario is not realistic on its own.

Another potential means of creating a radius valley with exclusively rocky planets in our model would require the

existence of two classes of rocky planets with hydrogen-rich atmospheres. As before, the first class—the progenitor of super-Earths—would consist of rocky planets with hydrogen-rich atmospheres and masses of $\sim$1–3 $M_\oplus$ planets. The second class of planets—the progenitors of mini-Neptunes—would consist of hydrogen-rich rocky planets with masses of $\gtrsim$6–9 $M_\oplus$. Before instabilities, both classes of planets have atmospheres and radii that put them about the radius valley. Collisions between 1 and 3 $M_\oplus$ planets would create bare rocky planets with masses between 2–6$M_\oplus$ and radii below 1.8 $R_\oplus$. These planets would therefore populate the super-Earth side of the radius valley (as in our previous scenario). Vice versa, collisions involving $\gtrsim$6–9$M_\oplus$ planets would produce bare rocky planets with masses $\gtrsim$12–18$M_\oplus$ and radii larger than 1.8 $R_\oplus$, which will be above the radius valley. In principle, one could assume that these two classes of planets could come from systems with two very distinct types of planetary system architectures. However, observations show that super-Earths and mini-Neptunes coexist in the same planetary systems (e.g., Carrera et al. 2018; Millholland & Winn 2021; Hawthorn et al. 2022) and have orbital configurations such that super-Earths are usually found closer in as compared to mini-Neptunes (e.g., Millholland & Winn 2021; Hawthorn et al. 2022). This implies that these envisioned populations of progenitors should also coexist in the same system, instead of making two distinct classes of planetary systems. None of the existing planet formation models self-consistently accounting for migration predict a strong dichotomy in mass and orbital radii for rocky planets in the same system (e.g., Ogihara et al. 2015; Lambrechts et al. 2019). Building planetary systems with such generic features—if possible—would require very specific and perhaps unrealistic assumptions on the distribution of rocky material in the inner parts of protoplanetary disks.

In light of these issues, we argue that if close-in exoplanets formed via disk migration and subsequent dynamical instabilities that lead to giant impacts—the breaking-the-chains evolution—then a large fraction of mini-Neptunes (and their cores) are water-rich planetary objects, as predicted by model C.

5. Conclusion

In this work, we have revisited the breaking-the-chains scenario for the formation of super-Earths and mini-Neptunes (Izidoro et al. 2017, 2021a) with the goal of testing if this model is consistent with the exoplanet radius valley (Fulton et al. 2017; Fulton & Petigura 2018) and peas-in-a-pod feature (Weiss et al. 2018; Weiss & Petigura 2020). We model the formation and dynamical evolution of planetary systems. We used simulations from Izidoro et al. (2021a) that include the effects of disk evolution, pebble accretion, gas-driven planet migration, eccentricity and inclination damping due to planet–gas tidal interactions, and mutual gravitational interactions of planetary embryos (see also Lambrechts et al. 2019 and Bitsch et al. 2019a). We selected three different setups from that study, which produced planetary systems dominated by rocky planets, water-rich planets, and planets of mixed compositions (water rich and rocky). We assume that at the end of the gas-disk-phase planets have atmospheres corresponding to 0.1%, 0.3%, 1%, or 5% of their masses, as suggested by observations (e.g., Lopez & Fortney 2013; Zeng et al. 2021), numerical simulations (e.g., Lambrechts & Lega 2017; Moldenhauer et al. 2022), and analytical atmospheric accretion models (e.g., Ginzburg et al. 2016).

The breaking-the-chains model proposes that after gas disk dispersal, more than 90%–95% of the planetary systems become dynamically unstable, which leads to a phase of giant impacts. We assume that late giant impacts strip primordial atmospheres of low-mass planets (Biersteker & Schlichting 2019). By using the MRR from Zeng et al. (2016, 2019) for different compositions and equilibrium temperatures, we show that the breaking-the-chains model is consistent with the planet-radius distribution, which has peaks at $\sim$1.4 $R_\oplus$ and $\sim$2.4 $R_\oplus$, and a valley at $\sim$1.8$R_\oplus$. Our model is also consistent with the peas-in-a-pod feature (Weiss et al. 2018; Weiss & Petigura 2020). We also tested our model using the empirical MRR of Otegi et al. (2020), and the results do not qualitatively change (see the Appendix).

Our results do not support an exclusively rocky composition for the cores of mini-Neptunes. A rocky composition—for super-Earths and mini-Neptune cores—is favored by photo-evaporation (Owen & Wu 2013; Mordasini 2014) and core-powered atmospheric mass-loss models (Gupta & Schlichting 2019). Instead, we predict that planets larger than $\sim$2 $R_\oplus$ (mini-Neptunes or their cores) are mostly water rich (>10% water by mass) and planets smaller than $\sim$1.6 $R_\oplus$ (super-Earths) are mostly rocky. We also suggest that orbital instabilities and late giant impacts are the dominant processes sculpting the orbital architecture of super-Earths and mini-Neptunes. Photoevaporation or other subsequent atmospheric mass loss processes play a minor role if giant impacts after gas disk dispersal are common as suggested by our model (see the Appendix).

Our results suggest that planet formation starts early (e.g., <0.5–1 Myr) in the inner disk. This is consistent with models of dust coagulation and planetesimal formation in the solar system (e.g., Izidoro et al. 2021b; Morbidelli et al. 2022) and the estimated ages of meteorites (e.g., Kruijer et al. 2017). Our results also suggest that planet formation beyond the snow line may not be as efficient as currently thought (e.g., Drążkowska & Alibert 2017), or large-range inward migration is slower or less efficient than usually considered (e.g., Paardekooper et al. 2011). Indeed, if planets beyond the snow line grow quickly, they migrate inward and destroy the inner rocky systems (see Model B and Izidoro et al. 2014) or become gas giants (Bitsch et al. 2019a, 2020). This is inconsistent with observations. Finally, we predict that a fraction of planets larger than $\sim$2$R_\oplus$ should be water rich and have a primordial H-rich atmosphere. This prediction may be tested by future James Webb Space Telescope observations.

Appendix
## Simulated Transit Observations and Complementary Results

In order to further compare our results to observations, we conduct synthetic transit observations of our planetary systems
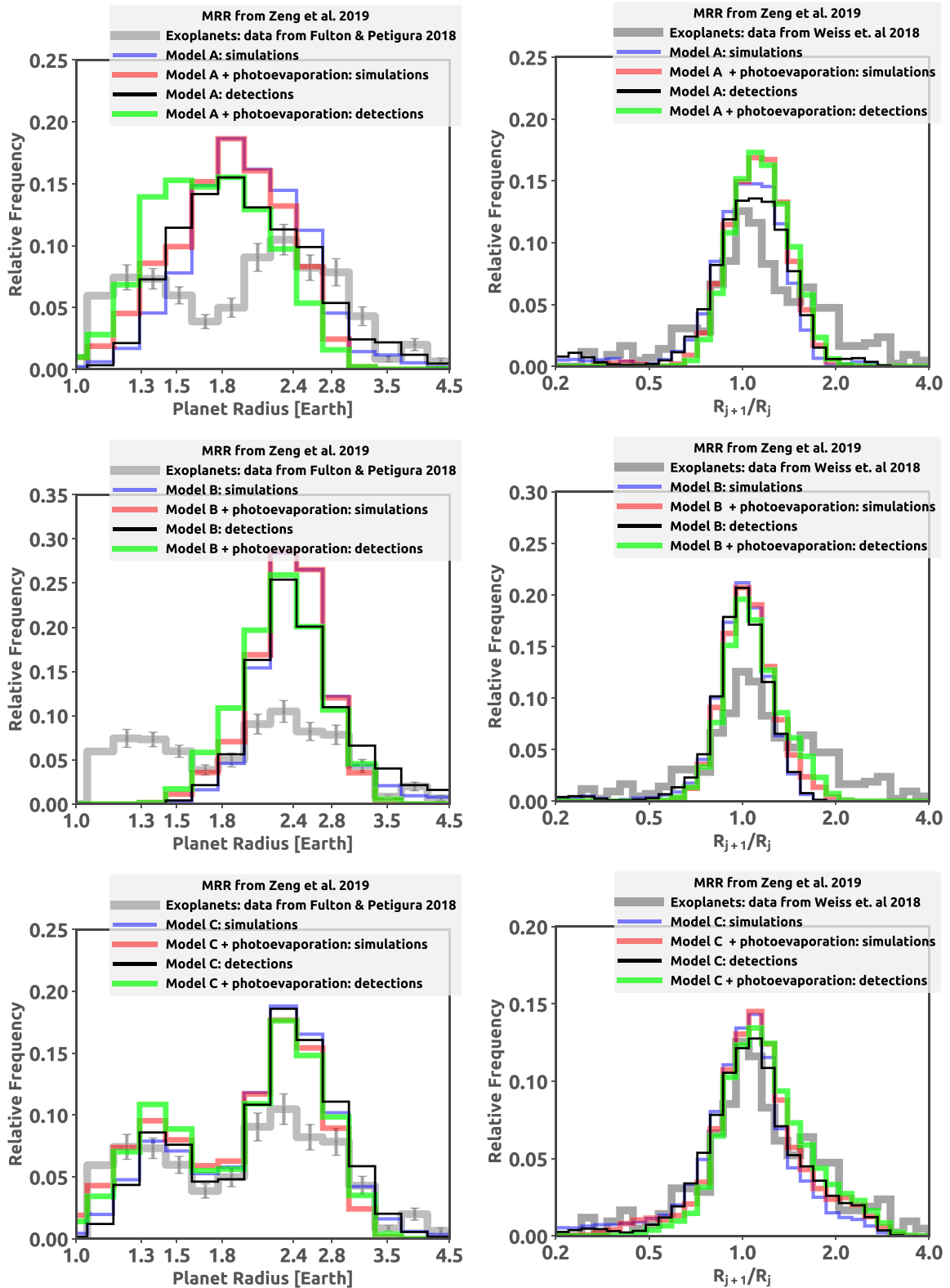
Figure A1. Left: planet-radius distribution. Right: planet size-ratio distribution of adjacent planet pairs. Observations are shown in gray. Blue shows the outcome of our planet-formation simulations. Red shows the outcome of our planet-formation simulations modeling the effects of photoevaporation. Black shows the synthetic transit observations of our simulations. Green shows the synthetic transit observations when we include the effects of photoevaporation. From top-to-bottom, the panel-rows show models A, B, and C. For all models, we use the MRR of Zeng et al. (2019) assuming an uncertainty of 7% in size and an initial atmosphere-to-core-mass ratio of 0.3%.

by following the procedure discussed in Izidoro et al. (2021a). In brief, each system is observed from different lines of sight characterized by inclination angles evenly spaced by 0°.1 from $-20°$ to $20°$ relative to the plane of the primordial gas disk. Viewing angles in the azimuthal direction are evenly spaced by 1° from 0° to 360°. For each line of sight, we check which
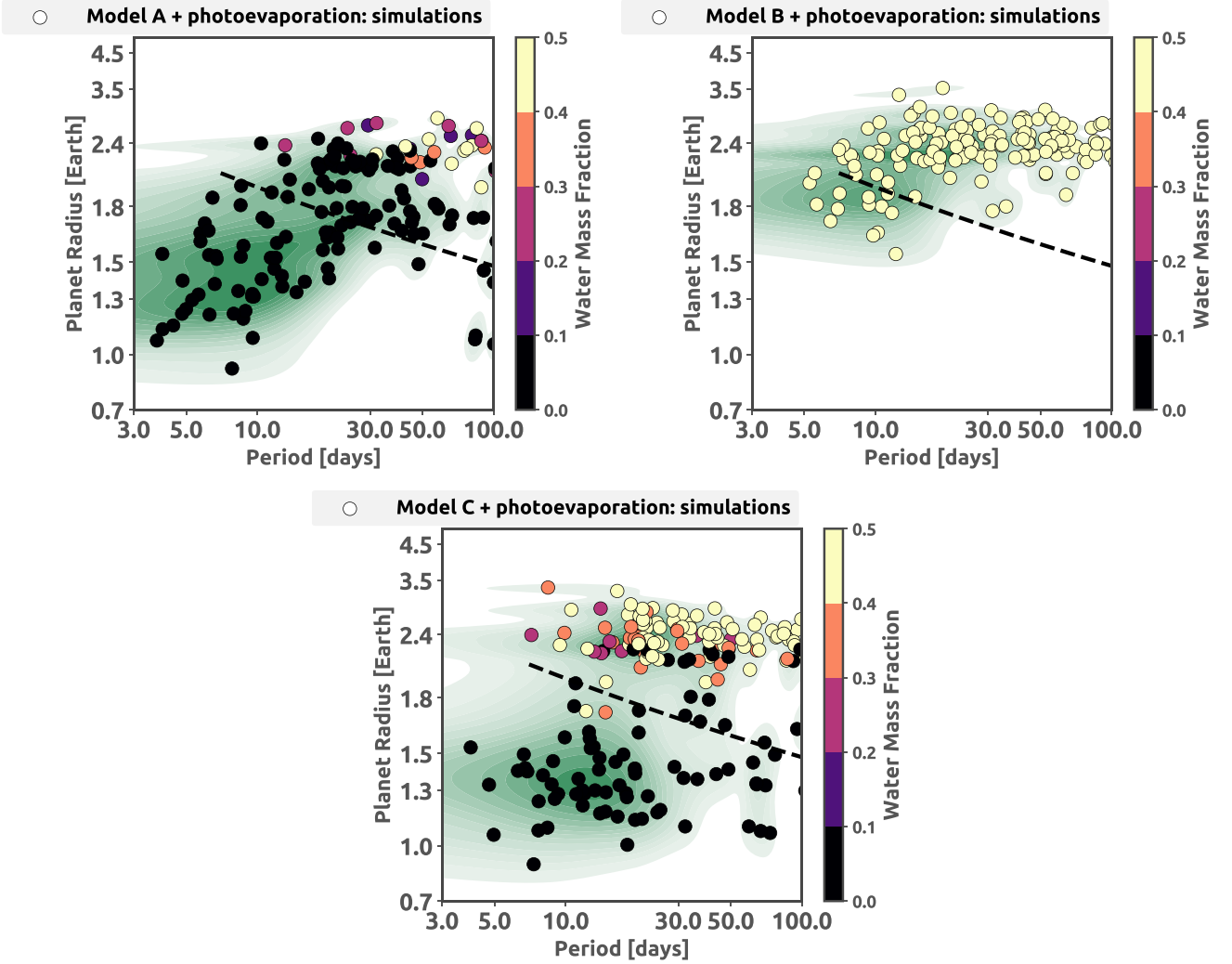
Figure A2. Planetary architecture of our systems at the end of the simulations when including the effects of photoevaporation. The assumed atmosphere-to-core-mass fraction is 0.3%. As in our nominal analysis, the MRR used to convert masses to planet radii comes from Zeng et al. (2019).

planets transit in front of their host star assuming 3.5 yr long observations (as in the case of Kepler observations; Weiss & Petigura 2020) and use these to create a list of detected planets. For example, let us assume that star A is orbited by planets b, c, and d, which have different orbital inclinations. If only planet b is observed to transit when the system is observed along a specific line of sight, then we add the system composed of A and b to our new list. Then, if planets c and d are observed to transit when the system is observed from a different point of view, we add the system composed of A, c, and d to the list of detected planets as well. In this way, one synthetic planetary system can result in many different observed planetary systems. In order to be considered detected, the signal-to-noise ratio (S/N) of a transiting planet must be S/N > 10. We calculate the S/N as

$$S/N = \frac{(R_p/R_\odot)^2 \sqrt{3.5 \text{ yr}/P}}{CDPP_{6hr}\sqrt{6 \text{ hr}/T}}, \qquad (3)$$

where T is the transit duration; $CDPP_{6hr}$ is the 6 hr Combined Differential Photometric Precision; a measurement of the stellar noise level (Christiansen et al. 2012); and R_p and P are the orbital physical radius and orbital period of the transiting planet. This equation takes into account the fact that most

Kepler stars were continuously observed for 3.5 yr. This simulator algorithm is more sophisticated than that used in Izidoro et al. (2017, 2021a), where detection relies on geometric transit only.

In our N-body numerical simulations, the central star is always a solar-type star ($R_\star = R_\odot$ and $\rho_\star = \rho_\odot$), and we do not need to make any assumption about the stellar photometric noise. However, to perform our synthetic transit observations, we do. We randomly pick $CDPP_{6hr}$ from direct measurements of the photometric noise for the CKS stellar sample (Weiss et al. 2018).

For every observed synthetic planetary system, we randomly assign one star from the CKS sample and take its corresponding $CDPP_{6hr}$ to compute the S/N for each transiting planet. Finally, the transit duration is calculated as

$$T = 13 \text{ hr}\left(\frac{P}{1 \text{ yr}}\right)^{1/3}\left(\frac{\rho_\star}{\rho_\odot}\right)^{-1/3}\sqrt{1 - b^2}, \qquad (4)$$

where b is the impact parameter.

Figure A1 shows the planet-radius distribution (left) and the planet size-ratio distribution for adjacent planets (right) of our simulations, synthetic detections, and exoplanets in the CKS
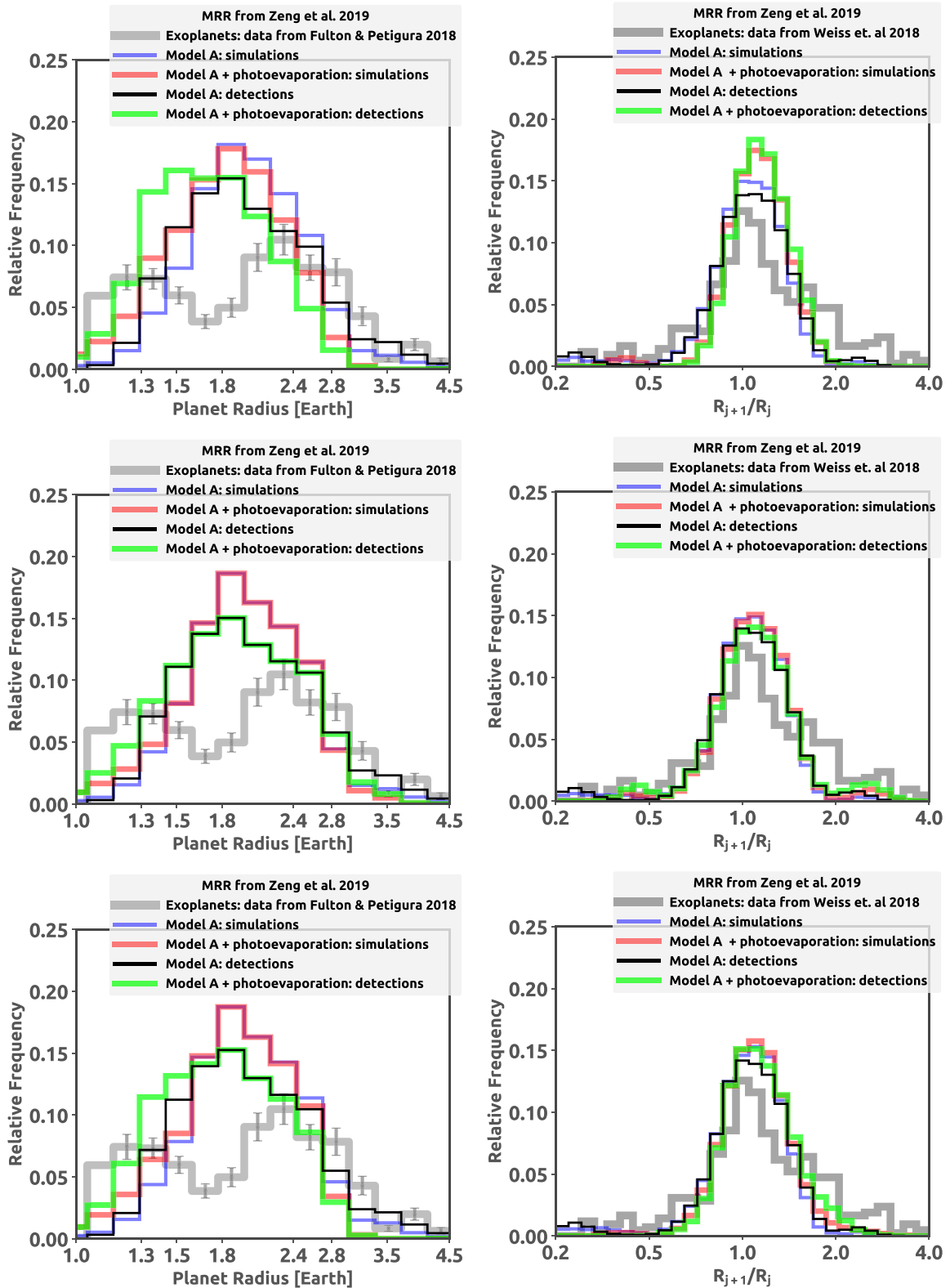
Figure A3. Planet-radius distribution (left) and planet size-ratio distribution (right) of model A assuming an initial atmospheric mass fractions of 0.1% (top panels), 1% (middle panels), and 5% (bottom panels). Observations are shown in gray. Blue shows the outcome of our planet-formation simulations. Red shows the outcome of our planet-formation simulations modeling the effects of photoevaporation. Black shows the synthetic transit observations of our simulations. Green shows the synthetic transit observations of our simulations including the effects of photoevaporation.

sample (Fulton & Petigura 2018; Weiss et al. 2018). In order to construct these two histograms (blue and red), we also account for uncertainties of 7% in planet size (see Section 2.3).

As also shown in the main paper, models A and B fail to reproduce the bimodal distribution of planet radii, even when observational biases are taken into account. In model A,
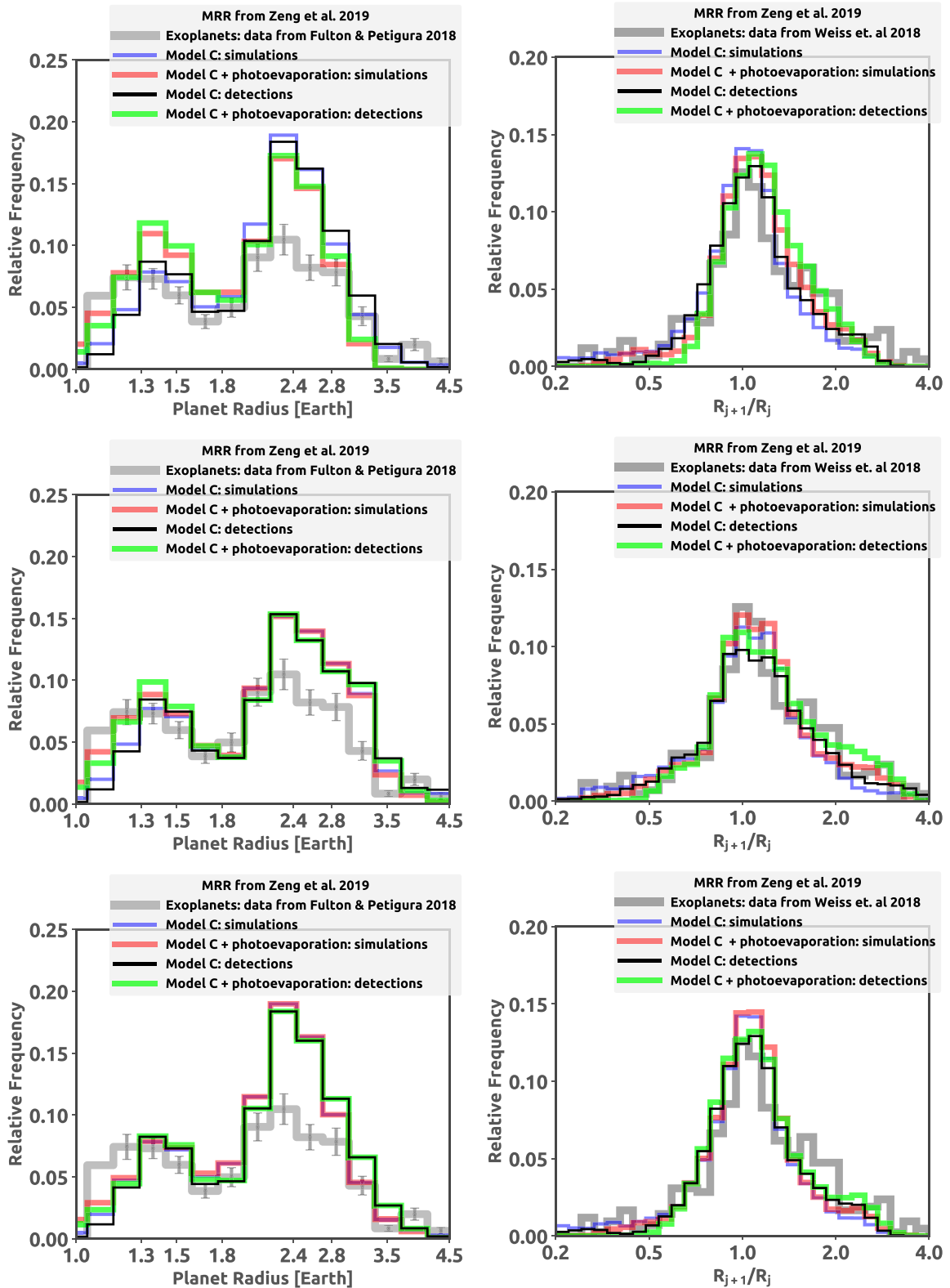
Figure A4. Planet-radius distribution (left) and planet size-ratio distribution (right) of model C assuming an initial atmospheric mass fractions of 0.1% (top panels), 1% (middle panels), and 5% (bottom panels). Observations are shown in gray. Blue shows the outcome of our planet-formation simulations. Red shows the outcome of our planet-formation simulations modeling the effects of photoevaporation. Black shows the synthetic transit observations of our simulations. Green shows the synthetic transit observations of our simulations including the effects of photoevaporation.

simulated detections that include the effect of photoevaporation (green histogram) have a flat-top distribution of planetary radii across the planet valley region (between 1.3 and 1.8 $R_{\oplus}$) but overpredict the relative frequency of these planets. In models B and C, we find that the inclusion of photoevaporation has a small effect on the distribution of planet radii.
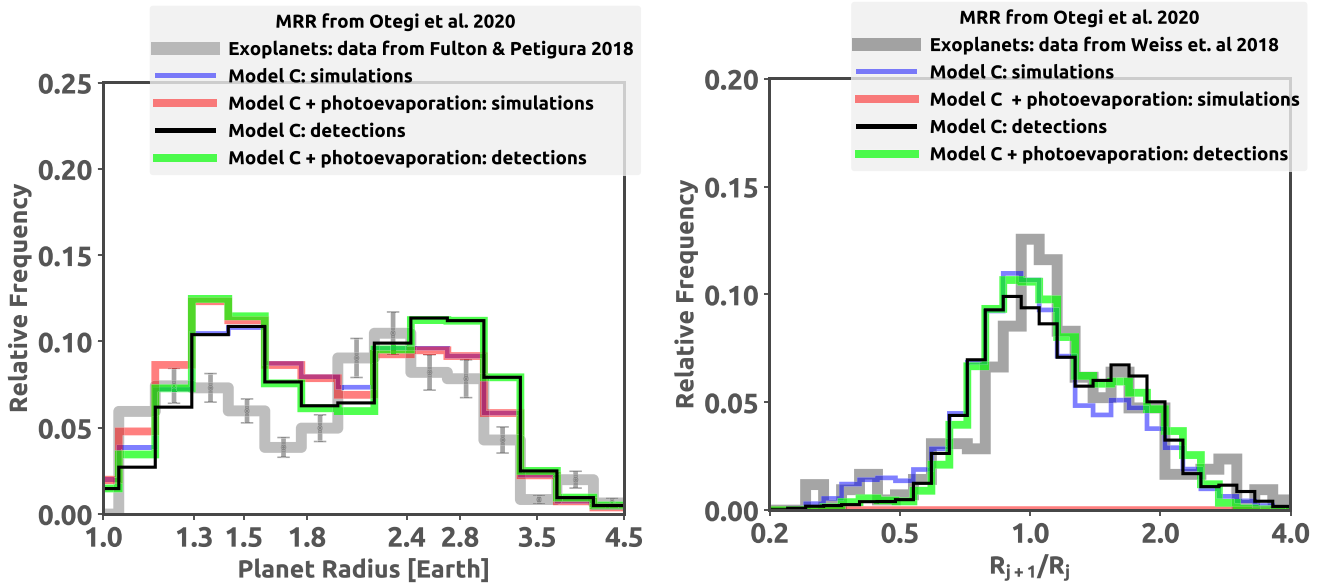
Figure A5. Planet-radius distribution (left) and planet size-ratio distribution (right) of model C assuming an initial atmospheric mass fraction of 0.3%. Planet radii are computed using the empirical MRR of Otegi et al. (2020). We use the high-density MRR equation of Otegi et al. (2020) to compute the sizes of rocky planets that experienced late impacts and the low-density equation for planets with water-rich compositions or rocky cores with atmospheres (avoided impacts). Observations are shown in gray. Blue shows the outcome of our planet-formation simulations. Red shows the outcome of our planet-formation simulations modeling the effects of photoevaporation. Black shows the synthetic transit observations of our simulations. Green shows the synthetic transit observations of our simulations including the effects of photoevaporation.
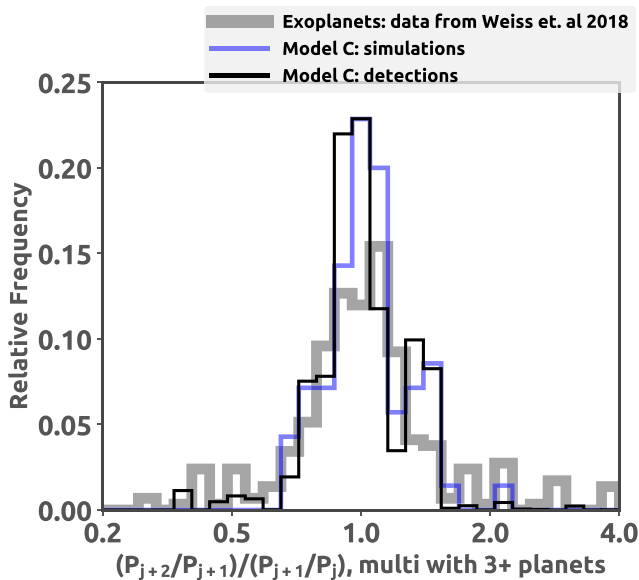
Figure A6. Ratio of orbital period ratios for adjacent triple of planets in the same planetary systems. Gray shows exoplanet data from Weiss et al. (2018). Blue and black show the results of our model C.

Figure A2 shows the final planetary architecture of our systems when we include the effects of photoevaporation. As one can see, the trends shown in Figure 2—where we do not include the effects of photoevaporation—do not change qualitatively.

Figures A3 and A4 show the planet-radius distribution and planet size-ratio distribution of adjacent planets when we assume that primordial atmospheres correspond to 0.1%, 1%, and 5% of the planet/core mass. The radius valley and peas-in-a-pod feature are matched well for model C but not for model A, regardless of the assumed atmosphere mass.

Figure A5 shows the planet-radius distribution and planet size-ratio distribution of adjacent planets when we assume the empirical MRR of Otegi et al. (2020). We assume that giant impacts make "high-density" rocky planets, and we used the MRR for high-density planets from Otegi et al. (2020). We use the low-density MRR to compute the radius of water-rich planets or water-rich planets with primordial atmospheres regardless of the occurrence of late giant impacts (compare with the bottom panel of Figure 3 and Figure A4). Compared to the results in the main paper, these two mass–radius determinations give very similar results in broadly matching the radius valley.

Figure A6 shows the ratio of orbital period ratios of adjacent triple planets in the same planetary system. This figure shows that our model C is also broadly consistent with the regular orbital spacing of observed adjacent planet pairs (Weiss et al. 2018).

ORCID iDs

André Izidoro ⓘ https://orcid.org/0000-0003-1878-0634
Hilke E. Schlichting ⓘ https://orcid.org/0000-0002-0298-8089
Andrea Isella ⓘ https://orcid.org/0000-0001-8061-2207
Rajdeep Dasgupta ⓘ https://orcid.org/0000-0001-5392-415X
Bertram Bitsch ⓘ https://orcid.org/0000-0002-8868-7649

References

Adams, E. R., Seager, S., & Elkins-Tanton, L. 2008, ApJ, 673, 1160
Baruteau, C., Crida, A., Paardekooper, S.-J., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al. (Tucson, AZ: Univ. of Arizona Press), 667
Bashi, D., Helled, R., Zucker, S., & Mordasini, C. 2017, A&A, 604, A83
Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013, ApJS, 204, 24
Bean, J. L., Raymond, S. N., & Owen, J. E. 2021, JGRE, 126, e06639
Béthune, W., & Rafikov, R. R. 2019, MNRAS, 488, 2365
Biersteker, J. B., & Schlichting, H. E. 2019, MNRAS, 485, 4454
Biersteker, J. B., & Schlichting, H. E. 2021, MNRAS, 501, 587
Bitsch, B. 2019, A&A, 630, A51
Bitsch, B., Izidoro, A., Johansen, A., et al. 2019a, A&A, 623, A88
Bitsch, B., Morbidelli, A., Johansen, A., et al. 2018, A&A, 612, A30

Bitsch, B., Raymond, S. N., & Izidoro, A. 2019b, A&A, 624, A109
Bitsch, B., Trifonov, T., & Izidoro, A. 2020, A&A, 643, A66
Bolmont, E., & Mathis, S. 2016, CeMDA, 126, 275
Carrera, D., Ford, E. B., & Izidoro, A. 2019, MNRAS, 486, 3874
Carrera, D., Ford, E. B., Izidoro, A., et al. 2018, ApJ, 866, 104
Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, Sci, 337, 556
Chance, Q., Ballard, S., & Stassun, K. 2022, ApJ, 937, 39
Chatterjee, S., & Ford, E. B. 2015, ApJ, 803, 33
Chen, J., & Kipping, D. 2017, ApJ, 834, 17
Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. 2012, PASP, 124, 1279
Cimerman, N. P., Kuiper, R., & Ormel, C. W. 2017, MNRAS, 471, 4662
Coleman, G. A. L., & Nelson, R. P. 2014, MNRAS, 445, 479
Coleman, G. A. L., & Nelson, R. P. 2016, MNRAS, 457, 2480
Cossou, C., Raymond, S. N., Hersant, F., & Pierens, A. 2014, A&A, 569, A56
Cresswell, N. R. P. 2008, A&A, 482, 677
Cresswell, P., & Nelson, R. P. 2006, A&A, 450, 833
Dorn, C., Khan, A., Heng, K., et al. 2015, A&A, 577, A83
Drążkowska, J., & Alibert, Y. 2017, A&A,, 608, A92
Esteves, L., Izidoro, A., Bitsch, B., et al. 2022, MNRAS, 509, 2856
Esteves, L., Izidoro, A., Raymond, S. N., & Bitsch, B. 2020, MNRAS, 497, 2493
Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, ApJ, 790, 146
Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, ApJ, 659, 1661
Fressin, F., Torres, G., Charbonneau, D., et al. 2013, ApJ, 766, 81
Fulton, B. J., & Petigura, E. A. 2018, AJ, 156, 264
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Natur, 542, 456
Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, ApJ, 825, 29
Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, MNRAS, 476, 759
Goldberg, M., & Batygin, K. 2022, AJ, 163, 201
Gupta, A., Nicholson, L., & Schlichting, H. E. 2022, MNRAS, 516, 4585
Gupta, A., & Schlichting, H. E. 2019, MNRAS, 487, 24
Gupta, A., & Schlichting, H. E. 2020, MNRAS, 493, 792
Hawthorn, F., Bayliss, D., Wilson, T. G., et al. 2022, arXiv:2208.07328
He, M. Y., Ford, E. B., & Ragozzine, D. 2019, MNRAS, 490, 4575
He, M. Y., Ford, E. B., & Ragozzine, D. 2021, AJ, 162, 216
Hellary, P., & Nelson, R. P. 2012, MNRAS, 419, 2737
Howard, A. W. 2013, Sci, 340, 572
Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, ApJS, 201, 15
Ida, S., & Lin, D. N. C. 2008, ApJ, 685, 584
Ida, S., & Lin, D. N. C. 2010, ApJ, 719, 810
Inamdar, N. K., & Schlichting, H. E. 2016, ApJL, 817, L13
Izidoro, A., Bitsch, B., Raymond, S. N., et al. 2021a, A&A, 650, A152
Izidoro, A., Dasgupta, R., Raymond, S. N., et al. 2021b, NatAs, 6, 357
Izidoro, A., Morbidelli, A., & Raymond, S. N. 2014, ApJ, 794, 11
Izidoro, A., Ogihara, M., Raymond, S. N., et al. 2017, MNRAS, 470, 1750
Jin, S., & Mordasini, C. 2018, ApJ, 853, 163
Johansen, A., & Lambrechts, M. 2017, AREPS, 45, 359
Johnson, J. A., Petigura, E. A., Fulton, B. J., et al. 2017, AJ, 154, 108
Kruijer, T. S., Burkhardt, C., Budde, G., & Kleine, T. 2017, PNAS, 114, 6712
Kuchner, M. J. 2003, ApJL, 596, L105
Kurosaki, K., Ikoma, M., & Hori, Y. 2014, A&A, 562, A80
Lai, D., & Pu, B. 2017, AJ, 153, 42
Lambrechts, M., & Johansen, A. 2012, A&A, 544, A32
Lambrechts, M., Johansen, A., & Morbidelli, A. 2014, A&A, 572, A35
Lambrechts, M., & Lega, E. 2017, A&A, 606, A146
Lambrechts, M., Morbidelli, A., Jacobson, S. A., et al. 2019, A&A, 627, A83
Lee, E. J., & Chiang, E. 2015, ApJ, 811, 41
Leinhardt, Z. M., & Stewart, S. T. 2012, ApJ, 745, 79
Leleu, A., Alibert, Y., Hara, N. C., et al. 2021, A&A, 649, A26
Levison, H. F., Morbidelli, A., Tsiganis, K., Nesvorný, D., & Gomes, R. 2011, AJ, 142, 152
Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011, ApJS, 197, 8
Liu, S.-F., Hori, Y., Lin, D. N. C., & Asphaug, E. 2015, ApJ, 812, 164

Lopez, E. D., & Fortney, J. J. 2013, ApJ, 776, 2
Lopez, E. D., & Fortney, J. J. 2014, ApJ, 792, 1
Lopez, E. D., Fortney, J. J., & Miller, N. 2012, ApJ, 761, 59
Luger, R., Barnes, R., Lopez, E., et al. 2015, AsBio, 15, 57
Luger, R., Sestovic, M., Kruse, E., et al. 2017, NatAs, 1, 0129
Marcus, R. A., Sasselov, D., Hernquist, L., & Stewart, S. T. 2010, ApJL, 712, L73
Marcy, G. W., Weiss, L. M., Petigura, E. A., et al. 2014, PNAS, 111, 12655
Matsumoto, Y., & Ogihara, M. 2020, ApJ, 893, 43
Mayor, M., Marmier, M., Lovis, C., et al. 2011, arXiv:1109.2497
McNeil, D. S., & Nelson, R. P. 2010, MNRAS, 401, 1691
Millholland, S. C., & Winn, J. N. 2021, ApJL, 920, L34
Mills, S. M., Fabrycky, D. C., Migaszewski, C., et al. 2016, Natur, 533, 509
Misener, W., & Schlichting, H. E. 2021, MNRAS, 503, 5658
Moldenhauer, T. W., Kuiper, R., Kley, W., & Ormel, C. W. 2022, A&A, 661, A142
Morbidelli, A., Baillié, K., Batygin, K., et al. 2022, NatAs, 6, 72
Morbidelli, A., Lambrechts, M., Jacobson, S., & Bitsch, B. 2015, Icar, 258, 418
Mordasini, C. 2014, A&A, 572, A118
Mordasini, C. 2020, A&A, 638, A52
Mousis, O., Deleuil, M., Aguichine, A., et al. 2020, ApJL, 896, L22
Mulders, G. D. 2018, Planet Populations as a Function of Stellar Properties (Berlin: Springer), 153
Mulders, G. D., Pascucci, I., Apai, D., & Ciesla, F. J. 2018, AJ, 156, 24
Nesvorný, D., & Morbidelli, A. 2012, AJ, 144, 117
Ogihara, M., Kokubo, E., Suzuki, T. K., & Morbidelli, A. 2018, A&A, 615, A63
Ogihara, M., Morbidelli, A., & Guillot, T. 2015, A&A, 578, A36
Ormel, C. W., & Klahr, H. H. 2010, A&A, 520, A43
Otegi, J. F., Bouchy, F., & Helled, R. 2020, A&A, 634, A43
Owen, J. E., & Campos Estrada, B. 2020, MNRAS, 491, 5287
Owen, J. E., & Wu, Y. 2013, ApJ, 775, 105
Owen, J. E., & Wu, Y. 2017, ApJ, 847, 29
Paardekooper, S.-J., Baruteau, C., & Kley, W. 2011, MNRAS, 410, 293
Petigura, E. A. 2020, AJ, 160, 89
Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, PNAS, 110, 19273
Petigura, E. A., Howard, A. W., & Marcy, G. W., et al. 2017, AJ, 154, 107
Petigura, E. A., Rogers, J. G., Isaacson, H., et al. 2022, AJ, 163, 179
Poon, S. T. S., Nelson, R. P., Jacobson, S. A., & Morbidelli, A. 2020, MNRAS, 491, 5595
Pu, B., & Wu, Y. 2015, ApJ, 807, 44
Raymond, S. N., Boulet, T., Izidoro, A., Esteves, L., & Bitsch, B. 2018, MNRAS, 479, L81
Raymond, S. N., Izidoro, A., Bolmont, E., et al. 2022, NatAs, 6, 80
Raymond, S. N., Izidoro, A., & Morbidelli, A. 2020, in Planetary Astrobiology, ed. V. S. Meadows et al. (Tucson, AZ: Univ. of Arizona Press), 287
Rogers, J. G., Gupta, A., Owen, J. E., & Schlichting, H. E. 2021, MNRAS, 508, 5886
Rogers, L. A., & Seager, S. 2010, ApJ, 712, 974
Spalding, C., & Batygin, K. 2016, ApJ, 830, 5
Terquem, C., & Papaloizou, J. C. B. 2007, ApJ, 654, 1110
Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, MNRAS, 479, 4786
Van Eylen, V., Albrecht, S., Huang, X., et al. 2019, AJ, 157, 61
Venturini, J., Guilera, O. M., Haldemann, J., Ronco, M. P., & Mordasini, C. 2020, A&A, 643, L1
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
Weiss, L. M., Millholland, S. C., Petigura, E. A., et al. 2022, arXiv:2203.10076
Weiss, L. M., & Petigura, E. A. 2020, ApJL, 893, L1
Wolfgang, A., Rogers, L. A., & Ford, E. B. 2016, ApJ, 825, 19
Zeng, L., Jacobsen, S. B., Hyung, E., et al. 2021, ApJ, 923, 247
Zeng, L., Jacobsen, S. B., Sasselov, D. D., et al. 2019, PNAS, 116, 9723
Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127
Zhang, M., Knutson, H. A., Dai, F., et al. 2022, arXiv:2207.13099
Zhu, W., Petrovich, C., Wu, Y., Dong, S., & Xie, J. 2018, ApJ, 860, 101