

# Planetary evolution with atmospheric photoevaporation

## I. Analytical derivation and numerical study of the evaporation valley and transition from super-Earths to sub-Neptunes

C. Mordasini

<sup>1</sup> Physikalisches Institut, University of Bern, Gesellschaftsstrasse 6, 3012 Bern, Switzerland  
e-mail: christoph.mordasini@space.unibe.ch

Received 25 March 2019 / Accepted 5 February 2020

### ABSTRACT

**Context.** Observations have revealed in the *Kepler* data a depleted region separating smaller super-Earths from larger sub-Neptunes. This can be explained as an evaporation valley between planets with and without H/He that is caused by atmospheric escape.

**Aims.** We want to analytically derive the valley's locus and understand how it depends on planetary properties and stellar X-ray and ultraviolet (XUV) luminosity. We also want to derive constraints for planet formation models.

**Methods.** First, we conducted numerical simulations of the evolution of close-in low-mass planets with H/He undergoing escape. We performed parameter studies with grids in core mass and orbital separation, and we varied the postformation H/He mass, the strength of evaporation, and the atmospheric and core composition. Second, we developed an analytical model for the valley locus.

**Results.** We find that the bottom of the valley quantified by the radius of the largest stripped core,  $R_{\text{bare}}$ , at a given orbital distance depends only weakly on postformation H/He mass. The reason is that a high initial H/He mass means that more gas needs to evaporate, but also that the planet density is lower, increasing mass loss. Regarding the stellar XUV-luminosity,  $R_{\text{bare}}$  is found to scale as  $L_{\text{XUV}}^{0.135}$ . The same weak dependency applies to the efficiency factor  $\varepsilon$  of energy-limited evaporation. As found numerically and analytically,  $R_{\text{bare}}$  varies a function of orbital period  $P$  for a constant  $\varepsilon$  as  $P^{-2p_c/3} \approx P^{-0.18}$ , where  $M_c \propto R_c^{p_c}$  is the mass-radius relation of solid cores. We note that  $R_{\text{bare}}$  is about  $1.7 R_{\oplus}$  at a ten-day orbital period for an Earth-like composition.

**Conclusions.** The numerical results are explained very well with the analytical model where complete evaporation occurs if the temporal integral over the stellar XUV irradiation that is absorbed by the planet is larger than the binding energy of the envelope in the gravitational potential of the core. The weak dependency on the postformation H/He means that the valley does not strongly constrain gas accretion during formation. But the weak dependency on primordial H/He mass, stellar  $L_{\text{XUV}}$ , and  $\varepsilon$  could be the reason why the valley is so clearly visible observationally, and why various models find similar results theoretically. At the same time, given the large observed spread of  $L_{\text{XUV}}$ , the dependency on it is still strong enough to explain why the valley is not completely empty.

**Key words.** planetary systems – planets and satellites: formation – planets and satellites: interiors – planets and satellites: atmospheres – planets and satellites: composition

## 1. Introduction

An intriguing result of the starting geophysical characterization of extrasolar planets is the large diversity in mean densities of low-mass and intermediate-mass planets. This has been detected observationally in the last few years (e.g., Weiss & Marcy 2014; Jontof-Hutter et al. 2016; Guenther et al. 2017). At planetary masses of about 1–20 Earth masses, observations suggest a variation of more than one order in magnitude in mean density (e.g., Hatzes & Rauer 2015; Jin & Mordasini 2018). This implies that the planetary bulk compositions must vary widely from rocky, Earth-like interiors for dense planets, such as CoRoT-7b or Kepler-93b (Léger et al. 2009; Dressing et al. 2015), to low-mass planets that contain significant amounts of H/He, such as the Kepler-11 planets (Lissauer et al. 2013). The analysis and interpretation of this transition from solid Earth-like rocky super-Earth (and potentially icy planets) to low-mass sub-Neptune planets with primordial H/He has subsequently attracted a lot of attention on the theoretical statistical side (e.g., Rogers 2015; Wolfgang & Lopez 2015; Chen & Kipping 2017).

In order to understand these planetary density measurements at the present day, it is necessary to understand how the planets

have evolved in time in the past. For this purpose, numerical models were developed that combine the long-term postformation thermodynamical evolution (cooling and contraction) of low-mass planets with atmospheric escape (Owen & Wu 2013; Lopez & Fortney 2013; Jin et al. 2014; Chen & Rogers 2016). Atmospheric escape is a process that has been observed for several types of planets, such as Hot Jupiters (e.g., Vidal-Madjar et al. 2003) or close-in Neptunian planets (e.g., Ehrenreich et al. 2015; Bourrier et al. 2018).

These theoretical models were then used in parameter studies to assess the effect of evaporation as a function of planet mass and orbital distance. This shows that the evaporation of primordial H/He envelopes is a process of prime importance for shaping the radii of typical *Kepler* planets, that is, close-in, low-mass planets with radii smaller than about  $4 R_{\oplus}$  at orbital distances of less than a few 0.1 AU. Even more remarkably, despite the different focus and independency of the aforementioned theoretical models, all of these studies similarly predicted in a rather rare congruence of theoretical models that on a population-level, atmospheric escape should lead to the following characteristic imprint: a depletion of the number of planets in a specific region of the orbital distance – planet radius plane that runs

diagonally downward with increasing distance. This feature was called the evaporation valley by [Jin et al. \(2014\)](#). The corresponding 1D radius distribution was found to be bimodal, with a local minimum separating smaller super-Earth planets that have lost all H/He from larger sub-Neptune planets that kept some of it. This also defines these two previously blurred planet types more clearly.

As pointed out by [Owen & Wu \(2013\)](#), the *Kepler* radius distribution known in 2013 already contained a hint of a bimodality compatible with the theoretically predicted one, but its significance was unclear, and the minimum was much less prominent than the deep minimum predicted theoretically, for example in [Jin et al. \(2014, their Figs. 13 and 14\)](#). Observationally, a limiting factor at that time were the relatively high uncertainties in the planetary radii because of poorly constrained host star properties in the original *Kepler* input catalog. The interest was therefore significant when [Fulton et al. \(2017\)](#) showed that for better-constrained stellar parameters ([Petigura et al. 2017; Johnson et al. 2017](#)), there is indeed a deep valley and minimum also in the observational data.

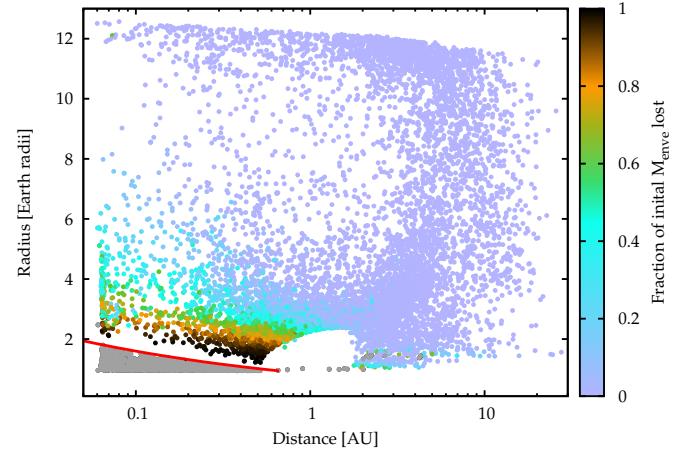
[Owen & Wu \(2017\)](#) and [Jin & Mordasini \(2018\)](#) then showed that the observed locus of the radius valley and the corresponding minimum in the 1D radius distribution not only agree with the previously theoretically predicted evaporation valley, but that the locus even shows that the dominant composition of the cores should be Earth-like (iron and silicates) without much ice, a strong constraint for formation models. Thus, this represents one of the not so numerous cases where a prediction from planet formation and evolution theory was later clearly confirmed by observations.

Besides photoevaporation, several alternative hypotheses have been proposed for the origin of the observed radius gap. First, in the core-powered mass-loss hypothesis, a planetary core’s internal luminosity drives the loss of its atmosphere ([Ginzburg et al. 2016](#)). Core-powered mass-loss is also able to reproduce the observed radius valley as a function of orbital period ([Ginzburg et al. 2018; Gupta & Schlichting 2019](#)). Similar to photoevaporation studies, [Gupta & Schlichting \(2019\)](#) furthermore also find that the observations are consistent with predominantly rocky cores. Second, the valley could also be caused by two distinct formation pathways for planets above and below the gap, with those above the valley being water-worlds ([Zeng et al. 2019](#)). Third, planetesimal impacts can also create a similar radius gap depending on whether atmospheres grow or deplete in collisions ([Wyatt et al. 2019](#)).

[Fulton et al. \(2017\)](#) have not determined the slope of the gap as a function of orbital distance or irradiation flux yet. This is however of central importance, as a transition from solid to planets with primordial H/He because of evaporation leads to a negative slope, whereas a formation in a gas-free environment after disk dispersal should lead to a positive slope ([Lopez & Rice 2018](#)). The paper of [Van Eylen et al. \(2018\)](#) who used astroseismology to even better constrain the stellar radii finally showed that the valley has a negative slope, consistent with an evaporative transition from super-Earth to sub-Neptunes, where the lower boundary of the valley which corresponds to the largest bare core  $R_{\text{bare}}$  as function of semimajor axis  $a$  is approximately found at

$$R_{\text{bare}}(a) \approx 1.6 \pm 0.1 \times \left( \frac{a}{0.1 \text{ AU}} \right)^{-0.08^{+0.02}_{-0.05}} R_{\oplus}. \quad (1)$$

For the middle of the gap, [Van Eylen et al. \(2018\)](#) found a steeper slope, scaling approximately as  $a^{-0.15 \pm 0.05}$ . For comparison, [Lopez & Rice \(2018\)](#) find theoretically a dependency like  $a^{-0.22}$ .



**Fig. 1.** Distance-radius diagram from an early population synthesis calculation around  $1 M_{\odot}$  stars at 5 Gyr. Colors give the fraction of the initial H/He envelope mass that was lost. Gray dots are planets that have lost all H/He in the triangle of evaporation. The red line shows the upper limit at  $R_{\text{bare}}(a)$  given by Eq. (2). Just above the line, there is the valley of evaporation. Adapted from [Jin et al. \(2014\)](#).

[Fulton & Petigura \(2018\)](#) confirmed the existence of the gap and quantified how empty it is using even better constrained stellar radii, thanks to new *Gaia* parallaxes (see also [Berger et al. 2018](#)). They also found that the characteristic features (maxima of Super-Earths and sub-Neptunes, valley) shift to larger radii for more massive stars. This trend is also seen if one considers a wider stellar mass range ([Wu 2019](#)). Recently, [Martinez et al. \(2019\)](#) find that the valley has a slope scaling as  $a^{-0.17 \pm 0.03}$  based on a spectroscopic analysis of the California-*Kepler* Survey, in agreement with the findings of [Van Eylen et al. \(2018\)](#). On the other hand, based on a machine learning approach, [MacDonald \(2019\)](#) report a much steeper slope, going as  $a^{-0.48^{+0.13}_{-0.17}}$ . The theoretical results found in the present paper are compared to the observational results below in Sect. 4.

Coming back to the predictions of theoretical models, one can consider for example the results of [Jin et al. \(2014\)](#). In this work, the population-wide effect of evaporation was studied with a model that does not only simulate the long-term evolution under the effect of evaporation but self-consistently couples this to a planet formation model for the accretion of the solid core and H/He, orbital migration, and disk evolution ([Alibert et al. 2005; Mordasini et al. 2012b](#)). This yields the initial conditions for the evolution (for as recent study also combining formation and evolution, see [Carrera et al. 2018](#)). We note that in the present paper, “core” is always used in the astrophysical, and not geophysical sense. It thus denotes the entire solid part of the planet, including the metallic, silicate, and possibly ice parts, but not the H/He envelope.

Figure 1 shows an adapted version of a  $a$ - $R$  plot of [Jin et al. \(2014\)](#). Analogous results were found by [Lopez & Fortney \(2013\)](#) and [Owen & Wu \(2013\)](#). The main result visible in the figure is that close-in low-mass planets lose all H/He whereas more massive planets at larger distances keep it. This leads to a diagonal separation of solid planets from planets with H/He in the orbital distance – planetary radius (or mass) plane, or in other words, a triangle-like region of planets without H/He envelopes. This “triangle of evaporation” ([Jin & Mordasini 2018](#)) with solid planets is clearly separated from the planets with H/He. These two types of planets have very different mean densities as even a tiny amount of H/He strongly increases the radius (e.g., [Valencia et al. 2013](#)).

Jin et al. (2014) and Jin & Mordasini (2018) found numerically in the population syntheses that for solar-like stars the most massive planet as a function of distance that has lost all H/He and was stripped to a bare solid planet has a radius  $R_{\text{bare}}$  of about

$$R_{\text{bare}}(a) \approx 1.6 \times \left( \frac{a}{0.1 \text{ AU}} \right)^{-0.27} R_{\oplus} \quad (2)$$

and a corresponding mass  $M_{\text{bare}}$  of about

$$M_{\text{bare}}(a) \approx 6 \times \left( \frac{a}{0.1 \text{ AU}} \right)^{-1} M_{\oplus}. \quad (3)$$

This is shown by the red line in Fig. 1. The quoted values are for an Earth-like composition of the core. For a core composition with 0.75% ice in mass, the separating values at 0.1 AU are shifted from 1.6 to 2.3  $R_{\oplus}$  and from 6 to 8  $M_{\oplus}$ , with a very similar distance dependency (Jin & Mordasini 2018). This transition mass (radius) can be derived analytically as is demonstrated in the second part of the present paper (Sect. 3) from essential energy comparison arguments<sup>1</sup>.

The analytical (and numerical) results show how the transition from solid planets to planets with H/He and thus the evaporation valley depend on orbital distance, stellar XUV-luminosity, efficiency of evaporation, and postformation envelope mass. These dependencies (and others, like the stellar mass) are typical for evaporation as the mechanism responsible for envelope loss. In general, they should differ from the imprints and dependencies of other mechanisms that can also lead to the loss of the envelope, like impacts (Liu et al. 2015; Schlichting et al. 2015; Inamdar & Schlichting 2015; Biersteker & Schlichting 2019; Wyatt et al. 2019) or mass loss powered by the luminosity of a planet’s cooling core (Ginzburg et al. 2018; Gupta & Schlichting 2019, 2020).

This means that it might become possible to disentangle the contribution of these different processes with (future) surveys, such as TESS (Ricker et al. 2014), CHEOPS (Fortier et al. 2014), and PLATO (Rauer et al. 2014), which will potentially be combined with RV observations. They allow one to map the mean planetary density in the mass or radius-distance plane, which would be important to improve formation, evolution, and evaporation theory. In the ideal case, they allow one to even understand the dependency on additional relevant parameters like the stellar type and activity ( $L_{\text{XUV}}$ ) or the age of the planets. This would render the observations even more constraining.

When considering impacts as the reason of envelope loss, it appears likely that it will lead to a fuzzy transition due to the stochastic nature of impacts. Also evaporation will in reality not be as deterministic and of homogeneous efficiency for all planets as in the idealized simulations presented here. Even at fixed mass and distance, it will differ from planet to planet as it depends on the ability of the gas to cool and thus its atmospheric composition (e.g., Johnstone et al. 2018). It will also differ from planetary system to planetary system, as stars can have a wide variety in their luminosity in the XUV-wavelength domain at young ages (Tu et al. 2015). We do however find numerically, and show analytically, that the transition is rather weakly dependent

<sup>1</sup> Owen & Wu (2017) have independently derived a comprehensive analytical theory for the evaporative transition. Except for the comparison with the subsequently published observational studies (like Fulton et al. 2017; Fulton & Petigura 2018 and Van Eylen et al. 2018), the analytical derivation and the numerical calculations shown in the paper here were completed before we became aware of their work. Comparisons with the Owen & Wu (2017) study can be found in Sect. 3.4.

on these parameters. Nevertheless, systems of several planets are of special interest, as there all planets have experienced the same irradiation history from the host star (cf. Owen & Estrada 2019). This makes in particular Kepler-36 (Carter et al. 2012) with its two closely spaced planets of very different densities an important benchmark case. As will be discussed in Sect. 4.2, this system can be well explained by the calculations presented here without any special model tuning (as previously shown by Lopez & Fortney 2013; Owen & Morton 2016). Connecting compositional constraints to the ones from the dynamic system architecture (existence of mean motion resonances, circular versus eccentric orbits, etc.) could be an interesting pathway to allow further insights into the mechanism that have led to the formation and evolution of a specific system (Quillen et al. 2013; Dawson et al. 2016; Carrera et al. 2018).

In this paper, we want to quantify and understand the shape and locus of the valley of evaporation and its dependency on the initial (postformation) properties of the planets like the envelope-to-core mass ratio, or the strength of evaporation. We work in the limiting assumptions that all planets start with H/He envelope. This is not what is expected because of impacts or a formation after the protoplanetary gas disk is gone, and thus we do not think that the actual transition from solid to planets with H/He will be as clean as reported here. Rather, by comparing the transition due to evaporation found here with observations, should help to disentangle different mechanism.

The structure of the paper is as follow: In the first part of the paper starting with Sect. 2, we conduct numerical simulations of the thermodynamical evolution of low-mass planets undergoing evaporation. We present our model in Sect. 2.1 and the initial conditions in Sect. 2.1.3. The results from a parameter study covering different planet masses and distances are shown in Sect. 2.3. We identify the location of the transition and test its dependency on the initial envelope mass, the strength of evaporation, the composition of the core (rocky or icy), the envelope opacity and the envelope metallicity. In the second part, Sect. 3, we develop a simply analytical model based on energy comparison to explain the numerical results and the location of the observed valley of evaporation. We derive the governing equations in Sect. 3.2 and give the final result in Sect. 3.3. Section 4 compares the theoretical results with observations. The summary and conclusions are finally given in Sects. 5 and 6.

## 2. Numerical study

The basic idea of the numerical study is simple: we use our evolutionary model that includes the effect of evaporation to conduct a parameter study of the evolution of planets lying in the distance and mass range affected by evaporation visible in Fig. 1. For this, we follow the evolution of a high number of simulated planets on grids of core masses and orbital distances, studying the impact of several important parameters on the transition from solid to planets with H/He. For the initial conditions, namely the postformation envelope-to-core mass ratio and luminosity, we use as a starting point the results from population synthesis calculations of the formation of planets via core accretion (Mordasini et al. 2012a), and then vary these initial conditions. Regarding these planet formation simulations, the goal is to see if by comparing the observed and simulated locus of the valley, we can put constraints on them. One could for example a priori think that the locus and slope of the valley depend on the postformation envelope mass as a function of a planet’s core mass and orbital distance (which is not the case, as we find below).



## 2.1. Simulation setup

We first summarize the numerical model, then we describe its initial conditions.

### 2.1.1. Planet evolution model with atmospheric escape

Our planet evolution model `complete21` was already described in several publications (Mordasini et al. 2012b; Jin et al. 2014; Linder et al. 2019); therefore, here, we only give a short overview. It models the temporal evolution of planets (their cooling and contraction) by integrating numerically their 1D interior structure, taking into account the loss of H/He because of atmospheric escape.

Our model assumes that planets consist of a solid differentiated part (the core) consisting itself of iron, silicates, and, if the planet accreted outside of the iceline, water ice. These substances are described by temperature-independent modified polytropic equations of state EOS giving the density  $\rho$  as a function of pressure  $P$  with the expression (Seager et al. 2007)

$$\rho(P) = \rho_0 + cP^n. \quad (4)$$

The parameters  $\rho_0$ ,  $c$ , and  $n$  for iron, perovskite, and water ice are taken from Seager et al. (2007). The radius of the core is then found by solving the 1D spherically symmetric equations of mass conservation and hydrostatic equilibrium

$$\frac{dm}{dr} = 4\pi r^2 \rho \quad \frac{dP}{dr} = -\frac{Gm}{r^2} \rho \quad (5)$$

where  $r$  is the distance from the planet's center,  $m$  the enclosed mass, and  $G$  the gravitational constant. When solving the equations, a possible pressure exerted by the envelope on the outer boundary (surface of the core) is taken into account which is however only important for massive envelopes (see Mordasini et al. 2012a). The dependency of the core's radius on the temperature is neglected because the influence is only small (Grasset et al. 2009).

The gaseous envelope consisting of H/He is described by the EOS of Saumon et al. (1995). The interior structure of the envelope is found in the 1D spherically symmetric approximation by solving the equations of mass conservation, hydrostatic equilibrium, energy generation, and energy transport (with  $T$  the temperature)

$$\frac{dm}{dr} = 4\pi r^2 \rho \quad \frac{\partial P}{\partial r} = -\frac{Gm}{r^2} \rho \quad (6)$$

$$\frac{\partial l}{\partial r} = 0 \quad \frac{\partial T}{\partial r} = \frac{T}{P} \frac{\partial P}{\partial r} \nabla(T, P) \quad (7)$$

where we make in the energy equation the approximation that the intrinsic luminosity  $l$  is radially constant at a given time, and use total energy conservation to follow the temporal evolution, following the method introduced in Mordasini et al. (2012b).

We use the Schwarzschild criterion to decide whether the energy transport occurs via radiative diffusion or convection, such that  $\nabla$  is always the smaller of the radiative and the adiabatic gradient. For the intrinsic luminosity of the planets, the cooling and contraction of the H/He envelope and of the core are included. For the latter we assume a constant heat capacity of  $10^7$  and  $6 \times 10^7$  erg g<sup>-1</sup> K<sup>-1</sup> for rocky material and ices, respectively. Our evolution model reproduces the usual results (Fortney et al. 2011) for the evolution of the gas and ice giant planets in the solar system (Linder et al. 2019): Jupiter's and Uranus' luminosity at

present day is recovered, but Saturn is less bright in the model than observed (potentially because of a helium rain which we do not include), whereas Neptune is predicted to be much more luminous than observed (which could be to a less efficient energy transport because of compositional gradients, which is also not considered in the model). The model also reproduces very well the evolutionary simulations of Lopez & Fortney (2014) for a  $5 M_{\oplus}$  planets with a 1% H/He envelope (Linder et al. 2019).

The outer boundary condition of the H/He envelope (the atmosphere) is described by a two stream double gray irradiated atmosphere in the form of an improved version of the Guillot (2010) analytical solution, as described in Jin et al. (2014). For a planet that has an intrinsic temperature  $T_{\text{int}}$  (resulting from the planet's intrinsic luminosity) and an equilibrium temperature  $T_{\text{equi}}$  (resulting from stellar irradiation), the resulting temperature as a function of IR optical depth  $\tau$  is

$$T^4 = \frac{3T_{\text{int}}^4}{4} \left( \frac{2}{3} + \tau \right) + \frac{3T_{\text{equi}}^4}{4} \left( \frac{2}{3} + \frac{2}{3\gamma} \right) \times \left[ 1 + \left( \frac{\gamma\tau}{2} - 1 \right) e^{-\gamma\tau} \right] + \frac{2\gamma}{3} \left( 1 - \frac{\tau^2}{2} \right) E_2(\gamma\tau) \quad (8)$$

where  $\gamma$  denotes the ratio of the mean opacity in the visual  $\kappa_v$  to the mean opacity in the thermal infrared  $\kappa_{\text{th}}$ , while  $E_2$  is an exponential integral. The parameter  $\gamma$  is tabulated in Jin et al. (2014). The opacities correspond to a condensate-free gas of solar composition (Freedman et al. 2014).

In our evaporation model (described in Jin et al. 2014), atmospheric escape is assumed to be hydrodynamic, and to occur at lower incident EUV-fluxes in the energy-limited regime (Watson et al. 1981) where

$$\frac{dM_{\text{UV,e-lim}}}{dt} = -\frac{\varepsilon_{\text{UV}} \pi F_{\text{UV}} R_{\text{UV}}^3}{GM}. \quad (9)$$

In this equation,  $M$  is the planet's mass,  $\varepsilon_{\text{UV}}$  is an efficiency factor, and  $R_{\text{UV}}$  the planetary radius where EUV radiation is absorbed which is estimated as in Murray-Clay et al. (2009). Following this work, the efficiency factor is set to 0.32, and assumed to be constant. At high EUV-fluxes, the radiation-recombination limited regime is considered, where the mass loss rate is (Murray-Clay et al. 2009)

$$\frac{dM_{\text{UV,r-lim}}}{dt} = -4\pi \rho_s c_s r_s^2 \quad (10)$$

where  $\rho_s$  and  $c_s$  are the density and speed of sound at the sonic point at a radius  $r_s$ . These quantities are also estimated as described in Murray-Clay et al. (2009).

Heating both by X-ray and EUV radiation is included, using the criterion of Owen & Jackson (2012) to decide which regime is dominant. In the X-ray driven regime, the escape rate is also modeled with an energy-limited formula. The stellar  $L_{\text{XUV}}$  as a function of time is taken from Ribas et al. (2005), but we explore in the simulations below scenarios with a 10 times higher and lower evaporation rate, which could be caused by correspondingly higher and lower stellar XUV luminosities, to account for the observed spread (Tu et al. 2015).

For some initial conditions of low-mass cores with massive envelopes, we (numerically) find at the beginning of the simulation an outer radius that exceeds the planet's Hill sphere radius, indicating that overflow occurs. In this case, we remove at each timestep gas outside of  $R_{\text{Hills}}$ , which leads to a very fast ( $\sim 10^4$  yr) reduction of the envelope mass and outer radius, until it falls



below  $R_{\text{Hills}}$ . This might be manifestation of the boil-off regime described by [Owen & Wu \(2016\)](#) (see also [Ginzburg et al. 2016](#) and [Fossati et al. 2017](#)).

The exact temporal behavior during this overflow might not be well described in our model because of (a) its underlying hydrostatic approximation, and (b) the radially constant luminosity. However, numerical experiments varying the treatment of this phase (for example the fraction of the envelope mass outside of  $R_{\text{Hills}}$  removed at each timestep) show that in any case, such planets suffer eventually a complete loss of the envelope, typically on short timescales of less than about 1 Myr, meaning that the final result at, for example, 5 Gyr, is not affected.

When comparing to observations, it should be kept in mind that despite the inclusion of several sub-regimes, our evaporation model still a simplified, parameterized (e.g., via the constant  $\varepsilon_{\text{UV}}$ ) description of the actual loss process. The evaporation model will be improved in future work to include the results of [Owen & Alvarez \(2016\)](#) on the occurrence of different evaporation regimes. Given the rather weak dependency on the details of the evaporation model found analytically and numerically, and the general agreement of our results with other more sophisticated evaporation models (see comparison in [Jin et al. 2014](#), and Sect. 3.4), we however do not expect these factors to fundamentally change the results found here.

### 2.1.2. Simulations with enriched atmospheres

For a subset of simulations, we take into account that the gaseous envelopes of the planets could be significantly enriched in heavy elements instead of being dominated by H/He. In these simulations, the model is modified relative to the description in the previous section with respect to three points:

First, the envelope is assumed to consist of a mixture of H/He described by the EOS of [Saumon et al. \(1995\)](#) and of  $\text{H}_2\text{O}$  with a mass fraction  $Z_{\text{enve}}$  described by the ANEOS equation of state ([Thompson 1990](#)). The results of ANEOS for  $\text{H}_2\text{O}$  are briefly described in Appendix B. The heavy element mass fraction  $Z_{\text{enve}}$  is assumed to be radially and temporally constant in the envelope. While its specific thermodynamic properties will obviously affect the results, the  $\text{H}_2\text{O}$  should in the current context be seen as representing heavy elements (metals) in general. The H/He and  $\text{H}_2\text{O}$  is mixed using the additive volume law (e.g., [Baraffe et al. 2008](#)).

Second, as in [Lopez \(2017\)](#), in the evaporation model the efficiency factor in the energy-limited regime is varied with  $Z_{\text{enve}}$  as  $Z_{\text{enve}}^{-0.77}$  ([Owen & Jackson 2012](#)). One should note that this scaling is taken from studies of the X-ray heating of protoplanetary disk atmospheres ([Ercolano & Clarke 2010](#)) rather than planetary atmospheres. While it seems likely that evaporation rates indeed decrease with increasing metal content because metal atomic lines are important coolants ([Salz et al. 2016](#); [Owen & Murray-Clay 2018](#)), the exact dependencies have not yet been explored. Our results based on this simple scaling are thus to be taken with caution. In the radiation-recombination limited domain, the mean molecular weights are also adjusted depending on  $Z_{\text{enve}}$  (see [Lopez 2017](#)), with the mean molecular weights taken directly from the mixed EOS.

Third, the atmospheric opacities ([Freedman et al. 2014](#)) are calculated for the  $[\text{M}/\text{H}]$  that corresponds to  $Z_{\text{enve}}$ . The conversion from metal mass fraction into  $[\text{M}/\text{H}]$  is made in an analogous way as described in [Valencia et al. \(2013\)](#). For the  $Z_{\text{enve}} = 0.1, 0.3, \text{ and } 0.5$  that we consider in the simulations below, this leads to  $[\text{M}/\text{H}] = 0.86, 1.43, \text{ and } 1.78$ .

### 2.1.3. Postformation H/He envelope mass

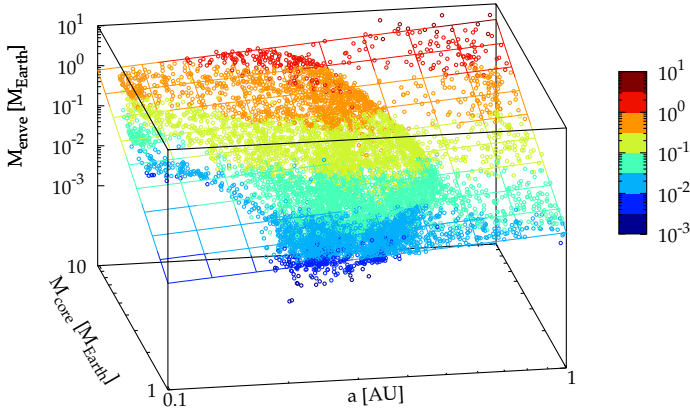
A central initial condition for the evolutionary calculations is the envelope mass of H/He  $M_{\text{e},0}$  as a function of planetary properties at the end of the formation epoch when the gas disk disperses (even if we see later that  $R_{\text{bare}}$  only depends very weakly on it). Here we start from the results of population synthesis calculations based on the core accretion paradigm that were presented in [Mordasini et al. \(2014\)](#).

In this paper, the effect of the atmospheric opacity during formation on the accreted H/He mass and associated mass-radius relation was studied. The envelope masses were found by explicitly solving numerically the standard 1D internal structure equations (e.g., [Mordasini et al. 2012b](#)). The accretion of planetesimals, orbital migration, and the evolution of the protoplanetary disk are also included in our global model ([Alibert et al. 2005](#)). Here we use the nominal population of [Mordasini et al. \(2012a\)](#). It is characterized by a fixed stellar mass of  $1 M_{\odot}$ , solar-composition H/He envelopes, and an opacity in the protoplanetary envelope that is given by [Freedman et al. \(2014\)](#) for the grain-free molecular opacities, and the [Bell & Lin \(1994\)](#) grain opacities reduced by a factor 0.003. This nominal reduction factor was determined in [Mordasini et al. \(2014\)](#) by calibrating the gas accretion timescales with the ones found with the detailed model of [Movshovitz et al. \(2010\)](#) for the grain dynamics. Such microphysical models for the dynamics of the grains ([Podolak 2003](#); [Movshovitz et al. 2010](#); [Ormel 2014](#); [Mordasini 2014](#)) consider the settling, coagulation, and evaporation of the grains in the outer radiative zone (atmosphere) of the protoplanets. They predict opacities that are much smaller than in the ISM, and on the order of a few  $10^{-3} \text{ cm}^2 \text{ g}^{-1}$  at the radiative-convective boundary rcb. This is comparable to, but still a bit higher than expected for a completely grain-free atmosphere.

For this work, we are interested in a mean analytical relation. We have therefore determined the mean postformation envelope mass  $M_{\text{e},0}$  as a function of the planet's core mass  $M_{\text{c}}$  and final semimajor axis  $a$  by fitting the numerical results with a least-square method. The core mass and orbital distance are the quantities that most clearly and systematically influence  $M_{\text{e},0}$ . There is a significant spread around the mean relation, as also other factors like the disk lifetime or the accretion rate of planetesimals prior to disk dispersal influence the final  $M_{\text{e},0}$ . For the fit, we have included planets with  $0.1 < a < 1 \text{ AU}$  and  $1 < M_{\text{c}} < 10 M_{\oplus}$ . They have envelopes with typical masses of  $\sim 1\text{--}10\%$  of their total mass. Such planets are a very frequent outcome of the formation simulations. This shows that the formation of close-in cores of  $5\text{--}10 M_{\oplus}$  with only relatively low-mass envelopes is a natural outcome when the structure equations are directly solved - not all these cores become giant planets. By solving the internal structure equations in the formation model, the decrease and/or limitation of the envelope mass due to the luminosity caused by solid accretion as well as because of the decreasing nebular pressure are automatically included. As impact stripping is not included, all planets are assumed to start with primordial H/He.

Figure 2 shows the envelope mass as function of core mass and semimajor axis together with the fit. One sees that the envelope mass increases with core mass and orbital distance, which is expected ([Ikoma & Hori 2012](#)). The imprint of the inner convergence zone of type I migration ([Dittkrist et al. 2014](#)) is visible in an arc like structure. Quantitatively, the following power law dependency is found for  $M_{\text{e},0}$ :

$$\frac{M_{\text{e},0}}{M_{\oplus}} = 0.024 \left( \frac{M_{\text{c}}}{M_{\oplus}} \right)^{2.23} \left( \frac{a}{1 \text{ AU}} \right)^{0.72}, \quad (11)$$



**Fig. 2.** Postformation envelope mass as a function of core mass and semimajor axis for synthetic close-in low-mass planets (colored points). The colored mesh is the least-square fit (Eq. (11)).

which corresponds to a roughly speaking quadratic increase with the core mass, and a weaker than linear increase with distance. In terms of gas-to-core mass ratio  $M_{e,0}/M_c$  and normalized at a more relevant semimajor axis of 0.1 AU, this corresponds to

$$\frac{M_{e,0}}{M_c} = 0.005 \left( \frac{M_c}{M_\oplus} \right)^{1.23} \left( \frac{a}{0.1 \text{ AU}} \right)^{0.72}. \quad (12)$$

It is interesting to compare this scaling relation with the analytical result of Lee & Chiang (2015). The analytical result is obtained by considering that the planets accrete gas on a timescale given by the envelope’s Kelvin-Helmholtz cooling timescale (Ikoma et al. 2000), and that the magnitude of cooling is set at the radiative-convective boundary (Lee & Chiang 2015). For the planets here, the regime of completely dust-free atmospheres in a gas poor nebula at 0.1 AU is most likely to apply. If we assume that the temperature at the rcb is approximately equal to the nebular temperature as suggested by Lee & Chiang (2015), that the later scales as  $a^{-0.5}$  (Ida & Lin 2004), and finally that the disk lifetime is 2 Myr (the mean lifetime of our synthetic disks, see also Haisch et al. 2001), Eq. (22) of Lee & Chiang (2015) predicts

$$\frac{M_{e,0}}{M_c} = 0.016 \left( \frac{M_c}{M_\oplus} \right)^{1.6} \left( \frac{a}{0.1 \text{ AU}} \right)^{0.95}. \quad (13)$$

We thus see that the power law exponents are not too different, but that the absolute mass found in the numerical calculations is about a factor of 3.5 lower. These lower envelope masses could be due to (a) a higher opacity (residual grain opacity, neglected in the analytical model), (b) some accretional heating by planetesimals (also neglected in the analytical model), especially as orbital migration brings the protoplanets into regions of the disk still containing planetesimals or (c) the decline of the ambient nebula over time. A detailed comparison of the analytical result and the direct solution of the structure equations is, however, beyond of this paper, especially in view of the weak dependency of  $R_{\text{bare}}$  on it that we find further down.

#### 2.1.4. Postformation luminosity

Besides the envelope mass fraction, one also needs to specify the luminosity at the end of the formation phase as an initial condition for the evolutionary simulations. This postformation

luminosity  $L_0$  was also taken from the aforementioned population synthesis calculations of planetary formation, considering the luminosity as a function of core mass for low-mass planets with masses between 1 to 10  $M_\oplus$  inside of 1 AU at an age of 10 Myr. One finds a fitting relation of

$$L_0/L_\oplus \approx 0.008 \times \left( \frac{M_c}{M_\oplus} \right)^{2.5}, \quad (14)$$

where  $L_\oplus$  is the present day intrinsic luminosity of Jupiter (about  $8.7 \times 10^{-10} L_\odot$ , Guillot & Gautier 2014). Most synthetic planets are within a factor two higher or lower than this mean relation. Given the rather weak dependency of the thickness of the convective zone of the H/He envelope on the age (i.e., luminosity) found by Lopez & Fortney (2014), especially when compared to the impact of the envelope mass, we did not investigate the consequences of varying  $L_0$ . The role of the luminosity for the evaporation valley, in particular when also considering a possible additional luminosity source like ohmic dissipation (that could be strong in low-mass planets, Pu & Valencia 2017) should however be addressed in future work.

In Appendix A we give as a side result for higher ages of the planets a fit for the luminosity as a function of time, core mass, and envelope mass. This fit may be used in time-independent internal structure calculations like Rogers et al. (2011); Dorn et al. (2017); Lozovsky et al. (2018) which need the luminosity as an input quantity.

#### 2.2. Simulations

We have calculated 14 grids of planetary evolution simulations in the  $M_c - a$  plane, varying the (1) the postformation envelope mass in several ways because of the motivation to understand whether gas accretion during formation can be constrained by the locus of the valley, (2) the strength of evaporation which could represent different efficiency factors and different stellar XUV-luminosities, (3) the Rosseland mean opacity in the atmosphere, (4) the metallicity (heavy element enrichment) of the gaseous envelope and (5) the ice mass fraction in the core.

Table 1 gives an overview of the simulations. In the second column, the postformation H/He envelope mass  $M_{e,0}$  (in  $M_\oplus$ ) as a function of core mass  $M_c$  (also in  $M_\oplus$ ) and potentially the semi-major axis  $a$  (in AU) is given. The last seven simulations use the same  $M_{e,0}$  as M3 (E1, E2, O1, I1) or M0 (Z1, Z2, Z3), but other parameters are varied, as indicated in the table. The 14 simulations are described as follows:

- Simulation M0 is the nominal simulation, where the envelope mass varies as described by Eq. (11). It thus increases roughly quadratically with the core mass and linearly with orbital distance.

- The next four simulations (M1-M4) vary the exponent  $p_e$  in the power law dependency of the envelope mass on core mass,  $M_e \propto M_c^{p_e}$  from 0 to 3 (Eq. (22)).

- The simulations N1 and N2 also investigate the impact of the primordial envelope mass. In M1-M4, the normalization constant  $M_{e,1}$  (Eq. (22)) which is the envelope mass of a 1  $M_\oplus$  core is 0.03  $M_\oplus$ . In N1 and N2 it is instead 0.06 and 0.3  $M_\oplus$  respectively, that is, we are studying the effects of envelope masses with are two and ten times as high than in M3.

- In simulations E1 and E2 the evaporation rate in all regimes is uniformly multiplied by a factors 0.1 and 10 relative to the evaporation rate normally given by the model, respectively.

- The simulation O1 quantifies the impact of a Rossland opacity in the H/He envelope that is increased uniformly by a

**Table 1.** Description and outcome of the 14 simulations.

Name	Characteristics	$M_{b,0p1}$ (num.)	$M_{b,0p1}$ (an.)	$e_m$	$R_{b,0p1}$ (num.)	$R_{b,0p1}$ (an.)	$e_r$
M0 (nominal)	$0.024 M_c^{2.23} a^{0.72}$	5.82	6.5	1.05	1.67	1.6	0.28
M1	0.03	8.30	6.5	1.49	1.82	1.6	0.36
M2	$0.03 M_c$	5.98	6.5	1.07	1.68	1.6	0.30
M3 (reference)	$0.03 M_c^2$	5.52	6.5	0.99	1.65	1.6	0.27
M4	$0.03 M_c^3$	5.21	6.5	0.87	1.60	1.6	0.25
N1	$0.06 M_c^2$	5.39	6.5	0.93	1.65	1.6	0.28
N2	$0.3 M_c^2$	5.38	6.5	0.93	1.66	1.6	0.28
E1	$\dot{M}_{\text{evap}} \times 0.1$	2.88	2.0	0.91	1.39	1.2	0.28
E2	$\dot{M}_{\text{evap}} \times 10$	10.6	20.5	0.99	1.97	2.2	0.27
O1	$\kappa_R \times 10$	7.09	–	0.97	1.78	–	0.28
Z1	$Z_{\text{enve}} = 0.1$	5.68	–	1.01	1.58	–	0.28
Z2	$Z_{\text{enve}} = 0.3$	3.97	–	1.10	1.43	–	0.30
Z3	$Z_{\text{enve}} = 0.5$	3.01	–	1.14	1.32	–	0.31
I1	Icy cores ( $f_{\text{ice}} = 1$ )	8.57	11.9	0.89	2.68	2.4	0.26

**Notes.** The second column gives for the first 7 simulations the initial envelope mass (in units of  $M_\oplus$ ) as a function of core mass  $M_c$  (in  $M_\oplus$ ) and orbital distance  $a$  (in AU). The last 7 simulations use the same initial envelope masses as M3 or M0 (for Z1, Z2, Z3), but differ from it in other characteristics shown in the table.  $M_{b,0p1}$  is the maximum bare core mass at 0.1 AU, found numerically (“num.”, Eq. (15)) and analytically (“an.”, Eq. (30)), while  $e_m$  is the slope of the valley in the mass-distance plane found numerically. The analytical model predicts  $e_m = 1$ .  $R_{b,0p1}$  and  $e_r$  are the equivalent quantities in the radius-distance plane (Eqs. (16) and (32)), meaning that  $R_{b,0p1}$  is the radius of the most massive planet that can be completely stripped. It corresponds to the bottom of the evaporation valley at 0.1 AU. The analytical model predicts  $e_r = p_c = 0.27$ .

factor of 10. Other quantities that likely also depend on the gas composition (EOS, evaporation rate) are unchanged.

– In the simulation Z1, Z2, and Z3 the composition of the gaseous envelope is changed from solar composition as in all other simulations to mixtures of H/He with  $\text{H}_2\text{O}$  with a mass fraction of  $Z_{\text{enve}} = 10, 30,$  and  $50\%$ . The EOS, opacity, and evaporation rate are all modified self-consistently, as described in Sect. 2.1.2.

– The last simulation I1 shows the impact of increasing the ice mass fraction in the core to 1. Such completely icy cores (without any silicates and iron) are certainly not expected in reality, but it is instructive for comparison with the analytical model and with Owen & Wu (2017); Jin & Mordasini (2018) who both investigated the valley’s position as function of ice mass fraction.

Regarding the grid, a range in core masses between 1 and  $20 M_\oplus$  and orbital distances between 0.03 and 0.6 AU was covered in most simulations. Because of computational time reasons, some simulations were conducted on a grid of reduced size.

### 2.3. Results

The main result of the numerical study is the location of the valley (transition from super-Earth to sub-Neptunes) as a function of orbital distance found in the 14 simulations. Given the simulation results, we quantify its location by numerically deriving a least-square power law fit to the highest mass  $M_{\text{bare}}$  and largest radius  $R_{\text{bare}}$  as a function of orbital distance  $a$  that has lost the entire H/He envelope at an age of 5 Gyr, normalized by the value at 0.1 AU. This means that the location of the valley is quantified with  $M_{b,0p1}$  and  $e_m$  in

$$M_{\text{bare}}(a) = M_{b,0p1} \left( \frac{a}{0.1 \text{ AU}} \right)^{-e_m} \quad (15)$$

in the mass-distance plane, and equivalently with  $R_{b,0p1}$  and  $e_r$  in

$$R_{\text{bare}}(a) = R_{b,0p1} \left( \frac{a}{0.1 \text{ AU}} \right)^{-e_r} \quad (16)$$

in the radius-distance plane. We note that the two fits to obtain the four quantities were made independently. These four quantities can be directly compared with the analytical predictions in the second part of the paper.

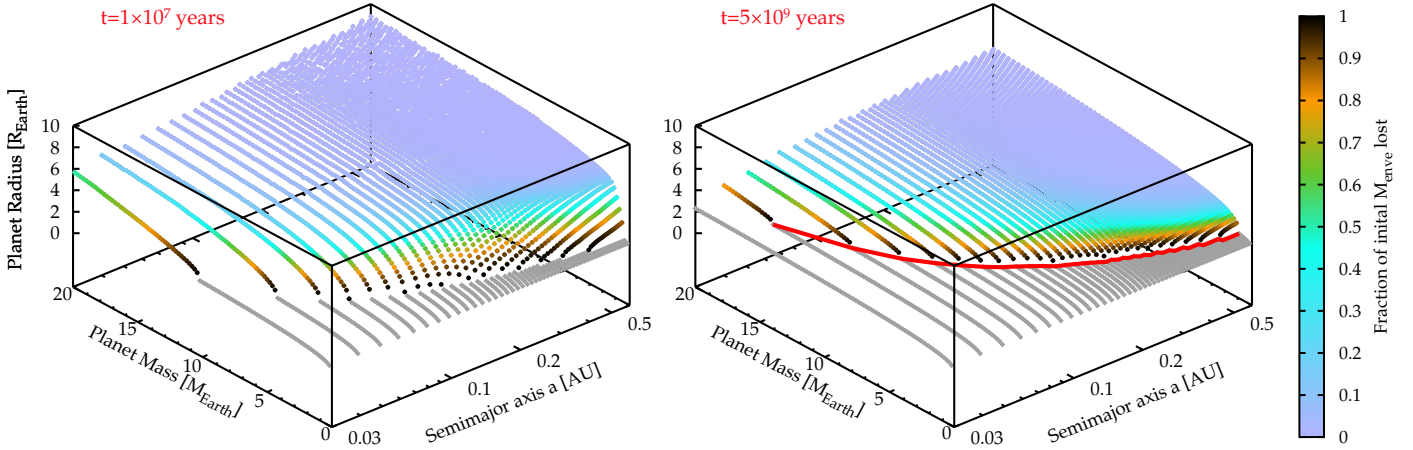
The choice of the specific age of 5 Gyr does not affect the results as long as we are considering Gyr-old planets, as most of the atmospheric loss occurs during the first  $\sim 100$  Myr anyway, after which the triangle of evaporation has already attained almost its final size (Jin et al. 2014). An observational determination of the temporal growth of the triangle at early times – if possible – would however represent an extremely interesting constraint for the various proposed processes of envelope loss. In this context it is interesting to note that the PLATO mission should be able to determine accurate stellar ages thanks to astro-seismology. We now discuss the outcomes of the 14 simulations.

#### 2.3.1. Nominal simulation: M0

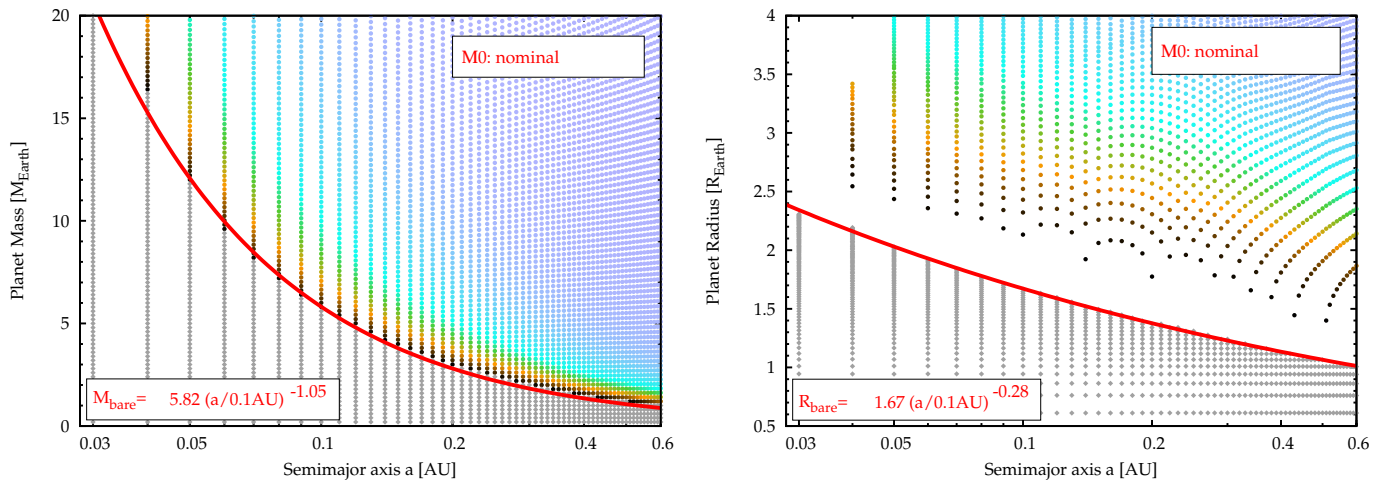
To illustrate the general outcome of the simulations, we show in Fig. 3 the radius of the planets in the grid as a function of their orbital distance and (total) mass at 10 Myr and at 5 Gyr. The overall contraction of the radii, the decrease of the H/He mass, and the growing bare core triangle is visible, extending to larger masses, radii, and distances as time goes on. At 5 Gyr, the red curve indicates  $R_{\text{bare}}$  as found from the least square fit (same as in Fig. 4).

Figure 4 shows the location of the valley in the mass-distance and radius-distance plane at 5 Gyr. As expected, there is no gap or valley in the mass-distance distribution, because first, the mass fraction of the H/He envelope is small compared to the core mass in any case (at least for the lower mass cores), and second, the envelope mass is reduced to zero in a continuous fashion, as indicated by the colors. In the radius, there is in contrast a gap or valley, as expected. As explained for example in Jin & Mordasini (2018), the valley comes into existence because first, even the addition of a very low-mass H/He envelope significantly increases the radius, and second, the loss of this last remaining





**Fig. 3.** Evolution of the planets of the nominal model M0 in the distance-mass-radius-time space. The radius ( $z$ -axis) at 10 Myr and 5 Gyr is shown as a function of semimajor axis and mass. The color code gives the fraction of the initial H/He envelope that was evaporated. Planets that have lost all H/He are plotted in gray. The overall contraction and the growing bare core triangle is visible, extending to larger masses and distances as time goes on. The red curve indicates  $M_{\text{bare}}$  and  $R_{\text{bare}}$  at 5 Gyr as in Fig. 4.



**Fig. 4.** Transition from solid planets to planets with H/He in the plane of mass (*left panel*) and corresponding radius (*right panel*) versus semimajor axis for the nominal simulation M0 at 5 Gyr. Colored points are planets that still have primordial H/He whereas gray diamonds are bare rocky planets. The color code shows the fraction of initial H/He that was evaporated (same color scale as in Fig. 3). The red line is a power law fit to the most massive respectively largest planet that has lost the envelope, representing the transition at  $M_{\text{bare}}$  respectively  $R_{\text{bare}}$ , with the fit parameters indicated in the inset ( $M_{\text{bare}}$  in  $M_{\oplus}$ ,  $R_{\text{bare}}$  in  $R_{\oplus}$ ). Planets below the red line are solid planets in the bare core triangle. In the mass plane the transition is continuous, whereas in the radius plane, there is a gap separating solid planets from planets with gas (the evaporation valley).

envelope occurs on a timescale of only  $\sim 10^5$  yr (Jin et al. 2014). This means that if we take a snapshot of the population at 5 Gyr, it is unlikely (but not impossible) to observe a planet just in this final phase, leading to the depleted region.

The location of the valley, quantified by  $M_{\text{b,0p1}}$ ,  $e_{\text{m}}$ ,  $R_{\text{b,0p1}}$ ,  $e_{\text{r}}$  (Eqs. (15) and (16)), is similar to the ones found in Jin et al. (2014) and Jin & Mordasini (2018). The four parameters are given in the figure, and in Table 1. This similarity is not surprising, because the new simulations shown here use the same evaporation model and similar initial conditions.

In Table 1 we also compare the numerical results for the locus of the transition with the analytical model of Sect. 3. For a constant efficiency factor  $\varepsilon$  in the energy-limited escape rate – as assumed in the numerical model –, the analytical model predicts for all numerical simulations a power law exponent for the transition mass as a function of orbital distance of  $e_{\text{m}} = 1$  (Eq. (30)) and for the radius  $e_{\text{r}} = 0.27$  (Eq. (32)). Numerically, for simulation M0,  $e_{\text{m}} = 1.05$ , and  $e_{\text{r}} = 0.28$  is found, corresponding to a very similar result.

### 2.3.2. Sub-Neptune desert versus (evaporation) valley

In Fig. 4 (and also other simulations shown below), besides the valley, we also note a complete absence of very strongly irradiated planets with H/He inside of 0.04 AU. Inside of this distance, even the most massive core considered in the grid ( $20 M_{\oplus}$ ) loses the entire envelope. Only much more massive giant planets could keep their H/He at these very small distances. This “photoevaporation or sub-Neptunian desert” which must be distinguished from the radius valley was explored observationally for example in Lundkvist et al. (2016); Mazeh et al. (2016); Bourrier et al. (2018). It is another characteristic consequence of atmospheric escape (e.g., Lecavelier des Etangs et al. 2004; Kurokawa & Nakamoto 2014; McDonald et al. 2019). In this desert, more massive planets are affected for which the H/He initially represents a significant part of the total mass, in contrast to the evaporation valley. Hence, the loss of the envelope leads for these more massive planets to a substantial reduction of the total mass (and not only of the radius as for the valley). This is why the desert shows

up also in the mass-distance plot, whereas the valley is only visible in the radii. While due to the same physical effect, this imprint is thus different in nature from the evaporation valley.

### 2.3.3. Scaling of the envelope mass with core mass: M1–M4

To understand whether the valley’s position can constrain the postformation core-envelope mass relation, it is interesting to vary the initial H/He mass. Figure 5 shows the location of the valley for four power law scalings of the initial envelope mass  $M_{e,0}$  with core mass,  $M_{e,0}/M_{\oplus} = 0.03(M_c/M_{\oplus})^{p_e}$  with  $p_e = 0, 1, 2$  (reference simulation), and 3.

The plot shows that while there is a correlation of a decreasing  $M_{\text{bare}}$  and  $R_{\text{bare}}$  for an increasing  $p_e$  (more massive envelopes for planets with  $M_c > 1 M_{\oplus}$ ), we always have  $R_{\text{bare}} \approx 1.6 \dots 1.8 \times (a/0.1 \text{ AU})^{0.25 \dots 0.35} R_{\oplus}$  despite the very large differences in initial (i.e., postformation) H/He envelope masses for the more massive cores (e.g., a factor 125 for the  $5 M_{\oplus}$  core between M1 and M4). This shows that the postformation envelope mass has no significant influence on the final locus of the valley, at least if the initial envelope masses are not very different from what we nominally assume based on formation models (Sect. 2.1.3). This has the important implication that unfortunately (from a formation point of view), the valley location does not constrain strongly envelope accretion models. We also see that the valley is at a very similar location as in the nominal simulation M0 (shown by gray lines in the plots) where the initial envelope mass depends on the semimajor axis also, in contrast to the four simulation shown here. Clearly, this distance dependency has no significant effect, neither.

The plot also shows that the relative difference is larger for  $M_{\text{bare}}$  than for  $R_{\text{bare}}$ . This is expected, as the core mass in the form of the local  $M_{\text{bare}}$  is the more fundamental quantity controlling whether a planet can hold on to its H/He envelope than the core radius, as will become clear in the analytical model (Sect. 3.3). Once  $M_{\text{bare}}$  is given, the  $R_{\text{bare}}$  then follows simply from the weak dependency inherent to the relation of a solid planet’s (core) radius and its (core) mass, approximately as  $R \propto M^{0.27}$  for an Earth-like composition (see Eq. (23)).

Compared to the weak dependency found here, the analytical model even predicts that there is no dependency of the valley’s location on  $p_e$  at all (Sect. 3.3). As will become obvious there, the physical reason is that at a given core mass, a higher (lower) postformation H/He mass on one hand means that there is more (less) material to evaporate, but on the other hand also that the planet has a larger (smaller) radius. Since the (total) mass is essentially given by the core mass, this means also that the planet has a smaller (larger) mean density  $\bar{\rho}$ . But as shown by Eq. (9),  $\dot{M}_{\text{evap}} \propto \bar{\rho}^{-1}$ , meaning that the planet with more (less) H/He also loses more (less). As we see here numerically, and show analytically in Sect. 3.2, the mass-radius relation  $R(M_c, M_e)$  of low-mass planets with H/He is such that the two effects nearly cancel.

For the distance dependency, the analytical model also predicts that  $M_{\text{bare}} \propto a^{-1}$  ( $e_m = 1$ ), and  $R_{\text{bare}} \propto a^{-0.27}$  ( $e_r = 0.27$ ), independently of the scaling of the envelope mass with core mass  $p_e$ , which is to good approximation also the case in most of the numerical simulations shown here. For M2 to M4,  $e_m = 0.87$ – $1.07$  and  $e_r = 0.25$ – $0.3$ . The largest discrepancy between analytical prediction and simulation occurs for M1 where  $e_m = 1.49$  and  $e_r = 0.36$ . In M1, the envelope mass is independent of core mass, meaning that also massive cores  $\geq 10 M_{\oplus}$  only have a very low-mass envelopes, in contrast to the predictions from formation theory. This case is addressed in Sect. 3.3.4.

### 2.3.4. Constraints on postformation envelope masses

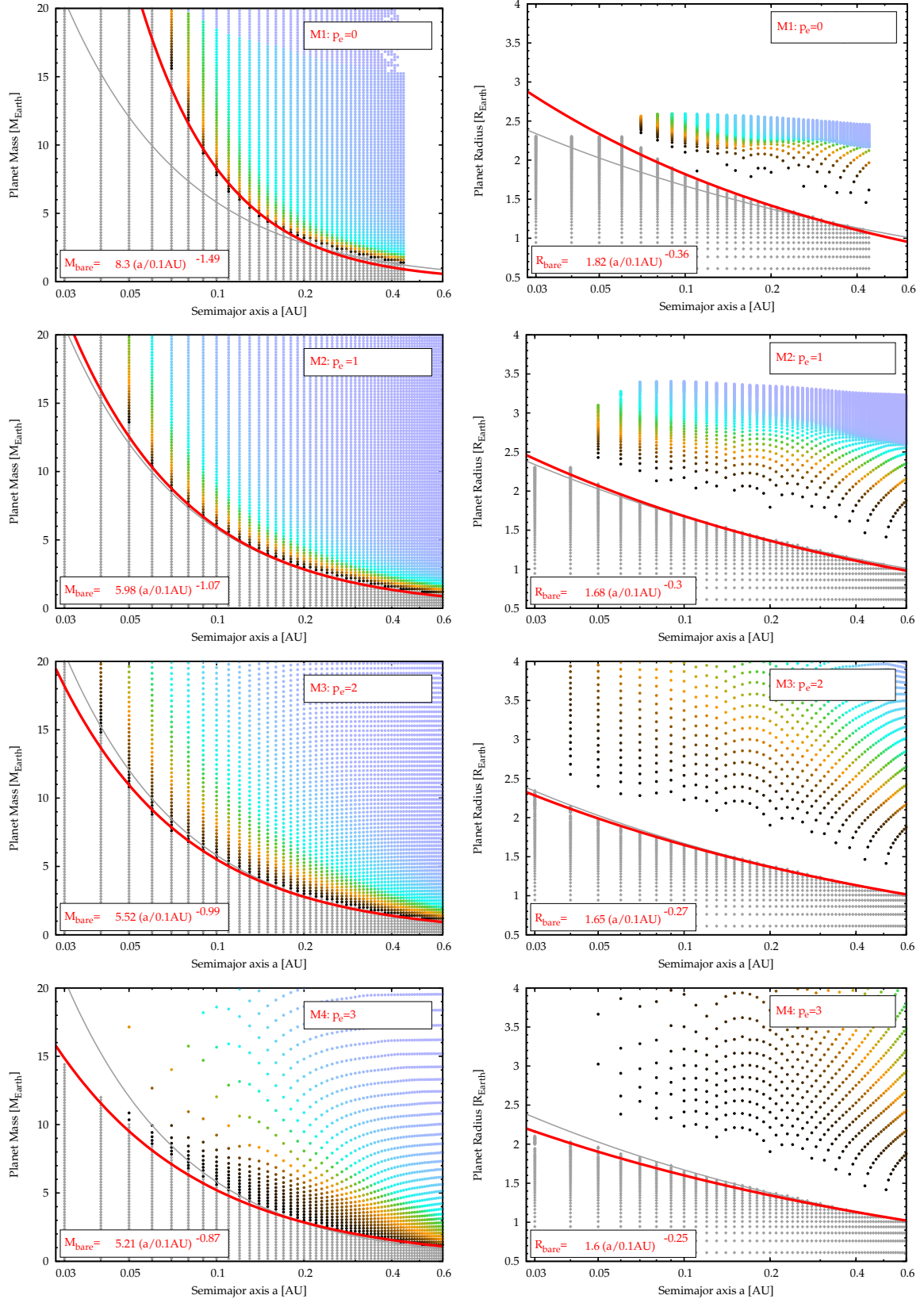
While the valley location is weakly dependent on the postformation envelope mass, far above the valley (i.e., at larger orbital distances or higher planetary masses), the influence of photoevaporation must be weaker and become eventually negligible, so that planets there still have the unaltered initial (postformation) envelope mass. In Figs. 4 and 5, planets that have lost less than about 10% of their initial envelope are shown by blue points.

To further investigate the relation between initial envelope mass and the one after 5 Gyr of evolution, we show in the upper six panels of Fig. 6 the envelope mass fraction  $M_e/M$  as a function of planet mass  $M$  for different orbital distances. Both the envelope mass fractions at 5 Gyr as well as the initial fractions are shown. Inspecting the curves shows that first, the smaller the orbital distance, the higher the planets’ mass relative to  $M_{\text{bare}}$  that still have lost a significant part of their original envelope. Second, the transition from full loss to nearly unaltered envelope masses is more brusque for lower postformation envelope masses (cases M1, M2) than for very high ones (M3, M4). For example, at 0.2 AU, in the case of very large initial  $M_e$  for planets more massive than  $1 M_{\oplus}$  in simulation M4, a planet needs a mass more than  $10 M_{\oplus}$  higher than the  $M_{\text{bare}}$  at this distance (about  $3 M_{\oplus}$ ) to approximately still have the initial envelope mass. In M2 with much lower initial envelope masses, a mass difference only about half as high is sufficient. The same picture is also seen in terms of the radii instead of masses (right panels of Figs. 4 and 5).

The lower six panels of Fig. 6 show the corresponding mass-radius relationships, comparing them to observations. In each panel, the dashed lines correspond to the dashed  $M_e/M$  lines in the upper group of panels. Black dots with error bars show the observed extrasolar planets orbiting stars with masses between 0.7 and  $1.3 M_{\odot}$  with semimajor axes in these six distance intervals. The observational data was downloaded from the NASA Exoplanet Archive on 6 December 2019 and includes all planets that have the radius, mass, semimajor axis and host star mass including the errors listed.

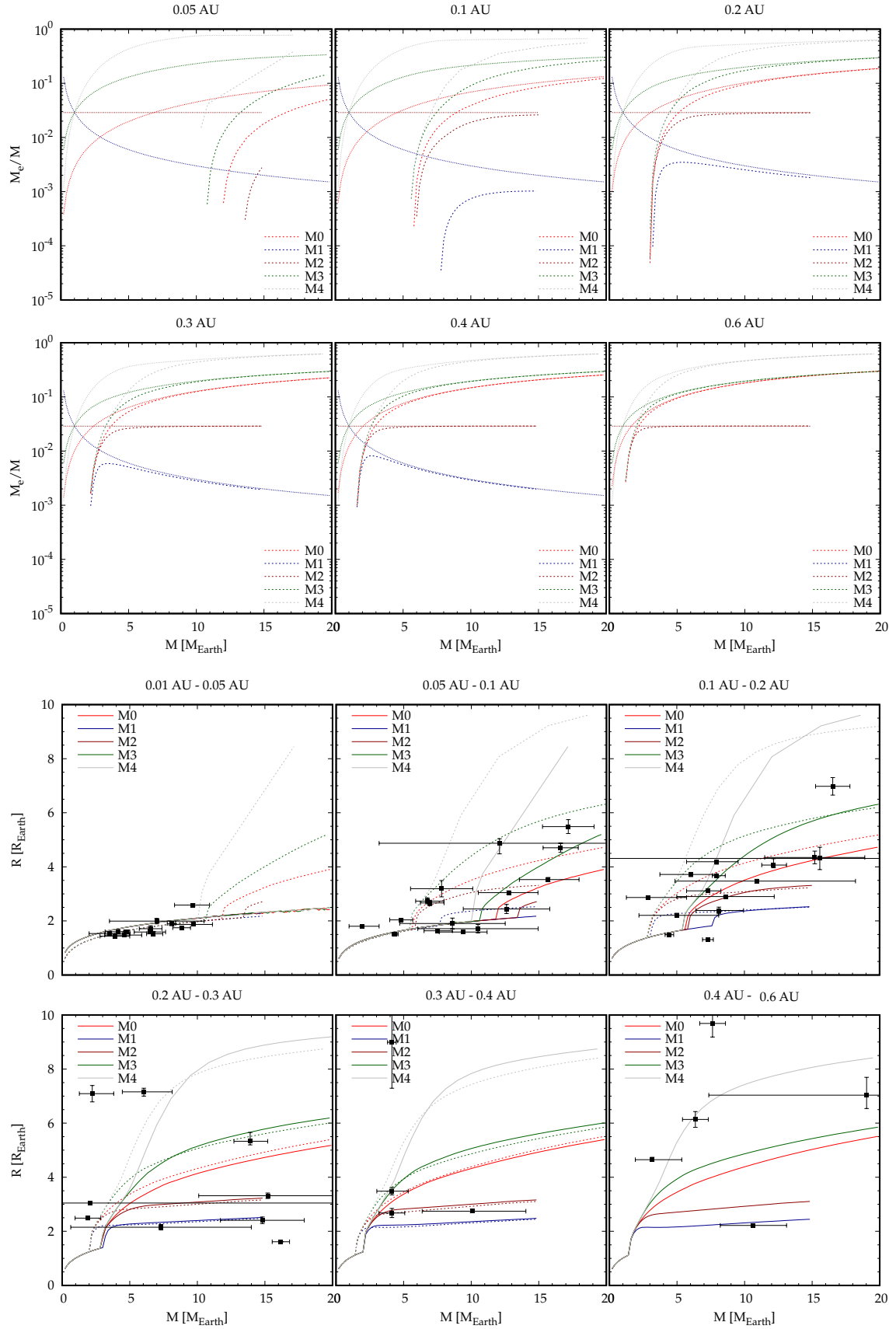
In the panel depicting planets with semimajor axes between 0.01 and 0.05 AU, one sees that with the exception of one planet, there are only objects which do not possess a voluminous envelope. This corresponds to the absence of close-in low-density planets of low and intermediate mass, that is, the (evaporation) desert. These very close-in planets therefore have a small radii. The observed planets approximately following the theoretical mass-radius relation of bare rocky cores, which is given by the solid line (at 0.01 AU, in all simulation M0–M4, all planets with masses less than  $20 M_{\oplus}$  have completely lost their envelopes).

In the distance interval of 0.05–0.1 AU, there is already a much larger spread visible in the  $M$ – $R$  relation. At smaller masses  $\lesssim 10 M_{\oplus}$ , there are planets which continue to approximately fall on the envelope-free  $M$ – $R$  relation. A general trend of increasing radius with increasing mass is seen both in the theoretical and observed data, but with a large scatter. To be compatible with one single theoretical postformation  $M_e(M_c)$  relation like for example the nominal relation of simulation M0, all actual exoplanets would have to fall between the dashed and solid red line. The plot shows that this is not the case. Instead it seems that the actual planets had a diversity of postformation envelope mass that are covered by a combination of the simulation M0, M1, M2, and M3. The very high postformation envelope masses for planets more massive than  $10 M_{\oplus}$  occurring in simulation M4 (which would result in planets with radii larger than  $6 R_{\oplus}$ ) are in contrast not occurring in the (small) observational sample.

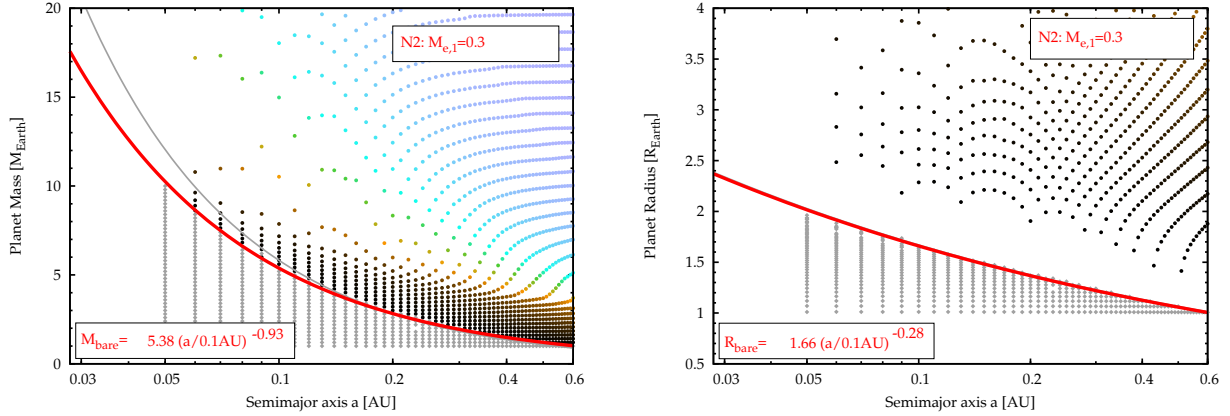


**Fig. 5.** Impact of the core-envelope mass scaling on the transition. The figure is analogous to Fig. 4, but for the simulations M1, M2, M3 (reference simulation), and M4. These 4 simulations are identical except for the power law exponent in the scaling of the initial envelope mass  $M_{e,0}$  with core mass  $M_c$ ,  $M_{e,0}/M_\oplus = 0.03(M_c/M_\oplus)^{p_e}$  with  $p_e = 0, 1, 2$  (reference simulation), and 3. In each case, the transition mass and radius were fit with a power law for  $M_{\text{bare}}$  and  $R_{\text{bare}}$ , as shown with the thick red line. The thinner gray line shows the transition mass and radius in the nominal simulation M0 to allow direct comparison.





**Fig. 6.** Upper six panels: envelope mass fraction  $M_e/M$  as a function of planet mass  $M$  for 6 different orbital distances. The dashed lines show the envelope mass fraction at 5 Gyr, while the dotted lines show the initial (postformation) envelope mass fraction  $M_{e,0}/M$ . The simulations M0–M4 from Figs. 4 and 5 are shown to display the consequences of different initial conditions. Lower six panels: corresponding mass-radius relations compared to observations. In each panel, the solid and dashed lines correspond to the lower and upper limit of the distance interval, respectively. Black dots with error bars show the observed extrasolar planets orbiting stars with masses between  $0.7$  and  $1.3 M_\odot$  in these distance intervals.



**Fig. 7.** Impact of the normalization of the initial envelope mass as found in simulation N2. The figure is analogous to Fig. 4. Compared to the reference case M3, the normalization constant of the envelope mass (i.e., the envelope mass for a  $1 M_{\oplus}$  planet) is here  $M_{e,1} = 0.3$  instead of  $0.03 M_{\oplus}$ . This means that all the initial envelope masses are increased uniformly by a factor 10.

The situation is similar for semimajor axes between 0.1 and 0.2 AU, but the number of planets without H/He decreases. There is also one planet which seems to have had a postformation envelope mass that was higher than in simulation M3 but lower than in M4. This is compatible with a trend of increasing envelope mass fraction with increasing orbital distance which is predicted theoretically (Ikoma & Hori 2012; Lee & Chiang 2015, Sect. 2.1.3).

Moving even further out, we see that there is still a large spread in the observed  $M-R$  relation. In contrast to smaller semimajor axes, however, there are now exoplanets that have very large radii for their mass, similar to or even larger than in simulation M4. In the context of the model, these planets would thus have started with envelope mass fractions even higher than in M4.

In summary, Fig. 6 shows that at distances of up to several 0.1 AU, it is necessary to take into account the reduction of the envelope mass during evolution when using the observed mass-radius relationship of sub-Neptunes to constrain the efficiency of H/He envelope accretion during formation (Mordasini et al. 2014; Lopez & Fortney 2014), as the envelope mass reduction may be significant also for planets above the valley (cf. Estrela et al. 2020). Only planets which are about 5–10  $M_{\oplus}$  more massive than  $M_{\text{bare}}$  at their orbital distance still have envelope masses that differ only weakly from the primordial value. The comparison with the observed mass-radius relation shows that the general trend of a postformation envelope mass which increases with planet mass and orbital distance is present also in the observed population, but that the actual planets were born with a significant spread of envelope masses.

### 2.3.5. Normalization of the initial envelope mass: N1, N2

In the context of the impact of the postformation H/He mass on the locus of the valley, we have also explored the consequences of varying the normalization constant  $M_{e,1}$  when expressing the postformation envelope mass as  $M_{e,0}/M_{\oplus} = M_{e,1}(M_c/M_{\oplus})^{p_e}$  (Eq. (22)). In simulation N1 and N2  $M_{e,0}$  is set to 0.06 and 0.3, corresponding to a factor two and ten increase relative to the value of  $0.03 M_{\oplus}$  used in M0 to M4.

Figure 7 shows the result for simulation N2 (tenfold increase of the envelope masses relative to M3). Despite this large increase, the locus of the transition again only changes marginally. This again shows that the postformation envelope mass is not important for the final position of the evaporation

valley. This is also predicted by the analytical model, where no dependency at all on  $M_{e,1}$  is found. The simulation N1 where the normalization mass is twice as large as in the reference case is not shown as it is even closer to nominal simulation (but it is included in Table 1).

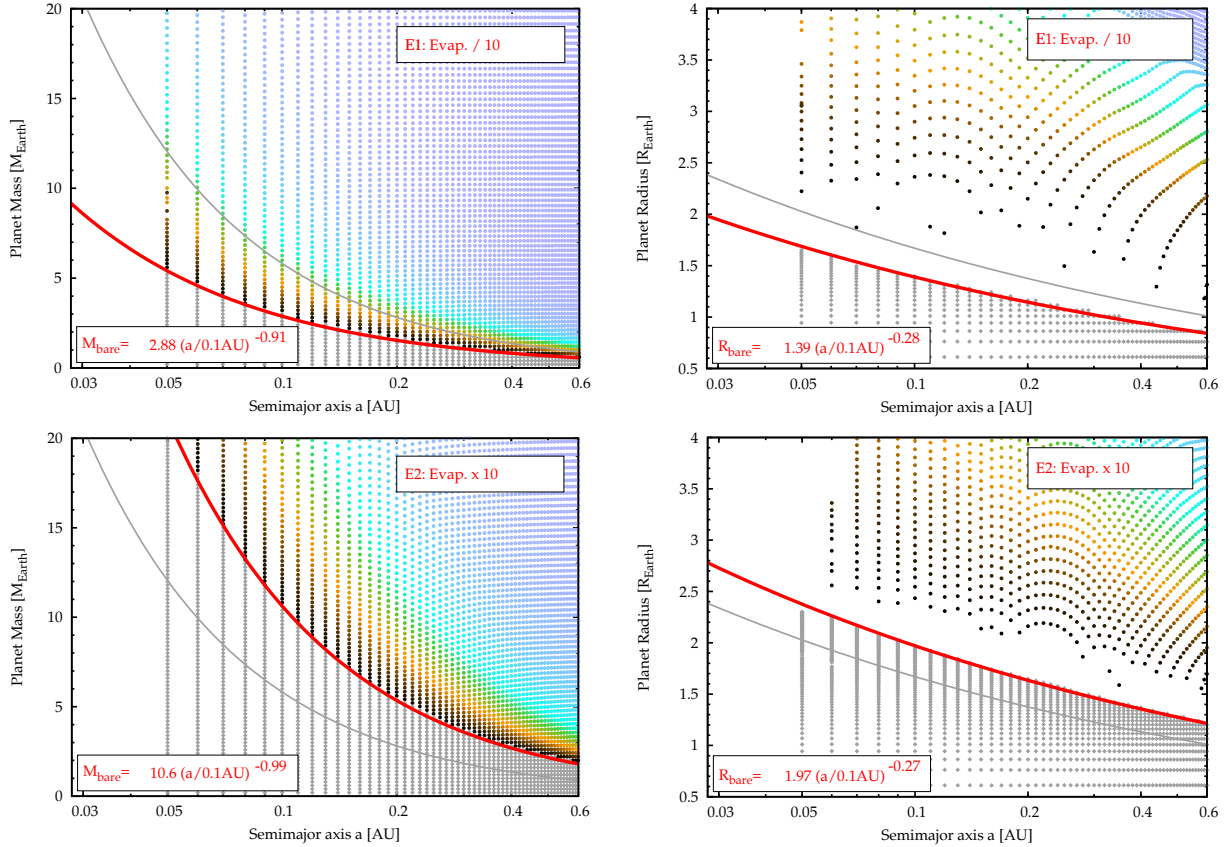
### 2.3.6. Evaporation rate: E1, E2

The evaporation model used in this work is relatively simple, as it does not directly solve the conservation equations to obtain the escape rate as for example in Murray-Clay et al. (2009), but instead uses the energy or radiation-recombination-limited formulae to calculate the loss rate, assuming constant efficiency factors. It also neglects the consequences of different atmospheric compositions or magnetic fields. It is clear that our escape rates are therefore only rough estimates.

Furthermore, it is well known that even at a fixed stellar mass of around  $1 M_{\odot}$ , stellar rotation rates (e.g., Johnstone et al. 2015) and associated  $L_{\text{XUV}}$  luminosities exhibit a spread of about a factor  $\sim 30$  at young ages (Tu et al. 2015) when most of the escape occurs. It is, however, also worth noting that most stars still cluster around a typical rotation period of about 5 days (Johnstone et al. 2015), where  $L_{\text{XUV}} \approx 10^{30} \text{ erg s}^{-1}$  during the first  $\sim 100$  Myr (Tu et al. 2015), and only a small fraction has markedly different rotation rates. Nevertheless, given this observed spread, it is therefore even more important to investigate the impact of a variable strength of the evaporation.

To account for this, we reduce in simulation E1 the evaporation rate predicted by the model uniformly by a factor 10, while in simulation E2 we increase the rate uniformly by such a factor, so that the consequences of an escape rate varying by two orders of magnitudes can be assessed. Physically, this variation could be due to the (combined) effects introduced by different stellar  $L_{\text{XUV}}$ , different durations of the saturated phase, or variations in the efficiency factors of escape. By simply scaling the evaporation rates, we do not need to specify this explicitly.

Figure 8 shows the result of the E1 and E2 simulations. The higher (lower) evaporation rate leads to a higher (lower) transition mass and radius, as expected. For the low evaporation rate, relative to the reference simulation M3 (from which E1 and E2 only differ by the evaporation rate), the  $M_{\text{b,0p1}}$  has decreased from 5.52 to 2.88  $M_{\oplus}$ , corresponding to a reduction by a factor 1.92. The radius  $R_{\text{b,0p1}}$  has decreased from 1.65 to 1.39  $R_{\oplus}$ , a factor 1.19 change.



**Fig. 8.** Impact of the strength of evaporation on the transition. The figure is analogous to Fig. 4 but for simulation E1 (*top panels*) and E2 (*bottom panels*) where the evaporation rate is uniformly reduced (E1) respectively increased (E2) by a factor 10 relative to the nominal case. The transition mass and radius were again fit with a power law for  $M_{\text{bare}}$  and  $R_{\text{bare}}$  as shown with the thick red line. The power law exponents for the dependency on the distance have remained similar as in the reference case M3, but the absolute value of  $M_{\text{bare}}$  is reduced in E1 (weak evaporation) relative to M3 by a factor 2.02, and a factor 1.20 for  $R_{\text{bare}}$ . In E2 (strong evaporation),  $M_{\text{bare}}$  and  $R_{\text{bare}}$  have increased by a factor 1.82 and 1.18, respectively. The thinner gray line shows the transition mass and radius in the nominal simulation M0 to allow direct comparison. M3 is in turn very similar to M0.

For the high evaporation rate,  $M_{\text{b},0\text{p}1}$  has increased from 5.52 to 10.6  $M_{\oplus}$ , corresponding to decrease also by a factor 1.92. The radius  $R_{\text{b},0\text{p}1}$  has increased from 1.65 to 1.97  $R_{\oplus}$ , again a factor 1.19 change. So the total change in  $M_{\text{b},0\text{p}1}$  by a factor 3.68 for a variation of a factor 100 in evaporation rate suggests a weak dependency on the factors determining the evaporation rate like  $L_{\text{XUV}}$  or  $\varepsilon$  that scales only as approximately  $M_{\text{b},0\text{p}1} \propto L_{\text{XUV}}^{0.28}$ . For comparison, the analytical model predicts a dependency like  $L_{\text{XUV}}^{0.5}$  (Eq. (30)). Similarly, the total variation in radius by a factor 1.97/1.39 = 1.42 suggests a very weak power law dependency just like approximately  $L_{\text{XUV}}^{0.08}$ . Analytically, we find for comparison  $R_{\text{b},0\text{p}1} \propto L_{\text{XUV}}^{0.135}$  (Eq. (32)). While the exact values of the exponents are certainly dependent on the details of the model used here, we consistently find weak dependencies, in particular for the transition radius.

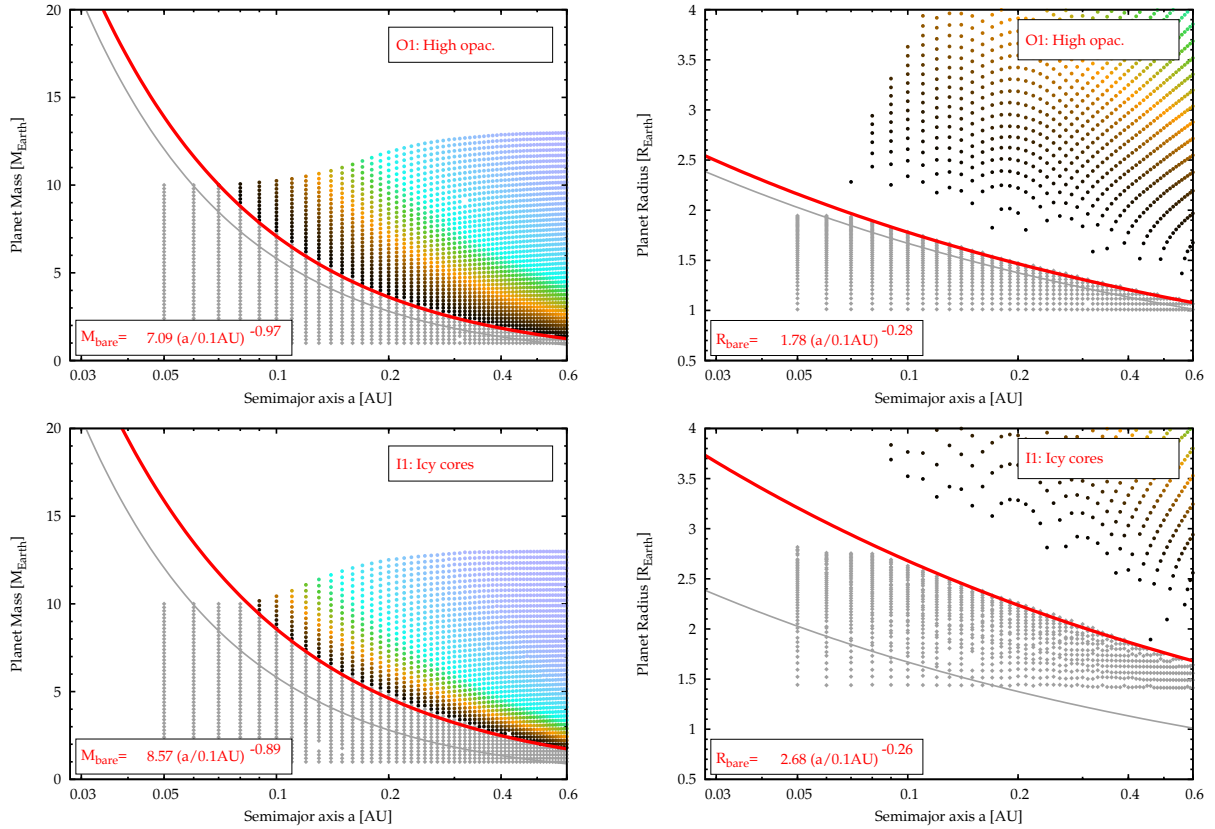
This has a very important observational implication: From the location and width of the valley in the two simulations it becomes apparent that if we would overlay the two simulations, a depleted region would appear, but it would not be completely empty: regions completely devoid of planets in E1 would be partially populated by planets from E2, and vice versa. This is the case because the change in the valley’s location (1.39 versus 1.97  $R_{\oplus}$  at 0.1 AU in E1 and E2, respectively, corresponding to a  $\Delta R = 0.58 R_{\oplus}$ ) is comparable to the intrinsic width  $\Delta W$  of the valley (about 0.5  $R_{\oplus}$ , see also Owen & Wu 2017). If we would instead have  $\Delta R \gg \Delta W$  (as it would be the case if the

dependency on  $L_{\text{XUV}}$  would be stronger, as one could maybe naively expect from Eq. (9)), the valley would not be well visible (or even not at all), as it would scatter too much from star to star. If we would have the other extreme,  $\Delta R \ll \Delta W$  (for example if the strength of evaporation would be identical in all systems), then the valley would be completely empty. Both these things are not observed. But if in nature the location of the valley varies from star to star (because of different  $L_{\text{XUV}}$ , but also different efficiency factors for example because of different atmospheric compositions) indeed by about 0.5  $R_{\oplus}$  around the mean value as we find here, while the intrinsic width  $\Delta W$  has a similar magnitude (again as found here), then we deduce that we should still see a depleted valley, but not a completely empty one. This corresponds to the observational result (Fulton & Petigura 2018). Clearly, in these *Kepler* observations, we see the overlay of all individual system-specific valleys.

### 2.3.7. Atmospheric opacity: O1

The top panels of Fig. 9 show the consequences for the locus of the valley resulting from increasing the Rossland mean opacity  $\kappa_{\text{R}}$  in all planets uniformly by a factor 10 relative to the nominal case. In the nominal case, we have used solar-composition grain-free opacities from Freedman et al. (2014) when calculating the cooling and contraction of the planets. In the solar system, the atmospheric (and bulk) enrichment in metals is increasing with





**Fig. 9.** Same as Fig. 4 but it shows the impact of the atmospheric opacity (O1, *top panels*) and of the composition of the solid core (I1, *bottom panels*). These simulations cover a smaller range in initial core masses than the other ones. In simulation O1 the atmospheric opacity is uniformly increased by a factor 10. In simulation I1 an ice mass fraction in the core of unity instead of an Earth-like composition was assumed. The thin gray line shows the transition mass and radius in the nominal simulation M0 to allow direct comparison.

decreasing planetary mass (Mordasini et al. 2016). This is also predicted by planet formation models based on the core accretion paradigm both for the bulk (Fortney et al. 2013) and (under the assumption of efficient mixing of atmosphere and envelope) atmospheric composition (Mordasini et al. 2016).

In extrasolar planets, the bulk metallicity also increases with decreasing planet mass, similarly as in the solar system (Thorngren et al. 2016). Regarding the atmospheric metallicity, the situation seems different: Fisher & Heng (2018) find no trend in the retrieved atmospheric water abundances across nearly two orders of magnitude in exoplanet mass. Wallack et al. (2019) also see no evidence for a solar-system-like correlation between planet mass and atmospheric metallicity. Both studies do however find a large spread in atmospheric enrichments. It should also be noted that the large majority of the planets studied in these works are more massive giant planets far from the valley. In any case, these findings indicate that there could be a large diversity in atmospheric compositions, meaning that the opacity in the low-mass planets we are studying here could be higher than solar.

A higher opacity in the atmospheres delays the cooling and contraction of the planets (e.g., Burrows et al. 2007), reducing their mean density, which leads to a higher evaporation rate (Eq. (9)). We thus expect that at high opacity, the valley should move to higher  $M_{\text{bare}}$  and  $R_{\text{bare}}$  at fixed distance compared to the nominal case. Figure 9 shows that this is indeed the case. But we also see that the difference is small. Relative to M3, the reference simulation which differs from O1 only by the enhanced opacity, the transition mass grows by a factor  $7.09/5.52 = 1.28$ , while for the radius forming the bottom of the valley, an increase by

just  $1.78/1.65 = 1.08$  is observed. These are small factors given the tenfold increase of  $\kappa_R$ , which would correspond to a power law dependency approximately  $\propto \kappa_R^{0.03}$  for the radius. This weak impact could be related to a certain auto-regulation in the sense that at fixed mass, a larger radius (caused by higher  $\kappa_R$ ) leads to a higher evaporation rate, which reduces the envelope mass, which in turn reduces the radius, and eventually the evaporation rate.

A tenfold increase of the opacity approximately corresponds to an atmospheric metal enrichment of about  $10\text{--}30 \times$  solar (Freedman et al. 2014). In reality, the exact enrichment level corresponding to such an opacity increase depends on pressure and temperature (Freedman et al. 2014) and is not uniform. For the  $5\text{--}10 M_{\oplus}$  planets we are mainly dealing with, the atmospheric enrichment could be even higher than  $10\text{--}30 \times$  solar (Mordasini et al. 2016). In the solar system for example, Uranus and Neptune are enriched in carbon (observed as  $\text{CH}_4$  in the atmosphere) by a factor of about 80 relative to the Sun (Guillot & Gautier 2014). So we may expect enrichments level on the order of  $100 \times$  solar. But the weak dependency on  $\kappa_R$  found in O1 indicates that even for such a high enrichments, there would not be a very large shift of the valley.

### 2.3.8. Metal enrichment of the gaseous envelope: Z1–Z3

In simulation O1, we have increased the opacity, but not modified two other quantities that in reality also change for a higher amount of metals: first, the mean molecular weight or more generally speaking, the equation of state. At low enrichments levels ( $\leq 10 \times$  solar), the opacity is already increased, but not yet significantly so the mean molecular weight. But if we go to

more significant enrichment levels (several  $10 \times$  solar), the effect on the mean molecular weight becomes important as well. For example, if we approximate all the metals as water vapor, we find a mean molecular weight of about 2.4, 2.6, and 5.3 for 1, 10, and 100 times solar. This clearly higher value for  $100 \times$  solar leads to a smaller planetary radius because of the increased density of the gas (e.g., Baraffe et al. 2008). This effect counteracts the increase of the radius that is found when just increasing the opacity, but not self-consistently increasing the mean molecular weight also (which is what we did in Simulation O1, because the EOS of the gas is still the pure H/He EOS of Saumon et al. 1995 independently of  $\kappa_R$ ).

Second, we have also not taken into account in simulation O1 that the higher opacities are caused by higher quantities of heavy elements (i.e., the composition of the atmosphere), and composition influences the atmospheric evaporation rate, too (e.g., Johnstone et al. 2018). Cooling via atomic lines of important metals like carbon or oxygen for example reduces the temperature in the XUV heated gas which leads to lower evaporation rates (Owen & Murray-Clay 2018). While the effect still needs to be systematically quantified in the context of strongly evaporating planetary atmospheres (Owen & Murray-Clay 2018), it appears likely that this effect also counteracts the stronger evaporation associated with higher opacities because of the aforementioned increase of the radius.

Clearly, given these opposing effects, the role of the atmospheric composition needs to be treated more self-consistently than in simulation O1 by linking the opacity, mean molecular weight (i.e., the EOS), the atmospheric and interior structure, and the efficiency of atmospheric escape self-consistently. This is done in simulations Z1, Z2, and Z3, where the opacity, EOS, and evaporation rate were self-consistently coupled, as described in Sect. 2.1.2. The three simulations study an envelope heavy element mass fraction  $Z_{\text{enve}} = 0.1, 0.3, \text{ and } 0.5$ , respectively.

These simulations are also of interest because observationally, it is possible to study the valley position as a function of host star [Fe/H] (Owen & Murray-Clay 2018). As discussed below, the difficulty in connecting the simulations shown here with these constraints lies in the question whether (or to what extent) the atmospheric compositions of the low-mass planets near the valley are correlated with the host star [Fe/H].

The resulting distance-mass and distance-radius plots are shown in Fig. 10. We see that in contrast to simulation O1, where the enhanced opacity has led to a (slight) increase of  $M_{\text{bare}}$  and  $R_{\text{bare}}$ , we now see that the higher  $Z_{\text{enve}}$ , the lower  $M_{\text{bare}}$  and  $R_{\text{bare}}$ . The valley thus shifts slowly downward with increasing metallicity. The two effects by which a higher  $Z_{\text{enve}}$  reduces the evaporation rate (first by increasing the mean density of a planet at fixed core and envelope mass, and second by reducing the efficiency factor of evaporation) overwhelm thus the effect that the associated higher opacity leads via a reduce cooling to less dense planets and thus more evaporation. This downward shift was expected given that Lopez (2017) found that for pure water envelopes ( $Z_{\text{enve}} = 1$ ), the valley should be at about  $1 R_{\oplus}$  at an orbital distance of 0.1 AU.

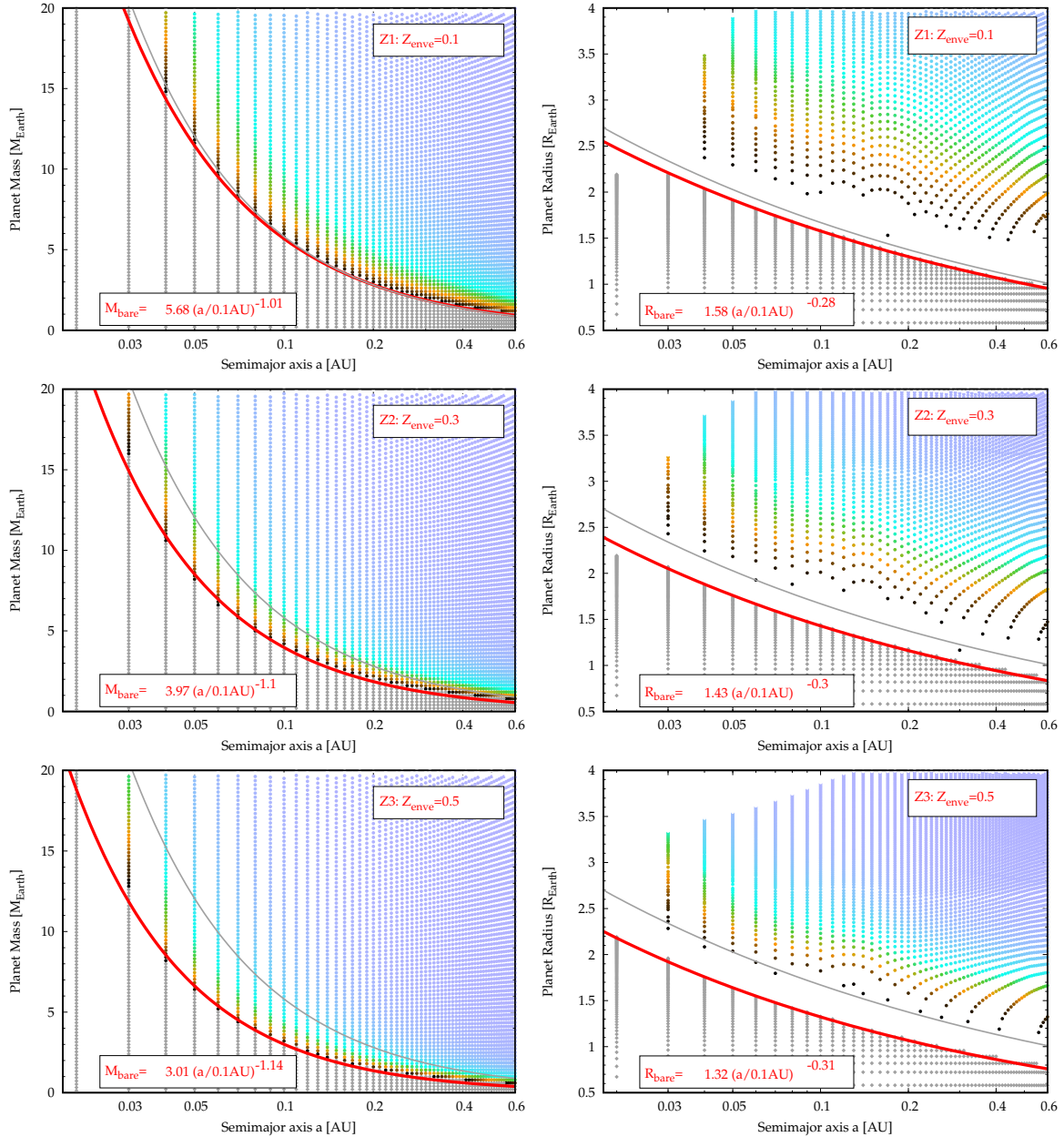
One also sees in Fig. 10 that at  $Z_{\text{enve}} = 0.1$ , the innermost orbital distance where planets with an envelope survive is 0.04 AU. The same result holds for the nominal simulation M0. At higher  $Z_{\text{enve}}$  in simulations Z2 and Z3, planets which have kept the envelope exist in contrast also at 0.03 AU. This is not surprising: the same mechanisms which reduce the efficiency of evaporation at high  $Z_{\text{enve}}$  not only shift the valley downward, but also push the limit of the sub-Neptune desert inward to smaller orbital distances. If the envelope metallicities  $Z_{\text{enve}}$  of planets are

indeed correlated with the stellar [Fe/H], this result is in agreement with the observational finding of Owen & Murray-Clay (2018) that planets hosting H/He envelopes are more common around higher metallicity stars at small orbital distances.

Owen & Murray-Clay (2018) also found that observationally, the locus of the valley is approximately independent of host star metallicity with a change of  $\lesssim 15\%$  in radius over the observed range of metallicities. For a valley position in the nominal simulation at about  $1.7 R_{\oplus}$ , 15% corresponds to a change of less than about  $0.3 R_{\oplus}$ , bringing the valley on the lower limit down to about  $1.4 R_{\oplus}$ . This value lies between the  $R_{\text{bare}}$  found in simulations Z2 and Z3 ( $1.43$  and  $1.32 R_{\oplus}$ ). Again under the assumption of a correlation of host star metallicity and planetary  $Z_{\text{enve}}$ , in order to be consistent with this limit, the envelopes in the valley region should not exhibit a systematic change in  $Z_{\text{enve}}$  with stellar [Fe/H] exceeding about 0.4. If we follow Gupta & Schlichting (2020) and assume that the stellar metallicities are directly representative of the metallicity of the planetary atmospheres, we typically expect a change in  $Z_{\text{enve}}$  because of the variation of stellar [Fe/H] (about  $-0.5$  to  $0.5$  dex in the solar neighborhood) of  $f_{\text{vol}} Z_{\odot} \times 10^{-0.5}$  to  $f_{\text{vol}} Z_{\odot} \times 10^{0.5}$ . Here,  $f_{\text{ref}}$  is the mass fraction of volatiles which are in the gas phase in the inner disk, and  $Z_{\odot}$  is the primordial solar heavy element mass fraction. The value of  $f_{\text{ref}}$  is about 0.67 and  $Z_{\odot} = 0.0149$  (Lodders 2003). Entering these number gives  $Z_{\text{enve}}$  ranging between 0.003 and 0.03. This is a very small range compared to the one needed to shift the valley significantly. Considering how  $R_{\text{bare}}$  varies across the simulations Z1–Z3, variations between 0.003 and 0.03 shift the valley position only by about  $0.02 R_{\oplus}$ . This is much less than the aforementioned observational limit of  $0.3 R_{\oplus}$ , meaning that if the host star [Fe/H] indeed sets the planetary  $Z_{\text{enve}}$ , the weak dependency of  $R_{\text{bare}}$  on  $Z_{\text{enve}}$  in the evaporation hypothesis for the valley is such that it is in agreement with the observed independency of the valley locus on host star metallicity.

An alternative explanation of the absence of a correlation of the valley position and the stellar [Fe/H] would be that there is no or only a weak correlation of host star metallicity and planetary envelope/atmospheric metallicity for planets in the valley region. For (massive) giant planets where gas is the clearly dominant component it seems likely that the disk gas metallicity, which should be correlated with the host star metallicity, is important for the atmospheric metal content (Mordasini et al. 2016). Observationally, Wallack et al. (2019) indeed find for giant planets a  $1.9\sigma$  trend of a correlation of planetary atmospheric metallicity and stellar metallicity. Surprisingly, Teske et al. (2019) find for the bulk composition in contrast no correlation between stellar [Fe/H] and planetary metallicity if one corrects for the effect that overall, planetary metallicity decreases with increasing planet mass. One should here keep in mind that there are limits to a situation where the atmospheric metallicity is correlated with host star [Fe/H] whereas the interior is not: at some point, such a configuration becomes unstable because density and pressure gradients would have opposite signs, which is Raleigh-Taylor unstable (Thorngren & Fortney 2019).

For the much lower mass planets near the valley, there are reasons to think why the envelope composition may not simply be set by the stellar metallicity: first, the pollution of the atmospheres by the accretion of planetesimals and pebbles (e.g., Podolak et al. 1988; Mordasini et al. 2006; Pinhas et al. 2016; Brouwers et al. 2018; Valletta & Helled 2019) is proportionally more important as these planets contain in mass only a small gas fraction. The amount of metals contained in the accreted disk gas without pollution is likely low, roughly on a 0.3–3% level, at least if the gas was accreted before the disk became



**Fig. 10.** Impact of the envelope enrichment on the transition. The figure is analogous to Fig. 4 but for simulations Z1, Z2, and Z3 where envelope is enriched with  $\text{H}_2\text{O}$  with a mass fraction  $Z_{\text{enve}} = 0.1, 0.3$  and  $0.5$ , respectively. Opacities and evaporation rates are modified accordingly, assuming a simple scaling law for the energy-limited evaporation rate  $\propto Z_{\text{enve}}^{-0.77}$ . The transition mass and radius were again fit with a power law for  $M_{\text{bare}}$  and  $R_{\text{bare}}$  as shown with the thick red line. The thinner gray line shows the transition mass and radius in the nominal simulation M0 to allow direct comparison.

chemically evolved (Guillot & Hueso 2006). This means that the pollution by a relatively low amount of metals is sufficient to dominate the composition. When impactors starts to pollute the envelope is however rather intrinsic to a planet (the envelope must be sufficiently massive) and not dependent on the host star  $[\text{Fe}/\text{H}]$ . Second, another form of pollution is the outgassing of various chemical species from the solid core (e.g., Bower et al. 2019) which would again be a process intrinsic to a planet. For terrestrial planets, secondary outgassed atmospheres differ completely from the host star composition and in this sense the stellar  $[\text{Fe}/\text{H}]$ . The case of the planets near the valley with a few percents of their mass in  $\text{H}/\text{He}$  is obviously not as extreme as they still posses (parts of) their nebular atmosphere. But it could still be that the fact that these envelopes are over Giga-years in contact

with a magna ocean influences the atmospheric composition more than the stellar  $[\text{Fe}/\text{H}]$ .

The finding of Owen & Murray-Clay (2018) that  $\text{H}/\text{He}$  atmosphere-hosting planets are more common around higher metallicity stars at short periods would then to be interpreted differently than that the high(er) atmospheric metallicity protects the atmosphere from photoevaporative loss. An alternative explanation could be the formation timescale: if planets around low  $[\text{Fe}/\text{H}]$  stars form later and therefore accrete less or even no  $\text{H}/\text{He}$ , this could potentially also explain this trend.

In the case that there is no correlation between stellar  $[\text{Fe}/\text{H}]$  and planetary atmospheric  $Z_{\text{enve}}$ , the variation in  $Z_{\text{enve}}$  from planet to planet – if sufficiently large – would then simply induce a spread in the transition from solid to planets with  $\text{H}/\text{He}$ , as



suggested by the results in simulations Z1–Z3. Observationally, this would make the valley less empty.

To finish this discussion it is important to stress that we have adopted here a very simple and uncertain scaling of the evaporation rate with  $Z_{\text{enve}}$ . This should be improved in future work. We have also used the same initial conditions for all  $Z_{\text{enve}}$ . If possible it should however be taken into account that different compositions affect already a planet’s formation, and therefore the initial conditions like for example the envelope mass fraction which depend on the envelope composition (Venturini et al. 2016) or the initial luminosity.

### 2.3.9. Rocky versus icy composition of the core: I1

The last simulation I1 is identical to M3 except for a mass fraction of (water) ice in the solid core that is equal to unity. In M3, an Earth-like composition without any ice was instead used. A core consisting exclusively of water ice like in I1 is certainly unlikely to occur in nature (if ice can condensate in some part of the protoplanetary disk, then also more refractory element). However, it is interesting to compare the results with earlier studies (Owen & Wu 2017; Jin & Mordasini 2018) and with the analytical model.

Owen & Wu (2017); Jin & Mordasini (2018) had both independently found that the locus of the valley as observed by Fulton et al. (2017) is consistent with a predominately Earth-like composition for the close-in low-mass planets probed by the valley, but not with a predominantly icy composition (see also Owen & Adams 2019 for the effects of magnetic fields). For an ice-rich composition, the valley would rather be located at about  $2.3 R_{\oplus}$  at 0.1 AU instead of the observed  $1.6 R_{\oplus}$  (Jin & Mordasini 2018). The reason for the larger  $R_{\text{bare}}$  for icy cores is that first, an ice-rich composition reduces the mean density of the planet (when the planet still posses H/He), making planets of a higher core mass more vulnerable to escape. Second, once the envelope is lost, at fixed core mass, icy cores have a larger radius.

A transition at  $1.6 R_{\oplus}$  as observed is instead consistent with an Earth-like iron-silicate composition. This indicates that these planets did not migrate to their current position from outside of the (water) iceline. As discussed extensively in Jin & Mordasini (2018), this does, however, also not mean that these planets did not migrate at all, but that their migration was confined to inside of the water iceline. This is also what is predicted by planet formation simulations that couple planetary accretion, (full non-isothermal) type I (and type II) migration, and disk evolution (e.g., Mordasini 2018). In these simulations, only more massive Neptunian planets migrate from beyond the iceline close to the star, as the type I migration timescale decreases with increasing mass. But these more massive planets can keep their H/He, except if they are very close to the star, like 55 Cnc e (Jin & Mordasini 2018).

It is also clear that the some individual planets can still have a high ice content as the observations only show the mean locus of the valley, but ice-rich should not be the dominant composition. It is interesting to note that thanks to evaporation and the resulting valley, it was possible for Owen & Wu (2017); Jin & Mordasini (2018) to infer Earth-like compositions based on the location of the valley alone, that is, based on radii only. The density of the planets which is usually used to infer such informations is not needed (and also not available for most Kepler planets).

As shown in Jin & Mordasini (2018), an Earth-like composition is also consistent with the (measured) densities of the (few) planets in the triangle of evaporation with a relatively

well-constrained density (i.e., planets for which both radius and mass were relatively precisely measured). The triangle of evaporation is the approximately triangular-shaped region (in a log–log plot) in the planet radius-orbital distance plane where planets completely lose the H/He, that is, planets with  $R < R_{\text{bare}}(a)$ . For such planets without H/He envelopes, the degeneracy of the mass-radius relation (Rogers & Seager 2010) is strongly reduced, and the density can be used to constrain the ice content. Unfortunately, only very few planets are currently known that have a density that is observationally constrained well enough to really infer the ice mass fraction (see also Dorn et al. 2017).

To improve this situation, several tens of planets with masses below  $10 M_{\oplus}$ , radii below  $4 R_{\oplus}$ , and inside of about 0.5 AU with very accurately measured masses and radii, for example from TESS (Ricker et al. 2014) and CHEOPS (Fortier et al. 2014) observations, would be very helpful.

Coming back to the two bottom panels of Fig. 9, we see that relative to the reference simulation M3,  $M_{\text{bare}}$  increases from 5.52 to  $8.57 M_{\oplus}$  (factor 1.55 increase) for a fully icy composition, while  $R_{\text{bare}}$  grows from 1.65 to  $2.68 R_{\oplus}$ , corresponding to a factor 1.62. We thus see that especially for the radius, the ice mass fraction has a strong influence on the transition locus. It leads, with a large margin, to the strongest increase of  $R_{\text{bare}}$  relative to the reference case of all simulations conducted in this work as can be seen from Table 1. The impact of an icy core is in particular also larger than the one of varying the evaporation rate by a factor of 10.

The shift from about 1.6 to  $2.7 R_{\oplus}$  is also more than twice as large as the intrinsic width of valley of about  $0.5 R_{\oplus}$ . This remains true also for a more realistic ice mass content of ice of 50–75% (Jin & Mordasini 2018). This is very different from the results for the impact of various initial envelope masses and strengths of evaporation presented in the previous sections. This strong impact of the composition explains why the valley is such a good probe for composition of the cores, as first suggested by Lopez & Fortney (2013).

While the (double) physical effect leading to a shift of  $R_{\text{bare}}$  to higher radii for an icy core composition was qualitatively explained earlier in this section, it will become possible to more quantitatively assess the impact of the ice mass fraction in the context of the analytical model (Sect. 3.3.2). There, it is found that the  $R_{\text{bare}}$  radius depends linearly on the ice mass fraction, whereas it depends only weakly on other quantities, for example only as  $L_{\text{XUV}}^{0.135}$  on the stellar XUV luminosity.

## 3. Analytical study

We now develop a simple analytical criterion for the evaporative transition from close-in super-Earth to sub-Neptunes that can explain the main results found in the numerical simulations. It explains in particular (i) the decrease of the transition mass from solid planets to planets with H/He  $M_{\text{bare}}$  with approximately  $a^{-1}$  or equivalently  $R_{\text{bare}}$  with  $a^{-0.27}$ , (ii) the weak dependency on the efficiency of atmospheric escape (Jin et al. 2014), and (iii) the independency of  $R_{\text{bare}}$  on the (power law) scaling of the initial envelope mass with core mass. The dependency of the transition on these quantities were previously investigated numerically in Lopez & Fortney (2013).

### 3.1. Energy comparison principle

The analytical model is based on the principle that the ability to keep an envelope is controlled by the comparison of the relevant energies: it compares the binding energy of the envelope

in the potential of the core with the energy deposited by the time-integrated XUV flux received by the upper atmosphere due to stellar XUV irradiation (the lifetime-integrated X-ray and extreme ultraviolet flux in the words of McDonald et al. 2019). Then, the mass of H/He envelope that is evaporated at some moment in time is given as the mass which has the same binding energy as the time integrated absorbed stellar XUV flux received up to that time by the planet (modulo the efficiency factor).

The same basic idea was first used by Lecavelier Des Etangs (2007) to study the evaporation of gaseous giant planets (for other energy-based studies for giant (i.e., H/He gas-dominated) planets, see also Ehrenreich & Désert 2011 and Jackson et al. 2012). In this work, we deal in contrast with low-mass, core-dominated planets. Therefore, we need to modify Lecavelier's approach in two crucial points:

(i) instead of the total binding energy of the planet (including all mass), the relevant quantity is in our case the binding energy of the gaseous envelope in the gravitational potential of the core (the core itself does not evaporate, but is much more massive than the envelope and thus dominates the gravitational potential).

(ii) the mass-radius relation is different from giant planets and more complex, as it depends on both the core and envelope mass. The radius of a low-mass planet depends in particular strongly on the envelope mass fraction (e.g., Lopez & Fortney 2014). This is very different from giant planets where the radius is always around 1 Jovian radius and only weakly dependent on the (total) mass over a large mass range because of increasing degeneracy in the interior (e.g., Baruteau et al. 2016).

These two differences lead to a more complex behavior than in the three aforementioned energy-based studies that addressed the evaporation desert important for (at least initially) gas-dominated planet, and not the evaporation valley of Fulton & Petigura (2018); Van Eylen et al. (2018) that is relevant for low-mass planets starting only with a low-mass H/He envelope.

### 3.2. Derivation of $M_{\text{bare}}$ and $R_{\text{bare}}$

Based on this idea, we now derive the analytical expressions for  $M_{\text{bare}}$  and  $R_{\text{bare}}$  which is the mass and radius of the most massive planet losing all its primordial H/He envelope at a given orbital distance. This corresponds to the bottom of the valley. The binding energy  $U_e$  of the gaseous envelope with mass  $M_e$  in the gravitational potential of the core with mass  $M_c$  can be approximated as

$$U_e = - \frac{k_{\text{pot}} G M_c M_e}{R}, \quad (17)$$

where  $G$  is the gravitational constant and  $k_{\text{pot}}$  a number on the order of unity which depends on the density structure. Since we are dealing with low-mass, core-dominated planets where  $M_e \ll M_c$ , we neglect the self-gravity of the envelope, but we do take into account that the (outer) radius of the planet can be much larger than the core radius even for  $M_e \ll M_c$ . Therefore, we use in the equation the total planetary radius  $R$  and not just the core radius  $R_c$ .

The cumulative (time integrated) XUV energy input that drives evaporation - assuming energy-limited escape - since the formation of the planet at  $t_{\text{form}}$  up to some time  $t$  is given as

$$V_{\text{XUV}} = \int_{t_{\text{form}}}^t \varepsilon \pi R^2 \frac{L_{\text{XUV}}}{4\pi a^2} dt. \quad (18)$$

In this equation,  $\varepsilon$  is the fraction of the incoming stellar XUV flux that drives evaporation (the efficiency factor),  $a$  the

planetary semimajor axis,  $L_{\text{XUV}}$  the stellar XUV luminosity, and  $R$  is again the planetary radius. For the analytical model, we assume that the radius where the XUV radiation is absorbed is the same as the radius in Eq. (17), neglecting a potential difference (Murray-Clay et al. 2009). All these quantities are variable in time: The stellar  $L_{\text{XUV}}$  decreases in time (e.g., Ribas et al. 2005; Tu et al. 2015),  $a$  could potentially change due to tides or  $N$ -body interactions, and  $R$  decreases due to the cooling and contraction of the planet, and because of the decrease of the envelope mass due to escape. For the analytical model, we retain only the temporal dependency of  $L_{\text{XUV}}$ , and assume that  $a$  is constant, while for  $R$  we then need to specify a representative mean radius (during the relevant time interval), as described below. For this temporal mean, we must keep in mind that early after formation, the planetary radii are much larger (Mordasini et al. 2012b).

Using the observational results for the EUV flux as a function of time of G and K stars (Ribas et al. 2005), we can evaluate the integral that now only runs over  $L_{\text{XUV}}$ . These studies have shown that the flux decreases in time as a power law  $L \propto t^{-\alpha}$ , where  $\alpha \sim 1.23$  for G type stars and 1–1200 Å, with a saturated (constant) EUV luminosity  $L_{\text{sat}}$  at ages lower than  $t_{\text{sat}} \sim 100$  Myr. Taking  $\alpha = 5/4$ , the integral then yields for  $t > t_{\text{sat}}$  (and thus also  $t \gg t_{\text{form}}$ )

$$V_{\text{XUV}} = \varepsilon \pi R^2 \frac{L_{\text{sat}} t_{\text{sat}}}{4\pi a^2} \left( 5 - 4 \left( \frac{t_{\text{sat}}}{t} \right)^{1/4} \right). \quad (19)$$

It is convenient to write this in the following form

$$V_{\text{XUV}} = \frac{\varepsilon \pi R^2}{4\pi a^2} E_{\text{XUV}}, \quad (20)$$

with the energy of the star emitted in the XUV since the formation of the planet given as (see also Jackson et al. 2012)

$$E_{\text{XUV}} = \int_{t_{\text{form}}}^t L_{\text{XUV}} dt = L_{\text{sat}} t_{\text{sat}} \left( 5 - 4 \left( \frac{t_{\text{sat}}}{t} \right)^{1/4} \right). \quad (21)$$

At this point, we could in principle set the binding energy and the integrated energy flux equal and solve for  $M_c$ . But this would not lead to a conclusive result, as  $R$  is itself a function of  $M_c$  and  $M_e$ .

#### 3.2.1. Envelope and core mass

Therefore, we next need to establish a relation of the postformation envelope mass and the core mass. If the Kelvin-Helmholtz cooling timescale of the gaseous envelope regulates the envelope accretion rate during the nebular phase, and if  $\tau_{\text{KH}} \propto M_c^{q_{\text{KH}}}$  as suggested by various past studies (e.g., Ikoma et al. 2000), we expect that the final envelope masses will also follow a power law, with  $M_e \propto M_c^{-q_{\text{KH}}+1}$  (see Appendix A in Mordasini et al. 2014). Indeed, population synthesis simulations of the formation of planets where the gas accretion rate is found by explicitly solving the planetary structure equations find roughly speaking such a power law dependency for low-mass, subcritical cores (Sect. 2.1.3 and Mordasini et al. 2014). Thus, we write

$$M_e = M_{e,1} \left( \frac{M_c}{M_{\oplus}} \right)^{p_e}. \quad (22)$$

In this equation,  $M_{e,1}$  is the primordial (i.e., postformation) H/He envelope mass of a 1  $M_{\oplus}$  core. Formation models (e.g., Ikoma et al. 2000; Ikoma & Hori 2012) show that its value

depends on the planet's orbital distance, the disk properties, and on the (grain) opacity in the protoplanetary atmosphere (Ormel 2014; Mordasini 2014). The latter can be represented by choosing different values of  $M_{e,1}$ . As the reference value, we use  $M_{e,1} = M_{e,1,r} = 0.03 M_{\oplus}$ , corresponding to a three percent envelope mass.

The power law exponent  $p_e$  gives the increase of the envelope mass with increasing core mass, that can be related to the KH exponent, as seen above. From the numerical results of Sect. 2.1.3, we use as the nominal value  $p_e = 2$ , but we also study  $p_e = 0, 1$ , and 3.

The exponent  $p_e$  transports important informations on the formation process. From a formation point of view, it would be very interesting if we could obtain it from the evaporative imprints. As we suspect from the numerical results exhibiting a weak dependency on the postformation envelope mass, this is unfortunately not the case: as we see later, in the analytical model,  $M_{\text{bare}}$  is even completely independent of  $p_e$  (or the power law normalization constant). From an evolution point of view, this is in contrast positive, as it means that the initial conditions are not important.

### 3.2.2. Mass-radius relation

The final step before we can calculate  $M_{\text{bare}}$  is to specify the mass-radius relation. This relation is crucial to understand why  $R_{\text{bare}}$  is independent of  $p_e$ . We first treat the core radius and the envelope radius separately. For the core radius we use the result of Valencia et al. (2006) or Lopez & Fortney (2014) who found that in the Super-Earth mass domain, planets with an Earth-like rocky composition (i.e., 1:2 iron-silicate by mass) closely follow a power law with

$$\frac{R_{c,r}}{R_{\oplus}} = \left( \frac{M_c}{M_{\oplus}} \right)^{p_c}. \quad (23)$$

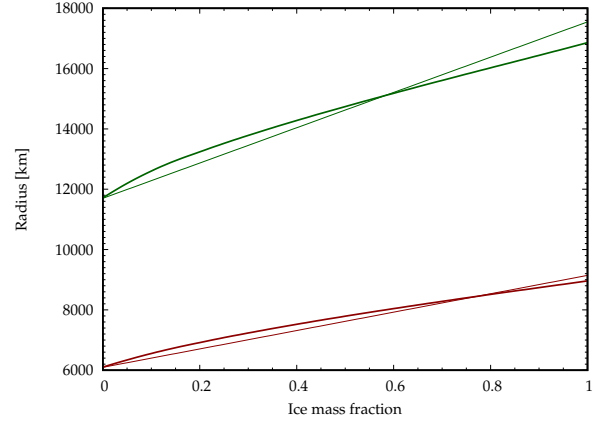
with  $p_c$  between 0.25 and 0.27.

Below we study how the location of the evaporation valley depends on the ice content of the core. We thus also need to specify how the presence of (water) ice increases the core radius. Considering the  $M$ – $R$  relation of rocky and icy low-mass planets (e.g., Mordasini et al. 2012a) as predicted from interior structure models suggest that to first order, we can write the following expression for the radius as a function the ice mass fraction  $f_{\text{ice}}$  (see also Zeng et al. 2019 for a similar expression)

$$R_c \approx R_{c,r} \left( 1 + \frac{1}{2} f_{\text{ice}} \right), \quad (24)$$

for masses between about 1 and  $10 M_{\oplus}$ . It is clear that this is not meant as an accurate expression for the impact of the ice mass fraction (see for example Fortney et al. 2007), but for our purpose it is sufficient. The increase of the core radius with increasing ice mass fraction as found from solving the interior structure equations is compared to this simple linear approximation in Fig. 11. The differences to the numerically obtained radii are  $\lesssim 4\%$  in this mass interval.

For the extent of the gaseous envelope we use the results of Lopez & Fortney (2014) who derived a fit to the thickness of the H/He envelope  $W_e$  as found in their evolutionary calculations. Their fit gives  $W_e$  as a function of core mass, envelope mass, semimajor axis, and time, where the latter two dependency are weaker. As mentioned above, for the analytical model we do not take into account explicitly the temporal dependency of the



**Fig. 11.** Radius of the solid core as a function of ice mass fraction (the rest has an Earth-like iron and silicate composition). The top and bottom pairs of lines are for a mass of 10 and  $1 M_{\oplus}$ , respectively. The thick lines show the radius as found from solving the interior structure equations, while the thin lines show the simple approximation from Eq. (24) with the linear dependency.

radius, but rather evaluate their expression at  $t = 100$  Myr as the representative timescale on which most of the evaporation occurs (e.g., Fig. 4 in Jin et al. 2014). Furthermore, we also neglect the weak dependency on the semimajor axis, but take again a representative distance, namely 0.1 AU. In this case, the envelope thickness  $W_e$  as a function of core and envelope mass given by Lopez & Fortney (2014) is

$$\frac{W_e}{R_{\oplus}} = 4.2 \left( \frac{M_c + M_e}{M_{\oplus}} \right)^{-0.21} \left( \frac{M_e / (M_c + M_e)}{0.05} \right)^{0.59}. \quad (25)$$

We see that the radius decreases with increasing total mass (approximately equal to  $M_c$  for the planets studied here), but increases with the envelope mass fraction. The total radius is then  $R = R_c + W_e$ , neglecting the usually small contribution of the radiative atmosphere (Lopez & Fortney 2014).

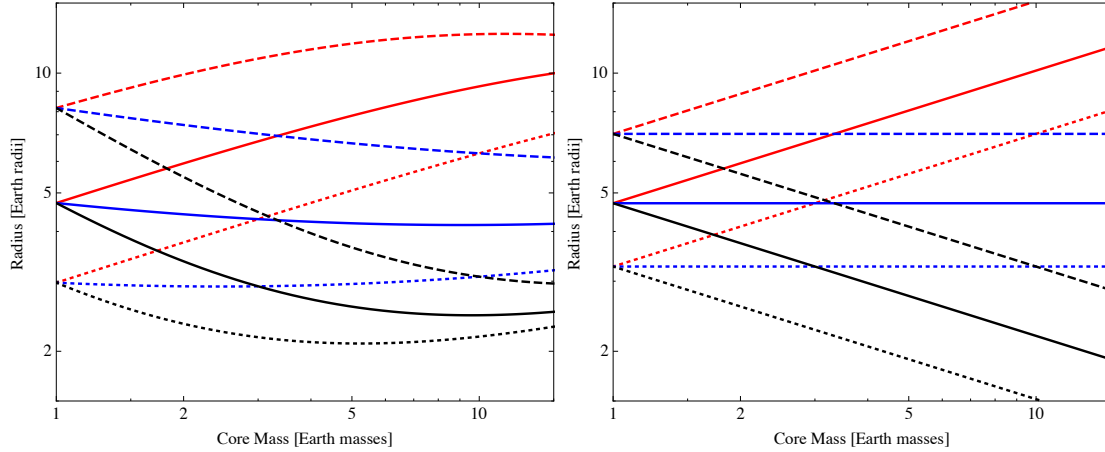
In Fig. 12, left panel, we plot the total radius  $R$  as a function of core mass. The envelope mass is given according to Eq. (22) with  $M_{e,1} = 0.01$  (less efficient gas accretion for example because of higher opacities), 0.03 ( $=M_{e,1,r}$ , the nominal value), and 0.1 (efficient H/He accretion) and for power law exponents of the increase of the envelope mass with core mass of  $p_e = 0$  (envelope mass independent of core mass), 1 (linear increase), and 2 (nominal case).

Despite being analytical, the expression for the total radius as the sum of core radius plus envelope thickness (Eqs. (23) and (25)), is unhandy for the present application. Fortunately, the plot suggest that for a large part of the parameter space, one can write the total radius with a simpler functional form that contains in particular no more an explicit dependency on  $M_e$ , but expresses it with  $M_c$  and  $p_e$ . Examining various expression eventually leads to the following form, shown in the right panel of Fig. 12:

$$R \approx R_{1,r} \left( \frac{M_{e,1}}{M_{e,1,r}} \right)^{\frac{1}{3}} \left( \frac{M_c}{M_{\oplus}} \right)^{\frac{p_e-1}{3}}. \quad (26)$$

In this expression, the normalization constant  $R_{1,r} \approx 4.6 R_{\oplus}$  is the total radius of a young  $1 M_{\oplus}$  planet at 0.1 AU with an envelope mass of  $M_{e,1,r}$ . Such a planet loses its envelope on a timescale of less than 100 Myr when including evaporation. This is however not a problem, as the specific mass of  $1 M_{\oplus}$  is just the arbitrary





**Fig. 12.** Radius as a function of core and envelope mass for planets at 100 Myr and 0.1 AU. The *left panel* shows the radius as found from the full expression (Eqs. (23) and (25)) while the *right one* uses the simple approximation of Eq. (26). The dotted, solid, and dashed lines show the radius for envelope masses of a  $1 M_{\oplus}$  core of  $M_{e,1} = 0.01, 0.03,$  and  $0.1 M_{\oplus}$ . Black, blue, and red lines show  $p_e = 0, 1,$  and  $2,$  the power law exponent representing how H/He envelope mass  $M_e$  scales with core mass,  $M_e \propto M_{c,1} M_c^{p_e}$ .

normalization mass. Therefore, Eq. (26) yields the relationship of core mass, envelope mass, and total radius for young low-mass planets as long as they still have an envelope, which is the quantity we need. The validity of the expression was checked by comparing it to the radii in the numerical simulations. A typical agreement on a  $\sim 20\%$  level (or better) was found.

The second term in Eq. (26) shows how the radius depends on the envelope mass at a fixed mass of  $1 M_{\oplus}$ . The third term finally shows the dependency on the core mass. From the left of Fig. 12, we see that we can write the dependency approximately also as a power law. Crucially, the exponent is itself a function of  $p_e$ . One finds a dependency as  $(p_e - 1)/3$ . This exponent shows an interesting behavior: if the envelope mass is independent of core mass ( $p_e = 0$ ), the total radius decreases with increasing core mass as approximately  $M_c^{-1/3}$ . The increasing core gravity compresses the envelope, and no additional gas must be accommodated with increasing core masses. For a linear increase of the envelope mass with core mass ( $p_e = 1$ ), the total radius is independent of core mass (see the approximately horizontal middle lines in the left panel of Fig. 12). Here the effect of the increasing gravity of the core leading to a smaller radius and the increase of the envelope mass and thus the necessary volume cancel each other. Finally, for a quadratically increasing envelope, the increase of the volume because of more gas wins over the increasing gravity of the core; the radius increases as  $M_c^{1/3}$ .

We comment that it appears possible that the (approximate) core mass-envelope mass-radius relation of Eq. (26) could be derived also analytically from (a simplified) solution of the structure equations. We have not tried to do this here. If such an analytical solution exists, it would make the model presented in this paper fully analytical.

### 3.2.3. Limits of the approximation

The approximation of Eq. (26) works well for core-envelope mass ratios suggested by the formation simulations (Sect. 2.1.3), which are a few percent of envelope in mass for cores of a few Earth masses, increasing with core mass to  $\sim 10\%$  in mass for cores approaching  $10 M_{\oplus}$ . As visible from comparing the left and right panel of Fig. 12, our approximation breaks down in the limit both of very low and high core-envelope mass ratios:

In the former case, this is due to the fact that nothing prevents the approximation for the total radius from becoming smaller

than the core radius, which is of course physically impossible. According to Eq. (23), the core radius of a  $10 M_{\oplus}$  planet of terrestrial composition is for example about  $1.9 R_{\oplus}$ , more than the value given by the black dotted line in the right panel of the figure. The approximation is thus not usable for cases where the core radius is not clearly smaller than the total radius, as the approximation would predict too small radii.

On the other end of the spectrum of envelope masses, it is also not applicable once the self-gravity of the (massive) envelope starts to be significant, compressing the planet. Here the approximation predicts too large radii. These two regimes manifested themselves in the largest differences between the full expression (left panel) and the approximation (right panel) occurring in the bottom and top right corners.

## 3.3. Locus of the valley of evaporation in the distance-mass and distance-radius plane

### 3.3.1. Maximum bare mass $M_{\text{bare}}$

With the expression for  $M_e(M_c)$  and  $R(M_c, M_e)$  we can finally set  $-U_e = V_{\text{XUV}}$  and solve for the largest core mass that loses its entire envelope. We follow here [Lecavelier Des Etangs \(2007\)](#) in neglecting the small contributions of the kinetic and thermal energy in the energy budget. We then have

$$\frac{k_{\text{pot}} G M_c M_e}{R} = \frac{\varepsilon \pi R^2}{4 \pi a^2} E_{\text{XUV}}. \quad (27)$$

Rearranging and inserting the equations for  $M_e$  (Eq. (22)) and  $R$  (Eq. (26)) yields

$$k_{\text{pot}} G M_c M_{e,1} \left( \frac{M_c}{M_{\oplus}} \right)^{p_e} = \frac{\varepsilon E_{\text{XUV}} R_{1,r}^3}{4 a^2} \left( \frac{M_{e,1,r}}{M_{e,1,r}} \right) \left( \frac{M_c}{M_{\oplus}} \right)^{p_e - 1}. \quad (28)$$

By solving for the  $M_c$ , which now becomes the maximal core mass that evaporates to a bare core  $M_{\text{bare}}$ , we obtain the final result

$$M_{\text{bare}} = \frac{1}{2a} \left( \frac{\varepsilon E_{\text{XUV}} R_{1,r}^3}{k_{\text{pot}} G} \right)^{\frac{1}{2}} \left( \frac{M_{e,1,r}}{M_{\oplus}} \right)^{-\frac{1}{2}}. \quad (29)$$

Together with the associated radius (Eq. (32)), this is the main result of the analytical part of this paper.

It is important to note that in the equation,  $R_{1,r}$  and  $M_{e,1,r}$  are constants from the mass-radius relation equivalent to the statement that a  $1 M_{\oplus}$  planet has a  $1 R_{\oplus}$  radius for an Earth-like composition, and not variables. There is in particular no dependency on  $p_c$  and  $M_{e,1}$  (i.e., the initial envelope mass) as the term on the RHS and LHS in Eq. (28) have cancelled. This independency is in excellent agreement with the numerical experiments presented above in Sect. 2.3 and explains why independently of the primordial H/He content, an always similar evaporation valley is found. It appears likely that it is actually because of this independency on the postformation envelope mass that the valley is so clearly appearing observationally: it means that the (likely) significant spread in the postformation envelope masses resulting from the formation epoch does not matter.

The reason behind this cancellation becomes clear with the analytical model: a more massive envelope (high  $p_c$  or  $M_{e,1,r}$ ) means more mass needs to be evaporated (on the LHS in Eq. (28)), but it also means that the radius is larger (at nearly fixed total ( $\approx$ core) mass) and thus the evaporation stronger, which appears on the RHS. Because of the particular functional form of the mass-radius relation of low-mass planets with H/He, the two cancel. We also see that  $M_{\text{bare}}$  scales as  $1/a$ , again as in the numerical simulations. From *Kepler's* law it follows that in terms of orbital period  $P$ ,  $M_{\text{bare}} \propto P^{-2/3}$ .

It is important to note that there is only a square root dependency on  $\varepsilon$  and  $E_{\text{XUV}}$ . This explains now analytically why, first, on the theoretical side various models of evaporation (which can be seen as different effective  $\varepsilon$ ) found similar locations of the valley, or in other words, why the details of these models do not matter so much. Second, on the observational side, the rather weak dependency on  $E_{\text{XUV}}$  explains why we see a clear valley at all: the large observed spread in the  $L_{\text{XUV}}$  of young stars does not so strongly alter the location of the valley around an individual star. This means that the valley is not blurred away when we consider the entire population of stars and planets. As we see in the following section, the dependency is even much weaker for  $R_{\text{bare}}$  than for  $M_{\text{bare}}$ .

Inserting the parameters  $M_{e,1,r} = 0.03 M_{\oplus}$  and  $R_{1,r} = 4.6 R_{\oplus}$ , setting  $k_{\text{pot}} = 1$ , and evaluating at  $4.5 \text{ Gyr} \gg t_{\text{sat}}$ , we get a value of

$$M_{\text{bare,late,rocky}} \approx 6.5 \left( \frac{a}{0.1 \text{ AU}} \right)^{-1} \left( \frac{\varepsilon}{0.1} \right)^{\frac{1}{2}} \times \left( \frac{L_{\text{sat}}}{10^{30} \text{ erg s}^{-1}} \right)^{\frac{1}{2}} \left( \frac{t_{\text{sat}}}{100 \text{ Myr}} \right)^{\frac{1}{2}} M_{\oplus}. \quad (30)$$

This is also quantitatively in quite good agreement with the numerical results (Table 1). The numerical value of  $6.5 M_{\oplus}$  is for an Earth-like rocky core composition. For ice-rich compositions, if we assume that the increase in the core radius (Eq. (24)) translates into the same increase of  $R_{1,r}$ , then an ice mass fraction of for example  $1/2$  should lead to an increase of  $M_{\text{bare,late}}$  by a factor  $5/4^{3/2} \approx 1.4$  larger, i.e.,  $9.1 M_{\oplus}$ .

### 3.3.2. Maximum bare radius $R_{\text{bare}}$

With the relations of Eq. (23) and (24) we convert  $M_{\text{bare}}$  trivially into the radius of the largest planet that becomes a bare core,  $R_{\text{bare}}$ . One finds

$$R_{\text{bare}} = \left( 1 + \frac{1}{2} f_{\text{ice}} \right) \left( \frac{1}{2a} \right)^{p_c} \left( \frac{\varepsilon E_{\text{XUV}} R_{1,r}^3}{k_{\text{pot}} G} \right)^{\frac{p_c}{2}} \times \left( \frac{M_{e,1,r}}{M_{\oplus}} \right)^{-\frac{p_c}{2}} \left( \frac{1}{M_{\oplus}} \right)^{p_c} R_{\oplus}. \quad (31)$$

The equation makes clear that the time-integrated X-ray and extreme ultraviolet luminosity of a star is key in determining the impact of atmospheric photoevaporation. This picture is supported by the finding of McDonald et al. (2019) that the (evaporation) desert exhibits much greater variability in the desert onset in the bolometric flux space compared to the integrated X-ray flux space.

Inserting the values of  $M_{e,1,r}$  and  $R_{1,r}$ , setting  $k_{\text{pot}} = 1$  and  $p_c = 0.27$  and evaluating  $E_{\text{XUV}}$  again at  $4.5 \text{ Gyr}$  yields the main result of the paper in terms of the location of the bottom of the evaporation valley as a function of the determining quantities

$$R_{\text{bare,late}} \approx 1.6 \left( 1 + \frac{1}{2} f_{\text{ice}} \right) \left( \frac{a}{0.1 \text{ AU}} \right)^{-0.27} \left( \frac{\varepsilon}{0.1} \right)^{0.135} \times \left( \frac{L_{\text{sat}}}{10^{30} \text{ erg s}^{-1}} \right)^{0.135} \left( \frac{t_{\text{sat}}}{100 \text{ Myr}} \right)^{0.135} R_{\oplus}. \quad (32)$$

As expected, in Table 1 we again find good agreement with the numerically obtained results, in particular for the dependency on the orbital distance  $\propto a^{-p_c} = a^{-0.27}$  (compare with  $e_r$  in Table 1). With *Kepler's* law, this corresponds in terms of orbital period  $P$  to  $R_{\text{bare}} \propto P^{-2p_c/3} \approx P^{-0.18}$ , where the normalization distance of  $0.1 \text{ AU}$  corresponds for a  $1 M_{\odot}$  star to  $P = 11.6 \text{ d}$ .

It is certainly remarkable that two fundamental properties of solid planets are imprinted into the valley's locus: first their mass-radius scaling in terms of  $d \log R / d \log M = p_c$  (Eq. (23)) which determines the valley's distance dependency. Second, the ice mass fraction (or more generally, the bulk composition), which determines the absolute position at a fixed distance. These dependencies can be compared with observational constraints (Sect. 4, see also Gupta & Schlichting 2019).

In this expression it becomes apparent that the transition radius has a very weak dependency on the strength of evaporation encapsulated in  $\varepsilon$ ,  $L_{\text{sat}}$ , and  $t_{\text{sat}}$ , even more so than the transition mass  $M_{\text{bare}}$ , which has important implications, as discussed in the last section and in Sect. 6.

The strongest dependency is the linear dependency on the ice mass fraction  $f_{\text{ice}}$  (or the composition of the core in a more general sense). This is the reason why the valley of evaporation is such a good diagnostic of the core composition. For an ice mass fraction of  $0.75$  as used in Jin & Mordasini (2018), Eq. (32) predicts that the transition radius at  $0.1 \text{ AU}$  increases from  $1.6$  to  $2.2 R_{\oplus}$ . This estimate neglects that  $M_{\text{bare}}$  is itself already larger as icy cores are more vulnerable to loss already when the envelope is still present, and not only that they are bigger once the envelope is lost. The numerically found value is thus somewhat larger ( $2.3 R_{\oplus}$  for  $75\%$  ice). Above, in Sect. 2.3.9 for  $100\%$  ice, numerically a  $R_{\text{bare}} = 2.7 R_{\oplus}$  at  $0.1 \text{ AU}$  was found (simulation I1), while Eq. (32) predicts about  $2.4 R_{\oplus}$ .

### 3.3.3. Incorporating a distance dependent efficiency factor

Up to now, we have assumed that the efficiency factor in the energy-limited escape formula is a constant. In reality, it depends on several factors (e.g., Kubyskhina et al. 2018b; Wu 2019), like the composition and temperature in the atmosphere which in turn depends on stellar irradiation and thus the orbital distance. Assuming that the latter dependency is the most important (and the one we can most easily study, given that *Kepler* planets forming the evaporation valley are found at a range of orbital distances), we may approximate the efficiency factor at least locally with a power law in orbital distance as

$$\varepsilon = \varepsilon_0 \left( \frac{a}{a_0} \right)^q, \quad (33)$$

where  $\varepsilon_0$  is the value of the efficiency factor at some normalization distance  $a_0$  which we set to 0.1 AU.

Repeating the analysis as before for the constant  $\varepsilon$ , we now find a transition mass at late times of

$$M_{\text{bare,late,rocky}} \approx 6.5 \left( \frac{a}{0.1 \text{ AU}} \right)^{-1+q/2} \left( \frac{\varepsilon_0}{0.1} \right)^{\frac{1}{2}} \times \left( \frac{L_{\text{sat}}}{10^{30} \text{ erg s}^{-1}} \right)^{\frac{1}{2}} \left( \frac{t_{\text{sat}}}{100 \text{ Myr}} \right)^{\frac{1}{2}} M_{\oplus}. \quad (34)$$

The corresponding transition radius becomes

$$R_{\text{bare,late}} \approx 1.6 \left( 1 + \frac{1}{2} f_{\text{ice}} \right) \left( \frac{a}{0.1 \text{ AU}} \right)^{-0.27(1-q/2)} \left( \frac{\varepsilon_0}{0.1} \right)^{0.135} \times \left( \frac{L_{\text{sat}}}{10^{30} \text{ erg s}^{-1}} \right)^{0.135} \left( \frac{t_{\text{sat}}}{100 \text{ Myr}} \right)^{0.135} R_{\oplus}, \quad (35)$$

i.e.,  $R_{\text{bare,late}} \propto a^{-p_c(1-q/2)}$ . The idea behind this approach is the following: given the recent observational determination of the power law slope of the valley (Van Eylen et al. 2018) as a function of orbital period which corresponds in the model to  $-(2/3)p_c(1-q/2)$ , we can determine from the observations  $q$  (assuming that  $p_c \approx 0.27$  is known), that is, how the efficiency factor varies a function of orbital distance (or in a more general sense, depends on planet properties like the escape velocity, see Wu 2019). This is an important constraint for evaporation models.

The slope of the middle of the valley varies as  $a^{-0.15 \pm 0.05}$  as found by Van Eylen et al. (2018), which translates to  $q = 0.89 \pm 0.26$ . This means that taken at face value, we would find that the efficiency factor varies roughly speaking a bit weaker than linearly with orbital distance, which is caused by the more horizontal valley in the observations compared to the predictions for a constant  $\varepsilon$ . At small orbital distances, such a lower (effective)  $\varepsilon$  could be caused by stronger radiative cooling, corresponding to a mass loss rate which is (partially) radiation-recombination-limited instead of energy-limited (Murray-Clay et al. 2009). As shown by Owen & Wu (2017), simple energy-limited evaporation models with a fixed  $\varepsilon$  indeed predict a steeper slope of the valley compared to more realistic models which directly calculate the evaporation rate from first principles (Owen & Jackson 2012).

At large distance, this inferred dependency formally predict very high values for  $\varepsilon$ . But here we have to take into account that at distances much outside of about 0.2 AU (Owen & Jackson 2012), the mass loss rate may in any case be much different (smaller) than assumed here: outside of such a distance, the evaporation should occur in the Jeans' regime, and no more in the hydrodynamic regime as assumed in the model, which would strongly reduce the escape rate. This should be tested in future with more detailed evaporation models (e.g., Kubyskhina et al. 2018b,a).

### 3.3.4. Comparison with the numerical simulations

Comparison of the locus of the valley as predicted by the analytical model with the numerically found values in Table 1 shows that the analytical model can very well explain the main results found with the simulations. The analytical model explains in particular:

- The dependency of the transition mass from solid to planets with H/He  $M_{\text{bare}}$  with orbital distance  $a$  as  $a^{-1}$  or equivalently  $R_{\text{bare}}$  with  $a^{-0.27}$ . This was found numerically to good approximation in all simulations except M1.

- The (at first) surprising independency or weak dependency of  $R_{\text{bare}}$  on the scaling of the envelope mass with core mass, that is, the postformation envelope mass. This was seen numerically with the simulations M0–M4, N1 and N2.

- The weak dependency on the strength of atmospheric escape encapsulated in the efficiency factor  $\varepsilon$  and  $L_{\text{XUV}}$ . This was investigated with the simulations E1 and E2.

- The comparatively strong dependency on the ice mass fraction or, more generally speaking, the composition of the solid core. Numerically, this was found with simulation I1.

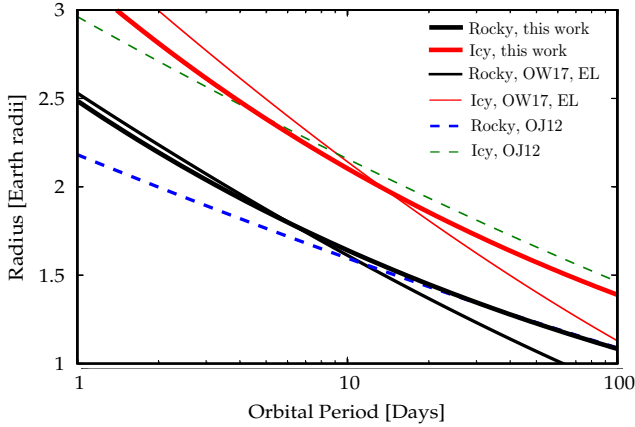
The largest discrepancy regarding the radial slope of the valley occurs for simulation M1, where a clearly steeper slope was found numerically with  $R_{\text{bare}} \propto a^{-0.36}$ , whereas the analytical model always predicts  $R_{\text{bare}} \propto a^{-0.27}$ . In this simulation,  $p_c = 0$  which means that the envelope mass is equal to  $0.03 M_{\oplus}$  independent of core mass. This leads to much lower core-envelope mass ratios for massive cores  $\geq 5-10 M_{\oplus}$  than in the other simulations. They are also much lower than the ones suggested by the formation model. The discrepancy can be explained when we recall that the analytical model uses an approximation for  $R(M_c, M_e)$  (Eq. (26)) which breaks down in the limit of both very low and high core-envelope mass ratios as discussed in Sect. 3.2.2. It is thus not surprising that the analytical model does not well recover the numerical results in M1: for massive cores with very low mass envelopes,  $R(M_c, M_e)$  predicted by Eq. (26) can become smaller than the core radius, which is obviously impossible.

Interestingly, however, it is possible to construct a more complex analytical model which uses the full expression for  $R(M_c, M_e)$  (Eqs. (24) and (25)) instead of the approximation of Eq. (26). For this model, it is no more possible to analytically solve for  $M_{\text{bare}}$  (and  $R_{\text{bare}}$ ), but instead one has to find the root of a more complex version of Eq. (28) numerically. But with this more complex analytical model, one finds excellent agreement with the locus of the valley also for simulation M1. This is a strong indication that the principle underlying the analytical model, namely that envelopes are lost when the temporal integral of the stellar XUV irradiation absorbed by the planet is larger than binding energy of the envelope in the potential of the core, indeed captures the governing physics.

### 3.4. Comparison to Owen & Wu (2017)

The analytical model presented here differs in a number of aspects from the one presented by Owen & Wu (2017). First, their model is fully analytical, while we use the fits to the numerically obtained envelope mass-core mass-radius relation from Lopez & Fortney (2013). Second, we only consider the regime of more massive envelopes (high  $X$  case in the terminology of Owen & Wu 2017), which is inspired by the postformation envelope masses found in our core accretion formation simulations (Sect. 2.1.3). Our mass-radius relation is not applicable for the case of massive cores with tiny envelopes: in this case, Eq. (26) can predict total radii that are smaller than the core radius, which is of course not possible. Third, the Owen & Wu (2017) model is first timescale-based, and uses binding vs. irradiated XUV energies in the final step only. The approach here is directly based on comparing these energies. Finally, the full model of Owen & Wu (2017) uses variable evaporation efficiencies (Owen & Jackson 2012) while we normally consider a constant efficiency factor, as Owen & Wu (2017) also do in their simplified model. The case of an efficiency factor changing with orbital distance is addressed in Sect. 3.3.3. In summary, the Owen & Wu (2017) model provides a more general and fully analytical view of





**Fig. 13.** Comparison of the radius  $R_{\text{bare}}$  which is the radius the largest stripped core as a function of orbital period as predicted by the analytical model of this paper (thick lines) and the model of Owen & Wu (2017). For the latter, the results for both the simpler energy-limited model with a constant efficiency factor and for the model of variable efficiency (based on Owen & Jackson 2012) are shown. The upper and lower three lines are for an Earth-like and ice-rich (1/3 in mass) core composition, respectively.

the mechanisms controlling envelope loss. The model presented here on the other hand makes the physical mechanism determining the boundary between Super-Earth and Sub-Neptunes very directly comprehensible, and illustrates well the dependencies of important parameters like the initial envelope mass or the stellar XUV-luminosity.

It is interesting to quantify the differences introduced by these point for the main result, which is the locus of the valley. Owen & Wu (2017) find a dependency of the radius of the largest stripped core  $R_{\text{bare}}$  as a function of period  $P$  of  $R_{\text{bare}} \propto P^{-0.25}$  for the simple energy-limited model with a constant efficiency factor, and of  $\propto P^{-0.16}$  in a model with a variable efficiency according to the Owen & Jackson (2012) evaporation models. The analytical model derived here leads for a constant  $\varepsilon$  to a similar dependency with  $R_{\text{bare}} \propto P^{-0.18}$ . This is between the two exponents found by Owen & Wu (2017).

Figure 13 compares the radius of the largest stripped planet  $R_{\text{bare}}$  (i.e., the bottom of the valley) as a function of orbital period  $P$  as found in Owen & Wu (2017) and analytically in this work. Two compositions of the core are shown: an Earth-like composition with a 2:1 silicate:iron mass ratio, and a planet consisting of silicates and 1/3 ice. Thin lines show the results of Owen & Wu (2017). Thick lines show the analytical results of the present paper using Eq. (32) with the nominal parameters, i.e.,  $R_{\text{bare}}/R_{\oplus} \propto 1.6p^{-0.18}$  for the Earth-like composition, and  $\propto 2.05p^{-0.18}$  for the case with ice. The latter normalization radius of  $2.05 R_{\oplus}$  is obtained with the aforementioned equation with  $f_{\text{ice}} = 1/3$ , and taking into account that the assumption of a pure silicate composition (without iron – in contrast to the assumption underlying our Eq. (23)) leads to a further increase of the radius by about 10% (Fortney et al. 2007).

Figure 13 shows that despite the aforementioned differences, the two models yield comparable results, regarding both the distance dependency of the valley and the impact of varying the core composition. We see that the differences between the analytical model presented here and the Owen & Wu (2017) results are comparable to the differences between the two approaches discussed by Owen & Wu (2017), namely the simple energy-limited approach with a constant efficiency factor, and the more realistic model with a variable efficiency (which is based on

Owen & Jackson 2012). This leads to a shallower slope Owen & Wu (2017); Wu (2019).

One also sees that at a period of around 10 days, which approximately corresponds to an orbital distance of about 0.1 AU around a solar-like stars, the results of the various models are very similar. This means that our results for the valley locus at the place where we normalize them (period of 10 days, resp. 0.1 AU), should not be strongly affected by our assumption of a constant efficiency factor. The dependency on orbital distance will in contrast be more affected.

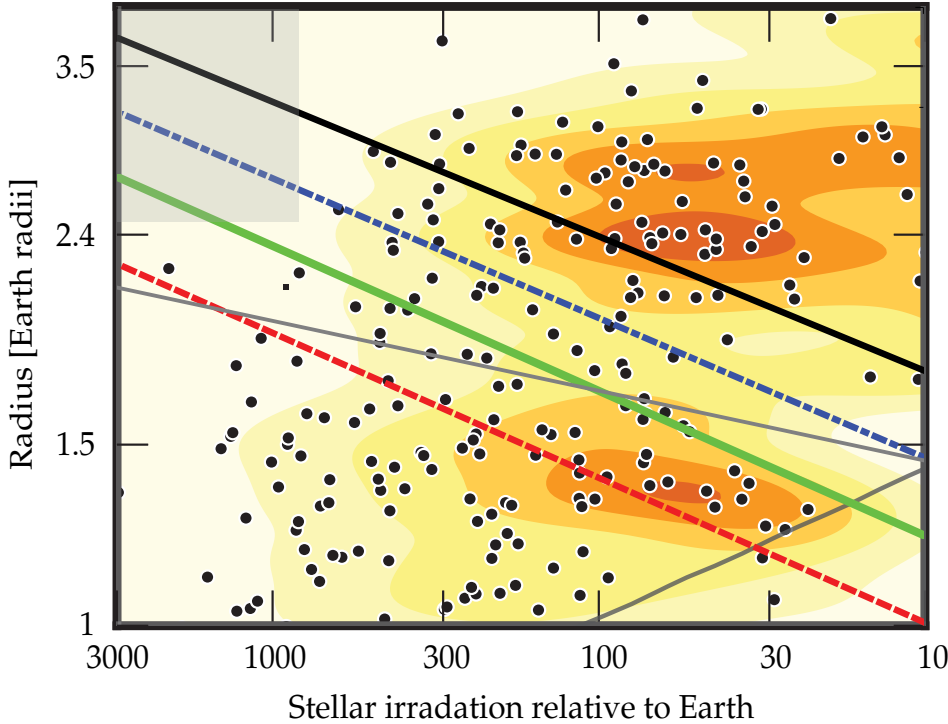
## 4. Comparison with observations

### 4.1. Comparison with Kepler data

Figure 14 shows the location of the valley as observed for *Kepler* planets around solar-like stars ( $0.96 \leq M/M_{\odot} \leq 1.11$ ) from Fulton & Petigura (2018) with points and yellow-brown completeness-corrected occurrence rate maps (dark colors indicate a high occurrence). This data is obtained with new *Gaia* parallaxes, *Kepler* photometry, and spectroscopically determined stellar temperatures from the California-*Kepler* Survey, resulting in a typical precision in the planetary radii of 5%. A similar study of Martinez et al. (2019) with a median internal uncertainty of 3.7% in the planetary radii finds results that are in overall agreement with Fulton & Petigura (2018).

The lines show the valley's location predicted by the simulations (using  $R_{\text{b,0p1}}$  and  $e_r$  from Table 1). The theoretical lines thus show the bottom of the valley, not its middle. The nominal simulation M0 shown by the green solid lines agrees quite well with the observed bottom of the valley, as noted previously (Owen & Wu 2017; Jin & Mordasini 2018). The red dashed and blue dashed-dotted lines show the simulations E1 and E2 (strong and weak evaporation). They give an impression of how much the valley's location shifts if the evaporation rate is varied by a factor 10 and 0.1 relative to the nominal case, for example because of the observed spread in stellar  $L_{\text{XUV}}$ . As discussed in Sect. 2.3.6, the valley is found to shift by a similar extent as its intrinsic width (about  $0.5 R_{\oplus}$ ), which explain why the valley is visible, but still not completely empty.

The black line is simulation I1 (ice cores) which is ruled out by the observations, again as already noted in Owen & Wu (2017); Jin & Mordasini (2018). The thin gray line finally assumes the linearly distance-dependent evaporation efficiency factor derived in Sect. 3.3.3 based on the slope of the valley determined observationally with astroseismology by Van Eylen et al. (2018). Comparing (by eye) with the radial dependency of the observed super-Earth over-density, it seems to provide a better reproduction of the observed slope also in this data set, but a quantitative statistical assessment using the exact occurrence maps would be required to really address this. Observations of the frequency of planets in the black triangle which is currently observationally unconstrained would obviously also be helpful for a better determination of the slope. Regarding the morphology of the valley at such larger distances, we recall that our theoretical model does not include the effect of the transition from hydrodynamic escape to the much slower Jeans escape. This could effectively put an end to the valley at around a stellar irradiation relative to Earth  $F_{\oplus}$  of roughly 25 (Owen & Jackson 2012) because the planets could keep the H/He. This of course only holds if also at such larger distances, most planets grow to their final mass during the presence of the gas disk. This becomes however increasingly unlikely, as the formation timescale via collisional growth increases with orbital



**Fig. 14.** Comparison of the locus of the valley as observed for *Kepler* planets around solar-like stars (points and yellow-brown occurrence maps, [Fulton & Petigura 2018](#), with permission) and as theoretically found in the simulations in Table 1 (lines). It is important to note that all the theoretical lines show the bottom of the valley, not its middle. The green solid line is the nominal simulation M0. The red dashed and blue dashed-dotted lines are E1 and E2 (strong and weak evaporation). The black line is I1 (ice cores). The thin gray line assumes a distance-dependent evaporation efficiency factor. The gray rectangle in the top left corner is the empty evaporation desert as found in the nominal simulation M0. The black triangle is observationally unconstrained.

distance (e.g., [Mordasini et al. 2009](#)), as exemplified by Earth’s last giant (moon-forming) impact at an age clearly higher than protoplanetary disk lifetimes ([Jacobson et al. 2014](#); [Haisch et al. 2001](#)). As discussed in detail in [Owen & Murray-Clay \(2018\)](#), it could therefore be that at such distance, terrestrial planets could become dominant that formed after disk dispersal, with a limiting radius that increases with distance ([Lopez & Rice 2018](#)). Simulations of combined planetary growth and self-consistently coupled evolution which includes the accretion of solids and of H/He, and the loss thereof via atmospheric escape and impacts along the lines of the population syntheses by [Mordasini \(2018\)](#) should help to clarify this picture.

Coming back to Fig. 14, the gray squared region in the top left corner is the empty evaporation desert as found in the nominal simulation M0. It is another characteristic imprint of atmospheric escape (e.g., [Kurokawa & Nakamoto 2014](#)). While there are smaller planets observed at such very close-in orbits, it is empty also in the observational data (cf. [Lundkvist et al. 2016](#); [Mazeh et al. 2016](#)). It is certainly a strong point of the evaporation hypothesis for the valley that evaporation naturally explains two a priori unrelated observed features, the valley and the desert.

#### 4.2. *Kepler-36*

Planets in multiple systems with dissimilar densities like the planets around *Kepler-36* ([Carter et al. 2012](#)), K2-106 ([Guenther et al. 2017](#)) or HD3167 ([Gandolfi et al. 2017](#)) are of particular interest as these planets have evolved in the same stellar XUV environment, allowing one to study the differential history of the planets under the action of evaporation with one variable less ([Owen & Estrada 2019](#)). A prime example in this context are the two planets around *Kepler-36*. They are characterized ([Carter et al. 2012](#)) by an extreme density difference (about 6.8 and 0.8 g/cm<sup>3</sup> for planets b and c, respectively) despite the very similar orbital distances (0.115 and 0.128 AU). But planet c is with 7.7  $M_{\oplus}$  clearly more massive than b (4.3  $M_{\oplus}$ ).

As shown in [Lopez & Fortney \(2013\)](#) (see also [Owen & Morton 2016](#)), this mass difference means that planet b with its weaker gravity cannot keep its H/He envelope, while c can, meaning that models of atmospheric escape can successfully explain this surprising system. This is also the case with the calculations presented here: The analytical model predicts a transition mass  $M_{\text{bare}} = 5.65 M_{\oplus}$  at b’s position, and  $M_{\text{bare}} = 5.08 M_{\oplus}$  at c’s position (Eq. (30)), which is between the masses of b and c. So planet b with a mass inferior to the local  $M_{\text{bare}}$  should lose its envelope, whereas c should keep it – exactly as observed. In this sense, the simple analytical model passes the *Kepler-36* acid test very well, without any model tuning. *Kepler-36* is thus a very interesting system that is at the border of the triangle of evaporation, and it should serve as a test case for any model trying to explain the observed properties of close-in low mass planets (see for example [Liu et al. 2015](#) for a scenario of devolatilized by giant impacts).

Comparing with the numerically obtained results (Table 1), we see that the minimum core masses  $M_{\text{bare}}$  for the planets to retain their envelopes is also here between 5 and 6  $M_{\oplus}$  (i.e., between the masses of b and c) for most simulations, including the nominal one, again in agreement with observations. The same result was found already also by [Owen & Wu \(2013\)](#); [Owen & Morton \(2016\)](#).

In the context of these numerical models it is possible to go a bit further with the analysis of *Kepler-36*, as simulations M1, E1, E2, and I1 predict – taken at face value – in contrast an incompatible  $M_{\text{bare}}$ . From this, we can infer the following.

First, from the incompatibility with simulation M1 (low envelope masses) we can deduce that the planet c initially had an envelope more massive than 0.03  $M_{\oplus}$  (corresponding to an envelope mass fraction of about 0.4%). This is consistent with the detailed analysis of [Lopez & Fortney \(2013\)](#) who demonstrated that both planets could have been born with H/He envelopes of approximately 22% of the planets’ initial mass. Similarly, [Owen & Morton \(2016\)](#) also find a higher initial envelope mass fraction of about 15–30% to explain the observations. For comparison,

our formation simulations predict via Eq. (11) an initial envelope mass fraction of about 7%, also much more than the 0.4% that is ruled out from simulation M1.

Second, from the incompatibility of the simulations E1 and E2 we can deduce that evaporation in the Kepler-36 system was not as strong as in E2 (where also planet c would lose its H/He), but also not as weak as in E1 (where also planet b would keep it), but within one order of magnitude of the nominal evaporation model employed here.

Third and finally, the simulation I1 (icy cores) also produces an incompatible  $M_{\text{bare}}$ . While it is not possible to constrain well the ice mass fraction of planet c because of the degeneracy with the H/He envelope mass fraction, the fact that the high density of planet b is consistent with a rocky Earth-like composition, with about 30% of its mass in iron (Carter et al. 2012), indicates that at least this planet indeed does not have an incompatible, ice-dominated composition.

Clearly, this simple analysis assumes that the other quantities and settings influencing the evaporative transition were at similar value as assumed in model. In reality degeneracies are possible, and one could construct models with modified evaporation rates and initial envelope masses that also lead to a  $M_{\text{bare}}$  consistent with observations. But the analysis nevertheless shows that one can derive interesting (joint) constraints like the postformation envelope mass or the strength of evaporation from such multiplanetary systems, making them very interesting (Owen & Estrada 2019). In the ideal case, studies investigate them should not only be limited to the interior structure aspect, but also take into account the global system architecture, like the dynamical state in terms of eccentricities and mean motion resonances (Poon et al. 2020). These additional aspects may differ depending on the proposed mechanism proposed to explain a specific system.

## 5. Summary

Precise observations of the radii of close-in extrasolar planets have revealed a depleted region at about 1.7 (1.5–2.0)  $R_{\oplus}$  for solar-like stars separating smaller super-Earths from larger sub-Neptunes (Fulton et al. 2017; Van Eylen et al. 2018; Fulton & Petigura 2018). This depletion can be explained as an evaporation valley between planets with and without H/He that is caused by atmospheric escape. This was predicted independently in a similar fashion by several theoretical models (Owen & Wu 2013; Lopez & Fortney 2013; Jin et al. 2014; Chen & Rogers 2016). The subsequently observationally found locus of the valley agrees with these theoretical predictions and is consistent with a mainly rocky, Earth-like composition without much ice as shown by Owen & Wu (2017); Jin & Mordasini (2018).

Building upon these works, in this paper we have conducted a systematic numerical and analytical study to further constrain the valley’s location, and to understand how it depends on important planetary properties like the orbital distance, the planet’s mass and composition, the H/He envelope mass, as well as the stellar XUV luminosity. An important goal of this evolutionary study is to derive constraints for planet formation models (e.g., Carrera et al. 2018; Mordasini 2018; Brügger et al. 2018; Poon et al. 2020). Regarding formation models, one could for example a priori think that the locus of the valley depends on the postformation envelope-to-core mass ratio and its dependency on orbital distance. This would constrain the efficiency of gas accretion during the nebular phase (e.g., Ikoma & Hori 2012; Lee & Chiang 2015; Cimerman et al. 2017) and the opacity in these

protoplanetary envelopes (Ormel 2014; Mordasini 2014) (which is not the case, as discussed below).

In the first part of the paper (Sect. 2) we have conducted systematic numerical simulations of the evolution of thousands of close-in low-mass planets starting with H/He that undergo atmospheric escape using the Bern planet evolution code `complete21` (Mordasini et al. 2012b; Jin et al. 2014; Linder et al. 2019). To understand the behavior in the relevant parameter space, we have simulated the evolution of model planets on grids of core mass and orbital separation. We have calculated 14 grids, varying the postformation H/He envelope mass, the strength of evaporation (i.e. different stellar  $L_{\text{XUV}}$ , durations of the saturated phase  $t_{\text{sat}}$  or efficiency factors  $\varepsilon$  of energy-limited evaporation), the opacity in the planetary atmosphere, the envelope metallicity, and the core composition. The results of the numerical study (Table 1) can be summarized as follows:

First, the position of the bottom of the valley in the  $a$ – $R$  diagram which corresponds to the largest core at a given orbital distance (or period) that has completely lost its H/He can in all grids be well approximated by a power law of the form  $R_{\text{bare}}(a)/R_{\oplus} = R_{b,0p1}(a/0.1 \text{ AU})^{-e_r}$  (Sect. 2.3). In the reference simulation M3,  $R_{b,0p1} = 1.65 R_{\oplus}$  and  $e_r = 0.27$ . Except for one case (M1) all other simulations exhibit similar exponents  $e_r$ . If the dependency on orbital period  $P$  instead of semimajor axis is considered, this corresponds with Kepler’s law to  $R_{\text{bare}} \propto P^{0.18}$ .

Second, the locus of the valley is only a weak function of the postformation H/He envelope mass, except for envelope masses that are much lower or higher than what is predicted by formation models (Sects. 2.3.3 and 2.3.5). We can write the postformation envelope mass as  $M_{e,0} \propto M_{e,1} M_c^{p_e}$  as suggested by the Kelvin-Helmholtz timescale of envelope cooling during the nebular phase (Sect. 3.2.1, Mordasini et al. 2014). We have then varied the normalization constant  $M_{e,1}$  by one order of magnitude (simulations N1, N2), and explored exponents  $p_e = 1, 2, 3$  (simulations M2, M3, M4). While we find a slight decrease of  $R_{b,0p1}$  with increasing envelope mass,  $R_{b,0p1}$  remained in all these simulations close to the nominal value, namely between 1.60 and 1.68  $R_{\oplus}$ . The reason is that at a given core mass, a higher envelope mass means that there is more gas to evaporate, but also that the planet has a larger radius. Since the (total) mass is for these planets essentially given by the (constant) core mass, this means that the planet has a lower mean density. A lower mean density implies a higher evaporation rate, meaning that the planet with more H/He also loses more.

Third, in the simulations E1 and E2 (Sect. 2.3.6) we have varied the evaporation rate relative to the nominal model by a factor 10. This change in the evaporation rate could be due to a variations of the stellar  $L_{\text{XUV}}$ , the duration of the saturated phase, or uncertainties in the evaporation model, like for example variations of the efficiency factor  $\varepsilon$  in the energy limited approximation (e.g., Kubyskhina et al. 2018b,a). Increasing (decreasing) the escape rate leads as expected to a shift of the valley to larger (smaller) radii, but the shift is moderate, namely to  $R_{b,0p1} = 1.97$  and 1.39  $R_{\oplus}$ , respectively. This corresponds to an increase of the bare core radius by only a factor 1.42 for a factor 100 increase of the evaporation rate. This is to be compared with the observed spread in the  $L_{\text{XUV}}$  levels of young solar-like stars (e.g., Tu et al. 2015) which exhibit a spread of about a factor  $\sim 30$  in the first  $\sim 100$  Myr. Most stars should however pile up around a mean value of  $L_{\text{XUV}} \approx 10^{30} \text{ erg s}^{-1}$  during the first  $\sim 100$  Myr (Johnstone et al. 2015).

Fourth, we have investigate the consequences of the atmospheric opacity. We have usually assumed that the opacity in the



planetary atmospheres corresponds to a condensate-free solar-composition gas (Freedman et al. 2014). However, planet formation models suggest an increasing enrichment with decreasing planet mass (Fortney et al. 2013). In Simulation O1 we have therefore studied how the valley locus changes if the atmospheric opacities are uniformly increased by a factor 10. A higher opacity means that the planets contract less, increasing via the reduced density mass loss. Therefore,  $R_{b,0p1}$  increases to  $1.78 R_{\oplus}$ .

Fifth, we have quantified the impact of the envelope enrichment on the valley position. In simulation O1 two important aspects were neglected: a higher metallicity atmosphere does not only imply a higher opacity, but also more coolants, which reduces the evaporation rate (e.g., Salz et al. 2016). It also implies a higher mean molecular weight. This leads to a higher planetary density, which reduces evaporation as well. To have a more self-consistent picture than in simulation O1, we have also conducted three simulations Z1, Z2, Z3 where the opacity, equation of state of the envelope gas, and evaporation rate were self-consistently changed. A wide range of envelope metal mass fractions  $Z_{\text{enve}} = 0.1, 0.3, \text{ and } 0.5$  were simulated. It is found that the evaporation valley shifts gradually downward with increasing  $Z_{\text{enve}}$  to  $R_{b,0p1} = 1.58, 1.43, \text{ and } 1.32 R_{\oplus}$ , respectively. The effects of the EOS and the reduced evaporation rate thus overwhelm the effect of the opacity. Whether this effect leads to an observable dependency of the valley position as a function of host star [Fe/H] depends on the unknown relation between atmospheric composition of planets near the valley and host star [Fe/H]. Under the (simplistic) assumption that the planetary atmospheric  $Z_{\text{enve}}$  is directly given by the stellar  $Z$ , a shift of the valley by just  $0.02 R_{\oplus}$  would occur, as  $Z_{\text{enve}}$  would in this case only cover a small range from about 0.3 to 3%. Observationally, the valley position does not change by  $\gtrsim 15\%$  in radius (about  $0.3 R_{\oplus}$ ) over a wide range of host star metallicity (Owen & Murray-Clay 2018). Our results show that for the photoevaporative hypothesis of the valley, the planetary  $Z_{\text{enve}}$  could thus change even much more strongly with stellar [Fe/H] and still remain consistent with the observed absence of a (clear) correlation of gap position and host star [Fe/H].

Finally, we have again studied (Owen & Wu 2017; Jin & Mordasini 2018; Gupta & Schlichting 2019) how the valley's location changes if we assume a (hypothetical) completely icy core composition. In this case, the bottom of the valley shifts to  $R_{b,0p1} = 2.68 R_{\oplus}$ , which is the largest difference relative to the nominal case of all simulations conducted here. This illustrates again that the valley position is a good diagnostic of the core composition (Lopez & Fortney 2013). As discussed in Sect. 2.3.6 and recapitulated below, also the fact that the valley is not completely empty, but to some extent filled (as observed, Fulton & Petigura 2018) does not mean that ice-rich compositions are necessarily needed to explain the observations; a nonempty gap is actually even expected also for a homogeneous (rocky) composition.

In the second part of the paper (Sect. 3) we have developed a simple analytical theory (see also Owen & Wu 2017) based on the comparison of the relevant energies, in the same spirit as originally done by Lecavelier Des Etangs (2007) for giant planets. In this analytical model complete evaporation of the envelope occurs if the temporal integral over the stellar XUV irradiation absorbed by a planet (modulo the efficiency factor  $\varepsilon$ ) is larger than the binding energy of the envelope in the gravitational potential of the core, given the special core mass – envelope mass – radius relation of close-in low-mass planets with H/He (Lopez & Fortney 2014).

This approach quickly leads to the main result of the analytical calculation which is the radius of the largest completely stripped core which corresponds to the bottom of the evaporation valley. As a function of orbital period  $P$ , it is given as (Sect. 3.3.2)

$$R_{\text{bare,late}} \approx 1.64 \left(1 + \frac{1}{2} f_{\text{ice}}\right) \left(\frac{P}{10 \text{ day}}\right)^{-0.18} \left(\frac{\varepsilon}{0.1}\right)^{0.135} \times \left(\frac{L_{\text{sat}}}{10^{30} \text{ erg s}^{-1}}\right)^{0.135} \left(\frac{t_{\text{sat}}}{100 \text{ Myr}}\right)^{0.135} R_{\oplus}. \quad (36)$$

for planets around a solar-like star and assuming a constant efficiency factor of evaporation. The middle of the valley lies about  $0.3 R_{\oplus}$  above this value. The numerical results listed above are recovered and understood very well with this expression:

First, regarding the position of the valley, the analytical model shows that the exponent of the period  $P$  dependency,  $P^{-0.18}$  is  $P^{-2p_c/3}$  where  $p_c \approx 0.27$  is the power law exponent in the mass-radius relation of the solid cores,  $R_c \propto M_c^{p_c}$  (Eq. (23)). It is certainly remarkable that two fundamental properties of solid planets are imprinted into the valley's locus (see also Owen & Wu 2017; Gupta & Schlichting 2019): first, their mass-radius scaling in terms of  $d \log R/d \log M = p_c$  that is predicted by interior structure model of solid planets (e.g., Valencia et al. 2006). It determines the valley's distance dependency. Second, the cores' ice mass fraction  $f_{\text{ice}}$ , or more generally speaking, the bulk composition, which determines the absolute position of the valley at a fixed period.

It is likely that in reality, the evaporation efficiency factor  $\varepsilon$  is not constant in contrast to what was assumed here in the nominal model. A variable  $\varepsilon$  leads to a shallower slope (Owen & Wu 2017; Wu 2019), closer to the observational result (Sect. 4). In reality, the situation might in any case be more complex, because of the following: XUV-driven atmospheric photoevaporation might not be the only mechanism driving the loss of H/He envelopes. Other mechanisms could be impacts (Liu et al. 2015; Schlichting et al. 2015; Wyatt et al. 2019) or mass loss powered by the luminosity of a planet's cooling core (Ginzburg et al. 2018; Gupta & Schlichting 2019, 2020). Additionally, some planets especially at larger distances may only formed after the dispersal of the gas disk, more similar to the way the Earth likely formed (Owen & Murray-Clay 2018; Swain et al. 2019). Lopez & Rice (2018) found that a such a formation in a gas-free environment should lead to a transition radius from solid super-Earth to sub-Neptunes that increases with orbital distance, that is to say, with the opposite dependency than evaporation. If these processes are at work in parallel with a certain contribution of each, this could lead to a different (in particular shallower) slope that could not be explained with one mechanism only.

Second, regarding the dependency on the envelope mass, in the analytical model the locus of the valley is not just weakly dependent on the postformation envelope mass (as found numerically), but even independent of it, as shown by the equation. As becomes clear when deriving an approximative core mass – envelope mass – radius relation for low-mass planets with H/He (Sect. 3.2.2), this relation is such that the two aforementioned effects occurring when increasing the envelope mass (more mass to evaporate, but also a higher escape rate) balance each other. It should be noted that these results are only valid in the (important) domain where the envelope is massive enough that the total planet radius is clearly larger than the core radius, but not so massive that the envelope's self-gravity becomes relevant (Sect. 3.2.3).

Third, for the dependency on the strength of evaporation, the analytical model shows that quantities describing the strength of evaporation ( $L_{\text{sat}}$ ,  $t_{\text{sat}}$  and  $\varepsilon$ ) enter the expression for  $R_{\text{bare,late}}$  only with a very weak exponent,  $0.135 = p_c/2$ , explaining why such a weak dependency was found numerically.

Finally, concerning the dependency on the core composition,  $R_{\text{bare}}$  is found analytically to increase linearly with ice mass fraction. This corresponds to a – in comparison – rather strong dependency.

## 6. Conclusions

From the points summarized above, we can draw three important conclusions.

1. Because of its weak dependency on the postformation envelope mass, the position of the valley cannot be used to constrain well the efficiency of gas accretion during the nebular phase (except of course for the fact that its sheer existence shows that these planets have accreted H/He). But it actually appears likely that it is because of this independence on the postformation envelope mass that the valley is clearly (or even at all) visible observationally: it means that the spread in the postformation envelope masses resulting from the formation epoch does not matter.
2. In a similar vein, the weak dependency of the valley's position on the strength of evaporation also has important implications: First, this explains why on the theoretical side various models of evaporation (like Owen & Wu 2013; Lopez & Fortney 2013; Jin et al. 2014; Chen & Rogers 2016) found similar locations of the valley, or in other words, why the details (one could also say uncertainties) of these models do not matter much. Second, on the observational side, the weak dependency on the strength of evaporation probably also explains why we see a valley at all: the observed spread in the  $L_{\text{XUV}}$  and saturation times of young stars does not alter the location of the valley to an extent that it completely blurs away. Combined, these two points mean that the evaporation valley is a robust feature.
3. At the same time, given the significant spread of  $L_{\text{XUV}}$  of young stars, the dependency of the valley's location on it is still strong enough to explain why the valley is not completely empty: as discussed in the context of the simulations M3, E1, E2 (Sect. 2.3.6) the location of the valley likely varies from star to star because of different  $L_{\text{XUV}}$ , but also different efficiency factors by about  $0.5 R_{\oplus}$  around the mean value. The intrinsic valley width has a similar magnitude (Owen & Wu 2017). Therefore, we can infer that by overlaying these simulations with different strength of evaporation, there would still be a depleted region, but not a completely empty one. This corresponds to the observational result (Fulton & Petigura 2018). The individual numerical simulations presented here give the misleading impression that the valley should be empty, but this is simply an artifact of assuming an identical  $L_{\text{XUV}}$  for all stars in one simulation. Clearly, in the actual observations, we in contrast see the overlay of all individual star-specific  $L_{\text{XUV}}$  and thus valleys which are not exactly at the same position. As we mentioned above, the fact that the valley is not completely empty, but to some extent filled does thus not mean that a compositional diversity of the cores is needed to explain the observations. Rather, a partially filled valley is (also) a natural consequence of evaporation given the spread of stellar rotation rates and thus  $L_{\text{XUV}}$ .

Looking ahead, to further understand the valley and its origin, and in particular atmospheric escape as a possible explanation, this hypothesis can be put to a number of observational tests: first, simply by determining precisely the density of transiting planets on both sides of the valley for example with TESS and CHEOPS plus high-accuracy radial velocities (e.g., Dumusque et al. 2019). This would reveal whether the densities are indeed consistent with on one side super-Earth planets with an Earth-like composition without H/He, and sub-Neptunes with H/He on the other side. While this test is obvious and simple in principle, it is not so trivial in practice, because the observational error bars on the masses and thus the density are often significant given the low masses of the planets. As shown in Jin & Mordasini (2018) where a simple analysis of the densities of planets on both sides of the valley was done, the currently available sample of planets with sufficiently well-constrained densities is small. But it at least does not contain planets with densities inconsistent with evaporation, namely planets with a density only explicable with H/He below the valley, or high density planets without volatile above it. The second test would be to obtain UV observations of atmospheric escape rates of planets on both sides with the HST. As shown for example in Jin et al. (2014, their Figs. 5 and 8), at an age of 5 Gyr, close-in planets just above the valley should have evaporation rates of  $10^9$ – $10^{10}$   $\text{g s}^{-1}$  (and much more at  $\sim 100$  Myr), while below it, the escape rates should be orders of magnitude lower. For comparison, the escape rate of the warm Neptune-mass planet GJ 3470b was measured to be about  $10^{10}$   $\text{g s}^{-1}$  (Bourrier et al. 2018). Third, studying the dependency of the valleys on the spectral type of the host star (see Fulton & Petigura 2018; Wu 2019; Gupta & Schlichting 2020; McDonald et al. 2019; Cloutier & Menou 2020) should in particular allow one to disentangle the different proposed loss mechanisms. Finally, with PLATO, it may even become possible to measure the valley's locus as a function of time, and to compare with the various theoretical predictions.

*Acknowledgements.* I would like to thank Sheng Jin, Paul Mollière, Gabriel-Dominique Marleau, Apurva Oza, Alexandre Emsenhuber, Jonas Haldemann, and Julia Venturini. I thank the referee for helpful comments and acknowledge the support from the Swiss National Science Foundation under grant BSSGI0\_155816 “PlanetsInTime”. Parts of this work have been carried out within the framework of the NCCR PlanetS supported by the Swiss National Science Foundation.

## References

- Alibert, Y., Mordasini, C., Benz, W., & Winisdoerffer, C. 2005, *A&A*, 434, 343
- Baraffe, I., Chabrier, G., & Barman, T. S. 2008, *A&A*, 482, 315
- Baraffe, I., Chabrier, G., Fortney, J., & Sotin, C. 2014, *Protostars and Planets VI* (Tucson: University of Arizona Press), 763
- Baruteau, C., Bai, X., Mordasini, C., & Mollière, P. 2016, *Space Sci. Rev.*, 205, 77
- Bell, K. R., & Lin, D. N. C. 1994, *ApJ*, 427, 987
- Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, *ApJ*, 866, 99
- Biersteker, J. B., & Schlichting, H. E. 2019, *MNRAS*, 485, 4454
- Bourrier, V., Lecavelier des Etangs, A., Ehrenreich, D., et al. 2018, *A&A*, 620, A147
- Bower, D. J., Kitzmann, D., Wolf, A. S., et al. 2019, *A&A*, 631, A103
- Brouwers, M. G., Vazan, A., & Ormel, C. W. 2018, *A&A*, 611, A65
- Brügger, N., Alibert, Y., Ataiee, S., & Benz, W. 2018, *A&A*, 619, A174
- Burrows, A., Hubeny, I., Budaj, J., & Hubbard, W. B. 2007, *ApJ*, 661, 502
- Carrera, D., Ford, E. B., Izidoro, A., et al. 2018, *ApJ*, 866, 104
- Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, *Science*, 337, 556
- Chen, J., & Kipping, D. 2017, *ApJ*, 834, 17
- Chen, H., & Rogers, L. A. 2016, *ApJ*, 831, 180
- Cimerman, N. P., Kuiper, R., & Ormel, C. W. 2017, *MNRAS*, 471, 4662
- Cloutier, R., & Menou, K. 2020, *AJ*, 159, 211
- Dawson, R. I., Lee, E. J., & Chiang, E. 2016, *ApJ*, 822, 54

- Dittkrick, K.-M., Mordasini, C., Klahr, H., Alibert, Y., & Henning, T. 2014, *A&A*, **567**, A121
- Dorn, C., Venturini, J., Khan, A., et al. 2017, *A&A*, **597**, A37
- Dressing, C. D., Charbonneau, D., Dumusque, X., et al. 2015, *ApJ*, **800**, 135
- Dumusque, X., Turner, O., Dorn, C., et al. 2019, *A&A*, **627**, A43
- Ehrenreich, D., Bourrier, V., Wheatley, P. J., et al. 2015, *Nature*, **522**, 459
- Ehrenreich, D., & Désert, J. M. 2011, *A&A*, **529**, A136
- Ercolano, B., & Clarke, C. J. 2010, *MNRAS*, **402**, 2735
- Estrela, R., Swain, M., Gupta, A., Sotin, C., & Valio, A. 2020, *ApJ*, submitted [arXiv:2001.01299]
- Fisher, C., & Heng, K. 2018, *MNRAS*, **481**, 4698
- Fortier, A., Beck, T., Benz, W., et al. 2014, *SPIE Conf. Ser.*, **9143**, 91432J
- Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, *ApJ*, **659**, 1661
- Fortney, J. J., Ikoma, M., Nettelmann, N., Guillot, T., & Marley, M. S. 2011, *ApJ*, **729**, 32
- Fortney, J. J., Mordasini, C., Nettelmann, N., et al. 2013, *ApJ*, **775**, 80
- Fossati, L., Erkaev, N. V., Lammer, H., et al. 2017, *A&A*, **598**, A90
- Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., et al. 2014, *ApJS*, **214**, 25
- French, M., Mattsson, T. R., Nettelmann, N., & Redmer, R. 2009, *Phys. Rev. B*, **79**, 054107
- 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
- Gandolfi, D., Barragán, O., Hatzes, A. P., et al. 2017, *AJ*, **154**, 123
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, *ApJ*, **825**, 29
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, *MNRAS*, **476**, 759
- Grasset, O., Schneider, J., & Sotin, C. 2009, *ApJ*, **693**, 722
- Guenther, E. W., Barragán, O., Dai, F., et al. 2017, *A&A*, **608**, A93
- Guillot, T. 2010, *A&A*, **520**, 27
- Guillot, T., & Gautier, D. 2014, in *Treatise on Geophysics*, 2nd edn., eds. T. Spohn, & G. Schubert (Amsterdam: Elsevier)
- Guillot, T., & Hueso, R. 2006, *MNRAS*, **367**, L47
- Gupta, A., & Schlichting, H. E. 2019, *MNRAS*, **487**, 24
- Gupta, A., & Schlichting, H. E. 2020, *MNRAS*, **493**, 792
- Haisch, K. E., Lada, E. A., & Lada, C. J. 2001, *ApJ*, **553**, L153
- Hatzes, A. P., & Rauer, H. 2015, *ApJ*, **810**, L25
- Ida, S., & Lin, D. N. C. 2004, *ApJ*, **616**, 567
- Ikoma, M., & Hori, Y. 2012, *ApJ*, **753**, 66
- Ikoma, M., Nakazawa, K., & Emori, H. 2000, *ApJ*, **537**, 1013
- Inamdar, N. K., & Schlichting, H. E. 2015, *MNRAS*, **448**, 1751
- Jackson, A. P., Davis, T. A., & Wheatley, P. J. 2012, *MNRAS*, **422**, 2024
- Jacobson, S. A., Morbidelli, A., Raymond, S. N., et al. 2014, *Nature*, **508**, 84
- Jin, S., & Mordasini, C. 2018, *ApJ*, **853**, 163
- Jin, S., Mordasini, C., Parmentier, V., et al. 2014, *ApJ*, **795**, 65
- Johnson, J. A., Petigura, E. A., Fulton, B. J., et al. 2017, *AJ*, **154**, 108
- Johnstone, C. P., Güdel, M., Brott, I., & Lüftinger, T. 2015, *A&A*, **577**, A28
- Johnstone, C. P., Güdel, M., Lammer, H., & Kislyakova, K. G. 2018, *A&A*, **617**, A107
- Jontof-Hutter, D., Ford, E. B., Rowe, J. F., et al. 2016, *ApJ*, **820**, 39
- Kamland Collaboration (Gando, A., et al.) 2011, *Nat. Geosci.*, **4**, 647
- Kubyskhina, D., Fossati, L., Erkaev, N. V., et al. 2018a, *ApJ*, **866**, L18
- Kubyskhina, D., Fossati, L., Erkaev, N. V., et al. 2018b, *A&A*, **619**, A151
- Kurokawa, H., & Nakamoto, T. 2014, *ApJ*, **783**, 54
- Lecavelier Des Etangs, A. 2007, *A&A*, **461**, 1185
- Lecavelier des Etangs, A., Vidal-Madjar, A., McConnell, J. C., & Hébrard, G. 2004, *A&A*, **418**, L1
- Lecointe, J. 2011, PhD thesis, Université de Bordeaux, France
- Lee, E. J., & Chiang, E. 2015, *ApJ*, **811**, 41
- Léger, A., Rouan, D., Schneider, J., et al. 2009, *A&A*, **506**, 287
- Linder, E. F., Mordasini, C., Mollière, P., et al. 2019, *A&A*, **623**, A85
- Lissauer, J. J., Jontof-Hutter, D., Rowe, J. F., et al. 2013, *ApJ*, **770**, 131
- Liu, S.-F., Hori, Y., Lin, D. N. C., & Asphaug, E. 2015, *ApJ*, **812**, 164
- Lodders, K. 2003, *ApJ*, **591**, 1220
- Lopez, E. D. 2017, *MNRAS*, **472**, 245
- Lopez, E. D., & Fortney, J. J. 2013, *ApJ*, **776**, 2
- Lopez, E. D., & Fortney, J. J. 2014, *ApJ*, **792**, 1
- Lopez, E. D., & Rice, K. 2018, *MNRAS*, **479**, 5303
- Lozovsky, M., Helled, R., Dorn, C., & Venturini, J. 2018, *ApJ*, **866**, 49
- Lundkvist, M. S., Kjeldsen, H., Albrecht, S., et al. 2016, *Nat. Commun.*, **7**, 11201
- MacDonald, M. G. 2019, *MNRAS*, **487**, 5062
- Martinez, C. F., Cunha, K., Ghezzi, L., & Smith, V. V. 2019, *ApJ*, **875**, 29
- Mazeh, T., Holczer, T., & Faigler, S. 2016, *A&A*, **589**, A75
- Mazevet, S., Licari, A., Chabrier, G., & Potekhin, A. Y. 2019, *A&A*, **621**, A128
- McDonald, G. D., Kreidberg, L., & Lopez, E. 2019, *ApJ*, **876**, 22
- Melosh, H. J. 2007, *Meteorit. Planet. Sci.*, **42**, 2079
- Mordasini, C. 2014, *A&A*, **572**, A118
- Mordasini, C. 2018, *Handbook of Exoplanets*, eds. J. A. Belmonte, & H. Deeg (Berlin: Springer International Publishing AG), 143
- Mordasini, C., Alibert, Y., & Benz, W. 2006, *Status of and Prospects for Hot Jupiter Studies*, eds. L. Arnold, F. Bouchy, & C. Moutou (Paris: Frontier Group), 84
- Mordasini, C., Alibert, Y., & Benz, W. 2009, *A&A*, **501**, 1139
- Mordasini, C., Alibert, Y., Georgy, C., et al. 2012a, *A&A*, **547**, A112
- Mordasini, C., Alibert, Y., Klahr, H., & Henning, T. 2012b, *A&A*, **547**, A111
- Mordasini, C., Klahr, H., Alibert, Y., Miller, N., & Henning, T. 2014, *A&A*, **566**, A141
- Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, *ApJ*, **832**, 41
- Movshovitz, N., Bodenheimer, P. H., Podolak, M., & Lissauer, J. J. 2010, *Icarus*, **209**, 616
- Murray-Clay, R. A., Chiang, E. I., & Murray, N. 2009, *ApJ*, **693**, 23
- Nettelmann, N., Holst, B., Kietzmann, A., et al. 2008, *ApJ*, **683**, 1217
- Ormel, C. W. 2014, *ApJ*, **789**, L18
- Owen, J. E., & Adams, F. C. 2019, *MNRAS*, **490**, 15
- Owen, J. E., & Alvarez, M. A. 2016, *ApJ*, **816**, 34
- Owen, J. E., & Estrada, B. C. 2019, *MNRAS*, **491**, 5287
- Owen, J. E., & Jackson, A. P. 2012, *MNRAS*, **425**, 2931
- Owen, J. E., & Morton, T. D. 2016, *ApJ*, **819**, L10
- Owen, J. E., & Murray-Clay, R. 2018, *MNRAS*, **480**, 2206
- Owen, J. E., & Wu, Y. 2013, *ApJ*, **775**, 105
- Owen, J. E., & Wu, Y. 2016, *ApJ*, **817**, 107
- Owen, J. E., & Wu, Y. 2017, *ApJ*, **847**, 29
- Petigura, E. A., Howard, A. W., Marcy, G. W., et al. 2017, *AJ*, **154**, 107
- Pinhas, A., Madhusudhan, N., & Clarke, C. 2016, *MNRAS*, **463**, 4516
- Podolak, M. 2003, *Icarus*, **165**, 428
- Podolak, M., Pollack, J. B., & Reynolds, R. T. 1988, *Icarus*, **73**, 163
- Poon, S. T. S., Nelson, R. P., Jacobson, S. A., & Morbidelli, A. 2020, *MNRAS*, **491**, 5595
- Pu, B., & Valencia, D. 2017, *ApJ*, **846**, 47
- Quillen, A. C., Bodman, E., & Moore, A. 2013, *MNRAS*, **435**, 2256
- Rauer, H., Catala, C., Aerts, C., et al. 2014, *Exp. Astron.*, **38**, 249
- Ribas, I., Guinan, E. F., Güdel, M., & Audard, M. 2005, *ApJ*, **622**, 680
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, *SPIE Conf. Ser.*, **9143**, 914320
- Rogers, L. A. 2015, *ApJ*, **801**, 41
- Rogers, L. A., & Seager, S. 2010, *ApJ*, **712**, 974
- Rogers, L. A., Bodenheimer, P. H., Lissauer, J. J., & Seager, S. 2011, *ApJ*, **738**, 59
- Salz, M., Czesla, S., Schneider, P. C., & Schmitt, J. H. M. M. 2016, *A&A*, **586**, A75
- Saumon, D., Chabrier, G., & van Horn, H. M. 1995, *ApJS*, **99**, 713
- Schlichting, H. E., Sari, R., & Yalinewich, A. 2015, *Icarus*, **247**, 81
- Seager, S., Kuchner, M. J., Hier-Majumder, C. A., & Militzer, B. 2007, *ApJ*, **669**, 1279
- Soubiran, F., & Militzer, B. 2015, *ApJ*, **806**, 228
- Swain, M. R., Estrela, R., Sotin, C., Roudier, G. M., & Zellem, R. T. 2019, *ApJ*, **881**, 117
- Teske, J. K., Thorngren, D., Fortney, J. J., Hinkel, N., & Brewer, J. M. 2019, *AJ*, **158**, 239
- Thompson, S. L. 1990, Sandia Natl. Lab. Doc.
- Thorngren, D., & Fortney, J. J. 2019, *ApJ*, **874**, L31
- Thorngren, D. P., Fortney, J. J., Murray-Clay, R. A., & Lopez, E. D. 2016, *ApJ*, **831**, 64
- Tu, L., Johnstone, C. P., Güdel, M., & Lammer, H. 2015, *A&A*, **577**, L3
- Turtle, E. P., & Pierazzo, E. 2001, *Science*, **294**, 1326
- Valencia, D., O'Connell, R. J., & Sasselov, D. 2006, *Icarus*, **181**, 545
- Valencia, D., Guillot, T., Parmentier, V., & Freedman, R. S. 2013, *ApJ*, **775**, 10
- Valetta, C., & Helled, R. 2019, *ApJ*, **871**, 127
- Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, *MNRAS*, **479**, 4786
- Venturini, J., Alibert, Y., & Benz, W. 2016, *A&A*, **596**, A90
- Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J.-M., et al. 2003, *Nature*, **422**, 143
- Wallack, N. L., Knutson, H. A., Morley, C. V., et al. 2019, *AJ*, **158**, 217
- Watson, A. J., Donahue, T. M., & Walker, J. C. 1981, *Icarus*, **48**, 150
- Weiss, L. M., & Marcy, G. W. 2014, *ApJ*, **783**, L6
- Wolfgang, A., & Lopez, E. 2015, *ApJ*, **806**, 183
- Wu, Y. 2019, *ApJ*, **874**, 91
- Wyatt, M. C., Kral, Q., & Sinclair, C. A. 2019, *MNRAS*, **491**, 782
- Zeng, L., Jacobsen, S. B., Sasselov, D. D., et al. 2019, *Proc. Acad. Sci.*, **116**, 9723



## Appendix A: Analytical fits for the luminosity of low-mass planets as a function of core mass, envelope mass, and time

In this appendix we use results from a high number of numerical calculations of planetary evolution to derive a fit for the mean luminosity of low-mass planets as a function of their core mass, envelope mass, and time.

In a similar spirit as the analytical expressions derived in Lopez & Fortney (2014) for the radii of close-in low-mass planets, these fits are not meant to be an exact replacement of a full cooling calculation of an individual planet. But such fits are of interest for a first rough estimate of the luminosity, and can in particular be used as input for static (time-independent) interior structure models like Rogers et al. (2011); Dorn et al. (2017); Lozovsky et al. (2018). Such models need the luminosity as a function of planet properties and time as an input parameter.

To derive the fit, we consider the luminosity of nonevaporating low-mass planets in a population that repeats the coupled formation and evolution calculations presented in Mordasini et al. (2012a), but now includes also the cooling model of the solid core described in Linder et al. (2019). The luminosity thus includes the contribution of the cooling and contraction of the core and envelope, plus the radiogenic luminosity. To derive the fit, planets with a core mass  $1 \leq M_c/M_\oplus \leq 20$ , an envelope mass  $M_e/M_\oplus \leq 20$  and an ice mass fraction of  $f_{\text{ice}} \leq 0.1$  are included.

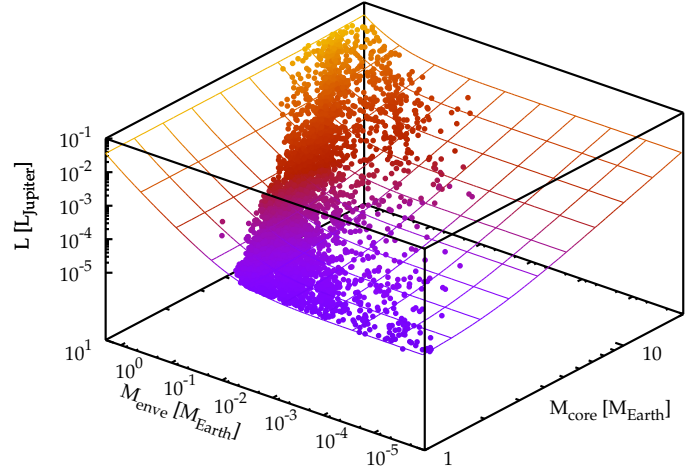
Figure A.1 shows the luminosity of the synthetic planets as a function of core and envelope mass at an age of 5 Gyr. The plot shows that the luminosity can be approximated with a power law. We chose a biquadratic fit in  $M_c$  and  $M_e$ , giving a fit of the form

$$\frac{L}{L_\star} = a_0 + b_1 \left( \frac{M_c}{M_\oplus} \right) + b_2 \left( \frac{M_c}{M_\oplus} \right)^2 + c_1 \left( \frac{M_e}{M_\oplus} \right) + c_2 \left( \frac{M_e}{M_\oplus} \right)^2. \quad (\text{A.1})$$

The parameters  $a_0$ ,  $b_1$ ,  $b_2$ ,  $c_1$  and  $c_2$  are derived by the least square method separately at 14 different moments in time between 0.1 and 10 Gyr. We note that despite the fact that the expression gives  $L$  as a function of  $M_c$  and  $M_e$  separately, these two quantities are statistically not independent of each other in the population that was used to derive it, as is visible in Fig. A.1. Since the cooling of the core and envelope are interdependent, this means that the fit is strictly speaking only applicable to planets that follow a similar  $M_c - M_e$  relation.

It should be noted that the expression is more complex than the scaling of the postformation luminosity (at 10 Myr) used in the parameter study presented above (Eq. (14)). However, except for very low-mass (a few Earth masses) planets with massive envelopes with a mass fraction of  $\geq 10\%$  which are not usually forming in the formation simulations, the expressions yield roughly speaking comparable luminosities at 100 Myr when accounting for the fact that the luminosities decreases by approximately one order of magnitude from 10 to 100 Myr. But the fit of Eq. (A.1) should of course be preferred when possible.

It is interesting to consider the luminosity predicted by the fit for specific planets. For an Earth-like planet ( $M_c/M_\oplus = 1$ ,  $M_e/M_\oplus = 0$ ), the expressions yields a luminosity of  $0.6 L_\oplus$  at 5 Gyr, where  $L_\oplus$  is the intrinsic luminosity of the Earth originating from radiogenic decay and delayed secular cooling, that is estimated to be about  $4 \times 10^{20} \text{ erg s}^{-1}$  (Kamland Collaboration 2011), or about  $10^{-4} L_\star$ . For a Neptune-like planet (assuming  $M_c/M_\oplus = 15.4$ ,  $M_e/M_\oplus = 1.7$ ) of the same age, the expression



**Fig. A.1.** Luminosity of low-mass planets with H/He as a function of core and envelope mass at 5 Gyr. The points are the luminosities of synthetic planets in a population around a solar-like star. The grid shows the corresponding biquadratic fit.

yields  $0.007 L_\star$ , while the observed value is about  $0.01 L_\star$  (Guillot & Gautier 2014). Finally, for a Jovian-like planet with  $M_c/M_\oplus = 17.8$  and  $M_e/M_\oplus = 300$  which is 15 times higher than the maximal envelope mass used to derive the fit (meaning that the fit is in principle not usable), it still predicts with  $2.78 L_\star$  at 4.5 Gyr a value that is not orders of magnitudes off.

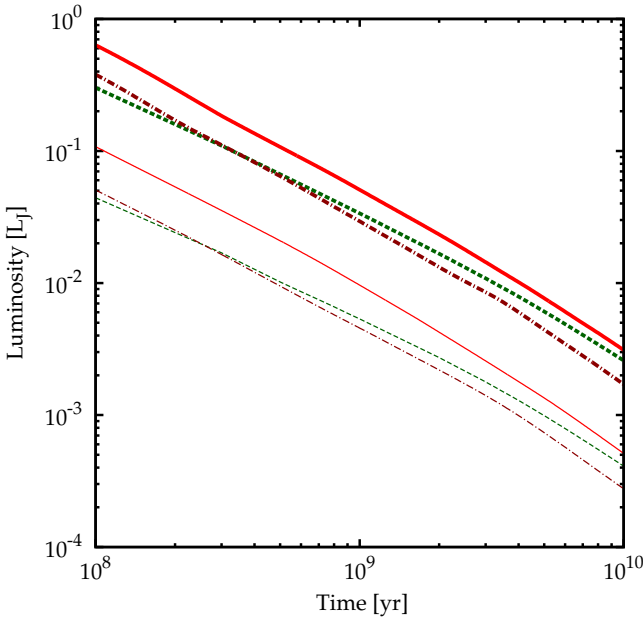
Finally, we compare the luminosity as a function of time as predicted by the analytic fit with the direct cooling calculations. One cannot expect that the two approaches yield identical results for an individual planet, because the fit reproduces only the statistical mean behavior, and a subpopulation of synthetic planets only was used to derive the  $M_c - M_e - t$  relation. This in particular means that individual planets may have different mass fractions of ice in the core, as well as different  $M_c - M_e$  relations than the ones used to derive the fit.

Figure A.2 shows the luminosity as a function of time for a 5 and 15  $M_\oplus$  planet (total mass) with an envelope of 0.5 and 2.7  $M_\oplus$ , respectively. For both planets, three lines are shown: the cooling curve obtained numerically for a core consisting of 50% ice and 50% rocky material, the cooling curve as found with identical settings except for a completely rocky core instead, and finally the luminosity predicted by Eq. (A.1), interpolating linearly in  $\log(L(\log(t)))$  using the parameters for the temporal dependency from Table A.1. We see that in both cases, the fit and numerical result yield roughly similar result, but that at early times, the luminosity predicted by the fit is approximately a factor two smaller than in the nominal model. The agreement with the numerical cooling curve is clearly better for planets with a rocky core, with relative differences that are on the order of 30% or less. This better agreement is not surprising, since the fit was derived for planets with  $f_{\text{ice}} \leq 0.1$ , that is, with (mainly) rocky cores. We thus see that for otherwise identical parameters, a planet with a rocky core has a lower luminosity. This is due to the lower thermal expansion coefficient and the lower specific heat capacity of silicates relative to ice (Baraffe et al. 2008; Linder et al. 2019). A simple Fortran program implementing the interpolation routine can be found online<sup>2</sup>.

<sup>2</sup> [https://www.space.unibe.ch/research/research\\_groups/planets\\_in\\_time/numerical\\_data/index\\_eng.html](https://www.space.unibe.ch/research/research_groups/planets_in_time/numerical_data/index_eng.html)

**Table A.1.** Parameters for the analytical fit of the luminosity of low-mass planets.

Time [Gyr]	$a_0$	$b_1$	$b_2$	$c_1$	$c_2$
0.1	0.002529060	-0.002002380	0.001044080	0.05864850	0.000967878
0.3	0.001213240	-0.000533601	0.000360703	0.02141140	0.000368533
0.5	0.000707416	-0.000394131	0.000212475	0.01381380	0.000189456
0.8	0.000423376	-0.000187283	0.000125872	0.00887292	0.000117141
1.0	0.000352187	-0.000141480	$9.94382 \times 10^{-5}$	0.00718831	$9.20563 \times 10^{-5}$
2.0	0.000175775	$-4.07832 \times 10^{-5}$	$4.58530 \times 10^{-5}$	0.00357941	$5.52851 \times 10^{-5}$
3.0	0.000114120	$-2.09944 \times 10^{-5}$	$2.91169 \times 10^{-5}$	0.00232693	$4.00546 \times 10^{-5}$
4.0	$8.81462 \times 10^{-5}$	$-2.07673 \times 10^{-5}$	$2.12932 \times 10^{-5}$	0.00171412	$2.90984 \times 10^{-5}$
5.0	$6.91819 \times 10^{-5}$	$-1.90159 \times 10^{-5}$	$1.62128 \times 10^{-5}$	0.00134355	$2.30387 \times 10^{-5}$
6.0	$5.49615 \times 10^{-5}$	$-1.68620 \times 10^{-5}$	$1.29045 \times 10^{-5}$	0.00109019	$1.96163 \times 10^{-5}$
7.0	$4.50320 \times 10^{-5}$	$-1.51951 \times 10^{-5}$	$1.05948 \times 10^{-5}$	0.00091005	$1.70934 \times 10^{-5}$
8.0	$3.80363 \times 10^{-5}$	$-1.40113 \times 10^{-5}$	$8.93639 \times 10^{-6}$	0.00077687	$1.50107 \times 10^{-5}$
9.0	$3.30102 \times 10^{-5}$	$-1.31146 \times 10^{-5}$	$7.69121 \times 10^{-6}$	0.000675243	$1.32482 \times 10^{-5}$
10.0	$2.92937 \times 10^{-5}$	$-1.24023 \times 10^{-5}$	$6.73922 \times 10^{-6}$	0.000595191	$1.17809 \times 10^{-5}$



**Fig. A.2.** Comparison of the luminosity as a function of time as predicted by direct cooling calculations and the fit. The upper three thicker lines are for a  $15 M_{\oplus}$  planet while the three thin lines below are for a  $5 M_{\oplus}$  planet. The green dashed line is the fit of Eq. (A.1). The other lines are direct simulations. The brown dashed-dotted and the red solid lines assume a rocky core, and a core consisting of 50% ice and 50% rocky material, respectively.

## Appendix B: ANEOS equation of state for $H_2O$

ANEOS (Thompson 1990) is an equation of state that yields the thermodynamic properties of various materials derived from an analytical expression of the Helmholtz free energy. It has been used widely in the context of planetary interiors (for example Baraffe et al. 2008; Thorngren et al. 2016; Lopez 2017; Lozovsky et al. 2018). While ANEOS has a number of limitations (see Baraffe et al. 2014 for a discussion), and while there are more modern water EOSs based on ab initio simulations (e.g., Nettelmann et al. 2008; French et al. 2009; Soubiran & Militzer 2015; Mazevet et al. 2019), we are here interested mainly in an EOS that covers a wide range of validity as opposed to a very

high degree of accuracy: in the current paper,  $H_2O$  is a placeholder for heavy elements in general. Additionally, ANEOS has the advantage of yielding a large number of thermodynamical quantities of interest, including the density  $\rho$ , specific internal energy  $u$ , specific entropy  $s$ , specific thermal capacity at constant volume  $c_v$ , sound speed  $c_s$ , and the radiative opacity  $\kappa$ . It also provides the derivative of the pressure with respect to temperature at constant density, and the derivative of the pressure with respect to density at constant temperature. This makes it in particular possible to calculate via the Maxwell relations the adiabatic gradient  $\nabla_{ad}$  which is needed for planetary interior structure calculations as (Leconte 2011)

$$\begin{aligned} \nabla_{ad} &= \left. \frac{\partial \ln T}{\partial \ln P} \right|_s = \frac{P}{T} \left( \left. \frac{\partial P}{\partial T} \right|_{\rho} + \left. \frac{\partial P}{\partial \rho} \right|_T \left. \frac{\partial \rho}{\partial T} \right|_s \right)^{-1} \\ &= \frac{P \left. \frac{\partial P}{\partial T} \right|_{\rho}}{c_v \rho^2 \left. \frac{\partial P}{\partial \rho} \right|_T + T \left. \frac{\partial P}{\partial T} \right|_{\rho}^2}. \end{aligned} \quad (B.1)$$

The fundamental approach of ANEOS is to write an analytical expression of the Helmholtz free energy  $F$  in the system.  $F$  is minimal at equilibrium for a system with constant  $T$  and  $V$  (or  $\rho$ ). Therefore, ANEOS takes  $T$  and  $\rho$  as input. For the current application for planetary interiors, where  $T$  and  $P$  are the independent variables, we have retabulated the ANEOS output. In ANEOS,  $F(\rho, T)$  is written as sum of three contributions (Melosh 2007)

$$F(\rho, T) = F_{cold}(\rho) + F_{thermal}(\rho, T) + F_{electronic}(\rho, T) \quad (B.2)$$

where the first term represents a ‘‘cold’’ component of atomic interactions independent of temperature such as the interatomic potential (modeled as a Morse potential). The second is a ‘‘thermal’’ component for the thermal energy of the nuclei and their associated electrons. The third term is an ‘‘electronic’’ contribution that represents the energies associated with the ionization of atoms. For some of these three contributions, internally different models can be used. The thermal contribution is a combination of the free energy of a generalized Debye solid and the monoatomic gas limit at high temperatures. In a Debye solid, the energy is calculated as the phonon energy times the average number of phonons times the number of modes. Ionization can be treated with either via a Saha or Thomas Fermi model. At very

high pressures, ANEOS approaches the Thomas-Fermi limit of a free electron gas with embedded nuclei, where the pressure of the degenerate electron gas is obtained via the Thomas-Fermi-Dirac approximation. An advantage of ANEOS is thus a correct behavior in the asymptotic limit of very high pressures.

Once  $F$  is known, all other thermodynamic quantities can be calculated. For example, the entropy and pressure are obtained as

$$S = - \left. \frac{\partial F}{\partial T} \right|_P \quad P = \rho^2 \left. \frac{\partial F}{\partial \rho} \right|_T. \quad (\text{B.3})$$

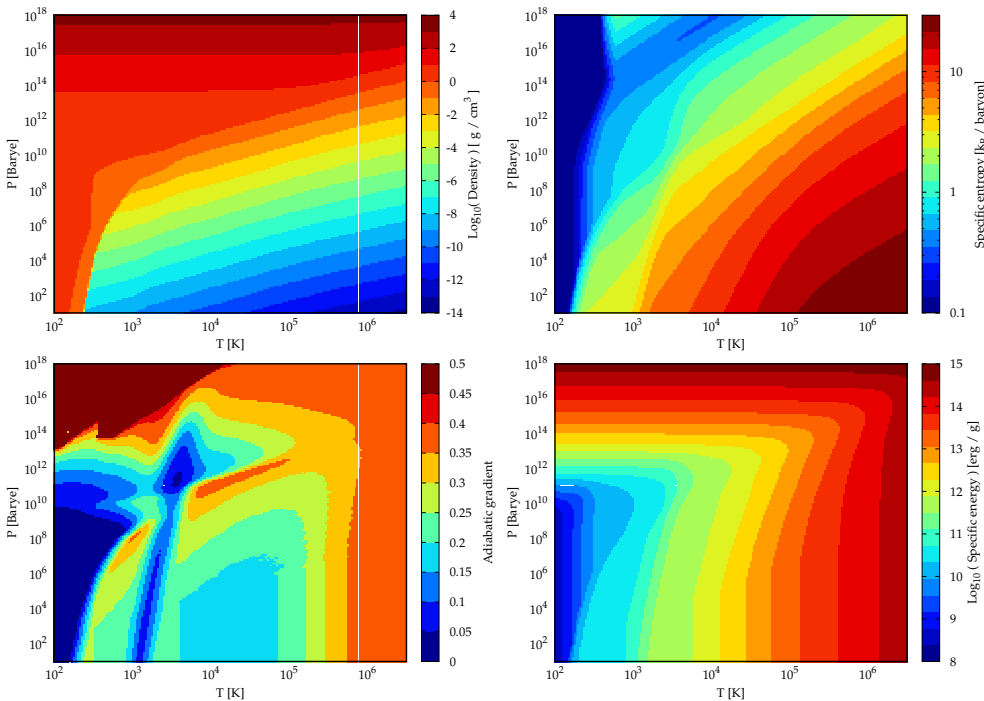
Because of its first principles approach, ANEOS includes a treatment of phase changes, which makes it, at least in principle, superior to some empirical EOS or tabulated EOS which are often only available for a certain phase or  $P-T$  domain. Several models for phase changes are included in ANEOS. Here we use the solid-liquid-gas with ionization mode (ANEOS Mode 4).

ANEOS requires an extended set of coefficients for any given material which can in part be derived from laboratory experiments, like shock wave data. For  $\text{H}_2\text{O}$ , we use the following coefficients (see also [Turtle & Pierazzo 2001](#)):

```

ANEOS1 -2 'H2O' THUG=-1 RHUG=-1 LONG
* nel neos rhoref tref pref Bref Grun. Tdebye
ANEOS2 2 4 1.11 0.02008 0 -1.7e5 0.58 -0.045
* Tg 3C24 Es Tmelt C53 C54 H0 C41
ANEOS3 .9 2 3.34e10 0.0224 0 0 0 0
* rhomin D1 D2 D3 D4 D5 Hf -rho1iq
ANEOS4 0.0 0.0 0.0 0.0 0.0 0.0 0.0 4.85e9 -1.
* Up Lo alpha beta gamma C60 C61 C62
ANEOS5 0.0 0.0 0.0 0.0 0.0 0.0 0.8 0.
* Ioniz. model Reactive chem. Molec. Clusters Pc
* Flag Eshift Sshift Rbond Ebind IntDOF flag exp
ANEOS6 1 0. 0. 0.E-8 3.2 0. 1. 1.8
* Zi fi
ANEOS7 1 2./3.
ANEOS8 8 1./3.
    
```

Figure B.1 shows the density  $\rho$ , specific entropy  $s$ , adiabatic gradient  $\nabla_{\text{ad}}$ , and specific internal energy  $u$  as a function of pressure and temperature. A very large range in  $P-T$  is covered, ranging from the low values found in planetary atmospheres ( $\sim 100$  K, milibars) to extremely high values exceeding those expected in the center of a 5 Gyr old 20 Jupiter mass planet (about  $3 \times 10^5$  K,  $2 \times 10^{16}$  Barye, [Mordasini et al. 2012b](#)). In the panel showing the density, one sees the various phases (solid, liquid, gaseous) as in a normal water phase diagram. It should be noted that ANEOS predicts that the density slightly decreases when passing at a pressure of 1 bar ( $10^6$  Barye) from the solid into the liquid state clearly, it cannot capture the special behavior (anomaly) of  $\text{H}_2\text{O}$  where the density of liquid water is higher than the one of ice. At the phase transition, we also see a strong increase of the entropy when passing from the liquid into the gaseous phase, as expected. In the density panel, one can also see the diagonal structure of the iso-density contours in the lower right part of the  $P-T$  domain. In this part, the water vapor (and plasma, after ionization) behaves like an ideal gas, meaning that  $p \propto T$  on iso-density contours. At very high pressures, we see that the density becomes in contrast independent of temperature, leading to horizontal iso-density lines. This is the expected behavior for completely degenerate matter. We have found that ANEOS fails to provide physical values for the adiabatic gradient at very high pressures and low temperatures, such that we have limited  $\nabla_{\text{ad}}$  in this regime to 0.5, as visible in the top left corner of the  $\nabla_{\text{ad}}$  panel. Fortunately, this regime is not occurring in planetary interiors at the current age of the Universe.



**Fig. B.1.** Thermodynamical properties of  $\text{H}_2\text{O}$  as predicted by ANEOS for a very large range in pressure  $P$  and temperature  $T$ . The density (*top left*), specific entropy (*top right*), adiabatic gradient (*bottom left*), and specific internal energy (*bottom right*) are shown. One sees the effects of phase changes and of the asymptotic behavior under extreme conditions (ideal gas limit at high  $T$  and low  $P$ ; fully degenerate limit at very high  $P$  and low  $T$ ).