



Chemical diversity of super-Earths as a consequence of formation

Jennifer Scora,¹ Diana Valencia,² Alessandro Morbidelli³ and Seth Jacobson⁴

¹Department of Astronomy and Astrophysics, University of Toronto, Toronto, ON, Canada

²Centre for Planetary Sciences, University of Toronto, 1265 Military Trail, Toronto, ON M1C 1A4, Canada

³Laboratoire Lagrange, Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Blvd de l'Observatoire, CS 34229, F-06304 Nice Cedex 4, France

⁴Department of Earth and Environmental Sciences, Michigan State University, East Lansing, MI 48824, USA

Accepted 2020 February 19. Received 2020 February 6; in original form 2019 October 18

ABSTRACT

Recent observations of rocky super-Earths have revealed an apparent wider distribution of Fe/Mg ratios, or core to mantle ratios, than the planets in our Solar system. This study aims to understand how much of the chemical diversity in the super-Earth population can arise from giant impacts during planetary formation. Planet formation simulations have only recently begun to treat collisions more realistically in an attempt to replicate the planets in our Solar system. We investigate planet formation more generally by simulating the formation of rocky super-Earths with varying initial conditions using a version of SYMBA, a gravitational N -body code, that incorporates realistic collisions. We track the maximum plausible change in composition after each impact. The final planets span a range of Fe/Mg ratios similar to the Solar system planets, but do not completely match the distribution in super-Earth data. We only form a few planets with minor iron-depletion, suggesting other mechanisms are at work. The most iron-rich planets have a lower Fe/Mg ratio than Mercury, and are less enriched than planets such as Kepler-100b. This indicates that further work on our understanding of planet formation and further improvement of precision of mass and radius measurements are required to explain planets at the extremes of this Fe/Mg distribution.

Key words: planets and satellites: composition – planets and satellites: formation.

1 INTRODUCTION

Before the discovery of exoplanets, planet formation theories were limited to explaining the Solar system and thus, were unintentionally biased. Now, with thousands of extrasolar planetary systems, there is a diverse set of data to test against formation theories. Models so far have been constructed to explain the masses and orbital architectures seen in the data (i.e. Hansen & Murray 2012; Ogihara, Morbidelli & Guillot 2015; Izidoro et al. 2017), where composition is just a by-product of the two. This is because composition is not a useful constraint for most classes of planets. For exo-jovians, the only possibility is a high H/He content, whereas for low-mass exoplanets the possibilities are unconstrained (Valencia, Sasselov & O'Connell 2007a; Rogers & Seager 2010; Vazan, Ormel & Dominik 2018). In fact, the degeneracy in the composition of mini-Neptunes arises from a trade-off between refractory material and water (Valencia et al. 2013). On the other hand, the composition of rocky super-Earths is a rich source of information with only minor degeneracies, and it has not yet been used to constrain formation scenarios. Therefore, with this study we propose to use composition of rocky super-Earths as another axes to constrain formation theories.

The reason for these fewer compositional degeneracies in rocky super-Earths is that the main determinant for the radius of the planet is the amount of iron (Valencia, Sasselov & O'Connell 2007b). In simple terms, rocky super-Earths are made of a silicate-oxide mantle overlaying an iron core, and therefore their relative masses can be inferred from measurements of the total mass and radius. Hence, we can estimate the core-mass fraction (CMF) of the planet, and from this we can infer the iron to silicon (Fe/Si) or the iron to magnesium (Fe/Mg) ratios, where silicon and magnesium are used as tracers of mantle material. In reality though, the above picture is more complex because: planets generally are not expected (1) to have pure iron cores – Si is a candidate among others for the Earth's core light alloy (Hirose, Labrosse & Hernlund 2013), or (2) to be completely differentiated – some iron may remain in the mantle (e.g. Earth's iron content in mantle minerals is about 10 per cent, Ringwood 1970). However, differences in total radius for a given mass due to these complexities are minor for rocky super-Earths. Plotnykov and Valencia (in preparation) calculate a difference in radius of 0.5 per cent and 0.8 per cent in the absence of a light alloy in the core, and 2.6 per cent and 2 per cent due to differentiation for Earth and a $5 M_E$ planet, respectively. Therefore, mass–radius data pairs for rocky super-Earths can provide valuable constraints on CMF.

Furthermore, observational efforts in the last decade have now yielded hundreds of rocky super-Earths with measured masses and radii, with the more precise data pairs amenable to compositional

* E-mail: jscora@sympatico.ca

inference. The first transiting super-Earths with measured masses were CoRoT-7b (Léger et al. 2009; Southworth 2011; Barros et al. 2014; Stassun, Collins & Gaudi 2017) and Kepler-10b (Batalha et al. 2011; Dumusque et al. 2014; Fogtmann-Schulz et al. 2014; Van Eylen et al. 2016), which early on demonstrated the difficulty of collecting and interpreting high-quality data. The stellar activity of CoRoT-7 made it difficult to infer the mass of the planet with high precision (Hatzes et al. 2011; Barros et al. 2014), while instead asteroseismology data on Kepler-10 allowed for an exquisite precision of $\pm 0.003 R_{\oplus}$ in planetary radius (Fogtmann-Schulz et al. 2014). Kepler yielded thousands of super-Earths' with measured radii, and hundreds of them have been followed up for mass estimates. Although, typical data are more precise in radius than it is in mass, we are finally at a stage where there is enough quality mass–radius data that we can infer the composition of rocky super-Earths as a population, and use it to constrain formation scenarios.

We propose to use this chemical data to inform planet formation theories. That is, any successful theory that forms super-Earths needs to explain their chemical diversity as seen from the Fe/Mg ratios, as well as the distribution of masses ($5\text{--}15 M_{\oplus}$) and semimajor axes (0.01–0.2 au) seen in rocky super-Earths from the NASA Exoplanet Archive as of 2019 July. Eccentricities and inclinations need to be accounted for as well, although in reality these quantities are highly unconstrained for most low-mass exoplanets.

When attempting to form super-Earth and mini-Neptune planets starting from the minimum-mass solar nebula (MMSN), followed by standard core accretion, fails to reproduce the exoplanet low-mass population (Raymond & Cossou 2014). Instead, efforts to explain the formation of these low-mass planets fall into two categories: either they form outside their current location and migrate inwards (Ogihara et al. 2018; Terquem & Papaloizou 2007; Ida & Lin 2010; Cossou et al. 2014; Izidoro et al. 2017, 2019) stopping at the current locations due to gas-dispersal or the disc's inner edge, or they form *in situ* from massive discs (Hansen & Murray 2012, 2013; Chatterjee & Tan 2014). These massive discs may have formed through inward drift of smaller material before later stages of planet formation (Hansen & Murray 2012; Chiang & Laughlin 2013; Chatterjee & Tan 2014). If planets migrate inwards, they can end up with substantial water in their envelopes or interiors. If they form *in situ*, they would have no water in either reservoir.

Much of the attention has been given to explaining why mini-Neptunes exist. That is, if planets could grow enough mass ($5\text{--}10 M_{\oplus}$ cores) in the presence of the gas to trigger runaway accretion, why did they not become gas giants? The answer usually invokes a gas-poor environment due to a timing issue: either because the gas disc lifetime is short (Alibert et al. 2006; Terquem & Papaloizou 2007) or the planets formed late when there was little gas around (Lee & Chiang 2016; Dawson, Chiang & Lee 2015). Alternatively, the planets may accrete gas more slowly than expected (Lambrechts & Lega 2017). Given that all these scenarios result in some gas envelope accretion, then how are bare rocky super-Earths formed? Two possibilities are (1) that rocky super-Earths formed so slowly that by the time planets grew massive enough to capture gas, there was no gas around (Dawson et al. 2015), or (2) that they did acquire an envelope but lost it, either due to atmospheric photo-evaporation (Owen & Wu 2017; Lopez & Fortney 2013) or giant impact collisions (Inamdar & Schlichting 2016). The first theory does seem to be challenged by the fast accretion time-scales that close-in super-Earths experience (Ogihara et al. 2015) (also see Section 4). No theory is without problems, especially once we consider multiplanet systems (e.g. Kepler-36, Carter et al. 2012) where both rocky super-Earths and mini-Neptunes exist with

similar masses, and, of course, the same disc lifetime and properties. However, any successful formation theory not only needs to explain the mass range and orbital periods of super-Earths as a population distinct from mini-Neptunes, but it also needs to account for their chemical diversity in terms of the Fe/Mg ratios seen in the data.

With this in mind, we focus on the formation of rocky super-Earths and the role of collisions in growing planets with the observed chemical make-up. Our interest in planetary collisions is highly motivated by the fact that we observe a spread in the Fe/Mg ratios for these planets which seems to be beyond the spread of stellar abundances (see Section 2), suggesting these planets do not have a primordial composition.

Collisions may be responsible for significantly altering the composition of growing planets (Asphaug, Agnor & Williams 2006), and can vary from super-catastrophic outcomes where the target is heavily disrupted to perfect mergers of the two involved bodies and anything in between. Typical formation models of growing planets by gravitational interactions involve N -body codes that only consider perfect mergers during encounters (for a summary see Raymond et al. 2014; Bond, O'Brien & Lauretta 2010). These models are severely limited for investigating the chemical diversity of planets. It is only recently that formation simulations have begun accounting for more realistic collisions (Kokubo & Genda ; Chambers 2013; Haghighipour & Maindl 2019). Obtaining the actual outcome of a collision requires high-resolution simulations and is typically done with smoothed-particle hydrodynamic (SPH) codes that are too computationally expensive to couple directly to N -body simulations. Instead, the above-quoted studies have used analytical collisional prescriptions (with varying complexity) that predict the outcomes of planetary encounters based on the impact parameters and velocities.

In this study, we use an N -body code to form planets from a compact disc and the analytic equations for collision outcomes by Stewart & Leinhardt (2012). Our assumptions consider the most favourable scenarios that produce the most chemical diversity as a limiting case. We find that impacts can explain iron enrichments similar to Mercury. However, planetary collisions among bodies that start with solar composition are insufficient to explain all the diversity seen in the super-Earth data.

Our manuscript layout is as follows: in Section 2, we explain the chemical information that can be inferred from the mass–radius data, interior structure models, and stellar abundances. In Section 3, we introduce our model for forming super-Earths and how we tracked composition during formation. In Sections 4–6, we present our results, discuss the implications, and outline our conclusions, respectively.

2 SUPER-EARTH DATA AND CHEMISTRY

From the 269 super-Earth planets with measured radius and measured masses below $20 M_{\oplus}$ on the NASA exoplanet archive as of 2019 July, we chose the ones that have low enough mass errors that a meaningful inference of composition is possible. Our cut-off was to consider planets with 50 per cent mass error estimates ($\Delta M/M < 50$ per cent).

Fig. 1 shows these planets colour-coded as a function of flux received by each planet. Highly irradiated planets are more susceptible to atmosphere evaporation, and thus are more likely to be rocky (Lopez & Fortney 2013; Owen & Wu 2017). Since many of this sample receive high fluxes, we can assume that most of these planets do not have significant atmospheres.

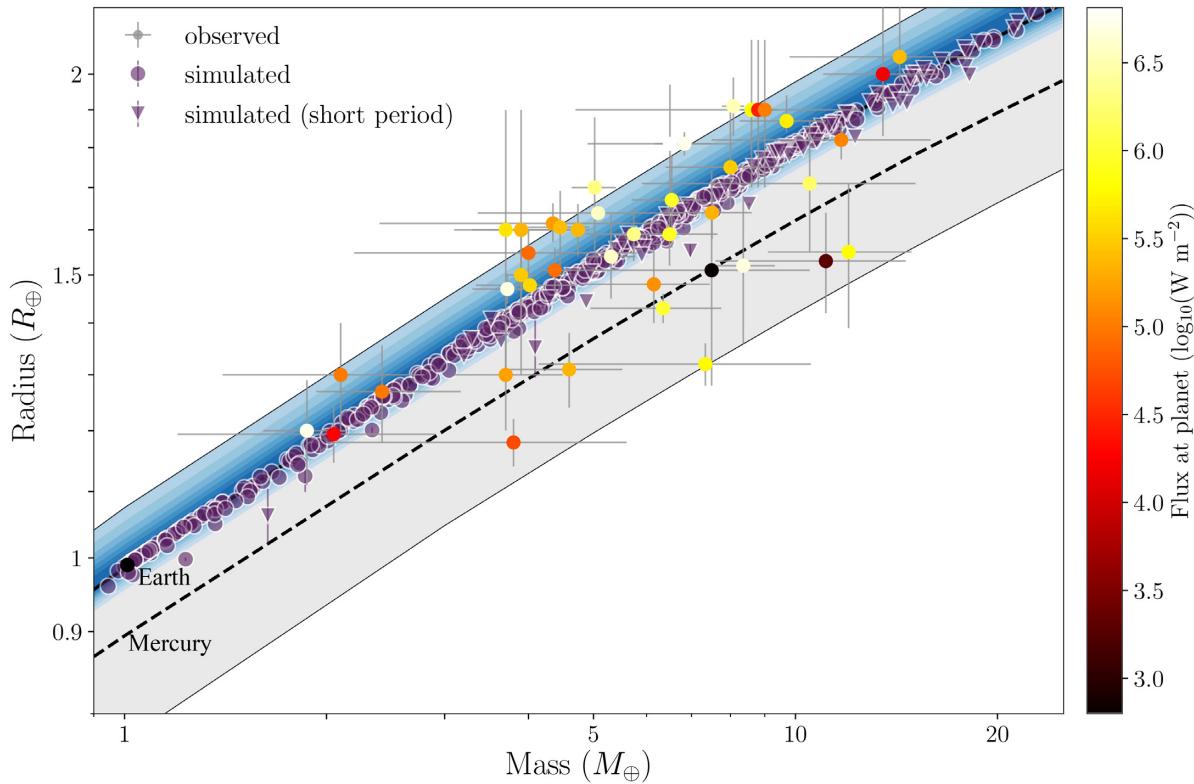


Figure 1. The masses and radii of the final planets from our simulations (in purple) compared to observed super-Earths with good measurements of mass and radius, which are colour-coded by the stellar flux received by the planet. Super-Earths from simulations that start with discs closer to the star (inner edge at 0.05 au) are plotted as triangles, and those from the main simulation suite are plotted as circles. The error bars on these points represent the maximum and minimum CMF values that are then averaged, as per the ‘merging disruption’ prescription described in Section 3.3. These are plotted over lines of constant composition [from top to bottom: 0 CMF, Earth (0.326), Mercury (0.69), 1 CMF]. The blue shaded region is the area that would be populated by planets if they had the Fe/Mg abundances of their host stars. The different shades of blue represent the probability distribution of stellar Fe/Mg abundances, as in Fig. 2, which shows the probability density of the Fe/Mg abundances of planet-hosting stars in the Hypatia Catalogue (Hinkel et al. 2014).

We calculated the mass–radius relationships for four different rocky compositions based on the model by Valencia, O’Connell & Sasselov (2006) and Valencia et al. (2007b).¹ These compositions are: (1) a planet devoid of iron, (2) an Earth like composition with $\text{Fe}/\text{Mg} = 2.0$ (CMF = 0.326, Stacey 2005), (3) an iron enriched composition (similar to Mercury) of $\text{Fe}/\text{Mg} \sim 8.45$ or CMF = 0.69 (Mercury’s is estimated at 0.6–0.75 by Hauck et al. 2013), and (4) a pure iron planet. The shaded region between the iron-free and pure iron planet are the boundaries within which rocky planets exist, with iron content increasing from top to bottom. Any planet above the iron-free line requires an envelope massive enough to modify the radius, thus, the planet is a mini-Neptune. We term this line the rocky-threshold radius (RTR). Any planet below the RTR may be rocky, but it is not definitive, because the degeneracy in composition means a planet can trade-off between iron and H/He or H_2O and still have the same mass and radius (Valencia et al. 2007a). Namely, a planet with the mass and size of the Earth could also have an H/He envelope if it had a larger iron core and less mantle. We do

not expect to see any planets below the pure iron line, as denser compounds are not abundant enough in the solar nebula to create whole planets.

We then investigated how planets with primordial compositions, or those that have the same Fe/Mg ratios as their host stars, compare to these mass–radius relationships and the super-Earth data. We use the Fe/Mg ratio instead of the Fe/Si ratio, as silicon is somewhat volatile (McDonough & Sun 1995), and thus may not be incorporated as readily into planets as magnesium. We consider the stellar absolute abundances by weight of stars with planets from the Hypatia Catalogue (Hinkel et al. 2014). Although there is a range in stellar Fe/Mg, (see Fig. 2), it translates to a very narrow spread in the relationship between mass and radius. In other words, planets that fall outside the blue shaded region in Fig. 1 do not have a primordial composition.

There are many planets that appear enriched and also depleted in iron with respect to the primordial composition distribution. The presence of significant amounts of water or an atmosphere could inflate planet radii, causing rocky planets to appear more iron-depleted than they are. However, some of the planets that appear iron-depleted are also highly irradiated by their stars, and thus should be stripped of any atmosphere or volatiles. In this study, we therefore assume that all the chosen super-Earths are completely rocky. Consequently, we consider the maximum range of refractory compositions that these planets could possibly have. We set out to investigate if the diversity in refractory composition seen in the data

¹We have updated the equations of state to reflect better data obtained by Stixrude & Lithgow-Bertelloni (2011) for the mantle and that of the core by Morrison et al. (2018). We chose the $\text{Fe}_{0.95}\text{Ni}_{0.05}$ equation of state for the core, but to calculate the chemical budget we include an unspecified light alloy in the core that makes up 10 per cent by mol and does not include magnesium, in line with the Earth (McDonough & Sun 1995), nor silicon, for simplification

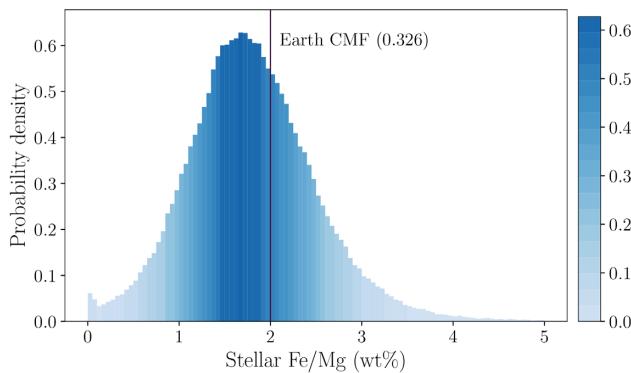


Figure 2. Distribution of Fe/Mg values for stars with planets, including their error bars. The values are taken from the Hypatia Catalogue (Hinkel et al. 2014) as of 2019 July. The colour bar corresponds to the probability density values for each stellar Fe/Mg value, and these colours are used in Fig. 1.

could be explained through chemical reprocessing of planets during the giant impact phase.

3 MODELLING FORMATION WITH AN N-BODY GRAVITATIONAL CODE

3.1 Initial conditions

We consider in this work the *in situ* formation scenario for super-Earths in the post-gas disc phase (Hansen & Murray 2012). We are aware that the consistency of this scenario is criticized (Raymond & Cossou 2014; Ogiwara et al. 2015). We adopt it as a best case scenario for high-velocity collisions that may remove mantle material and enhance the CMF of the final planets. If we do not find that this scenario produces planets that are sufficiently iron-enriched, it is unlikely that other scenarios, including planet collisional accretion at larger distances and smaller relative velocities or the *in situ* accretion of small drifting particles can explain the large Fe/Mg ratios inferred for some exoplanets.

Giant impacts occur after the oligarchic growth phase of terrestrial planet formation, once planetary embryos have formed. We use the basic initial conditions for *in situ* planet formation developed by Kokubo & Ida (2002). Half of the disc mass starts in embryos, or bodies that have grown large enough to dominate gravitational interactions, and the other half is in planetesimals, bodies roughly 1 km in diameter that interact mainly with the embryos. We distribute the embryos using the following equations for the isolation mass, or initial mass (M_{iso}), and mutual Hill radius (r_H) for planets around a $1 M_\odot$ star, developed by Kokubo & Ida (2002):

$$M_{\text{iso}} = 0.16 \left(\frac{b \Sigma_o}{10} \right)^\alpha \left(\frac{a}{1 \text{au}} \right)^{\alpha(2-\alpha)} M_\oplus \quad (1)$$

$$r_H = 6.9 \times 10^{-3} \left(\frac{b \Sigma_o}{10} \right)^{1/2} \left(\frac{a}{1 \text{au}} \right)^{(1/2)(4-\alpha)} \text{au} \quad (2)$$

Here, a is the semimajor axis of the embryo, and b is the number of hill radii between each embryo. Both the embryos and the planetesimals are distributed assuming some surface density profile for the disc. The general equation is:

$$\Sigma = \Sigma_o \left(\frac{a}{1 \text{au}} \right)^{-\alpha} \quad (3)$$

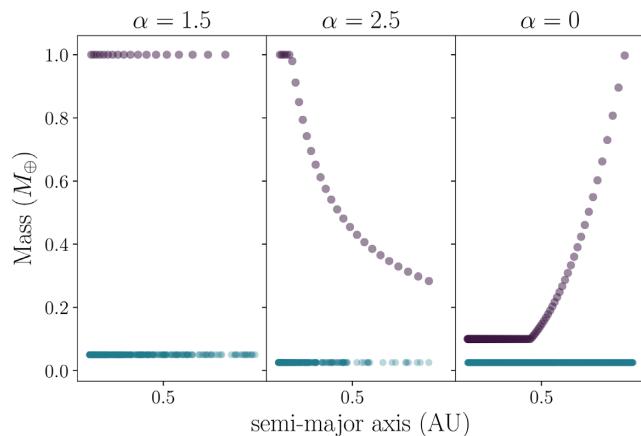


Figure 3. Mass distribution of the planetary disc for three surface density slopes: $\alpha = 3/2$, $5/2$, and 0 . $\alpha = 3/2$ is the slope of the MMSN. The upper (purple) slope corresponds to embryo masses, which has a minimum of $0.1 M_\oplus$ and is capped at $1 M_\oplus$. The lower (blue) line is the planetesimals, which all share the same mass.

where Σ_o is the normalization constant, and α is the slope (Hayashi 1981). For the MMSN, $\Sigma_o = 7$ and $\alpha = 1.5$, which gives $3.3 M_\oplus$ within 1 au. Since this is too low-mass to form super-Earths, we base our profiles off the initial conditions of Hansen & Murray (2012), as their simulations formed super-Earths with semimajor axes between 0.05 and 1 au, similar to those observed in the Kepler 11, HD69830, and 61 Virginis systems. In our fiducial simulations, we adopt $\Sigma_o = 68$ and $\alpha = 3/2$, with embryos and planetesimals placed at a range of distances between 0.1 and 1 au.

Equation (1) gives initial embryo masses that can be as large as $5 M_\oplus$, which is already a super-Earth. We place an upper limit of 0.5 or $1 M_\oplus$ on our embryos and a lower limit of $0.1 M_\oplus$ to more precisely investigate the formation process. In order to fit the same amount of mass between 0.1 and 1 au, we decrease the spacing of the embryos to a few Hill radii (r_H) apart (from the $\sim 10 r_H$ given by Kokubo & Ida 2002). We justify this by the fact that we are considering proto-embryos before they have reached the isolation stage described by equation (1).

Our simulations do not have high enough resolution for realistic-sized planetesimals, so instead we have larger bodies that represent a collection of planetesimals. Typically one collection of planetesimals is 3 per cent of the mass of an embryo. The planetesimal locations are drawn randomly from a probability distribution calculated from the surface density profile (equation 3). Both embryos and planetesimals are given initial random eccentricities less than 0.01 and inclinations less than 0.25° , creating a dynamically cold disc. Fig. 3 shows the initial distributions of embryos and planetesimals with different surface density slopes.

3.2 The simulations

We simulated the evolution of these initial conditions using a modified version of the gravitational N -body code SYMBA (Duncan, Levison & Lee 1998). The code uses a multistep symplectic integrator to efficiently resolve close encounters between bodies. It tracks the movement of particles within the disc based on gravitational interactions alone. Only embryos can gravitationally interact with each other; planetesimals have no gravitational effect on each other but do interact with the embryos. To keep the simulation from stalling we set an inner limit for the bodies'

Table 1. A description of the collision types used in our planet formation simulations. See Stewart & Leinhardt (2012) and Genda, Kokubo & Ida (2012) for the quantitative boundaries between these collision types and the fraction of the colliding mass ejected into space.

Number	Collision name	Description
1	Supercatastrophic disruption	Only debris is left
2	Catastrophic disruption	Small amount of target is left, mainly debris
3	Erosion	The target is eroded by projectile
4	Partial accretion	The projectile is partially accreted by the target
5	Hit and spray	The projectile is eroded
6	Hit and run	Two bodies graze each other and produce some debris
7	Graze	Target and projectile bounce inelastically at large angles
8	Perfect merger	Target and projectile merge
9	Graze and merge	Target and projectile impact and then merge

semimajor axis at 0.1 au, and at smaller distances the body is considered to fall into the star. We also set an outer limit at 10 au. We run the simulations to a maximum of 100 Myr, or until they are stable enough that there are no impacts for a few million years.

As mentioned in the Introduction, collisions beyond perfect mergers were implemented in *N*-body codes only recently. Kokubo & Genda (2010) only divided their collision into two types: perfect mergers and hit-and-run collisions, where the two bodies bounce off of each other (Asphaug et al. 2006). An improvement by Chambers (2013) considered each collision to result in fragmentation; the collision was only considered a perfect merger if the fragments were below a minimum mass. Both these simulations resulted in some minor improvements over perfect merger codes, but overall the process and length of planet formation were similar because most of the fragments and bodies which escaped the imperfect collisions remained in the same general orbits. So even if part of the projectile was not accreted during the first collision, the debris eventually collided with the same embryo and accreted on to it at a later time. Thus, the planets formed out of essentially the same material and in the same places (Chambers 2013). The aim of these simulations was to reproduce the Solar system, so they never tested higher velocity discs to produce a broader range of compositions.

We improve upon this work in two aspects: we use a more realistic collisional outcome map, and we include the production and loss of debris.

The version of SYMBA we use is updated to use the analytic collision prescriptions of Stewart & Leinhardt (2012), Leinhardt & Stewart (2012), and Genda et al. (2012) to realistically treat collisions between large bodies. During each collision, the algorithm in the appendix of Leinhardt & Stewart (2012) is used to determine the outcome with only small modifications. We sort the collision outcomes into nine types, as described in Table 1, which are similar to those in Leinhardt & Stewart (2012). Leinhardt & Stewart (2012) divide the collision outcomes into two regions, the disruption and grazing regimes. The disruption regime is subdivided into the supercatastrophic, erosion, and partial accretion subdivisions. We rename the supercatastrophic regime as the catastrophic regime, and define the supercatastrophic regime when the mass of the largest remnant is no longer distinct from the power-law distribution of debris from the collision, an outcome not in Leinhardt & Stewart (2012). We also name the divisions in the grazing regime the graze and merge, graze, hit and run, and hit and spray regimes to keep them distinct from the disruption regime subdivisions. The criteria from Genda et al. (2012) is used to more accurately define the boundary between graze-and-merge and hit-and-run.

Table 2. Parameters varied in initial conditions. For the most part, we perform simulations with each permutation of the first three parameters. Table A1 details the parameter combinations for each simulation set.

Maximum initial embryo mass (M_{\oplus})	Disc mass (M_{\oplus})	α	Debris mass lost (%)
0.1	25	0	0
0.5	50	3/2	100
1.0		5/2	

In the case of a hit-and-spray, graze-and-merge, perfect merger, or any of the disruption regime collisions, there is distinct from the debris a single surviving largest remnant made mostly from target materials, which we place at the centre of mass and momentum of the system of colliding material. In the other cases, there is also a surviving second largest remnant made mostly from projectile material, which we then place a mutual Hill radii away from the largest remnant along the collision trajectory. The mutual velocities are determined by assuming that the collision is inelastic with a coefficient of restitution of 0.5 along the collision vector, but elastic in the two other directions. The debris particles are placed in a ring orthogonal to the collision vector a Hill radius away from the largest remnant and moving away from it with a relative velocity of 5 per cent greater than their mutual escape speed (similar to Chambers 2013). For computational tractability, we limit the number of debris particles to 38 and set a minimum mass of debris particles to $1.5 \times 10^{-5} M_{\oplus}$. If there is not enough debris mass, the number of debris particles is decreased accordingly. The total mass and momentum of the system is conserved unless some debris is removed because it is considered to have been turned into dust.

We also include the effects of radiation pressure on the debris formed in the collisions. Extending the debris size function developed in Stewart & Leinhardt (2012) to smaller scales suggests that ~ 3 per cent of the debris mass is small enough ($< 0.1 \mu\text{m}$) to be pushed away by radiation pressure. Continued collisional grinding, a process by which debris particles repeatedly collide with each other and fragment into smaller and smaller pieces, should increase this fraction significantly. As much as 20–50 per cent of the debris could be ground down small enough to be blown away by radiation pressure in just a few orbits (Jackson & Wyatt 2012). We simulate this effect by immediately removing a fraction of the debris mass created in each collision. Table 2 shows the main two fractions of debris we remove (0 and 100 per cent). Test simulations were performed with 3 per cent of the debris mass lost, however we

found 3 per cent was such a small fraction it had little effect. We remove 100 per cent (instead of 50 per cent, for example) in order to test the maximum impact of collisional grinding. Additionally, comparing the two allows us to compare the impact of debris on the outcome of the simulations. If the removal of debris turned out to be the crucial ingredient to reproduce the high Fe/Mg ratio of some exoplanets, we would focus in a subsequent work on modelling more precisely the collisional cascade within the debris population.

3.3 Composition

We track the collisions that each body experiences used to this to calculate the composition of the bodies in post-processing. All bodies are given the same initial composition based on the solar photosphere Fe/Mg abundance ratio, as we assume a solar-type star for our exoplanet systems. For consistency, we use SUPEREARTH (Plotnykov and Valencia, in preparation) to calculate the corresponding CMF, as well as all of the final planets' CMFs. The mantle is made of olivine and pyroxenes in the upper mantle, olivine, and garnet in the transition zone, bridgemanite and ferropericlase in the lower mantle and post-perovskite and ferropericlase in the lowermost mantle. For simplicity, and to consider the most favourable case for iron enrichment of planets through collisions, we consider iron to be only in the core and use the Mg end-member for all minerals in the mantle (e.g. complete differentiation). We find that a solar Fe/Mg abundance ratio of 2.001 (Palme, Lodders & Jones 2014) gives a CMF of ~ 0.333 . This is larger than the Earth's (0.326, Stacey 2005) as the Earth is slightly depleted in iron with respect to the Sun (McDonough & Sun 1995).

We assume our embryos to start out completely differentiated into an iron core and mantle. Our initial embryos have masses between 0.1 and $1 M_{\oplus}$, for which we can expect a large degree of differentiation. For comparison, Mars has a mass of $0.1 M_{\oplus}$, and Mars and the Earth are thought to have 20 per cent and 10 per cent by mol of iron in the mantle, respectively (Waenke & Dreibus 1988). In addition, giant impacts deliver energy that raises the temperature of the mantle. The temperatures can rise beyond the solidus to melt the mantles of growing planets, which enables some of the remaining iron in the mantle to sink to the core, enhancing differentiation (Tonks & Melosh 1992; Nakajima & Stevenson 2015). Thus, we consider our assumption of complete differentiation to be a valid one that yields the most variation in composition expected.

To calculate the final composition of the planets, we create prescriptions for how each collision type changes the compositions of the interacting bodies. Core and mantle material are conserved throughout the collisions, but due to the loss of material (i.e. via radiation pressure or planets falling into the star) there is no overall conservation of the abundance ratios in the disc. Accretion events, where small bodies (debris or planetesimals) accrete on to an embryo, are treated separately. The core and mantle mass of the small body is added to that of the embryo, and it is assumed that the embryo remains completely differentiated. For the collisions between embryos, our compositional map is split up into two main categories: disruptive and grazing. The disruptive events are those that leave one embryo and some debris.

We chose the prescription for the disruptive collisions to follow the prescription developed in Marcus et al. (2010), where the collisions either always result in the cores of both bodies merging, or only result in core mergers if the remaining body is more massive than the target body. The final planet CMFs are calculated assuming the first scenario, giving a maximal CMF, and then with the second scenario, giving a minimal CMF for each planet. We take the

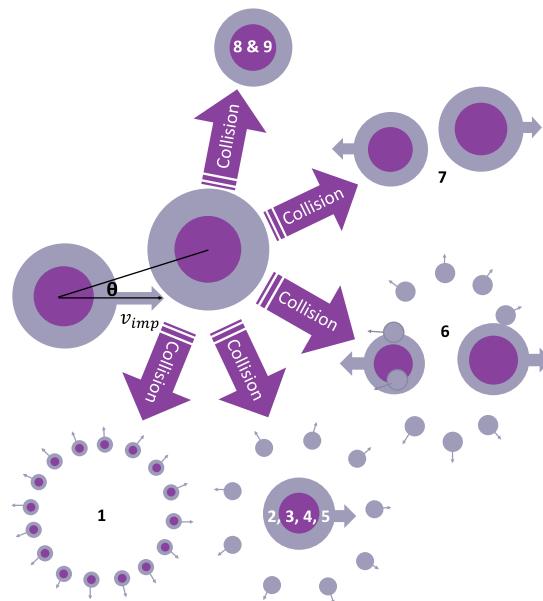


Figure 4. Visualization of the morphology and change in CMF for all collision types. The numbers correspond to the collision types in Table 1. Purple represents the core, and grey the mantle. The moving body is the least massive, and termed the ‘impactor’. The body it hits is called the ‘target’. Sizes are not representative of actual mass ratios.

average of these two outcomes, and call this the ‘merging disruption’ prescription. In the case of the most disruptive events, where there is no embryo left (supercatastrophic), we simply assume that the core material is evenly distributed throughout the debris.

On the other extreme, as disruptive events become closer to the hit-and-run collisions, their energy decreases, and two embryos are left instead of one. As per the simulations of Asphaug (2010) and Leinhardt & Stewart (2012), the projectile can strip mantle material from the target or the projectile depending on the velocity of the impact. This is the only significant way that the collision can impact the CMF of the final body (or bodies, Jackson, 2020, Private communication). So our prescription still follows that of Marcus et al. (2010), where the largest remnant has the same core mass as the target body, and only the mass of the mantle changes. Similarly, the second largest remnant will keep the same core mass as the impactor, and thus the debris from these collisions is typically mantle material. Fig. 4 provides a visual representation of the collision geometry and collision types and their possible outcomes. For a more quantitative and detailed description of the compositional prescriptions we use, see Appendix B.

4 RESULTS OF SIMULATIONS

We ran 120 simulations around $1 M_{\odot}$ stars that span the parameters listed in Table 2: disc surface density power-law indexes of 2.5, 1.5, and 0, maximum embryo masses of 0.5 and $1 M_{\oplus}$, total disc masses of 25 and $50 M_{\oplus}$, and debris mass loss of 0 and 100 per cent. Table A1 shows the breakdown of these simulations. The typical time-scale of accretion for a planet is 3 Myr. This short time-scale is a consequence of massive discs concentrated in small annuli with fast orbital time-scales. Fig. 5 shows the evolution of a typical system, where planets have evolved to close to final mass in 1 Myr.

We stop simulations at 10–50 Myr, when there are no collisions of any type in the last ~ 2 Myr. Although a small mass of debris and

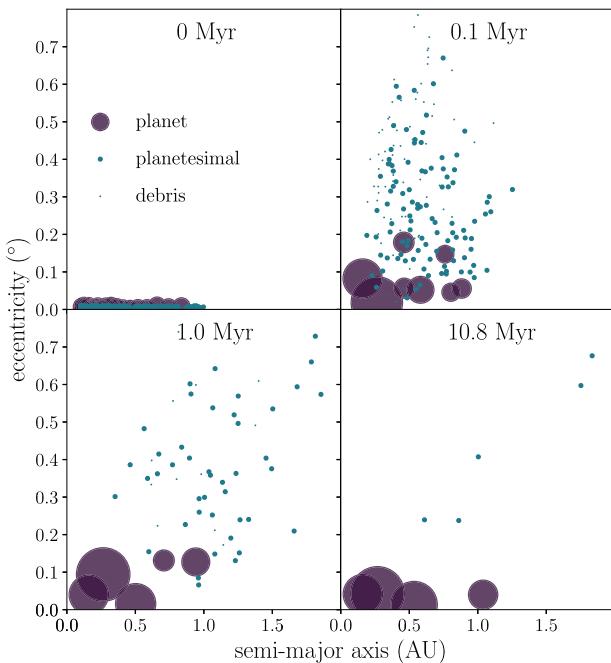


Figure 5. The evolution of a fiducial (maximum $M_{\text{embryo}} = 1.0 M_{\oplus}$, $M_{\text{disc}} = 25 M_{\oplus}$, $\alpha = 1.5$, and no debris loss) simulation over 10 Myr. The size of the circle in the plot corresponds to the mass of the body. The embryos are purple, and the planetesimals and debris are blue. The mass begins concentrated inside 1 au, and large planets quickly form close to the star. Over time they consolidate down to a few planets inside 1 au, and most of the small debris is accreted by the planets or thrown out of the system through orbital interactions. The planetesimals achieve such high eccentricities as there is no drag from the gas disc to damp the perturbations from close encounters.

planetesimals still exists in the systems at the end of the simulations, much of it is on increasingly wide, eccentric orbits, and moving further from the Sun. One of these planetesimals can be seen in the final panel of Fig. 5, and typically debris will extend out to larger distances (i.e. 4 au). This leaves a negligible mass of debris that would accrete on to the planets, barely affecting the CMF. As such, it can be ignored for our purposes.

On average across the whole set of runs, each simulation forms four planets, with an average mass of $5 M_{\oplus}$. The average total mass in planets per simulation is $22 M_{\oplus}$, and the total average mass of debris produced is $1.14 M_{\oplus}$. The majority of the planets are between 0.1 and 1 au, though planets extend out to about 4 au.

By design, and to keep computational costs reasonable, we run the majority of our simulations with discs between 0.1 and 1 au. However, we also ran a few simulations with discs that extended from 0.05 to 1 au to investigate shorter period planets and see how their collisional histories would differ from the planets forming further out. We ran one simulation for each of the parameter configurations with maximum embryo masses of 0.5 and $1.0 M_{\oplus}$ (so 24 in total). Although the comparison size was small, we did not find any significant differences. Fig. 6 shows the planets we form (discs starting at 0.1 au are circles, discs starting at 0.05 au are triangles) compared to the 44 rocky super-Earths with good mass estimations (blue), which are located between 0.01 and 0.3 au.

As can be seen, we can successfully grow planets to $5-15 M_{\oplus}$ between 0.1 and 1.0 au in our simulations, and there is strong indication from these test simulations that the same formation

scenario would form planets at shorter periods if the disc is extended closer to the central star.

The planets have eccentricities from 0 to 0.7, with an average of 0.09. Fig. 7 compares the spread of eccentricities to those measured for super-Earths (with mass $< 20 M_{\oplus}$, radius $< 2.1 R_{\oplus}$, and around K type or larger stars) from the NASA Exoplanet Catalogue as of 2019 July, demonstrating that they roughly overlap. The high eccentricity spike seen at lower masses in our simulations is likely due to the lack of damping otherwise caused by a gas disc. Including a gas disc would reduce the excitation of eccentricities due to close encounters, resulting in lower planet eccentricities overall. This would also lead to lower impact velocities and less debris.

4.1 Collisions and final compositions

After all the simulations are complete we analyse the collisional history of each planet formed and the population as a whole. The collisions in these simulations are divided into two types as mentioned in Section 3.3: those between planetesimals and embryos (perfect accretion), and those between two embryos (giant collisions). The small collisions account for ~ 95 per cent of the collisions in the simulations overall, and have little effect on the composition of the final planets. All the planetesimals are given the same initial CMF ratio as the embryos, so the only effect on composition occurs when the small body is a debris particle from a previous collision that may have a different composition (i.e. 0 CMF).

As we are interested in how composition is affected, we focus on the giant collisions. Fig. 8 shows all the giant collisions that our simulations produced as a function of the ratio between impact and escape velocity and of the impact angle. This collisional map can be compared to the one proposed by Stewart & Leinhardt (2012). It is clear that during planet formation all impact angles are well sampled during collisions, with frequency peaking around 45° . In contrast, reaching impact velocities exceeding more than 50 per cent of the escape velocity of the combined mass of the colliding bodies (i.e. mutual escape velocity) is less common. This is as expected, since the impact velocities depend on the velocity dispersion of the disc, and the velocity dispersion of these discs is low. As there are no giant planets to perturb the disc in these simulations, velocity dispersion depends on the strength of the largest close encounters that perturb the orbits of bodies before they collide. The strength of such encounters are set by the mutual escape velocity, and so the relative velocities (and thus impact velocities) of embryos scale with the escape velocities of the embryos. The lack of collisions at high impact to mutual escape velocity ratios translates to very few erosion, catastrophic and supercatastrophic disruption collisions.

We show more quantitatively the prevalence of each type of collision in the overall formation history of our simulations, as well as the average amount of debris created by each type of collision, in Fig. 9. A strong conclusion from our simulations is that the most common collisions during planet formation were those that produce little or no debris, and that the collisions with the most potential to change planetary composition were very infrequent.

More specifically, Fig. 9 shows that graze and merge collisions were the most predominant collision type, and like perfect mergers they do not modify planet composition. Partial accretion was the next most common collisional outcome, followed by perfect mergers, graze, hit-and-run, and hit-and-spray. Finally, as previously mentioned, the disruptive collisions such as erosion, catastrophic and supercatastrophic were rare events.

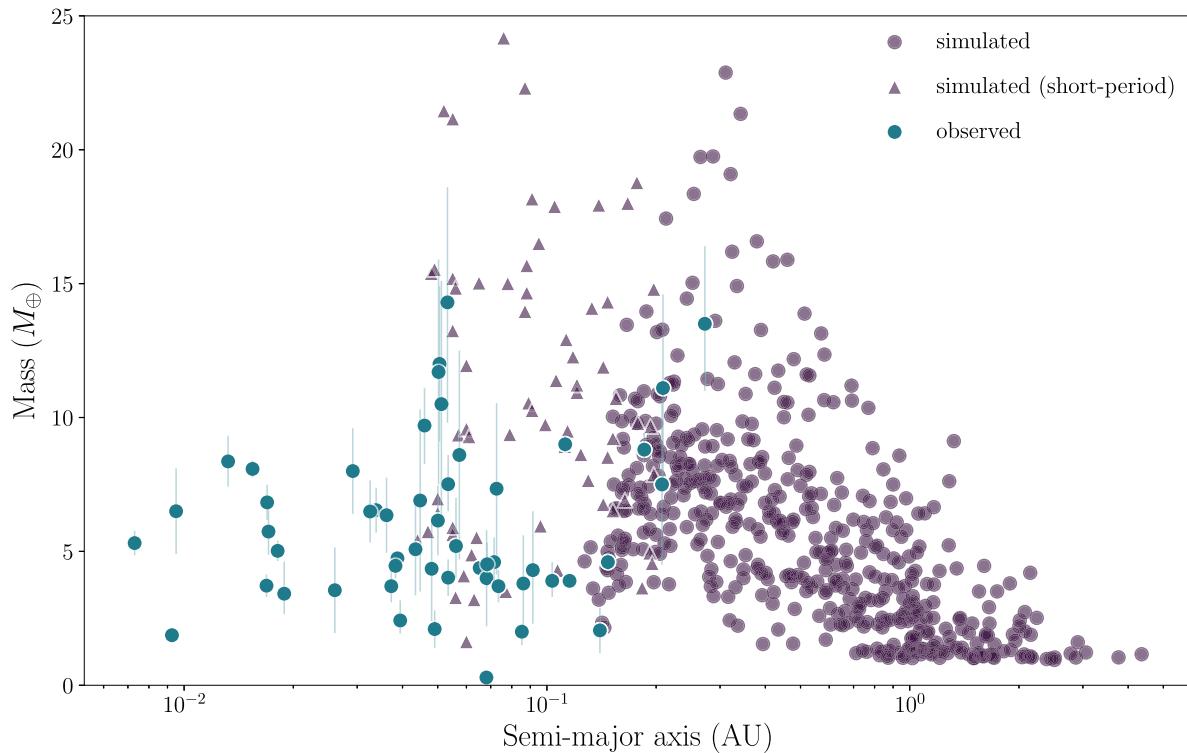


Figure 6. The mass and semimajor axis of observed (blue) and simulated (purple) super-Earths. Circular points are from simulations with discs from 0.1 to 1 au. Triangular points are from a smaller suite of simulations of super-Earths formed from discs from 0.05 to 1.0 au, where only the planets formed inside 0.2 au are plotted.

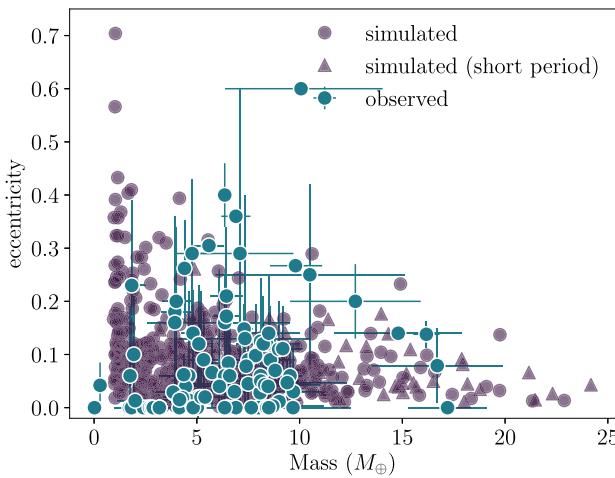


Figure 7. The eccentricities of the simulated planets (purple), compared to measured eccentricities of super-Earths from the NASA Exoplanet Archive (blue).

By tracking debris for each collision we find that the average debris mass created in a collision less disruptive than erosion is less than 10 per cent of the total colliding mass (see Fig. 9). If all of this debris were mantle material, this would only produce a 10 percent increase in CMF. This is contrasted with the catastrophic disruption and erosion collisions, where 50–90 per cent of the colliding mass becomes debris. This would produce a much more significant change in CMF of the remaining embryo, increasing it by a factor of 2 or more, and could even reduce the embryo to almost a bare core. If instead we consider a fixed mass of debris, then the

change in the final embryo's CMF will be the largest for a small embryo, and will decrease as the embryo mass increases. This is because a smaller mass fraction of mantle is being removed. So it is also the case that the larger embryos will tend to have a smaller CMF change.

In sum, the predominance of gentle collisions and scarcity of disruptive collisions translates to a difficulty in changing the composition of the initial embryos. We show the cumulative effects of giant impact collisions for both composition prescriptions in Fig. 10 by obtaining a histogram of the core-mass enrichment or depletion after planet formation. The spread in the resulting distribution of CMF is produced by the stochasticity of giant impact collisions during planet formation, and the spread is nevertheless narrow due to the dominance of collisions that produced very little debris.

4.2 Mass–radius from growing planets

We translate the mass and composition of the planets to a final radius using the internal structure code SUPEREARTH (Plotnykov and Valencia, in preparation), and show the results in Fig. 1. The radii for the maximum and minimum CMF scenarios are plotted as error bars on the averaged CMF. We compare our simulated planets to those super-Earths for which there are good mass estimates ($\Delta M/M < 0.5$), as discussed in Section 2.

It is clear that giant impact collisions can increase the core component up to CMFs around Mercury's. This is only the case when considering the maximum CMF enrichment from the prescription, where cores always merge. The average values from this prescription give a maximum CMF of 0.6, which translates to an

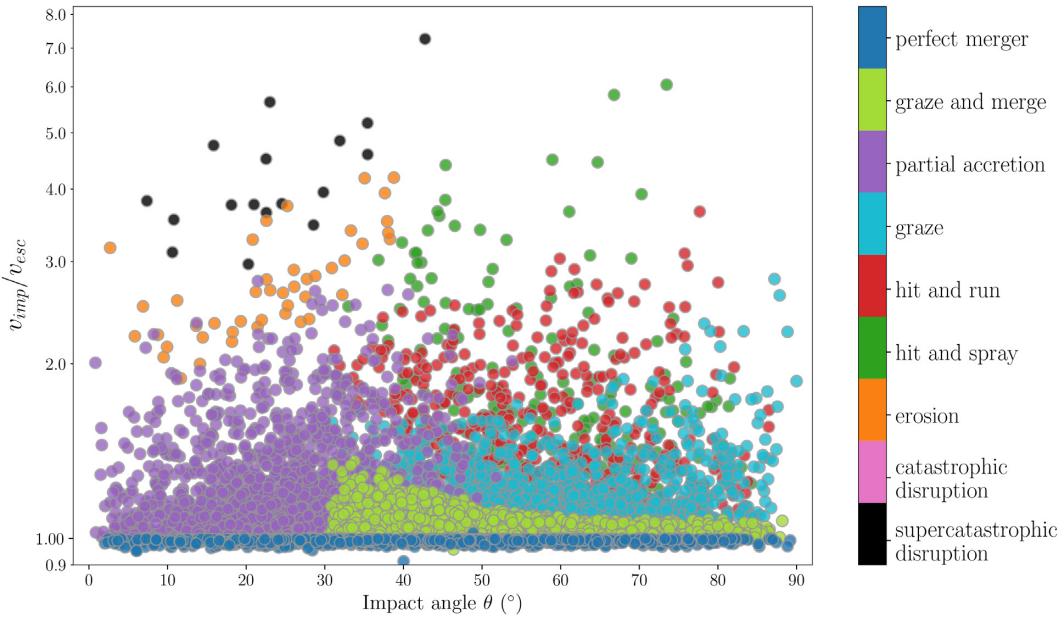


Figure 8. Collisional map for embryos during planet formation. The impact velocity is scaled by the escape velocity of the combined mass of the two impacting bodies. There is a wide spread of collisions across angles, but the majority of collisions occur at low impact velocity ratios.

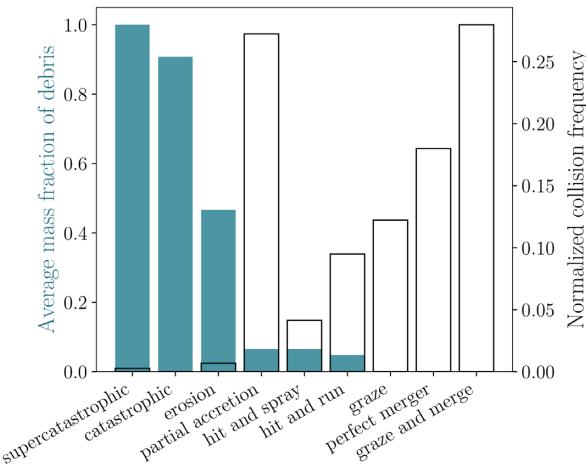


Figure 9. The normalized frequency of each collision type over the simulation suite (outlined purple), combined with the average fraction of the collision mass that becomes debris for each collision type (filled-in blue). The least frequent collisions (i.e. catastrophic disruption) create the most debris.

enrichment of Fe/Mg by a factor of ~ 3 . In comparison, the super-Earth data show some planets that appear to be even more enriched.

For the planets for which we obtained the most iron enrichment, the collisional history shows that they tend to be formed by one or two CMF changing collisions, such as majorly erosive collisions or those with core merging, with additional collisions performing minor CMF enrichment. On the other end, not many iron-depleted planets were produced, since the prescription did not allow for much of a reduction in CMF. The only pathway to iron depletion is for a planet to accrete large amounts of mantle debris.

Roughly 29 out of the 44 observed super-Earths fall within the mass-radius range of the simulated planets, when considering their error bars. It is thus reasonable to conclude that giant impact collisions can explain some of the iron enrichment seen in the rocky

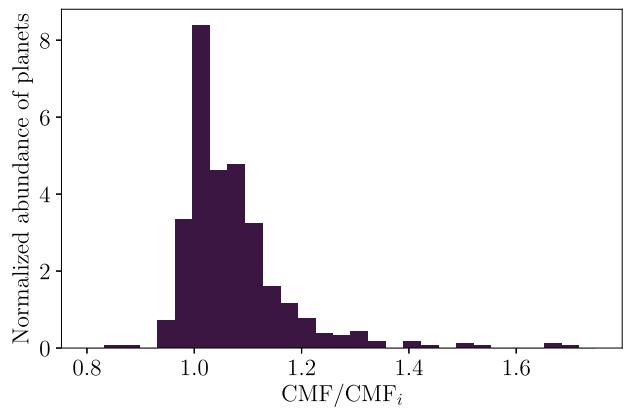


Figure 10. A histogram of the CMFs of the final planets in all the simulations. The CMFs are given as a ratio of final CMF over the initial CMF (0.333) to show the change in CMF of each planet. Using the maximum and minimum CMF produced instead of the average CMF only results in a < 6 per cent difference in the histogram bin heights.

super-Earth population. The simulations have difficulty achieving even the lowest CMF estimated for Mercury, so it is clear that more work needs to be done to understand how high CMF planets form. Planets like K2-38b (Sinukoff et al. 2016), K2-3d (Crossfield et al. 2015; Almenara et al. 2015; Dai et al. 2016), and Kepler-100b (Marcy et al. 2014; Borucki et al. 2011; Batalha et al. 2013) have data consistent with almost pure iron content. Forming a pure iron planet is difficult. Condensation temperatures of iron alloys and silicate oxides are similar, so condensation of these materials from the nebula happens within a few hundred Kelvin of each other. In addition, the giant impact phase does not seem to be energetic enough to completely remove the mantle of embryos. In fact, even at very large impact velocities of five times that of mutual escape of $8 M_{\oplus}$ planets, there is always some mantle left over (Marcus et al. 2009). One would have to invoke a series of these catastrophic events to produce an iron planet. However, as seen in

our calculations, these super energetic collisions are rare. Further observational work for narrowing the error in mass of the extremely iron-enriched planets like Kepler-100b is therefore fundamental for testing formation theories. With current error bars, these planets could be either close to the simulated CMF range or far outside it, and thus pose a challenge to formation theories.

4.3 Effect of initial CMF

We assume an initial Fe/Mg ratio that is constant across the planetary disc. This is consistent with the assumption that the disc formed from the same material that formed the star. Our simulations do not start until later in the disc's evolution, however, and by this time the embryos have formed through accretion and collisions. This formation process could have already altered the CMF of the embryos from the primordial composition. Indeed, since some of our embryos start at $1 M_{\oplus}$, it seems that these may have undergone some collisional evolution already. Thus, instead of a constant initial CMF, these embryos could have some initial distribution of CMFs.

A test run that began with $0.1 M_{\oplus}$ embryos confirmed this. Once the embryos reached $1 M_{\oplus}$, they had a CMF range of 0.3–0.4. To replicate this effect without the extra simulation time needed to start all embryos at $0.1 M_{\oplus}$, we use the simulations that start at 1 and $0.5 M_{\oplus}$ embryos, but perform multiple post-processing runs, one with initial CMF = 0.3 and another with CMF = 0.4. Each planet then has two possible CMFs, one from each run. These are considered to define the range of CMFs for each planet. This of course results in a wider spread of CMFs for planets in these runs than those with initial CMF = 0.333. As expected, the increase in the spread depends on the spread of the initial CMF given to the embryos. Given our data, we can only plausibly expect $\Delta \text{CMF}_i = 0.1$, which corresponds to an increase in the maximum CMF of ~ 0.1 . Simulations by Carter et al. (2015) of the collisional formation of embryos show that their variation in CMF could reach up to a maximum of $\Delta \text{CMF}_i = 0.1$, corroborating our results.

When given a singular initial CMF of 0.333, the final planets have a maximum CMF enhancement of $\text{CMF}/\text{CMF}_i = 1.7$ or an enrichment of Fe/Mg of 2.8. When given a spread of initial CMF, the maximum CMF enhancement increases to CMF/CMF_i to 2, and an Fe/Mg enrichment of 4. This is still not enough to form the most iron-rich planets, like Kepler-100b, making planets with much higher enrichment than 4 seem puzzling to form.

4.4 Effect of parameters

The parameters varied in our simulations are to determine the importance of different disc structures, embryo mass, and debris on composition. We find that, while most of these parameters have some effect on the CMF distribution, the size of the effect tends to be small. The overall shape of the CMF distribution as seen in Fig. 10 remains roughly the same. Since the distribution is so sharply peaked, we use the height of the peak to identify the effect of each of the following parameters.

4.4.1 Embryo mass

The mass of an embryo is related to the stage of its evolution. Smaller embryos ($0.5 M_{\oplus}$) are less evolved, and as such require more collisions to form the same mass planets as larger ($1.0 M_{\oplus}$) embryos. This is further illustrating the effect of initial CMF as discussed in Section 4.3. In this case, as the smaller embryos are

a factor of 2 less massive than the $1 M_{\oplus}$ embryos, there are also two times as many embryos packed into the same disc size, which increases the chances of collisions. Therefore, we expect smaller embryos to increase the number of giant collisions. We see a roughly twofold increase in the number of giant impacts in the simulations, which increases the chance of CMF-changing collisions. The final CMF distribution seen in Fig. 11 is correspondingly slightly wider for the smaller embryos.

4.4.2 Planetary disc mass

Different disc masses are meant to emulate the variation in total exoplanet system masses. Increased disc mass translates to more massive planetesimals and embryos, in addition to an increase in the number of embryos inside the disc radius. As was the case with smaller embryos, there are more collisions between embryos and thus more opportunities for CMF change. This caused $50 M_{\oplus}$ disc simulations to have a slightly wider distribution than the $25 M_{\oplus}$ simulations, as can be seen from Fig. 11. The $50 M_{\oplus}$ disc also produces planets that are $20 M_{\oplus}$ and larger, which is beyond the maximum mass of rocky super-Earths we observe.

4.4.3 Disc slope

Disc surface density slope affects the distribution and size of embryos, which causes the small differences seen in the CMF distribution for each slope. Discs with slopes of $\alpha = 0$ and 2.5 have more small embryos than those with a slope of $\alpha = 1.5$, as can be seen in Fig. 3. The smaller embryos must accrete more mass to form planets, thus there are more giant collisions in runs with $\alpha = 0$ and 2.5 . $\alpha = 2.5$ runs have two times more collisions than $\alpha = 1.5$, and $\alpha = 0$ discs have four times more collisions than $\alpha = 1.5$.

In accordance with the above sections, more collisions translate to a wider spread of CMF, and thus the $\alpha = 1.5$ runs have the most strongly peaked CMF distribution. Fig. 11 shows that the $\alpha = 0$ runs are slightly less peaked, and the $\alpha = 2.5$ runs have the widest spread of CMF. Otherwise, there seems to be little correlation between disc slope and the CMF distribution of the final planets.

4.4.4 Debris loss

Removing some (or all) of the debris mass created in a collision is meant to emulate the removal of very small debris particles via radiation pressure, as discussed in Section 3.2. We removed 100 percent of the debris mass as soon as it was created in about half of the simulations. While more than expected due to collisional grinding, this was done to test the maximum impact of collisional grinding, and as a test of the importance of the debris to planet formation and composition. Debris provides dynamical friction, which could change the orbital evolution of the planetary system. It is also key for the composition of the planets. Debris is mostly mantle material, as this is easier to strip during collisions than the core. Consequently, its removal should allow for more efficient mantle stripping of the bodies during the simulation, as the debris mass can no longer re-accrete on to its original parent bodies. This is an effect that was observed in the few simulations of planet formation that have been performed with debris-producing collisions (Chambers 2013).

We can see some evidence that there was re-accretion of debris on to the original bodies. The CMF distribution for the 0 per cent debris loss runs, seen in Fig. 11, is more peaked around the

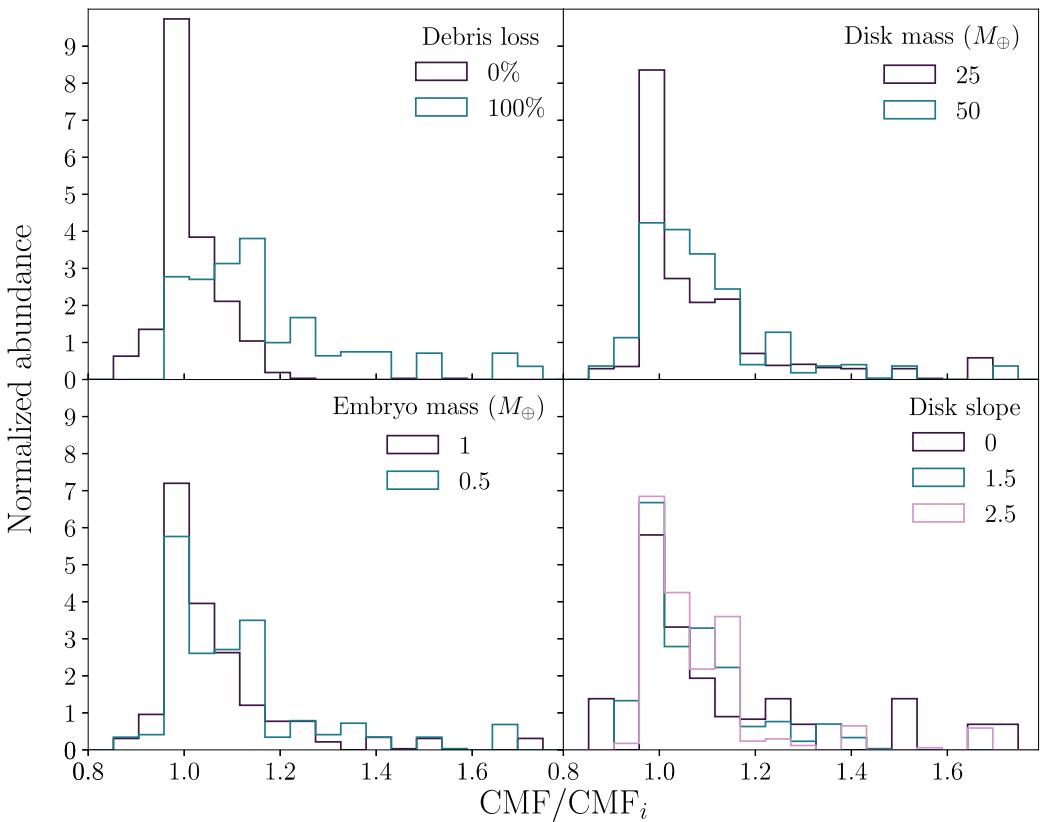


Figure 11. Effect of parameters on final planet CMF distributions. Each panel is a normalized CMF histograms, where CMF is shown as a ratio of final planet CMF over initial CMF (0.333). Each panel shows the effect of a parameter on the histogram. Top left: debris-loss changes the distribution, decreasing the peak at the initial CMF and preferentially enriching bodies. Top right: increased disc mass widens the distribution slightly. Bottom left: decreased embryo mass similarly decreases the peak of the distribution. Bottom right: disc slopes other than 1.5 slightly decrease the main peak of the distribution.

initial distribution, and the 100 per cent debris loss histogram peaks at a slightly enriched CMF due to the preferential loss of mantle material in the form of debris. In addition, the 100 per cent debris-loss histogram extends to higher CMFs, and includes almost all of the high CMF planets. Thus, it is clear that the accretion of mantle debris particles is a major factor in determining the CMF of final planets. Without that debris, there are also no planets that have CMF ratios lower than the initial, as accretion of mantle debris is the only way for a planet to decrease its CMF. The importance of debris in determining the range of final CMFs points to the importance of understanding collisional grinding of debris and the subsequent loss of debris mass it causes.

5 DISCUSSION

In this section, we discuss the implications of some of the assumptions we have made in our work.

5.1 Stellar Fe/Mg abundance

Throughout the simulations, we assume that all the embryos start with a single primordial composition determined by the stellar composition, and we use solar values to compute the radius of the final planets (as discussed in Section 3.3). However, there is a spread in the stellar Fe/Mg abundances for stars that host planets (see Fig. 2). Thus, it may be that the most iron-rich observed super-Earths (i.e. Kepler-100b) come from a star that is more iron rich than

the Sun. However, one would need to invoke a substantially iron-rich primordial composition to explain the most iron-rich observed super-Earths, and stars that differ beyond four times more Fe/Mg content from solar are much less common. It is beyond the scope of this project to address the statistical relations between the stellar Fe/Mg distribution and the distribution of observed super-Earth compositions, we leave this for future work.

5.2 Gaseous disc

The simulations do not include the effect of the gas disc on the dynamics. We have assumed, as with most terrestrial planet formation simulations (i.e. Chambers 2013; Kokubo & Genda 2010), that the oligarchic growth stage occurs late enough that the gas disc has already dissipated. However, our simulations show that super-Earths grow very quickly, so most of the action should occur within the lifetime of the disc of gas (see also Ogiwara et al. 2015). The gas disc damps the eccentricities and inclinations of the planets, thus lowering the mutual collision speed. So, our simulations maximize the possible CMF enhancement and therefore provide an upper bound to the maximal CMF that can be achieved.

Although super-Earths probably form within gaseous discs, Izidoro et al. (2017, 2019) and Ogiwara et al. (2018) showed that dynamical instabilities of super-Earth systems are likely after the removal of the gas. This leads to high-speed giant impacts among massive planets, similar to those simulated in this paper. However, we pointed out in Section 4.1 that the collisions among the most

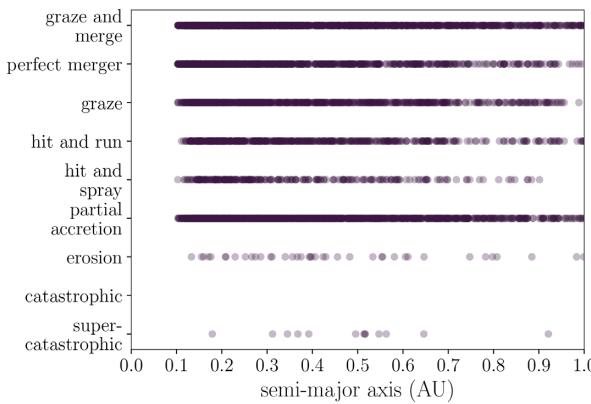


Figure 12. The frequency of each collision type with respect to semimajor axis.

massive planets typically lead to the smallest variations in CMF. Thus, compared to systems with discs, we consider our simulations a favourable case to producing the largest range of CMF, or bulk compositions, from collisions.

One caveat to the above is that during the dynamical instability phase after gas has dispersed, it is likely that there will be fewer planetesimals, as planet formation has progressed while the gaseous disc was in place. This could reduce the dynamical friction experienced by the planets, and actually increase their relative (and thus impact) velocities, increasing CMF change. Further work will explore the effect of the gaseous disc and the mass of planetesimals on the CMF distribution.

5.3 Planet semimajor axis

For computational reasons, our fiducial simulations remove objects inwards of 0.1 au. Thus, we do not reproduce exoplanets observed within this threshold distance from the central star. However, we expect that planets forming within 0.1 au should not be radically different from those we simulated here. In fact, by investigating the type of collisions as a function of semimajor axis (illustrated by Fig. 12), we find that all types of collisions happen at every location, except for a drop in supercatastrophic disruption collisions beyond 0.6 au. We attribute this feature to the lack of significant material beyond the outer edge of our initial distribution at 1 au. The only important feature is then that the frequency of all collision types increases as semimajor axis decreases, due to the increase in number density of embryos dictated by the surface density profiles chosen (see Fig. 14 for an example). Despite this, there was no corresponding correlation between CMF and semimajor axis in our simulated planets between 0.1 and 1 au. It is thus reasonable to expect that there will continue to be no correlation for planets that form between 0.01 and 0.1 au.

Orbital speeds increase the closer in the embryo is to the star, and thus it would seem that impact velocities should increase as semimajor axis decreases. Fig. 13 shows that the average impact to escape velocity ratio does appear to increase slightly as semimajor axis decreases. However, as discussed in Section 4.1, the relative velocities of embryos actually depend on the escape velocities of those embryos, and not on the distance to the star. Thus, this trend is likely due to the tendency of more massive, fully formed embryos to be closer in. Consequently, planet formation even closer to the disc should not change drastically, as it will scale with the escape velocity of the embryos like everywhere else in the disc. Even a

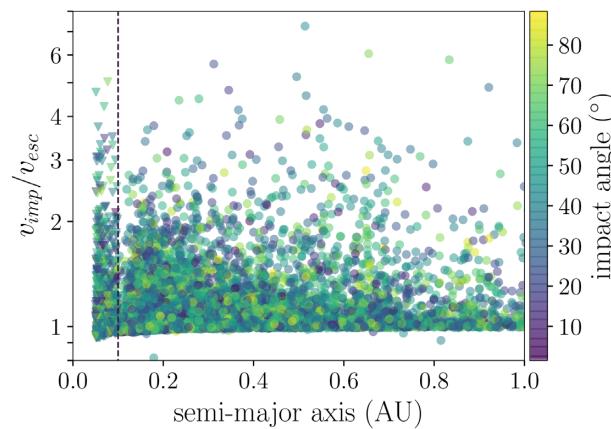


Figure 13. The impact velocity and angle of collisions with respect to semimajor axis. Collisions are shown for the main super-Earth simulations (circles), which stop at 0.1 au. To extend the picture inward, we add collisions interior to 0.1 au that occur in the short-period super-Earth simulations (triangles). The majority of collisions cluster at low impact velocity to escape velocity ratios. The maximum velocity ratio increases only slightly with decreasing semimajor axis.

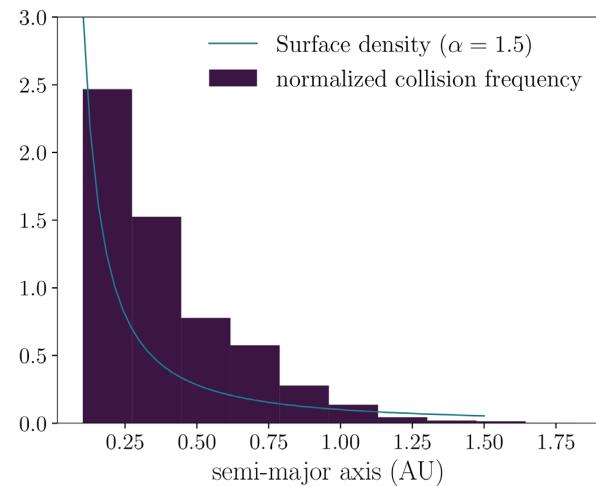


Figure 14. Partial accretion collision frequency compared to typical disc surface density profile.

slight increase in impact velocity to escape velocity ratio should not drastically impact the CMF distribution's range or shape.

To test this hypothesis we ran simulations of planetary discs with inner edges at 0.05 au, and we compare these to those that terminate at 0.1 au. We are interested in the formation of planets interior to planets formed in our main simulation suite, thus we only consider planets formed inside 0.2 au. We stopped the short-period simulations as soon as planets that formed inside 0.2 au were at least $10 r_H$ apart at their closest approach. This satisfies general stability criteria, indicating a lower probability of further collisions with embryos. These are the triangle planets shown in Figs 6 and 13. There was no significant difference in the number or type of collisions that occurred at this smaller semimajor axis, and the planets' CMFs largely fell within the range generated in the large suite of simulations. Therefore, the comparison between simulated and observed planets in this paper is strict to planets beyond 0.1 au, but these preliminary runs seem to extend our conclusions to planets inside 0.1 au as well.

6 CONCLUSIONS

We simulate the formation of rocky super-Earths during the giant impact stage including realistic collisions via an N -body code (modified SYMBA) that includes a parametrized map of collisional outcomes to investigate whether the compositional diversity of rocky super-Earths can be explained from giant collisions during formation alone. We track the change in composition after each impact by monitoring how the CMF of the embryos changes over time, assuming the embryos are differentiated into an iron core and silicate–magnesium mantle. We obtain planetary masses and compositions after planet formation which we translate to planetary radius to enable a comparison of mass–radius data of our simulated planets to those for which we have measurements. A summary of our conclusions follows:

(I) We find that although all types of collisions from the least to the most energetic happen during formation, the collisions tend to not occur at impact velocities beyond five times the mutual escape velocity, even when occurring at high orbital velocities close to the star. This is due to the fact that the relative velocities between embryos scale with mutual escape velocity, and thus are largely independent of the orbital velocity. As a result, the most prevalent type of collision is the low-energy graze-and-merge, which is equivalent to perfect merger in terms of adding the composition of the interacting embryos. The more disruptive collisions – catastrophic and erosion – with the highest impact velocity ratios and the most potential to change the composition, occur infrequently. This leaves the intermediate collision types of partial accretion, hit and spray and hit and run, as the main means to diversify composition. A single one of these collisions can rarely produce a large enough amount of debris to change composition considerably. However, they are common enough that through a series of collisions, cumulative effects can be seen in the formed planets. Thus, although the most likely fate is that composition changes little from a primordial start, the deviations from primordial are mostly from multiple intermediate collisions. The outliers tend to be produced by a collision on the more disruptive end of one of these intermediate collisions, in combination with the removal of debris material that would normally re-accrete on to the planet. This highlights the importance of understanding debris behaviour during planet formation, and further investigating the extent of collisional grinding.

(II) We find overall that collisions can provide a roughly twofold increase in CMF of a planet. Disc structure and embryo size do not affect this range, only the details in the final CMF distribution. This distribution is strongly peaked near the initial composition, and thus there is a low probability of extreme changes in CMF. Even the doubling of the initial CMF is not quite large enough to produce some of the most dense planets discovered, such as Kepler-100b. Thus, we find that realistic collisions starting from a solar composition can explain some, but not all, of the spread in composition seen in rocky super-Earths.

(III) Two possible scenarios may help explain the mismatch between our predicted composition distribution and the one inferred from mass–radius super-Earth data. Either the stellar hosts of the observed super-Earths have compositions that differ significantly from solar, and/or the planets’ mass or radius is somewhat over or underestimated. We can test either scenario with better observational data. Thus, our work highlights the need (1) to obtain the refractory ratios of the host stars with super-Earth planets as it can help us compare stellar to planet chemistry for each system and build a statistical sample overtime, and (2) to measure planetary

mass and radius at high precision to constrain the composition of the planets, especially the ones that seem to be outliers.

ACKNOWLEDGEMENTS

We thank Alan Jackson for useful discussions on compositional outcomes after planetary collisions, and the anonymous reviewer for their comments. JS and DV are supported by the Natural Sciences and Engineering Research Council of Canada (grant RBPIN-2014-06567). The research shown here acknowledges use of the Hypatia Catalog Database, an online compilation of stellar abundance data as described in Hinkel et al. (2014), which was supported by the National Aeronautics and Space Administration’s (NASA) Nexus for Exoplanet System Science (NExSS) research coordination network and the Vanderbilt Initiative in Data-Intensive Astrophysics (VIDA). This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. We would like to acknowledge that our work was performed on land traditionally inhabited by the Wendat, the Anishnaabeg, Haudenosaunee, Metis, and the Mississaugas of the New Credit First Nation.

REFERENCES

- Alibert Y. et al., 2006, *A&A*, 455, L25
- Almenara J. M. et al., 2015, *A&A*, 581, L7
- Asphaug E., 2010, *Chem. Erde/Geochem.*, 70, 199
- Asphaug E., Agnor C., Williams Q., 2006, *Nature*, 439, 155
- Barros S. C. C. et al., 2014, *A&A*, 569, A74
- Batalha N. M. et al., 2011, *ApJ*, 729, 27
- Batalha N. M. et al., 2013, *ApJS*, 204, 24
- Bond J. C., O’Brien D. P., Lauretta D. S., 2010, *ApJ*, 715, 1050
- Borucki W. J. et al., 2011, *ApJ*, 736, 19
- Carter J. A. et al., 2012, *Science*, 337, 556
- Carter P. J., Leinhardt Z. M., Elliott T., Walter M. J., Stewart S. T., 2015, *ApJ*, 813, 72
- Chambers J. E., 2001, *Icarus*, 152, 205
- Chambers J. E., 2013, *Icarus*, 224, 43
- Chatterjee S., Tan J. C., 2014, *ApJ*, 780, 53
- Chiang E., Laughlin G., 2013, *MNRAS*, 431, 3444
- Cossou C., Raymond S. N., Hersant F., Pierens A., 2014, *A&A*, 569, A56
- Crossfield I. J. M. et al., 2015, *ApJ*, 804, 10
- Dai F. et al., 2016, *ApJ*, 823, 115
- Dawson R. I., Chiang E., Lee E. J., 2015, *MNRAS*, 453, 1471
- Dumusque X. et al., 2014, *ApJ*, 789, 154
- Duncan M., Levison H., Lee M., 1998, *AJ*, 116, 2067
- Fogtmann-Schulz A., Hinrup B., Van Eylen V., Christensen-Dalsgaard J., Kjeldsen H., Silva Aguirre V., Tingley B. O., 2014, *ApJ*, 781, 67
- Genda H., Kokubo E., Ida S., 2012, *ApJ*, 744, 137
- Haghhighipour N., Maindl T. I., 2019, in de Sanctis M. C., Marchi S., eds, *The Main Belt: A Gateway to the Formation and Early Evolution of the Solar System*, Villasimius, Italy, p. 90
- Hansen B. M. S., Murray N., 2012, *ApJ*, 751, 158
- Hansen B. M. S., Murray N., 2013, *ApJ*, 775, 53
- Hatzes A. P. et al., 2011, *ApJ*, 743, 75
- Hauck S. A. et al., 2013, *J. Geophys. Res. (Planets)*, 118, 1204
- Hayashi C., 1981, *Prog. Theor. Phys. Suppl.*, 70, 35
- Hinkel N. R., Timmes F. X., Young P. A., Pagano M. D., Turnbull M. C., 2014, *AJ*, 148, 54
- Hirose K., Labrosse S., Hernlund J., 2013, *Annu. Rev. Earth Planet. Sci.*, 41, 657
- Ida S., Lin D. N. C., 2010, *ApJ*, 719, 810
- Inamdar N. K., Schlichting H. E., 2016, *ApJ*, 817, L13

- Izidoro A., Ogihara M., Raymond S. N., Morbidelli A., Pierens A., Bitsch B., Cossou C., Hersant F., 2017, *MNRAS*, 470, 1750
- Izidoro A., Bitsch B., Raymond S. N., Johansen A., Morbidelli A., Lambrechts M., Jacobson S. A., 2019, preprint ([arXiv:1902.08772](https://arxiv.org/abs/1902.08772))
- Jackson A. P., Wyatt M. C., 2012, *MNRAS*, 425, 657
- Kokubo E., Genda H., 2010, *ApJ*, 714, L21
- Kokubo E., Ida S., 2002, *ApJ*, 581, 666
- Lambrechts M., Lega E., 2017, *A&A*, 606, A146
- Lee E. J., Chiang E., 2016, *ApJ*, 817, 90
- Léger A. et al., 2009, *A&A*, 506, 287
- Leinhardt Z. M., Stewart S. T., 2012, *ApJ*, 745, 79
- Lopez E. D., Fortney J. J., 2013, *ApJ*, 776, 2
- Marcus R. A., Stewart S. T., Sasselov D., Hernquist L., 2009, *ApJ*, 700, L118
- Marcus R. A., Sasselov D., Stewart S. T., Hernquist L., 2010, *ApJ*, 719, L45
- Marcy G. W. et al., 2014, *ApJS*, 210, 20
- McDonaldough W. F., Sun S. S., 1995, *Chem. Geol.*, 120, 223
- Morrison R. A., Jackson J. M., Sturhahn W., Zhang D., Greenberg E., 2018, *J. Geophys. Res. (Solid Earth)*, 123, 4647
- Nakajima M., Stevenson D. J., 2015, *Earth Planet. Sci. Lett.*, 427, 286
- Ogihara M., Morbidelli A., Guillot T., 2015, *A&A*, 578, A36
- Ogihara M., Kokubo E., Suzuki T. K., Morbidelli A., 2018, *A&A*, 615, A63
- Owen J. E., Wu Y., 2017, *ApJ*, 847, 29
- Palme H., Lodders K., Jones A., 2014, Solar System Abundances of the Elements. Elsevier, Amsterdam, p. 15
- Raymond S., Cossou C., 2014, *MNRAS*, 440, L11
- Raymond S. N., Kokubo E., Morbidelli A., Morishima R., Walsh K. J., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI, University of Arizona Press Tucson, Tucson, p. 595 ([arXiv:1312.1689](https://arxiv.org/abs/1312.1689))
- Ringwood A. E., 1970, *Phys. Earth Planet. Inter.*, 3, 109
- Rogers L. A., Seager S., 2010, *ApJ*, 712, 974
- Sinukoff E. et al., 2016, *ApJ*, 827, 78
- Southworth J., 2011, *MNRAS*, 417, 2166
- Stacey F. D., 2005, *Rep. Prog. Phys.*, 68, 341
- Stassun K. G., Collins K. A., Gaudi B. S., 2017, *AJ*, 153, 136
- Stewart S. T., Leinhardt Z. M., 2009, *ApJ*, 691, L133
- Stewart S. T., Leinhardt Z. M., 2012, *ApJ*, 751, 32
- Stixrude L., Lithgow-Bertelloni C., 2011, *Geophys. J. Int.*, 184, 1180
- Terquem C., Papaloizou J. C. B., 2007, *ApJ*, 654, 1110
- Tonks W. B., Melosh H. J., 1992, *Icarus*, 100, 326
- Valencia D., O'Connell R. J., Sasselov D., 2006, *Icarus*, 181, 545
- Valencia D., Sasselov D. D., O'Connell R. J., 2007a, *ApJ*, 656, 545
- Valencia D., Sasselov D. D., O'Connell R. J., 2007b, *ApJ*, 665, 1413
- Valencia D., Guillot T., Parmentier V., Freedman R. S., 2013, *ApJ*, 775, 10
- Van Eylen V. et al., 2016, *ApJ*, 820, 56
- Vazan A., Ormel C. W., Dominik C., 2018, *A&A*, 610, L1
- Waenke H., Dreibus G., 1988, *Phil. Trans. Roy. Soc. Lond. Ser. A*, 325, 545

APPENDIX A: BREAKDOWN OF SIMULATION PARAMETERS AND RESULTS

The parameter space of simulations completed is discussed in Section 3.1, and the results of these simulations are explained in Section 4. Table A1 provides a more detailed breakdown of the numbers of simulations completed for each parameter set, and the results from each of these sets of simulations. The angular momentum deficit is calculated as per the equations in Chambers (2001).

Table A1. Breakdown of simulations done with discs from 0.1 to 1 au by parameters. The number of runs and initial parameters are in the left-hand section. In the right-hand section are the averaged outcomes of each group of simulations. The angular momentum deficit (AMD) is calculated for all planets above $0.9 M_{\oplus}$ as AMD_{pl} , and for all remaining bodies for AMD_{tot} .

Number of runs	Maximum initial embryo mass (M_{\oplus})	Disc mass (M_{\oplus})	α	Percent of debris mass lost (%)	Number of planets	Average planet mass (M_{\oplus})	AMD_{pl} (planets)	AMD_{tot} (total system)
1	0.1	25	1.5	0	4	4.78	0.0039	0.014
6	0.5	25	3/2	0	4	5.30	0.0054	0.015
1	0.5	25	3/2	3	4	5.1	0.0064	0.018
23	0.5	25	3/2	100	3.7	5.07	0.0097	0.029
5	0.5	50	3/2	0	4	7.62	0.011	0.033
5	0.5	25	5/2	0	4.2	3.50	0.0025	0.014
3	0.5	50	5/2	0	4.3	5.1	0.0018	0.013
5	0.5	25	0	0	4.4	6.29	0.0045	0.034
3	0.5	50	0	0	4.7	8.03	0.010	0.026
6	1.0	25	3/2	0	4	5.72	0.0089	0.017
1	1.0	25	3/2	3	5	4.5	0.0027	0.011
5	1.0	25	3/2	50	4.6	5.22	0.0028	0.013
15	1.0	25	3/2	100	4.2	5.42	0.0048	0.022
5	1.0	50	3/2	0	4.6	5.90	0.024	0.030
10	1.0	50	3/2	100	5	5.74	0.022	0.037
5	1.0	25	5/2	0	3.6	4.25	0.0094	0.027
5	1.0	50	5/2	0	3	6.29	0.0051	0.017
10	1.0	50	5/2	100	3.5	5.16	0.0071	0.028
2	1.0	25	0	0	4	7.02	0.0035	0.017
5	1.0	50	0	0	4.2	8.94	0.0044	0.021

APPENDIX B: DETAILED COMPOSITIONAL PRESCRIPTION

Section 3.3 provides a brief overview of the collision prescriptions we use. We choose prescriptions that maximize the possible change in CMF. In each collision, there is a projectile and a target. The projectile is always the less massive body. The remaining bodies are called the largest remnant and/or the second largest remnant (if there is a second embryo), and the debris.

B1 Disruptive collisions

Disruptive collisions produce one embryo and some debris. The possible outcomes of such a collision are: (1) the cores merge, (2) the core of the target remains unchanged, or the (3) core of the target is eroded by the projectile. Remaining core mass is distributed evenly throughout the debris particles. As mentioned in Section 3.3, the more disruptive the collision, the larger the iron fraction (Marcus et al. 2009). Marcus et al. (2009) find that the CMF increases following a power law:

$$\text{CMF}_{\text{new}} = \text{CMF}_{\text{init}} + 0.25(Q_R/Q_{RD}^*)^{1.65} \quad (\text{B1})$$

Above, Q_R is the centre of mass specific energy of the impact and Q_{RD}^* is the catastrophic disruption energy threshold, as defined in Stewart & Leinhardt (2009). Thus, the CMF increases as the energy of the collision increases. This is simply due to the fact that the more disruptive the collision, the more mass is stripped off the target. As we are assuming differentiated bodies, the majority of the mass eroded will be mantle material. We use the more general prescription, developed in Marcus et al. (2010), that provides results that roughly follow this power law. In this prescription, we average the outcomes of two extreme prescriptions. The first produces the

maximum CMF, and in this case cores always merge. If the largest remnant is less massive than the two cores, the remaining core and all the mantle become debris. The second prescription minimizes the final CMF, and in this case the cores of the projectile and target only merge if the largest remnant is larger than the target; essentially, if it is a partial accretion. Otherwise, the core of the largest remnant is the same mass as the target's core, and all that changes is the mass of mantle material. The projectile core and mantle become debris.

B2 Grazing collisions

Grazing collisions, or larger angle collisions, such as hit-and-run, hit-and-spray, and graze collisions fall outside the disruption criteria, and are therefore treated differently. Hit-and-spray collisions have a similar outcome to disruptive collisions, with only one embryo and some debris remaining. The rest of these graze collisions have two remaining embryos; in hit-and-run collisions they are accompanied by debris. The simulations of Asphaug et al. (2006) show that the typical outcome of a hit-and-run collision results in the erosion of the projectile, with minor erosion of the target. Since the collision angle is high, the projectile just ‘grazes’ the target, and simulations show that they just strip mantle from the target and projectile. Thus, for these types of collisions the core of the largest remnant is the same mass as the target core. If there is one embryo, the projectile core is split up in the debris. If there is another remaining embryo, it has the projectile core. The only CMF change from these collisions comes from the mantle stripping of the embryos.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.