

# The New Generation Planetary Population Synthesis (NGPPS)

## II. Planetary population of solar-like stars and overview of statistical results<sup>\*</sup>

Alexandre Emsenhuber<sup>1,2,3</sup>, Christoph Mordasini<sup>2</sup>, Remo Burn<sup>2,4</sup>, Yann Alibert<sup>2</sup>, Willy Benz<sup>2</sup>, and Erik Asphaug<sup>1</sup>

<sup>1</sup> Lunar and Planetary Laboratory, University of Arizona, 1629 E. University Blvd., Tucson, AZ 85721, USA

<sup>2</sup> Physikalisches Institut, Universität Bern, Gesellschaftsstrasse 6, 3012 Bern, Switzerland

<sup>3</sup> Universitäts-Sternwarte München, Ludwig-Maximilians-Universität München, Scheinerstraße 1, 81679 München, Germany  
e-mail: emsenhuber@usm.lmu.de

<sup>4</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

Received 7 July 2020 / Accepted 5 October 2021

### ABSTRACT

**Context.** Planetary formation and evolution is a combination of multiple interlinked processes. Constraining the mechanisms observationally requires statistical comparison to a large diversity of planetary systems.

**Aims.** We want to understand global observable consequences of different physical processes (accretion, migration, and interactions) and initial properties (like disc masses and metallicities) on the demographics of the planetary population. We also want to study the convergence of our scheme with respect to one initial condition, the initial number of planetary embryo in each disc.

**Methods.** We select distributions of initial conditions that are representative of known protoplanetary discs. Then, we use the Generation III Bern model to perform planetary population synthesis. We synthesise five populations with each a different initial number of Moon-mass embryos per disc: 1, 10, 20, 50, and 100. The last is our nominal population consisting of 1000 stars (systems) that we use for an extensive statistical analysis of planetary systems around  $1 M_{\odot}$  stars.

**Results.** The properties of giant planets do not change much as long as there are at least 10 embryos in each system. The study of giants can thus be done with simulations requiring less computational resources. For inner terrestrial planets, only the 100-embryos population is able to attain the giant-impact stage. In that population, each planetary system contains on average 8 planets more massive than  $1 M_{\oplus}$ . The fraction of systems with giant planets at all orbital distances is 18%, but only 1.6% are at  $>10$  au. Systems with giants contain on average 1.6 such planets. The planetary mass function varies as  $M^{-2}$  between 5 and  $50 M_{\oplus}$ . Both at lower and higher masses, it follows approximately  $M^{-1}$ . The frequency of terrestrial and super-Earth planets peaks at a stellar  $[\text{Fe}/\text{H}]$  of  $-0.2$  and  $0.0$ , respectively, being limited at lower  $[\text{Fe}/\text{H}]$  by a lack of building blocks, and by (for them) detrimental growth of more massive dynamically active planets at higher  $[\text{Fe}/\text{H}]$ . The frequency of more massive planets (Neptunian, giants) increases monotonically with  $[\text{Fe}/\text{H}]$ . The fast migration of planets in the  $5\text{--}50 M_{\oplus}$  range is reduced by the presence of multiple lower-mass inner planets in the multi-embryos populations. To assess the impact of parameters and model assumptions, we also study two non-nominal populations: in-situ formation without gas-driven migration, and a different initial planetesimal surface density.

**Conclusions.** We present one of the most comprehensive simulations of (exo)planetary system formation and evolution to date. For observations, the syntheses provides a large data set to search for comparison synthetic planetary systems that show how these systems have come into existence. The systems, including their full formation and evolution tracks are available online. For theory, they provide the framework to observationally test the global statistical consequences of theoretical models for specific physical processes. This is a important ingredient towards the development of a standard model of planetary formation and evolution.

**Key words.** Planets and satellites: formation — Planetary systems — Planet-disk interactions — Protoplanetary disks

### 1. Introduction

Exoplanets are common. Results from the *Kepler* survey show that, on average, there are more exoplanets than stars, at least in the galactic environment probed by the Kepler satellite (e.g., Mulders et al. 2018; Zhu & Dong 2021). The number of discovered exoplanets, principally through large surveys, either radial velocity (RV), such as HARPS (Mayor et al. 2011) and Keck & Lick (Fulton et al. 2021), or transit surveys, as *CoRoT* (Moutou et al. 2013) or *Kepler* (Borucki et al. 2010; Thompson et al. 2018), permits to constrain properties of exoplanetary systems, about their mass, radii, distances, eccentricities, spacing and mutual inclinations (e.g. Winn & Fabrycky 2015). In addition,

various correlations with stellar properties have also been determined (Santos et al. 2003; Mayor et al. 2011; Petigura et al. 2018).

Yet, understanding how the formation and evolution of these planets work remains a challenge. Observations of the progenitors (circumstellar discs) are plentiful, but only few forming planets are known, such as PDS 70b (Keppler et al. 2018; Müller et al. 2018). Reliance on theoretical modelling for the formation stage is then necessary. A model that reproduces the final systems accounting for the initial state can provide valuable information about how planetary systems form and evolve.

For the constrains on planets, we can divide them in three main categories: 1) the characteristics of the planets themselves, for example their mass, radii, distances and eccentricities, 2) the properties of planetary systems and their diversity in terms of architecture, such as their multiplicity, mutual spacing and cor-

<sup>\*</sup> The data supporting these findings is available online at <http://dace.unige.ch> under section “Formation & evolution”.

relations between occurrences of different planet types, and 3) the correlations between the previous items and stellar properties, such as its metallicity.

Giant planets within 5–10 au around FGK stars have a frequency of 10–20% (Cumming et al. 2008; Johnson et al. 2010; Mayor et al. 2011). Earlier works based on radial velocity surveys found that giants have increasing probability of occurrence in  $\log(P)$  between 2 and 2000 days (Cumming et al. 2008), with an excess of hot-Jupiters, as they occur on 0.5–1% of Sun-like star (Howard et al. 2010; Mayor et al. 2011; Wright et al. 2012). More recent results based on the Kepler satellite survey also find an increase with distance (Dong & Zhu 2013; Santerne et al. 2016). There could be a peak at intermediate distances, possibly near the snow-line (Fernandes et al. 2019; Fulton et al. 2021). Then there is a decrease in the occurrence rate with distance, where the onset of the reduction could be already at 3–10 au (Bryan et al. 2016; Nielsen et al. 2019) and a  $\approx 1\%$  occurrence rate for detectable distant (tens to hundreds of AUs), massive planets (Bowler 2016; Galicher et al. 2016; Vigan et al. 2021).

System-level statistics provide additional information about properties in a system versus the whole population level. Diversity within each system compared to the whole population is a good example. For instance planets in small-mass systems have similar masses (Millholland et al. 2017), sizes and spacing (Weiss et al. 2018). Planet multiplicity tend to decrease for systems that host more massive planets (Latham et al. 2011). For giant planets, hot Jupiters do not usually have nearby companions (Steffen et al. 2012), but roughly half of them have more distant ones (Knutson et al. 2014). Conversely, distant giants also have a multiplicity rate of roughly 50% (Bryan et al. 2016; Wagner et al. 2019). There are also correlations between Super Earths and Jupiter analogs (Bryan et al. 2019) and between Super Earths and cold giants (Zhu & Wu 2018). This will be the subject of a companion work (Schlecker et al. *in press*). In addition, Bryan et al. (2016) observed that planets in multiple systems have on average a higher eccentricity than single giant planets; a different result from previous studies that found that planets in multiple systems had on average lower eccentricities (Howard 2013; Limbach & Turner 2015).

Correlations between stellar and planetary properties provide important information on the formation mechanism. Protoplanetary discs properties, especially their heavy-elements content, is linked to the host star’s metallicity (Gáspár et al. 2016), as they form from the same molecular cloud. Giant planets are preferentially found around metal-rich stars (e.g. Gonzalez 1997; Santos et al. 2004; Fischer & Valenti 2005; Adibekyan 2019). For low-mass planets, such a correlation still exists although it is weaker (Sousa et al. 2008, 2011; Buchhave et al. 2012, 2014; Wang & Fischer 2015; Petigura et al. 2018).

Further, we now have correlations between architecture and metallicity, with compact multi-planetary systems being more common on metal-poor stars (Brewer et al. 2018) while systems around metal-rich stars are more diverse (Petigura et al. 2018). Also, the eccentricities of giant planets around metal-rich stars tend to be higher than the one around metal-poor stars (Buchhave et al. 2018).

From the survey of star forming regions, we can determine the distribution of some characteristics of protoplanetary disc. The percentage of stars with a disc decreases with age in an exponential fashion with a characteristic time of a few Myr (Mamajek 2009; Fedele et al. 2010). Correlations were also found between disc masses and sizes (Andrews et al. 2010, 2018; Tripathi et al. 2017; Hendl et al. 2020), stellar masses (Andrews et al. 2013; Ansdell et al. 2016; Pascucci et al. 2016) and accre-

tion rate onto the star (Manara et al. 2016b, 2019; Mulders et al. 2017).

With these observations, it is possible to retrieve the characteristics at early stages of disc evolution (Tychoniec et al. 2018; Tobin et al. 2020), which are relevant for the initial conditions, and constraints on the transport mechanism in effect (Mulders et al. 2017).

To link protoplanetary discs to final systems, we need to use a formation model. Several approaches can be used: the study of individual Rosetta Stone systems, statistical studies on the population level as in the case here, or also studies of disc chemistry imprints for formation (e.g. Öberg et al. 2011; Mordasini et al. 2016). However, the constraints derived from observation for a single exoplanetary system compared to the model parameters does not permit to fully understand planetary formation at the individual system level. In addition, the diverse outcomes of  $N$ -body simulations (e.g. Hansen & Murray 2012, 2013) renders the task even more difficult. Working at the population level, with planetary population synthesis (Ida & Lin 2004a; Mordasini et al. 2009a,b, 2012a) is a much more powerful tool to understand planetary formation in general. This allows to determine how the different mechanisms that occur during planetary systems formation of interact.

Modelling planetary formation is a complex task, as many physical effects occur concurrently: growth of micron-size dust to planetary-sized bodies, the accretion of gas, orbital migration and dynamical interactions for multi-planetary systems. In Emshuber et al. (*in press*, hereafter Paper I), we present an update of the *Bern* model of planetary formation and evolution. This is a global end-to-end model, i.e. it includes the relevant processes that occur from the initial accretion of the protoplanets starting at the planetesimal-embryo stage up to their long-term evolution, trying to address as many relevant physical processes as possible. By using an approach that is rich in physics, but low-dimensional numerically to keep the computational cost acceptable, this model can be used to compute synthetic planet populations. Our formation model is based on the core accretion paradigm with planetesimals. The early phases of the evolution of the solids from dust to pebbles to planetesimals to embryos and pebble accretion (e.g. Ormel & Klahr 2010) are currently not included, but will be taken into account in future work based on Voelkel et al. (2020).

Theoretical models that are able to reproduce the characteristic of the observed exoplanets can be used to make predictions about the real population, which is helpful when designing future observations and instruments. For discovered planets, they can be used to propose a pathway for their formation (Armstrong et al. 2020), or point to other formation mechanisms if they cannot be reproduced at all (Morales et al. 2019).

In this work, we apply the Generation III *Bern* model of planetary formation and evolution described in Paper I to obtain synthetic populations of planetary systems. We provide the methods that we use to perform population synthesis, which are an update from Mordasini et al. (2009a, hereafter M09a).

We then present five synthetic planet populations for solar-like stars where we vary the initial number of embryos per system, which represent the oligarchs at the end of the planetesimals runaway growth. They act similarly to the large bodies in  $N$ -body studies, such as O’Brien et al. (2006) or Raymond et al. (2009). As we do not model their formation in our work, we treat their number as a free parameter. The goal is to test the convergence of our model with respect to this parameter. The populations with a larger number of embryos are capable to follow the formation of terrestrial planets (Paper I) but they are expensive to compute.

On the other hand, the populations with a lower number of embryos are much cheaper to compute (with the extreme case of a single embryo per system), but fails to follow properly terrestrial planets. This test will be useful for future works in this series about the effects of the parameters of the model or physical processes, which requires the computation of multiple populations.

## 2. Formation and evolution model

The model is described in [Paper I](#); so we will give here only a brief summary. In our coupled formation and evolution model, we first model the planets' main formation phase for a fixed time interval (set to 20 Myr) during which planets accrete solids and gas, migrate, and interact via the N-body. Afterwards, in the evolutionary phase, we follow the thermodynamical evolution of each planet individually to 10 Gyr.

The formation model derives from the work of [Alibert et al. \(2004, 2005\)](#). It follows the evolution of a viscous accretion disc ([Lüst 1952](#); [Lynden-Bell & Pringle 1974](#)). The turbulent viscosity is provided by the standard  $\alpha$  parameter ([Shakura & Sunyaev 1973](#)). Solids are represented by planetesimals, whose dynamical state is given by the drag from the gas and the stirring from the other planetesimals and the growing protoplanets ([Rafikov 2004](#); [Chambers 2006](#); [Fortier et al. 2013](#)). This disc provides gas and solids from which the protoplanets can accrete while also affecting the bodies that are inside it, by gas-driven planetary migrations.

The formation of the protoplanets is based on the core accretion paradigm ([Perri & Cameron 1974](#); [Mizuno 1980](#)), assuming planetesimal accretion in the oligarchic regime ([Ida & Makino 1993](#); [Ohtsuki et al. 2002](#); [Inaba & Ikoma 2003](#); [Chambers 2006](#); [Fortier et al. 2013](#)). Gas accretion is initially governed by the ability of the planet to radiate away the potential energy ([Pollack et al. 1996](#); [Lee & Chiang 2015](#)), and so the envelope mass is determined by solving the internal structure equations ([Bodenheimer & Pollack 1986](#)). Once the planet is massive enough (on the order of  $10 M_{\oplus}$ ), cooling becomes efficient, and runaway gas accretion can occur. In that situation, the envelope is no longer in equilibrium with the surrounding gas disc and contracts ([Bodenheimer et al. 2000](#)) while gas accretion is limited by the supply of the gas disc.

Multiple embryos can form concurrently in each system, and the gravitational interactions are modelled using the mercury *N*-body package ([Chambers 1999](#)).

Once the formation stage is finished, the model transitions to the evolutionary phase, where planets are followed individually to 10 Gyr. The planetary evolution model is based on [Mordasini et al. \(2012c\)](#) and includes atmospheric escape ([Jin et al. 2014](#)) and migration due to tides raised on the star ([Benítez-Llambay et al. 2011](#)).

## 3. Population synthesis

To perform a population synthesis of planetary systems, we use a Monte Carlo approach for the initial conditions of the discs, in a similar fashion that has been performed in [M09a](#) and [Mordasini et al. \(2012b\)](#). The Monte Carlo variables are selected as:

- The initial mass of the gas disc  $M_g$
- The external photo-evaporation rate  $\dot{M}_{\text{wind}}$
- The dust-to-gas ratio  $f_{D/G} = M_s/M_g$
- The inner edge of the gas disc  $r_{\text{in}}$
- The initial location of the embryos

**Table 1.** Fixed parameters for the formation and evolution model.

Quantity	Value
Stellar mass $M_{\star}$	$1 M_{\odot}$
Disc viscosity parameter $\alpha$	$2 \times 10^{-3}$
Power law index of the gas disc $\beta_g$	0.9
Power law index of the solids disc $\beta_s$	1.5
Characteristic radius of the solids disc $r_{\text{cut,s}}$	$r_{\text{cut,g}}/2$
Planetesimal radius	300 m
Planetesimal density (rocky)	$3.2 \text{ g cm}^{-3}$
Planetesimal density (icy)	$1 \text{ g cm}^{-3}$
Embryo mass $M_{\text{emb},0}$	$1 \times 10^{-2} M_{\oplus}$
Opacity reduction factor $f_{\text{opa}}$	$3 \times 10^{-3}$

Here,  $M_s$  is the initial mass of solids in a disc. The other fixed parameters used in this study are provided in Tab. 1. These are taken to remain the same in all systems.

In the rest of this section, we discuss each Monte Carlo variable and their distributions, as well as the related fixed parameters. The significant number of parameters in global end-to-end models like the one used here is a notoriously difficult aspect of this approach. The issue naturally results from global models combining many sub-models, each coming with its own model parameters. Some of these parameters are at least to some extent constraint by observations, while others are based on theoretical considerations only, and some are merely educated guesses. When interpreting the results presented in this work, like for example the key demographic predictions of planet occurrence rates or the general shape of the planet mass-distance diagrams, it is important to keep in mind that these results are clearly functions of the chosen parameters and base assumption underlying the formation model. Thus, these results always have to be seen as the predictions made in the context of the current model and for the chosen (nominal) parameter values, and that large systematic uncertainties exist.

Ideally, one would quantitatively assess the impact of all these parameters by running numerous syntheses where the values of the parameters, as well as important underlying model assumptions, are varied systematically. This would give an understanding of the systematic uncertainties in the model predictions. In practice, this is not easily feasible, because the computational cost of the multi-embryo syntheses is very significant ( $\sim 1$  M CPU h), especially for a high number of initial embryos per disc. To still elucidate the impact of parameters at least for two of them (besides the initial number of embryos per disc), we present in [Appendix A](#) two non-nominal populations: one, where the initial solid surface density of planetesimals has a different slope, and one where gas-driven orbital migration is neglected. In the appendix, we study how this changes the mass-distance diagram, and key demographic properties.

### 3.1. Gas disc mass

It is very difficult to observe directly  $\text{H}_2$  in protoplanetary discs, and so the most reliable method to determine disc masses remains the measurement of the continuum emission of the dust. To recover the gas mass, a dust-to-gas ratio similar to the interstellar medium is applied ([Beckwith & Sargent 1996](#); [Andrews & Williams 2005](#); [Andrews et al. 2010](#)).

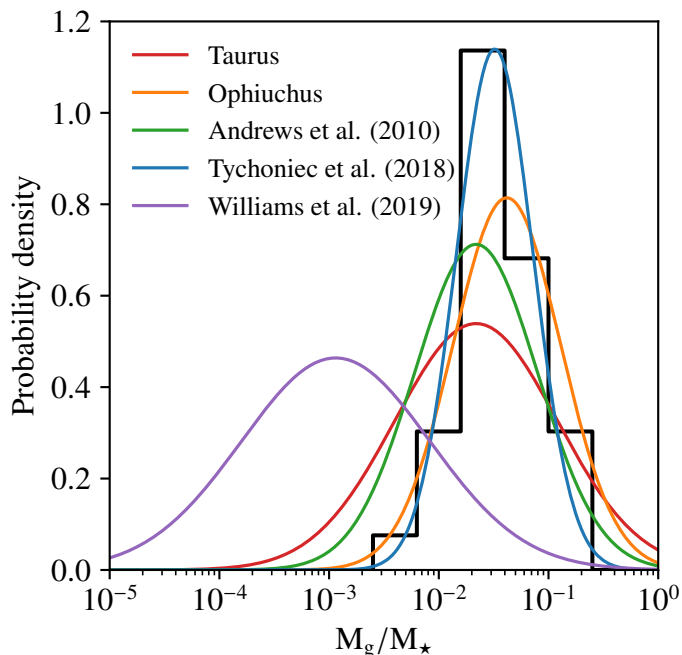
Several observational data for protoplanetary disc masses are reported in [Table 2](#) and plotted in [Fig. 1](#). The first two values, for the fits on the distributions of Taurus and Ophiuchus star-forming regions were obtained by [M09a](#) by fitting log-normal



**Table 2.** Mean and standard deviation of the normal distribution of the disc mass for different observational sample.

Source	$\mu$	$\sigma$
Fit to Taurus <sup>a</sup>	-1.66	0.74
Fit to Ophiuchus <sup>a</sup>	-1.38	0.49
Andrews et al. (2010)	-1.66	0.56
Fit to class I from Tychoniec et al. (2018)	-1.49	0.35
Class I from Williams et al. (2019)	-2.94	0.86

<sup>a</sup> Fit to the values obtained by Beckwith & Sargent (1996) performed by M09a.



**Fig. 1.** Probability density functions for the different distributions given in Table 2. In addition, we show the histogram of Class I discs from Fig. 12 of Tychoniec et al. (2018) in black. All the curves are normalised so that the surface below them is unity.

distributions on the results of Beckwith & Sargent (1996). The third value was directly given in Andrews et al. (2010), while for the fourth one, we applied the same procedure as for the first two, but using the histogram of Class I disc masses reported in Fig. 12 of Tychoniec et al. (2018). Finally, we provide ALMA data of Class I discs in the Ophiuchus star-forming region from Williams et al. (2019). The latter was converted to gas masses using a gas-to-dust mass ratio of 100:1 (as in Tychoniec et al. 2018).

There is more than one order of magnitude difference between the results from ALMA for the Ophiuchus star-forming region (Williams et al. 2019) and others, such as those obtained with the VLA for Perseus (Tychoniec et al. 2018). These differences are discussed in Tychoniec et al. (2020), where the authors argue that 1) their median masses from VLA are more complete and 2) Class 0/I objects are more likely to be representative of the discs at early stage of planetary formation. The second point is related to our modelling, as the model used in this work begins once the protoplanetary disc is formed and dust has grown into planetesimals. Class I discs are hence the most relevant for our study. Thus, the work of Tychoniec et al. (2018) is then the best suitable for our initial conditions, and this is the one we select. To avoid extreme values, we only allow disc masses between

$4 \times 10^{-3}$  and  $0.16 M_{\star}$ . With this upper mass limit, the discs are always self-gravitationally stable.

Compared to the populations obtained with earlier versions of the model, our disc masses are smaller than the ones from M09a, which used the parameters derived from fitting the values in the Ophiuchus star-forming region from Beckwith & Sargent (1996). It should be noted that unlike M09a, we model the entire disc and not only the innermost 30 au, so we do not need to scale the disc masses to obtain only the innermost region. However, the distribution we adopted has a higher mean than what was obtained by Andrews et al. (2010); so we have overall larger disc masses than in the works of Mordasini et al. (2012c,b), A13, Fortier et al. (2013), and Thiabaud et al. (2014, 2015) also used the results from Andrews et al. (2010), albeit in a different fashion, where initial masses were bootstrapped from the specific values of the observed discs.

### 3.2. Initial gas surface density: Spatial distribution

With spatially resolved discs it is possible to estimate the distribution of the material with respect to the distance from the star. The surface density typically goes with  $r^{-1}$  until a characteristic radius where it relates more to an exponential decrease (Hughes et al. 2008; Andrews et al. 2009, 2010). While in principle both the index of the power law and the characteristic radius would require their own distributions, we decided against adding more parameters for the initial conditions of our populations.

The power law index is fixed to  $\beta_g = 0.9$ , which is consistent with the results from Andrews et al. (2010). For the characteristic radius  $r_{\text{cut,g}}$  as a function of disc mass, we use the following relationship, which is taken from Fig. 10 of Andrews et al. (2010),

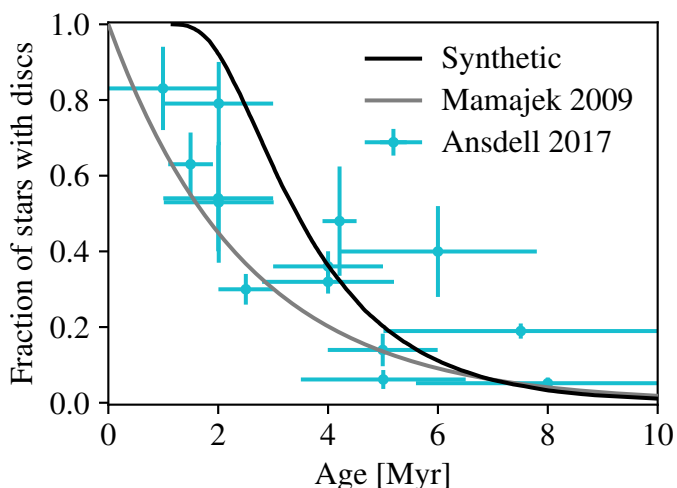
$$\frac{M_g}{2 \times 10^{-3} M_{\odot}} = \left( \frac{r_{\text{cut,g}}}{10 \text{ au}} \right)^{1.6}. \quad (1)$$

The relationship is somewhat different than the  $M_g \propto r_{\text{cut,g}}^2$  found in more recent work (Tripathi et al. 2017; Andrews et al. 2018). Further, the latter is however not universal across different stellar forming regions with various ages (Hendler et al. 2020). The results of that work also suggest that the relationship becomes shallower with age, and the power-law index we use is similar to the youngest stellar forming regions, and thus more appropriate as an initial condition.

A complication arises from the fact that the observational relation of Eq. 1 was derived for dust disc radii of Class II discs, and not for gas discs at early times (e.g., after the end of infall and potential gravitational instabilities). On one hand, this could mean that our approach leads to too small initial gas disc radii in our synthetic disc population given the effect of inward drift of dust (Ansdell et al. 2018). On the other hand, the discs observed by Andrews et al. (2010) were specifically selected for good observability with SMA and span the upper half of the millimeter continuum luminosity distribution only. This could mean that the disc radii are on the large side compared to a more representative sample. These effects could partially cancel each other. To elucidate this, we present in Sect. 3.8 a comparison of our theoretical disc gas radii with the dust radii of younger discs, as found with the more recent VANDAM survey (Tobin et al. 2020).

### 3.3. External photo-evaporation rate

The photo-evaporation rate  $\dot{M}_{\text{wind}}$  and the viscosity parameter  $\alpha$  are the main parameters that determine the life time of the gas



**Fig. 2.** Fractions of stars with a protoplanetary disc as function of their age. The black line shows our results, while the blue line follow the exponential decay with a time scale of 2.5 Myr from Mamajek (2009). The purple points are from Ansdell (2017).

discs. This is a degenerate problem, as increasing either  $\alpha$  or  $\dot{M}_{\text{wind}}$  leads to shorter disc life times. However,  $\alpha$  also constrains the mass that is accreted onto the star, which we can use to lift the degeneracy. Our aim is then to find combinations of  $\alpha$  and  $\dot{M}_{\text{wind}}$  that provide accretion rates onto the star and disc life times that are in agreement with observations.

Mulders et al. (2017) combined the ALMA observations of the disc mass  $M_{\text{disc}}$  from Pascucci et al. (2016) and the X-Shooter accretion rate onto the star  $\dot{M}_{\text{acc}}$  from Manara et al. (2016a, 2017) for the Chamaeleon I star-forming region and ALMA from Ansdell et al. (2016) and X-Shooter from Alcalá et al. (2014, 2017) for the Lupus region. The  $M_{\text{disc}}-\dot{M}_{\text{acc}}$  relation obtained by the combination of the two region is shallower than linear, indicating that another effect than viscous dissipation is potentially at play. Nevertheless, they obtained that for  $\alpha$  values between  $10^{-3}$  and  $10^{-2}$ , it is possible to find relations that are comparable with observation.

Manara et al. (2019) compared the  $M_{\text{disc}}-\dot{M}_{\text{acc}}$  relation predicted in a population synthesis obtained with an earlier version of the formation used in this work to an extended sample relative to Mulders et al. (2017). The synthetic disc population for a constant  $\alpha$  fails to reproduce the whole scatter observed in the actual  $M_{\text{disc}}-\dot{M}_{\text{acc}}$  relationship. Nevertheless, the synthetic population of discs is able to retrieve the observed correlation of  $M_{\text{disc}}$  and  $\dot{M}_{\text{acc}}$ . Thus, to avoid introducing one more Monte Carlo variable in our population synthesis scheme, we will stick to a single  $\alpha$  value for all discs. We selected a value of  $\alpha = 2 \times 10^{-3}$ , which is the same as the comparison shown in Manara et al. (2019). This leaves only the value of the external photo-evaporate to determine the life times of the discs.

Protoplanetary discs have a lifetime in the 3–7 Myr range (Haisch et al. 2001; Fedele et al. 2010; Richert et al. 2018). Fitting the results with an exponential law gives time constants of 2.5 Myr (Mamajek 2009) or 2.7 Myr (Ansdell 2017).

Given the fixed  $\alpha = 2 \times 10^{-3}$  and the fixed distribution of initial disc masses described above, we determine an empirical distribution of external photoevaporation rates that leads to a distribution of the lifetimes of the synthetic discs that is in agreement with the observed distribution of disc lifetimes.

In this way, we find a log-normal distribution with parameters  $\log_{10}(\mu/(M_{\odot}/\text{yr})) = -6$  and  $\sigma = 0.5$  dex. Note that these

**Table 3.** Mean and standard deviation of the normal distribution of [Fe/H] for different observational sample.

Source	$\mu$	$\sigma$
Santos et al. (2005)	-0.02	0.22
Petigura et al. (2017)	0.03	0.18

rates would give the actual photoevaporation rates if the modelled discs would have a size of 1000 au (Paper I). In reality, their outer radius are smaller ( $\sim 100$  au) and given dynamically by the equilibrium of viscous spreading that acts to increase the outer radii and external photoevaporation which reduces the radii.

The selection of those values was made so that we have a cluster of disc life times at about 3 Myr. We show in Fig. 2 the corresponding life times obtained using our model for the disc masses,  $\alpha$  and  $\dot{M}_{\text{wind}}$  that we selected. While we miss the short-lived discs (less than 1 Myr), our distribution is more able to reproduce some longer-lived clusters in the range of 4 Myr to 6 Myr.

### 3.4. Dust-to-gas-ratio

The initial mass of the solids disc is linked to that of the gas disc by a factor  $f_{\text{D/G}}$ . To determine the distribution of this parameter, we assume that stellar and disc metallicities are identical. Hence we have the relation (Murray et al. 2001)

$$\frac{f_{\text{D/G}}}{f_{\text{D/G},\odot}} = 10^{[\text{Fe}/\text{H}]}. \quad (2)$$

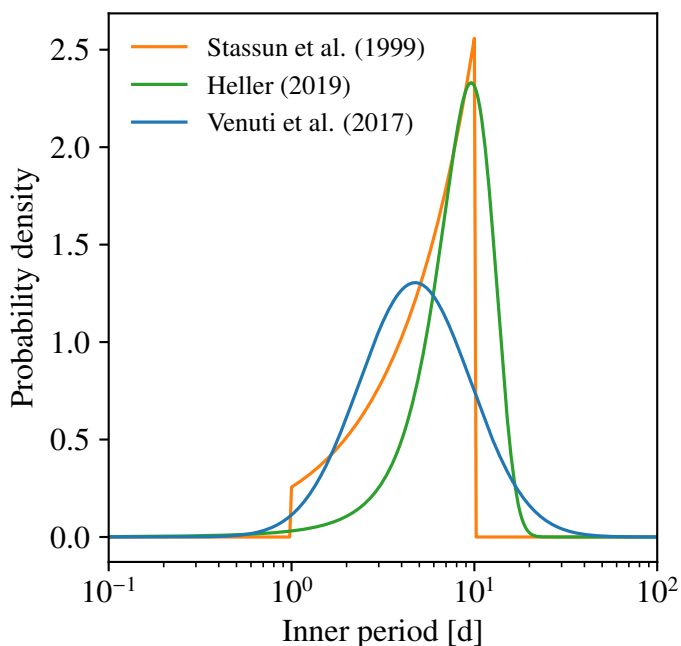
Furthermore, we now assume that the dust-to-gas of the Sun,  $f_{\text{D/G},\odot} = 0.0149$  (Lodders 2003). It should be noted that this value is quite lower than in the first generation of our planetary population syntheses (Mordasini et al. 2009a, 2012b), where it was taken to be a factor roughly three times greater.

There are multiple possibilities for the distribution of the parameter; as stellar metallicities vary among different regions in the galaxy. The choice depends on the kind of observational survey we aim to compare to. RV surveys will favour stars in the neighbourhood of the Sun, while transit and in particular microlensing surveys can reach greater distances. For instance, the Kepler survey targets stars only in one specific direction towards Cygnus and Lyra. We provide the parameters of a normal distribution from different sources in Table 3. The two distribution are similar, with the difference in their mean corresponding to 12%. The selection of either distribution should therefore not affect significantly our results. For the population syntheses presented below, we use the distribution from Santos et al. (2005) for the Coralie RV search sample.

One thing to mention is that the normal distribution is unbound on both sides. Hence to avoid modelling system that have metallicities not occurring in the solar neighbourhood given galactic chemical evolution, we restrict the selection of the parameter to the  $-0.6 < [\text{Fe}/\text{H}] < 0.5$  range.

### 3.5. Inner edge

The position of the inner edge of the gas disc plays an important role for the final location of the close-in planets. For planets that form and then migrate inward, migration will stall when the planet reaches a location where gas is no longer present. If planets rather form in-situ, then the inner edge is also linked to where planets are able to accrete.



**Fig. 3.** Probability density functions for the different distributions of inner radius as given in the text. All the curves are normalised so that the surface below them is unity.

There are various possible ways to determine the inner edge of protoplanetary discs, for example 1) determining stellar rotation rate and assuming the disc is truncated at the corotation radius, 2) from the continuum near infra-red (NIR) emissions, and 3) from emission lines.

We chose to use the corotation radius to determine the inner edges of protoplanetary discs. Apart from the good agreement to observations, the main reason of this choice over a prescription for the magnetospheric truncation radius is that the magnetic field strengths of young stars are not very well constrained. Heller (2019) recently investigated planet formation scenarios using either inner edges at the corotation or at a prescribed magnetospheric truncation radius. In any case, the two radii lie very close to each other: Heller (2019) used a magnetic truncation period of 4 days motivated by works of Romanova & Lovelace (2006) and Kuchner & Lecar (2002).

The continuum NIR mainly originates from the hot dust and not from dust-depleted gas. However, as indicated by observations from Eisner et al. (2005, 2009) and Isella et al. (2008) and in detail modeled by Flock et al. (2017), the temperatures at the inner edge of the disc are larger than the evaporation temperature of silicate dust grains. Therefore, the gas extends closer to the star than the silicate evaporation line. Thus, NIR might not be able to trace the inner edge of the gas discs. This is the likely reason put forward by Eisner et al. (2005) who found consistently larger radii by NIR interferometry than the corotation and the magnetospheric truncation radii.

As for emission lines, they are able to trace the gas disc. Carr (2007) found a factor of 0.7 smaller disc radii (using the CO  $v = 1 - 0$  transition near  $4.7 \mu\text{m}$ ) than the corotation radius. This is a reasonable agreement given the scatter of the distribution. The largest dataset of this sort by Eisner et al. (2009, 2010) consists of 15 discs around stars of various masses (including 7 T tauri stars). This is a low number of observation, from which it is difficult to extract a full distribution. It can nevertheless be noted that the values are in good agreement with magnetic corotation truncation discussed below.

**Table 4.** Adopted values for our calculation of the minimum-mass solar nebula (MMSN).

Items	Min. [ $M_{\oplus}$ ]	Max. [ $M_{\oplus}$ ]	Ref.
Terrestrial	2	2	
Jupiter	24	46	Wahl et al. (2017)
Saturn	16	23	Militzer et al. (2019)
Uranus	10	14	Podolak et al. (2000)
Neptune	14	16	Podolak et al. (2000)
Migration	–	50	Hahn & Malhotra (1999)
Total	66	151	

By using the corotation radius, the location of the inner edge can be derived from rotation rates of young stellar objects (YSOs). We show several distributions of those values in Fig. 3: a uniform distribution in the period between 1 and 10 d that is compatible with the results of Stassun et al. (1999), a normal distribution with parameters  $\mu = 8.3$  d and  $\sigma = 5$  d derived by Heller (2019) based on the work of Irwin et al. (2008), and a log-normal distribution with a mean  $\log_{10}(\mu/\text{d}) = 0.67617866$  and deviation  $\sigma = 0.3056733$  dex that is derived from the work of Venuti et al. (2017).

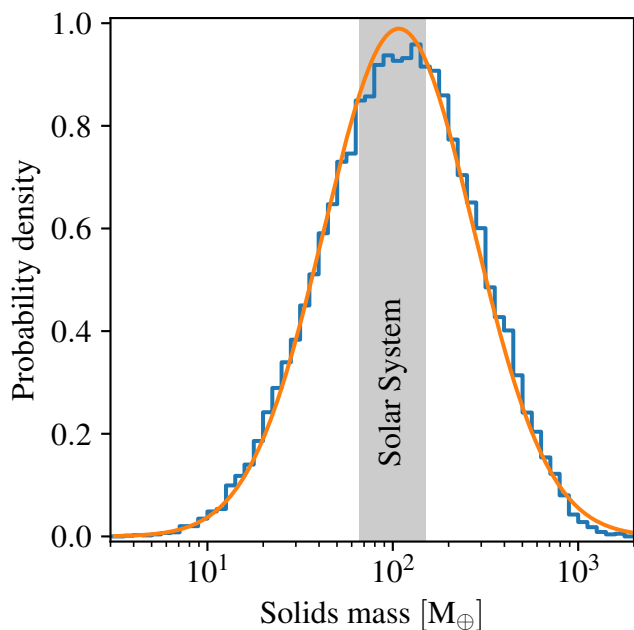
In the present work, we adopt the last one, based on Venuti et al. (2017). Here, the mean corresponds to a rotation period of 4.7 days or a distance of 0.055 au. To avoid that some discs have inner edges that are smaller than the initial stellar radius predicted by the stellar evolution model (Paper I), we truncate the distribution so that no inner edge can be within  $1.65 \times 10^{-2}$  au, which corresponds to a period of 0.77 days. We use the period as the main variable to obtain the inner radius as it is largely independent of the stellar mass at young ages (e.g. Henderson & Stassun 2012).

It should be noted that the means of all the distributions presented here are lower than what is obtained in other works, such as 10 days in Lee & Chiang (2017). The value of 10 days also correspond the peak of the location of the innermost planet as found by Kepler (Mulders et al. 2018).

Concerning the solids disc, we do not place planetesimals inside the iron evaporation line as given by our condensation model (Paper I). Therefore, if the iron evaporation line is further out than the inner edge of the gas disc, a disc has two edges: one for the gas and one for the planetesimals. The region inside the planetesimal inner edge will not contribute to solid accretion, but it can be an important region for orbital migration. However, it is found that the inner part is hot enough only for the most massive gas discs with small inner edges. In most cases, the temperature at the inner edge of the gas discs is less than about 1700 K, meaning that the inner edge of gas and solid disc coincide.

### 3.6. Planetesimal disc masses

The total mass in solids is not itself a Monte Carlo variable, but the product of the gas disc mass  $M_g$  with the dust-to-gas ratio  $f_{D/G}$ . However, it is one of the most important quantities that determines the types of planets that will be formed. Thus, it is still worth discussing. The distribution of the total mass in solids is shown in Fig. 4. The disc masses were computed using the disc model, in a similar fashion than for the disc life times (Sect. 3.3). As the distribution of solids mass is the product of two log-normal distributions (the gas disc mass and the dust-to-gas ratio), it also close to a log-normal distribution (because the two underlying distributions are truncated plus the reduction of



**Fig. 4.** Distribution of initial planetesimals disc masses. The blue curve is an histogram of the actual values while the yellow curve shows a log-normal fit to the data, whose mean (in log-space) is  $108 M_{\oplus}$  and a standard deviation of 0.40 dex. The gray area denotes the possible range of values for the minimum-mass solar nebula (MMSN).

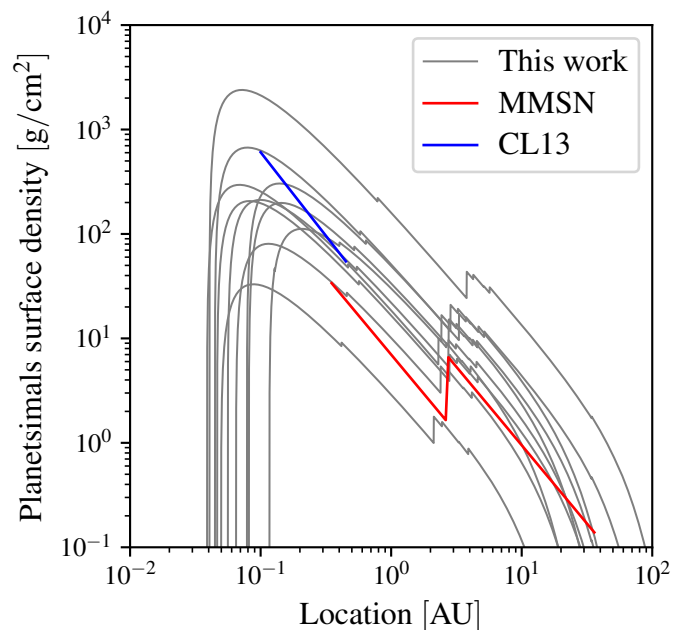
solids mass by volatiles being in the gas inside the corresponding ice lines). We therefore fitted a log-normal distribution, whose parameters are a mean (in log-space) of  $108 M_{\oplus}$  and a standard deviation of 0.40 dex.

To compare the obtained masses with the solar system, we overlay the distributions with a possible range of values for the minimum-mass solar nebula. The lower boundary was chosen according to the lowest estimates for the core masses of the giant planets, at  $66 M_{\oplus}$  while the upper boundary was calculated as  $101 M_{\oplus}$ , from the higher estimates, plus  $50 M_{\oplus}$  needed for the outward planetesimals-driven migration of Neptune (Fernandez & Ip 1984; Malhotra 1993; Hahn & Malhotra 1999). The calculation of these values is provided in Table 4.

### 3.7. Initial solids surface density: spatial distribution

To account for the inward drift of solids and the effect of planetesimal formation, we select an initial profile of the planetesimal disc that is different to that of the gas disc. The first difference is that the characteristic radius of the planetesimals disc is set to half that of the dust disc that was observed by Andrews et al. (2010). This follows the planetesimal formation simulations of Voelkel et al. (2020), who found that the planetesimals disc is smaller than that of the dust. In this work, we chose to still keep the relationship of Andrews et al. (2010) to provide the characteristic radius of the gas disc and use a smaller radius for the planetesimals disc. In future work, thanks to the addition of the dust-pebble-planetesimal growth phase in Voelkel et al. (2020), such approximations will no longer be necessary. The second difference is that the power law index is steeper for the solids disc ( $\beta_s = 1.5$ ) than that of the gas disc ( $\beta_g = 0.9$ ) following Lenz et al. (2019) and Voelkel et al. (2020).

To show how the difference in the characteristic radius affects the planetesimal surface density distribution in the discs, we provide in Fig. 5 a comparison between the synthetic discs of



**Fig. 5.** Initial planetesimals surface density profiles for 10 discs, which were selected using the quantiles of the disc mass distribution, to be representative of the entire population. The top and bottom grey lines thus show the most and least massive disc. The red line is the minimum-mass solar nebula (MMSN, Weidenschilling 1977; Hayashi 1981), while the blue line is the minimum-mass extrasolar nebula (Chiang & Laughlin 2013, CL13).

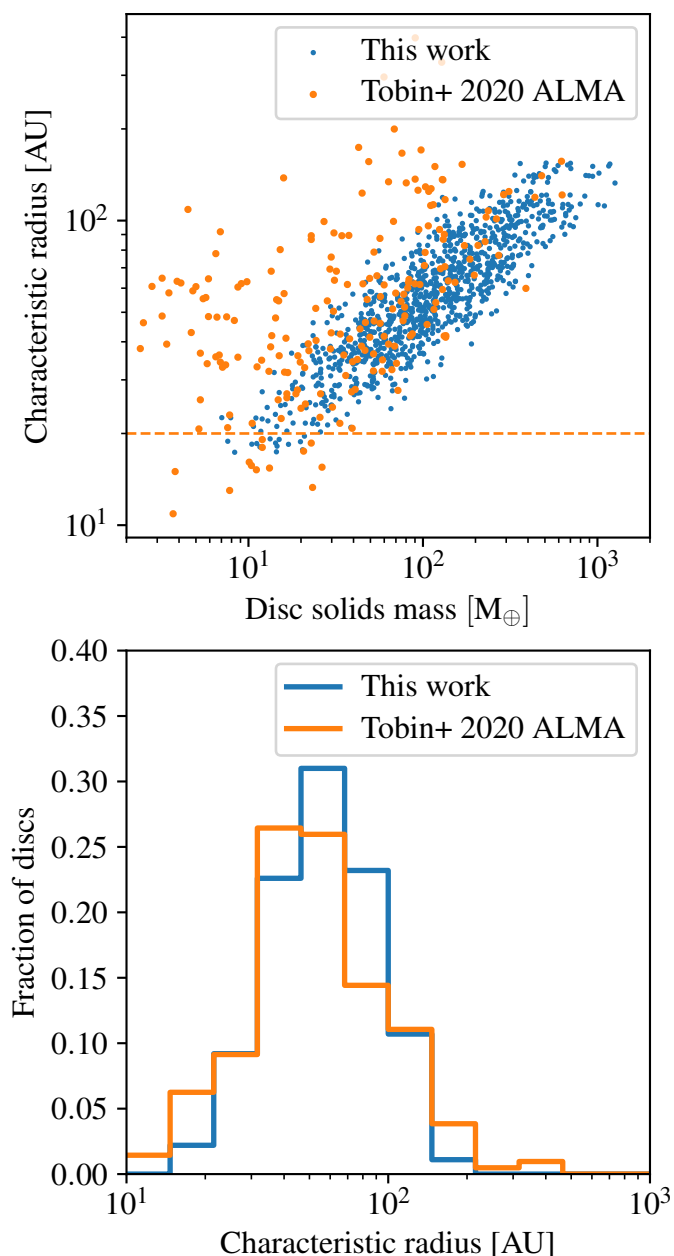
our populations, the minimum-mass solar nebula (MMSN; Weidenschilling 1977; Hayashi 1981) and the minimum-mass extrasolar nebula of (Chiang & Laughlin 2013). The ten discs were selected using the quantiles of the planetesimals disc mass distribution so that they are representative of the overall distribution of the discs in our populations. Outside the ice line, the median surface density is larger by a factor of roughly two compared with the MMSN. Due to the larger jump in surface density at the ice line in the MMSN compared to our populations, we find a larger difference inside the ice line. Nevertheless, our profiles are compatible or even smaller than the minimum-mass extrasolar nebula obtained by Chiang & Laughlin (2013). It is derived from close-in planets discovered by the Kepler satellite, assuming as for the MMSN in-situ formation.

Thus, despite the relatively small characteristic radii we selected for our planetesimal discs, which increase the surface density at given total mass, the planetesimals surface density are, on average, only larger than the MMSN by a factor of roughly 2 outside the ice line, and there are also discs with lower surface densities. The region outside of the ice line is a location of great importance for the formation of giant planets, as most of the cores of these planets are formed there. Regarding the regions close to the star, most synthetic discs have surface densities lower than the minimum-mass extrasolar nebula.

### 3.8. Comparison with Tobin et al. (2020)

As mentioned already in Sect. 3.2, the definition of initial conditions from disc observations is not trivial because of the differences of dust, planetesimal, and gas discs and different ages. The picture is further complicated because dust radii found in simulations versus those observed may substantially differ because





**Fig. 6.** Comparison of characteristic initial gas disc radii versus disc masses (top) and disc radii alone (bottom) between this work (in blue) and the observational results for Class 0/I/flat spectrum dust discs of non-multiple protostars using ALMA (Tobin et al. 2020, in orange). The dashed orange line represents half the typical spatial resolution of the survey.

observations (also) depend on sufficient opacity to detect matter (Rosotti et al. 2019).

With the properties of both the synthetic gas and planetesimal disks introduced, we here compare our approach with a more recent observational paper, Tobin et al. (2020). In their multi-wavelength VANDAM survey (ALMA and VLA), they observed several hundred protostellar discs. These younger discs should be more representative of initial conditions in which we are interested here than older Class II discs. The ages of the Class 0 and I/flat spectrum discs should roughly be 100 and 200 kyr, respectively (Tobin et al. 2020). At these early times, the chances are higher that evolution has not yet led to significant differences in the dust and gas radii. To which extent this holds is a function

of several parameters like the turbulence level or the strength of external photoevaporation, as indicated by simulation of dust evolution (Birnstiel et al. 2012; Voelkel et al. 2020). For the time being, we follow Tobin et al. (2020) and compare our initial gas disk radii with their observed dust radii of Class 0/I/flat spectrum disks of non-multiple protostars. We furthermore make the rough assumption that our initial planetesimal masses are representative of their observed dust masses. This obviously only holds if the planetesimal formation process is efficient. Some planetesimal formation models do produce such a high efficiency of dust-to-planetesimal conversion if the turbulence level in the discs is low (Lenz et al. 2019; Voelkel et al. 2020), whereas others rather find a  $\sim 10\%$  efficiency (Coleman 2021). In the absence of a description for the early phases of the growth from dust over pebbles to planetesimals in our current model, this is the comparison that can currently be made that at least does not involve additionally also converting dust masses into gas masses, which would add even more uncertainties.

The result is shown in Fig. 6. One sees that our disc radii overlap well with that of (Tobin et al. 2020), even though we are using the Class II relation of Andrews et al. (2010). We also note that our solid mass distribution does not extend to the lowest masses seen in VANDAM. At these low masses, many observed discs have still significant radii of about 40-50 au, while our relation would predict sizes of 10-20 AU. Here, it might be relevant that the spatial resolution of the survey was about 40 AU, meaning that it could be incomplete at the small sizes. However, we note that there is a discrepancy between our disc masses (which are based on the VLA measurements of Perseus by Tychoniec et al. 2018) and those of Tobin et al. (2020) for Orion. Tobin et al. (2020) discuss this difference and suggest that the opacity law used in previous studies needs to be revised. However, they can not rule out an underlying discrepancy between the two regions.

### 3.9. Embryos

The embryos are initialised in the following way: we place a pre-determined number of bodies of initial mass  $M_{\text{emb},0} = 10^{-2} M_{\oplus}$  with a uniform probability in the logarithm of the distance between  $r_{\text{in}}$  and 40 au. This spacing was selected to reproduce the outcomes of  $N$ -body studies of runaway and oligarchic growth where embryos have a constant spacing in terms of Hill radius (Kokubo & Ida 1998). We further enforce that no pair of embryos can lie within 10 Hill radii of each other, which is the usual spacing at the end of runaway growth (Kokubo & Ida 1998; Chambers 2006; Kobayashi et al. 2010; Walsh & Levison 2019). We thus begin with planetesimals plus embryos, as other studies by, for instance, O’Brien et al. (2006) or Raymond et al. (2009), although the planetesimals in our case are treated in a fluid-like description (surface density with a dynamic state).

The starting mass was selected such that 1) it is somewhat larger to where embryos start to repulse each other giving the 10 Hill radii separation (Kokubo & Ida 2000), 2) the mass is below the threshold where gravitational interactions start to play a role, as we found in Paper I, and 3) the mass is also below the onset of envelope effects (such as the increase of the planetesimals capture radius, Paper I). The selected starting mass is, however, larger than the transition from planetesimal runaway to oligarchic growth from Ormel et al. (2010), which for our planetesimals size is usually between  $10^{-4}$  and  $10^{-3} M_{\oplus}$ .

The embryos start right at the beginning of the simulation. This means we assume that they form in a negligible time compared to evolution of the gas disc. This is obviously a strong as-



sumption and will be revised in future generations of the model by addressing the evolution of the solids at early times (drift, planetesimal formation, embryo formation, see Voelkel et al. 2020). In our populations, we place a maximum of 100 lunar-mass embryos per system. With this number of embryos, the mean separation is roughly 28 Hill radii. This also means that a maximum of 280 lunar-mass embryos per systems could be placed while enforcing a minimum separation of 10 Hill radii.

### 3.10. Other parameters

The formation model has several other parameters that are kept constant throughout this work (Table 1). The grain opacity reduction factor in protoplanetary atmospheres, which is important for the efficiency of gas accretion in the attached phase, was set to  $f_{\text{opa}} = 0.003$ . This was selected according to the numerical simulations in Mordasini et al. (2014), which showed that this reduction factor produces the best agreement with detailed numerical simulations of grain dynamics and resulting opacities (Movshovitz et al. 2010). This value also leads to the best match of planetary metal content between numerical models and observations (Mordasini et al. 2014). The choice of the planetesimal radius follows Fortier et al. (2013), who found that small planetesimals are required to reproduce the occurrence rate of exoplanets. This is a strong assumption of the model and deviates from studies that found that planetesimals are formed big, for instance Morbidelli et al. (2009). For expanded discussions of the comparison of this values with constraints of the Solar system, the reader is referred to Paper I and Schlecker et al. (in press). The density of the planetesimals inside the ice line is set to be  $3.2 \text{ g cm}^{-3}$ , similar to values used in Hills & Goda (1993) or Podolak et al. (1988), while the density outside the ice line is taken to be  $1 \text{ g cm}^{-3}$  following Podolak et al. (1988).

Overall, the most uncertain parameters of our population synthesis are the planetesimals radius and the opacity reduction factor  $f_{\text{opa}}$ . These are the least constrained by observations, and were selected according to previous theoretical studies and population syntheses. Underlying theoretical model concepts that likely also come with large uncertainties are the general description of the gas disc as a constant- $\alpha$  viscous accretion disc, neglecting recent result on wind-driven accretion (Turner et al. 2014), or the description of gas-driven orbital migration, a process that is still not fully understood (e.g. Baruteau et al. 2016). The treatment of the early phases of the evolution of the solids is, as mentioned, also a model aspect that will be improved in future work.

### 3.11. Results

In this work, we perform five population syntheses, that differ only by the initial number of planetary embryos per system: 100 (NG76), 50 (NG75), 20 (NG74), 10 (NG84), and 1 (NG73). Here, per system also means per star and per disc. We will use the terms interchangeably in the following discussion. The names in parentheses refer to populations identifiers on the online archive DACE<sup>1</sup>.

For the populations with multiple embryos per system, we model  $N_{\text{sys,tot}} = 1000$  systems, whereas the single embryo population includes  $N_{\text{sys,tot}} = 30\,000$  systems to compensate the overall lower number of embryos. In the remainder of this section, we will discuss results at the population level without taking into

account how planets are distributed in the systems. System-level statistics will be discussed in Sect. 7.

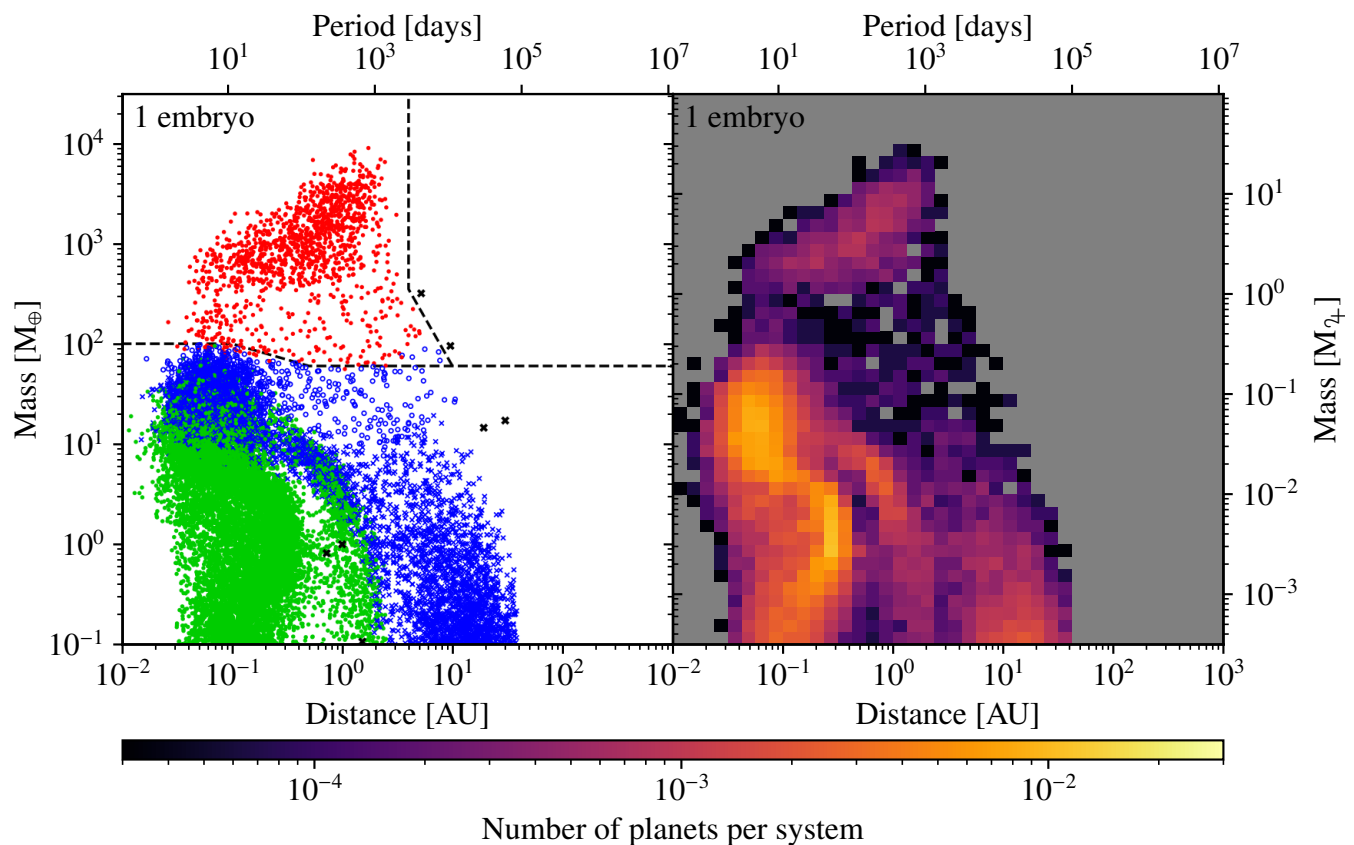
In addition, we compute two non-nominal populations with initially 100 embryos per disc but varied parameters (one with a different power-law index of the planetesimal disc and one without gas-driven migration) that we discuss in Appendix A. We use them to assess to what extent the relative results obtained with the nominal populations (like, for instance, the relative occurrence rates of super Earths versus giant planets) and general emerging trends (like correlations with stellar metallicity), are robust when changing model parameters and assumptions. The main results of this analysis are that a steeper power law index (which result in an increased concentration of mass near the star) results in more super Earths being ice-poor. We think that such relative trends should be less affected by the specific chosen model parameters than absolute results like the (absolute) occurrence rates and multiplicities of certain planet types. However, it seems still important to report them also, as first, this allows to calculate the relative frequencies, and second, when keeping the caveat in mind that these are results for given parameters only, they can still be directly confronted with observations. We also find that gas-driven migration affects the mass distribution and location of the planets. For instance, migration is necessary to bring giant planets close to the observed peak near 2–3 au (Fernandes et al. 2019; Fulton et al. 2021), however the giant planets in our nominal model are too close-in comparatively. Further, the inclusion of migration reduces the number of planets whose masses are between that of Neptune and Jupiter, which is in contradiction with certain analysis of radial velocity results (e.g. Bennett et al. 2021). Both cases contain elements that are consistent with observations, which suggests that migration is weaker than previously thought, as Ida et al. (2018) already pointed out.

## 4. Mass-distance diagrams and formation tracks

A key result of synthetic populations is the mass-against-distance diagram of the final planets. It shows what kind and where the formed planets are. This and the corresponding 2D histogram for the single embryo population are provided in Fig. 7. For the four populations with multiple embryos per system, the diagrams are shown in Fig. 8, and the corresponding histograms in Fig. 9. To generate these snapshots, we used the state at 5 Gyr. For the mass-distance diagrams, the time at which the results are plotted has a limited effect, as long as it is during the evolution stage (after 20 Myr). Only the close-in planets may be affected, either by tidal migration or photo-evaporation.

To better understand the differences between the populations and how the interactions between embryos affect planetary formation, we also analyse how different types of planets form in two populations. We therefore show in Fig. 10 formation tracks in the mass-distance diagram of selected groups of planets, for two populations: the one with a single embryo per system and the one with 100. In that figure, there are nine groups, divided in 3 series. The first series in the top panels shows Earth-mass planets close to the inner edge of the gas disc (group A in green, about 0.1 au and  $1 M_{\oplus}$ ), Earth-like planets (group B in light blue, about 1 au and  $1 M_{\oplus}$ ), and at the end of the region where such planets are found (group C in dark blue, about 40 au and  $1 M_{\oplus}$ ). The middle panels show intermediate-mass planets in the “planetary desert” (see below), in the pile-up at the inner edge of the disc (group D in red, about 0.05 au and  $30 M_{\oplus}$ ), at the position of the Earth (group E in orange, about 1 au and  $50 M_{\oplus}$ ), and at large separation (group F in yellow, about 20 au and  $100 M_{\oplus}$ ,

<sup>1</sup> <https://dace.unige.ch/populationSearch/>



**Fig. 7.** Mass-distance diagram (*left*) and the corresponding histogram (*right*) for the population with a single embryo per system. The colours and shapes of the symbols show the bulk composition: Red points are giant planets with  $M_{\text{env}}/M_{\text{core}} > 1$ . Blue symbols are planets that have accreted more than 1 % by mass of volatile material (ices) from beyond the ice line(s). The remainder of the planets are shown by green circles. Open green and blue circles have  $0.1 \leq M_{\text{env}}/M_{\text{core}} \leq 1$  while filled green points and blue crosses have  $M_{\text{env}}/M_{\text{core}} \leq 0.1$ . Black crosses show the Solar system planets. The dashed black line highlights the change of planet regime (from core-dominated, blue, to envelope-dominated, red) at  $100 M_{\oplus}$  inside 0.1 au to  $60 M_{\oplus}$  beyond 0.5 au. The vertical dashed line shows the outer limit for giant planets (4 au above  $350 M_{\oplus}$  to 10 au at  $60 M_{\oplus}$ ).

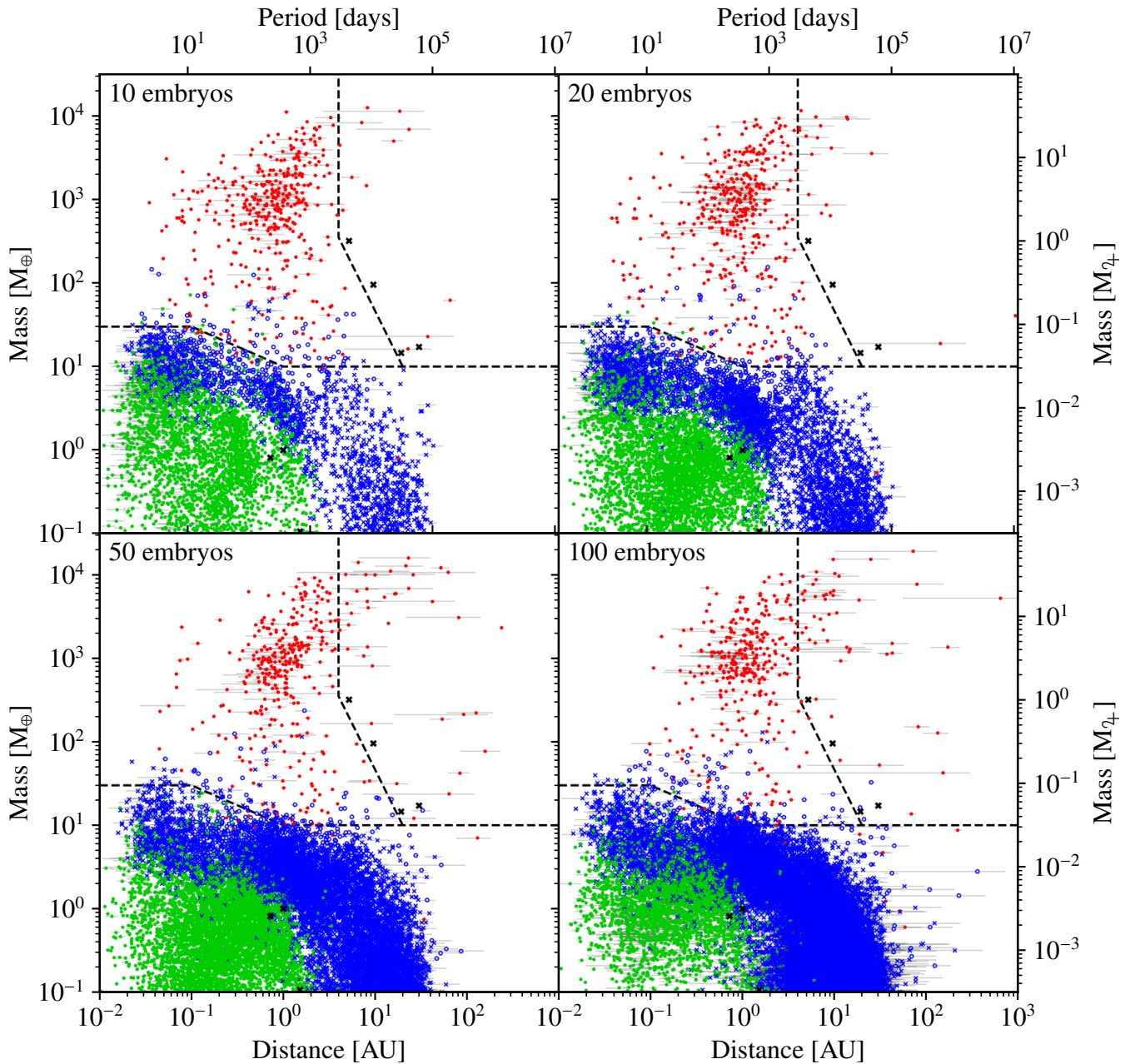
with a minimum mass of  $50 M_{\oplus}$ ). The bottom panels show giant planets, with hot-Jupiters (group G in maroon, about 0.1 au and  $800 M_{\oplus}$ ), at the location of the Earth (group H in purple, about 1 au and  $2 \times 10^3 M_{\oplus}$ ), and distant giants (group I in pink, about 40 au and  $5 \times 10^3 M_{\oplus}$ , with a minimum of  $2 \times 10^3 M_{\oplus}$ ). For selecting the planets that belong in each group, we use the following procedure: we search for the ten closest planets to the given point, the metric being the difference in the logarithm of both quantities (possibly with a second criterion on the minimum mass). This ten planets are highlighted and their formation tracks are superimposed on the overall mass-distance diagram.

In all populations, planets whose masses are between that of Neptune and Jupiter are less common than smaller or larger planets. This result is contrary to results from radial velocity (Bennett et al. 2021) and microlensing (e.g. Suzuki et al. 2016) surveys and is an area where the model could be improved in the future. As this range is where planets reach the critical mass to undergo runaway gas accretion. Planets accrete mass rather quickly here, and it is therefore unlikely that the gas disc vanishes during the short period of time planets spend in this mass range. Ida & Lin (2004a) called this deficit of planets the “planetary desert”. Another common feature is the gradual inward migration of icy planets (shown in blue symbols on the diagrams) for intermediate masses causing planets with masses higher than 3 to  $10 M_{\oplus}$  to reach the inner edge of the disc. This formation of this morphological feature is similar to the “horizontal branch” of planets found first in Mordasini et al. (2009a), as we will see in Sect. 4.1.

As the Type I migration rate is proportional to the planet’s mass (e.g. Tanaka et al. 2002; Baruteau et al. 2014), more massive planets will tend to end up at locations that are further inward from their original position than lower-mass planets, as long as the planets are not so massive that they migrate in slower Type II migration regime. An important consequence of this is that the ice content of planets when starting in the left bottom corner increases not only when going outwards to larger orbital separations as it is trivially expected, but also when moving upwards to higher masses.

Coming to the differences between the populations, we see that the single-embryo population stands out compared to the others. Among the major differences we can cite: 1) the presence of a pile up of planets between 4 and  $100 M_{\oplus}$  at the inner edge of the disc (about 0.02 to 0.2 au), 2) a different mass for the transition to envelope-dominated planets as visible in the transition from the blue to the red points ( $60$  to  $100 M_{\oplus}$  in single-embryo population compared to  $10$  to  $30 M_{\oplus}$  in the multi-embryos case, as shown by the horizontal dashed lines on Figs. 7 and 8), 3) the effect of the convergence zone for Type I migration (see for instance Lyra et al. 2010; Dittkrist et al. 2014; Paper I) which are most visible in the single-embryo population and less as the number of embryos per system increases, and 4) a total lack of distant giant planets in the single embryo population (the upper right region on the left panel of Fig. 7).

The first two effects are due to the intricate link between accretion and migration that we discuss in Sect. 4.1. The following



**Fig. 8.** Mass-distance diagrams of the populations with initially 10 (top left), 20 (top right), 50 (bottom left) and 100 (bottom right)  $10^{-2} M_{\oplus}$  embryos per disc. The symbols are identical as in the left panel of Fig. 7. In addition, the grey horizontal bars go from  $a - e$  to  $a + e$ . Dashed black lines show distinct regions in the diagrams, with the change from core-dominated (blue or green) to envelope dominated (red) at  $30 M_{\oplus}$  inside 0.1 au to  $10 M_{\oplus}$  outside 10  $M_{\oplus}$ . The vertical dashed line show the same division for giant planets in Fig. 7.

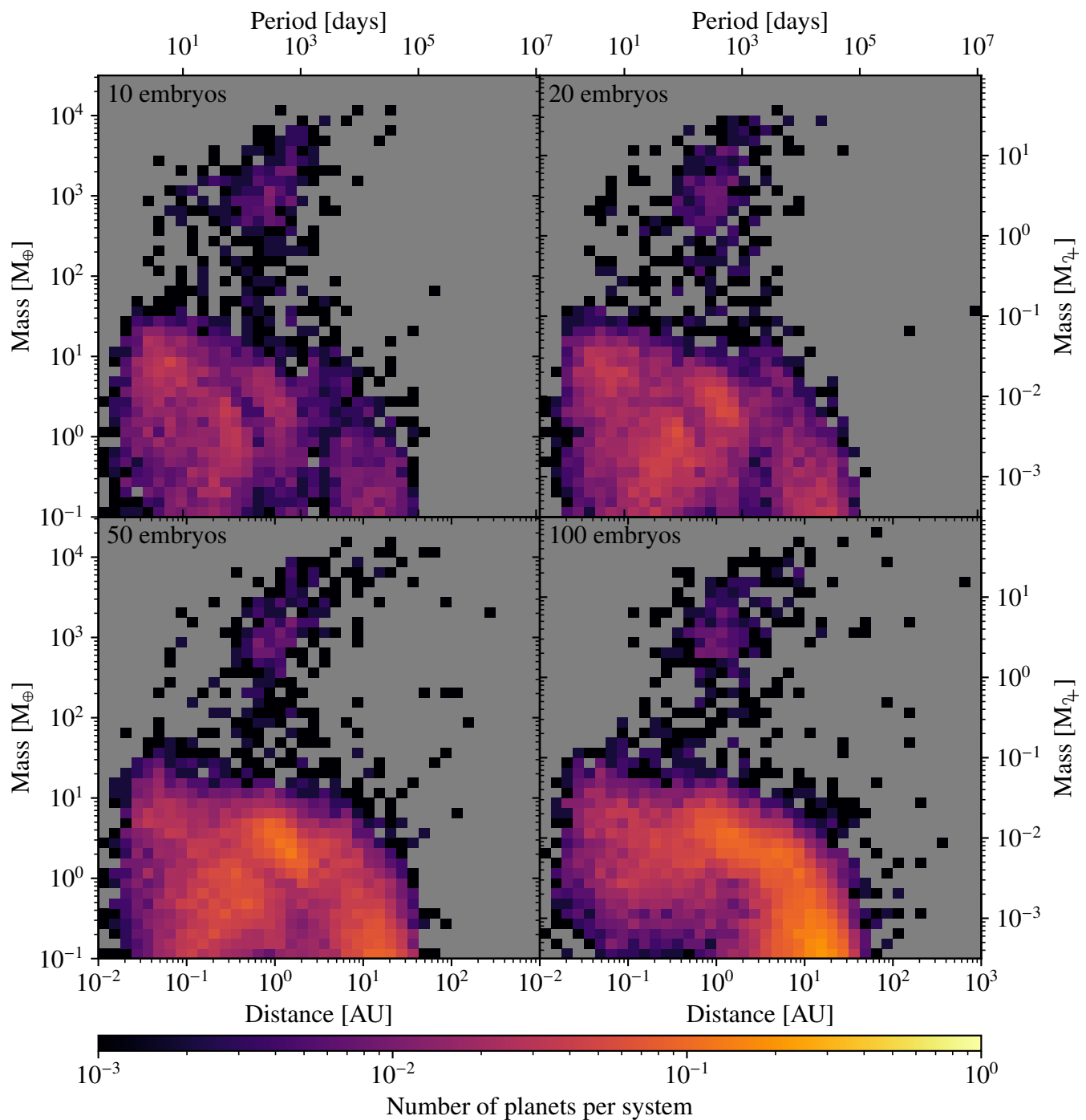
two effects are extension of the changes we see in the multi-embryos populations. If we look at all of them, we see gradual changes in the imprint of migration, the masses and locations of the giant planets. These will be discussed in Sect. 4.2. The last effect is due to close encounters resulting in planet-planet scatterings that cannot happen in the presence of only a single protoplanet. In addition, only the 100-embryos population shows one important feature about the inner low-mass planets, namely that inside of 1 au, there are less planets of very low mass ( $\sim 0.1 M_{\oplus}$ ) than planets of  $\sim 1 M_{\oplus}$ . In the populations with less than 100 embryos, there are in contrast many embryos inside 1 au that have not grown significantly. We will discuss this in Sect. 4.3.

#### 4.1. The interplay between migration and accretion: single versus multiple embryos per disc

The single-embryo population stands out with respect to the multi-embryos in several ways. A major difference is that the single-embryo population completely lacks dynamical interactions. This means that the only possibility for a planet to change its location is through orbital migration, and migration is not hampered by the presence of other embryos.

To compare the formation tracks of planets in the single-embryo population with that of the multi-embryos populations, we chose only particular to study. We came down to system 438, which has a solids disc of  $287 M_{\oplus}$  and can form planets up to about  $2 M_{\oplus}$  in the single-embryo population, depending on the





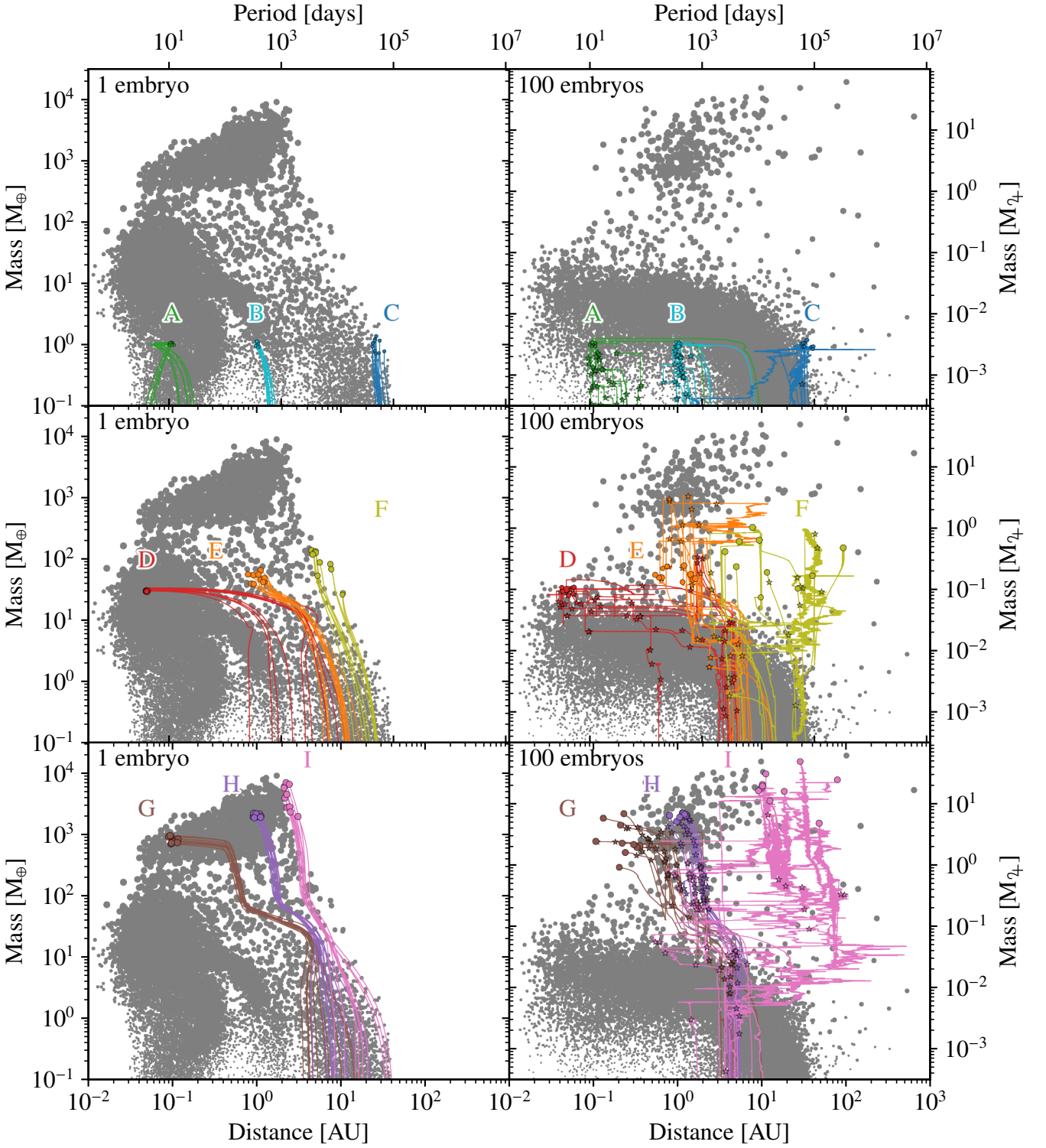
**Fig. 9.** Two dimensional histogram of mass-distance relationship of the populations with initially 10 (top left), 20 (top right), 50 (bottom left) and 100 (bottom right)  $10^{-2} M_{\oplus}$  embryos per disc (as in Fig. 8). The number of planets has been normalised by the number of systems. The colour scale is the same in all populations, but different than in Fig. 7. Gray regions have no planets.

initial location of the embryo. The possible formations tracks of a single embryo in this system are provided Fig. 11. There, we show a grid of 100 distinct systems whose initial conditions are identical, except for the initial location of the embryo. The initial location was set using a uniform spacing in the logarithm of the distance, from the inner edge to 40 au. This means that each point of the grid represents the same probability in the population.

It can be seen in Fig. 11 that many of the intermediate-mass planets, between 10 and  $100 M_{\oplus}$ , end up at the inner edge of the protoplanetary disc. This is due to migration being most effi-

cient in this mass range. The low- and middle-mass planets (up to a few tens of Earth masses, see Fig. 10 of Paper I) will undergo type I orbital migration, whose rate is proportional to the planet’s mass (e.g., Tanaka et al. 2002; Baruteau et al. 2014). Thus, the least massive planets (below  $1 M_{\oplus}$ ) will remain close to their original location.

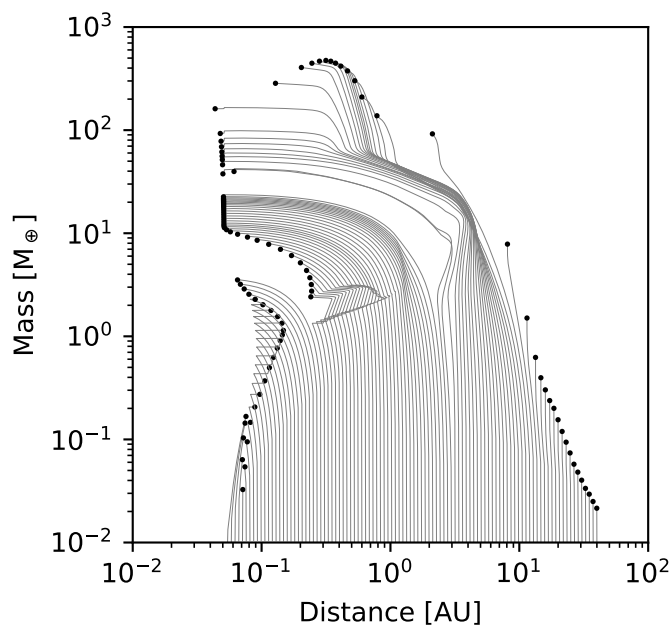
Now, in the single-embryo population, this fast migration will only stop under two conditions: 1) when the planet reaches the inner edge of the disc, or 2) when the planet grows sufficiently to open a gap in the disc and switches to type II migration, which is significantly slower.



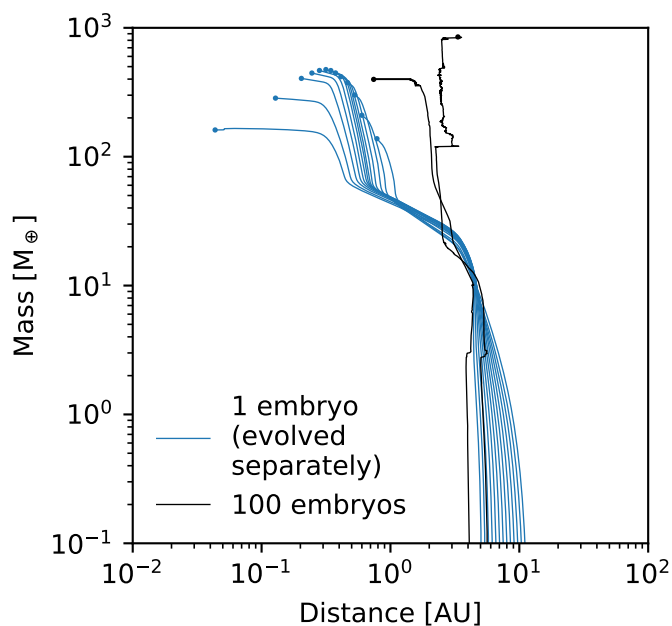
**Fig. 10.** Comparison of the formation tracks between the population with initially 1 (*left*) and 100 (*right*) embryos per system and 9 different group of planets labelled A through I, each shown with a different colour. The positions of the groups in the mass-distance diagram are explained in the text. The stars in the 100-embryos population denote the instant at which the planets were hit by other protoplanets (giant impacts).

Thus, to avoid being taken to the inner edge, planets must grow rapidly while they are in the 10 to 100  $M_{\oplus}$  range. The planets are still in the planet-limited gas accretion regime at this epoch (that is, the attached phase): gas accretion is limited by the ability of the planet to radiate away the gravitational energy gained by accretion of both solids and gas. Thus, if the planet is still accreting solids, its ability to bind a large amount of gas

is severely limited. To be able to undergo runaway gas accretion, the planet must either strongly decrease its solids accretion rate or attain a mass large enough so that cooling (and therefore contraction which allows gas accretion) become efficient, hence being able to accrete gas despite the solids accretion rate remaining large.



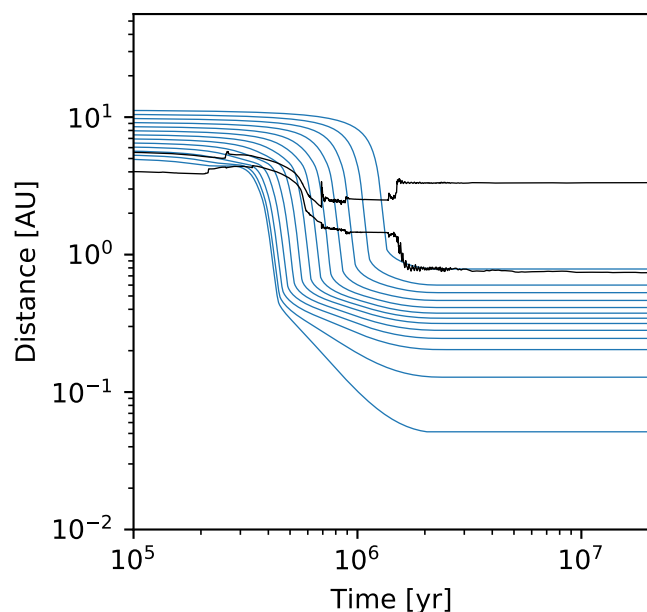
**Fig. 11.** Possible formation tracks for the case on a single embryo per system for one given disc. This figure shows 100 systems with the same initial conditions except for the initial location of the embryo. The initial location was varied from the inner edge of the disc to 40 au with an even spacing in the logarithm of the distance to reflect our choice of initial embryos location in the overall population.



**Fig. 12.** Comparison of the formation tracks between single embryos (in blue) and the corresponding 100-embryos system (in black). Only planets whose masses are larger than  $100 M_{\oplus}$  are shown.

In contrast, multi-embryos populations have additional mechanisms to prevent rapid inward migration:

- mean-motion resonances with closer-in protoplanets that will slow migration down (because the torque is spread over multiple bodies when planets migrate in resonances, see e.g. Lee & Peale 2002; Kley & Nelson 2012), and
- a decrease of the accretion rate of solids that will enable the protoplanet to undergo runaway gas accretion.



**Fig. 13.** Comparison of the distance as function of time between single embryos (in blue) and the corresponding 100-embryos system (in black). Only planets whose masses are larger than  $100 M_{\oplus}$  are shown.

The combined effect of these two processes can be seen in Fig. 12. The protoplanets migrate more slowly in the multi-embryos populations, which prevents the rapid inward migration and at the same time reduces the planetesimals accretion rate, enabling runaway gas accretion at lower core masses. The resulting giant planets are located further out and have smaller core masses than the planets formed in the single-embryo case.

#### 4.1.1. Mean-motion resonances as a way to reduce inward migration

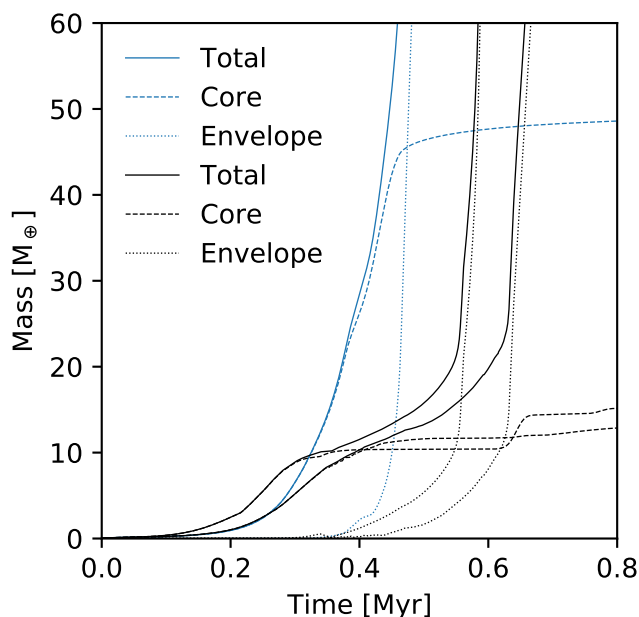
In the single-embryo population, planets follow precisely the migration prescription, as shown in Fig. 11. For the multi-embryos populations however, dynamical interactions have to be taken into account too. One possible dynamical interaction is the trapping in mean-motion resonances (MMRs). What we found is that the trapping in MMRs can significantly reduce the inward migration of intermediate-mass planets; we show an example of this in Fig. 13. That figure shows the same planets as in Fig. 12, but instead provides the time evolution of the planet's distances.

Figure 11 shows regions of outward migration for planet masses between 1 and  $20 M_{\oplus}$ . These are caused either by opacity transition near the iceline (Lyra et al. 2010) or structure in the gas disc (Kretke & Lin 2012). A migration map depicting these features is shown in Fig. 10 of Paper I.

In Fig. 13, it can be seen that the two giant planets formed in the 100-embryos system (in black) migrate much more slowly than planets that are alone in their system (in blue). A sketch of how the effect happens is the following:

- embryo growth happens inside out, as accretion time scale is shorter in the inner region of the disc,
- more distant embryos become larger than closer-in ones as the isolation mass increases with distance, and
- once embryos start to feel the torque from the disc that leads to migration, they will be trapped into MMRs with closer-in, lower-mass planets, which will result in a reduced migration speed for the largest embryos.





**Fig. 14.** Comparison of the mass (core, envelope, and total) as function of time between one single embryos (in blue) and the corresponding 100-embryos system (in black, with only planets whose masses are larger than  $100 M_{\oplus}$  are shown). The scales are linear to compare to Pollock et al. (1996).

As closer-in, lower-mass planets will have a lower intrinsic torque, they would migrate more slowly than outer, more massive planets. Thus, to allow the outer planets to migrate, the latter have to transfer torque to the inner, low-mass planets through MMRs. This will lead to a reduced migration speed because the torque has to be spread across multiple planets.

A consequence is that planets in the  $20\text{--}50 M_{\oplus}$  will have more time available before ending in the inner region of the disc. Another consequence is that planets can be pushed out of the convergence zones of orbital migration. For instance, we observe in the single-embryo population, three different zones with a lack of planets in between. The first two are for rocky planets, while the latter contains mostly icy planets. In the 10, 20, and 50-embryos we still see some imprint of the convergence zones, each time with a decreased intensity. In the 100-embryos population, the effects of the convergence zones have nearly vanished.

#### 4.1.2. A reduction of the solids accretion rate

With a single embryo per system, there can be no reduction of the accretion rate of solids once the planet starts to migrate. This is because the planet will always find new material to accrete from as it enters regions full of pristine planetesimals. It will most likely end only when the planet comes to the inner edge of the disc. The thermal support of the envelope because of strong continuous planetesimal accretion is sufficient to prevent runaway gas accretion except for the most massive cores. Hence, giant planets in that population always have a massive core, because it is the only way for them to undergo runaway gas accretion quickly. This effect also requires that a large amount of solids is present where the planet forms, so that it can accrete a very massive core without migrating too much.

On the other hand, when multiple embryos are present, the competition for solids provides a different pathway for giant planets to form. In this scenario, the initial part of the accretion

of the core, until planets start to migrate, remains similar as in the single-embryo case. However, once the core experiences inward type I migration, it will enter at some point a region where another embryo has grown and depleted the planetesimals. This will deprive the first core of material to accrete from and cause a sudden decrease in the accretion rate of solids. As consequence, there will be a drop in the luminosity released by the accretion of solids, which opens the pathway to trigger runaway gas accretion at lower masses.

This difference is able to explain the first two items we mentioned above, namely the pile-up of massive close-in planets at the inner edge in the single-embryo population and the difference for the transition to envelope-dominated planets ( $\sim 100 M_{\oplus}$  versus 10 to  $30 M_{\oplus}$ ). Also, the more embryos there are, the less migration is needed to enter the region where another embryo has already accreted, as the embryos are more tightly packed. This results in a lower extent of Type I migration in the many-embryos populations, as the planets will undergo gas runaway more rapidly and switch to the slower Type II migration.

This effect also means that the multi-embryos populations have a way to limit the accretion of planetesimals as it would occur if the embryos “shepherd” the planetesimals while they migrate (e.g., Tanaka & Ida 1999). Thus, the single-embryo populations does not represent the true situation with the efficient accretion of planetesimals during planetary migration.

To see such effects, it is necessary to calculate the interior structure to get the accretion rate dependent on the core accretion rate and the corresponding luminosity. With a model where the envelope mass depends only on the core mass, such an effect cannot be reproduced. It should also be noted that collisions with other embryos are included in our model. An additional contribution to the luminosity by collisions is included in the internal structure calculation (Paper I), and it does not provide a continuous luminosity source. They do not hinder gas accretion on the long term as does a relatively continuous accretion of planetesimals (Broeg & Benz 2012).

#### 4.2. Other effects of the number of embryos

There other, gradual changes that arise as the number of embryos increase. These include the distant giants and planets with masses below  $2 M_{\oplus}$  and distance below  $0.02 \text{ au}$  (i.e., inside the inner edge of the disc). These effects are mainly due to gravitational interactions between the protoplanets.

MMRs can push planets inside the inner edge of the disc by the inward migration of another planet which is still located within the disc, hence we find planets closer to the star than the inner edge of the disc in the corresponding populations.

We set the limit for the transition between ice-poor (rocky) and ice-rich planets at 1% of volatiles by mass in the core. This is to avoid having planets with extremely low amount of volatile appearing as icy in the diagram. The limit was set according to the amount of water (the main component of volatiles) to obtain high-pressure ice at the bottom of oceans of a  $1 M_{\oplus}$  planet (Alibert 2014). There is nothing particular happening in the model at this limit, it is only set for visualisation. Comparing the different population, we find that in the 50 and 100-embryos populations, ice-rich planets are found in regions populated only with ice-free planets in the 10-embryos population. This can be seen at the position of the Earth on Fig. 8. In the 10-embryos population, the Earth lies in a region harbouring only ice-free planets. In the 20 and 50-embryos populations, the Earth lies at the transition between the two, while in the 100-embryos population, it is in the

icy-rich region. Further, dynamical interactions are able to send icy low-mass planets in the inner region of the disc (inside 1 au)

As the number of embryos increases, we observe a greater mixing of the rocky and icy planets. In the single embryo population, the two are well separated, while as the number of embryos grows, we note more and more icy planets of a few Earth masses in the inner part of the disc. This affects only planets of more than a few Earth masses, or regions directly inside of the ice line; for instance, we do not obtain icy planets less than an Earth mass within fractions of an AU. As the single embryo population shows, bringing icy low-mass bodies to the inner part of systems is not possible by migration only, so there must be multi-body effects, such as close encounter and capture in resonance, that send part of the icy planets forming outside of the ice line in the inner part of the disc.

At large orbital distances, the populations with multiple embryos per systems contain planets located outside the outer limit of embryo starting locations (40 au) while the population with a single embryo does not. The only possibility for planets to end at those position are scattering events due to close encounters with other planets, as outward migration does not happen at these locations. The black horizontal bars of these planets on Fig. 8 show their eccentricities. We see all of these planets have a periastron inside 40 au, indicating that these planets come in a location where other planets are present at some point during their orbit. We then might find planets formed by core accretion at large separation, but with our model these planet remain on eccentric orbit, as circularisation does not happen on a sufficiently short time scale before the dispersal of the gas disc. We could thus explain directly-imaged planets at large separation, such as HIP 65426b (Chauvin et al. 2017) only if they have a significant eccentricity to have a periastron at a distance where core accretion is efficient, i.e., inside of  $\sim 10$  au. This formation scenario was studied extensively in Marleau et al. (2019).

#### 4.3. Formation of low-mass planets

The formation tracks of the low-mass planets in the single-embryo case (top left panel of Fig. 10) is straightforward. As we already mentioned in Paper I and Sect. 4, gas-driven migration is weak for these planets, so that they end close to where they started, with minimal inward migration. We can still note that the close-in group (A, in green), there is either outward migration all the way through, or inward migration followed by an episode of outward migration without accretion. This effect is caused by the presence of the innermost outward migration zone for low-mass planets (see Fig. 7 of Paper I). We are in the presence of two scenarios that depend of the disc characteristics: either planets are inside the outward migration zone from the beginning and they will move out while they accrete, or they are in the inward migration zone at the beginning and pass in the outward zone later on during the disc's evolution. In the latter case, there is no accretion during the second pass in a region because all planetesimals were previously accreted.

In the 100-embryos population, the formation of the same resulting planets are more varied. For the two innermost groups (A and B), we divide the planets in two groups. First, for 16 planets, there is growth by giant impacts that we had anticipated and discussed in Paper I. These planets have starting distances of 0.1 to 0.3 au. The second (4 planets) is growth by accretion of planetesimals at a much larger distance (starting distances of 2 to 10 au) followed by a strong inward migration combined with limited accretion. This pathway is unseen in the single-embryos population because the inward migration is caused by the trapping

in resonance chains with other more massive planets (around  $10 M_{\oplus}$ ) that experience stronger migration. Clearly, these different formation pathways could result in diversity in terms of the composition of the planets, as we will discuss below. For the outermost group, we see that there are also two formation pathways, but they are not the same as for the inner groups. The first pathway is the same as in the single-embryo population, where the only effect is limited inward migration. The second is growth stirred by more massive planets, which causes jitter in their location and occasional scattering events. However, we see that these planets undergo much less giant impacts than their close-in counterparts. This implies that these planets growth mostly from the accretion of planetesimals, in a similar way than planets in the single-embryo case.

There is also a specific feature absent in the other population, for planets within 1 au: a decrease of the occurrence rate with decreasing mass at masses less than  $3 M_{\oplus}$ , with a near total absence of bodies of the mass of Mars ( $\sim 0.1 M_{\oplus}$ ). This feature relates to the formation of the terrestrial planets we discussed in Paper I. It applies to systems with a low metallicity, where migration is unimportant because growth is slow. Only in the 100-embryos population, the inner region of the disc is fully depleted in planetesimals and the embryos end their growth with a “giant impact” stage, similar to the terrestrial planets in the solar system (Wetherill 1985; Kokubo & Ida 2002). In the other populations, the spacing between the embryos is too large and they end up growing as if they were isolated. This means that they grow only to masses that are much less compared to the case that all solids in the inner disc end up in planets as in the 100-embryos population, instead of remaining in planetesimals.

The 100-embryos populations should hence be representative of the formation of planets spanning the entire mass range, at least with 1 au. This will enable use to compare architectures of low-mass (i.e. terrestrial) systems with observations to determine if planet pairs have similar masses (Millholland et al. 2017), radii and spacing (Weiss et al. 2018); see Schlecker et al. (in pressb). Accretion through giant impacts is stochastic in nature, and planets may well have collided with bodies originate from beyond the ice lines, or that have themselves had giant impacts with such distant embryos. This explains why we obtain close-in planets that have some content of icy material. For the second pathway (forced migration), we see that those planets have accreted most of their mass before experience the strong migration. We can thus expect that they harbour a large amount of icy material.

There are two caveats with our model for the formation of terrestrial planets. Firstly, we set the limit for the formation stage to 20 Myr, where planets can accrete and dynamically interact via  $N$ -body integration. As the accretion time scale increases with distance, this limitation affects more the outer part of the system. In the case of a system with an initial mass comparable to MMSN, we found in Paper I that by about 20 Myr the instability phase had mostly finished inside of 1 au. This is comparable to what Chambers (2001) found is required for Earth-like planets to accrete half of their mass for a similar scenario. Growth is more rapid in systems with a higher solid content though. It follows that for a disc with MMSN content of solids, the low-mass planets obtained in our population are mostly at the end of their formation in the inner region, roughly a factor two too small around 1 au, and that much longer times are required for more distant planets. The limited time for the  $N$ -body interactions (which in our model occur only during the formation) can also mean that we miss dynamical instabilities at late times.

For instance, Izidoro et al. (2021) found that it can take up to 100 Myr for systems to go unstable.

The second caveat is that terrestrial planets accrete predominantly from other embryos, rather than planetesimals as it the case for the giant planets. However, giant impacts, due to similar size of the involved bodies have a variety of possible outcomes (Asphaug 2010). Accretion (or merging) is not the most common result of such a collision (Stewart & Leinhardt 2012; Chambers 2013; Quintana et al. 2016). Despite this, our collision model is unconditional merging when it comes to terrestrial planets (Paper I). For instance, we see that we are unable to form equivalents of the smaller terrestrial planets of the Solar System, Mercury and Mars as a majority of terrestrial-planet forming systems give planet masses similar to  $1 M_{\oplus}$ . Mercury might be the result of a series of erosive collisions (Benz et al. 2007; Asphaug & Reufer 2014; Chau et al. 2018) that are not part of our model. Using a more realistic collision model, Chambers (2013) found that the resulting planets have slightly lower mass and eccentricities, and that the overall time required to form these planets increases, as collisional accretion is not as efficient.

#### 4.4. Formation of intermediate-mass planets

The formation of the intermediate-mass bodies in the single-embryo population (left middle panel of Fig. 10) has already been largely discussed in Sect. 4.1 for the innermost group (D) and in Sect. 8 of Paper I for the more distant groups (E and F). They are in the range where migration is most efficient, massive enough to undergo strong migration, but not massive enough to open a gap and migrate in the slow Type II regime. The inner group (D, in red) is similar to the “horizontal branch” of Morasini et al. (2009a).

In the 100-embryos population, some of the planets in the innermost group (D) form in a similar fashion as in the single-embryos case, with the exception of some giants impact near the end of their migration. However, a few of these planets had their envelope removed in the aftermath of giant impacts that occurred at about 2 au. Were it not for the impacts, these planets would have most certainly ended up being giants at larger separation.

The more distant groups (E and F) have, however, a different formation history. Half of the planets in the mid-distance group had at some point a mass larger than  $100 M_{\oplus}$  while their final mass is close to  $50 M_{\oplus}$ . The mass loss is due to the removal of the envelope following a giant impact, due to the burst of luminosity caused by the sudden accretion of the impactor. The mass loss is delayed by about  $3 \times 10^4$  yr from the time of impact due to the timescale of release of the supplementary luminosity following Broeg & Benz (2012). It is then possible for the giant impact to not be marked exactly at the location of the mass loss in the formation tracks. We also see that some of these planets spent time around 10 au after beginning the runaway gas accretion closer to 1 au. The formation tracks of these planets also show sudden change in their position, both outward and inward. Thus, not only migration is responsible for their final locations, but also close encounters. Actually, migration plays a lesser role in the 100-embryos population, as most of these planets began inside 8 au, while in the single-embryo population most of the embryos come from outside 10 au.

It is obvious that the effects induced by the concurrent formation of several planets introduces strong additional diversity in the formation pathways of planets. In the single-embryo case, planets may undergo gas runaway only once, and this must be late in the evolution of the gas disc to not accrete too much gas. With multiple embryos, the possibility of giant impacts means

that it is possible to undergo gas runaway multiple times, provided the envelope is removed in between.

#### 4.5. Formation of giant planets

The formation of giant planets in the single-embryos population (left bottom panel of Fig. 10) has also been discussed in Sect. 8 of Paper I. They follow a similar pattern than for the intermediate masses at the beginning, but accretion dominates over migration, as indicated by the different slopes of the tracks in the mass-distance plane. Starting with roughly  $10 M_{\oplus}$ , the formation tracks of the different group begin to be well separated from each other. Then, the planets that finish closer-in will undergo larger migration up to about  $50 M_{\oplus}$  before undergoing runaway gas accretion. Migration takes over again later and, during later times, the planet migrate with only little accretion, as indicated by the different slopes of the tracks in the mass-distance plane. For the more distant planets (groups H and I), a similar structure of the formation tracks is observed, but migration remains overall less efficient. This is because accretion is very fast, as they must undergo runaway gas accretion in the early times so that they can accrete gas during most of the life time of the protoplanetary disc. As accretion is so fast, it leaves only little time for migration to act.

For the 100-embryos population, we first note that gas-driven migration is overall less efficient than in the single-embryo population. For instance, there are no giants inside 0.1 au, so planets in the innermost group (G) are on average further away than in the single-embryo case. These planet undergo gas runaway at little bit further out than in the single-embryo case: the former are slightly outside 1 au, while the latter are inside of that mark. The difference comes later, with the planets in the 100-embryo population not migrating as much afterwards. The initial location of the seeds forming these planets is also slightly different: in the 100-embryos population they are concentrated between 3 and 6 au (and one case at about 10 au) while in the single-embryo close-in giant form from seeds between 3 and 15 au, only half of them being inside of 6 au. For the intermediate-distance planets (group H), in the case of the 100-embryos population, we see many giant impacts occurs. However, since these are giant planets, most of them did not loose their envelope. The few that did loose it have accreted a secondary envelope, and since they had already migrated inward, ended closer-in that we would have if they had not lost the envelope. The final part of their formation track is otherwise very similar to the single-embryo population. The initial part is different, with the same observation as the for the close-in giants: the seeds come from closer-in. Actually, the seeds for the close-in and intermediate-distance giants are from the same region of the disc in the 100-embryos population (between 3 and 6 au). In the single-embryo population however, these are centred at about 8 au.

For the giant planets ending up at 0.1 au and 1 au, we see by the numerous stars along the brown and violet tracks that once the giant planets have triggered runaway gas accretion and masses  $>100 M_{\oplus}$ , they are hit by numerous planets. The mass of the impactors are mostly  $5 M_{\oplus}$  and lower. These impactors also mostly come from nearby the giant planet they collide with. The collisions are due to the destabilisation of other less massive planets by the forming giant because of its strong mass growth coupled with migration. The effect of the impacts on the luminosity both during the formation and evolution phases will be investigated in a future publication of this series.

Distant giant planets in the 100-embryos population (group I) grow even less smoothly than the other groups. Interestingly



enough, they come again from the same region of the disc as the two other groups in the 100-embryos population. They are then scattered outward by other massive planets in the same system. The fact that there are other massive planets in the same system is recognisable by the jitter in their distance. The large diversity of formation tracks of planets with very similar final mass and orbital distance in the 100-embryos population illustrates how difficult it is to infer the formation tracks of a specific planet just from the final position of an individual planet in the mass-distance diagram.

#### 4.6. Formation time

To further characterise the differences between the populations, we seek to determine the time it takes for the planets to form. For this, we plot in Fig. 15 the time at which the core’s mass is half of its final value, itself given in terms of the life time of the protoplanetary disc. We show only the two end members of our populations: the single-embryo one of the left, and the 100-embryos population on the right.

In all populations, we note that the core of the giant planets formed early. In the case of the single-embryo population it is not particularly different from other types of planets (principally the close-in ones), but in the 100-embryos population it does stand out compared to the rest of the population, where most of the other planets form much later. Further, the single-embryo planet shows a consistent trend for the formation time, with the most massive giant planets forming their core earliest, while the less massive ones, and the planets in the desert forming late. This trend is much less perceptible in the 100-embryos population. Some of the planets in the desert formed quite late and show a formation history similar to what is obtained in the single-embryo population, that is, a smooth growth and slight inward migration. Others however had their core formed early, in the same way as for the most massive giant planets. However the planet-planet interactions enable other pathways for the formation of these planets. For instance, some have had lost partly or entirely their envelope following a giant impact, leaving the planets with a limited time to accrete gas again. Others have been trapped in mean-motion resonances with massive planets still in the disc. These resonances will prevent the fast inward migration, thus enabling a formation process that is somewhat similar to one of giant planets without migration (e.g., Pollack et al. 1996), that is, with a significant delay between the accretion of the bulk of the core and the onset of the runaway gas accretion.

For the intermediate-mass planets, we observe that the ones that are close-in had their core formed quite early. In the single-embryo population, these are the planets that are close to the inner boundary of the disc, while for the 100-embryos population, this concerns some close-in planets in the 1 to 30  $M_{\oplus}$  range. This is related to the discussion in Sect. 4.1, about the planets that formed close to or beyond the ice line and rapidly migrated inward without undergoing runaway gas accretion, similar to the “horizontal branch” of M09a.

For the inner low-mass planets, (up to 1 au and 30  $M_{\oplus}$ ), we obtain that the time taken for the formation increases along with the number of embryos. This can be seen in the right panel of Fig. 15, where many of these planet only attain half of their final mass after the dispersal of the gas disc. Thus a large part of the accretion process occurs in a gas-free environment. This helps explain for instance the disappearing of the features related to the gas disc, such as the migration traps.

We hence come back to the discussion of Sect. 4.3 about the integration time required to model the formation of these sys-

tems. Our limit of 20 Myr is still too short for planets beyond  $\sim 2$  au, as here we have a increasing fraction of planets that still accreted most of their mass while the gas disc was still present. This indicates that there have not been interactions after the dispersal of the latter, hence that a longer integration time could lead to fewer, more massive planets. This process however takes a long time to happen, on the order of 100 Myr or more, as we have seen with the terrestrial planets. Thus the direct modelling would be a very expensive task computationally in the context of a population synthesis, i.e. for around 1000 planetary systems.

One consequence of the late formation of the low-mass planets, which turns out to be more similar to the formation of the solar system’s terrestrial planets is the (in)ability of these planets to retain an envelope. Giant impacts that occur after the dispersal of the gas disc may lead to the ejection of the planet’s envelope, but it cannot be reaccreted any more. Atmospheric escape is then not the only mean to loose the envelope, and the evaporation valley (Lammer et al. 2009; Lopez et al. 2012; Owen & Jackson 2012; Jin et al. 2014; Jin & Mordasini 2018) is not as clear as in the populations with a lower initial number of embryos per disc.

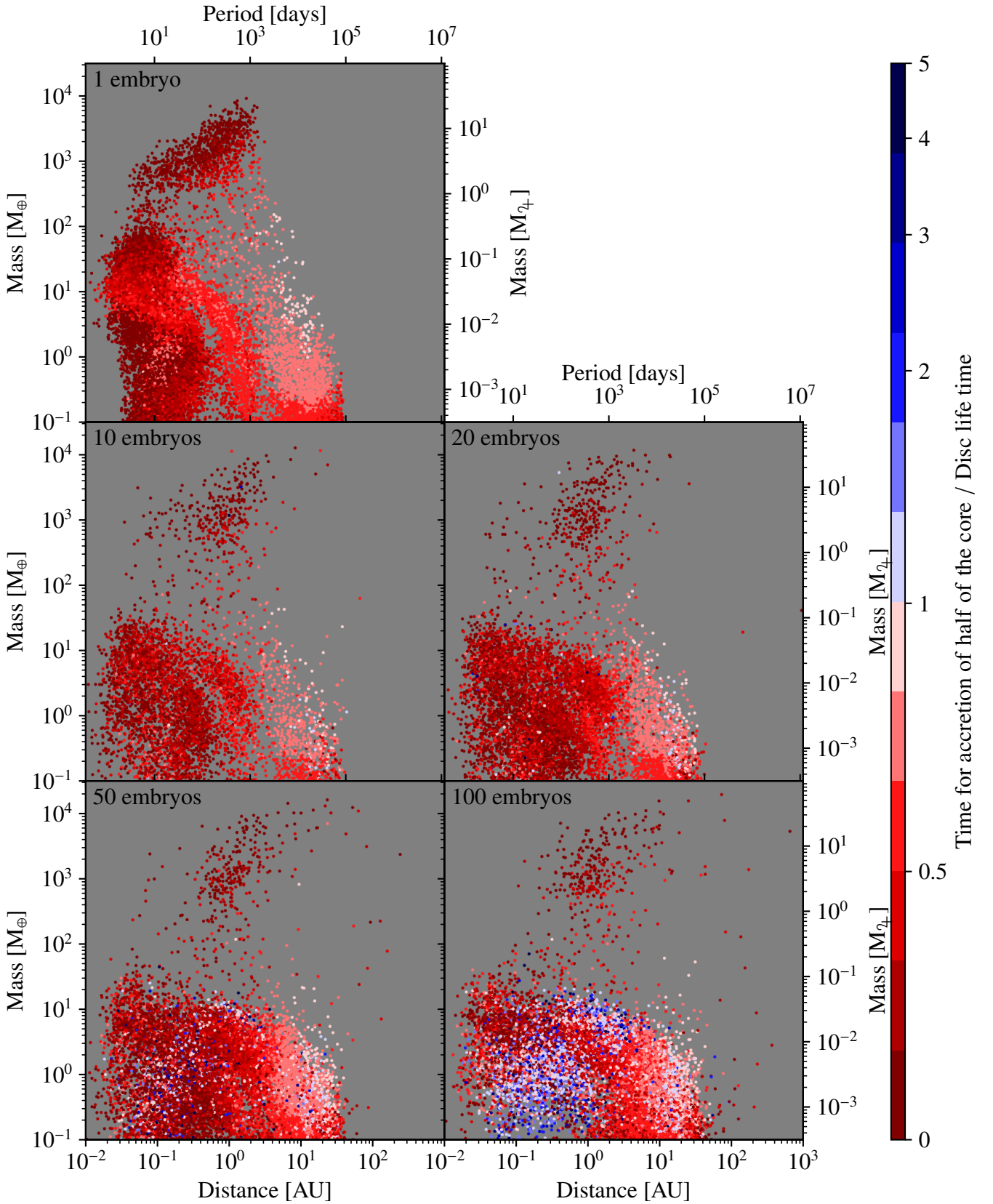
In future work, we will improve the way how impact stripping of gaseous envelopes is dealt with (Schlichting & Mukhopadhyay 2018; Denman et al. 2020). As described in Paper I, at the moment the impact energy is added into the internal structure calculation. This however neglects the mechanical removal of some gas during the impact via momentum exchange, and also assumes that the entire impact energy is deposited evenly deep in the planet’s envelope, at the core-envelope boundary. In reality, only a part of the envelope close to the impact location might be strongly heated. Both these effects affect the efficiency of impact stripping. Interestingly enough, in a population synthesis, the emptiness of the valley can be used to observationally constrain the efficiency of impact stripping. The fact that the valley seems to be too populated compared to observations in the 100-embryos model is an indication that the current model for impact stripping in the Bern model is too efficient.

## 5. Planet radii and luminosity

### 5.1. Mass-radius relationship

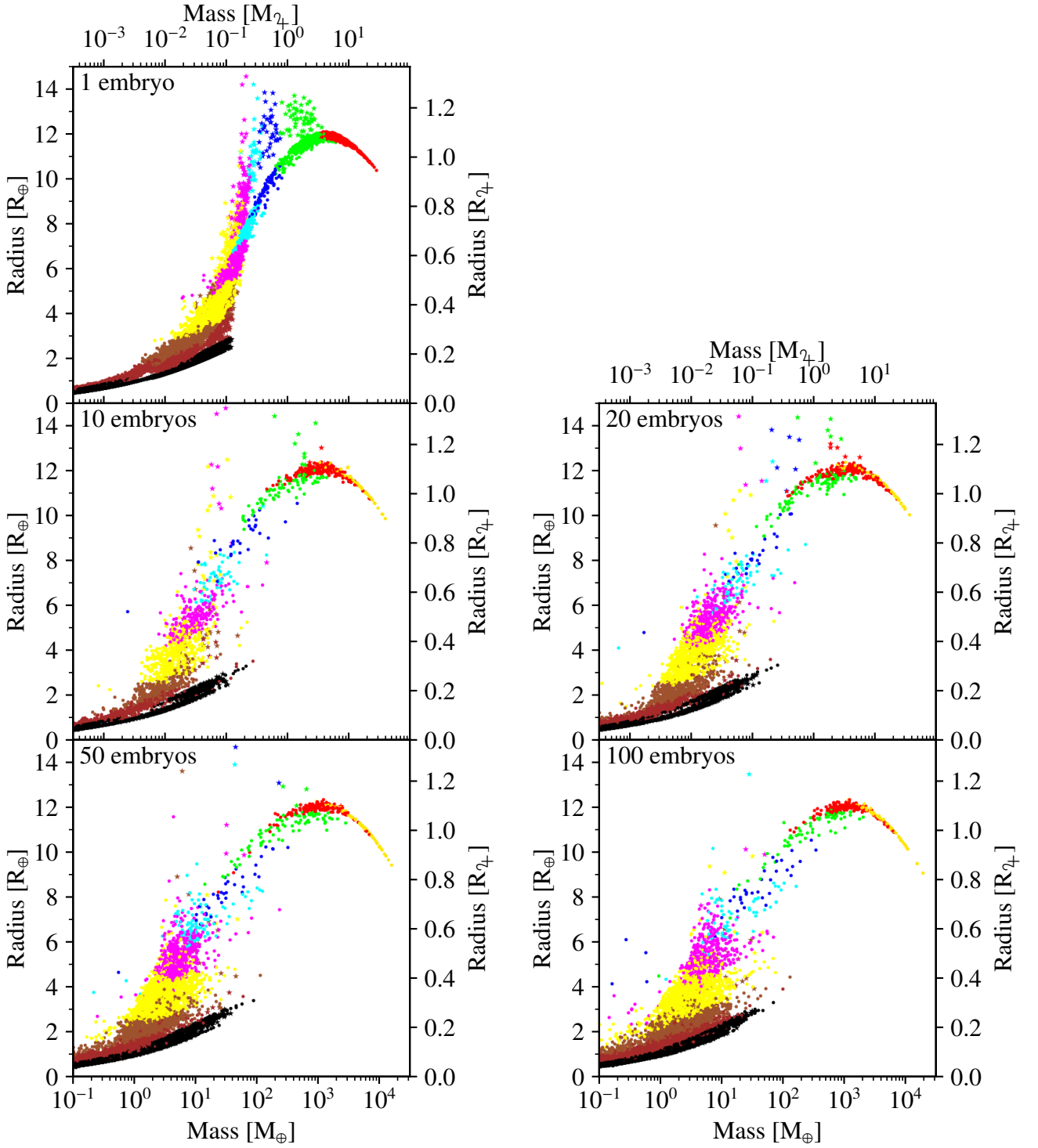
The mass-radius relation in the context of formation and evolution of planetary systems with 1 embryo per disc was extensively discussed in Mordasini et al. (2012b) and Mordasini et al. (2014). To compare that scenario with the case of many embryos per disc, we show in Fig. 16 the five population that are part of this study.

As discussed in Mordasini et al. (2012b), the global structure of the mass-radius relationship is caused by the combined effects of the properties of the equation of state of the main planetary forming materials (iron, silicates, ices, H/He), and the increase of the H/He mass fraction with the planet mass. The overall lower core masses in the population with multiple embryos per disc result in comparatively increased radius and lower metallicity for a given planet mass as the number of embryos increases. The spread of radii for a given mass is due to both different planet metallicities,  $M_{\text{core}}/M$ , and distance to the central star. The last effect is more important in our work than in Mordasini et al. (2012b) due to the prescription for the bloating of the close-in planets following Thorngrén & Fortney (2018) with a criterion for a minimum stellar flux of  $2 \times 10^8 \text{ erg s}^{-1} \text{ cm}^{-2}$  from Demory & Seager (2011). Planets that satisfy this criterion and have their symbol set to a star in Fig. 16. We note in the single-embryo



**Fig. 15.** Two dimensional histogram of mass-distance relationship of the five populations shown in this study, with the initial number of embryos per system given in each panel

. The dot colours denote the time needed for each planet to accrete half of their final core mass, given in terms of the disc life times. The scale is linear from 0 to 1, and then logarithmic. The ratio can be larger than unity because we model formation to 20 Myr while the median disc life time is about 4 Myr.



**Fig. 16.** Mass-radius diagram at 5 Gyr obtained from following the long-term thermodynamic evolution of the five populations, as marked in each panel. Atmospheric loss by impacts, atmospheric escape driven by XUV photoevaporation, and bloating are included. The iron and ice mass fraction in the core is considered when calculating the radius. The colours show the bulk solid mass fraction  $M_{\text{core}}/M$ , the rest being the H/He envelope: orange:  $<0.01$ , red:  $0.01-0.05$ , green:  $0.05-0.2$ , blue:  $0.2-0.4$ , cyan:  $0.4-0.6$ , magenta:  $0.6-0.8$ , yellow:  $0.8-0.95$ , sienna:  $0.95-0.99$ , brown:  $0.99-1$ . Black: no H/He envelope. Star symbols indicate planets who have additional luminosity from bloating, while circles indicate planets that do not.

population, there are two branches in the mass-distance diagram, with the bloated planets having a radius few  $0.1 R_{\text{J}}$  larger than their more distant counterpart. This branch does not continue for masses above  $10^3 M_{\oplus}$  because we do not have that massive planets at close-in locations. The 100-embryos population, however, does not show the second branch in the mass-distance diagram

for the bloated planets because there are no giant planets at close-in locations, and only few for less massive bodies.

The most bloated planets have a mass of about  $60 M_{\oplus}$ . Observationally, the most bloated planets have in contrast masses larger than  $100 M_{\oplus}$ . This reflects that using the empirical bloating model of [Thorngren & Fortney \(2018\)](#) for planets of any



mass leads in our model to a M-R relation that differs from the observed one. There could be several reasons for this: the actual physical bloating mechanism has a mass dependency (or dependency on a parameter linked to the mass like the magnetic field strength or the metallicity) not accounted for in the empirical model which was derived mostly on giant planets. The discrepancy could also be due rather to the evaporation model in the sense that atmospheric loss for bloated  $\sim 60 M_{\oplus}$  planets is more efficient than predicted by our evaporation model. This would reduce the radii of these planets. The morphology of the close-in population will be studied further in a dedicated NGPPS publication.

The presence of multiple embryos in a disc lead to more diverse formation tracks, as we have seen in Fig. 10. This is reflected in the bigger spread of radii and envelope mass fraction for a given mass. The spread works in both directions. Planets in the multi-embryos population can have higher envelope contents, for instance, at  $M = 10 M_{\oplus}$ , the largest radius is around  $5 R_{\oplus}$  for the single embryo population, whereas for 100-embryos case, planets can have radii up to  $8 R_{\oplus}$ . Planets in the single-embryo population have a smooth formation and similar tracks for similar final positions and masses. It follows that these planets have similar core masses, which limits the core-mass effect. The 100-embryos, however, has two more effects that can change the core mass fraction in opposite ways, that we will now discuss in more details.

The first effect to alter the core mass fraction in the 100-embryos population is the competition for solids. We have discussed in Sect. 4.1 that giant planets in the multiple-embryos populations have lower core masses. This is reflected in the mass-radius diagram by the difference in the core mass fractions. For instance, there are no planets with a core mass fraction of less than 1 % in the single-embryo population while we frequently obtain this value in the 100-embryo population for planets above  $10 M_{\oplus}$ . Even if the radius is only weakly dependent on the metallicity for these masses, we see that the maximum radius for the non-bloated planets is slightly larger in the 100-embryos population.

The second effect is giant impacts. Due to their random nature, they add some spread to the core mass fractions at a given mass. Some planets suffered from collisions with other bodies relatively late during their formation. These collisions can lead to the loss of a significant part of the planet's envelope. As consequence, these planets have a lower envelope mass fraction for a given mass, because there is no loss of solids during such an event. Thus a collision does not simply reset the planet back to an earlier time, rather it can induce an increase of the bulk metallicity. We can see examples of such planets at intermediate masses. In the single-embryo population, the minimum envelope mass fraction increases significantly starting with about  $40 M_{\oplus}$  and no planet more massive than that values remains without envelope. In the 100-embryos however, we have several examples of planets with higher masses that exhibit small radii, including one roughly  $70 M_{\oplus}$  core without any surrounding envelope. One can also note a few giant planets in the 100-embryos population that have smaller radii than in the single embryo case. They are also caused by giant impacts.

The timing of the collision is important. Early-on events when the gas disc is still important, may even lead to a more massive envelope that there could have been if no collision occurred, because the collision enables the core to cool more efficiently, thus increasing the gas accretion rate (Broeg & Benz 2012). Otherwise, the lack of a gas reservoir prevents the re-accretion of an envelope, namely when the collision occurs during the late stage

of the gas disc presence in which case the envelope will not grow back to its previous mass.

Collisions are also the reason why there is a more extended range of planet masses without any envelope in the populations with multiple embryos per system. In the single embryo population, where only atmospheric escape acts, there are no planets without an envelope past  $40 M_{\oplus}$ , whereas we do have such cases in the other populations.

## 5.2. Distance-radius plot

In Fig. 17 we show the population NG73 (single embryo) and NG76 (100 embryos initially) as they would appear to transit and direct imaging surveys, that is, by showing the planes of orbital distance versus radius and apoastron distance versus absolute H magnitude.

A first major goal of the New Generation Planetary Population Synthesis was to predict directly and self-consistently all important observable characteristics of planets in multi-planetary systems, and not only masses and orbital elements as in previous generations of the Bern model. To achieve this, we have included (see Paper I) in the Generation 3 Bern model the calculation of the internal structure of all planets in all phases, in particular also in the detached phase, which was not done in Generation 2. We have also coupled the formation phase to the long-term thermodynamic evolution phase (cooling and contraction) over Gigayear timescales. With this we can predict also the radius, the luminosity, and the magnitudes for each planet, from its origins as a  $0.01 M_{\oplus}$  seed to potentially a massive deuterium burning super-Jupiter. In this way it becomes possible to compare one population to all major observational techniques (radial velocity, transits, direct imaging, but also microlensing). These techniques probe all distinct parameters spaces of the planetary populations, and thus constrain different aspects of the theory of planet formation and evolution. Taken together, they lead to compelling combined constraints, and help to eventually derive a standard model of planet formation that is able to explain all major observational findings for the entire population, as opposed to a theory that is tailored to explain a certain sub-type of planets, but fails at other planets.

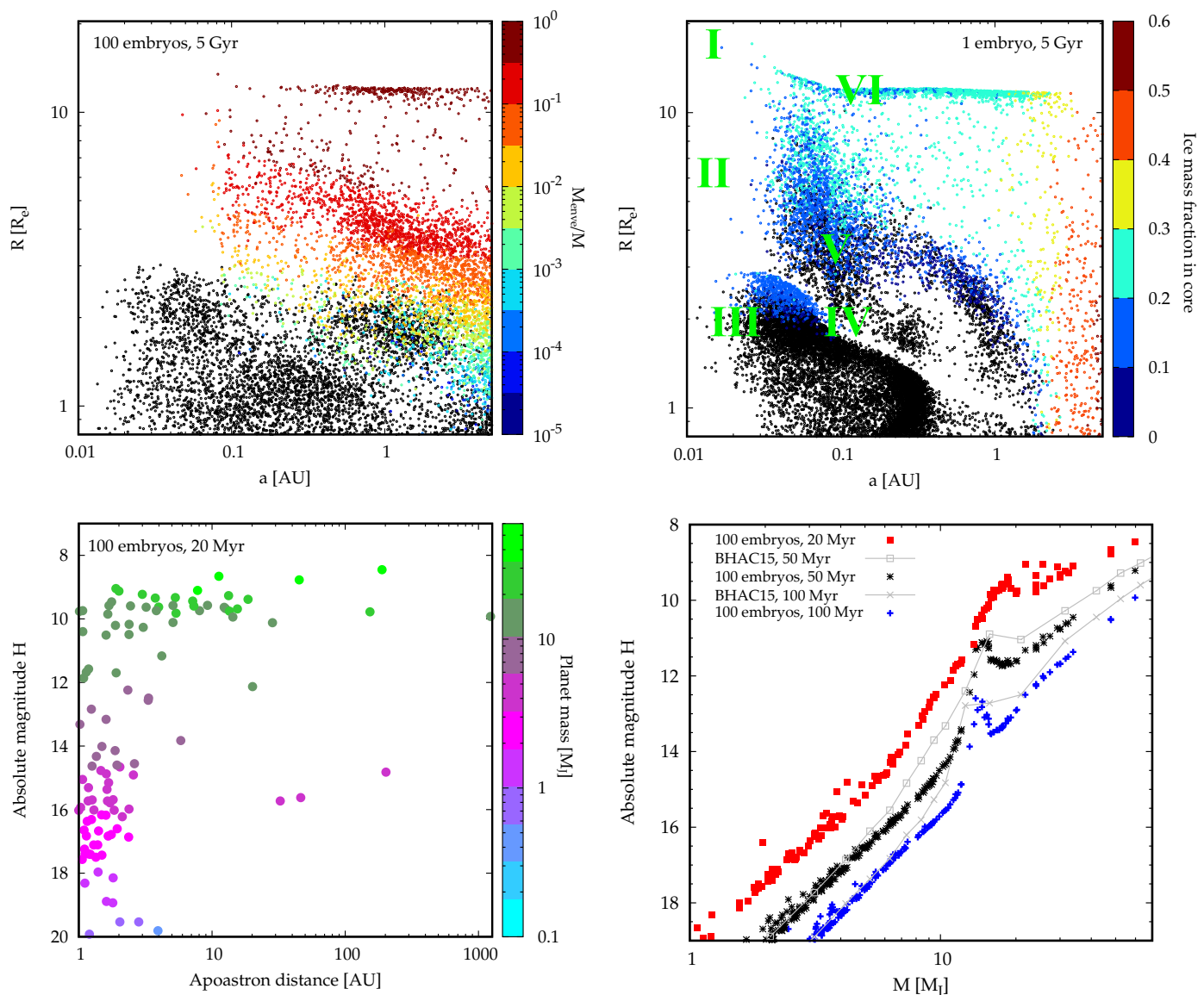
A second major goal of the new generation population synthesis was to be able to simulate planets ranging from Mars mass to super-Jupiters, and from star-grazing to very distant, or even rogue planets. For close-in planets for which the stellar proximity strongly influences the evolution, this meant that we had to include the effects of atmospheric escape, bloating, and stellar tides. Only then it becomes possible to meaningfully link formation and observations at an age of typically several Gigayears.

The top panels of the figure show the  $a - R$  diagram of the two populations, at an age of 5 Gyr. The quantitative description of the radius distribution, the formation tracks leading to the radii, and the statistical comparison to transit surveys will be the subject of a dedicated NGPPS paper (Schlecker et al. in pressb, see also Mulders et al. 2019), therefore we here only give a short qualitative overview.

In the right plot, the roman numerals shown important morphological features of the close-in population.

(I) are the bloated hot Jupiters. The bloating model is the empirical model of Thorngren & Fortney (2018), leading to an increase of the radii with decreasing orbital distance inside of about 0.1 au (Demory & Seager 2011; Sestovic & Demory 2020).

(II) is the (sub)Neptunian desert, which is an absence of very close intermediate mass planets that was observationally char-



**Fig. 17.** The synthetic populations in the eye of transit and direct imaging surveys. *Top left:* distance-radius plot of the 100-embryos population at 5 Gyr. The color code gives the H/He envelope mass fraction. Planets without H/He are shown in black. *Top right:* distance-radius plot of the single-embryo population, also at 5 Gyr. Roman numerals show important morphological features. The color code shows here the ice mass fraction in a planet’s core. Black points are planets that do not contain any ice. *Bottom left:* Apoastron distance versus absolute magnitude in the H band for the 100-embryos population at an age of 20 Myr. The colors show the planets’ mass. *Bottom right:* H band absolute magnitude as a function of planet mass and age. The isochrones of Baraffe et al. (2015) are also shown with grey lines.

acterized for example by Lundkvist et al. (2016), Mazeh et al. (2016), or Bourrier et al. (2018). It is likely a consequence of atmospheric escape (e.g., Lecavelier Des Etangs 2007; Kurokawa & Nakamoto 2014; McDonald et al. 2019). In the plot, it is not very well visible, but the hot Jupiters “above” are indeed found to down to smaller orbital distances than the intermediate mass planets.

(III) corresponds to the the hot and ultra hot solid planets like Corot-7b (Léger et al. 2009) or Kepler-10b (Batalha et al. 2011).

(IV) is the evaporation valley (Fulton et al. 2017) which was predicted theoretically by several planet evolution models including atmospheric escape (Owen & Wu 2013; Lopez & Fortney 2013; Jin et al. 2014). Super-Earth planets below the valley have lost their H/He as the temporal integral over the stellar XUV irradiation absorbed by these planets exceeded the gravi-

tational binding energy of their envelope in the potential of the core (Mordasini 2020).

(V) are the Neptunian and sub-Neptunian planets above the valley. They appear numerous in the single embryo population, but less so in the 100-embryos population. In the analysis of Mulders et al. (2019) of an earlier generation of the Bern model, it was found that this class is the only one occurring with a significantly different rate (a lower one) than inferred observationally from the Kepler survey.

(VI) are the giant planets. In the synthetic population, outside of about 0.1 au (that is, where no bloating is acting), the giant planets lead to an almost horizontal, thin pile-up of radii (but note the logarithmic y-axis). This concentration is the consequence of the following Mordasini et al. (2012b): the mass-radius relationship in the giant planet mass range has a maximum at around 3 Jovian masses, and is relatively flat. This

causes many planets from a quite wide mass range to fall in a similar radius range, close to  $1 R_{\oplus}$ . In the synthetic population, this concentration effect is artificially accentuated: during both the formation and evolutionary phase, the molecular and atomic opacities (from [Freedman et al. 2014](#)) correspond to a solar-composition gas, identically for all planets. In reality, the atmospheric compositions and thus opacities differ, inducing via different contraction timescales ([Burrows et al. 2007](#)) a certain spread in the mass-radius relation that cannot occur in the synthesis. Similarly, in reality planets do not have all exactly the same age.

Several of these features are also visible in the top left panel showing the 100-embryos population, albeit often in a less clear way. This is a consequence of the stochastic nature of the  $N$ -body interactions. Giant impacts that strip the H/He envelope are an additional effect that is important for the radii that cannot occur in the single embryo population. Two consequences of giant impacts are obvious: first, they populate the evaporation valley with cores that would otherwise be too massive for the envelope to be lost only via atmospheric escape. The fact that the valley appears rather too blurred in the 100 embryo population compared to observations ([Fulton et al. 2017](#); [Petigura et al. 2018](#)) could be an observational hint that impact stripping might be overestimated in the model and should be improved in further model generations, for example along the lines of [Denman et al. \(2020\)](#). Second, in the 100 embryo population, in the group of planets at around  $a = 1$  au and with radii between about  $1.5$  and  $2 R_{\oplus}$  (which is above the evaporation valley), there is a region of mixed planets with some possessing H/He, and others without it. In the single embryo population, all planets above the evaporation valley possess in contrast H/He envelopes. The black points in the 100-embryos population are thus the results of giant impact envelope stripping.

The quantitative comparison of the populations with transit surveys is as mentioned beyond the scope of this overview paper, but we note that many similar features are also found in the observed population. This reflects that the Generation III Bern is in contrast to older model generations able to simulate the formation and evolution also of close-in planets that are observationally particularly important.

### 5.3. Distance-magnitude plot, mass-magnitude relation and giant planets at large orbital distances

While transit surveys probe the planetary population at close-in orbital distance, direct imaging surveys like GPIS ([Nielsen et al. 2019](#)), NACO LP ([Vigan et al. 2017](#)), or SPHERE SHINE ([Vigan et al. 2021](#)) probe young giant planets at large orbital distances.

In the bottom left panel of Figure 17, the population with 100 embryos (NG76) is shown in the plane of apoastron distance versus absolute magnitude in the H band. The magnitude was calculated using the AMES COND atmospheric grid ([Allard et al. 2012](#)) assuming a solar-composition atmosphere. Magnitudes are a strong function of planet mass, therefore they accentuated massive giant planets.

Note that similarly to the radii, also here, the magnitudes were not obtained from some pre-computed mass-time-luminosity (or magnitude) relation or fit, but from solving the planetary internal structure of all planets during their entire “life”, i.e., from a planet’s birth as a lunar-mass embryo to present day, possibly as a massive deuterium-burning planet. To the best of our knowledge, the Generation III Bern model is currently the only global model predicting self-consistently besides

the orbital elements and masses also the radii, luminosities, and magnitudes.

The plot shows that the synthesis only predicts few giant planets outside of about 5 AU. The main cause for this absence is rapid inward migration, explaining why in the single embryo population, there are no giant planets at all outside of about 3 au (Fig. 7). As mentioned above (see the formation tracks of group I in Fig. 10), in the multiple embryo populations, giant planets at larger orbital distances are the result of violent scattering events among several concurrently forming giant planets ([Marleau et al. 2019](#)). Such events take preferentially place in very massive and metal-rich discs, explaining why the distant giant planets are massive (about 3 to several ten Jovian masses), in particular more massive than the “normal” giants in the pile-up at about 1 au. This is reflected in the apoastron-magnitude plot by an absence of distant planets with higher magnitudes (i.e., fainter planets). Compared to the mass-distance plot, the clustering is amplified by another effect ([Mordasini et al. 2017](#)): we see that there is a pile up of planets that have similar magnitudes of 11 to 9. To understand this pile-up, we need to consider the bottom right panel showing the mass - H magnitude relation at 20, 50, and 100 Myr. This plot is equivalent to the mass-radius plot shown above in connecting fundamental observable characteristic (here mass and luminosity).

Besides the expected general decrease of the magnitude with mass, we also see that there is a bump in the relation. It is caused by deuterium burning which is modelled as described in [Mollière & Mordasini \(2012\)](#). At 20 Myr, the bump is centered at around  $20 M_{\oplus}$ , and at around  $13-15 M_{\oplus}$  at later times. Deuterium burning delays the cooling, and causes planets of a relatively large mass range (at 20 Myr about  $15$  to  $35 M_{\oplus}$ ) to fall in the same aforementioned magnitude range. This leads to the pile-up seen on the left.

In terms of the statistical properties and frequency of distant giants in the 100 embryo population, we find (see Table 7) that giant planets ( $>300 M_{\oplus}$ ) are found outside of 5, 10, and 20 au for only 3.5, 1.6, and 0.8% of all stars (compared to 18% for all orbital distances). For comparison, in the SPHERE SHINE survey, the observed fraction of stars with at least one planet with  $1 - 75 M_{\oplus}$  and  $a = 5 - 300$  au is  $5.9^{+4.5}_{-2.7}$  % ([Vigan et al. 2021](#)). In this paper, a statistical analysis of the NG76 population in the context of the SPHERE SHINE survey can be found. The distant synthetic planets are also eccentric (mean eccentricity of about 0.4-0.6), found around high [Fe/H] stars (mean: 0.2-0.3), and their multiplicity is unity, i.e., there is only one distant giant planet. They do, however, often have another massive companion closer in. For example, of the 8 giant planets with  $a > 20$  au in the population, 5 have a giant companion inside of 5 au. These properties are all signposts of the violent formation pathway of these planets.

In the future, comparisons with direct imaging surveys should include besides the planet frequency such architectural aspects and also that due to different formation histories, at a given moment in time, there is no unique mass - magnitude conversion ([Mordasini et al. 2017](#)) as it is traditionally often assumed. This is again visible in the bottom right plot, where the magnitudes as a function of mass obtained in the synthesis are compared to the well known [Baraffe et al. \(2015\)](#) models. They start from arbitrary hot initial conditions. The general trend is as expected similar in the two cases, and the magnitudes are very similar at lower masses  $<5 M_{\oplus}$  at 50 and 100 Myr, but above there are differences of almost 1 mag. The peak caused by D-burning is clearly sharper in our simulations. This is partially due to the coarse sampling in the [Baraffe et al. \(2015\)](#) tables,



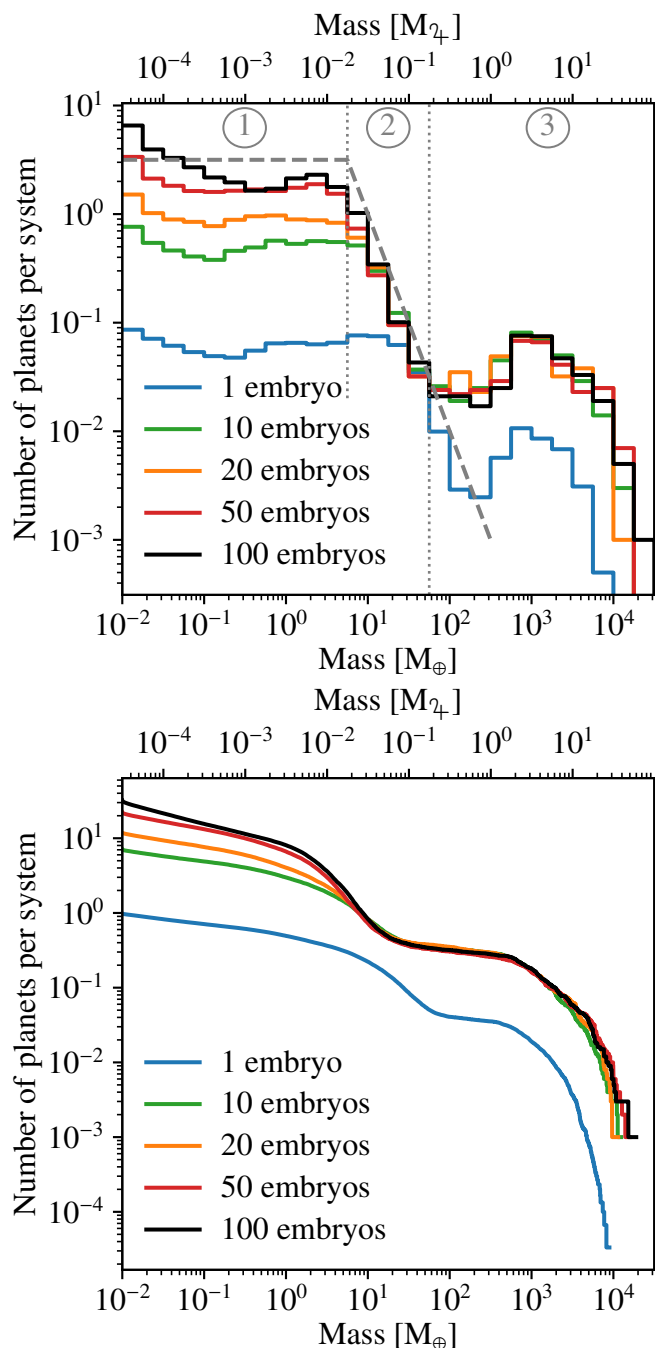
but not only. This could affect that analyses of direct imaging surveys. We also see the intrinsic spread in the self-consistent model population model especially at young ages which comes from the different formation histories. The spread includes now in particular also the effects of giant impacts. The spread means that there is no 1-to-1 conversion from magnitude to mass, even if all other complexities (like cold vs. hot start, atmospheric composition, clouds etc.) would be solved. At 20 Myr, the spread induces a fundamental uncertainty in the mass-magnitude relation of maximum  $1 M_{\text{Jup}}$  at lower masses without D-burning. In the mass range where D burning occurs, the impact is much larger, inducing an uncertainty of up to  $10 M_{\text{Jup}}$ .

## 6. The planetary mass function (PMF)

The prediction of the planetary mass function (PMF; [Ida & Lin 2004a](#); [Mordasini et al. 2009b](#)) is a fundamental outcome of any population synthesis. The PMF is a key quantity because of its observability and because it bears the imprint of the formation mechanism. We show the PMF and its reverse cumulative distribution of the different populations in [Fig. 18](#). Both give the average number of number planets per systems, i.e. the total number of planets divided by the number of systems in each population. The intersection of the curves with the left axis gives the average number of planets per system in each population. In the single-embryo population, the number is close to one, as all planet reach the end of the formation stage and can only be lost by tidal migration during the evolution phase. In the multi-embryos populations, giant impacts lead to the loss of embryos, especially which is especially important in the populations with the largest initial number of embryos. For instance, in the 100-embryos population, on average, only 32 embryos per system reach 5 Gyr. To improve clarity, we will stick to the same colour code throughout the remainder of this article when comparing the populations: black curve denotes the 100-embryos population, red the 50-embryos population, orange the 20-embryos population, green the 10-embryos population, and blue the single-embryo one.

When comparing the overall results, we may divide the relative behaviour in three different regions, as shown with the grey dotted lines on [Fig. 18](#). Region 1, a relatively flat region in the histogram, with its upper boundary depending on the population: about 5 to  $10 M_{\oplus}$  for the multi-embryos populations and  $30 M_{\oplus}$  for the single-embryo population. Region 2 shows a drop in the occurrence rates, up to the  $50 M_{\oplus}$  where the cumulative distribution indicates that we have an increased percentage of planets with the population with 50 embryos compared to the one with 20. Finally, region 3 of the giant planets, where there is first a minimum of occurrence rate at about  $200 M_{\oplus}$  followed by a local maximum located in the  $1000\text{--}2000 M_{\oplus}$  range.

In the first region, the increase of the number of embryos results in the corresponding increase of planets, there is thus virtually no other effect occurring. The only different is the end of this region, which gradually tends toward lower masses as the number of embryos increases. For the population with a single embryo, we observe a steeper drop of the cumulative curve in the  $20\text{--}80 M_{\oplus}$  range. Planets that are contributing to this feature are actually located at the inner disc edge; these are planets that migrated inward without accreting substantial material during their migration. Furthermore, we note that the first bin in the histogram has a greater value than the other ones; this is due to the far out embryos that do not grow, or only very little during the formation process.

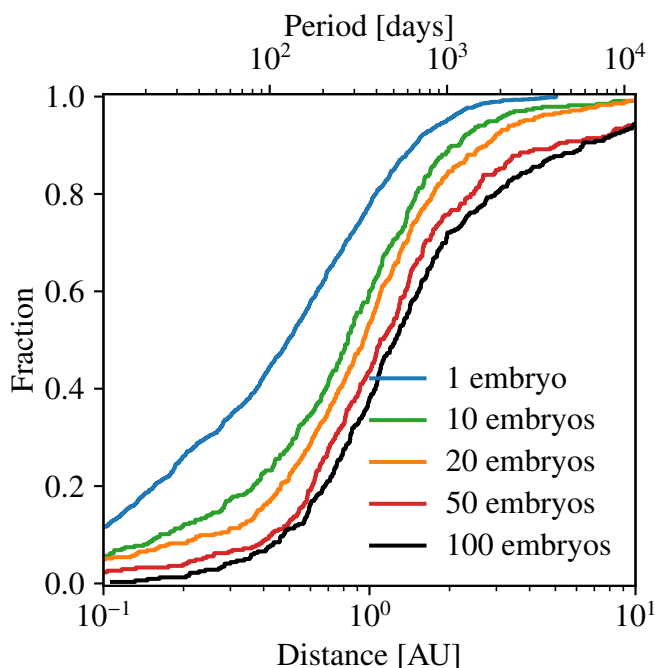


**Fig. 18.** Histogram (*top*) and reverse cumulative distribution (*bottom*) of the planet masses for the four populations presented in this study. The values are normalised by the number of systems in each population. Only planets that reached the end of the formation stage are counted; the maximum number of planets per system (the top left ending of the cumulative curves) can then be lower than the initial number of embryos.

### 6.1. Independence on the number of embryos for the giant planets

For the planets above  $10 M_{\oplus}$ , the mass function shows limited variations in all the populations with multiple embryos. The highest mass achieved in each populations shows a trend with the number of embryos. Except for that, the results we obtain are robust. This includes the common slope in the histogram





**Fig. 19.** Cumulative distribution of the distance of the giant planets (mass greater than  $300 M_{\oplus}$ ) for the five populations presented in this study. The higher the number of embryos, the more distant the giant planets.

for masses below  $100 M_{\oplus}$  and the “planetary desert” (Ida & Lin 2004a) for planets around  $200 M_{\oplus}$ .

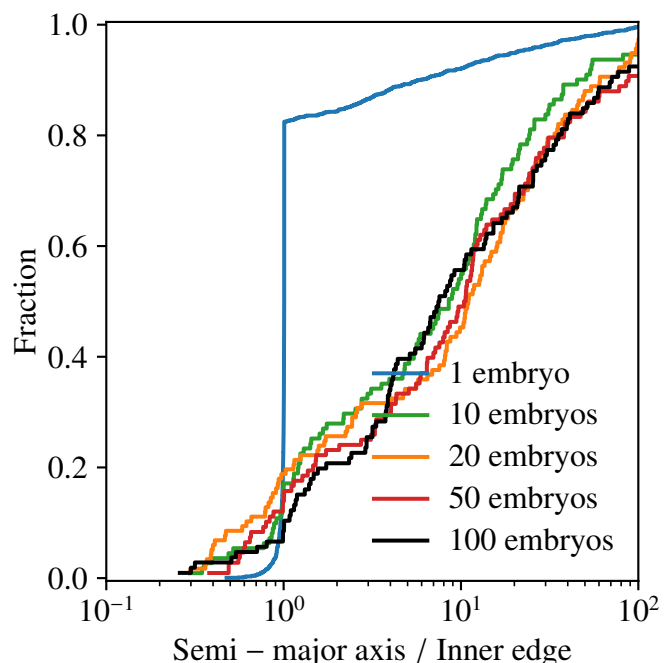
Thus, to obtain a mass function for planets above  $\sim 50 M_{\oplus}$ , the number of embryos is unimportant. The single-embryo population show an overall lower number of planets, but this is due to missed opportunities to form giant planets because it is unlikely that the embryo will start at a location which is needed to form these planets. Applying a correction factor on the outcome of that population is also a possibility to retrieve the mass function obtained from multi-embryos populations while limiting the computational needs (because the  $N$ -body is the most resource-intensive part of the model). We will use this to study the effects of the model parameters in the subsequent papers of this series.

## 6.2. Location of the giant planets

However, while the mass function of the giant planets is similar between the populations, the location of the giant planets is not. We find that there is a steady increase in the distance as the number of embryos grows. To illustrate this effect, we provide cumulative distributions of the giant planet’s distances for the different populations in Fig. 19. Also, both the 50- and 100-embryos population have 5% of the giant planets beyond 10 au.

Nevertheless, all the populations show a similar pattern in the distribution of these planets. We observe a pileup of planets around 1 au, which is consistent with results that suggest a maximum occurrence rate close to the ice line (Fernandes et al. 2019). In our populations, the median location of the ice line is at 2.81 au, while the median location of the giant is 0.49, 0.83, 0.95, 1.10 and 1.26 au in the 1, 10, 20, 50 and 100-embryos populations respectively. The giant planets are further in than the ice line, which is caused by the gas-driven migration.

We note that there are two causes for this change. First is the reduction of the importance of migration. We have already discussed in Sects. 4.4 and 4.5 that in the 100-embryos population



**Fig. 20.** Cumulative distribution of the location of the planets between  $30$  and  $300 M_{\oplus}$  with respect to the inner edge of the disc for the five populations presented in this study. If only one embryo per disc is present, more than 80% of all planets in this mass range end up at the inner edge of the gas disc.

the final location of the planets is closer to the starting location of the embryos than in the single-embryo case. The second cause is the increase of the close-encounters that put planets on wide orbits. This effect is responsible for the increase of the distant planets.

All the populations have a similar percentage of planets in the region between 0.7 and 2 au. The differences remain in the inner or outer locations, where the populations with a higher number of embryos have more planets beyond 2 au, while the populations with less embryos have more planets inside 0.7 au. Thus, the number of planets in the middle region, between 0.7 and 2 au is independent of the initial number of embryos. This feature will be useful, as it allows to study this region using populations with a limited number embryos, which are less computationally expensive.

It was discussed in Sect. 4.1 that the single-embryo population exhibit a different accretion pattern than the multi-embryos populations. In the former case, only very massive cores ( $\geq 50 M_{\oplus}$ ) can undergo runaway gas accretion because the luminosity due to the accretion of solids does not drop during the inward migration. This means many planets will end up at the inner edge of the disc. To illustrate the effect, we show in Fig. 20 the location of the planets in the  $30$  to  $300 M_{\oplus}$  range, normalised to the inner edge of the gas disc. It can be observed that for the single-embryo population more than 80% of the planets are located within or at the inner edge of the gas disc, while we see no special pile-up of planets at the inner edge for the other populations. For the multi-embryos populations, unlike for the giant planets, we do not obtain any systematic shift between the populations. They are also closer-in, with the median distance being 0.6 to 0.8 au.

### 6.3. A common slope for medium-mass planets

For planets between 5 and 50  $M_{\oplus}$ , it can be seen that all populations show the same behaviour in the histogram. To highlight this point, an additional dashed grey line with a slope  $\partial \log N / \partial \log M = -2$  has been superimposed. Here  $N d \log M$  is the number of planets whose masses are between  $\log M$  and  $\log M + d \log M$  (the bin sizes being constant in the logarithm of the planet mass). This corresponds to  $N \propto M^{-2}$ , as well as  $P \propto M^{-2} + C$ , where  $P$  is the total number of planets whose masses are larger than  $M$  (the cumulative distribution) and  $C$  a constant of integration. In the case of the single-embryo population, the mass range where the number of planets is similar to the ones of the other populations is limited to masses above  $\sim 30 M_{\oplus}$ , but on the other hand the distribution follows the line to larger masses, up to  $200 M_{\oplus}$ .

The bulk of the planets in this range are those that migrated close to the inner edge of the gas disc due to the fast inward migration in this mass range. There are two effects that are needed to obtain this peculiar slope: 1) gas accretion and 2) planetary migration.

Concerning the first point, these planets hold a significant amount of H/He, although they have not undergone runaway gas accretion. In the majority of these planets, the dominant component is the core. This slope cannot be achieved only with solids accretion, because without gas, the most massive planets in this range cannot be reproduced. Nevertheless, this is the range where the most massive planets would be found, were it not for gas accretion.

The second effect is planetary migration as the planets that cause the slope are almost all located inside 0.2 au. In the in-situ case (without any migration at all), the decrease in the occurrence rate begins before  $5 M_{\oplus}$ , because planets can accrete only up to their isolation mass. With migration however, planets can access a larger mass reservoir. Changing the planetary migration prescription in the models also affects the slope. It is however unclear to us the mechanism that causes this precise slope.

In the four multi-embryos populations, the end location where the common slope is encountered is similar, but not for the beginning location. The single-embryo population is different, first because it starts to follow the slope at higher masses (about  $30 M_{\oplus}$ ) and second because the end is also for larger masses, at about  $300 M_{\oplus}$ . As we mentioned before, the slope only occurs where planets are core-dominated. What is different with the single-embryo population with respect to the others is the different behaviour of accretion and migration, as we saw in Sect. 4.1. The resulting planets are mostly located at the inner edge of the disc (Fig. 20). This being the case, the maximum gas accretion of those planets remains low, as the gas surface density is low (and as we use the Bondi rate to compute the maximum gas accretion rate, Paper I). This explains the shift to larger masses for the change in behaviour of the single-embryo population compared to the multi-embryos one. This interaction, which can also be seen as a competition for solids, can shift the location of where this common slope is found, but will not change it fundamentally.

### 6.4. Convergence for the lower masses

For small masses, the histograms flatten, which is the expected behaviour with our setup. To highlight this, let us remember that the initial surface density of solids follows  $\Sigma_s \propto r^{-\beta_s}$  and define  $b$  the half-width of the feeding zone given in terms of the Hill radius (Paper I). Then, let us assume that all bodies grow to their

isolation mass,

$$M_{\text{iso}} = \frac{(4\pi b r^2 \Sigma_s)^{\frac{3}{2}}}{(3M_{\star})^{\frac{1}{2}}} \propto r^{\frac{3}{2}(2-\beta_s)} \quad (3)$$

(Lissauer 1987). As we place the embryos with a uniform probability in the logarithm of the distance  $r$ , we have  $dP \propto d \log r$ . Substituting for the isolation mass, we have  $dP \propto (3/2)(2 - \beta_s) d \log M_{\text{iso}}$ . So as long as  $\beta_s \neq 2$ , we have  $dP \propto d \log M_{\text{iso}}$ , i.e.  $dP \propto (1/M_{\text{iso}}) dM_{\text{iso}}$ . This relationship results in flat histogram when the bin sizes are uniform in the logarithm of the mass, as it is the case in Fig. 18.

There are other mechanisms affecting the mass distribution. For instance, not all planets will grow to their isolation mass, especially the ones at large separation. This results in the number of planets decreasing with the mass, as distant planets, the ones with the largest isolation mass, will need more time. Close-in planets, whose isolation mass is low, will have short accretion times compared to that of the protoplanetary disc and will not suffer from time constraints. However, this effect alone is not able to explain the shape of the distribution for the small-mass planets.

Another mechanism that will affect the distribution is planetary migration. The consequence is that embryos will have access to a larger reservoir of solids that they can accrete. As migration efficiency increases with the mass in the range under consideration here (e.g. Tanaka et al. 2002; Baruteau et al. 2014). This will lead to planets that would attain a mass of  $\sim 1 M_{\oplus}$  to migrate and have access to new planetesimals. This pushes the mass distribution toward larger values, which should tend to flatten the curve. The reduced occurrence rate of planets that occur at about  $10^{-1}$  (for the populations with 1, 10, 20, and 50 embryos) and  $1 M_{\oplus}$  (for the 100 embryos population) are due to planetary migration.

## 7. Planet types and system-level analysis

So far, we have only performed population-level analysis, disregarding the properties planets of planetary systems. Here, we will define different planet types (or categories). This allows to separate the diverse planets from our population and analyse them separately. In addition, this will help to quantitatively compare certain regimes of our populations with the known exoplanets.

The results presented in this section assume that all stars have had a protoplanetary disc during their formation. This assumption can lead to an overall overestimation of the frequency of planets; however observation results shown in Fig. 2 show that roughly 80% of stars have such a disc until about 2 Myr. Thus, while our models miss short lived discs, they represent a small fraction of the total and will not significantly affect the analysis presented here.

### 7.1. Definitions of planet categories

The planet categories were selected as follows: we first have a series that are constrained only by the planet masses: Earth-like planets are between 0.5 and  $2 M_{\oplus}$ , then super Earth up to  $10 M_{\oplus}$ , Neptunian to  $30 M_{\oplus}$ , Sub-giant to  $300 M_{\oplus}$  and giant above. In addition we also provide Deuterium-burning planets for masses larger than  $13.6 M_{\text{J}}$ , which overlaps with the giant planets category. The mass range of the sub-giants was chosen so that the category is located where the planetary desert discussed in

**Table 5.** Constraints for the different planet categories

Type	Min. mass [ $M_{\oplus}$ ]	Max. mass [ $M_{\oplus}$ ]	Min. dist. [au]	Max. dist. [au]
Mass $>1 M_{\oplus}$	1	...	...	...
Earth-like	0.5	2	...	...
super Earth	2	10	...	...
Neptunian	10	30	...	...
Sub-giant	30	300	...	...
Giant	300	...	...	...
D-burning	4322	...	...	...
Earth-like $<1$ au	0.5	2	...	1
super Earth $<1$ au	2	10	...	1
Neptunian $<1$ au	10	30	...	1
Sub-giant $<1$ au	30	300	...	1
Giant $<1$ au	300	...	...	1
Habitable zone	0.3	5	0.95	1.37
Kepler (Petigura et al. 2018)	see Eq. (4)	...	...	0.88
Kepler (Zhu et al. 2018)	see Eq. (5)	...	...	1.06
Hot Jupiter	100	...	...	0.15
Jupiter analogues	105.3	953.4	3	7
Giant $>5$ au	300	...	5	...
Giant $>10$ au	300	...	10	...
Giant $>20$ au	300	...	20	...

Sect. 6 is found in the multi-embryo population. We also set categories for the same masses, but for planets inside 1 au. The presence of the second series of categories is to avoid counting embryos that did not finish growth during the formation stage of our model (see the discussion in Sect. 4.3).

We defined planets in the habitable zone as planets between 0.3 and  $5 M_{\oplus}$  in mass and located between 0.95 and 1.37 au (Kasting et al. 2014). We also include two different category that relate to the Kepler’s observatory biases. The first follows Petigura et al. (2018), which contains planets with a period  $P < 300$  d (0.88 au) that also satisfy

$$\frac{R_{\text{tot}}}{R_{\oplus}} > 1.37 \left( \frac{P}{100 \text{ d}} \right)^{0.19}, \quad (4)$$

with  $R_{\text{tot}}$  the planet’s radius. The second follows Zhu et al. (2018), with planets that have a period  $P < 400$  d (1.06 au) and satisfy

$$\frac{R_{\text{tot}}}{R_{\oplus}} > 2 \left( \frac{a_{\text{planet}}}{0.7 \text{ au}} \right)^{0.31}. \quad (5)$$

Finally we have several categories related to giant planets: hot Jupiters have more than  $100 M_{\oplus}$  and are located within 0.1 au, Jupiter analogues have masses between  $1/3$  and  $3 M_{\text{J}}$ , and semi-major axis between 3 and 7 au, and three categories for giant planets (mass above  $300 M_{\oplus}$ ) further out than 5, 10 and 20 au. These categories were chosen to identify planets that lie outside of the bulk of giants. Such planets are prime targets for direct imaging surveys. For instance, there are no giant planets outside 5 au in the single-embryo population (see Fig. 7), so they ended up there because of planet-planet interactions in the multi-embryos populations. All these definitions are summarised in Table 5.

## 7.2. Occurrence rates and multiplicity as function of the number of embryos

One of the goal of this work is to determine the convergence of our formation model with respect to the initial number of embryos. For this, we provide the occurrence rates and the multiplicity of these categories of planets in Table 6. These quantities are computed as follows. The total number of systems in each population is  $N_{\text{sys,tot}}$ , whose value is 1000 in the multi-embryos population and 30 000 in the single-embryo population. The number of planet in each category is  $N_{\text{pla}}$  and the number of systems where at least one such planet is present is  $N_{\text{sys}}$ . From these, we define the occurrence rate  $o_{\text{p}} = N_{\text{pla}}/N_{\text{sys,tot}}$ , the fraction of systems harbouring such planets  $f_{\text{s}} = N_{\text{sys}}/N_{\text{sys,tot}}$  and the mean multiplicity of such planets  $\mu_{\text{p}} = N_{\text{pla}}/N_{\text{sys}}$ . It follows that  $o_{\text{p}} = f_{\text{s}}\mu_{\text{p}}$ .

A graphical representation of the values for the categories of planets as function of their masses for any location is provided in Fig. 21, while the same information for planets inside 1 au is provided in Fig. 22. In the latter case, the category of Deuterium-burning planets has been left out as it is always empty. In addition to the overall occurrence of these kinds of planets (as we discussed in Sect. 6 for the mass function), this gives additional insights on the distribution of planet types in the different systems.

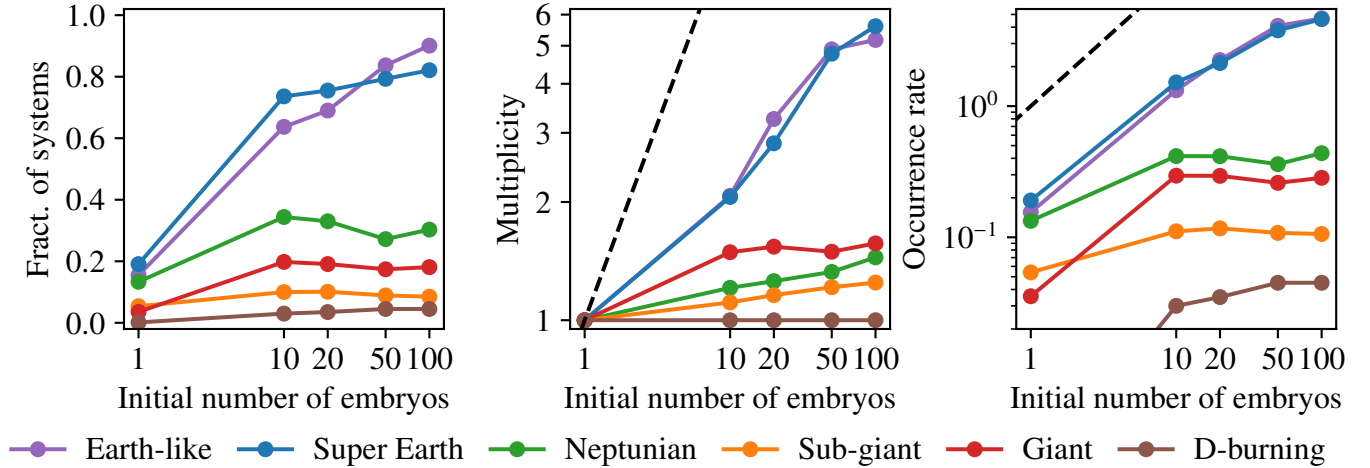
Overall, the results confirm what we discussed in the previous section: convergence is achieved with a smaller number of embryos for the most massive planets than for the lower-mass ones. In the low-mass range (habitable zone, Earth-like and Super-Earth planets) the trend is an increasing number of planets along with the number of embryos. As we already discussed in Paper I and Sect. 4.3, the growth of the planetary bodies is not finished for larger separation by the time our model switched from the formation stage to evolution. Thus, the bodies that are further out may not reflect the end state of planetary systems. For this reason, we also provide categories accounting bodies that are inside 1 au, where growth should be mostly finished at the end of the formation stage of our model (20 Myr), and whose result we show in Fig. 22. In that plot, we may note that the multiplicity of the Earth-like planets drops in the 100-embryos population compared to the 50-embryos one. This effect is related to how the growth of small-mass planets is followed up to a giant-impact phase only in the 100-embryos population. With less embryos, the planets do not disturb their orbits to the same extent, and the final phase of planetary growth via giant impacts is missing. This is corroborated by the median mass of the planets in this category: in the 20 and 50-embryos populations, these value is 0.95 and 0.96  $M_{\oplus}$  while in the 100-embryos population it increases to 1.14  $M_{\oplus}$ .

For the most massive planets (Neptunian, sub-giants, giants and Deuterium-burning) however, we obtain similar numbers in the populations that have at least 10 embryos. Nevertheless, we still see some trends. The first three categories (Neptunian, sub-giants and giants) have slight reductions in their fraction of systems as the initial number of embryos increase while the multiplicity slightly increases so that the overall number of such planets remain quite constant. On the other hand, for the last category (deuterium-burning) we observe first an increase of the occurrence rate along with the number of embryos. Then it becomes constant at 4.5 % for both the 50 and 100-embryos populations.

For the location of the giant planets, the different categories based on the separation show results that are consistent with what is shown on Fig. 19. The fraction of systems with Hot-Jupiters peaks for the 10-embryos population at 2.8 % of sys-

**Table 6.** Percentage of systems/stars with specific planetary types ( $f_s$ ) and their mean multiplicity ( $\mu_p$ ) for the different populations. The multiplicity is the mean number of planets of a given type per star for those stars which have at least one planet of this type. The multiplicity of the single-embryo population is not given because it is 1 per definition.

Type	Initial number of embryos									
	1	10		20		50		100		
	$f_s$	$f_s$	$\mu_p$	$f_s$	$\mu_p$	$f_s$	$\mu_p$	$f_s$	$\mu_p$	
Mass $>1 M_\oplus$	49.3 %	91.8 %	3.3	93.2 %	4.3	94.6 %	7.0	96.0 %	8.4	
Earth-like	15.5 %	63.7 %	2.1	69.0 %	3.3	83.7 %	4.9	90.1 %	5.2	
Super Earth	19.1 %	73.6 %	2.1	75.5 %	2.8	79.3 %	4.8	82.1 %	5.6	
Neptunian	13.3 %	34.4 %	1.2	33.0 %	1.3	27.2 %	1.3	30.3 %	1.4	
Sub-giant	5.4 %	10.0 %	1.1	10.1 %	1.2	8.9 %	1.2	8.5 %	1.2	
Giant	3.6 %	19.8 %	1.5	19.1 %	1.5	17.4 %	1.5	18.1 %	1.6	
D-burning	0.1 %	3.0 %	1.0	3.5 %	1.0	4.5 %	1.0	4.5 %	1.0	
Earth-like $<1$ au	12.9 %	52.8 %	1.8	53.3 %	2.9	58.7 %	3.7	57.2 %	2.8	
Super Earth $<1$ au	17.8 %	66.9 %	1.9	63.1 %	2.5	62.0 %	3.7	66.1 %	3.7	
Neptunian $<1$ au	13.1 %	32.8 %	1.2	30.9 %	1.2	24.8 %	1.2	26.2 %	1.4	
Sub-giant $<1$ au	5.1 %	7.7 %	1.0	6.9 %	1.1	6.7 %	1.1	6.5 %	1.1	
Giant $<1$ au	2.8 %	14.0 %	1.2	12.6 %	1.2	9.3 %	1.1	9.2 %	1.1	
Habitable zone	1.0 %	14.4 %	1.2	29.9 %	1.3	49.4 %	1.5	43.7 %	1.3	
Kepler (Petigura et al. 2018)	40.8 %	78.2 %	3.0	76.8 %	3.5	72.2 %	4.6	76.7 %	4.4	
Kepler (Zhu et al. 2018)	37.2 %	78.1 %	2.8	77.1 %	3.3	71.6 %	4.3	65.7 %	4.5	
Hot Jupiter	0.8 %	2.8 %	1.0	2.2 %	1.1	0.9 %	1.1	0.3 %	1.0	
Jupiter analogue	0.0 %	0.5 %	1.0	0.4 %	1.0	0.2 %	1.0	0.8 %	1.0	
Giant $>5$ au	0.0 %	0.7 %	1.0	1.2 %	1.0	2.6 %	1.0	3.5 %	1.0	
Giant $>10$ au	0.0 %	0.3 %	1.0	0.3 %	1.0	1.4 %	1.0	1.6 %	1.0	
Giant $>20$ au	0.0 %	0.1 %	1.0	0.1 %	1.0	0.9 %	1.0	0.8 %	1.0	

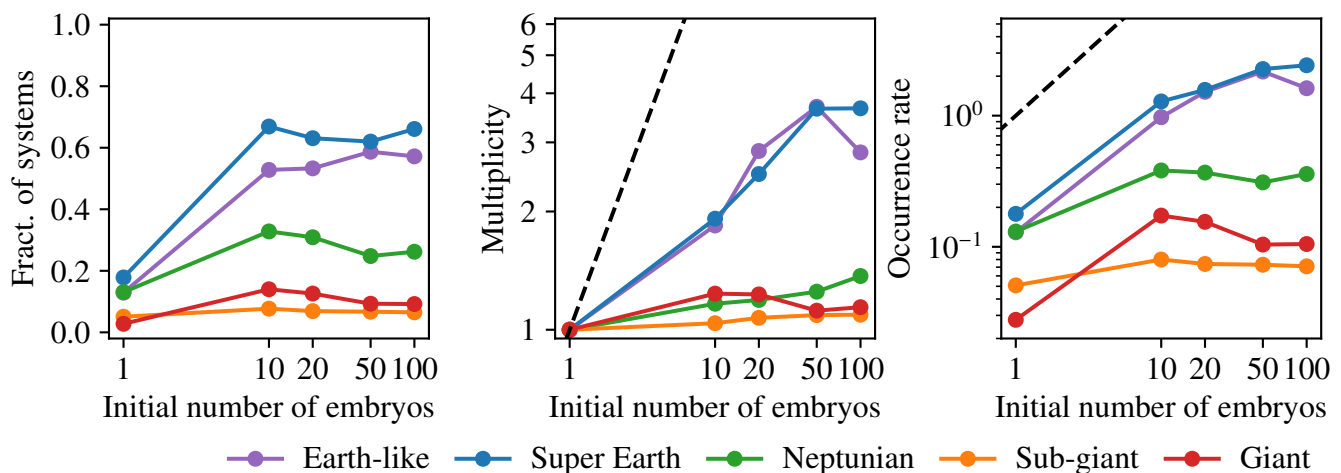


**Fig. 21.** Graphical representation of the fraction of systems (stars) containing at least one planet of this category ( $f_s$ ), multiplicity ( $\mu_p$ , mean number of planets of this category per star including only these stars with at least 1 planet of this category), and occurrence rate ( $o_p = f_s \mu_p$ ) as function of the number of embryos for six planet categories that depend on the masses. The underlying data is provided in Table 6. The dashed black lines in the two last panels show the identity function.

tems, down to 0.3 % for the 100-embryo population. For comparison, the observed occurrence rate of these planets is 0.5–1 % (e.g. Howard et al. 2012). Thus, only the 50-embryo with 0.9 % shows a value that is consistent with the observations. As we have discussed previously, the overall separation of the giant planets increase along with the number of embryos (see Sect. 4.2), so that 100-embryos population has very few inner giant planets. The decrease of the number of hot-Jupiters is consistent with the decrease of the efficiency of migration with increasing embryo number that we observed in Sect. 4.5, as the embryos forming hot-Jupiters come mostly from beyond the ice-line.

Conversely, for the most distant giants, their number increase along with the initial amount of embryos. The fraction Jupiter analogues increases, with an occurrence rate of up to 0.8 % in the 100-embryos population. Observational estimates for this class of planets are:  $(3.3 \pm 1.4)\%$  (Wittenmyer et al. 2011),  $(2.7 \pm 0.8)\%$  (Cumming et al. 2008) and  $\approx 3\%$  (Rowan et al. 2016). We thus obtain values lower than the observational results for this class, even for the 100 embryos population. The same increase with the initial number of embryos applies for the distant giant planets (beyond 5 au). It should be noted that there is a value that is the same for all categories and populations: there is never more than a single distant giant planet in any system.





**Fig. 22.** Graphical representation of the fraction of systems (stars) containing at least one planet of this category ( $f_s$ ), multiplicity ( $\mu_p$ , mean number of planets of this category per star including only these stars with at least 1 planet of this category), and occurrence rate ( $o_p = f_s \mu_p$ ) as function of the number of embryos for five planet categories that depend on the masses, but only accounting for the inner planets, i.e. inside 1 au. The underlying data is provided in Table 6. The dashed black lines in the two last panels show the identity function.

### 7.3. Multiplicity of the different types of planets

To investigate the distributions of multiplicities in a more detailed fashion than just the mean values shown in Figs. 21 and 22, and Table 6, we provide in Fig. 23 histograms of these for five types of planets. These categories are the five ones defined in Sect. 7.1 that have a mass criterion. For the first three (Earth-like, Super-Earth and Neptunian) we use the categories that are limited to planets inside 1 au while for the last two (Sub-giant and Giant) we selected the categories without restriction on the planet’s distance. This choice is consistent with our discussion about the lack of convergence for the smaller-mass planets at larger distances.

The results here are very similar to our previous discussion. For the giant and sub-giant categories, all the multi-embryos populations show a similar distribution. Although we do not show it, this is valid for both set of categories (all distances or only within 1 au). Thus, for the most massive planets, the number of embryos does play a role for the final multiplicity, as long as that number is at least around 10. This result is in line with A13. It can be noted that there are roughly the same number of systems with giant planets, that have a multiplicity of 1 and 2. This result is consistent with the results of Bryan et al. (2016) that half of the systems with a giant planet inside 5 au have a companion planet.

For Neptunian and super Earth planets inside 1 au, we also see that the distributions of multiplicity converge. The Neptunian category does not show much variation between the population, as for the sub-giant and giants planets. However, the convergence of the Super Earths is only achieved between the two populations with the most embryos per system. In the 10-embryos population a steady decrease of the number of systems for higher multiplicities, while in the populations with more embryos it is more likely to find systems with several such planets than lower counts. The Earth-like category shows a similar behaviour, except for the 100-embryos population. Here, the 100-embryos shows less systems with high multiplicity than the 50-embryos population. This is most likely related to the formation of the terrestrial planets that we discussed in Paper I and Sect. 4.3. Thus, for planets above  $0.5 M_{\oplus}$ , increasing the further the number of embryos would not increase the planet count further.

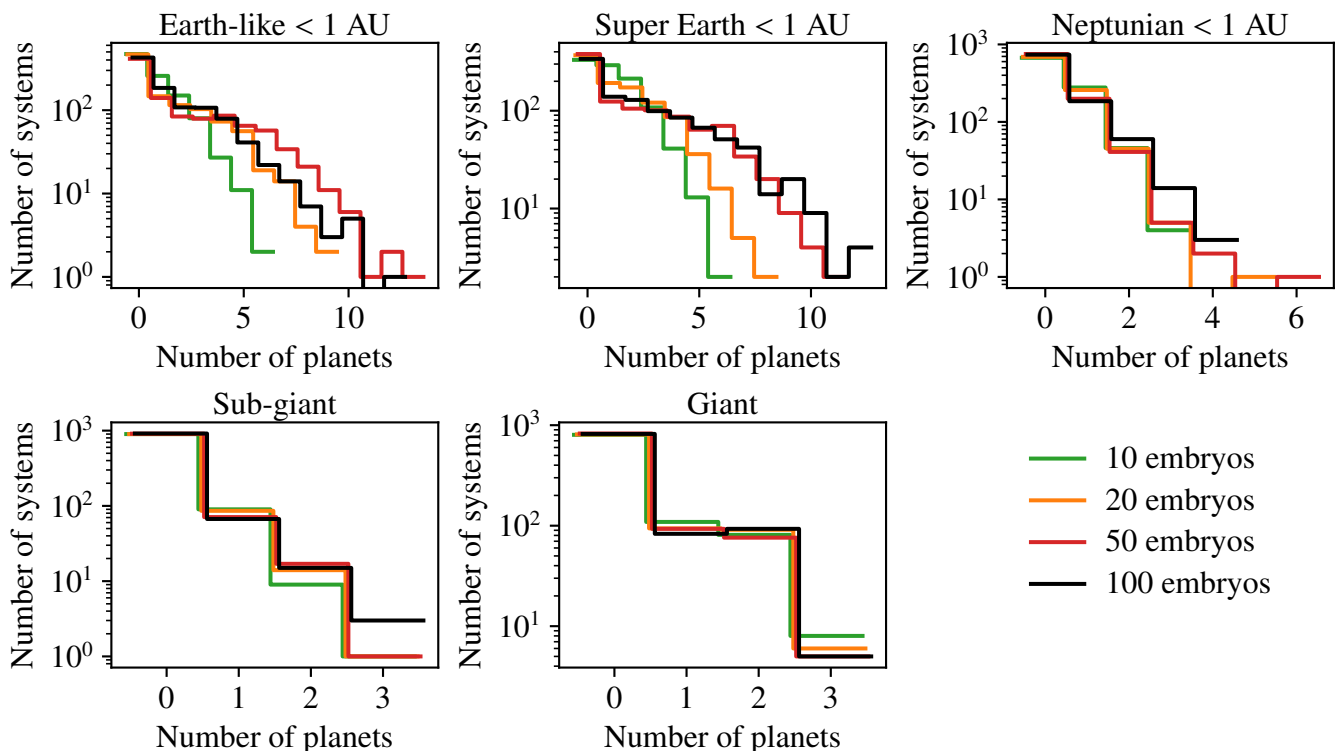
It should also be noted that unlike for the other categories or other populations, the Earth-like and super Earth categories in the 50- and 100-embryos populations show a plateau for the low-multiplicity counts. Here, the multiplicities between 1 and 3 have similar probabilities and they account for 32 % of the systems with Earth-like planets and 21 % of the systems with Super-Earths in the 100-embryos population.

In summary, we find that convergence for the overall multiplicity (that is, the total number of planets of a given type divided by the number of systems having such planets) is a good indicator for the convergence of the underlying distribution of multiplicities. The multiplicity of the sub-giant and giant planets at all locations are similar in all multi-embryos populations (though not their locations, see Sect. 6.2); the same applies for the Neptunian planets inside 1 au. For the inner Super-Earths, only the 50 and 100-embryos populations show similar results while for inner Earth-like planets, the 100-embryos population show a decrease of the multiplicity of the Earth-like planets. The 100-embryos population should be the only one used to analyse Earth-like planets.

### 7.4. Statistical results on the 100 embryos population

For the 100-embryos population, we provide key statistical characteristics of the different kinds of planets in Table 7, which constitutes the overall demographic predictions of our formation model. The column fraction of systems is the same as in Table 6. The mean [Fe/H] column denotes the mean host star metallicity of systems where the relevant kinds of planets are found. We provide an annotated graphical view in Fig. 24. This figure shows the same as the bottom right panel of Fig. 8, but the colouring has been removed and dot sizes go with the logarithm of the planets’ physical radii. Following the discussion Sects. 4.3 and 4.6, the Earth-like, super Earth, and Neptunian at all distances should be taken cautiously. For the two Kepler-related criterion that use the radius rather the mass, we provide their graphical representation in Fig. 25. We remind the reader that these (absolute) results depend on the model parameters, likely in a stronger fashion than general trends and relative correlations, as discussed at the beginning of Sect. 3 and in Sect. 3.11.

The values of the occurrence rate column for the “All”, “Mass  $> 1 M_{\oplus}$ ” and “Giant” categories give of the cumulative



**Fig. 23.** Histogram of the multiplicity of different types of planets, for the four multi-embryo populations presented in this study. In each panel, the curves are slightly shifted horizontally from another to make them more visible. We see for example that for the giant planets, out of the 1000 systems about 800 do not contain any giant planets. About an equal number (100 each) have one or two giants. Less than ten out of the 1000 systems contain 3 giants, which is the highest number per system that occurs.

distribution shown in the bottom panel of Fig. 18 at 0.01, 1 and  $300 M_{\oplus}$  respectively. Out of the initial 100 000 embryos (1000 systems with 100 embryos each), only 32 030 remain at 5 Gyr. Most of the embryos (63 124) were lost due to giant impacts, 2675 were ejected, 1869 ended in the central star following close-encounters during formation stage, 292 ended in the central star due to tidal migration in the evolution stage, and 10 were fully evaporated during the evolution stage. Thus, on average, 32 embryos per disc remain. But these are mainly embryos that did not grow, in outer parts of disc where accretion is very slow. Of the average of 32 embryos per disc that remain, only 8.4 have a mass larger than  $1 M_{\oplus}$ , as indicated by the “Mass  $> 1 M_{\oplus}$ ” category. For comparison, the solar system has five planets matching the same criterion (plus Venus that has a mass of  $0.8 M_{\oplus}$ ). The values are hence not different. The multiplicity is larger for systems with only terrestrial planets, as giants will usually lead to the removal of terrestrial planets (Paper I).

Most of the sub-giants are also found to be in the inner part of the disc, with 69% of them being within 1 au. These planets either form late or had their envelope ejected just before the dispersal of the gas disc to have a limited time in the runaway gas accretion. They spent then more time when their masses were in the  $10 M_{\oplus}$  range, which means they experienced more migration than giant planets that had to form quicker or terrestrial planets that are largely unconstrained by the life time of the protoplanetary gas disc. It can also be seen that the multiplicity of sub-giant planets is the lowest, as it is unlikely for two planets to be in the same situation in the same system.

The multiplicity of the distant giant planets is always unity. This means that in no systems we find two (or more) giant planets beyond 5 au. As we discussed Sect. 4.5, these planets mostly originate from seeds that were initially positioned within 10 au.

They are then moved to their final location by one or more close encounters with other massive planets. Out of those systems, nearly the half only have one giant planet remaining, that is, the one beyond 5 au while the other have one (or in one case, two) other giant planets further in. Nevertheless all these systems had two giant planets at some point, some of which were subsequently lost, mostly by ejections. The study of systems with giant planets will be the subject of further work (Emsenhuber et al. in prep., Paper XI).

The comparison between the inner categories and the others allows to recover some information about the location of these planets. Only few systems have multiple Sub-giant and giant planets inside 1 au, as we can see in Table 7. What we can learn in addition here is that these do not occur for systems with the highest metallicity, but rather for moderate values.

### 7.5. Effect of metallicity

The occurrence rate of giant planets is known to be correlated with the host star’s metallicity (Gonzalez 1997; Santos et al. 2004; Fischer & Valenti 2005). Lower-mass planets on orbits of less than 10 d are also preferentially found around metal-rich stars, but the correlation is weaker for other planets (Mulders et al. 2016; Petigura et al. 2018). This finding, particularly in the case of the giant planets, has been an argument to promote the core accretion paradigm, as the formation of a sufficiently massive core takes less time when more solids are present, leaving more time for gas accretion (Ida & Lin 2004b; Mordasini et al. 2009b; Wang et al. 2018).

The mean stellar metallicity of the systems harbouring the different kind of planets is provided in Table 7. For both sets of categories that depend of the planet masses (all distances and in-

**Table 7.** Properties of different planet kinds from the population with 100 embryos per system. See main text for the precise definition of the kinds.

Type	Number of planets	Number of systems	Fraction of systems	Occurrence rate	Multiplicity	Mean [Fe/H] ± std. dev.	Mean ecc. ± std. dev.
All	32030	1000	100.0 %	32.03	32.03	$-0.03 \pm 0.20$	$0.05 \pm 0.13$
Mass $>1 M_{\oplus}$	8065	960	96.0 %	8.06	8.40	$-0.02 \pm 0.20$	$0.05 \pm 0.10$
Earth-like	4660	901	90.1 %	4.66	5.17	$-0.04 \pm 0.20$	$0.07 \pm 0.12$
Super Earth	4603	821	82.1 %	4.60	5.61	$-0.03 \pm 0.19$	$0.04 \pm 0.08$
Neptunian	438	303	30.3 %	0.44	1.45	$0.05 \pm 0.17$	$0.05 \pm 0.09$
Sub-giant	106	85	8.5 %	0.11	1.25	$0.12 \pm 0.16$	$0.06 \pm 0.16$
Giant	284	181	18.1 %	0.28	1.57	$0.14 \pm 0.15$	$0.12 \pm 0.18$
D-burning	45	45	4.5 %	0.04	1.00	$0.19 \pm 0.15$	$0.14 \pm 0.24$
Earth-like $<1$ au	1618	572	57.2 %	1.62	2.83	$-0.09 \pm 0.18$	$0.07 \pm 0.11$
Super Earth $<1$ au	2421	661	66.1 %	2.42	3.66	$-0.01 \pm 0.18$	$0.03 \pm 0.08$
Neptunian $<1$ au	359	262	26.2 %	0.36	1.37	$0.05 \pm 0.17$	$0.05 \pm 0.10$
Sub-giant $<1$ au	71	65	6.5 %	0.07	1.09	$0.13 \pm 0.15$	$0.07 \pm 0.11$
Giant $<1$ au	105	92	9.2 %	0.10	1.14	$0.11 \pm 0.16$	$0.13 \pm 0.15$
Habitable zone	560	437	43.7 %	0.56	1.28	$-0.11 \pm 0.18$	$0.03 \pm 0.06$
Kepler (Petigura et al. 2018)	3344	767	76.7 %	3.34	4.36	$0.00 \pm 0.19$	$0.03 \pm 0.10$
Kepler (Zhu et al. 2018)	2934	657	65.7 %	2.93	4.47	$0.02 \pm 0.18$	$0.03 \pm 0.09$
Hot Jupiter	3	3	0.3 %	0.00	1.00	$0.19 \pm 0.05$	$0.07 \pm 0.04$
Jupiter analogues	8	8	0.8 %	0.01	1.00	$0.19 \pm 0.11$	$0.14 \pm 0.20$
Giant $>5$ au	35	35	3.5 %	0.04	1.00	$0.22 \pm 0.12$	$0.35 \pm 0.26$
Giant $>10$ au	16	16	1.6 %	0.02	1.00	$0.24 \pm 0.12$	$0.58 \pm 0.24$
Giant $>20$ au	8	8	0.8 %	0.01	1.00	$0.25 \pm 0.14$	$0.55 \pm 0.28$

side 1 au only), the mean metallicity increases with the masses. The means for Earth-like ( $-0.04$ ) and Super-Earths ( $-0.03$ ) planets are close to the one of the overall population ( $-0.03$ ), so it is for the inner Super-Earths. This means that there is almost no metallicity effect for these planet kinds. However, systems with Earth-like planets inside 1 au and habitable zone are more metal-poor ( $-0.09 \pm 0.18$  and  $-0.11 \pm 0.18$ ); these are the only two categories whose mean is lower than the one of the overall population. The mean of the systems with Neptunian and Sub-giant increase, but they are similar each for all distance and inner planets. This suggests that there are no dependency on the stellar metallicity for the location of these planets. Giant planets behave similarly, although the mean of the systems with giants inside 1 au is slightly lower than the one for all distances. On the other hand, the 3 hot-Jupiters have again a higher mean metallicity hosts than distant giant planets, as have the Jupiter analogues and those beyond 5 au. Our results are consistent with observational results for hot Jupiters (e.g. Buchhave et al. 2018; Moe & Kratter 2019) and distant, eccentric giants (Buchhave et al. 2018). However, we are unable to reproduce the observation of Buchhave et al. (2018) for cold and circular giants forming around stars with near-Solar metallicities. This suggests that another formation channel exists for this planets, such as pebble accretion of gravitational instability.

The trend of increasing stellar metallicity with planet mass continues to the brown dwarfs (deuterium-burning). This is compatible with the results of Adibekyan (2019), who found that the brown dwarfs can be explained by the core accretion paradigm, as we do in this work. They also found that it is possible for massive brown dwarfs to form around star with solar-like metallicity, but this is for more massive stars that we do not model in this work.

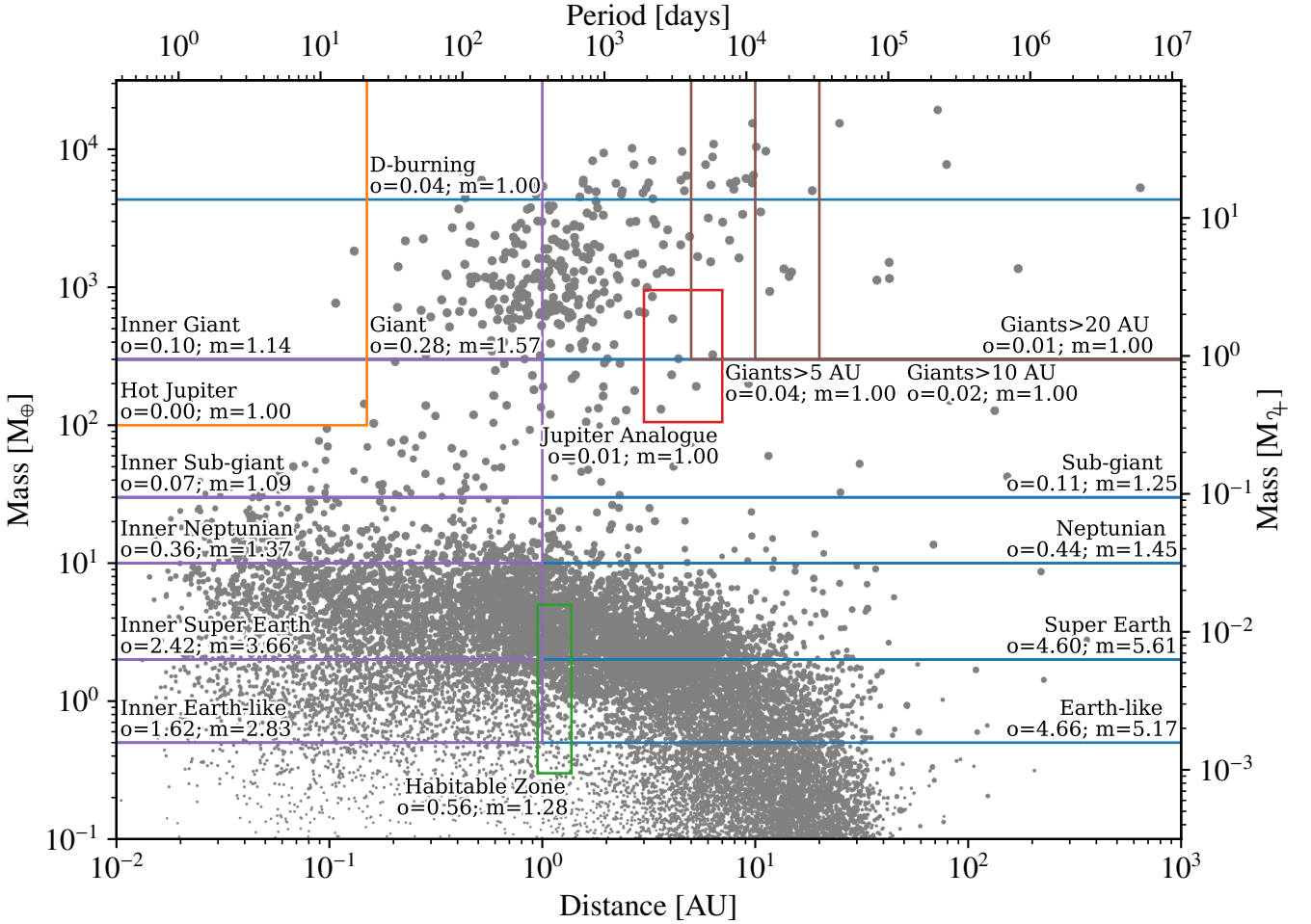
We also note that there is a trend of the metallicity for giant planets at intermediate and large orbital distances. The ones at larger separation are found still over more metal-rich stars than

the general population:  $0.22 \pm 0.12$  and  $0.25 \pm 0.14$  for the ones beyond 5 au and 20 au versus  $0.14 \pm 0.15$  otherwise. We remember that all these systems formed more than one giant planet, some of which were subsequently lost (see Sect. 7.4). Most of the distant giants were brought to their distant orbits after one or more close encounter with other massive planets. These encounters happen after the planets have undergone runaway gas accretion, though the planets may continue to accrete after being sent on wide orbits. Hence, it is necessary for multiple giant planets to form in a single systems for close encounters to strong enough to alter the orbits from  $\approx 1$  au to more than 20 au.

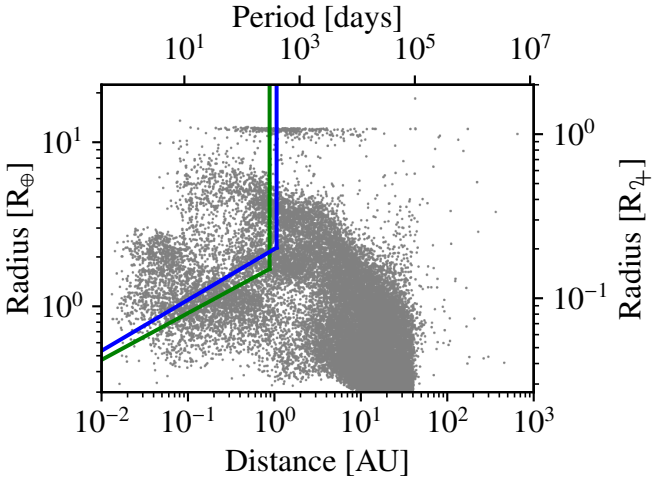
### 7.6. Correlation between multiplicity and metallicity

Another way to check for a metallicity effect is to look at the correlation between the numbers of certain types of planets as a function of the stellar metallicity. The results of this analysis of the 100-embryos population are provided in Fig. 26 for the categories encompassing all distances and Fig. 27 for the ones restricted to planets inside 1 au. The systems are divided in six equally spaced metallicity bins spanning metallicities from  $-0.6$  to  $0.5$  dex.

The results for the most massive planets exhibit the expected behaviour: the fraction of systems with massive planets (Neptunian, sub-giants, and giants) increases monotonically with stellar metallicity. The lowest-metallicity bin does not have any system with Neptunian planets or above. The second bins has some systems with Neptunian planets, very few systems with Sub-giants and none with giants. The next bins show a gradual increase of the fraction of systems with these kind of planets, with roughly the half of the systems in the highest metallicity bin. Additionally, we can see the dependency of the multiplicity on the metallicity. For the sub-giants, we observe that as the metallicity increase, there are first systems with only one such planet, and the further on systems with two and for a few systems even three



**Fig. 24.** Mass-distance diagram of the population with 100 embryos per system overlaid with the different planet categories. The point size is related to the logarithm of the planet’s physical size. The different categories provided in Table 6 have their boundaries marked, and the principal characteristics of the planets falling within each category are repeated:  $o$  denotes the occurrence rate and  $m$  the multiplicity.



**Fig. 25.** Radius-distance diagram of the population with 100 embryos per system overlaid with two Kepler-related categories. The green line shows the criterion following Petigura et al. (2018), while the blue line shows the criterion following Zhu et al. (2018).

appear, starting roughly with a solar metallicity. For the giant planets however, the story is interestingly different. In this case, we have that, at the metallicities high enough to form giant planets, the percentage of systems with a single giant planet with

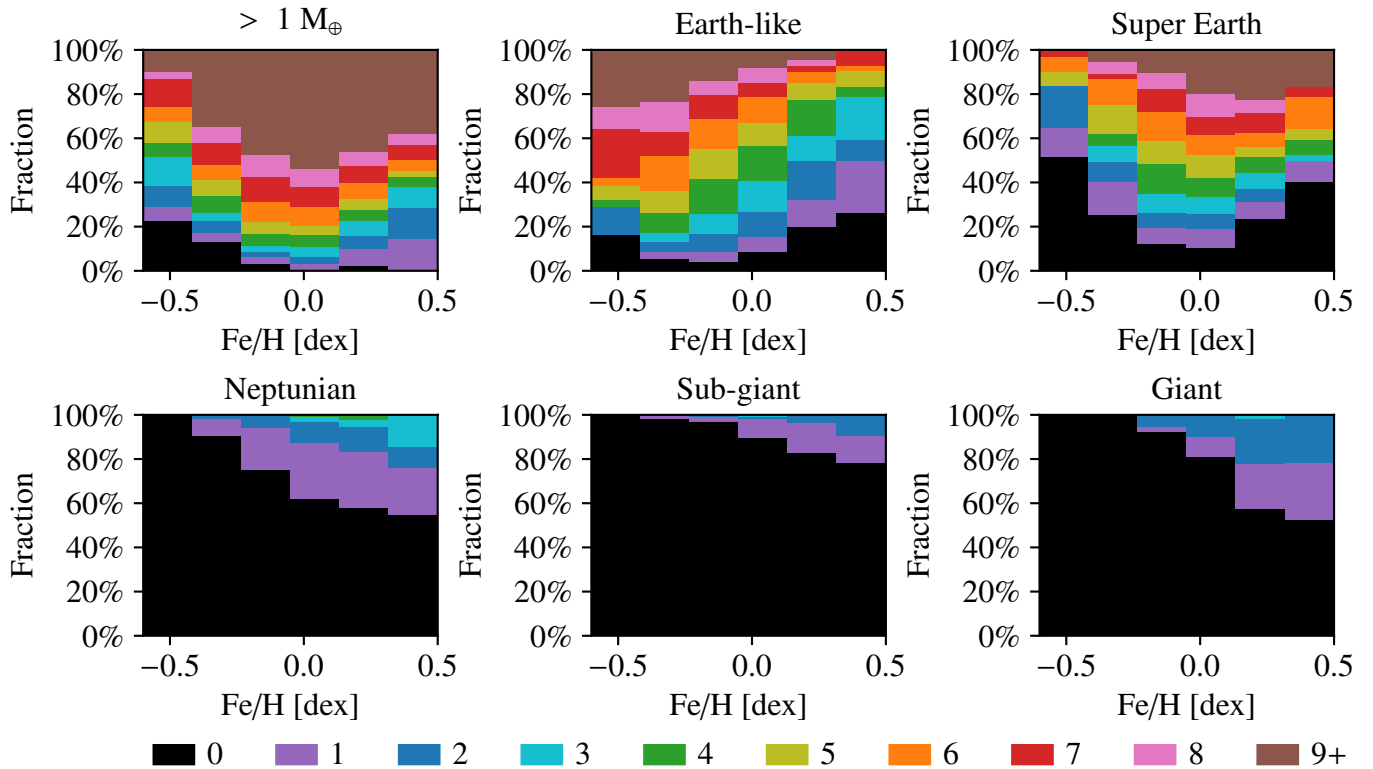
respect to the population of systems with any number of giant planets increases. This comes to say that the mean multiplicity is anticorrelated to the metallicity. This is visible by the fact that systems with two giant planets are less frequent in the highest metallicity bin than in the one below. Similarly, the five systems with three giants are not in the highest metallicity bin.

Giant planet formation is then a self-limiting process. The more giant planets are formed, the more likely is that these systems will be unstable. When an instability occurs, it will lead to the loss of planets, by collisions between planets, ejections, or, in small fraction of the cases, accretion by the central star.

In Paper XI, dedicated to giant planets, we will quantify the number of giants lost in collisions with other planets and the star, and by the ejection out of the system where they become rogue planets.

A effect is happening for the systems with the highest metallicity: the number of inner planets decreases. All the low-metallicity systems have some inner planets, although they can be very low mass (as there quite less planets that are Earth-like or more). However, this does not mean that these systems do not form planets; it can be seen that all these systems have at least one Earth-mass planet at least (top left panel of Fig. 26). What happens in these systems form several massive planets; due their number, the systems become dynamically unstable and the inner planets are lost. Much of these planets collide or are ejected, some fall in the star. In all but one of the resulting systems, a gi-





**Fig. 26.** Histogram of the multiplicities of different planet categories versus the stellar metallicity for the 100-embryos population. All these categories do not have constraints on the locations of the planets. The top-left panel shows the histogram of all planets larger than  $1 M_{\oplus}$  while the others are the categories discussed previously. They were defined in Sect. 7.1.

ant planet remains beyond 1 au. In the last case, a smaller planet remains, but its low mass is due to envelope ejection.

## 8. Summary and conclusions

In this work, we use the new Generation III Bern model of planetary formation and evolution presented in Paper I to compute synthetic planetary populations of solar-like stars. The model assumes that planets form according to the core accretion paradigm. During the formation stage (0 to 20 Myr), the model self-consistently evolves a 1D radial constant- $\alpha$  gas disc with internal and external photoevaporation, and the dynamical state of planetesimals under viscous stirring and damping by gas drag. Accretion of solids by the protoplanets includes both planetesimals and giant impacts, while gas accretion is obtained by solving the 1D spherically-symmetric internal structure equations. The model also includes gas-driven planetary migration and gravitational interactions between the protoplanets by means of the mercury  $N$ -body integrator. During the evolutionary phase (20 Myr to 10 Gyr) we follow the thermodynamic evolution (cooling and contraction) of the individual planets including the effects of atmospheric escape, bloating, and stellar tides.

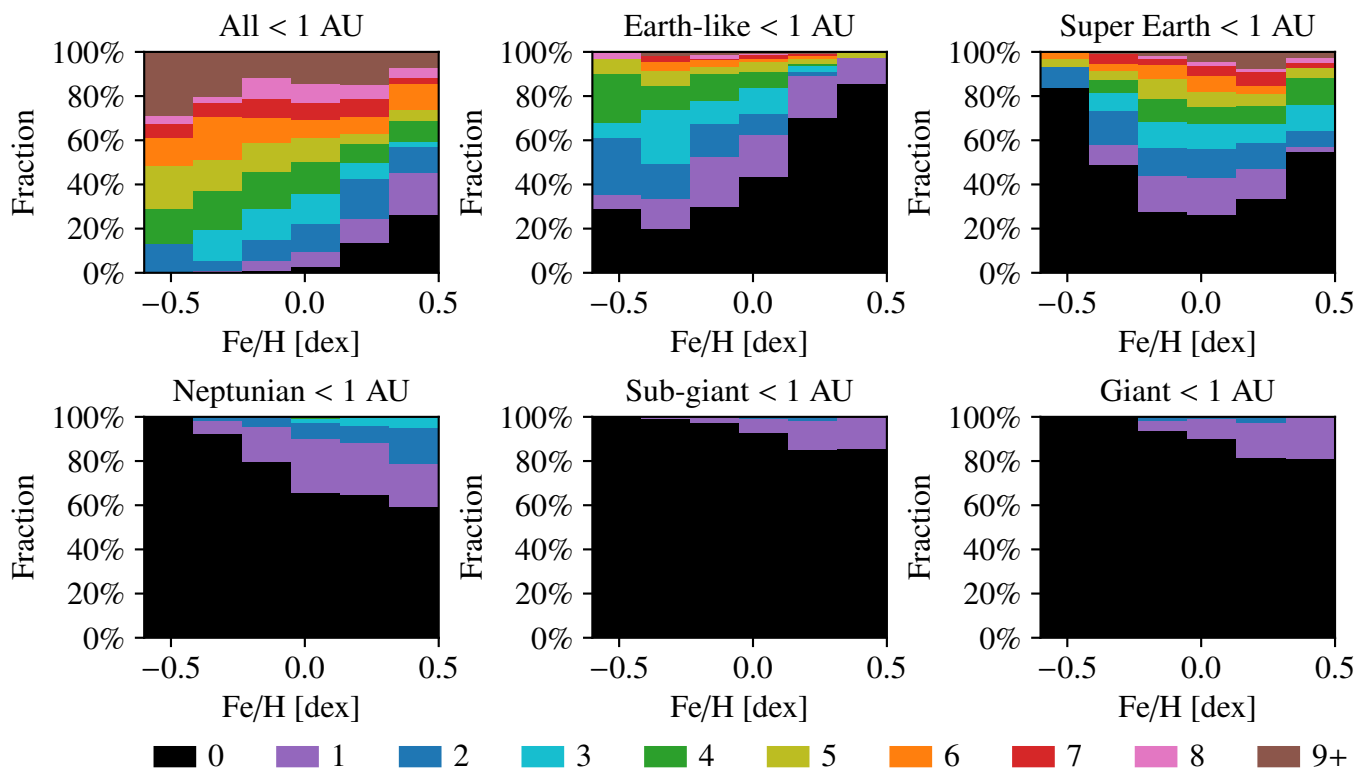
To synthesise populations, we vary four disc initial conditions of the the model according to observed (or observationally motivated) distributions. These Monte Carlo variables are: the initial mass of the gas disc (Tychoniec et al. 2018), the dust-to-gas ratio which is tied to the stellar [Fe/H] (Santos et al. 2005), the external photoevaporation rate which is distributed such that the synthetic discs have a lifetime distribution compatible with the observed one (see Sect. 3.3), and the inner edge of the protoplanetary disc (Venuti et al. 2017). Lunar-mass ( $10^{-2} M_{\oplus}$ ) plan-

etary seeds are put with a uniform probability in the logarithm of the distance into the disc. We compute five populations, each with a different the initial number of seeds per system (or disc).

One aim of this study is to determine the convergence of the model with respect to this free parameter. Our results for this part are:

- There is a strong difference between the single and multi-embryos populations. We find that migration in the single-embryo is more effective than in the multi-embryos population.
- The properties of the giant planets are only weakly affected by the number of embryos, as long as the latter is at least about 10, consistent with previous work (A13). For example, the fraction of stars with giant planets and their multiplicity is 19.8 % and 1.5 in the 10 embryo case, and 18.1 % and 1.6 in the 100 embryo case.
- For the lower-mass planets, a higher number of embryos is necessary. Only the 100-embryos population is able to track the formation of the lower-mass planets up to giant impacts stage (large embryo-embryo collisions).

There are two main reasons for these changes. The first is the dynamical interactions between the embryos, as we discussed in Paper I. A tighter spacing between the embryos increases their mutual gravitational interactions, which gives them access to more planetesimals to accrete. This helps small-mass systems to accrete a large percentage of the planetesimals at small separation during the time of our formation models (20 Myr). For the larger-mass planets however, the increased number of embryos results in more competition for solids. When the embryos grow to several Earth masses, they undergo gas-driven migration, which result in access to a larger mass reservoir. However, other embryos will have accreted planetesimals at different



**Fig. 27.** Histogram of the fraction of systems hosting planets of five categories versus the stellar metallicity for the populations presented in this study. The categories have the same boundaries as in Fig. 26 (with the exception of the top left panel that show planets of all masses), but they only account for the inner planets, i.e. inside 1 au.

places of the disc, resulting in migrating embryos experiencing a sudden drop in their growth rate. The more embryos there are, the less migration embryos must have performed before experiencing this effect. This in turn can trigger runaway gas accretion (see discussion in Sect. 4.1). The last effect is presence of multiple large embryos. With many embryos, it is more likely to form multiple giant planets. This means that the protoplanets can experience giant impacts. They can lead to envelope stripping of some giant planets. Thus, we find a small proportion of massive cores with a tiny envelope compared to the usual scenario provided by the core accretion paradigm. Systems with many embryos offer a greater diversity of envelopes mass fractions. The increase of dynamical interactions with the number of embryos has repercussions on the formation tracks, with planets being scattered to wide and eccentric orbits.

One of the reason for this study is to determine if results of the population with many embryos per systems can be recovered by populations with a lower initial embryo count. The more embryos are put in each systems, the larger the computational requirements are (mainly due to the  $N$ -body). For future work where we want to study the effects of model parameters, it is then more efficient to run the simulations with a lower number of embryos. From this study, we find that planets whose masses are roughly  $10 M_{\oplus}$  or more are insensitive to this parameter provided there are at least 10 embryos per system. There some effects of including more embryos, such as an overall increased distance for the giant planets (see Sect. 6.2 and Fig. 19), but this is small compared to other changes in the outcomes of the model (for instance, between the single and multi-embryos populations), so it should not constitute a major problem. The single-embryo population is different from the others and most of its properties are not recovered in multi-embryos populations. Nevertheless, some

outcomes, such as the mass function for planets above roughly  $10 M_{\oplus}$  can be retrieved. This means that the study of gas accretion in the detached phase or the overall fraction of giant planets (provided a correction factor is taken into account) can be done with these simple populations that require very limited computational resources.

Based on our population with the highest number of embryos per system (100), we computed properties of different planet kinds that are provided in Table 7 and graphically in Fig. 24. These values represent the key demographic predictions of our formation model. The main points are:

- Overall, planetary systems contain on average 8 planets larger than  $1 M_{\oplus}$ . The fraction of systems with giants planets at all orbital distances is 18 %, but only 1.6 % have one further than 10 au. System with giants contain on average 1.6 giants. This value is consistent with observations (Bryan et al. 2016).
- Inside of 1 au, the planet type with the highest occurrence rate and multiplicity are super Earth (2.4 and 3.7), followed by Earth-like planets (1.6 and 2.8). They are followed by Neptunian planets, but with an already clearly reduced occurrence rate and multiplicity (0.4 and 1.4).
- The planet mass function varies as  $M^{-2}$  between 5 and  $50 M_{\oplus}$ . Both at low and high masses, it follows approximately  $M^{-1}$ .
- The frequency of terrestrial and super Earth planets peaks at a stellar metallicity of -0.2 and 0.0 respectively. At lower metallicities, they are limited by a lack of building blocks and at higher metallicities by detrimental growth of more massive, potentially dynamically active planets, which results in accretion or ejection of terrestrial planets and super Earths. The frequency of more massive planet types (Neptunian, giants) increases in contrast monotonically with [Fe/H].

These results support observations about the metallicity effect for giant planets (see Figs. 26 and 27). It should be noted that the quantitative demographic results presented here like the (absolute) occurrence rates and multiplicities are functions of the chosen parameters and underlying model assumptions. We think that relative results and trends are more robust against parameter variations than such absolute results. To assess the impact of at least two model parameters and assumptions, we also studied two non-nominal populations starting with 100 embryos per disc (Appendix A): in-situ formation without gas-driven orbital migration, and a population with a steeper slope of the initial planetesimal surface density, inspired by recent planetesimal formation models (Voelkel et al. 2020).

In future work, we will compare these populations with observational data, in a similar fashion that was already done for radial-velocity surveys (Mordasini et al. 2009b) and transit (Mulders et al. 2019). This will determine how our populations statistically compare to the known exoplanet population. This should allow us to make steps towards the development of a standard model of planetary system formation and evolution. Observationally, the syntheses represent a large data set that can be searched for comparison synthetic planetary systems that show how observed systems may have come into existence. The systems, including their full formation and evolution tracks are available online. Knowing the underlying population will also help to understand the pathways certain categories of system follow to reach their final stage and the initial conditions they require. It would also permit to make predictions on the yet-unobserved regions of the parameter space, which is important for the development of future exoplanet discovery and characterisation missions.

*Acknowledgements.* The authors thank Ilaria Pascucci, Rachel B. Fernandes, and Barbara Ercolano for fruitful discussions. We also thank the anonymous reviewer, whose comments helped improve this manuscript. A.E. and E.A. acknowledge the support from The University of Arizona. A.E. and C.M. acknowledge the support from the Swiss National Science Foundation under grant BSSG10\_155816 “PlanetsInTime”. R.B. and Y.A. acknowledge the financial support from the SNSF under grant 200020\_172746. Parts of this work have been carried out within the frame of the National Center for Competence in Research PlanetS supported by the SNSF. Calculations were performed on the Horus cluster at the University of Bern. The plots shown in this work were generated using *matplotlib* (Hunter 2007).

## References

Adibekyan, V. 2019, *Geosciences*, 9, 105  
 Alcalá, J. M., Manara, C. F., Natta, A., et al. 2017, *A&A*, 600, A20  
 Alcalá, J. M., Natta, A., Manara, C. F., et al. 2014, *A&A*, 561, A2  
 Alibert, Y. 2014, *A&A*, 561, A41  
 Alibert, Y., Carron, F., Fortier, A., et al. 2013, *A&A*, 558, A109  
 Alibert, Y., Mordasini, C., & Benz, W. 2004, *A&A*, 417, L25  
 Alibert, Y., Mordasini, C., Benz, W., & Winisdoerffer, C. 2005, *A&A*, 434, 343  
 Allard, F., Homeier, D., & Freytag, B. 2012, *Philosophical Transactions of the Royal Society of London Series A*, 370, 2765  
 Andrews, S. M., Rosenfeld, K. A., Kraus, A. L., & Wilner, D. J. 2013, *ApJ*, 771, 129  
 Andrews, S. M., Terrell, M., Tripathi, A., et al. 2018, *ApJ*, 865, 157  
 Andrews, S. M. & Williams, J. P. 2005, *ApJ*, 631, 1134  
 Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., & Dullemond, C. P. 2009, *ApJ*, 700, 1502  
 Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., & Dullemond, C. P. 2010, *ApJ*, 723, 1241  
 Ansdell, M., Williams, J. P., Trapman, L., et al. 2018, *ApJ*, 859, 21  
 Ansdell, M., Williams, J. P., van der Marel, N., et al. 2016, *ApJ*, 828, 46  
 Ansdell, M. C. 2017, PhD thesis, University of Hawai’i at Manoa  
 Armstrong, D. J., Lopez, T. A., Adibekyan, V., et al. 2020, *Nature*, 583, 39  
 Asphaug, E. 2010, *Chemie der Erde / Geochemistry*, 70, 199  
 Asphaug, E. & Reufer, A. 2014, *NatGeo*, 7, 564  
 Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, *A&A*, 577, A42

Baruteau, C., Bai, X., Mordasini, C., & Mollière, P. 2016, *Space Sci. Rev.*, 205, 77  
 Baruteau, C., Crida, A., Paardekooper, S. J., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 667  
 Batalha, N. M., Borucki, W. J., Bryson, S. T., et al. 2011, *ApJ*, 729, 27  
 Beckwith, S. V. W. & Sargent, A. I. 1996, *Nature*, 383, 139  
 Benítez-Llambay, P., Masset, F., & Beaugé, C. 2011, *A&A*, 528, A2  
 Bennett, D. P., Ranc, C., & Fernandes, R. B. 2021, arXiv e-prints, arXiv:2104.05713  
 Benz, W., Anic, A., Horner, J., & Whitby, J. A. 2007, *Space Sci. Rev.*, 132, 189  
 Birnstiel, T., Klahr, H., & Ercolano, B. 2012, *A&A*, 539, A148  
 Bodenheimer, P., Hubickyj, O., & Lissauer, J. J. 2000, *Icarus*, 143, 2  
 Bodenheimer, P. & Pollack, J. B. 1986, *Icarus*, 67, 391  
 Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, 327, 977  
 Bourrier, V., Lecavelier des Etangs, A., Ehrenreich, D., et al. 2018, *A&A*, 620, A147  
 Bowler, B. P. 2016, *PASP*, 128, 102001  
 Brewer, J. M., Wang, S., Fischer, D. A., & Foreman-Mackey, D. 2018, *ApJ*, 867, L3  
 Broeg, C. H. & Benz, W. 2012, *A&A*, 538, A90  
 Bryan, M. L., Knutson, H. A., Howard, A. W., et al. 2016, *ApJ*, 821, 89  
 Bryan, M. L., Knutson, H. A., Lee, E. J., et al. 2019, *AJ*, 157, 52  
 Buchhave, L. A., Bitsch, B., Johansen, A., et al. 2018, *ApJ*, 856, 37  
 Buchhave, L. A., Bizzarro, M., Latham, D. W., et al. 2014, *Nature*, 509, 593  
 Buchhave, L. A., Latham, D. W., Johansen, A., et al. 2012, *Nature*, 486, 375  
 Burrows, A., Hubeny, I., Budaj, J., & Hubbard, W. B. 2007, *ApJ*, 661, 502  
 Carr, J. S. 2007, in *Star-Disk Interaction in Young Stars*, ed. J. Bouvier & I. Appenzeller, Vol. 243, 135–146  
 Chambers, J. 2006, *Icarus*, 180, 496  
 Chambers, J. E. 1999, *MNRAS*, 304, 793  
 Chambers, J. E. 2001, *Icarus*, 152, 205  
 Chambers, J. E. 2013, *Icarus*, 224, 43  
 Chau, A., Reinhardt, C., Helled, R., & Stadel, J. 2018, *ApJ*, 865, 35  
 Chauvin, G., Desidera, S., Lagrange, A. M., et al. 2017, *A&A*, 605, L9  
 Chiang, E. & Laughlin, G. 2013, *MNRAS*, 431, 3444  
 Coleman, G. A. L. 2021, *MNRAS*, 506, 3596  
 Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, *PASP*, 120, 531  
 Dawson, R. I. & Johnson, J. A. 2018, *ARA&A*, 56, 175  
 Demory, B.-O. & Seager, S. 2011, *ApJS*, 197, 12  
 Denman, T. R., Leinhardt, Z. M., Carter, P. J., & Mordasini, C. 2020, *MNRAS*, 496, 1166  
 Dittkrist, K. M., Mordasini, C., Klahr, H., Alibert, Y., & Henning, T. 2014, *A&A*, 567, A121  
 Dong, S. & Zhu, Z. 2013, *ApJ*, 778, 53  
 Eisner, J. A., Graham, J. R., Akeson, R. L., & Najita, J. 2009, *ApJ*, 692, 309  
 Eisner, J. A., Hillenbrand, L. A., White, R. J., Akeson, R. L., & Sargent, A. I. 2005, *ApJ*, 623, 952  
 Eisner, J. A., Monnier, J. D., Woillez, J., et al. 2010, *ApJ*, 718, 774  
 Emsenhuber, A., Mordasini, C., Burn, R., et al. in prep., *A&A*  
 Emsenhuber, A., Mordasini, C., Burn, R., et al. in press, *A&A*[arXiv:2007.05561]  
 Fedele, D., van den Ancker, M. E., Henning, T., Jayawardhana, R., & Oliveira, J. M. 2010, *A&A*, 510, A72  
 Fernandes, R. B., Mulders, G. D., Pascucci, I., Mordasini, C., & Emsenhuber, A. 2019, *ApJ*, 874, 81  
 Fernandez, J. A. & Ip, W. H. 1984, *Icarus*, 58, 109  
 Fischer, D. A. & Valenti, J. 2005, *ApJ*, 622, 1102  
 Flock, M., Fromang, S., Turner, N. J., & Benisty, M. 2017, *ApJ*, 835, 230  
 Fortier, A., Alibert, Y., Carron, F., Benz, W., & Dittkrist, K. M. 2013, *A&A*, 549, A44  
 Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., et al. 2014, *ApJS*, 214, 25  
 Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, *AJ*, 154, 109  
 Fulton, B. J., Rosenthal, L. J., Hirsch, L. A., et al. 2021, *ApJS*, 255, 14  
 Galicher, R., Marois, C., Macintosh, B., et al. 2016, *A&A*, 594, A63  
 Gáspár, A., Rieke, G. H., & Ballering, N. 2016, *ApJ*, 826, 171  
 Gonzalez, G. 1997, *MNRAS*, 285, 403  
 Hahn, J. M. & Malhotra, R. 1999, *AJ*, 117, 3041  
 Haisch, Karl E., J., Lada, E. A., & Lada, C. J. 2001, *ApJ*, 553, L153  
 Hansen, B. M. S. & Murray, N. 2012, *ApJ*, 751, 158  
 Hansen, B. M. S. & Murray, N. 2013, *ApJ*, 775, 53  
 Hayashi, C. 1981, *Progress of Theoretical Physics Supplement*, 70, 35  
 Heller, R. 2019, *A&A*, 628, A42  
 Henderson, C. B. & Stassun, K. G. 2012, *ApJ*, 747, 51  
 Hendler, N., Pascucci, I., Pinilla, P., et al. 2020, *ApJ*, 895, 126  
 Hills, J. G. & Goda, M. P. 1993, *AJ*, 105, 1114  
 Howard, A. W. 2013, *Science*, 340, 572  
 Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, *ApJS*, 201, 15  
 Howard, A. W., Marcy, G. W., Johnson, J. A., et al. 2010, *Science*, 330, 653  
 Hughes, A. M., Wilner, D. J., Qi, C., & Hogerheijde, M. R. 2008, *ApJ*, 678, 1119  
 Hunter, J. D. 2007, *Computing in Science and Engineering*, 9, 90

- Ida, S. & Lin, D. N. C. 2004a, *ApJ*, 604, 388
- Ida, S. & Lin, D. N. C. 2004b, *ApJ*, 616, 567
- Ida, S. & Makino, J. 1993, *Icarus*, 106, 210
- Ida, S., Tanaka, H., Johansen, A., Kanagawa, K. D., & Tanigawa, T. 2018, *ApJ*, 864, 77
- Inaba, S. & Ikoma, M. 2003, *A&A*, 410, 711
- Irwin, J., Hodgkin, S., Aigrain, S., et al. 2008, *MNRAS*, 384, 675
- Isella, A., Tatulli, E., Natta, A., & Testi, L. 2008, *A&A*, 483, L13
- Izidoro, A., Bitsch, B., Raymond, S. N., et al. 2021, *A&A*, 650, A152
- Jin, S. & Mordasini, C. 2018, *ApJ*, 853, 163
- Jin, S., Mordasini, C., Parmentier, V., et al. 2014, *ApJ*, 795, 65
- Johnson, J. A., Aller, K. M., Howard, A. W., & Crepp, J. R. 2010, *PASP*, 122, 905
- Kasting, J. F., Kopparapu, R., Ramirez, R. M., & Harman, C. E. 2014, *Proceedings of the National Academy of Science*, 111, 12641
- Kepler, M., Benisty, M., Müller, A., et al. 2018, *A&A*, 617, A44
- Kley, W. & Nelson, R. P. 2012, *ARA&A*, 50, 211
- Knutson, H. A., Fulton, B. J., Montet, B. T., et al. 2014, *ApJ*, 785, 126
- Kobayashi, H., Tanaka, H., Krivov, A. V., & Inaba, S. 2010, *Icarus*, 209, 836
- Kokubo, E. & Ida, S. 1998, *Icarus*, 131, 171
- Kokubo, E. & Ida, S. 2000, *Icarus*, 143, 15
- Kokubo, E. & Ida, S. 2002, *ApJ*, 581, 666
- Kretke, K. A. & Lin, D. N. C. 2012, *ApJ*, 755, 74
- Kuchner, M. J. & Lecar, M. 2002, *ApJ*, 574, L87
- Kurokawa, H. & Nakamoto, T. 2014, *ApJ*, 783, 54
- Lammer, H., Odert, P., Leitzinger, M., et al. 2009, *A&A*, 506, 399
- Latham, D. W., Rowe, J. F., Quinn, S. N., et al. 2011, *ApJ*, 732, L24
- Lecavelier Des Etangs, A. 2007, *A&A*, 461, 1185
- Lee, E. J. & Chiang, E. 2015, *ApJ*, 811, 41
- Lee, E. J. & Chiang, E. 2017, *ApJ*, 842, 40
- Lee, M. H. & Peale, S. J. 2002, *ApJ*, 567, 596
- Léger, A., Rouan, D., Schneider, J., et al. 2009, *A&A*, 506, 287
- Lenz, C. T., Klahr, H., & Birnstiel, T. 2019, *ApJ*, 874, 36
- Limbach, M. A. & Turner, E. L. 2015, *Proceedings of the National Academy of Science*, 112, 20
- Lissauer, J. J. 1987, *Icarus*, 69, 249
- Lodders, K. 2003, *ApJ*, 591, 1220
- Lopez, E. D. & Fortney, J. J. 2013, *ApJ*, 776, 2
- Lopez, E. D., Fortney, J. J., & Miller, N. 2012, *ApJ*, 761, 59
- Lundkvist, M. S., Kjeldsen, H., Albrecht, S., et al. 2016, *Nature Communications*, 7, 11201
- Lüst, R. 1952, *Zeitschrift Naturforschung Teil A*, 7, 87
- Lynden-Bell, D. & Pringle, J. E. 1974, *MNRAS*, 168, 603
- Lyra, W., Paardekooper, S.-J., & Mac Low, M.-M. 2010, *ApJ*, 715, L68
- Malhotra, R. 1993, *Nature*, 365, 819
- Mamajek, E. E. 2009, in *American Institute of Physics Conference Series*, Vol. 1158, American Institute of Physics Conference Series, ed. T. Usuda, M. Tamura, & M. Ishii, 3–10
- Manara, C. F., Fedele, D., Herczeg, G. J., & Teixeira, P. S. 2016a, *A&A*, 585, A136
- Manara, C. F., Mordasini, C., Testi, L., et al. 2019, *A&A*, 631, L2
- Manara, C. F., Rosotti, G., Testi, L., et al. 2016b, *A&A*, 591, L3
- Manara, C. F., Testi, L., Herczeg, G. J., et al. 2017, *A&A*, 604, A127
- Marleau, G.-D., Coleman, G. A. L., Leleu, A., & Mordasini, C. 2019, *A&A*, 624, A20
- Masset, F. & Snellgrove, M. 2001, *MNRAS*, 320, L55
- Mayor, M., Marmier, M., Lovis, C., et al. 2011, *arXiv e-prints*, arXiv:1109.2497
- Mazeh, T., Holczer, T., & Faigler, S. 2016, *A&A*, 589, A75
- McDonald, G. D., Kreidberg, L., & Lopez, E. 2019, *ApJ*, 876, 22
- Militzer, B., Wahl, S., & Hubbard, W. B. 2019, *ApJ*, 879, 78
- Millholland, S., Wang, S., & Laughlin, G. 2017, *ApJ*, 849, L33
- Mizuno, H. 1980, *Progress of Theoretical Physics*, 64, 544
- Moe, M. & Kratter, K. M. 2019, *arXiv e-prints*, arXiv:1912.01699
- Mollière, P. & Mordasini, C. 2012, *A&A*, 547, A105
- Morales, J. C., Mustill, A. J., Ribas, I., et al. 2019, *Science*, 365, 1441
- Morbidelli, A., Bottke, W. F., Nesvorný, D., & Levison, H. F. 2009, *Icarus*, 204, 558
- Mordasini, C. 2020, *A&A*, 638, A52
- Mordasini, C., Alibert, Y., & Benz, W. 2009a, *A&A*, 501, 1139
- Mordasini, C., Alibert, Y., Benz, W., Klahr, H., & Henning, T. 2012a, *A&A*, 541, A97
- Mordasini, C., Alibert, Y., Benz, W., & Naef, D. 2009b, *A&A*, 501, 1161
- Mordasini, C., Alibert, Y., Georgy, C., et al. 2012b, *A&A*, 547, A112
- Mordasini, C., Alibert, Y., Klahr, H., & Henning, T. 2012c, *A&A*, 547, A111
- Mordasini, C., Klahr, H., Alibert, Y., Miller, N., & Henning, T. 2014, *A&A*, 566, A141
- Mordasini, C., Marleau, G. D., & Mollière, P. 2017, *A&A*, 608, A72
- Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, *ApJ*, 832, 41
- Moutou, C., Deleuil, M., Guillot, T., et al. 2013, *Icarus*, 226, 1625
- Movshovitz, N., Bodenheimer, P., Podolak, M., & Lissauer, J. J. 2010, *Icarus*, 209, 616
- Mulders, G. D., Mordasini, C., Pascucci, I., et al. 2019, *ApJ*, 887, 157
- Mulders, G. D., Pascucci, I., Apai, D., & Ciesla, F. J. 2018, *AJ*, 156, 24
- Mulders, G. D., Pascucci, I., Apai, D., Frasca, A., & Molenda-Zakowicz, J. 2016, *AJ*, 152, 187
- Mulders, G. D., Pascucci, I., Manara, C. F., et al. 2017, *ApJ*, 847, 31
- Müller, A., Keppler, M., Henning, T., et al. 2018, *A&A*, 617, L2
- Murray, N., Chaboyer, B., Arras, P., Hansen, B., & Noyes, R. W. 2001, *ApJ*, 555, 801
- Nielsen, E. L., De Rosa, R. J., Macintosh, B., et al. 2019, *AJ*, 158, 13
- Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, *ApJ*, 743, L16
- O'Brien, D. P., Morbidelli, A., & Levison, H. F. 2006, *Icarus*, 184, 39
- Ohtsuki, K., Stewart, G. R., & Ida, S. 2002, *Icarus*, 155, 436
- Ormel, C. W., Dullemond, C. P., & Spaans, M. 2010, *ApJ*, 714, L103
- Ormel, C. W. & Klahr, H. H. 2010, *A&A*, 520, A43
- Owen, J. E. & Jackson, A. P. 2012, *MNRAS*, 425, 2931
- Owen, J. E. & Wu, Y. 2013, *ApJ*, 775, 105
- Pascucci, I., Testi, L., Herczeg, G. J., et al. 2016, *ApJ*, 831, 125
- Perri, F. & Cameron, A. G. W. 1974, *Icarus*, 22, 416
- Petigura, E. A., Howard, A. W., Marcy, G. W., et al. 2017, *AJ*, 154, 107
- Petigura, E. A., Marcy, G. W., Winn, J. N., et al. 2018, *AJ*, 155, 89
- Podolak, M., Podolak, J. I., & Marley, M. S. 2000, *Planet. Space Sci.*, 48, 143
- Podolak, M., Pollack, J. B., & Reynolds, R. T. 1988, *Icarus*, 73, 163
- Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, *Icarus*, 124, 62
- Quintana, E. V., Barclay, T., Borucki, W. J., Rowe, J. F., & Chambers, J. E. 2016, *ApJ*, 821, 126
- Rafikov, R. R. 2004, *AJ*, 128, 1348
- Raymond, S. N., O'Brien, D. P., Morbidelli, A., & Kaib, N. A. 2009, *Icarus*, 203, 644
- Richert, A. J. W., Getman, K. V., Feigelson, E. D., et al. 2018, *MNRAS*, 477, 5191
- Romanova, M. M. & Lovelace, R. V. E. 2006, *ApJ*, 645, L73
- Rosotti, G. P., Tazzari, M., Booth, R. A., et al. 2019, *MNRAS*, 486, 4829
- Rowan, D., Meschiari, S., Laughlin, G., et al. 2016, *ApJ*, 817, 104
- Santerne, A., Moutou, C., Tsantaki, M., et al. 2016, *A&A*, 587, A64
- Santos, N. C., Israelian, G., & Mayor, M. 2004, *A&A*, 415, 1153
- Santos, N. C., Israelian, G., Mayor, M., et al. 2005, *A&A*, 437, 1127
- Santos, N. C., Israelian, G., Mayor, M., Rebolo, R., & Udry, S. 2003, *A&A*, 398, 363
- Schlecker, M., Mordasini, C., Emsenhuber, A., et al. in pressa, *A&A*[arXiv:2007.05563]
- Schlecker, M., Pham, D., Burn, R., et al. in pressb, *A&A*[arXiv:2104.11750]
- Schlichting, H. E. & Mukhopadhyay, S. 2018, *Space Sci. Rev.*, 214, 34
- Sestovic, M. & Demory, B.-O. 2020, *A&A*, 641, A170
- Shakura, N. I. & Sunyaev, R. A. 1973, *A&A*, 500, 33
- Sousa, S. G., Santos, N. C., Israelian, G., Mayor, M., & Udry, S. 2011, *A&A*, 533, A141
- Sousa, S. G., Santos, N. C., Mayor, M., et al. 2008, *A&A*, 487, 373
- Stassun, K. G., Mathieu, R. D., Mazeh, T., & Vrba, F. J. 1999, *AJ*, 117, 2941
- Steffen, J. H., Ragozzine, D., Fabrycky, D. C., et al. 2012, *PNAS*, 109, 7982
- Stewart, S. T. & Leinhardt, Z. M. 2012, *ApJ*, 751, 32
- Suzuki, D., Bennett, D. P., Sumi, T., et al. 2016, *ApJ*, 833, 145
- Tanaka, H. & Ida, S. 1999, *Icarus*, 139, 350
- Tanaka, H., Takeuchi, T., & Ward, W. R. 2002, *ApJ*, 565, 1257
- Thiabaud, A., Marboeuf, U., Alibert, Y., et al. 2014, *A&A*, 562, A27
- Thiabaud, A., Marboeuf, U., Alibert, Y., Leya, I., & Mezger, K. 2015, *A&A*, 574, A138
- Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, *ApJS*, 235, 38
- Thorngrén, D. P. & Fortney, J. J. 2018, *AJ*, 155, 214
- Tobin, J. J., Sheehan, P. D., Megeath, S. T., et al. 2020, *ApJ*, 890, 130
- Tripathi, A., Andrews, S. M., Birnstiel, T., & Wilner, D. J. 2017, *ApJ*, 845, 44
- Turner, N. J., Fromang, S., Gammie, C., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 411
- Tychoniec, Ł., Manara, C. F., Rosotti, G. P., et al. 2020, *A&A*, 640, A19
- Tychoniec, Ł., Tobin, J. J., Karska, A., et al. 2018, *ApJS*, 238, 19
- Venuti, L., Bouvier, J., Cody, A. M., et al. 2017, *A&A*, 599, A23
- Vigan, A., Bonavita, M., Biller, B., et al. 2017, *A&A*, 603, A3
- Vigan, A., Fontanive, C., Meyer, M., et al. 2021, *A&A*, 651, A72
- Voelkel, O., Klahr, H., Mordasini, C., Emsenhuber, A., & Lenz, C. 2020, *A&A*, 642, A75
- Wagner, K., Apai, D., & Kratter, K. M. 2019, *ApJ*, 877, 46
- Wahl, S. M., Hubbard, W. B., Militzer, B., et al. 2017, *Geophys. Res. Lett.*, 44, 4649
- Walsh, K. J. & Levison, H. F. 2019, *Icarus*, 329, 88
- Wang, J. & Fischer, D. A. 2015, *AJ*, 149, 14
- Wang, W., Wang, L., Li, X., Chen, Y., & Zhao, G. 2018, *ApJ*, 860, 136
- Weidenschilling, S. J. 1977, *Ap&SS*, 51, 153
- Weiss, L. M., Marcy, G. W., Petigura, E. A., et al. 2018, *AJ*, 155, 48
- Wetherill, G. W. 1985, *Sci.*, 228, 877
- Williams, J. P., Cieza, L., Hales, A., et al. 2019, *ApJ*, 875, L9
- Winn, J. N. & Fabrycky, D. C. 2015, *ARA&A*, 53, 409
- Wittenmyer, R. A., Tinney, C. G., O'Toole, S. J., et al. 2011, *ApJ*, 727, 102
- Wright, J. T., Marcy, G. W., Howard, A. W., et al. 2012, *ApJ*, 753, 160
- Zhu, W. & Dong, S. 2021, *arXiv e-prints*, arXiv:2103.02127
- Zhu, W., Petrovich, C., Wu, Y., Dong, S., & Xie, J. 2018, *ApJ*, 860, 101
- Zhu, W. & Wu, Y. 2018, *AJ*, 156, 92



**Table A.1.** Percentage of systems/stars with specific planetary types ( $f_s$ ) and their mean multiplicity ( $\mu_p$ ) for the different populations. Similar to Table 6, but comparing the non-nominal populations.

Type	Nominal		slope -2		In-situ	
	$f_s$ [%]	$\mu_p$	$f_s$ [%]	$\mu_p$	$f_s$ [%]	$\mu_p$
Mass >1 $M_\oplus$	96.0	8.4	96.2	8.7	96.2	4.7
Earth-like	90.1	5.2	89.9	5.3	54.9	3.8
Super Earth	82.1	5.6	77.6	5.8	78.9	2.7
Neptunian	30.3	1.4	32.3	1.7	33.2	1.2
Sub-giant	8.5	1.2	12.8	1.2	28.8	1.3
Giant	18.1	1.6	21.3	1.5	45.4	1.4
D-burning	4.5	1.0	3.8	1.0	9.3	1.0
Earth-like <1 au	57.2	2.8	49.6	3.8	37.4	1.8
Super Earth <1 au	66.1	3.7	69.5	4.3	63.3	1.4
Neptunian <1 au	26.2	1.4	30.3	1.6	8.0	1.1
Sub-giant <1 au	6.5	1.1	10.7	1.2	0.6	1.0
Giant <1 au	9.2	1.1	14.0	1.1	0.0	...
Habitable zone	43.7	1.3	49.4	1.3	26.5	1.1
Kepler P18	76.7	4.4	90.0	4.7	61.9	1.4
Kepler Z18	65.7	4.5	83.2	4.4	37.8	1.2
Hot Jupiter	0.3	1.0	1.4	1.0	0.0	...
Jupiter analogue	0.8	1.0	0.3	1.0	22.0	1.0
Giant >5 au	3.5	1.0	2.2	1.0	30.0	1.0
Giant >10 au	1.6	1.0	1.3	1.0	6.9	1.0
Giant >20 au	0.8	1.0	0.7	1.0	2.1	1.0

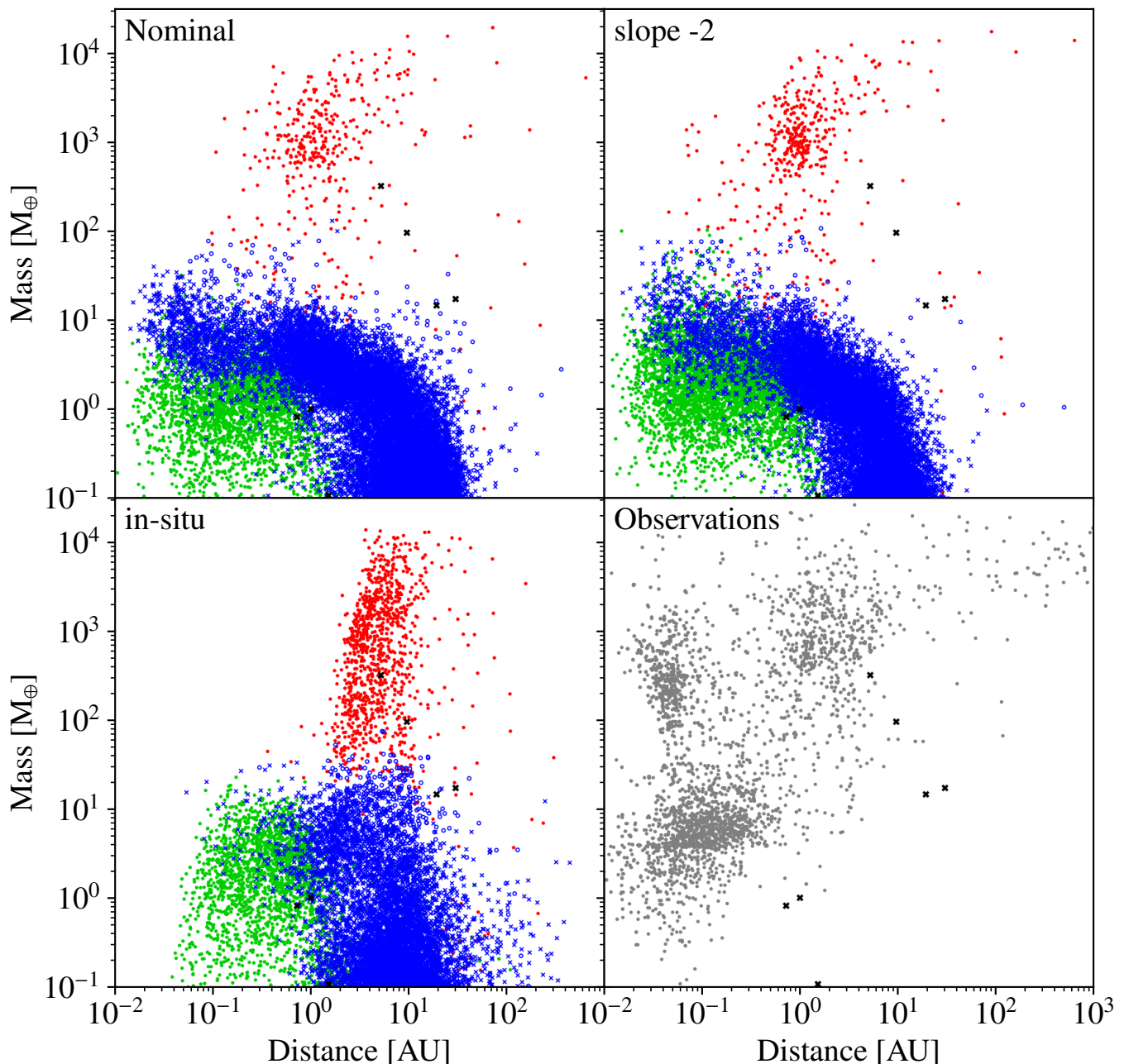
## Appendix A: Influence of model parameters

To study the effect of some model parameters, we performed additional populations with initially 100 embryos per system. Two additional populations were generated with 1000 stars/systems each; the first one has the power-law index of the radial slope of initial planetesimals distribution was set to  $\beta_s = 2$  (instead of the nominal 1.5) so that the planetesimal isolation mass remains constant with distance (Lissauer 1987) except for the jumps at the ice lines, and in line with the results of Lenz et al. (2019) regarding the slope of the planetesimal disc as predicted by their planetesimal formation model. In the second population, gas-driven migration is not included, though  $N$ -body interactions remain; we refer to this population as “in-situ”. These two populations, along with the nominal one are shown in Fig. A.1. For comparison, we also plot the confirmed exoplanets as of 18 June 2021, without accounting for observational biases. Table A.1 lists the fundamental demographic results.

Comparing first the the nominal population and that with the modified power-law index, we note little differences for the giant planets. The number of such planets is 315 in the population with the modified slope compared to 284 in the nominal one, or a 11 % increase. These planets are still piled up around 1 au in the synthetic populations, or slightly closer-in than the observed population. Some differences remain, like a larger number of hot Jupiters, and of giant planets inside of 1 au in the populations with the modified slope. The number of distant giants is in contrast reduced, meaning that the more centrally condensed distribution of solids also leads to more compact planetary systems, as one might naively expects. In terms of the  $f_s$  and  $\mu_p$  listed in Table A.1, these differences are certainly visible, but still do not correspond to a really fundamental change of the demographic predictions. Thus, we conclude the power-law index has a rather limited effect on the giant planets, affecting mostly

the occurrence of such planets as a function of orbital distance in the form of an inward shift. Similar, rather limited changes are seen in the  $f_s$  and  $\mu_p$  of several other planet types. An interesting difference between the two populations is seen in the hot and warm-Neptune-mass planets (about 10  $M_\oplus$  and inside 0.5 au). In the nominal population, most of these planets come outside the ice line, as there is limited mass reservoir in the inner region of the disc. With a power-law index of  $\beta_s = 2$  however, even planets of several Earth masses are able to grow locally. Migration still brings planets from regions beyond the ice line, but it is no longer the only mechanism able to form such planets. This difference is visible in Figure Fig. A.1 by the high number of green points in the aforementioned part of the  $a-M$  diagram compared to the nominal case, where this part is predominantly populated with blue points. The percentage of stars with planets in the region probed by Kepler is increased in the slope -2 population, whereas their multiplicity is not significantly changed.

Conversely, the in-situ population shows large differences to the others in many aspects, which is visible both in the  $a-M$  diagram and the demographic data of Table A.1. For a start, the number of giant planets is increased by a factor of more than two to 621. The fraction of stars with giants is about 45%, compared to about 20% in the other populations. The increase of the number of giant planets is due to the large efficiency of the in-situ model at forming these planets; for reference only 8 giants were accreted by the star in the nominal population, while the in-situ population has 3 such occurrences. Secondly, the planetary desert between 30 and 300  $M_\oplus$  is no longer observed. There are still slightly less planets in this range than giants, but the difference is much smaller. The reason for these differences (number of planets and the depth of the desert) with the mechanism of interplay between accretion and migration discussed in Sect. 4.1. In the in-situ population, the mechanism is not effective at all, as there is no migration. This means that the formation pattern observed by Pollack et al. (1996), with core growth first followed by a long intermediate stage with little accretion before the final gas runaway, is possible in this case. In contrast, when migration is included the planets cannot remain at intermediate masses ( $\sim 10-50 M_\oplus$ ) for too long or they will end up taken to the inner edge of the gas disc where gas accretion is impossible. Thirdly, without migration, giant planets are found mostly beyond the ice line, and thus further away than the detected sub-population of warm giants. The in-situ population does not contain a single giant inside of 1 au, in clear contrast to observations. This is because only the high-eccentricity migration channel remains (Dawson & Johnson 2018) and the we do not account for it in our populations. Further, the in-situ population also fails to reproduce the warm Neptune and hot-Jupiter sub-populations. This is due to the low mass of solid building blocks available in the inner region of the disc. This means that gas-driven migration is necessary to reproduce the observed population (in particular the location of the giant planets, the presence of close-in massive planets, and the super Earth and sub-Neptunes inside of about 0.1 au). Our nominal prescription results in contrast in too large migration. A certain reduction of the gas-driven migration would lead to a better match with the observations. This could be explained by the following: 1) a weaker migration than what is usually assumed (Ida et al. 2018), 2) by lower viscosity, which would reduce migration speed, or 3) a combination of strong and weak migration. The last point stems out from the fact that the observed population shows characteristics that are reproduced partly in either cases. We do not account for the last possibility in our populations, as the viscosity parameter is taken to be constant across all discs. This effect could be reproduced however



**Fig. A.1.** Mass-distance diagram for the comparison of 100 embryos populations generated with different model parameters. The upper left panel shows the nominal population, the upper right panel shows one where the index of the power law for the initial distribution of solids was changed to  $\beta_s = 2$  (so that the isolation mass is constant with distance), while the lower left panel shows a population where no gas-driven migration is included. The bottom right panel shows the known exoplanets as of 18 June 2021. It should be noted that this does not account for detection biases, which favour the discovery of hot-Jupiters. This gives the incorrect impression that the model severely fails to reproduce those planets.

by the parameter being varied, such as if were treated as one of the Monte Carlo variables of our populations.

Finally, it is worth mentioning that the in-situ simulation populates the positions and masses of Jupiter and Saturn much more than the populations with orbital migration. The latter populations do contain systems with an architecture that is qualitatively the same as the Solar System with low-mass planets inside, followed by two giants, and then ice giant planets outside; however, in such synthetic systems the giants are shifted inward relative to the Solar System, with the inner giant planet residing at about 2-3 au and the outer at 5-8 au. The in-situ population provides a better match, but fails, as we have just seen, in repro-

ducing several fundamental constraints coming from the extra-solar planets. An obvious candidate for a mechanism that would help to conciliate Solar System and exoplanets is the Masset-Snellgrove mechanism (Masset & Snellgrove 2001) which leads to outward migration of two giant planets in a mean motion resonance. This is not possible in the Type II migration model we currently use. Including outward migration of two giants in the 1D disc/migration framework is thus another important line of future work.