SOLAR SYSTEM FORMATION IN THE CONTEXT OF EXTRA-SOLAR PLANETS

Sean N. Raymond
Laboratoire d’Astrophysique de Bordeaux, CNRS and Université de Bordeaux, Allée Geoffroy St. Hilaire, 33165 Pessac, France; rayray.sean@gmail.com

Andre Izidoro
UNESP, Univ. Estadual Paulista - Grupo de Dinâmica Orbital Planetologia, Guaratinguetá, CEP 12.516-410, São Paulo, Brazil

Alessandro Morbidelli
Laboratoire Lagrange, UMR 7293, Université de Nice Sophia-Antipolis, CNRS, Observatoire de la Côte d’Azur, Boulevard de l’Observatoire, 06304 Nice Cedex 4, France

Exoplanet surveys have confirmed one of humanity’s (and all teenagers’) worst fears: we are weird. If our Solar System were observed with present-day Earth technology – to put our system and exoplanets on the same footing – Jupiter is the only planet that would be detectable. The statistics of exo-Jupiters indicate that the Solar System is unusual at the ∼1% level among Sun-like stars (or ∼0.1% among all main sequence stars). But why are we different?

This review focuses on global models of planetary system formation. Successful formation models for both the Solar System and exoplanet systems rely on two key processes: orbital migration and dynamical instability. Systems of close-in ‘super-Earths’ or ‘sub-Neptunes’ cannot have formed in-situ, but instead require substantial radial inward motion of solids either as drifting mm- to cm-sized pebbles or migrating Earth-mass or larger planetary embryos. We argue that, regardless of their formation mode, the late evolution of super-Earth systems involves migration into chains of mean motion resonances anchored at the inner edge of the protoplanetary disk. The vast majority of resonant chains go unstable when the disk dissipates. The eccentricity distribution of giant exoplanets suggests that migration followed by instability is also ubiquitous in giant planet systems. We present three different models for inner Solar System formation – the low-mass asteroid belt, Grand Tack, and Early Instability models – each of which invokes a combination of migration and instability. We discuss how each model may be falsified.

We argue that most Earth-sized habitable zone exoplanets are likely to form much faster than Earth, with most of their growth complete within the disk lifetime. Their water contents should span a wide range, from dry rock-iron planets to water-rich worlds with tens of percent water. Jupiter-like planets on exterior orbits may play a central role in the formation of planets with small but non-zero, Earth-like water contents. Water loss during giant impacts and heating from short-lived radioisotopes like $^{26}$Al may also play an important role in setting the final water budgets of habitable zone planets.

Finally, we identify the key bifurcation points in planetary system formation. We present a series of events that can explain why our Solar System is so weird. Jupiter’s core must have formed fast enough to quench the growth of Earth’s building blocks by blocking the flux of pebbles drifting inward through the gaseous disk. The large Jupiter/Saturn mass ratio is rare among giant exoplanets but may be required to maintain Jupiter’s wide orbit. The giant planets’ instability must have been gentle, with no close encounters between Jupiter and Saturn, also unusual in the larger (exoplanet) context. Our Solar System system is thus the outcome of multiple unusual, but not unheard of, events.

1. INTRODUCTION

The discovery of extra-solar planets demonstrated that the current Solar System-inspired paradigm of planet formation was on the wrong track. Most extra-solar systems bear little resemblance to our well-ordered Solar System. While the Solar System is radially segregated, with small inner rocky worlds and more distant giant planets, few known exo-systems follow the same blueprint. Models designed with the goal of reproducing the Solar System failed spectacularly to understand why other planetary systems looked different than our own.

Yet exoplanets represent a huge sample of outcomes of planet formation, and new ideas for Solar System formation and evolution borrow liberally from models designed...
to explain the exoplanet population. While we are far from a complete picture, just a handful of processes may explain the broad characteristics of most exoplanet systems and the Solar System.

We review the current thinking in how Solar System formation fits in the larger context of extra-solar planetary systems. We first (§1) review observational constraints on the frequency of Solar System-like systems, pointing out specific characteristics of the Solar System that don’t fit within a simple formation picture. We then (§2) briefly summarize the stages of planet formation – from dust to full-sized planets – as they are currently understood, with liberal references to more detailed recent reviews of different steps. Next (§3) we discuss current models for the different populations of extra-solar planets and how they match quantifiable constraints. We then turn our attention to the Solar System (§4). We present the empirical constraints and a rough timeline of events in Solar System formation that includes a discussion of the Nice model for the Solar System’s (giant planet) dynamical instability. We discuss the classical model and its shortcomings, then present three newer competing models to match the important constraints of the inner Solar System. A challenge for all current models is to explain the mass deficit interior to Venus’ orbit. In §5 we extrapolate to Earth-mass planets around other stars, discussing the various formation pathways for such planets and their expected water contents. We conclude that most exo-Earths are unlikely to be truly Earth-like. Finally, in §6 we first synthesize these models into a large-scale picture of planetary system evolution, highlighting the key bifurcation points that may explain the observed diversity and the events that must have taken place to produce our own Solar System. We lay out a path for future research by showing how to use theory and observations to test current models for both exoplanet and Solar System formation.

1.1 How common are Solar Systems?

To date, radial velocity (RV) and transit surveys have discovered thousands of extra-solar planets. Figure 2 shows a sample of the diversity of detected exo-systems. These surveys have determined occurrence rates of planets as a function of planet size/mass and orbital period around different types of stars (Howard et al. 2010, 2012; Mayor et al. 2011; Fressin et al. 2013; Dong and Zhu 2013; Pettigra et al. 2013; Fulton et al. 2017). Meanwhile, gravitational microlensing and direct-imaging surveys have placed constraints on the properties of outer planetary systems (Cassan et al. 2012; Biller et al. 2013; Morz et al. 2017; Bowler and Nielsen 2018).

1 Microlensing may actually be the most sensitive method for detecting analogs to our Solar System’s giant planets. Indeed, microlensing observations have found a Jupiter-Saturn analog system (Gaudi et al. 2008; Bennett et al. 2010) as well as rough analogs to the ice giants (Poleski et al. 2013; Sumi et al. 2014). However, given that microlensing requires a precise alignment between a background source and the star whose planets can be found (e.g. Gould and Loeb 1992), it cannot be used to search for planets around a given star. Rather, its power is statistical in nature (see Gould et al. 2010; Clanton and Gaudi 2014, 2016; Suzuki et al. 2016b). Nonetheless, upcoming microlensing surveys – especially space-based surveys such as WFIRST – are expected to find hundreds to thousands of planets in the Jupiter-Saturn regions of their stars (Penny et al. 2018).

To put the Solar System on the same footing as the current sample of extra-solar planets we must determine what our system would look like when observed with present-day Earth technology. The outcome is somewhat bleak. The terrestrial planets are all too small and too low-mass to be reliably detectable. Although sub-Earth-sized planets were discovered by Kepler (e.g. Barclay et al. 2013), and ~Earth-mass planets have been found by radial velocity monitoring (e.g. Anglada-Escudé et al. 2016), they were all on close-in orbits. Strong observational biases make it extremely challenging to detect true analogs to our terrestrial planets (e.g. Charbonneau et al. 2007; Fischer et al. 2014; Winn 2018). However, a decade-long radial velocity survey would detect Jupiter orbiting the Sun. Indeed, several Jupiter analogs have been discovered (e.g. Wright et al. 2008; Saturn, Uranus and Neptune are too distant to be within the reach of radial velocity surveys. Figure 1 shows seven known Jupiter analog systems for scale.

The exo-Solar System is therefore just the Sun-Jupiter system. Observations of the Solar System as an exoplanet system would provide a decent measurement of Jupiter’s mass (really, its $m \sin i$, where $i$ is the angle between our line of sight and its orbital plane) and semimajor axis, with modest constraints on its orbital eccentricity.

Based on current data, the Sun-Jupiter system is rare at the one-in-a-thousand level. The Sun – a member of the G dwarf spectral class – is much more massive than most stars; among nearby stars only ~5% have similar masses (e.g. Chabrier 2003). Roughly 10% of Sun-like stars have gas giant planets (defined as having masses $M > 50 M_\oplus$; Butler et al. 2006; Udry and Santos 2007;Cumming et al. 2008; Mayor et al. 2011; Clanton and Gaudi 2014). However, most have orbits that are either significantly closer-in or more eccentric. Using a relatively broad definition for Jupiter-like planets – as planets with orbital radii larger than 2 AU and orbital eccentricities below 0.1 – only 10% of giant exoplanets are Jupiter-like. This puts Jupiter as a 1% case among Sun-like stars, or ~0.1% overall.

The Solar System’s peculiarity can also be considered in terms of planets that are present in other systems but absent in ours. At least 30-50% of main sequence stars have planets smaller than 4 $R_\oplus$ (or less massive than ~10 – 20 $M_\oplus$) on orbits closer-in than Mercury’s (Mayor et al. 2011; Howard et al. 2012; Fressin et al. 2013; Pettigra et al. 2013; Dong and Zhu 2013; Hsu et al. 2018; Zhu et al. 2018). Recent modeling suggests that less than 8% of planetary systems have their innermost planet on an orbit as wide as Mercury’s, and less than 3% have an innermost planet on an orbit as wide as Venus’ (Mulders et al. 2018). This reinforces the Solar System’s standing as an outsider.
Gaudi et al. (2005) star (microlensing and contains a pair of giant planets with broad ly similar properties to Jupiter and Saturn orbiting a rough ly half-Solar mass solars (see formation that does not invoke large-scale radial drift of solids (see §3.2). Raymond et al. (2005) through the planets' present-day orbits to reconstruct a disk from which they may have formed (neglecting any significant radial motion like orbital migration). This style of growth leads to systems in which planets on adjacent orbits have similar sizes (e.g., Kokubo and Ida 2002). The observed super-Earths do appear to have similar sizes within a given system (Millholland et al. 2017; Weiss et al. 2018), although it is hard to imagine a scenario for their formation that does not invoke large-scale radial drift of solids (see §3.2).

When we compare the classical model blueprint with the actual Solar System, several discrepancies emerge:

- **Why is Mars so much smaller than Earth?** Simulations of terrestrial planet formation from a smooth disk tend to produce Earth and Mars analogs with similar masses (e.g., Wetherill 1978; Chambers 2001; Raymond et al. 2006a), in contrast with the actual 9:1 mass ratio between the two planets. This is called the ‘small Mars’ problem, which was first pointed out by Wetherill (1991) and has motivated a number of models of terrestrial planet formation (see §4).

- **Why is Mercury so much smaller than Venus?** Although it receives far less attention, the large Venus/Mercury mass ratio is an even bigger problem than the ‘small Mars’ problem. The Venus/Mercury mass ratio is 14:1 but simulations again tend to produce planets with similar masses and with more compact orbital configurations than the real one. It is worth noting that, in the context of extra-solar planetary systems (in which super-Earths are extremely common), the mass deficit interior in the very inner Solar System is in itself quite puzzling.

- **Why are the asteroid and Kuiper belts so low-mass yet dynamically excited?** The asteroid and Kuiper belts contain very little mass: just \(\sim 5 \times 10^{-4} \text{ M}_\oplus\) and \(\sim 0.01\text{–}0.1 \text{ M}_\oplus\), respectively (DeMeo and Carry 2013; Gladman et al. 2001). However, both belts are dynamically excited, with much higher eccentricities and inclinations than the planets. Yet the mass required to self-excite those belts exceeds the present-day mass by many orders of magnitude (e.g.,


Fig. 2.— A sample of extra-solar systems chosen to illustrate their diversity (but not the true distribution of systems; for instance, most super-Earth systems contain only a single detected planet). The top systems were discovered by transit surveys and the bottom systems by radial velocity, although some planets in the transit systems also have radial velocity constraints (e.g., WASP-47; Sinukoff et al. 2017) and some RV-detected systems host transiting planets (e.g., 55 Cnc e; Demory et al. 2011). The size of each planet is proportional to its true size (but not on the orbital scale); for planets with only mass (or \( m \sin i \)) measurements, we used the \( M \propto R^{2.06} \) scaling from Lissauer et al. (2011). For planets more massive than 50 M\(_{\oplus}\), with eccentricities larger than 0.1 (as well as for Jupiter and Saturn), the horizontal error bar represents the planet's radial excursion over its orbit (from pericenter to apocenter). Given the logarithmic x axis, the separation between adjacent planets is a measure of their period ratio regardless of their orbital radii. This plot includes a number of systems of close-in super-Earths, including two – TRAPPIST-1 (Gillon et al. 2017; Luger et al. 2017) and Kepler-223 (Mills et al. 2016) – in which the planets have been shown to be in long chains of orbital resonances (note that the GJ 876 system also includes a 3-planet Laplace resonance among more massive planets; Rivera et al. 2014). There are systems with gas giants on eccentric orbits such as Ups And (Ford et al. 2009). The central stars vary dramatically in mass and luminosity for the different systems; for instance, the TRAPPIST-1 system orbits an ultracool dwarf star of just 0.08 M\(_{\odot}\) (Gillon et al. 2017). Some systems include planets that are smaller than Earth (e.g., the Kepler-444 system; Campante et al. 2015) and others include planets far more massive than Jupiter (e.g., HD 168443; Marcy et al. 2001). There are systems with roughly Earth-sized planets in their star’s habitable zones, notably Kepler-186 (Quintana et al. 2014), TRAPPIST-1 (Gillon et al. 2017), and GJ 667C (Ansdal-Escudé et al. 2013). Some of these planets are in multiple star systems (e.g., 55 Cnc; Fischer et al. 2008).

The diagram shows a variety of exoplanet systems, including transiting systems (M&nbs...
observed to be very abundant in young protoplanetary disks (e.g. Briccio et al. 2001; Haisch et al. 2001). Dust particles growing by coagulation in a gaseous disk encounter a number of barriers to growth such as fragmentation and bouncing (Brauer et al. 2008; Güttler et al. 2010). Once they reach roughly mm-size (or somewhat larger) particles very rapidly drift inward, leading to what was historically called the “meter-sized barrier” to growth (Weidenschilling 1977; Birnstiel et al. 2012). New models suggest that, if they are initially sufficiently concentrated relative to the gas, the streaming instability can produce clumps of drifting particles that are bound together by self-gravity and directly form 100 km-scale planetesimals (Youdin and Goodman 2005; Johansen et al. 2009, 2015; Simon et al. 2016; Yang et al. 2017). Planetesimals are the smallest macroscopic bodies that do not undergo rapid aerodynamic drift, and are often considered the building blocks of planets. Exactly where and when planetesimals form depends itself on the dynamics and structure of the disk (Drążkowska et al. 2016; Carrera et al. 2017), and some recent studies suggest that planetesimal growth may be favored near the inner edge of the disk (Drążkowska et al. 2016) and just past the snow line (Armitage et al. 2016; Drążkowska and Alibert 2017; Schoonenberg and Ormel 2017). (For reviews of dust growth/drift and planetesimal formation, see: Blum and Wurm 2008; Chiang and Youdin 2010; Johansen et al. 2014; Birnstiel et al. 2016).

Pebble- and planetesimal accretion. Planetesimals can grow by accreting other planetesimals (Greenberg et al. 1978; Wetherill and Stewart 1993; Kokubo and Ida 2000) or pebbles that continually drift inward through the disk (Johansen and Lacerda 2010; Lambrechts and Johansen 2014). Pebbles are defined as particles for which the gas drag timescale is similar to the orbital timescale, and are typically mm- to cm-sized in the terrestrial- and giant-planet-forming regions of disks (see Johansen and Lambrechts 2017). Pebbles are thought to continually grow from dust and drift inward through the disk, such that growing planetesimals see a radial flux of pebbles across their orbits (Lambrechts et al. 2014; Chambers et al. 2016; Ida et al. 2016). At low relative speeds, a large planetesimal efficiently accretes nearby small particles (either pebbles or small planetesimals) because the large planetesimal’s gravity acts to increase its effective collisional cross section (a process known as gravitational focusing; Safronov 1969; Rafikov 2004; Chambers 2006). This triggers a phase of runaway growth (Greenberg et al. 1978; Wetherill and Stewart 1993; Kokubo and Ida 1998). At later stages, growth by the accretion of other planetesimals is self-limited because the growing planetesimal excites the random velocities of nearby planetesimals, decreasing the efficiency of gravitational focusing (Kokubo and Ida 2000; Leinhardt and Richardson 2005). However, gas drag acts much more strongly on pebbles and maintains their low velocities relative to larger bodies. The efficiency of pebble accretion increases with the growing planetesimal’s mass (Lambrechts and Johansen 2012; Morbidelli and Nesvorny 2012), and pebble accretion outpaces planetesimal accretion for bodies more massive than roughly a lunar mass (0.012 M⊕; although the exact value depends on the parameters of the disk (Johansen and Lambrechts 2017). Above roughly a lunar mass these objects are generally referred to as “planetary embryos”. When an embryo reaches a critical mass called the *pebble isolation mass* it generates a pressure bump in the disk exterior to its orbit, which acts to block the inward flux of pebbles (Morbidelli and Nesvorny 2012; Lambrechts et al. 2014; Bitsch et al. 2018b). This acts to quench not only the embryo’s growth but also the growth by pebble accretion of all objects interior to the embryo’s orbit. Later growth must therefore rely on the accretion of planetesimals, other embryos or gas. We note that there is some debate about whether pebble accretion remains efficient for planets with significant gaseous envelopes below the pebble isolation mass (see Alibert 2017; Brouwers et al. 2018). (For reviews of pebble- and planetesimal accretion, see: Johansen and Lambrechts 2013; Kokubo and Ida 2002, respectively)

Gas accretion and giant planet growth. Once planetary embryos become sufficiently massive they accrete gas directly from the disk (Pollack et al. 1996; Ida and Lin 2004; Alibert et al. 2005). Gas accretion operates to some degree for Mars-mass planetary embryos, and there is evidence from noble gases in Earth’s atmosphere that a portion of the atmospheres of Earth’s constituent embryos was retained during Earth’s prolonged accretion (Dauphas 2003). Gas accretion onto a growing planet depends on the gaseous envelope opacity (Koma et al. 2005; Hubicky et al. 2005; Machida et al. 2010) and temperature, which is determined in part by the accretion rate of energy-depositing solid bodies (Rice and Armitage 2003; Broeg and Benz 2012). The dynamics of how gas is accreted onto a growing planet’s surface is affected by small-scale gas flows in the vicinity of the planet’s orbit (Fung et al. 2015; Lambrechts and Leiva 2017) as well as the structure of the circum-planetary disk (if there is one; e.g., Ayliffe and Bate 2003; Szulágyi et al. 2016). When the mass in a planet’s envelope is comparable to its solid core mass it undergoes a phase of runaway gas accretion during which the planet’s expanding Hill sphere – the zone in which the planet’s gravity dominates the star’s – puts it in contact with ever more gas, allowing it to grow to a true gas giant on the Kelvin-Helmholtz timescale (Mizuno 1980; Ida and Lin 2004; Thommes et al. 2008). This culminates with the carving of an annular gap in the disk, which slows the accretion rate (Bryden et al. 1999; Crida et al. 2006; Lubow and D’Angelo 2006). (For reviews of giant planet growth, see: Lissauer and Stevenson 2007; Hellic et al. 2014)

Orbital migration. Migration is an inevitable consequence of gravitational interactions between a growing planet and its natal gas disk. Planets launch density waves in the disk, whose flow is determined by
the disk’s dynamics. These density perturbations impart torques on the planets’ orbits (Goldreich and Tremaine 1980, Ward 1986). These torques damp planets’ eccentricities and inclinations (Papaloizou and Larwood 2000, Tanaka and Ward 2004) and also drive radial migration. For planets low enough in mass not to carve a gap in the disk (i.e., for planetary embryos and giant planet cores) the mode of migration is sometimes called type I. Multiple torques are at play. The differential Lindblad torque is almost universally negative, driving planets inward at a rate that is proportional to the planet mass (Ward 1997, Tanaka et al. 2002). In contrast, the corotation torque can in some situations be positive and overwhelm the differential Lindblad torque, leading to outward migration (Paardekooper and Mellema 2006, Kley and Crida 2008, Masset and Casoli 2013, Paardekooper et al. 2011, Benítez-Llambay et al. 2015). The disk structure plays a central role in determining the regions in which outward migration can take place (Bitsch et al. 2013, 2015a). Planets that carve a gap enter the type II migration regime, which is generally slower than type I. In this regime a planet’s migration is determined in large part by the radial gas flow within the disk and thus its viscosity (Ward 1997, Durmann and Kley 2015). Type II migration is directed inward in all but a few situations (Masset and Snellgrove 2001, Veras and Armitage 2004, Crida et al. 2009, Pierens et al. 2014). (For reviews of migration, see: Kley and Nelson 2012, Baruteau et al. 2014.)

Late-stage accretion. The gaseous component of planet-forming disks is observed to dissipate after a few million years (Haisch et al. 2001, Hillenbrand 2008). When the gas disk is gone gas accretion, pebble accretion, gas-driven migration and dust drift cease, and the final phase of growth begins. Late-stage accretion consists of a protracted phase of collisional sweep-up of remaining planetary embryos and planetesimals (Wetherill 1978, 1996, Chambers 2001, Raymond et al. 2004). Whereas some late accretion should occur for all planets, this process is thought to have played a predominant role in the growth of Earth and Venus. For these planets, this phase was characterized by giant impacts between Mars-mass or larger planetary embryos (e.g., Agnor et al. 1999, Stewart and Leinhardt 2012, Quintana et al. 2016). The very last giant impact on Earth (and perhaps in the whole inner Solar System) is thought to have been the Moon-forming impact (Benz et al. 1986, Canup and Asphaug 2001, Čuk and Stewart 2012). Dynamical friction from remnant planetesimals – which leads to an energy equipartition between planetesimals and embryos – keeps the orbits of embryos relatively circular during the early parts of late-stage accretion (O’Brien et al. 2006, Raymond et al. 2006b). While dynamical friction dwindles in importance as the planetesimal population is eroded by impacts and dynamical clearing, it is nonetheless important in setting the planets’ final orbits. Growth generally proceeds inside to out. The timescale for Solar System-like systems to complete the giant impact phase is \( \sim 100 \) Myr (Chambers 2001, Raymond et al. 2009b, Jacobson et al. 2014). Analogous terrestrial planet systems around different types of stars would have different durations of this phase (Raymond et al. 2007b, Lissauer 2007). However, some systems (e.g., those in resonant chains) may never undergo a late-stage phase of giant impacts. (For reviews of late-stage accretion see: Morbidelli et al. 2012, Raymond et al. 2014, Izidoro and Raymond 2018.)

3. EXOPLANET FORMATION MODELS

We now review formation models for extra-solar planets. We focus on two specific categories of exoplanets: super-Earth systems and giant exoplanets. For each category we summarize the observational constraints, then present the relevant models. It is interesting to note that, despite the very different regimes involved, leading models for each category rely heavily on two processes: migration and dynamical instability.

3.1 Systems of close-in low-mass planets (‘super-Earths’)

The abundance of close-in low-mass/small planets is one of the biggest surprises to date in exoplanet science. Both radial velocity and transit surveys find an occurrence rate of 30-50% for systems of planets with \( R < 4 \, \text{R}_\oplus \) or \( M < 10-20 \, \text{M}_\oplus \) and \( P < 50-100 \) days (see Winn and Fabrycky 2015, for a compilation of measured rates and a comparison between them). These planets are commonly referred to as ‘super-Earths’, even though a significant fraction appear to be gas-rich and thus closer to ‘sub-Neptunes’ (Rogers 2015, Wolfgang et al. 2016, Chen and Kipping 2017).

Super-Earth formation models are constrained by several lines of observations:

- Their high occurrence rate (Howard et al. 2010, 2012, Mayor et al. 2011, Fressin et al. 2013, Dong and Zhu 2013, Petigura et al. 2013), and the stellar mass-dependence of their properties. Low-mass stars appear to have a similar overall abundance of close-in small planets but, compared with Sun-like stars, they have more super-Earths (with \( R < 2 \, \text{R}_\oplus \)), fewer sub-Neptunes (with \( R > 2 \, \text{R}_\oplus \)) and a higher average total mass in planets (Dressing and Charbonneau 2015, Mulders et al. 2015b).

- The distribution of orbital period ratios of adjacent planets (Lissauer et al. 2011, Fabrycky et al. 2014, Steffen and Hwang 2015). Only a small fraction (perhaps 5-10%) of pairs or neighboring planets appear to be in resonance.

- The multiplicity distribution, or how many planets are detected around each star (Lissauer et al. 2011, Batalha et al. 2013, Rowe et al. 2014, Fabrycky et al. 2014). There is a large peak in the transit detection of singleton super-Earth systems compared with multiple planet systems (the so-called Kepler dichotomy;
The simplest model of super-Earth formation – in-situ – was proposed by Raymond et al. (2008b) who subsequently discarded it as unrealistic (see also Ogihara et al. 2015a). Given its recent revival (Chiang and Laughlin 2013; Hansen and Murray 2012, 2013) and surprising popularity in the exoplanet community we think it worth explaining exactly why it cannot be considered a viable model. The simplest argument against in-situ accretion is simply that it is not self-consistent. If super-Earths accreted in-situ close to their stars, then their natal planet-forming disks must have been quite dense (Raymond et al. 2008b; Chiang and Laughlin 2013; Schlaufman 2014). The timescales for accretion of these planets are very short because their disks are massive and the orbital time is short (Sakuraba 1969). Simulations demonstrate that planets similar to the observed super-Earths indeed accrete on thousand- to hundred-thousand-year timescales (Raymond et al. 2008b; Hansen and Murray 2012; Bolmont et al. 2014), long before the dispersal of gaseous disks. Thus, super-Earths must have been massive enough to gravitationally interact with the gaseous disk, which itself must have been extremely dense to accommodate the planets’ in-situ growth. The planets must therefore have migrated (Ogihara et al. 2015a). In fact, the disks required for in-situ accretion are so dense that even aerodynamic drag alone would have caused their orbits to shrink (Inamdar and Schlichting 2015).

The in-situ model is thus caught in a logical impossibility. If the planets formed in-situ then they must have migrated. But if they migrated, they did not form in-situ. In other words, super-Earths simply cannot have formed in-situ. Solids must have drifted relative to the gas, either at small scales (pebble drift) or large scales (migration). Meanwhile, there is abundant circumstantial evidence for inward planet migration. One extreme case is the existence of planets interior to the silicate sublimation radius (Swift et al. 2013). Another piece of evidence is the existence of systems of super-Earths in resonant chains in systems such as Kepler-223 (Mills et al. 2016) and TRAPPIST-1 (Gillon et al. 2017; Luger et al. 2017), given that it is extremely improbable for planets to end up in low-eccentricity resonant configurations without invoking migration.

Two models remain viable candidates to explain the origin of most super-Earths: the drift and migration models.

The drift model (Boley and Ford 2013; Chatterjee and Tria 2014, 2015; Hu et al. 2016, 2017) proposes that inward-drifting pebbles are trapped in the inner parts of the disk, perhaps at a pressure bump associated with a region toward the inner edge of the disk where its properties (e.g., its viscosity) change abruptly. Pebbles accumulate at the pressure bump until a threshold is reached for them to form planetesimals (e.g., Yang et al. 2017) and accrete into full-sized planets. Chatterjee and Tria (2014) proposed that the pressure bump itself would respond to the first planet’s presence and retreat to an external orbit, providing a new nexus of material to the in-situ model. Solids must have drifted relative to the gas, either at small scales (pebble drift) or large scales (migration).

Meanwhile, there is abundant circumstantial evidence for inward planet migration. One extreme case is the existence of planets interior to the silicate sublimation radius (Swift et al. 2013). Another piece of evidence is the existence of systems of super-Earths in resonant chains in systems such as Kepler-223 (Mills et al. 2016) and TRAPPIST-1 (Gillon et al. 2017; Luger et al. 2017), given that it is extremely improbable for planets to end up in low-eccentricity resonant configurations without invoking migration.
richment took place within the inner disk, presumably by a mechanism such as dust/pebble drift. They further assume that planetesimals were uniformly distributed across the inner disk, usually as a simple, power-law profile. Such initial conditions – young disks with a broad close-in planetesimal distribution but little gas – are hard to reconcile with current thinking. If planetesimals formed quickly then the gas density should still be high and migration should be very fast. If planetesimals formed more slowly, presumably from pebbles that drifted inward to supplement the inner disk’s solid reservoir, then they are unlikely to have a smooth radial distribution (e.g. [Chatterjee and Tan 2014; Drackowska et al. 2016]). Thus, while these studies have provided interesting insights into various aspects of the accretion process, it seems unlikely that their starting conditions reflect reality. The mechanism by which inner disks are enriched in solids seems to be central to understanding the origins of super-Earths.

The migration model (Terquem and Papaloizou 2007; Ogihara and Ida 2009; McNeil and Nelson 2010; Ida and Lin 2010; Cossou et al. 2014; Izidoro et al. 2017, 2018) proposes that large planetary embryos form throughout the disk and migrate inward. Inward migration is counteracted by the positive surface density gradient at the disk’s inner edge (Massey et al. 2006). Embryos thus migrate inward and pile up into long resonant chains anchored at the disk’s inner edge. Collisions are common during this phase, leading to a breaking of resonance followed by continued migration and rapid re-formation of the resonant chain in a new configuration. After the gas disk dissipates many resonant chains become unstable, leading to a phase of giant impacts between growing planets that is not unlike the final phase of in-situ accretion. If 90-95% of resonant chains become unstable, the resulting systems provide a quantitative match to the observed period ratio and multiplicity distributions (Izidoro et al. 2017, 2018). In this model, the Kepler dichotomy is an observational artifact: the broad distribution of mutual inclinations in multiple super-Earth systems naturally produces a peak in systems with a single transiting planet. The resonant chains that remain stable are associated with observed multi-resonant systems such as TRAPPIST-1 (Gillon et al. 2017; Luger et al. 2017) and Kepler-223 (Mills et al. 2016).

Ormel et al. (2017) proposed a hybrid scenario in which planetesimals form first at the snow line, undergo pebble accretion and then migrate inward. This idea is consistent with the migration model and also connects with dust growth and drift models, which find that planetesimals tend to form fastest just past the snowline (Armitage et al. 2016; Drackowska and Alibert 2017; Schoonenberg and Ormel 2017).

The late stages in super-Earth evolution should be the same for both the drift and migration models. Of course, the two models invoke different formation modes and feeding zones for the planets. However, once there is a population of planets massive enough to migrate, the subsequent evolution is independent of how the planets formed. Whatever the processes responsible for creating a population of such planets, they migrate. And the outcome of this migration is a well-studied problem. As long as there is a disk inner edge, a system of migrating planets invariably organizes itself into a chain of mean motion resonances (e.g. Snellgrove et al. 2001; Lee and Peale 2002; Papaloizou and Terquem 2006; Creswell et al. 2007). Magnetohydrodynamic simulations of disk accretion onto young stars show that for most plausible parameters disks should indeed have inner edges (Romanova et al. 2003; Bouvier et al. 2007). Although the strength of the (positive) corotation torque depends on the local disk properties (Masset and Casoli 2010; Paardekooper et al. 2011), migration is likely to be directed inward during the late phases of disk evolution (Bitsch et al. 2014, 2015). Another factor is that a system of many super-Earths in resonance acts to excite their mutual eccentricities, decreasing the strength of the corotation torque (Bitsch and Kley 2010; Endyke and Nelson 2014) and potentially leading to inward migration of the cohort (Cossou et al. 2013).

We therefore argue that super-Earths must migrate regardless of how they formed. Late in the disk lifetime all formation scenarios converge on the evolution envisioned in the “breaking the chains” model of Izidoro et al. (2017, 2018). Migration should produce resonant chains with the innermost planets anchored at the inner edge of the disk. When the gas disk dissipates along with its stabilizing influence, the vast majority of resonant chains become unstable. While many details remain to be resolved, this evolution matches the key observed super-Earth constraints.

Raymond et al. (2008b) proposed that super-Earth formation models could be differentiated with two observables: the planets’ compositions (via their densities) and the systems’ orbital architectures. We have just argued that the late phases of the two viable models should converge to the same dynamical pathway, i.e., the “breaking the chains” model. We therefore do not expect the orbital architectures of super-Earth systems to provide a means of differentiation between models.

The compositions of super-Earths should in principle be different for the drift and migration models. In the drift model, all planet-building takes place close-in. Super-Earths should therefore be purely rocky, because tempera-

---

2The disk inner edge thus provides a built-in stopping mechanism for inward migration that plays a central role in the migration model. One may then wonder whether the in-situ formation model could also have a mechanism for avoiding migration. Indeed, in-situ growth within a region of slow or stopped migration – such as near the disk’s inner edge – would naturally reduce the importance of migration. However, there are two important caveats. First, if planets grew in-situ in slow-migration regions, this would not remove the importance of planet-disk interactions, which also affect the growing planets’ eccentricities and inclinations (e.g. Papaloizou and larwood 2003; Tanaka and Ward 2004). Second, regions of reduced migration are thought to be narrow (e.g. Hasegawa and Pudritz 2013; Bitsch et al. 2015a; Baillie et al. 2015), such that even if some super-Earths did indeed grow more or less in-situ, the bulk of growing super-Earths would still have experienced migration.
tures so close-in are so hot that the local building blocks should not contain any volatiles such as water. In contrast, a simple view of the migration model would predict ice-rich planets. Indeed, some models of dust growth and drift first produce planetesimals at or beyond the snow line, the radial distance beyond which water vapor can condense as ice (Armitage et al. 2016; Drążkowska and Alibert 2017). Immediately past the snow line, embryos are thought to grow faster and larger than closer-in because of pebbles that are most likely to migrate and thus, super-Earths should themselves be ice-rich in the migration model.

There are three problems with the idea of the migration model producing exclusively ice-rich super-Earths (see Raymond et al. 2018b). First, in some cases rocky embryos may grow large enough to migrate (see Jacobson et al. 2018). Indeed, in some models planetesimals form first in the terrestrial planet-forming region (Drążkowska et al. 2016; Surville et al. 2016), which would give rocky embryos a head-start in their growth. Second, migrating embryos must pass through the building blocks of terrestrial planets on their way to becoming super-Earths. Their migration can act to pile up rocky material in inner resonances with the migrating embryos (Izidoro et al. 2014b) and catalyze the rapid formation of rocky super-Earths, which preferentially end up interior to the migrating ice-rich embryos. Indeed, simulations of the migration model show that the innermost super-Earths are often built entirely from inner planetary system material and should be purely rocky (Raymond et al. 2018b). Third, it is not clear that embryos that migrate inward from beyond the snow line must be ice-rich. It is the fastest-forming embryos that are most likely to migrate, for simple timescale reasons. Yet rapid growth implies massive volatile loss.

Thermal evolution models find that any planetesimals that form within 1 Myr are completely depleted by strong $^{26}$Al heating (Grimm and McSween 1993; Monteux et al. 2018). This short-lived radionuclide (half-life of $\sim 700,000$ years) is thought to have been injected into the Sun’s planet-forming disk from a nearby massive star (e.g. Hester et al. 2004; Gounelle and Meibom 2008; Gaidos et al. 2009; Ouellette et al. 2010) and to have played a central role in the thermal evolution of the fastest-forming planetesimals. In addition, giant collisions between ice-rich bodies preferentially strip outer icy mantles and leave behind rocky/iron cores (Marcus et al. 2010). Likewise, later giant impacts can lead to substantial water loss for ocean planets (Genda and Abe 2005).

It is the very closest-in planets that are easiest to characterize. Within the so-called “photo-evaporation valley”, the atmospheres of any super-Earth-sized planets are thought to be rapidly evaporated by UV irradiation from the central star (Lammer et al. 2003; Baraffe et al. 2004; Hubbard et al. 2007). Planets in this region should necessarily have lost their gaseous envelopes and thus be “naked” (Lopez and Fortney 2013; Owen and Wu 2013; Sanchis-Ojeda et al. 2014), while planets just past the valley may have larger radii due to atmospheric heating (Carrera et al. 2018). From their observed sizes and masses, these very close-in planets appear to be rocky in nature, not icy (Owen and Wu 2017; Lopez and Mordasini 2018). Some studies have suggested that the observed dip in planet occurrence between $1.5 - 2 R_\oplus$ (Fulton et al. 2017; Teske et al. 2018; Fulton and Petigura 2018) may provide evidence that super-Earths are rocky (Jin and Mordasini 2018; Van Eylen et al. 2018; Gupta and Schlichting 2018) suggest that in this context, ‘rocky’ planets should have less than 20% water by mass (which is comparable to Europa’s water content by mass). Given that both the migration and drift models are consistent with rocky super-Earths (Raymond et al. 2018b), observations cannot yet differentiate between the drift and migration models. Interpreting the mean densities of planets beyond the distance for atmospheric photo-evaporation is challenging because of the degeneracies that arise once gas is included as a third potential constituent (along with rock/iron and water; see Selsis et al. 2007; Adams et al. 2008). The main difference between the migration and drift models is that the migration model predicts that some super-Earths should be ice-rich, in particular those planets that formed relatively late or in disks with little $^{26}$Al.

Embryos embedded in the disk should also accrete gas (e.g. Ikoma et al. 2000; Rogers et al. 2011). Indeed, many super-Earths are observed to have very low densities (e.g. Marcy et al. 2014) and this has been interpreted as planets larger than $1.5 R_\oplus$ having gaseous envelopes that are typically 0.1-10% of their total mass (Lopez and Fortney 2014; Weiss and Marcy 2014; Rogers 2015; Wolfgang et al. 2016; Chen and Kipping 2017). Recent work has shown that gas accretion is far more complex than previously assumed. For few Earth-mass planets, currents of gas often pass within a small fraction of the planet’s Hill sphere before exiting (Fung et al. 2015; Lambrechts and Lega 2017), casting doubts on the simple picture that gas within a planet’s Hill sphere must simply cool sufficiently to approach the planet’s surface. Nonetheless, growing super-Earths must accrete gas during the disk phase (Lee et al. 2014; Inamdar and Schlichting 2015) but should rarely reach the 50% gas-by-mass threshold for runaway gas accretion (Pollack et al. 1996) because the occurrence rate of hot Jupiters is more than an order of magnitude smaller than that of super-Earths (Howard et al. 2010; Mayor et al. 2011). Super-Earths may be in a constant state of gas accretion (Ikoma et al. 2000; Lambrechts and Lega 2017) moderated by loss processes related to collisions (Schlichting et al. 2015; Inamdar and Schlichting 2016) as well as the dissipation of the disk itself (Ikoma and Hori 2012; Ginzburg et al. 2016). Models invoking atmospheric loss from a few large impacts — such as those characteristic of late instabilities in the breaking the chains model — can broadly match the ob-
Fig. 3.— How orbital migration and dynamical instabilities can explain the properties of exoplanet populations. **Left:** Evolution of the “breaking the chains” migration model for the origin of super-Earths (Izidoro et al. 2017, 2018). Embryos within the snow line are entirely rocky and much smaller than those that form past the snow line, which also incorporate ice. Presumably ice-rich embryos migrate inward through the rocky material, catalyzing the growth of purely rocky planets interior to the ice-rich ones (Raymond et al. 2018b). Planets migrate into long chains of mean motion resonances, with the innermost planet at the inner edge of the disk. The vast majority (90-95%) of resonant chains become unstable when the gas disk dissipates. The resulting planets match the distributions of known super-Earths (Izidoro et al. 2017, 2018). Given various loss process (e.g. Grimm and McSween 1993; Genda and Abe 2005; Marcus et al. 2010; Monteux et al. 2018) the water/ice contents of these planets may be drastically overestimated. **Right:** Evolution of the planet-planet scattering model for the origin of giant exoplanets (e.g. Adams and Laughlin 2003; Chatterjee et al. 2008; Raymond et al. 2010). Several embryos grow quickly enough to accrete gas and grow into gas giants. They subsequently migrate into a resonant chain without drastically affecting the orbits of nearby growing rocky planets (or outer planetesimal disks). After the disk dissipates, the vast majority (75-90%) of giant planets systems become unstable. The resulting systems match the correlated mass-eccentricity distribution of known giant exoplanets (e.g. Ford and Rasio 2008; Wright et al. 2009).
served distribution of gaseous envelope masses ([Inamdar and Schlichting 2016]), although the initial, pre-impact atmospheric masses remain uncertain.

It is difficult to understand why close-in super-Earths exist in so many systems but not in all systems. Models of super-Earth formation struggle not to form them whereas microlensing observations show that similar-mass planets are extremely abundant past the snow line ([Beauregard et al. 2016; Gould et al. 2010; Suzuki et al. 2016]). A simple explanation is to invoke a timing constraint: if large embryos form too slowly then they would not have time to migrate all the way to the inner edge of the disk. However, most studies find that migration is fast (although new simulations by McNally et al. 2018, find that migration may be slower) so this requires a fine-tuned delay such that most super-Earths grow large just before the gas disk dissipates. And late accretion of smaller embryos beyond a few AU after the dissipation of the disk would be quite inefficient (e.g., Levison and Stewart 2001; Thommes et al. 2003).

Lower disk masses also cannot explain why many stars do not host close-in super-Earths. There is an observed super-linear correlation between the disk mass and stellar mass ([Passucchi et al. 2016], albeit with large scatter at a given stellar mass (e.g., Scholz et al. 2006; Williams and Cieza 2011). This means that M dwarf stars – with masses between roughly 8% and 60% of the Sun’s – have on average significantly lower-mass disks than Sun-like stars. However, the occurrence rate of super-Earths around M dwarfs is at least as high as around Sun-like stars (Howard et al. 2012; Fressin et al. 2013; Dressing and Charbonneau 2015) and systems of close-in low-mass planets contain on average a higher total mass around M dwarfs (Mulders et al. 2015b). If lower-mass disks could not produce super-Earths, M dwarfs should naively have lower average occurrence rates.

Wide-orbit giant planets may potentially explain why some systems do not have close-in super-Earths. Once a planet accretes enough gas it carves an annular gap in the disk, slows its migration and becomes a true gas giants (Lin and Papaloizou 1986; Bryden et al. 1999; Crida et al. 2006). The inward migration of more distant large embryos are then blocked by the gas giant (Izidoro et al. 2015b). In this context, Uranus and Neptune (and perhaps Saturn’s core) may represent failed super-Earths, embryos whose migration was blocked by the young Jupiter (Izidoro et al. 2015b). This predicts an anti-correlation between systems with many close-in super-Earths and those with wide-orbit gas giants (Izidoro et al. 2015b). However, the occurrence rate of wide-orbit gas giants is \( \sim 10\% \) for Sun-like stars (Cumming et al. 2008; Mayor et al. 2011; Wittenmyer et al. 2016; Rowan et al. 2016) and is likely far lower for M dwarfs (Johnson et al. 2007; Lovis and Mayor 2007; Dressing and Charbonneau 2015). For FGK stars this is a factor of 3-5 lower than the occurrence rate of super-Earths, and the problem is even worse for M dwarfs. Thus, the Jupiter migration barrier does not appear capable of explaining why most systems do not have close-in super-Earths.

Wide-orbit planets with masses comparable to the ice giants’ \((10-20 M_\oplus)\) may help solve this problem by stunning the growth of planetary embryos. The largest planetesimals are thought to represent the seeds of planetary embryos and to grow by accreting pebbles that drift inward through the disk (e.g., Johansen and Lambrechts 2017, see §2). Once an embryo reaches the pebble isolation mass it creates a pressure bump in the gas disk exterior to its orbit that acts to trap drifting pebbles and shut off the pebble flux ([Morbidelli and Nesvorny 2012; Lambrechts et al. 2014; Bitsch et al. 2018b]). This not only starves the embryo but all other embryos interior to its orbit, which may continue to accrete planetesimals (but not pebbles). For typical disk parameters the pebble isolation mass is on the order of \(20 M_\oplus\) (Lambrechts et al. 2014; Bitsch et al. 2018b). A fast-growing wide-orbit planet with a mass similar to Neptune’s \((17 M_\oplus)\) may starve the inner disk and prevent closer-in embryos from reaching large enough masses to undergo rapid migration and following the “breaking the chains” pattern discussed above. This mechanism is especially promising given that the abundance of wide-orbit ice giant-mass planets inferred from microlensing is on the same order as the occurrence rate of close-in super-Earths (e.g., Gould et al. 2010; Petigura et al. 2013; Clanton and Gaudi 2016; Winn and Fabrycky 2015). However, given that pebble-blocking outer Neptunes must form quickly (to starve inner embryos), it is not clear how such planets could avoid migrating inward and becoming super-Earths themselves.

It is of course possible that migration may not always follow the pattern that we have laid out. In one type of disk model that invokes winds as an angular transport mechanism, the surface density in the inner 1-2 AU of the disk increases steeply with radius ([Suzuki et al. 2016]). In such a disk, type I migration is significantly slowed and may even be quenched ([Oshihara et al. 2015b]), such that co-migrating super-Earths may not always form resonant chains ([Oshihara et al. 2018]). Magnetic stresses in the mid-plane of certain disk models may generate positive torques that drive outward planet migration but only in specific situations, for example if the star’s spin vector is aligned with the magnetic field ([McNally et al. 2017; 2018]). Such effects could in principle prevent inward migration in a subset of disks.

To conclude this subsection, we reiterate that the breaking the chains scenario ([Izidoro et al. 2013, 2018]) provides a match to the observed super-Earth systems. The late evolution of that model should hold whether the bulk of super-Earths’ mass comes from pebbles that drifted inward (the drift model; Chatterjee and Tan 2014) or from cores that formed past the snow line (the migration model; Terquem and Papaloizou 2007). What remains unexplained is why so many – but not all – stars have close-in super-Earths.

### 3.2 Systems with giant exoplanets
Giant exoplanets are found around ~10% of Sun-like stars (see discussion in §1.2: [Cumming et al. 2008; Mayor et al. 2011; Foreman-Mackey et al. 2016]). Most giant exoplanets have orbital radii larger than 0.5-1 AU (Butler et al. 2006; Udry and Santos 2007) and only ~1% of stars have hot Jupiters (Howard et al. 2010; Wright et al. 2012). Giant planet masses follow a roughly \( dN/dM \propto M^{-1.1} \) distribution of minimum masses (Butler et al. 2006). The median eccentricity of this population is \( \approx 0.25 \), roughly five times larger than Saturn and Jupiter’s long-term average eccentricities (Quinn et al. 1991). More massive giant planets have statistically higher eccentricities than lower-mass giant planets (typically, the division between high-mass and low-mass giant planets is at roughly Jupiter’s mass; Jones et al. 2006; Ribas and Miralda-Escude 2007; Ford and Rasio 2008; Wright et al. 2009).

As for super-Earths, the population of giant exoplanets is thought to have been sculpted in large part by migration and instability.

A subset of giant exoplanets has been found to be in mean-motion resonance. Strong resonances with small-amplitude libration of resonant angles are thought to be a clear signature of migration (e.g., Papaloizou and Terquem 2006; Kley and Nelson 2012); weaker resonances can be produced by instabilities – see [Raymond et al. 2008a]). Notable examples of giant exoplanets in resonance include the GJ 876 system in which three planets are locked in 4:2:1 Laplace resonance (Rivera et al. 2010) and the HR 8799 system (Marois et al. 2008, 2010), for which stability considerations indicate that the outer three or perhaps all four super-Jupiter-mass planets may be in resonance (Reidemeister et al. 2009; Fabrycky and Murray-Clay 2010; Goździewski and Migaszewski 2014; Görgen et al. 2016). These systems are thought to represent signature cases of migration, and a recent analysis found that a significant fraction may be in resonance (25% of the 60 giant exoplanet systems studied in Boisvert et al. 2018).

Giant planets are thought to form on circular orbits. However, a large fraction of giant exoplanets are found to have significant orbital eccentricities, which are thought to be an indicator of dynamical instability (see, e.g. Ford and Rasio 2008). The eccentricity distribution of giant exoplanets can be reproduced by the planet-planet scattering model, which proposes that the observed planets are the survivors of system-wide instabilities (Rasio and Ford 1996; Weidenschilling and Marzari 1996; Lin and Ida 1997). Giant planets are assumed to form in systems of two or more planets. After the gaseous disk dissipates the planets’ orbits become dynamically unstable, leading to a phase of close gravitational scattering events. Scattering events involve orbital energy and angular momentum exchange between the planets and tend to increase their eccentricities and inclinations. During this phase of dramatic orbital excitation, nearby small bodies such as the building blocks of terrestrial planets are generally driven to such high eccentricities that they collide with the host star (Veras and Armitage 2005, 2006; Raymond et al. 2011, 2012; Matsumura et al. 2013). More distant planetesimals are preferentially ejected (Raymond et al. 2011, 2012, 2018; Marzari 2014). Giant planet instabilities typically conclude with the ejection of one or more planets into interstellar space (Veras and Raymond 2012). The surviving planets have eccentric orbits (Ford et al. 2003; Adams and Laughlin 2003; Chatterjee et al. 2008; Ford and Rasio 2008; Raymond et al. 2010). The observed eccentricity distribution can be matched if at least 75% – and probably closer to 90% – of all giant exoplanet systems represent the survivors of instabilities (Jurić and Tremaine 2008; Raymond et al. 2010). Scattering can also match the observed orbital spacing of giant planet systems, in particular with regards to their proximity to the analytically-derived boundary for orbital stability (also called ‘Hill stability’; Raymond et al. 2009a), as well as the secular structure of observed systems (Ford et al. 2003; Timpe et al. 2013).

Gravitational scattering produces an energy equipartition among planets in the same system such that the lowest-mass planets have the highest eccentricities. This is in disagreement with observations, which show that higher-mass planets have higher eccentricities than lower-mass, with a statistically significant difference (Ribas and Miralda-Escude 2007; Ford and Rasio 2008; Wright et al. 2009). This discrepancy is resolved if systems with very massive giant planets \( (M \gtrsim M_{\text{Jup}}) \) systematically form multiple, very massive giant planets with near-equal masses (Raymond et al. 2010; Ida et al. 2013). The eccentricities of surviving planets are highest in systems with the most massive, equal-mass planets (Ford et al. 2001; Raymond et al. 2010).

All that is needed to trigger instability is the formation and migration of 2-3 gas giants\(^1\). The timescale of instability is a function of the planets’ initial separations (Chambers et al. 1996; Marzari and Weidenschilling 2002), so most studies simply started planets in unstable configurations to determine the outcome of the instability. A more self-consistent approach invokes a prior phase of orbital migration. While migration is often thought of as a dynamically calm process, several studies have shown that migration of multiple giant planets often generates instabilities after, or even during, the gaseous disk phase (Moeckel et al. 2008; Matsumura et al. 2010; Marzari et al. 2010; Moeckel and Armitage 2012; Lega et al. 2013). The planets that emerge from migration-triggered instability provide a match to the observed giant exoplanets (Adams and Laughlin 2003; Moorhead and Adams 2005).

Hot Jupiters present an interesting melding of migration and instability (see Dawson and Johnson 2018, for a review). In recent years it has been debated whether hot Jupiters migrated into their current locations (Lin et al. 1996; Armitage 2007) or were scattered to such high eccentricities (and such small pericenter distances) that tidal dis-

\(\text{\footnote{Wide binary stars can also trigger instabilities, as torques from passing stars and the galactic tidal field occasionally shrink their pericenters to approach the planetary region (Kaib et al. 2013).}}\)
sipation within the star (also called ‘tidal friction’) shrank their orbits \cite{Nagasawa11,Beaug2012}. The Kozai effect – in which a perturbing planet or star on a highly-inclined orbit induces large-scale, anti-correlated oscillations in a planet’s eccentricity and inclination – may play a role in generating the high eccentricities needed to produce hot Jupiters by tidal friction \cite{Fabrycky07,Naoz2011} but only in situations in which such large mutual inclinations arise (e.g., after an instability). Of course, it is possible that hot Jupiters underwent rapid gas accretion close-in and thus represent the rare hot super-Earths that grew fast enough to trigger runaway gas accretion \cite{Bodenheimer00,Boley16,Barbin2016}. Yet even if hot Jupiters accreted their gas close-in, the arguments presented in §3.1 still indicate that their building blocks must have either migrated or drifted inward.

Another constraint comes from observations of the Rossiter-McLaughlin effect, which measures the projected stellar obliquity in systems with transiting planets \cite{Winn05,Gaudi07}. This translates to a projection of the planet’s orbital inclination with respect to the stellar equator. For a significant fraction of measured hot Jupiters, the orbital plane is measured to be strongly misaligned with the host stars’ equators. More massive stars, for which the timescale for tidal dissipation are much longer \cite{Zahn77}, are far more likely to host misaligned hot Jupiters \cite{Winn10,Triaud10,Albrecht2012}. This may be explained by planets being scattered onto very eccentric and inclined orbits before their orbits are shrunk by tidal friction \cite{Fabrycky07,Nagasawa11,Naoz2011,Beaug2012,Lai12}.

Migration predicts that hot Jupiters should remain aligned with their birth disks. However, disks themselves can be torqued into configurations that are misaligned with respect to the stellar equator \cite{Lai11,Batygin12}. If planet-forming disks are themselves misaligned then both the migration and close-in growth models can in principle explain hot Jupiters’ misaligned orbits. In the Kepler-56 system two massive planets share an orbital plane that is misaligned with the stellar equator \cite{Huber13}. While this is suggestive of the planets having formed in a tilted disk, other dynamical mechanisms can plausibly explain such tilting \cite{Innanen97,Mardling10,Kaib11,Boue14,Gratia17}. In addition, there is as yet no sign of debris disks – dust disks observed around older stars whose gas disks have already dissipated – that are misaligned with the equators of their host stars \cite{Greaves14}.

While the general picture of migration and instability appears to match the broad characteristics of giant exoplanets, questions remain. While slower than for low-mass planets, the timescale for type 2 migration – which is thought to be controlled in large part by the disk’s viscosity \cite{Lin86,Ward97} – is still in many cases faster than the disk lifetime. Why, then are there so few gas giants interior to 0.5-1 AU? Photo-evaporation of the disk produces inner cavities of roughly that size \cite{Alexander14,Ercolanop17} \cite{Alexander12}; but see \cite{Wise18} and discussion in \cite{Morbidelli18}. The cavity is only generated late in the disk’s lifetime, so if it is to explain the deficit of gas giants within 0.5-1 AU this would require very slow migration and thus very low-viscosity disks \cite{Alexander12}.

To conclude, the general evolution of gas giant systems thus appears to follow a similar pattern as super-Earths’ breaking the chains evolution. Gas giants form in cohorts and migrate into resonant configurations \cite{Kaib11}. After the disk dissipates (or sometimes before) the vast majority of systems become unstable and undergo a violent phase of planet-planet scattering that generates the observed gas giant eccentricities \cite{Juric08,Raymond10} and also disrupts the growth of any smaller planets in the systems \cite{Raymond11}. The origins of hot Jupiters remain debated, but viable formation models invoke a combination of migration and instability \cite{Dawson18}.

4. MODELS FOR SOLAR SYSTEM FORMATION

We now turn our attention to the origin of the Solar System. Given the astrobiological context of this chapter, we emphasize the origin of the inner Solar System. However, it is important to keep in mind that dynamical perturbations from Jupiter during its growth, migration and early evolution played an important part in shaping the terrestrial planets.

In this section we first lay out the observational constraints (§4.1). Next, in §4.2 we present a rough timeline of events, including a discussion of the late instability in the giant planets’ orbits (the so-called ‘Nice model’). In §4.3 we describe the so-called classical model of terrestrial planet formation and explain its shortcomings. In §4.4 we present and contrast three models for the early evolution of the inner Solar System. In §4.5 we explore a feature of the Solar System that remains hard to explain with all current models: the mass deficit in the very close-in Solar System.

4.1 Solar System constraints

A successful formation model must match the Solar System’s broad characteristics (see §7.3 for a philosophical discussion). We now lay out the central constraints to be matched, in rough order of importance. We start with constraints related to the inner Solar System (roughly from most- to least-stringent) and conclude with constraints related to the outer Solar System (Jupiter and beyond).

The masses and orbits of the terrestrial planets. The terrestrial planets follow an odd pattern, with two large central planets (Venus and Earth) flanked by much smaller ones
(Mercury and Mars). Roughly 90% of all the rocky material in the Solar System is thus concentrated in a ring that is only 0.3 AU in width (encompassing the orbits of Venus and Earth). In addition, the terrestrial planets’ orbits are remarkably close to circular. These constraints are often quantified using statistics on the radial mass concentration (Chambers 2001) and degree of orbital excitation (often using the angular momentum deficit, defined as the fractional deficit in angular momentum of a system of planets relative to an identical system on perfectly circular, coplanar orbits; Laskar 1997; Chambers 2001). In addition, the Earth/Mars and Venus/Mercury mass ratios offer simple, surprisingly strong constraints.

The mass, orbital and compositional structure of the asteroid belt. The main belt (between roughly 2 and 3.2 AU) covers a surface area more than three times larger than that of the terrestrial planet region, yet the entire belt contains only $\sim 4.5 \times 10^{-4}$ M$_{\oplus}$ (Krasinsky et al. 2002; Kuchynka and Folkner 2013; DeMeo and Carr 2013). Yet the asteroids’ orbits are excited, with eccentricities that span from zero to 0.3 and inclinations that extend above 20°. The belt contains a diversity of spectroscopically-distinct types of objects (Bus and Binzel 2002). The belt is also radically segregated: the inner main belt (interior to ~ 2.7 AU) is dominated by S-types whereas the main belt beyond 2.7 AU is dominated by C-types (Gradie and Tedesco 1982; DeMeo and Carr 2013, 2014). S-types are identified with ordinary chondrites, which are relatively dry (with water contents below 0.1% by mass), whereas C-types are associated with carbonaceous chondrites, which typically contain ~ 10% water by mass (Robert et al. 1977; Kerridge 1985; Alexander et al. 2018).

The cosmochemically-constrained growth histories of Earth and Mars. Isotopic analyses of different types of Earth rocks, lunar- and Martian meteorites constrain the growth histories of Earth and Mars. Isotopic systems such as Hf-W with half-lives comparable to the planets’ formation timescales are particularly useful (the half-life of radioactive $^{188}$Hf is 9 Myr; see Alexander et al. 2001). These studies indicate that Earth’s core formation did not finish until at least $\sim 40 - 100$ million years after the start of planet formation (Touboul et al. 2007; Kleine et al. 2009). The final episode of core formation on Earth is generally assumed to have been the Moon-forming impact (Benz et al. 1986; Canup and Asphaug 2001). Mars’ growth is directly constrained to have been far faster than Earth’s (Nimmo and Kleine 2007). Indeed, Mars’ accretion was complete within 5-10 Myr (Dauphas and Pourmand 2011).

The abundance and isotopic signature of water on Earth. Despite being mostly dry and rocky, Earth still contains a small fraction of water by mass, and is thought to be essential for life. The exact amount of water on Earth remains only modestly-well constrained. An “ocean” of water is defined as the total amount of water on Earth’s surface, roughly $1.5 \times 10^{24}$ grams (or $\sim 0.025\%$ of an Earth mass). The mantle is thought to contain between a few tenths of an ocean (Hirschmann 2008; Panero and Caracas 2017), and 5-10 oceans (Léveillé et al. 1998; Marty 2012; Hallidie 2013). The core is generally thought to be very dry (Badro et al. 2014) but Nomura et al. (2014) inferred a very large reservoir of water exceeding 50 oceans. Assuming a total water budget of four oceans, Earth’s bulk water content is thus 0.1% by mass.

The isotopic signature of Earth’s water – the D/H and $^{15}$N/$^{14}$N ratios – is a key discriminant of different models of water delivery (see §6). Earth’s water is a good match to carbonaceous chondrite meteorites, specifically the CM subgroup (Marty and Yokochi 2003; Alexander et al. 2012). Earth’s water is isotopically distinct from nebular and cometary sources (see data compiled in Morbidelli et al. 2001; note that there are two comets observed to have Earth-like D/H ratios but both have non-Earth-like $^{15}$N/$^{14}$N ratios; see discussion in §6).

The late veneer on Earth, Mars and the Moon. Highly-siderophile elements are those that are thought to have a chemical affinity for iron rather than silicates. Most of these elements are thus thought to be sequestered in a planet’s core during core-mantle segregation. All of the highly-siderophile elements in Earth’s crust must therefore have been delivered by impacts after the Moon-forming impact (Kimura et al. 1974). From the abundance of highly-siderophile elements, and assuming the impactors to be chondritic in compositions, it has been inferred that the last ~ 0.5% of Earth’s accretion took place after the Moon-forming impact (Day et al. 2007; Walker 2009; Morbidelli and Wood 2015). Meteorite constraints indicate that Mars accreted ~ 9 times less material than Earth during the late veneer, and that the Moon accreted 200-1200 times less than Earth (Day et al. 2007; Walker 2009).

The orbits and masses of the giant planets. The giant planets are radially segregated by mass, with the most massive planets closest-in. As discussed in §1.1, Jupiter’s orbit is wider than most known giant exoplanets’ and it is only barely detectable by long-duration radial velocity surveys. Jupiter and Saturn each have low-eccentricity but non-circular orbits, each with Myr-averaged radial velocity surveys. Uranus’ average eccentricity is comparable to the gas giants’ but Neptune’s is only ~ 0.01. There are no mean motion resonances among the giant planets. The Jupiter/Saturn, Saturn/Uranus and Uranus/Neptune period ratios are 2.48, 2.85, and 1.95, respectively.

The total mass and orbital structure of the outer Solar System’s small body populations. The Kuiper belt extends outward from just Neptune’s orbit. It has a complex orbital structure that includes a population of objects such as Pluto that are locked in mean motion resonances with Neptune. The Kuiper belt has a broad eccentricity and inclination distribution and includes a population of very dynamically cold objects from 42-45 AU often called the cold classical belt. The total mass in the Kuiper belt has been estimated at a few to ten percent of an Earth
mass (Gladman et al. 2001). The scattered disk is a subset of Kuiper belt objects whose orbits cross those of the giant planets. The Oort cloud is the source of long-period (isotropic) comets and extends from roughly 1,000 AU out to the Sun’s ionization radius (currently at ∼ 200, 000 AU; see Tremaine 1993 for details).

4.2 A rough timeline of events

Theory and observations can combine to provide a rough timeline of the events that must have taken place in the Solar System. Time zero is generally assumed to be the time of formation of CAIs (Calcium and Aluminum-rich Inclusions), roughly mm-sized components of primitive Solar System. Time zero is generally assumed to be the time of formation of CAIs (Calcium and Aluminum-rich Inclusions), roughly mm-sized components of primitive (chondritic) meteorites that are well-dated to be 4.568 Gyr old (e.g., Bouvier and Wadhwa 2010).

• Within 100,000 years planetesimal formation was underway. CAIs and mm-scale chondrules had started to form (e.g., Connelly et al. 2008, Nyquist et al. 2009) and coalesce into larger objects (e.g., Dauphas and Chauveteau 2011, Johansen et al. 2015).

• Within 1 million years large embryos had formed. Ages of iron meteorites indicate that embryos had formed in the inner Solar System (e.g., Halliday and Kleine 2006, Krot et al. 2014, Schiller et al. 2015). Meanwhile, the segregation of the parent bodies of carbonaceous and non-carbonaceous meteorites indicates that at least one ∼ 10 M⊕ embryo – presumably Jupiter’s core (Krot et al. 2017) – had formed in the giant planet region. From this point onward, this core blocked the inward drift of pebbles and thus starved the inner Solar System (e.g., Bitsch et al. 2018b).

• Within a few million years the gaseous planet-forming disk had dissipated. Evidence for the timescale of disk dissipation comes from two sources. First, observations of hot dust – thought to trace the gas – around stars in young clusters with different ages indicate a typical dissipation timescale of 2-5 Myr (Haisch et al. 2001, Hillenbrand 2008, Pascucci et al. 2009, Mamajek 2009). Second, given that all chondrule formation models require the presence of the gas disk, the latest-forming chondrules provide a lower limit on the gas disk lifetime of 4-5 Myr (the CB chondrites; see Kita et al. 2005, Krot et al. 2005, Johnson et al. 2016). The existence of a hot Jupiter around a 2 million-year-old T Tauri star (Donati et al. 2016) demonstrates that giant planet formation and migration can happen on an even shorter timescale. The gas- and ice giant planets were fully-formed and likely in a compact resonant resonant chain (Morbidelli et al. 2007, Iciderro et al. 2015a). By this time Mars was close to fully-formed (Nimmo and Kleine 2007, Dauphas and Pourmand 2011) but Earth (and presumably Venus) were still actively accreting via planetesimal and embryo impacts.

• While the Sun was still in its birth cluster it underwent a relatively close encounter with another star. Such an encounter has been invoked to explain the orbits of the Sednoids (Morbidelli and Levison 2004, Kenyon and Bromley 2004, Jilková et al. 2015, Pfalzner et al. 2018) – named after Sedna, the first one discovered (Brown et al. 2004) – whose semimajor axes are greater than 250 AU and whose perihelia are detached from the planets’. The encounter may have either excited existing Solar System planetesimals onto Sedna-like orbits or captured the objects from the passing star. The encounter distance was likely at a few hundred to a thousand AU. While the exact properties of the Sun’s birth cluster remain a matter of debate (Adams 2010, Gounelle and Meynet 2012, Portegies Zwart 2018), such encounters are expected to be a common occurrence (Malmberg et al. 2011). It is interesting to note that the Sun must have left its parent cluster before the giant planet instability, since that is when the Oort cloud would have formed (Brasser et al. 2013) and it would be much more compact had it formed in a cluster environment, given the stronger tidal field (Tremaine 1993, Kaib and Quinn 2008).

• Roughly 50-100 Myr after CAIs, Earth suffered its final giant impact (Touboul et al. 2007, Kleine et al. 2009). This impact triggered Earth’s final core formation event and the formation of the Moon (Benz et al. 1986, Canup and Asphaug 2001). Only ∼ 0.5% of Earth’s mass was accreted after this point (Dav et al. 2007, Walker 2009, Morbidelli and Wood 2015).

• Within 500 Myr the outer Solar System went unstable. The instability – thought to have been generated by interactions between the giant planets and an outer planetesimal disk, essentially the primordial Kuiper belt – is commonly referred to as the Nice model. Starting from a compact resonant chain originally formed as a consequence of a previous phase of migration in the gaseous disk (Morbidelli et al. 2007), the giant planets underwent a series of close encounters. Interactions with the outer planetesimal disk caused the giant planets’ orbits to radially spread out and destabilized the planetesimal disk (Levison et al. 2011), which led to a phase of impacts throughout the Solar System that was originally proposed to correspond to the so-called late heavy bombardment (Tera et al. 1974, Gomes et al. 2005), the event often associ-

8The origin of chondrules is hotly debated, and some models suggest that they are the outcomes of collisions between planetesimals and planetary embryos rather than their building blocks (e.g., Asphaug et al. 2011, Johnson et al. 2015, Lichtenberg et al. 2018).
ated with many of the oldest craters on Mercury, the Moon and Mars. The instability can explain a number of features of the Solar System including the giant planets’ present-day orbits (Tsiganis et al. 2005; Nesvorný and Morbidelli 2012), the orbital distribution of Jupiter’s co-orbital asteroids (in particular their large inclinations) (Morbidelli et al. 2005; Nesvorný et al. 2013), and the characteristics of the giant planets’ irregular satellites (Nesvorný et al. 2007). Simulations that invoke that the young Solar System had 1-2 additional ice giants that were ejected during the instability have a much higher success rate in matching the present-day Solar System (Nesvorný and Morbidelli 2012; Batygin et al. 2012). The gas giants’ relatively low eccentricities constitute a dynamical constraint: Jupiter and Saturn never underwent a close mutual encounter, although they must have had encounters with one or more ice giants (Morbidelli et al. 2009).

While the existence of the instability remains in favor, the late timing has recently been challenged (Boehnke and Harrison 2016, Morbidelli et al. 2018, Michael et al. 2018, Nesvorný et al. 2018). An early giant planet instability is easier to understand from a dynamical perspective. The dispersal of the gaseous disk is the natural trigger for instabilities (e.g., Matsumura et al. 2010) and simulations had substantial difficulty in delaying the onset of instability (Gomes et al. 2005; Levison et al. 2011). Simulations of giant planets interacting with outer planetesimal disks indeed show that most instabilities happen early, although there is a tail of instabilities that extends to much longer timescales (Thommes et al. 2008b; Raymond et al. 2010). A key input in such models – the inner edge location and orbital distribution of planetesimals in the outer primordial disk – remains poorly-constrained.

The giant planet instability must have had a significant impact on the inner Solar System. The changing dynamical environment caused by changes in Jupiter and Saturn’s orbits caused secular resonances to sweep and/or jump across the inner Solar System, exciting anything in their path (Brasser et al. 2009; Agnor and Lin 2012). A late instability tends to excite and often to destabilize the orbits of the already-formed terrestrial planets (Brasser et al. 2013; Roig et al. 2016; Kaib and Chambers 2016). An early instability – triggered shortly after the gas disk’s dispersal and before the final assembly of the terrestrial planets – has the potential to resolve this problem, and constitutes the basis for one of the terrestrial planet formation models we will discuss in §4.4 (the Early Instability model of Clement et al. 2018a).

- For the past 4 billion years, the orbital architecture of the Solar System has remained roughly constant. Most impacts on the terrestrial planets come from asteroids that are disrupted and whose fragments end in unstable resonances (often after drifting due to the Yarkovsky effect; see Gladman et al. 1997; Bottke et al. 2006b; Granvik et al. 2017). The planets’ orbits undergo secular oscillations due to long-range gravitational perturbations (e.g. Quinn et al. 1991). The oscillations in Earth’s orbit and spin are called Milankovitch cycles and play a key role in its climate evolution (e.g., Berger et al. 2005). The inner Solar System is chaotic with a Lyapunov timescale of a few Myr (Laskar 1990; Batygin and Laughlin 2015), but it is unknown whether the outer Solar System’s evolution is chaotic or regular, as both types of solutions exist within the current error bars on the giant planets’ positions (Haves 2007). Regardless of whether the outer Solar System is chaotic or regular there is no chance of future instability. In contrast, the terrestrial planets have a ∼ 2% chance of becoming unstable before the Sun becomes a red giant in 4.3 The classical model of terrestrial planet formation

The so-called classical model of terrestrial planet formation was pioneered by a series of papers by George Wetherill spanning 2-3 decades (e.g., Wetherill 1978, 1985, 1996). It has succeeded in explaining a large number of features of the inner Solar System, and its shortcomings have served to point newer models in the right direction. The classical model remains to this day the basis of comparison with more recent models (e.g. Morbidelli et al. 2012; Raymond et al. 2014; Jacobson and Walsh 2015; Izidoro and Raymond 2018).

The central assumption in the classical model is that giant planet formation can be considered separately from terrestrial accretion. At face value this appears to be a reasonable assumption. Gas-dominated protoplanetary disks are observed to dissipate in a few Myr (Haisch et al. 2001; Hillenbrand 2008), setting an upper limit on the timescale of gas giant formation. In contrast, cosmochemical studies have demonstrated that Earth’s accretion lasted 50-100 Myr (Kleine et al. 2009). Simulations of the classical model start from a population of rocky building blocks (planetary embryos and planetesimals) and fully-formed gas giants, inherently assuming that there was no prior interaction between these different populations and that the gas had already been removed.

Figure 4 shows the evolution of a characteristic simulation of the classical model (from Raymond et al. 2006b). The simulations is gas-free and so its time zero effectively corresponds to the dissipation of the gaseous disk. The population of rocky embryos (initially ~Ceres- to Moonmass in this case) self-excites by mutual gravitational forcing from its inner regions outward, producing larger embryos with a characteristic spacing (Kokubo and Ida 1998).
that other embryos are growing concurrently, the largest accreting planetesimals as well as other embryos, and given Raymond et al. O'Brien et al. ties; the embryos’ eccentricities and inclinations are gener-

ational scattering. There is a long chaotic phase charac-
cited bodies transmit this disturbance through mutual grav-

There are some circumstances under which the classical model can produce small Mars analogs. For ex-
ample, if Jupiter and Saturn’s orbits were more excited \( \epsilon_{\text{Jup}} \approx \epsilon_{\text{Sat}} \approx 0.07 - 0.1 \) during terrestrial accre-
tion than they are today, then secular resonances would have been far stronger and could have acted to clear

material from the Mars zone without depleting Earth’s feeding zone (the EEJS, or ‘Extra Eccentric Jupiter and Saturn’ configuration from Raymond et al. 2009; Morishima et al. 2010; Kaib and Cowan 2015). However, the EEJS setup has its own Achilles heel: its initial conditions are not consistent with the evolution of Jupiter and Saturn in the gaseous disk. Simulations universally show that planet-disk interactions tend to drive the planets into resonance (in this case, specifically the 3:2 or 2:1 resonances Pierens et al. 2014). However, if Jupiter and Saturn were in a resonant configuration, the location of their secular resonances within the terrestrial disk would not help to produce a small Mars (e.g. Izidoro et al. 2016). A similar model invokes secular resonance sweeping during the dispersal of the gaseous disk to explain the depletion of the asteroid belt and Mars region (Nagasawa et al. 2005; Thommes et al. 2008; Bromley and Kenyon 2017). However, this model suffers from the same problem as the EEJS model: the gas giants’ orbits are not consistent with the evolution of the disk, and using appropriate (generally lower-eccentricity, resonant) orbits removes the desired depletion. An early giant planet instability may, however, produce a giant planet configuration similar to the EEJS configuration as we discuss in §4.4.

The small Mars problem is inherently coupled with the asteroid belt’s orbital excitation Izidoro et al. 2015c. While very low in total mass, the asteroids’ orbits are much more excited than the planets’, with a broad range of eccen-
crities and inclinations. The current amount of mass in the belt cannot account for its excitation because there is not enough mass for gravitational self-stirring to be ef-

Fig. 4.— A simulation of the classical model of terrestrial planet formation (adapted from Raymond et al. 2006b). The simula-
tion started from 1886 self-gravitating planetary embryos, re-
presented as dots with size proportional to its mass \( \frac{1}{3} \). Jupiter is fully-formed (large black dot) on a near-circular orbit at the start of the simulation. The color of each embryo represents its water content (see color bar at the bottom), with red objects being dry and the darkest blue containing 5% water by mass (see Raymond et al. 2004). This simulation produced quite good Earth and Venus analogs, a very poor Mars analog, and a plausible, albeit far too massive, asteroid belt. Note that time zero for this simulation corresponds to the dissipation of the gaseous disk, a few Myr after CAIs. A movie of this simulation can be viewed here: https://youtu.be/m7hNlg9Gxvo.

The outer parts of the rocky disk are ex-
cited by secular and resonant forcing from Jupiter, and ex-
cited bodies transmit this disturbance through mutual grav-
itational scattering. There is a long chaotic phase charac-
terized by excitation of planetesimals to high eccentricities; the embryos’ eccentricities and inclinations are gener-

ally kept lower by dynamical friction \( O’Brien et al. 2006; \) Raymond et al. 2006b). During this phase embryos grow by accreting planetesimals as well as other embryos, and given that other embryos are growing concurrently, the largest impacts tend to happen late (e.g. Agnor et al. 1999). By roughly 100 Myr after the start of the simulation most rem-

nant planetesimals have been cleared out and three planets have formed.

The simulation from Fig. 4 illustrates the successes of the classical model as well as its shortcomings. The two inner surviving planets bear a strong likeness to Venus and Earth. Their orbital separation and eccentricities are similar and their masses are reasonably close. In addition, their feeding zones are wide enough to extend into the outer asteroid belt and have accreted water-rich mate-

rals (see Morbidelli et al. 2000). Earth’s accretion happened on a geochemically-appropriate timescale of \( \sim 100 \) Myr and included late giant impacts suitable for Moon forma-
tion. However, the third planet bears little resemblance to Mars. Its orbit is somewhat wide of Mars’ but the big problem is that the planet is as massive as Earth.

Mars is the classical model’s Achilles heel. Simulations of the classical model systematically fail to match Mars’ small mass and instead form Mars analogs that are a factor of 5-10 too massive \( \) Wetherill 1978; Chambers 2001; Raymond et al. 2006b, 2009b; Morishima et al. 2010; Fischer and Ciesla 2014; Kaib and Cowan 2015. This was first pointed out by Wetherill (1991) and is commonly referred to as the ‘small Mars’ problem.

The small Mars problem can be understood in a very simple way. If we assume that the disk of rocky building blocks extended smoothly from within 1 AU out to the giant planet region, then there was roughly the same amount of mass in Mars’ feeding zone as in Earth’s. In the absence of large perturbations, bottom-up accretion therefore produces Mars analogs that are as massive as Earth.

There are some circumstances under which the classical model can produce small Mars analogs. For ex-
ample, if Jupiter and Saturn’s orbits were more excited \( \epsilon_{\text{Jup}} \approx \epsilon_{\text{Sat}} \approx 0.07 - 0.1 \) during terrestrial accre-
tion than they are today, then secular resonances would have been far stronger and could have acted to clear

material from the Mars zone without depleting Earth’s feeding zone (the EEJS, or ‘Extra Eccentric Jupiter and Saturn’ configuration from Raymond et al. 2009; Morishima et al. 2010; Kaib and Cowan 2015). However, the EEJS setup has its own Achilles heel: its initial conditions are not consistent with the evolution of Jupiter and Saturn in the gaseous disk. Simulations universally show that planet-disk interactions tend to drive the planets into resonance (in this case, specifically the 3:2 or 2:1 resonances Pierens et al. 2014). However, if Jupiter and Saturn were in a resonant configuration, the location of their secular resonances within the terrestrial disk would not help to produce a small Mars (e.g. Izidoro et al. 2016). A similar model invokes secular resonance sweeping during the dispersal of the gaseous disk to explain the depletion of the asteroid belt and Mars region (Nagasawa et al. 2005; Thommes et al. 2008; Bromley and Kenyon 2017). However, this model suffers from the same problem as the EEJS model: the gas giants’ orbits are not consistent with the evolution of the disk, and using appropriate (generally lower-eccentricity, resonant) orbits removes the desired depletion. An early giant planet instability may, however, produce a giant planet configuration similar to the EEJS configuration as we discuss in §4.4.

The small Mars problem is inherently coupled with the asteroid belt’s orbital excitation Izidoro et al. 2015c. While very low in total mass, the asteroids’ orbits are much more excited than the planets’, with a broad range of eccen-
crities and inclinations. The current amount of mass in the belt cannot account for its excitation because there is not enough mass for gravitational self-stirring to be ef-
It is not at all clear that planetesimals should form uni-
eral mechanisms such as vortices can also act to strongly
concentrate particles at a relatively smooth radial distributions, ALMA observa-
ations show that dust in young disks is concentrated into
a ring of planetesimals containing roughly \(2 \, \text{M}_\oplus\) centered
between V enus’ and Earth’s present-day orbits. The ter-
restrial planets that accrete from such a planetesimal annu-
lus provide a good match to the terrestrial planets’ radial
mass distribution (Hansen, 2009; Kaib and Cowan, 2013;
Walsh and Levison, 2016; Raymond and Izidoro, 2017b). In this
case, Mars’ growth was stunted when it was scattered
out of the dense ring of embryos. This naturally ex-
plains why Mars stopped accreting early and Earth’s growth was more prolonged, lasting up to
\(\sim 100\, \text{Myr}\). Earth’s growth was more prolonged, lasting up to
\(\sim 100\, \text{Myr}\).

4.4 Viable models for the inner Solar System

We now discuss three successful models for the origin of
the terrestrial planets and S-types. However, measured ages imply that non-carbonaceous objects formed in several generations over the disk’s lifetime (e.g., \(\text{Kruijer et al.}, 2017\)). This problem may in principle be solved if another generation of planetesimals formed past Earth’s orbit but still interior to the asteroid belt.

It is interesting to note that the implantation of planetes-
imals from the terrestrial planet forming region into the as-
teroid belt happens regardless of the formation model. In
the Empty Asteroid belt model this represents the main
source of volatile-depleted asteroids (Raymond and Izidoro, 2017b). However, classical model simulations have also found that terrestrial planetesimals are implanted into the main belt, and with a similar efficiency (Bottke et al, 2006a;
starts with Jupiter and Saturn’s
3 M
terrestrial planet formation. It remains to be under-
jects into the asteroid belt and to-
ke a population of 10^4 planetesimals assumed to be 100 km in diameter for the gas
drag calculation. Jupiter grew (and carved a gap in the disk) from
100-200 kyr and Saturn from 300-400 kyr, and the disk dissipated on a 200 kyr exponential timescale (uniformly in radius). The colors of planetesimals represent their starting location. This case is the least dynamic possible scenario as it neglects migration of the gas giants and the formation and migration of the ice giants. Including those factors the source region of planetesimals implanted into the main belt extends out to ∼ 20 AU [Raymond and Izidoro 2017a]. An animation of this simulation can be viewed here: https://youtu.be/Ji5ZC7CP5to.

Mastrobuono-Battisti and Perets 2017. This suggests that the present-day belt must contain a population of leftovers of terrestrial planet formation. It remains to be understood whether meteorites from such objects already exist in our collection or whether for unlucky reasons they are extremely rare (e.g., if there have not been any recent breakups of such asteroids).

The Low-mass Asteroid Belt model’s weakest point is its initial conditions. Planetesimal formation models in the context of dust concentration and streaming instability in disks with realistic structures struggle to produce planetesimal distributions consistent with the Solar System. Some models do produce rings of planetesimals at ∼ 1 AU suitable for terrestrial planet formation but no outer planetesimals that may have produced the giant planets’ cores [Drążkowska et al. 2016, Surville et al. 2016]. Other models produce planetesimals in outer planetary systems – in particular just past the snow line – but none in the terrestrial region [Armitage et al. 2016, Carrera et al. 2017, Drążkowska and Alibert 2017]. It remains unclear what conditions or processes are needed to produce planetesimal disks that are plausible precursors to the Solar System. The abundance of new studies shows that this issue may be resolved in the near-term. From a geochemical standpoint, if the building blocks of the terrestrial planets were concentrated in a narrow ring then it is difficult to understand observed differences between Earth and Mars and also why the Earth’s chemistry is consistent with accretion from a heterogenous reservoir of material [Rubie et al. 2011].

The Low-mass Asteroid Belt model is robust to a modest degree of orbital migration of Jupiter and Saturn. Once the terrestrial ring of planetesimals has formed it is mostly separated from the giant planets’ dynamical influence. In addition, there are disk-planet configurations for which Jupiter and Saturn’s migration is slow or negligible [Morbidelli and Crida 2007, Pierens et al. 2014].

The Grand Tack model invokes large-scale migration of Jupiter to sculpt the inner Solar System [Walsh et al. 2011]. The disk of rocky embryos and planetesimals is assumed to have extended smoothly from 0.5-0.7 AU out to the giant planet-forming region. Jupiter is assumed to have formed at 3-4 AU, opened a gap and started to type II migrate inward, shepherding and scattering the rocky bodies in its path (as in pervious studies focusing on exoplanets; Forg and Nelson 2005, Raymond et al. 2006a, Mandell et al. 2007). Meanwhile, Saturn grew on an exterior orbit and migrated inward in the very rapid, gap-clearing type III regime [Masset and Papaloizou 2003]. Saturn became locked in Jupiter’s exterior 2:3 resonance in a shared gap in the disk. This configuration changes the balance of disk torques felt by the coupled Jupiter-Saturn system, causing both planets to migrate outward while maintaining the resonance [Masset and Snellgrove 2001, Morbidelli and Crida 2007, Pierens and Nelson 2008, Zhang and Zhou 2013, Pierens and Raymond 2011, Pierens et al. 2014]. If Jupiter’s turnaround, or “tack” point was at 1.5-2 AU then the disk of rocky material was truncated at ∼ 1 AU, creating an edge reminiscent of the outer edge of embryos in the Low-mass Asteroid Belt model. Jupiter and Saturn’s migration continued until the disk itself started to dissipate, stranding the planets on resonant orbits consistent with the later giant planet instability (see Nesvorný and Morbidelli 2012).

The disk of rocky material sculpted by Jupiter’s migration can match the terrestrial planets’ orbital and mass distributions and formation timescales [Walsh et al. 2011, Jacobson et al. 2014, Jacobson and Morbidelli 2014, Brassier et al. 2016]. Despite having traversed the asteroid belt twice, the belt is re-populated by both scattered inner disk material (linked here with S-types) and implanted outer disk planetesimals (linked with C-types) and its orbital distribution and total mass match the present-day belt [Walsh et al. 2011, 2012, Deienno et al. 2017]. During the giant planets’ outward migration, some C-type ma-
Fig. 6.— Illustration of three models that can each match the large-scale properties of the inner Solar System. **Left:** The Grand Tack model (Walsh et al. 2011; Raymond and Morbidelli 2014; Brasser et al. 2016), invokes Jupiter’s inward-then-outward migration to truncate the inner disk of rocky material, to explain the large Earth/Mars mass ratio. The asteroid belt is emptied then re-populated from two different source regions. S-types from interior to Jupiter’s initial orbit were scattered outward during Jupiter’s inward migration then back inward during its outward migration and implanted back onto similar orbits to their original ones, albeit with a low efficiency. C-types were implanted from exterior to Jupiter’s initial orbit, also during the outward migration phase. **Center:** The Low-mass Asteroid Belt model (Hansen 2009; Drążkowska et al. 2016; Raymond and Izidoro 2017b), which proposes that the bulk of rocky planetesimals in the inner Solar System formed within a narrow ring that only extended from roughly 0.7 to 1 AU. The ring produced terrestrial planets with the correct properties. Meanwhile, the asteroid belt was contaminated with C-type planetesimals implanted during the giant planets’ growth (Raymond and Izidoro 2017a). Later dynamical evolution – perhaps explained by chaotic evolution of Jupiter and Saturn (Izidoro et al. 2016) or by secular forcing during the giant planet instability (Deienno et al. 2018) – is required to explain the asteroids’ level of excitation. **Right:** The Early Instability model (Clement et al. 2018a), which is based on the giant planet instability happening within ∼ 10 Myr after the dispersal of the gaseous disk. The instability acts to excite and deplete the asteroid belt and Mars’ feeding zone, leading to realistic Earth/Mars mass ratios and terrestrial planets. Simulations find that the terrestrial system is best matched when the surviving giant planets match the real ones.
The Grand Tack’s weakness is the outward migration mechanism. Hydrodynamical simulations of Jupiter-mass and Saturn-mass planets embedded in isothermal disks universally find that the two planets become locked in 3:2 resonance and migrate outward.\cite{Mattis2014,Peale2010,Pierens2011}. \textit{Outward migration can extend across long stretches of planet-forming disks} by \cite{Crida2007}. When the disk mass and viscosity are varied and a more realistic thermal structure is accounted for, there is a spectrum of evolutionary pathways for the gas giants’ orbits.\cite{Pierens2014}. In disks with relatively low masses (smaller than the minimum-mass solar nebula model of \cite{Weidenschilling1977,Hayashi1981}, introduced in §1.2) and moderate viscosities (viscous stress parameters of $10^{-4} \lesssim \alpha \lesssim 10^{-2}$, where $\alpha$ is the poorly-constrained parameter that controls the rate at which the disk evolves via angular momentum transport; \cite{ShakuraSunyaev1973}) the gas giants are trapped in 2:1 resonance and maintain roughly stationary orbits. In somewhat more massive disks with moderate viscosities, Jupiter and Saturn are trapped in 3:2 resonance and migrate outward as in the isothermal case. In very low-viscosity disks ($\alpha \approx 10^{-9}$) Jupiter and Saturn are trapped in 2:1 resonance but migrate outward.\cite{Pierens2014}. This diversity of potential outcomes encompasses the regimes appropriate for the three models presented in this section.

An additional issue is whether outward migration of Jupiter and Saturn can be maintained in the face of gas accretion (see discussion in \cite{Raymond2014}). In isothermal disks outward migration happens when two conditions are met \cite{Masset2001}. First, Saturn must be at least half of its present-day mass in order to open a partial gap in the disk. Second, the Jupiter-to-Saturn mass ratio must be between 2 and 4. The outward migration of Jupiter and Saturn in the Grand Tack spans a wide range in orbital distance (from $\sim$1.5 to $>5$ AU for Jupiter) and takes a significant amount of time (likely on the order of 0.5-1 Myr for the entire outward migration phase). The question is, can the gas giants maintain the requisite conditions for outward migration during this entire phase when gas accretion onto the planets is taken into account? A definite answer to this question requires a good understanding of the disk’s structure and evolution and of the planets’ gas accretion. Both of these processes remain too poorly understood to allow for a clear evaluation at this point, although given the large amount of interest in the Grand Tack there is hope for progress in the near-term.

The Early Instability model proposes that terrestrial planet formation was strongly affected by the giant planet instability \cite{Clement2018a}. This constrains the instability to have taken place within $\sim 10$ Myr of the disk’s dissipation \cite{Morb2018} in order to have an impact on stunting Mars’ growth \cite{Nimmo2007}. Within the inner Solar System, an early instability effectively causes a rapid transition from a dynamically calm state to an excited one that bears a strong resemblance to the EEJS (‘Extra Eccentric Jupiter and Saturn’) configuration discussed in §4.2 from \cite{Raymond2009b}. An early instability thus explains how the giant planets could have reached EEJS-like orbits in a self-consistent way at an early enough time to make a difference for terrestrial accretion.

Simulations by \cite{Clement2018a} show that an early instability acts to excite the asteroid belt and to clear out the Mars zone. This is in contrast to simulations of a late giant planet instability that often over-excite the terrestrial planets’ orbits (or destabilize them entirely; \cite{Brasser2013,Robig2016,Knib2016}). With an early instability, the depletion is so strong that a significant fraction ($\sim 20\%$) of simulations leave Mars’ present-day orbital region completely empty and the median Mars analog is close to Mars’ true mass. In many simulations planets grow much larger than a Mars mass at Mars’ orbital distance, and are then excited to significant eccentricities by the instability, and then collide at pericenter with the growing Earth or Venus. Mars analogs in these simulations are often stranded embryos that avoid colliding with the larger embryos when these are scattered inward.

The Early Instability model can match the asteroid belt and Earth’s water content. The belt is pre-excited before the instability by scattering of planetesimals by resident embryos \cite{Pet2001,Chan2001,OBrien2007}. During the instability the belt is depleted \cite{Robig2015,Clement2018b} and embryos are removed, and surviving planetesimals provide a decent match to the present-day belt’s eccentricity and inclination distributions \cite{Clement2018b}. As in the classical model, the belt’s S-/C-type compositional dichotomy is assumed to be matched due to previous events (and recall that Jupiter’s growth invariably implants C-types into the outer main belt \cite{Raymond2017a}). Water-rich material is delivered to Earth by the same mechanism as in the classical model \cite{Morb2000,Raymond2009,Raymond2009b}, via impacts from water-rich planetesimals and embryos originating in the outer main belt.

An appealing aspect of the Early Instability model is that it simplifies the Solar System’s timeline. Rather than invoking separate mechanisms to sculpt the terrestrial- and giant planet systems, a single key event can explain them both. In addition, the terrestrial planets are best reproduced when the giant planets reach their actual configuration (as measured by the giant planets’ orbits and the strength of different secular modes \cite{Clement2018}).
The Early Instability model’s weak points are related to the timing of the instability. Xenon isotopes are the first issue. Xenon in Earth’s atmosphere is fractionated due to hydrodynamic escape and is enriched in heavier isotopes relative to Xenon found in chondritic meteorites or in the Solar wind (Ozima and Podosek 2002). When this fractionation is taken into account, the isotopic signature of presumably primordial atmospheric Xenon is still significantly different from that of Xenon in the mantle. This suggests that another source of atmospheric Xenon had to exist with a different isotopic signature than chondrites or the Solar wind. Comet 67P/Churyumov-Gerasimenko matches the missing Xenon signature and Earth’s atmosphere can be matched with a 20–80 mixture of cometary and chondritic Xenon (Marty et al. 2017). In contrast, the mantle’s Xenon was purely chondritic (e.g., Mukhopadhyay 2012; Caracausi et al. 2016). The long-standing mystery of Earth’s Xenon thus appears to be solved if Earth’s atmospheric Xenon included an additional cometary component. In this scenario the mantle’s Xenon came from the inner Solar System as Earth was accreting. The atmosphere’s Xenon came from a late bombardment of comets necessarily related to the giant planet instability, which represent a significant source of noble gases but not of water (Marty et al. 2016). This elegant argument relates the relative timing of Earth’s accretion and the cometary bombardment linked with the giant planet instability. In principle, a very early instability would cause a cometary bombardment at the same time as accretion such that one might expect Earth’s mantle Xenon to have the same isotopic signature as its fractionation-corrected atmospheric Xenon. In contrast, a late cometary bombardment would only affect the atmospheric Xenon signature.

At face value, the Xenon constraint would seem to rule out a giant planet instability earlier than the Moon-forming impact. The relative timing of the instability and Moon-forming impact affects the ability of all terrestrial planet formation models to match the Xenon constraint, not just the Early Instability model. However, there are two caveats. First, the particular Xenon signature has only been measured in a single comet (although it is the only comet in which the signature could have been detected). Second, the repartition of the Xenon signature between the mantle and atmosphere during Earth’s impact-driven evolution is uncertain and certainly depends on factors related to Earth’s growth history and chemical evolution.

Another potential conflict comes from the orbital structure of the Kuiper belt. Nesvorný (2015) found that the high-inclination classical Kuiper belt can be matched by a smooth phase of migration of Neptune that was interrupted by the instability. If true, this would restrict the earliest possible timing of the instability to be ~ 20 Myr after gas disk dissipation, too late to stunt Mars’ growth. This constraint is relatively indirect, as other models may potentially explain the high-inclination population.

4.5 An outstanding issue: the mass deficit in the very inner Solar System

The very low amount of mass in the very inner Solar System (interior to Mercury’s orbit) is hard to understand. This feature is important because it represents a divide between the Solar System and exoplanet systems. Roughly half of all stars have Earth-sized or larger planets interior to Mercury’s orbit (e.g., Howard et al. 2010, 2012; Mayor et al. 2011). However, our terrestrial planets are consistent with having formed from a narrow ring of planetesimals and embryos between Venus’ and Earth’s present-day orbits (Hansen 2009; Walsh and Levison 2016; Raymond and Izidoro 2017). Models such as those presented in §4.4 can explain the outer edge of this ring of rocky material, for example by invoking dynamical truncation by the migrating Jupiter (Walsh et al. 2011). However, the inner edge of this ring – and the absence of other planets closer-in than Mercury – remains challenging to explain. Of course, given its very large iron core and the reduced oxidation state of its crust and mantle, Mercury itself is a challenge to explain (e.g., Ebel and Stewart 2017).

It was proposed by Leake et al. (1987) that a population of planetesimals formed on orbits interior to Mercury’s and later bombarded Mercury. Indeed, there is a belt of dynamically stable orbits between 0.06 and 0.21 AU (sometimes called the “Vulcanoid zone”; Evans and Tabachnik 1999). However, a belt of planetesimals on such close-in orbits would undergo vigorous collisional grinding (Stern and Durda 2000). Efficient removal of small bodies via radiative transport would remove the bulk of the population’s mass, and the surviving planetesimals would themselves be further depleted by Yarkovsky-effect-driven drift into unstable orbital configurations (Vokrouhlík et al. 2000). Very few km-scale planetesimals are expected to survive. Only planetesimals large enough to have significant self-gravity (D > 100 km) would have survived, and to date none have been found. This suggests that a belt of 100 km-scale planetesimals typical of the streaming instability (Simon et al. 2016; Schäfer et al. 2017) did not form closer-in than Mercury or, if it did, it was dynamically – not collisionally – removed.

Ida and Lin (2008) proposed that the rocky mass interior to roughly Venus’ orbit was removed by inward migration. They assumed that accretion proceeds roughly as a wave sweeping outward in time and that the large embryos produced by accretion migrated inward and fell onto the young Sun. Embryos massive enough to undergo gas-driven orbital migration only had time to form interior to roughly Venus’ orbit. While appealing, this model ignores the fact that disks have inner edges (as demonstrated by simulations of magnetic accretion onto young stars; see Romanova et al. 2004). Embryos can migrate to the inner edge of the disk, where they are trapped by a strong positive torque (Masset et al. 2006; Romanova and Lovelace 2006), but they remain a great distance from the surface of the star (roughly an order of magnitude larger than the stellar radius).

Raymond et al. (2016) proposed that outward migration could explain the very inner Solar System’s mass deficit.
Assuming planetesimals to form throughout the inner disk, they invoked the rapid formation of a large core of a few Earth masses close to the Sun. This object could plausibly have formed by trapping a fraction of inward-drifting pebbles, perhaps at a pressure bump in the very inner disk such as found in the disk models of [Flock et al. 2017]. For objects of a few Earth masses embedded in the inner parts of radiative viscous disks, migration is often directed outward (the details depend on parameters such as the disk metallicity and accretion rate; see [Bitsch et al. 2015a]). The large core’s outward migration through a population of planetesimals and embryos is analogous to the case of a large planet migrating inward through similar objects discussed in §3.2 (see also [Fogg and Nelson 2008; Raymond et al. 2006a; Idörö et al. 2014b]). For outward migration timescales of ~ 105 years, [Raymond et al. 2016] found that the core shepherds embryos and planetesimals in exterior resonances and clear out the inner Solar System. However, this mechanism typically broke down when the core reached 0.5-1 AU due to scattering between shepherded bodies. The core’s migration would have continued to a zero-torque location past the snow line [Bitsch et al. 2015a], often contaminating the primordial asteroid belt with material from the very inner Solar System. In the context of this model, the migrating core represents Jupiter’s core and therefore it is expected to contain a large fraction of rock. The weakness of this idea is the setup, as it requires 1) a large, close-in core that forms much faster than more distant, smaller embryos, and 2) a disk in which migration is directed outward. Yet this setup (at least point 1) appears to be plausible as many disk models have an inner pebble/dust trap (e.g. [Flock et al. 2017]) and this idea is at the heart of the drift model for super-Earth formation [Chatterjee and Tai 2014].

Morbidelli et al. (2016) proposed that the edge in the presumed disk of rocky building blocks at 0.7 AU corresponds to the location of the silicate condensation line at early times when the disk was hot. Rocky planetesimals that formed early could not have formed closer than the silicate condensation line. That early generation of planetesimals continued to grow by mutual collisions or by accreting pebbles but the lack of closer-in material would be preserved if no other planetesimals formed closer-in. There are two main uncertainties in this model. First, why did no planetesimals form closer-in at later times? Given that several generations of planetesimals are thought to have formed in the inner Solar System (e.g. [Kruijer et al. 2017]), it remains to be understood why none would form closer-in. Second, what happens to the pebbles that drift inward past the growing planetesimals? Tens of Earth masses in pebbles are likely to have drifted inward past the rocky planetesimals (e.g. [Lambrechts and Johansen 2014]). These are generally assumed to have reached the inner parts of the disk and simply sublimed. If even a small fraction is trapped, then it could lead to the formation of a large core as in the Raymond et al. (2016) model.

Two papers proposed that the early Solar System contained a population of super-Earths that were destroyed. Volk and Gladman (2015) proposed that very energetic collisions ground the planets to dust. However, examination of the collisional parameters in the simulated collisions suggest that they are far below the catastrophic destruction threshold (Leinhardt and Stewart 2012; Wallace et al. 2017) and it is hard to understand how all of the planets’ mass could have been removed. Unlike km-scale planetesimals – which, as discussed above, would indeed be ground to dust and removed on orbits closer-in than Mercury’s (Stern and Durda 2000; Vokrouhlickýs et al. 2000) – planets’ self-gravity prevents their total destruction.

In contrast, Batygin and Laughlin (2015) invoked collisional debris generated by the Grand Tack to push a population of primordial super-Earths onto the young Sun. While there are key issues related to the mechanism at play (see discussion in Raymond et al. 2016), the main problem with this model is that, as described above, planets cannot simply migrate onto their stars. Rather, their migration is blocked by the inner edge of the disk (Masset et al. 2006; Romanova and Lovelace 2006) and their evolution should be similar to the breaking the chains model (Idörö et al. 2017, 2018). See also Fig. 3 and discussion in §3.1.

To conclude, in our minds the origin of the mass deficit closer-in than Venus’ orbit remains unexplained. While successful models to match the deficit do exist, there is no clear theory that does not have significant counterarguments or require specific assumptions. Based on our current understanding it is not plausible to invoke migration onto the Sun as a mechanism for losing close-in material. Rather, close-in material may have been swept outward by a migrating core (Raymond et al. 2016), or the inner disk may simply never have produced planetesimals (Morbidelli et al. 2016). It is also entirely possible that another mechanism may be responsible.

5. EXTRAPOLATION TO EXO-EARTHS: FORMATION TIMESCALES AND WATER CONTENTS

We now turn our attention to the more general question of Earth-like planets around other stars. Just how ‘Earth-like’ should we expect Earth-sized planets in the habitable zones of their host stars to be? And how does this depend on other properties of these systems (in particular observable ones) such as the planetary system architecture, the planet size/mass and the stellar type? In this section we address the formation and water contents of potentially habitable planets. We do not address the question of what conditions are needed for habitability. Rather we simply assume that ~ Earth-mass planets in the habitable zones of their parent stars are viable candidates.

The Solar System’s terrestrial planets are thought to have accreted in steps (see §2). First, planetesimals formed from drifting pebbles and dust. Then planetesimals grew into planetary embryos by accreting pebbles and other planetesimals. Embryos grew slowly enough and remained small enough not to have undergone any significant migration.
Finally, there was an extended phase of giant collisions between embryos lasting \( \sim 100 \) Myr.

To generalize the formation of Earth-mass planets, we want to know how universal each of these steps is. Do all planet-forming disks follow the same general pattern as ours?

While great strides have been made in understanding how 100 km-scale planetesimals form, models disagree on where and when they form (e.g., Drążkowska et al. 2016; Carrera et al. 2017; Drążkowska and Alibert 2017). We can imagine that planetesimal disks might be quite diverse in their structures; for instance, the Low-mass Asteroid Belt (Greenberg et al. 1978; Wetherill and Stewart 1993; Kokubo and Ida 2000). For a minimum-mass disk embryo to grow to 1 AU at which point they excite the orbits of nearby planetesimals, decrease the effects of gravitational focusing and their growth from planetesimal accretion slows down drastically (e.g. Kokubo and Ida 1998; Leinhardt and Richardson 2005).

Yet pebble accretion should continue and even accelerate (e.g. Ormel and Klahr 2010; Lambrechts and Johansen 2012). Matching the terrestrial planets therefore requires a quenching of pebble accretion to prevent embryos at 1 AU from growing too massive (see Jacobson et al. 2018 for simulations of the pebble flux-governed bifurcation between terrestrial planets and rocky super-Earths). This may have happened as a consequence of the growth of Jupiter’s core to the pebble isolation mass, at which point it created a barrier to inward pebble drift (Morbidelli and Nesvorný 2012; Lambrechts et al. 2014; Bitsch et al. 2018b). Kruijer et al. (2017) used the temporal co-existence of meteorites with different nucleosynthetic signatures (carbonaceous vs. non-carbonaceous) to infer that Jupiter’s core did indeed provide a barrier within 1 Myr after CAIs (see also Desch et al. 2018). From that point on, pebble accretion was shut off in the inner Solar System and the terrestrial planets grew by accreting planetesimals and embryos.

This line of thinking implies that the timing of the growth of Jupiter’s core was critical (see, e.g. Bitsch et al. 2015b; 2018a). If it had grown much more slowly, pebble accretion would have generated more massive terrestrial embryos. These large embryos would then have migrated and likely followed the breaking the chains evolution discussed in §3.1 (see also Fig. 3). It is not clear that faster growth of Jupiter’s core would have had much of an effect, as pebble accretion for sub-Mars-mass embryos is relatively slow. If we assume that the growth of large outer cores varies significantly from disk to disk, these differences in timing can have big consequences. Systems with fast-growing cores may preserve their small inner rocky embryos, whereas in systems with slower-growing cores terrestrial embryos grow sufficiently fast that they cannot avoid migration.

Migration must play an important role in the formation of many habitable zone planets. This happens if a) an outer core (analogous to Jupiter’s core) formed slowly enough for large embryos to grow by pebble accretion, or b) the central star is low enough in mass that the formation timescale in the habitable zone is very short. The accretion timescale depends on the local disk surface density (in planetesimals) and the orbital timescale (Saffronov 1969). Given the strong scaling of the habitable zone with stellar type (because of the strong stellar mass-luminosity scaling: Scalò et al. 2007; Mulders et al. 2015b) the accretion timescales for planets in the habitable zones of low-mass stars are much shorter than for Sun-like stars (Raymond et al. 2007b; Lissauer and Stevenson 2007; Dawson et al. 2015). Extrapolating from Earth’s 50-100 Myr formation timescale, planets in the habitable zones of stars less massive than 1/2 to 1/3 of a Solar mass should form quickly, with most of the assembly taking place during the gas disk phase even in the absence of pebble accretion (Raymond et al. 2007b).

In these systems embryos should follow the breaking the chains behavior described above: they should migrate into resonant chains anchored at the inner edge of the disk, most of which go unstable when the gas dissipates. Habitable zone planets would undergo a final phase of giant impacts shortly after disk dispersal (Ogihara and Ida 2009; Izidoro et al. 2017, 2018). The compositions of planets in these systems would vary (Raymond et al. 2018b). Some could be made up entirely of large rocky embryos. However, given that embryos should still eventually grow large past the snow line, some planets would likely contain a significant fraction of their mass in volatiles.

What about systems with more massive central stars in which an outer core forms quickly? Fast-growing cores are themselves likely to migrate. If the disk properties are such that there is a zero-torque migration trap that lasts for close to the entire disk lifetime (Lyra et al. 2010; Bitsch et al. 2015a), the inner planetary system may be protected from the core’s migration. In some cases cores may accrete gas to become gas giants and transition to slower, type 2 migration. In those cases terrestrial planet formation should in principle follow the same pattern as in the Solar System provided the giant planets do not migrate all the way into the habitable zone (see Fogg and Nelson 2005; Raymond et al. 2006a). When migration is not stopped, the large core would plow into (or through, depending on the migration timescale) the growing terrestrial planets (Izidoro et al. 2014b; Raymond et al. 2018b). Planets can still form in the habitable zones in such systems but they are not rocky...
worlds like Earth. Rather, such planets should have compositions representative of the region past the snow line, presumably with large water contents.

Some simple analysis can hint at the distribution of outcomes. Low-mass (M) stars are the most common by number (e.g., Chabrier 2003). The disk mass is observed to be a steeper than linear function of the stellar mass (scaling roughly as $M_{\text{disk}} \propto M_*^{1.6}$ albeit with large scatter; e.g., Scholz et al. 2006; Pascucci et al. 2010). For low-mass stars, the snow line is farther-removed from the habitable zone than for FGK stars, as measured simply by the snow line to habitable zone distance ratio ($R_{\text{snow}}/R_*$). Given that the growth timescale scales with the disk mass (e.g., Safronov 1969) we expect outer cores to grow slowly around low-mass stars. Assuming that there is sufficient local material to build an Earth-mass planet in situ within the habitable zone, then rocky embryos should form quickly (Raymond et al. 2007b; Lissauer 2007; Dawson et al. 2015), undergo pebble accretion, and migrate. Embryos from past the snow line are also likely to migrate later in the disk lifetime, as disks around low-mass stars are observed to have longer lifetimes than around Sun-like stars (Pascucci et al. 2009). The location of the disk’s inner edge depends on the rotation rate of young stars and is unlikely to be a strong function of the stellar mass. This means that the habitable zones of low-mass stars are closer to the inner edge of the disk than those of higher-mass stars. Rocky planets that form near the habitable zone do not necessarily migrate far away. Given the late instabilities characteristic of breaking the chains behavior, the final planets should be a mixture of embryos that started with terrestrial compositions and those with ice-rich compositions from past the snow line. This should result in a diversity of planetary compositions, from pure rock planets to planets with tens of percent ice by mass (neglecting various water/ice loss processes; e.g., Grimm and McSween 1993; Genda and Abe 2005; Marcus et al. 2010; Monteux et al. 2018).

The main difference between this scenario for low-mass and higher-mass stars is the accretion timescale in the habitable zone (Raymond et al. 2007b). For FGK stars, terrestrial embryos in the habitable zone are less likely to grow fast enough to migrate if they only accrete planetesimals. Compared with the same setting around low-mass stars, growing rocky planets in the habitable zones of FGK stars are more likely to be protected from pebble accretion by a fast-accreting core, and to have a large ice-rich core migrate into the terrestrial zone. Given the much faster core accretion compared with terrestrial accretion, close-in planets that result from the inward migration of large cores are likely to have higher average volatile contents than for low-mass stars. Of course, in some situations outer cores will accrete into gas giants and (in some cases) remain on wide orbits. In these cases the terrestrial planets’ accretion is protected, although the gas giants’ growth may shower the terrestrial zone with volatile-rich planetesimals (Raymond and Izidoro 2017b).

Compared with FGK stars, low-mass stars are found to have more super-Earths smaller than $2R_\oplus$ but fewer sub-Neptunes between 2 and $4R_\oplus$ and a higher total average planet mass on close-in orbits (Mulders et al. 2015b,c). This can be explained by the reasoning presented above if the smaller super-Earths preferentially formed from migrating rocky embryos and larger sub-Neptunes formed mainly from migrating ice-rich embryos. Higher-mass stars have more gas giants (Johnstone et al. 2007; Lovis and Mayor 2007). While growing gas giant cores block pebbles from drifting past (Morbidelli and Nesvorny 2012; Lambrechts et al. 2014; Bitsch et al. 2018b), gas giants themselves block embryos from migrating past (Izidoro et al. 2015b). This might tilt the scales in favor of low-mass stars having a higher average total mass in close-in planets.

Let’s put the pieces together. Given that M stars dominate by number, the formation pathway of their habitable zone planets likewise dominates. The habitable zones of low-mass stars are so close-in that we expect planets to grow rapidly from both rocky and ice-rich material and to follow the breaking the chains evolution described in §3.1. The bulk of these planets’ growth took place during the gas disk phase, with last giant impacts happening during a late instability shortly after the dissipation of the disk. Their compositions are likely to span a wide range from pure rock to ice-rich depending on the objects’ individual growth histories.

Earth-mass planets in the habitable zones of FGK stars are likely to have followed one of two pathways. In systems in which cores of $10 - 20M_\oplus$ or more grow quickly past the snow line, the flux of pebbles toward the inner system is shut off. In some systems (like our own) these large cores grow into gas giants, which migrate slowly and may remain isolated from the terrestrial zone. Further growth involves planetesimal and embryo accretion, and rocky embryos are unlikely to reach high enough masses to migrate within the gas disk’s lifetime. In other systems outer, ice-rich cores do not become gas giants but instead migrate inward into the terrestrial zone. In that case habitable zone planets may typically be very volatile-rich (e.g., Kuchner 2003).

How can we use observations to constrain these ideas? Systems that follow the breaking the chains evolution should commonly form habitable zone planets. However, the habitable zone planets themselves should have a diversity of compositions, encompassing systems in which rocky embryos grew large enough to migrate and those in which they did not. Low-mass stars should be more likely to have migrating rocky embryos and, since their habitable zones are likely to be closer to the inner edge of the disk, are also more likely to retain rocky embryos/planets in the habitable zone. Finally, systems with outer giant planets (beyond 2.5-3 AU for Sun-like stars; see Raymond 2006) and no inner giants are the best candidates for having habitable zone planets with small but non-zero, Earth-like water contents.
6. DISCUSSION

6.1 Central processes that sculpt planetary systems

Two processes appear to be ubiquitous in planet formation: migration and instability. These are essential ingredients in explaining the origins of exoplanet systems as well as the Solar System. The breaking the chains model proposed by [Dražkowska et al. 2017, 2018] invokes inward migration of large embryos into long chains of mean motion resonances anchored at the inner edge of the disk. The vast majority of resonant chains become unstable when the disk dissipates, leading to a late phase of giant collisions. Given that virtually all observed super-Earths are massive enough to undergo gas-driven migration, we argued in §3.1 that, regardless of when and how they form, all super-Earth systems converge to the breaking the chains evolution.

The population of giant exoplanets may be explained by invoking the formation of multiple gas giants that also migrate into compact resonant configurations (see §3.2). As for super-Earths, the vast majority of systems undergo instabilities when the disk dissipates or perhaps even during the late phases of the disk lifetime. The outcome of an instability correlates with the Safronov number, which is the ratio of the planets’ escape speed to the local escape speed from the system [Safronov 1969, Ford and Rasio 2008]. For high Safronov numbers the planets impart strong enough gravitational kicks that scattering is favored over collisions. Giant planet instabilities thus lead to planet-planet scattering and the surviving planets match the observed giant exoplanet eccentricity distribution (as well as the correlated mass-eccentricity distribution: [Raymond et al. 2010].

Solar System formation models likewise invoke different combinations of migration and instability. All current evolutionary pathways rely on the Nice model instability in the giant planets’ orbits, although the timing of the instability is uncertain [Morbidelli et al. 2013]. The Grand Tack model uses a Jupiter’s specific migration path to deplete the asteroid belt and Mars’ feeding zone [Walsh et al. 2011]. In contrast, the Early Instability model proposes that an early giant planet instability was responsible for depleting the asteroid belt and Mars region [Clement et al. 2018].

Two additional processes are central in setting the stage: planetesimal formation and pebble accretion. They determine where and when embryos large enough to migrate can form. The Low-Mass Asteroid belt model proposes that the mass depletion in the Mars- and asteroid belt regions was inherited from the planetesimal formation stage [Dražkowska et al. 2016]. Likewise, the timing of the formation of Jupiter’s core – which acted to block the flux of pebbles to the inner Solar System – was a key moment in keeping the terrestrial planets terrestrial (see discussion in §5).

Different processes have different philosophical implications. Migration and instability both act to erase the initial conditions. Many formation pathways lead to a phase of migration, but once a system migrates it starts to forget its initial conditions, as all pathways converge to the same evolution. It is for that reason that density constraints on super-Earth compositions are so important: they are our only clue as to the planets’ origins, and even for accurate measurements density is a weaker diagnostic than one would like (see discussion in §3.1). Likewise, the chaotic nature of instabilities makes it impossible to rewind the clock on a system of planets with eccentric orbits to uncover their pre-instability configuration (although statistical studies try to do just this; e.g., [Nesvorny and Morbidelli 2012]).

In contrast, bottom-up accretion retains a memory of its initial conditions. The planets that form from a disk with a given surface density still follow that same profile [Raymond et al. 2005]. This is the central argument behind the ‘minimum-mass solar nebula’ model [Weidenschilling 1977b, Hayashi 1981]. This allows us to put strong – but not unique – constraints on the properties of the precursor disks that formed systems of small planets, assuming that accretion was the main process involved. For instance, a few different initial distributions of planetesimals and embryos can match the terrestrial planets but they all share common properties such as a strong mass deficit past Earth’s orbit and interior to Venus’ [Hansen 2009, Idorio et al. 2013].

How well do planetesimal formation and pebble accretion remember their initial conditions? In the current paradigm, planetesimals form when drifting dust and pebbles are sufficiently concentrated to trigger the streaming instability [Johansen et al. 2014]. Exactly where and when this happens depends on the underlying disk model (e.g., [Dražkowska et al. 2016, Carrera et al. 2017]). However, once planetesimal formation is triggered, objects form in a characteristic size distribution [Simon et al. 2017, Schafer et al. 2017]. Thus, while accretion may preserve a trace of where planetesimals formed and in what abundance, planetesimal formation itself does not retain a memory of dust coagulation and drift. Pebble accretion is too parameter-dependent to retain a memory of the historical pebble flux. If the pebble size (or Stokes number) and the initial planetesimal/embryo mass were known this might be possible.

6.2 Planet formation pathways and bifurcations: How did our Solar System get so weird?

There are a few key bifurcation points in planetary system formation. At these points, small differences in outcome lead to very different evolution. We consider the key bifurcation points to be: 1) disk properties, 2) planetesimal formation (where? when?), 3) giant planet formation, 4) instability trigger (timing). We now go over each of these bifurcation points, then discuss which pathway the Solar System must have followed.

A star’s protoplanetary disk is its cradle, where its planetary system is born and raised. The characteristics of the disk and its evolution are perhaps the single most important
factor in planet formation. While the detailed structure and evolution of disks are themselves poorly understood (see Morbidelli and Raymond, 2016), observations and theory demonstrate that there is a diversity in disk mass, structure and lifetime (e.g., Haisch et al. 2001; Williams and Cieza 2011; Bate 2018). The disk mass may itself be the key determinant of planetary system evolution (Greaves et al. 2006; Thommes et al. 2008). More massive disks should more readily form planetesimals, embryos and gas giants. Given their higher abundance of solids, higher-metallicity stars should also more readily form planetesimals and planets.

Where and when planetesimals form is of vital importance. If planetesimals form early then they are bathed in a flux of pebbles and can quickly grow into large embryos/cores and perhaps even gas giants (e.g., Lambrechts and Johansen 2014; Bitsch et al. 2015b). However, if planetesimals only form late – perhaps triggered by the dissipation of the gas disk and the accompanying increase in dust/gas ratio (Throop and Bally 2005; Carrera et al. 2017) – then the bulk of their growth must appeal to gas-free processes such as planetesimal accretion and no gas giants can form. The radial distribution of planetesimals is naturally of vital importance to planet formation. The Low-mass Asteroid Belt model relies on planetesimals forming in a narrow ring in the inner Solar System (Dražkowska et al. 2016) whereas the Grand Tack (Walsh et al. 2011) and Early Instability (Clement et al. 2018a) models are devised assuming that planetesimals did indeed form in the Mars region and asteroid belt and that a depletion mechanism was needed.

The formation of a giant planet is essentially two bifurcation points. Once a planet’s core reaches the pebble isolation mass (of roughly $20 M\oplus$ at Jupiter’s orbit for characteristic disk models; Lambrechts et al. 2014; Bitsch et al. 2018b), the pebble flux is blocked and the entire planetary system interior to the core is cut off from further pebble accretion. We argued in §5 that the timing of the formation of a pebble-blocking core relative to the growth of inner embryos is the central parameter that determines whether most super-Earths are likely to be rocky or ice-dominated. After a core undergoes rapid gas accretion to become a giant planet and carves a gap in the disk, it also blocks the inward migration of any other large cores that form on exterior orbits (Izidoro et al. 2015b). This is a second way in which a wide-orbit planet cuts off its inner planetary system from the inward-drifting/migrating mass. Of course, if the wide-orbit planet migrates inward then it itself becomes that inward-migrating mass.

Triggering instability is the final and perhaps most dramatic bifurcation point. While instability appears to be near-ubiquitous, the impact of instabilities can vary. For instance, the Solar System’s giant planets are thought to have undergone an instability but only a very weak one when compared with the instabilities in most exoplanet systems. Indeed, the instabilities characteristic of giant exoplanet systems often drive the growing terrestrial planets into their host star (Raymond et al. 2011, 2012). The late stage accretion of the terrestrial planets also represents a form of instability that concluded with the Moon-forming impact. The breaking the chains (Izidoro et al. 2017, 2018) evolution characteristic of super-Earth systems causes much more dramatic late instabilities that involve collisions between much larger (typically $\sim 5 M\oplus$) objects and, given their close-in orbits, much higher collision speeds. Of course, a small minority of systems avoid instability. Stable systems easily recognized by their resonant orbits, which are systematically destroyed by instabilities (although scattering does generate resonances in a small fraction of cases; Raymond et al. 2008a).

What path must the Solar System have taken with regards to these bifurcations? The Sun’s planet-forming disk may have been somewhat more massive than average, with enough mass to form the cores of several gas giants (which total $\sim 40 - 50 M\oplus$), but not enough to form multiple Jupiter-mass planets. A few Earth-masses worth of rocky planetesimals must have formed early in the terrestrial planet region, either in a smooth disk or in one or more rings. Planetesimals also formed beyond the snow line and produced the giant planets’ cores. Jupiter’s core grew fast enough to starve the inner Solar System of pebbles within $\sim 1 M\text{yr}$, keeping the precursors of carbonaceous and non-carbonaceous meteorites physically separated (Kruijer et al. 2017), preventing further growth of the terrestrial planets’ constituent embryos, and fossilizing the snow line (Morbidelli et al. 2016). Jupiter’s growth also stopped the ice giants and Saturn’s core from migrating into the inner Solar System (Izidoro et al. 2015b). When the disk dissipated, the inner disk of terrestrial embryos entered its late instability, which lasted $\sim 100$ Myr but during which Mars remained mostly isolated and protected. The giant planets’ orbits became unstable sometime within the 500 Myr following disk dissipation. While the disk’s dispersal is the main natural trigger for instability, some geochemical arguments (e.g., the atmospheric Xenon constraint; see §4.3) point to a later trigger. While the instability did clear out the primordial Kuiper belt, it was far less dramatic than instabilities characteristic of extra-solar systems.

The Solar System’s presumed evolution contains multiple unusual occurrences. First, the gas giants’ masses are quite different. The fact that the most massive giant exoplanets have the highest eccentricities (Wright et al. 2009) indicates that massive gas giants (of roughly $1 M_{\text{Jup}}$ or above) typically form in systems with other, roughly equal-mass gas giants that go unstable (Raymond et al. 2010). Second, Jupiter’s orbit remained wide of the terrestrial region. This may be because of the dynamical influence of Saturn; the Jupiter-Saturn system can migrate outward or remain on near-stationary orbits depending on the disk properties (Pierens et al. 2014). However, avoiding inward migration depends on a Jupiter/Saturn mass ratio close to its current value (Masset and Snellgrove 2001), so this unusual occurrence may be intrinsically linked with the previous one. Third, the giant planet instability did not include
any close encounters between Jupiter and Saturn. Simulations show that such an encounter would likely have ejected Saturn and stranded Jupiter with an eccentricity of $\sim 0.2$ (Morbidelli et al. 2007). In other words, Jupiter’s eccentricity would be typical of giant exoplanets if its instability had proceeded in typical fashion. However, in that case the terrestrial planets may well have been driven into the Sun (Raymond et al. 2011).

We interpret these unusual occurrences as why our Solar System is weird. These evolutionary steps explain why Sun-Jupiter systems are rare within the known exoplanet sample (see §1.1). This in turn implies that terrestrial planet systems like ours are also rare, although an understanding of the timescales of different processes is needed to assess that assertion in a more careful way (see §5).

6.3 A digression on the significance of models

A successful formation model is expected to match a planetary system (or a distribution of systems) in broad strokes and with a suitable success rate. But what exactly constitutes the ‘breadth’ of the ‘strokes’ needed for success? And at what rate is a model deemed successful? These inherently philosophical questions are central to models of planet formation. If a model matches the Solar System in 1% of simulations, should we consider the problem solved? Or should we continue to test other models? And if model A provides a match in 30% of cases and model B in 10%, can we be confident that model A is truly preferred over model B?

We do not pretend to have a concrete solution, but we think it important to keep such considerations in mind. We expect that in the future, global planet formation modeling may make use of more rigorous statistical methods to address these issues.

6.4 Paths Forward

There remain a plethora of outstanding problems in planet formation. As described in §4.4, studies will strongly constrain (and may falsify some) models of Solar System formation in the coming years. Nonetheless, we encourage the development of new models. NASA’s OSIRIS-REx and JAXA’s Hayabusa2 missions will return samples of carbonaceous asteroids (the B-type Bennu and the C-type Ryugu) that will certainly improve our understanding of the formation conditions of such objects and provide additional constraints on their origins. NASA’s Lucy mission will study Jupiter’s co-orbital asteroids, thought to have been captured during the giant planet instability (Morbidelli et al. 2005), and NASA’s Psyche mission will study an apparently metallic asteroid that may have originated in the inner parts of a differentiated planetary embryo. Meanwhile, upcoming exoplanet-focused instruments such as NASA’s TESS and ESA’s PLATO and ARIEL missions will deepen our understanding of the orbital architecture of planetary systems as well as their more detailed characteristics. This will provide additional constraints on exoplanet formation models.

Among our ever-increasing stockpile of extremely valuable data, we conclude this chapter by emphasizing the need for global models to connect the dots. Models should not be constrained by dogma or current paradigms. Of course, a model is only viable if matches observations or measurements. A model is most useful if it is testable in the near term. And a model is most relevant when it lays the broadest possible foundation. This means putting Solar System formation in the context of extra-solar planets.

Acknowledgments.

We thank referees John Chambers and Kevin Walsh for constructive reports, and are grateful to all of our colleagues who helped develop the ideas presented here. We each thank the Agence Nationale pour la Recherche for funding and support via grant ANR-13-BS05-0003-002 (grant MOJO). A. I. acknowledges financial support from FAPESP (grants 16/12686-2 and 16/19556-7). S. N. R. also acknowledges NASA Astrobiology Institute’s Virtual Planetary Laboratory Lead Team, funded under solicitation NNH12ZDA002C and cooperative agreement no. NNA13AA93A.

REFERENCES

Adams E. R., Seager S., and Elkins-Tanton L. (2008) Astrophys. J., 673, 1160.
Adams F. C. (2010) Anna. Rev. Astron. Astrophys., 48, 47.
Adams F. C., Hollenbach D., Laughlin G. et al. (2004) Astrophys. J., 611, 360.
Adams F. C. and Laughlin G. (2003) Icarus, 163, 290.
Agnor C. B., Canup R. M., and Levison H. F. (1999) Icarus, 142, 219.
Agnor C. B. and Lin D. N. C. (2012) Astrophys. J., 745, 143.
Aglol E. and Fabrycky D. C. (2017) Transit-Timing and Duration Variations for the Discovery and Characterization of Exoplanets, p. 7.
Albrecht S., Winn J. N., Johnson J. A. et al. (2012) Astrophys. J., 757, 18.
Alexander C. M. O., , Bowden R., Fogel M. L. et al. (2012) Science, 337, 721.
Alexander C. M. O., Boss A. P., and Carlson R. W. (2001) Science, 293, 64.
Alexander C. M. O., McKeean K. D., and Altwegg K. (2018) Science, 363, 142.
Alexander R., Pascucci I., Andrews S. et al. (2014) Protostars and Planets VI, pp. 725–745.
Alexander R. D. and Pascucci I. (2012) Mon. Not. R. Astron. Soc., 422, L82.
Alibert Y. (2017) Astron. Astrophys., 606, A69.
Alibert Y., Mordasini C., Benz W. et al. (2005) Astron. Astrophys., 434, 343.
ALMA Partnership, Brogan C. L., Pérez L. M. et al. (2015) Astrophys. J. Lett., 808, L3.
Andrews S. M., Wilner D. J., Zhu Z. et al. (2016) Astrophys. J. Lett., 820, L40.
Stern S. A. and Durda D. D. (2000) Mon. Not. R. Astron. Soc., 374, 1092.

Steffen J. H. and Hwang J. A. (2015) Mon. Not. R. Astron. Soc., 453, 15.

Snellgrove M. D., Papaloizou J. C. B., and Nelson R. P. (2001) Mon. Not. R. Astron. Soc., 322, 1012.

Stern S. A. and Durda D. D. (2000) Icarus, 143, 360.

Stewart S. T. and Leinhardt Z. M. (2012) Astrophys. J., 751, 32.

Sumi T., Udalski A., Bennett D. P. et al. (2016) Astrophys. J., 825, 112.

Surville C. and Mayer L. (2018) ArXiv e-prints.

Szulágyi J., Masset F., Lega E. et al. (2016) Mon. Not. R. Astron. Soc., 460, 2853.

Touboul M., Kleine T., Bourdon B. et al. (2007) Icarus, 187, 689.

Tout K. J., Ogihara M., Morbidelli A. et al. (2016b) Astron. Astrophys., 596, A74.

Swift J. J., Johnson J. A., Morton T. D. et al. (2013) Astrophys. J., 764, 105.

Szułagyi J., Masset F., Lega E. et al. (2016) Mon. Not. R. Astron. Soc., 460, 2853.

Tanaka H., Takeuchi T., and Ward W. R. (2002) Astrophys. J., 565, 1257.

Tanaka H. and Ward W. R. (2004) Astrophys. J., 602, 388.

Tera F., Papanastassiou D. A., and Wasserburg G. J. (1974) Earth and Planetary Science Letters, 22, 1.

Terquem C. and Papaloizou J. C. B. (2007) Astrophys. J., 654, 1110.

Teske J. K., Ciardi D. R., Howell S. B. et al. (2018) arXiv:1804.10170.

Thommes E., Nagasawa M., and Lin D. N. C. (2001) Mon. Not. R. Astron. Soc., 322, 1012.

Thommes E. W., Bryden G., Wu Y. et al. (2008b) Astrophys. J., 675, 1538.

Thommes E. W., Duncan M. J., and Levison H. F. (2003) Icarus, 161, 431.

Thommes E. W., Matsumura S., and Rasio F. A. (2008c) Science, 321, 814.

Thommes E. W., Matsumura S., and Rasio F. A. (2008d) Science, 321, 814.

Throop H. B. and Bally J. (2005) Astrophys. J. Lett., 623, L149.

Timpe M., Barnes R., Kopparapu R. et al. (2013) Astron. J., 146, 63.

Touboul M., Kleine T., Bourdon B. et al. (2007) Nature, 450, 1206.

Tremaine S. (1993) in: Planets Around Pulsars, vol. 36 of Astronomical Society of the Pacific Conference Series, (edited by J. A. Phillips, S. E. Thorsett, and S. R. Kulkarni), pp. 335–344.

Triaud A. H. M. J., Collier Cameron A., Queloz D. et al. (2010) Astron. Astrophys., 524, A25+.

Tsiganis K., Gomes R., Morbidelli A. et al. (2005) Nature, 435, 459.

Turner N. J., Fromang S., Gammie C. et al. (2014) Protostars and Planets VI, pp. 411–432.

Udry S. and Santos N. C. (2007) Annu. Rev. Astron. Astrophys., 45, 397.

Van Eylen V., Agentoft C., Lundkvist M. S. et al. (2018) Mon. Not. R. Astron. Soc., 479, 4786.

Veras D. and Armitage P. J. (2004) Mon. Not. R. Astron. Soc., 347, 613.

Veras D. and Armitage P. J. (2005) Astrophys. J. Lett., 620, L111.

Veras D. and Armitage P. J. (2006) Astrophys. J., 645, 1509.

Veras D. and Raymond S. N. (2012) Mon. Not. R. Astron. Soc., 421, L117.

Vokrouhlický D., Farinella P., and Bottke W. F. (2000) Icarus, 148, 147.

Volk K. and Gladman B. (2015) Astrophys. J. Lett., 806, L26.

Walker R. J. (2009) Chemie der Erde / Geochemistry, 69, 101.

Wallace J., Tremaine S., and Chambers J. (2017) Astron. J., 154, 175.

Walsh K. J. and Levison H. F. (2016) Astron. J., 152, 68.

Walsh K. J., Morbidelli A., Raymond S. N. et al. (2011) Nature, 475, 206.

Walsh K. J., Morbidelli A., Raymond S. N. et al. (2012) Meteoritics and Planetary Science, 47, 1941.

Ward W. R. (1986) Icarus, 67, 164.

Ward W. R. (1997) Icarus, 126, 261.

Warren P. H. (2011) Earth and Planetary Science Letters, 311, 93.

Weidenschilling S. J. (1977a) Mon. Not. R. Astron. Soc., 180, 57.

Weidenschilling S. J. (1977b) Astrophys. Space Sci., 51, 153.

Weidenschilling S. J. and Marzari F. (1996) Nature, 384, 619.

Weiss L. M. and Marcy G. W. (2014) Astrophys. J. Lett., 783, L6.

Weiss L. M., Marcy G. W., Petigura E. A. et al. (2018) Astron. J., 155, 48.

Wetherill G. W. (1978) in: IAU Colloq. 52: Protostars and Planets, (edited by T. Gehrels), pp. 565–598.

Wetherill G. W. (1985) Science, 228, 877.

Wetherill G. W. (1991) in: Lunar and Planetary Institute Science Conference Abstracts, vol. 22 of Lunar and Planetary Inst. Technical Report, p. 1495.

Wetherill G. W. (1996) Icarus, 119, 219.

Wetherill G. W. and Stewart G. R. (1993) Icarus, 106, 190.

Williams J. P. and Cieza L. A. (2011) Annu. Rev. Astron. Astrophys., 49, 67.

Winn J. N. (2018) Planet Occurrence: Doppler and Transit Surveys.

Winn J. N., Fabrycky D., Albrecht S. et al. (2010) Astrophys. J. Lett., 718, L145.

Winn J. N. and Fabrycky D. C. (2015) Annu. Rev. Astron. Astrophys., 53, 409.

Winn J. N., Noyes R. W., Holman M. J. et al. (2005) Astrophys. J., 631, 1215.

Wise A. W. and Dodson-Robinson S. E. (2018) Astrophys. J., 855, 145.

Wittenmyer R. A., Butler R. P., Tinney C. G. et al. (2016) Astrophys. J., 819, 28.

Wolfgang A., Rogers L. A., and Ford E. B. (2016) Astrophys. J., 825, 19.

Wright J. T., Fakhouri O., Marcy G. W. et al. (2010) ArXiv e-prints.

Wright J. T., Marcy G. W., Butler R. P. et al. (2008) Astrophys. J. Lett., 683, L63.

Wright J. T., Marcy G. W., Howard A. W. et al. (2012) Astrophys. J., 753, 160.

Wright J. T., Upadhyay S., Marcy G. W. et al. (2009) Astrophys. J., 693, 1084.

Yang C.-C., Johansen A., and Carrera D. (2017) Astron. Astrophys., 606, A80.

Youdin A. N. and Goodman J. (2005) Astrophys. J., 620, 459.

Zahn J.-P. (1977) Astron. Astrophys., 57, 383.

Zhang H. and Zhou J.-L. (2010) Astrophys. J., 714, 532.

Zhou J.-L., Aarseth S. J., Lin D. N. C. et al. (2005) Astrophys. J.
Zhu W., Petrovich C., Wu Y. et al. (2018) Astrophys. J., 860, 101.