



Collision Chains among the Terrestrial Planets. III. Formation of the Moon

Erik Asphaug¹, Alexandre Emsenhuber^{1,2,3}, Saverio Cambioni¹, Travis S. J. Gabriel⁴, and Stephen R. Schwartz¹¹Lunar and Planetary Laboratory, University of Arizona, 1629 E. University Boulevard, Tucson, AZ 85721, USA; asphaug@lpl.arizona.edu²Space Research and Planetary Science, University of Bern, Gesellschaftsstrasse 6, 3012 Bern, Switzerland³Universitäts-Sternwarte München, Ludwig-Maximilians-Universität München, Scheinerstraße 1, D-81679 München, Germany⁴School of Earth and Space Exploration, Arizona State University, 781 E. Terrace Mall, Tempe, AZ 85287, USA

Received 2021 April 20; revised 2021 July 29; accepted 2021 July 30; published 2021 September 23

Abstract

In the canonical model of Moon formation, a Mars-sized protoplanet “Theia” collides with proto-Earth at close to their mutual escape velocity v_{esc} and a common impact angle $\sim 45^\circ$. The “graze-and-merge” collision strands a fraction of Theia’s mantle into orbit, while Earth accretes most of Theia and its momentum. Simulations show that this produces a hot, high angular momentum, silicate-dominated protolunar system, in substantial agreement with lunar geology, geochemistry, and dynamics. However, a Moon that derives mostly from Theia’s mantle, as angular momentum dictates, is challenged by the fact that O, Ti, Cr, radiogenic W, and other elements are indistinguishable in Earth and lunar rocks. Moreover, the model requires an improbably low initial velocity. Here we develop a scenario for Moon formation that begins with a somewhat faster collision, when proto-Theia impacts proto-Earth at $\sim 1.2v_{\text{esc}}$, also around $\sim 45^\circ$. Instead of merging, the bodies come into violent contact for a half hour and their major components escape, a “hit-and-run” collision. N -body evolutions show that the “runner” often returns ~ 0.1 – 1 Myr later for a second giant impact, closer to v_{esc} ; this produces a postimpact disk of ~ 2 – 3 lunar masses in smoothed particle hydrodynamics simulations, with angular momentum comparable to canonical scenarios. The disk ends up substantially inclined, in most cases, because the terminal collision is randomly oriented to the first. Moreover, proto-Earth contributions to the protolunar disk are enhanced by the compounded mixing and greater energy of a collision chain.

Unified Astronomy Thesaurus concepts: Earth-moon system (436); Planet formation (1241); Lunar origin (966)

1. Introduction

This is the third paper in our study of nonaccretionary giant impacts and their consequences, now considering the origin of the Moon. We begin with a synopsis of the prevailing theories for lunar formation and the underlying physics and chemistry, emphasizing the challenges and ongoing developments that motivate our effort.

1.1. Parameters of the Giant Impact

Analysis of the first lunar samples and geophysical and remote sensing data from Apollo missions (Cummings 2019) led to independent arguments for the giant impact origin of the Moon (Asphaug 2014; Hartmann 2014; Melosh 2014; Canup et al. 2021, and other reviews). Hartmann & Davis (1975) proposed that instead of growing through the accumulation of planetesimals, Earth finished its accretion with several calamitous mergers involving similar-sized bodies. Consistent with this theory, Cameron & Ward (1976) showed that if a Mars-sized protoplanet (eventually called “Theia”; Halliday 2000) was accreted by proto-Earth at a typical impact angle of $\theta_{\text{coll}} \sim 45^\circ$, with a velocity at contact $v_{\text{coll}} \approx v_{\text{esc}}$ (Equation (1)), this would account for the high angular momentum of the Earth–Moon system. It was further recognized that the Moon, if born at the interface of colliding terrestrial mantles, would end up with a predominantly silicate composition, explaining its small core (Goldstein et al. 1976), at most a few percent of a lunar mass M_{C} . This makes it a unique body in the solar system, although possibly akin

(Asphaug & Reufer 2013) to the massive Saturnian ice moons Tethys and Iapetus that also do not have cores.

A collision between two planetary bodies of masses m_{tar} and m_{imp} and radii r_{tar} and r_{imp} begins on a hyperbolic trajectory. The velocity at contact v_{coll} is thus at least the mutual escape velocity

$$v_{\text{esc}} = \sqrt{2G(m_{\text{tar}} + m_{\text{imp}})/(r_{\text{tar}} + r_{\text{imp}})}, \quad (1)$$

where G is the gravitational constant. The parameters of the canonical model of Moon formation are nominally $m_{\text{tar}} \simeq 0.9 M_{\oplus}$ and $m_{\text{imp}} \simeq 0.1 M_{\oplus}$ ($\gamma = m_{\text{imp}}/m_{\text{tar}} = 1/9$), in which case $v_{\text{coll}} \gtrsim 9 \text{ km s}^{-1}$, around twice the sound speed in geologic materials.

Shock waves from the colliding interfaces span the globes (Tonks & Melosh 1993), and other violent processes are set in motion. The energetics are especially complex for giant impacts that end in merger (Carter et al. 2020), and depend on scale (Asphaug 2010) and on the starting composition and temperature (Genda et al. 2012). The kinetic energy is the change in gravitational potential up to contact, plus the energy of original velocities and rotations. For accretions there is additionally the change in gravitational potential that plays out over hours or even days as the cores and mantles merge and attain equilibrium. Collisional energy is applied to momentum transfer and spin-up of the target and is dissipated to internal energy by shocks, viscous heating, and friction. Heat is also consumed and released in phase transformations, as well as to surface energies in the case of expanding plumes (Asphaug et al. 2011).

Then, there are advective losses, insofar as planetary collisions are not perfect mergers. The hottest, highest-velocity debris escapes, even in slow accretions like the canonical model, removing energy from the largest remnant. High angular

momentum material escapes, as well as interfacial material lost to jetting and spallation (e.g., Pierazzo & Melosh 2000). Counter-intuitively this can mean that a faster, more lossy collision can contribute *less* total energy to the largest remnant of a giant impact, i.e., Earth, than a slower, more accretionary collision.

Change in enthalpy dH occurs in materials that start out at depth inside colliding bodies. Volumes of materials that were stable at kilobars to megabars of pressure (hydrostatic conditions in the midmantle of Theia) find themselves, hours later, inside the droplets, clumps, and filaments of the protolunar disk and the escaping debris, at a greatly reduced pressure. Decompression makes available an enthalpy $dH = VdP + PdV$, where P is the pressure and $\rho \equiv 1/V$ is the density, analogous to an ascending volcanic plume but happening globally (Asphaug et al. 2006, 2011).

In summary, while the specifics are not well understood and depend on currently unknown initial states and conditions, the thermodynamic aspects of giant impacts readily account for the igneous geology of the Moon, the compelling evidence for a lunar magma ocean (LMO; Wood et al. 1970; Warren 1985), and the depletion and fractionation of volatiles (Wolf & Anders 1980).

As for the dynamical scenario, that of a Mars-sized body straying in orbit to collide with proto-Earth, the traditional scientific context is the “late stage” of terrestrial planetary growth (e.g., Chambers & Wetherill 1998; Agnor et al. 1999; Raymond et al. 2004), when dozens of planet-sized “oligarchs” (Kokubo & Ida 2002) became gravitationally excited, collided, and *overall* merged until there were four terrestrial planets and the Moon. Three-dimensional N -body simulations of the late stage (e.g., Chambers & Wetherill 1998; Kenyon & Bromley 2006; Kokubo & Genda 2010), starting with dozens of oligarchs orbiting the Sun, can produce final architectures resembling the inner solar system, a handful of finished planets between 0.5 and 2 au, the consequence of dozens of giant impacts in an epoch lasting ~ 10 – 100 Myr. These simulations are evocative but seldom end up with terrestrial planetary systems closely resembling our own. And as we describe below, the physics of collisions implemented in these models is incomplete (Emsenhuber et al. 2020).

An alternative theory (Youdin 2011; Bitsch et al. 2015; Chambers 2016; Matsumura et al. 2017) is “pebble accretion,” where planets form directly from the sweep-up of small particles concentrated by aerodynamic sorting in the nebula. Pebble accretion has become the accepted framework for giant planet core formation (e.g., Lambrechts & Johansen 2012); today the debate is on the extent that it applies to terrestrial planet formation, given that the process must happen in the presence of the solar nebula, in the first ~ 2 Myr. Johansen et al. (2021) propose, based on simulations, that Theia accreted directly as a fifth terrestrial planet, rapidly attaining a final mass of $\sim 0.4 M_{\oplus}$, forming between proto-Earth ($\sim 0.6 M_{\oplus}$) and Mars. In contrast to there being a late stage involving dozens of collisions, Moon formation would be the *only* giant impact event, after proto-Earth and Theia somehow got dislodged from their starting orbits.

Ages of Moon formation obtained from geochemistry are broadly consistent with estimates for the duration of the late stage based on N -body simulations and gas-free dynamics (Safronov 1969; Wetherill 1980). The oldest geochemical age is ~ 10 – 30 Myr after solar system formation, derived from Hf–W isotopic analysis (Jacobsen 2005). This may measure the

core–mantle differentiation of precursor bodies; on the other hand, formation ages $\gtrsim 60$ Myr are obtained when the cosmogenic production of ^{182}W is accounted for in rocks from the lunar surface (Touboul et al. 2007). Pristine zircons in Apollo 14 samples crystallized ~ 67 Myr after solar system formation (Barboni et al. 2017); if they solidified directly from the LMO, this would place the giant impact somewhat earlier. This is within the error bars of $\sim 90 \pm 30$ Myr estimated by Jacobson et al. (2014) based on the abundance of accreted siderophiles after lunar solidification and the $\gtrsim 70$ Myr age (Kruijer & Kleine 2017) estimated from the absence of radiogenic W variations in lunar samples. Still later solidification ages, ~ 150 – 200 Myr (Carlson et al. 2014) or later (Maurice et al. 2020), are obtained from combined Sm–Nd isotopic analyses of suites of crustal rocks. These may be consistent with the ~ 150 Myr age obtained from U–Pb isotope systematics assuming that Pb was fractionated by the giant impact (Connelly & Bizzarro 2016).

These estimates and others (e.g., Nemchin et al. 2009) date different aspects of the Moon-forming giant impact and its magmatic aftermath. A molten Moon-sized silicate body radiating at the liquidus will solidify in < 1 Myr, except that it forms an insulating crust that regulates deeper cooling (Solomon 1977), leading to an estimated ~ 10 – 100 Myr solidification timescale, absent other factors. If solidification is relatively rapid (e.g., Elkins-Tanton et al. 2011; Perera et al. 2018), then geochemical closure might be close to the time of the giant impact. Otherwise, LMO solidification might have been delayed by radioactive elements (incompatibles) concentrated in the residual melt or KREEP and by tidal friction during the Moon’s orbital evolution, starting from a few radii away. Early, powerful tides may have sustained a partially melted state inside the LMO for $\gtrsim 100$ Myr according to Meyer et al. (2010), which might be consistent with the theory of forced convection sustaining a lunar dynamo (Dwyer et al. 2011).

We must be cautious in tying these closure ages, and the associated timescales for LMO solidification, to the Moon’s and the Earth’s formation. A Ceres-sized asteroid colliding with the Moon after it had solidified (even a late-returning remnant of the giant impact itself as discussed below) would create a gigantic basin whose melt volume would exceed the crater volume (Melosh 1989), producing a new magma ocean spanning the impacted hemisphere. Crystallization sequences of lunar samples might in that case date a later, regional solidification (Borg et al. 2011) and not the Moon-forming giant impact.

1.2. Collisional Dynamics

A giant impact starts with a dynamical perturbation that leads to a collision and its aftermath. The collision begins hours ahead of contact, when the planets approach within their tidal influence, several planetary radii apart. The smaller body is subjected to increasing tidal deformation and torque, which changes its cross section and rotation, and hence changes the impact parameters—considerations that are astrophysical as much as geophysical. It is therefore natural that the modern understanding of the giant impact process began with the application of 3D stellar astrophysics simulations, modified for geophysical equations of state (EOSs; Benz et al. 1986)—the smoothed particle hydrodynamics (SPH) method from which

our present code derives (see Section 3 and Emsenhuber et al. 2021, hereafter Paper II).

Early suites of SPH modeling experiments (e.g., Cameron & Benz 1991; Cameron 1997) showed that massive silicate disk formation is favored by “graze-and-merge” collisions (Leinhardt et al. 2010); these are common for giant impacts within a few percent of the escape velocity, at impact angles in the range of $\theta_{\text{coll}} \gtrsim 30^\circ$. These off-axis accretions bring significant angular momentum and, for most angles, merge the metallic cores so that the disk is made almost entirely of silicates. Calculations using the same code at higher resolution (Canup & Asphaug 2001) confirmed that the scenario postulated by Cameron & Ward (1976), a collision by a Mars-sized Theia at $\sim 45^\circ$ at close to v_{esc} into an almost-finished proto-Earth, leads to outcomes that can match the mass distribution, composition, and angular momentum of the Earth–Moon system.

The projectile is not stopped abruptly in a giant impact the way a bullet is stopped in a block of wood, or an asteroid striking Earth. Giant impacts are similar-sized collisions, meaning they are usually significantly off-center, like a rugby ball hitting a basketball, or two pool balls. A direct hit is not typical. When the sine of the impact angle exceeds the fractional radius of the metallic core, i.e., $\sin^{-1}(1/2) = 30^\circ$ in the case of chondritic compositions, the colliding cores do not touch along the projected impact vector. This is more head-on than the average impact angle of 45° , so that most giant impacts are grazing by this definition (Asphaug 2010).

The impact angle specified in the canonical model (Cameron & Ward 1976), about 45° , is required to provide sufficient Earth–Moon angular momentum, assuming zero rotation to start with. As noted, most of the projectile “misses” the target (e.g., Movshovitz et al. 2016), and the bodies do not slow down enough to end in immediate accretion. What happens next depends on the impact velocity v_{coll} . If close to v_{esc} , most giant impacts are graze-and-merge collisions like the canonical case, drawn-out pinwheel mergers that are good at stranding mantle material (in this case silicates) in orbit. If somewhat faster, the projectile can escape, continuing as a “runner” discussed below. The latter, called “hit-and-run” collisions, are generally poor candidates for Moon formation because they fail to accrete the angular momentum, although Reufer et al. (2012) identified some potential scenarios.

Accretion at higher v_{imp} can occur, but this requires a relatively head-on impact to decelerate the projectile so as to end in capture, which in turn brings insufficient angular momentum to explain the Earth–Moon system. As for lower velocities, the database of simulated satellite-forming giant impacts includes examples of direct capture, when just the right fraction of angular momentum is lost and dissipated. In this manner a slow, highly grazing projectile can end up relatively intact and in orbit. But this also captures the core into orbit, so it is contrary to the composition of the Moon but is consistent with Pluto–Charon formation (Canup 2011).

For the canonical giant impact parameters the runner does not escape but loops back for a reimpact several hours later, along with shredded outer-mantle and crustal/oceanic debris. The reimpact is, in turn, a slower grazing merger, although sometimes it can fly over the target in the loop-back orbit, which can then lead to potential direct-capture scenarios. More often, the gravitationally bound cores of the colliding bodies rapidly combine into a lopsided, spinning central body that transfers gravitational torques to protodisk silicates, and to

protolunar clumps, creating one or more tidal spiral arms that expand upon release from shock and from the pre-collision hydrostatic pressure inside Theia.

In summary, giant impact accretions are complex processes whose detailed outcomes depend sensitively on the mass ratio, impact angle, velocity, composition, and thermal state of the colliding bodies. Yet despite this complexity, numerical simulations of graze-and-merge collisions starting with canonical giant impact parameters often lead to reproducible outcomes that are consistent with Moon formation, although the specific results depend on the numerical method and resolution as discussed below.

The canonical scenario is widely appealing. It is compatible with astrophysical predictions of oligarchic growth (e.g., Kokubo & Ida 2002), where two planetary embryos get into crossing orbits owing to late gravitational perturbations and collide. It satisfies the constraints of Earth–Moon dynamics, producing a stable disk that is more massive than the Moon and a disk and an Earth–Moon system with the right angular momentum (Canup et al. 2001). The disk is made of molten silicates, satisfying the riddle of the Moon’s small iron core (Goldstein et al. 1976) and providing a thermodynamic explanation for the LMO. A hot start might also explain the volatile-poor nature of the Moon, although volatile escape is perhaps not as significant as originally thought (see Lock et al. 2018; Nakajima & Stevenson 2018), and the Moon is not as volatile depleted as originally thought (e.g., Hauri et al. 2011).

Giant impacts play a significant role in planetary habitability (e.g., Zahnle et al. 2007), so it is important to identify the scenario that made the Moon. Major outcomes (such as hit-and-run or graze-and-merge) are sensitive to the impact angle and velocity, and these parameters are stochastic, so overall there is a predicted diversity of giant impact outcomes that could account for the diversity of potentially habitable planets (Asphaug 2010). If giant impacts were rare, or if Moon formation required special conditions, this might explain why Venus lacks a moon and rotates slowly, lacks a magnetic field, and is overall “un-Earth-like.” One can speculate that perhaps the final giant impact into Venus was retrograde, robbing angular momentum and destroying an existing moon. Or, Venus never suffered a giant impact (Jacobson et al. 2017; Johansen et al. 2021), or it accreted from systematically different giant impacts than Earth (the subject of Paper II). Giant impact modeling does not answer these questions but allows us to better understand and evaluate the assumptions and quantitative implications of proposed scenarios.

1.3. Limits of Simulations

Meaningful simulations of giant impacts must accurately model multiple stages of physics: the dynamics leading up to the collision, the formation and release of impact shocks and compressions, the global long-range interactions of self-gravity, and melt-vapor thermodynamics, which depends sensitively on nonlinear EOSs and precise treatment of energy and entropy by the code. The EOS, in turn, assumes that we know the compositions and constitutive properties of the colliding bodies.

As for modeling more complicated planets, such as an ocean on the proto-Earth, this is resolvable in 1D (Genda & Abe 2005) and 2D, but not yet in routine 3D simulations of giant impacts. An ocean would have to be more than a few hundred kilometers deep to be adequately resolved in our

current models. A target with such a large water fraction would respond quite differently to a giant impact for reasons of moment of inertia and EOS (e.g., Haghighipour et al. 2018). For now we stick to the well-studied two-component planets (rock and metal), focusing on the physics in 3D and relying on thermodynamically consistent EOSs for forsterite and iron.

SPH is accurate in computing global self-gravity, and unlike most grid-based codes, it conserves angular momentum precisely. This, as well as its relative ease of use and analysis, has made it a standard method for simulating giant impacts. Modern implementations of SPH are accurate in energy conservation to better than 1%. As with any hydrocode, the results are limited by resolution, since self-gravity and off-axis dynamics require 3D evolution for several gravitational timescales, up to days. Artifacts such as numerical viscosity and numerical tension can masquerade as physical, and sparse-particle regions can experience fictional shear forces. Heating (e.g., Gabriel & Allen-Sutter 2021) and disk stability (e.g., Raskin & Owen 2016) are sensitive to the artificial viscosity used to resolve shocks.

It is encouraging that simulations of the canonical model by various groups (e.g., Reufer et al. 2012; Nakajima & Stevenson 2015), even using diverse techniques and resolutions (Marcus et al. 2009; Canup et al. 2013; Barr 2016), often give comparable results for the mass, composition, and angular momentum of the protolunar disk, for similar initial parameters. However, this is not universally the case, and disagreements might influence the viability of the hypothesis. For example, Kraus et al. (2012) argue for significantly greater vapor production compared to the “flying magma ocean” predicted earlier (Stevenson 1987), factors that also depend on the thermal state of the target (Genda et al. 2012; Hosono et al. 2019). Gravitational instability (coagulation) and the onset of shocks that might disperse accreting regions of the disk both depend on the local sound speed in the disk. The phase, pressure, and temperature of the postimpact disk may thus determine whether the Moon can accrete before the material disperses. Multi-million-zone Eulerian simulations by Wada et al. (2006) end with a disk comparable to other simulations but indicate that it would be subject to internal shocks that could severely frustrate the growth of the Moon.

A comparative study by Canup et al. (2013) using different hydrocodes (CTH and SPH) at various resolutions (e.g., 10^4 – 10^6 particles) found first-order similarities in terms of disk mass and angular momentum, but they identified substantial structural differences. Some simulations ended with one or two massive protolunar clumps totaling about one lunar mass, embedded in a less massive disk, while others ended with a clump-free disk. They found no trend in clumping with numerical resolution. According to Hosono et al. (2017), numerical predictions for circumterrestrial postimpact structures have not converged even with millions of particles.

Even when giant impacts are properly and accurately computed by a code, they are sensitive to relatively minor changes in impact angle, velocity, composition, and mass ratio (roughly in that order; see Timpe et al. 2020). Strong variation is observed within common ranges of those parameters (Gabriel et al. 2020). Predictive sensitivity extends to the static compressibility of the EOS (Wissing & Hobbs 2020). The EOS for forsterite and iron used in our ongoing research (ANEOS; see Section 3) gives a more realistic and complete treatment of shocked material than the simple Tillotson EOS that was used in earlier studies (e.g., Canup & Asphaug 2001;

Agnor & Asphaug 2004). Yet despite its shortcomings (Stewart et al. 2020), an Earth-mass planet constructed using Tillotson can end up with approximately the correct target radius and central density, whereas planets constructed using ANEOS end up a fraction too compressed. This makes a giant impact slightly more grazing ($\sin \theta_{\text{coll}} = b/r_{\text{tar}}$) for a given impact parameter b , and the bodies start and end more gravitationally bound. It therefore remains quite challenging to relate differences in giant impact outcomes to differences in material physics, versus differences in implementation.

1.4. Isotopic Contradictions

The canonical model is a slow ($v_{\text{imp}} \approx v_{\text{esc}}$), near-perfect merger with an accretion efficiency $\xi \gtrsim 0.98$ in most simulations, where

$$\xi = (m_{\text{lr}} - m_{\text{tar}})/m_{\text{imp}} \quad (2)$$

equals 1 for perfect merger. Here m_{lr} is the mass of the final largest remnant (Earth in this case) and $m_{\text{tar}} > m_{\text{imp}}$ are the target and projectile masses (i.e., proto-Earth and Theia). Earth ends up being made out of both bodies proportional to their contributions, so about $1/(1 + \xi\gamma)$ proto-Earth. The disk, though, ends up deriving mostly from Theia, specifically the mantle fraction that is approximately opposite its initial contact with Earth. This distal mid- to outer mantle of Theia has the greatest captured angular momentum and so contributes more than two-thirds of the protolunar disk mass in simulations (e.g., Canup & Asphaug 2001; Reufer et al. 2012). The fraction $1 - \xi$, also mostly from Theia, remains in heliocentric orbit, unaccreted by Earth for now.

The provenance of protolunar disk material is a serious problem for the canonical model. The Moon was not blasted out of Earth’s mantle, as is often envisioned; it is instead a lossy collisional capture of a silicate portion of Theia. Lunar rocks should therefore be readily distinguishable from Earth rocks, in the way that meteorites from asteroids or from Mars are distinct from each other and from Earth in isotopic composition. Theia would be distinct from proto-Earth according to oligarchic growth (Kokubo & Ida 2002), where embryos are born in well-separated feeding zones of the nebula, where temperature, pressure, and composition variations would give rise to isotopic differences (e.g., Burkhardt et al. 2012).

Yet Apollo samples from the nearside, as well as lunar meteorites from all over the Moon, have oxygen isotopic ratios ($\delta^{17}\text{O}$) that are indistinguishable from Earth (Wiechert et al. 2001; Spicuzza et al. 2007; Hallis et al. 2010), even to a few ppm (Young et al. 2016). Other rock-forming species are indiscernible as well (Dauphas et al. 2014; Dauphas 2017), for instance, Ti (Zhang et al. 2012) and Cr (Qin et al. 2010), which have quite different thermophysical and petrological behaviors and associations; radiogenic W (Touboul et al. 2007); and Si (Armytage et al. 2012). It presents a striking dilemma: if the canonical model is correct, how could Theia have been so alike Earth?

The most convenient explanation would be that Theia and proto-Earth were born in the same feeding zone, as this would also be consistent with a low-velocity collision. The idea that Theia was a Trojan planet that co-accreted in a 1:1 resonance with proto-Earth (Belbruno et al. 2005) has been ruled out (Kortenkamp & Hartmann 2016), but the idea of Theia being a

dynamical neighbor of proto-Earth has some basis (Mastrobuono-Battisti et al. 2015; Quarles & Lissauer 2015). Other arguments are to the contrary (Kaib & Cowan 2015). The explanation would require Theia to be born so close in time and space to proto-Earth as to have their compositions indistinguishable, yet to have their collision deferred by ~ 100 Myr after their formation. Perhaps the inner solar system is more isotopically homogeneous than meteorites suggest; until we have samples from Mercury and Venus, this dilemma will be mired in speculation.

The same-feeding-zone explanation is not without geochemical paradoxes of its own. If the feeding zones were the same, then Si would have started off the same. But deep inside proto-Earth, Si would have partitioned more effectively to the Fe–Ni core under pressures that were 10 times greater than inside Theia. According to Armytage et al. (2012), there would be heavier Si remaining in Theia’s mantle and hence the Moon, yet lunar Si is the same. Another puzzle is FeO, which appears to be significantly more abundant in lunar mantle than in Earth. This can be accounted for by FeO disproportionation at higher pressure (Frost & McCammon 2008), and possibly by changes in oxidation in the aftermath of giant impacts (Cambioni et al. 2021). It could also indicate a more oxidized Theia (Budde et al. 2019); however, in that case an identical isotopic reservoir is even less likely. And lastly, there is radiogenic W, where noncosmogenic $^{182}\text{W}/^{184}\text{W}$ ratios are the same in rocks from Earth and the Moon (Touboul et al. 2007) and could require coincident differentiation in Theia and proto-Earth.

The sampled lunar crust is not the bulk Moon (Gross et al. 2014), and moderately volatile elements such as Rb and Zn show evidence for isotopic fractionation in crustal rocks (Nie & Dauphas 2019). Contrary to prior studies, Cano et al. (2020) report that lunar rocks do have small variations in $\delta^{17}\text{O}$ around an average value that matches that of Earth. They argue that their reported variations correlate systematically with petrology, and they propose that the interior of the Moon has heavier oxygen than Earth (being mostly Theia), while the crustal suite happens to overlap the bulk silicate Earth owing to fractionation between the magma ocean and an escaping lunar silicate atmosphere. In any case, isotopic composition is perhaps not so straightforward a constraint on scenarios of Moon formation.

1.5. Mixing, Diffusion, and Layering

One way to reconcile the Earth–Moon isotopic similarity is to invoke widespread compositional diffusion between the protolunar disk and the postimpact Earth’s silicate atmosphere and magma ocean. This would apply to any mechanism of Moon formation, not just the canonical model, and would be more effective in alternative models that involve greater collisional heating and mixing.

Pahlevan & Stevenson (2007) calculate that proto-Earth and a vapor-rich protolunar torus could have approached 90% diffusive equilibrium in ~ 100 yr. This is longer than the ~ 0.1 – 1 yr accretion timescale obtained by Ida et al. (1997) and Kokubo et al. (2000b), although in those approaches the Moon forms from an N -body disk of large solid particles subject to gas-free coagulation. Even for the relatively gentle canonical model, simulations predict >20 wt% silicate vapor production (e.g., Nakajima & Stevenson 2014). A melt-vapor torus would coagulate on longer timescales, limited by cooling (Stevenson 1987).

Diffusion in this manner requires a material connection between the disk and postimpact Earth. Simulations of the

canonical model produce a postimpact disk that straddles the Roche limit R_{Roche} , about 2.9 Earth radii for silicate moonlets, and that spreads onto Earth. Bodies coagulating interior to R_{Roche} tend to shear apart; this could lead to heating, fluidization, and viscous spreading in the model of Salmon & Canup (2012). Exterior to R_{Roche} , moonlets would grow under stable conditions and rapidly accrete to form a massive satellite (e.g., Kokubo et al. 2000b).

Oxygen represents almost 40% of the mass of silicates, so its widespread diffusion would ultimately require the exchange of most of the angular momentum in the disk. Angular momentum transfer would lead to the dispersal of the disk (Melosh 2014), a contradiction. Furthermore, wholesale oxygen exchange between Earth’s early mantle and the disk requires a subsequent explanation for why the Moon ended up several-fold depleted in water (Hauri et al. 2011). One possibility is the subsequent escape of volatiles at the Hill radius of the accreted Moon (Charnoz et al. 2021).

A compromise advocated by Salmon & Canup (2012) is that the massive inner disk would more rapidly attain isotopic equilibration with Earth. This equilibrated disk would spread, a fraction condensing into bodies that would be accreted by the proto-Moon outside R_{Roche} . The deep mantle and core of the Moon would be made mostly of Theia, while the outer mantle and crust (that we sample) would be made of more Earth-like inner-disk material. This would contrast with how lunar formation geology is usually interpreted, a complex plagioclase cumulate above a global magma ocean (e.g., Warren 1985).

Another idea for equilibration invokes a series of not-quite-giant impacts that each launched mostly proto-Earth-derived silicates into orbit (Rufu et al. 2017). If impacting at high enough velocity and working together, these could build up the Moon sequentially. This approach has the benefit of averaging out the heterogeneities contributed by individual projectiles, allowing for a realistically diverse bombarding population. The critical problem is that each event would have to add to the previous disk or proto-Moon and not erode it, or cause it to de-orbit, implying that the collisions would somehow have to be aligned in angular momentum.

1.6. System Angular Momentum

An original tenet of the giant impact hypothesis (Cameron & Ward 1976) is that the Earth–Moon system ended with approximately the same angular momentum L_{EM} as it has today. The Moon currently accounts for $\sim 4/5$ of the system angular momentum, acquired from Earth during its outward tidal migration. Projecting back in time to Moon formation, Darwin (1879) calculated that most of L_{EM} was in the early Earth that spun with a period $P_{\text{rot}} \approx 5$ hr, much faster than the other planets. This quantity is generally thought to be conserved during coupled Earth–Moon evolution.

This is different from the related question of how much angular momentum L_{disk} has to end up in the disk orbiting the postimpact Earth at the end of a collision. A minimum is sometimes taken to be $\sim 0.18L_{\text{EM}}$, that of a $1 M_{\oplus}$ body in circular orbit at the Roche limit (Canup 2004). Although this is useful as a comparative criterion, L_{disk} continues to evolve beyond the time frame of SPH simulations owing to the mass asymmetry of the spinning postimpact Earth and the ongoing evolution of massive clumps.

Like isotopic abundances, the system angular momentum L_{EM} has been taken as a fundamental quantitative constraint on

giant impact scenarios. But this approach is not straightforward either. As the Moon migrated out, the Sun increased in relative gravitational influence. At some point the Moon encountered the evection resonance (Touma & Wisdom 1998), where its orbital precession around Earth equals 1 yr. The angle between the Sun and the line of apsides would be constant, so a torque would accumulate that could drain the Earth–Moon angular momentum (Ćuk & Stewart 2012). If the Moon migrated slowly, then it could have been trapped in the evection for some time. Slow migration, in turn, requires low tidal friction and thus severely constrains Earth’s interior and crustal evolution (Zahnle et al. 2015).

Wisdom & Tian (2015) showed that capture into evection can occur only for a narrow range in relative tidal dissipation inside Earth and the Moon; otherwise, orbital eccentricity increases and the Moon escapes the resonance. Ward et al. (2020) obtained the result that evection could not be maintained for most tidal parameters. It is possible that other perturbations (Ćuk et al. 2016) and quasi-resonances (Tian et al. 2017; Rufu & Canup 2020) may have applied (see, however, Tian & Wisdom 2020). So with significant caveats the door remains open to higher and lower angular momentum scenarios.

One such scenario is the merger, at close to v_{esc} , of equal-mass semi-Earths. Cameron (1997) demonstrated the idea but did not consider it further, as it accretes twice the L_{EM} . It was examined more carefully by Canup (2012) in the context of isotopic similarity. Its appeal is that if the merging bodies are very close in bulk composition (e.g., both chondritic), then Earth and the Moon would accrete in equal fraction out of each body, guaranteeing isotopic similarity no matter how diverse their origins. However, if the twin planets varied by several percent in mass or moment of inertia, or rotated much differently, then the collision would be lopsided and the compositions would be distinct, so the proposed solution applies for a quite narrow range of cases. Still, equal-mass collisions are 2–3 times as energetic as canonical collisions (Lock et al. 2020), so postimpact mixing, diffusion, and layering described above would be more effective at reducing isotopic differences following the collision.

In terms of evidence, there are no half-Earth planets remaining, but two Theia-mass planets, Mars and Mercury, that could be leftovers (Asphaug & Reufer 2014). But absence of evidence is not evidence of absence, especially in this case, as the semi-Earths would be beneath our feet. By whatever process two equal-mass progenitors may have come about, the idea faces two major dynamical challenges. One is the high L_{EM} of the merger. The other is to explain how two major planets would be perturbed into colliding orbits ~ 100 Myr after their formation. In the scenario of pebble accretion (e.g., Johansen et al. 2021) they would form on rather circular orbits at least $1/3$ au apart, though there may be a tendency for resonant pairs (Lambrechts et al. 2019).

If the Earth–Moon system can shed angular momentum, then perhaps the proto-Earth could have been spinning 10 times faster than today. In this case the target’s equator would be already almost escaping owing to centrifugal forces, so that Moon formation could be an “impact-triggered fission,” the scenario proposed by Ćuk & Stewart (2012). In their Figure 1 a projectile impacts the equator of the oblate proto-Earth at 20 km s^{-1} , about twice v_{esc} , somewhat counter to its rotation ($\theta_{\text{coll}} = -20^\circ$, $\varphi = 180^\circ$, $P_{\text{rot}} = 2.3 \text{ hr}$). The equator spins

toward the projectile at up to $\sim 10 \text{ km s}^{-1}$, greatly increasing the impact energy while lowering the preimpact binding energy. The result is a global-scale explosion and expansion (“synestia”; Lock et al. 2018), much of which recondenses onto the spinning core, leaving a vaporized torus of Earth-derived material to make the Moon.

Theia must come from the outer solar system to have such high velocity (Jackson et al. 2018), and from there the delivery probabilities are low (e.g., Morbidelli et al. 2000). The projectile, twice the mass of Ganymede, would be from a lost or undiscovered population. On the other hand, the composition of Theia would be inconsequential because the projectile is dispersed; the scenario perfectly addresses the isotopic similarity. It begs the question of how proto-Earth could have spun up so fast in the first place. Terrestrial planets end up spinning relatively slowly under pebble accretion (Visser et al. 2020). As for late-stage accretion, Agnor et al. (1999) found that perfect merging would lead to fast-spinning planets; however, their conclusion was that perfect merging is a wrong assumption. To spin proto-Earth to $P_{\text{rot}} < 3 \text{ hr}$ would seem to require one or more grazing mergers ahead of the proposed collision.

2. Hit-and-run Return

Giant impacts at $\gtrsim 1.2v_{\text{esc}}$ are common in N -body studies of late-stage planet formation and accord with classical theory (e.g., Wetherill 1980), yet few studies have considered Moon formation in this velocity range (e.g., Reufer et al. 2012; Rufu et al. 2017). These are in a different regime than canonical graze-and-merge scenarios but much slower than impact-triggered fission.

Giant impacts at higher velocity convey more angular momentum than the canonical model, as well as more kinetic energy, so in principle they are promising for Moon formation. But for a similar-sized projectile to convey the requisite angular momentum at $\gtrsim 1.2v_{\text{esc}}$, the impact angle must still be $\gtrsim 30^\circ$, so the target does not manage to capture the angular momentum, although there is more of it coming in. Such hit-and-run collisions can at first resemble graze-and-merge events—the energetics and dynamics are initially quite similar. But instead of returning on a bound orbit, the impactor (now runner) escapes, often barely. Some of the projectile may be accreted by the target in a hit-and-run, and there can be significant mixing during the half hour of violent contact. But the high angular momentum material that is needed to explain L_{EM} and the protolunar disk is lost.

The idea that Theia might escape from a giant impact presents another opportunity for explaining the isotopic similarity, as in the scenario by Ćuk & Stewart (2012) but at lower energy. Reufer et al. (2012) demonstrate that relatively head-on (less than $\sim 40^\circ$) hit-and-runs by massive projectiles can dredge up Earth-derived mantle and retain it as a protolunar disk, with most of Theia’s contribution escaping. Focusing on mass ratios about twice that of the canonical model ($\gamma \sim 1:5$) and a range of projectile–target compositions including icy and metallic bodies, they identify scenarios that improve by a factor of two the isotopic fit. Their scenario applies to a narrow range of parameters, however, and they do not end up with quite a lunar mass in the disk (although using the stricter periape criterion; see below). And a hit-and-run collision leaves open the question (Figure 1), what happens to the runner?

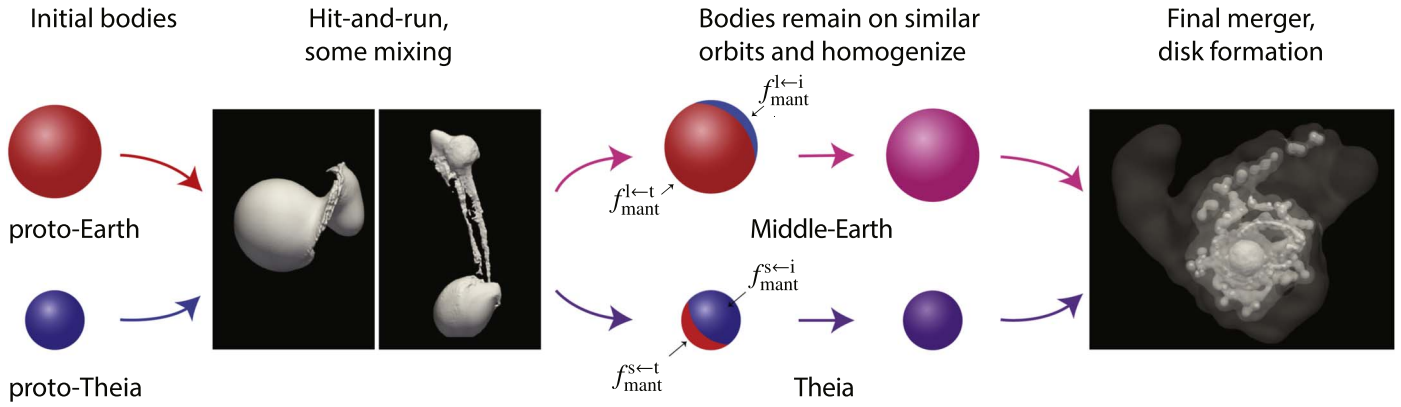


Figure 1. Schematic representation of the hit-and-run return scenario for Moon formation. Grayscale images are renderings of representative SPH results of the two giant impacts. The first is a hit-and-run, producing largest (l , aka middle-Earth) and second-largest (s , the runner, aka Theia) remnants. The second collision is slower, a graze-and-merge that forms a massive disk. Colors indicate mixing, where t and i refer to target and impactor (proto-Earth and proto-Theia), respectively, per Equation (11).

2.1. Accretion Efficiency

Moon formation is such a persistent problem in computational physics in part because of the precision that is required to obtain a meaningful solution. A realistic calculation must resolve the dynamics and thermodynamics of the hottest, most tenuous few percent of the total colliding mass and determine its fraction that neither escapes nor is accreted but ends up in stable orbit around the spinning Earth.

A much simpler and more robust computation is the *accretion efficiency* $\xi \leq 1$ (Equation (2)) of a giant impact, whose determination focuses on the coldest, least tenuous, most massive components of a collision outcome that are gravitationally bound. The major bound masses are the postimpact target and the runner, if there is one; these are reliably obtained to within a few percent in comparative simulations except around the sensitive transition between graze-and-merge and hit-and-run.

Accretion efficiency discriminates strongly between scenarios of Moon formation. For example, $\xi \sim 0$ for Reufer et al. (2012; hit-and-run), $\xi < 0$ for Čuk & Stewart (2012; fission), and $\xi \sim 1$ for the canonical model. Perhaps because of the decades of focus on the latter, and for the simplicity of the assumption, perfect merging ($\xi = 1$) has been a common expediency in N -body studies of planet formation: two bodies come in (often with an “expansion factor” so they collide even if they miss by several radii) and a combined body goes out with the equivalent mass and momentum. Agnor et al. (1999) examined the implication of this assumption, tracking the accretion of angular momentum along with the accretion of matter. They found that planets end up spinning well beyond the disruption limit assuming perfect merging, with Earth-mass planets acquiring rotation periods shorter than 1 hr. This is impossible (Chandrasekhar 1969); the reality is that accretion is inefficient for similar-sized collisions when the product of impact velocity and impact angle is large.

For planets of a given composition, such as differentiated chondritic bodies, ξ is approximately a function of the dimensionless parameters

$$\xi = \xi(\gamma, v_{\text{coll}}/v_{\text{esc}}, \theta_{\text{coll}}), \quad (3)$$

where $\gamma = m_{\text{imp}}/m_{\text{tar}}$ is the mass ratio and $\sin(\theta_{\text{coll}})$ is equivalent to the normalized impact parameter $b/(r_{\text{imp}} + r_{\text{tar}})$. Composition, as represented by a variable core mass fraction Z ,

introduces another dimensionless parameter that affects collisions systematically (Timpe et al. 2020).

Departures from scale invariance occur because EOSs are nonlinear, and because constitutive properties (strength, friction, porosity) are disproportionately significant at smaller masses (Emsenhuber et al. 2018). In collisions between accreting planetesimals (“small giant impacts”) v_{esc} is subsonic, so the linear (e.g., elastic) response of the EOS matters most, plus crushing behavior and friction (Jutzi 2015) that depend on gravity and pressure. The nonlinear EOS responses become dominant at larger scales for two main reasons: the escape velocity is higher, so the collisions are faster and the shocks are more intense (Carter et al. 2020; Gabriel & Allen-Sutter 2021), and more massive bodies begin and end a collision more centrally condensed, changing the gravitational energetics, making it increasingly difficult to exhume core material, and increasing the preponderance of hit-and-run collisions (Gabriel et al. 2020).

The mass ratio γ is associated with each Moon-formation scenario: $\sim 1/9$ in the canonical model, $\sim 1/20$ in the model of Čuk & Stewart (2012), $\sim 2/3$ for pebble accretion (Johansen et al. 2021), and ~ 1 for semi-earths. The closer in mass, the more physically grazing a collision ends up being, for a given impact angle, making hit-and-runs more common for higher γ . As for impact angle θ_{coll} , this parameter is stochastic for most giant impacts (e.g., Emsenhuber & Asphaug 2019b) and follows the familiar Shoemaker (1962) probability distribution $dp(\theta_{\text{coll}}) = \sin 2\theta_{\text{coll}} d\theta_{\text{coll}}$. Half of giant impacts are between 30° and 60° . The present study focuses on $\sim 45^\circ$ collisions, as they have maximum probability.

Impact velocity $v_{\text{coll}} = \sqrt{v^2 + v_{\text{esc}}^2}$, where v is the relative velocity before the encounter, which depends on the starting orbits of the colliding bodies. In a self-stirred system (not excited by external perturbers) v is stochastic, with an average value that may be estimated by the statistics of encounters (e.g., Safronov 1969). In the absence of nebular gas and particle drag, mutual gravitational encounters increase random velocities until, over time, they might approach the escape velocity of the major perturbers (Wetherill 1980), in which case $\langle v_{\text{coll}} \rangle \rightarrow \sqrt{2}v_{\text{esc}}$ depending on the distribution, with slower and faster collisions as outliers. Gravitational drag causes the relative velocities of the major bodies to decrease with an increasing mass fraction in planetesimals (O’Brien et al. 2006; Chambers 2013; Kaib & Cowan 2015).

The collisions that characterize the late stage depend sensitively on whether $\langle v_{\text{coll}} \rangle$ is $1.1v_{\text{esc}}$, $1.2v_{\text{esc}}$, or faster, as

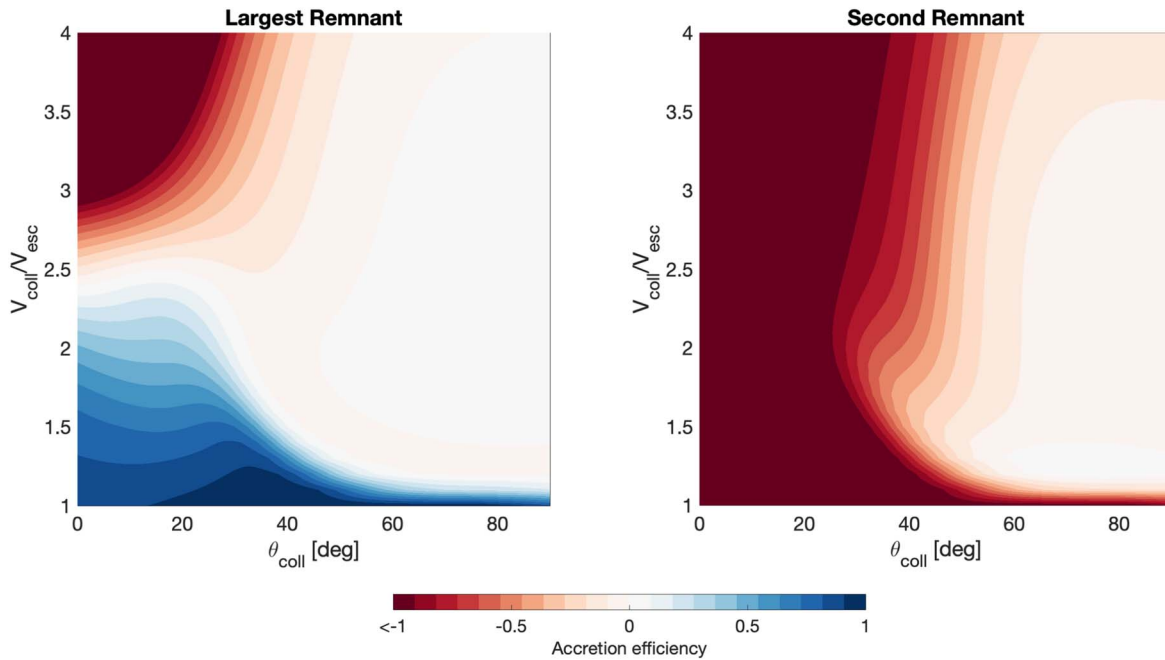


Figure 2. Left: accretion efficiency ξ of giant impacts (Equation (2)) plotted for chondritic planets, mass ratio $\gamma = 1/6$, and target mass $0.9 M_{\oplus}$. ξ is evaluated using the surrogate model of Emsenhuber et al. (2020) and differs slightly from the simulations using our updated SPH code. Perfect merger ($\xi \approx 1$) is dark blue; the canonical model plots around the bottom center. Increases in velocity or angle (collision angular momentum) lead to a transition from graze-and-merge to hit-and-run ($\xi \approx 0$, white). Right: the accretion efficiency of the projectile is computed using the associated surrogate model of Emsenhuber et al. (2020), showing fractional mass loss from the surviving runner. The lower-velocity hit-and-runs studied here produce relatively intact runners, $\xi_{\text{imp}} \approx 0$.

this represents the transition from accretion to hit-and-run for expected mass ratios (see, e.g., Genda et al. 2012; Leinhardt & Stewart 2012, and other studies). This transition was first explored by Agnor & Asphaug (2004) on the basis of dozens of SPH simulations using the same code and EOS as Canup & Asphaug (2001) but starting with Mars-mass progenitors. They noted a precipitous drop in ξ with impact angle over the velocity range associated with moderate self-stirring, establishing that realistic accretion efficiency must be accounted for in N -body studies of planet formation.

The strong tendencies with impact angle and velocity can be appreciated analytically (Gabriel et al. 2020) and are the basis for physically motivated scaling laws (e.g., Stewart & Leinhardt 2012; Movshovitz et al. 2016) and semiempirical models (Kokubo & Genda 2010). However, detailed 3D numerical simulations are required to assess accretion efficiency in detail. But while thousands of giant impact simulations have been published at high resolution in 3D, this is still a small sampling of the parameter space, so machine learning has been applied to learn generalized underlying behavior from the published simulations.

Cambioni et al. (2019) and Emsenhuber et al. (2020) obtained surrogate models for accretion efficiency ξ and other giant impact outcomes by training neural networks on 800 giant impact simulations of terrestrial planets, for a variety of impact angles and velocities and masses between 0.001 and $1 M_{\oplus}$. The training data (Reufer 2011; Gabriel et al. 2020) are high-resolution simulations, about 200,000 SPH particles total, of giant impacts starting from nonrotating bodies of 30/70 wt% iron/forsterite composition. A more expansive parameter sampling was generated by Timpe et al. (2020), although at lower resolution (10,000 particles in the projectile). This enabled a more general assessment of machine-learning approaches, applied to a data set that includes preimpact rotation and compositional variation. They obtained the important result that

γ , $v_{\text{coll}}/v_{\text{esc}}$, θ_{coll} , and core mass fraction Z are the most significant parameters in predicting giant impact outcomes. Preimpact rotation is significant, but mostly to the postimpact rotation.

Surrogate models make rapid predictions of known confidence for the outcomes of collisions within the parameter range of the training validation set. They are effectively functions mapping inputs to outputs and thus allow for inversion of evolution models. However, for collisions around the transitions, specific predictions made by surrogate models are less reliable than for major regions (erosion, hit-and-run, merger) because the physical response is nonlinear and the parameter sampling so far is coarse.

The surrogate model for ξ from Emsenhuber et al. (2020) is plotted in Figure 2, evaluated for $\gamma = 1/6$ and for $0.9 M_{\oplus}$ targets. Because the training data were generated using the prior version of our code (Reufer 2011), this surrogate model computes slightly different outcomes than our updated simulations, but that is not important here. Note that accretion is more than 50% efficient only in the slowest giant impacts, or collisions that are close to head-on. Only a subset of slow collisions have $\xi \sim 1$ (dark blue), including the canonical model, which plots near the bottom center. At somewhat faster velocities the projectile becomes an escaping runner for impact angles greater than $\sim 40^\circ$, the hit-and-run regime to the right of the graze-and-merge boundary, $\xi \sim 0$. The plot is for a specific mass ratio, and it would feature more hit-and-run outcomes for larger γ and fewer for smaller γ for the reasons described above.

We can discriminate the kinds of hit-and-run collisions in terms of the accretion efficiency of the *projectile*,

$$\xi_{\text{imp}} = (m_{\text{run}} - m_{\text{imp}})/m_{\text{imp}}, \quad (4)$$

where m_{run} is the mass of the runner. This is analogous to the accretion efficiency of the target (Equation (2)), except it is

negative because the projectile loses mass. A surrogate model for ξ_{imp} was similarly obtained by training on the same database of simulations (Emsenhuber et al. 2020), plotted in the right panel of Figure 2, also for $\gamma = 1/6$.

Any material not in the largest two remnants is called the debris mass ξ_{esc} , where by mass conservation $\xi + \xi_{\text{imp}} + \xi_{\text{esc}} \equiv 1$ when debris is measured in units of m_{imp} . In a low-velocity hit-and-run, the debris mass is usually much less than 1% of the total colliding mass (Paper II); the rest is in the postimpact target and runner. In a canonical merger it can be $\sim 1\%$ – 2% (in terms of projectile mass, $\xi \gtrsim 10\%$), although this is sensitive to the velocity and angle of the collision and the thermal states of the colliding bodies.

At higher velocities and intermediate angles, the escaping projectile from a hit-and-run can be destroyed ($\xi_{\text{imp}} < -1/2$), producing voluminous debris even while causing only minor erosion of the target. Multiple massive runners can be derived from one projectile, with highly varied compositions (Asphaug et al. 2006; Sekine & Genda 2012). Significant erosion of the target requires still higher energy events, in which the projectile escapes as a plume of debris. Relatively head-on impacts are required (upper left; $\xi \ll 0$) because otherwise the projectile glances off with little momentum and energy transfer.

Target disruption, a collision removing more than half the target mass, is outside the plot, far upper left. For similar-sized collisions, target disruption requires impact velocity several times v_{esc} and a nearly head-on impact to achieve sufficient coupling. This is faster than can be commonly expected in the late stage, or during any accretionary epoch; otherwise, merger would be rare and too frequently undone. Catastrophic disruptions of planet-sized targets during the late stage, such as invoked for the origin of Mercury (Benz et al. 1988), require the scattering influence of migrating gas giants (e.g., Carter et al. 2015) and are then further unusual because of the required small impact angle.

In summary, the simplification that pairwise collisions either result in efficient merger, or above some energy threshold destroy the target, ignores the vast diversity of outcomes in the middle. A corollary, once acknowledging that hit-and-runs are common during pairwise accretion, is that even if velocities are too slow for projectiles to destroy their targets, targets frequently destroy projectiles, almost as often as they accrete them, as represented by $\xi_{\text{imp}} \leq -1/2$ in Figure 2.

Hit-and-run disruption has been invoked to explain how some iron meteorite parent bodies were disrupted under low-shock conditions (Yang et al. 2007) and, overall, for the diversity of meteorite progenitors (Asphaug 2017). At planet-forming scales, the loss of Mercury’s mantle can be explained by a violent hit-and-run by proto-Mercury into proto-Venus or proto-Earth (Asphaug & Reufer 2014; Chau et al. 2018), an event that would plot to the left-middle of Figure 2. Given the preservation of volatiles in Mercury (e.g., Peplowski et al. 2011) and the preponderance of lower-speed collisions, the preferred scenario is two nominal-velocity, nominal-impact-angle hit-and-runs in a row (Asphaug & Reufer 2014; Jackson et al. 2018)—a “stranded runner” as demonstrated in Paper II.

The *net* accretion efficiency of a giant impact increases if the runner or debris eventually get accreted by the target. The first study to consider the fate of post-Moon-formation debris, by Jackson & Wyatt (2012), modeled the dynamical evolution of escaping material obtained in a canonical simulation by Marcus et al. (2009), where a bit more than one lunar mass escaped

($\xi_{\text{esc}} \sim 0.13$). Of this, they estimated that 17% of the debris would end up on Venus and 20% on Earth within 10 Myr, apart from losses to collisional grinding. This would be sufficient to accrete a >10 km layer onto each planet.

In the case of a hit-and-run, the runner may reimpact the target; this too is not uncommon considering the proximity of their orbits. For energetic hit-and-runs with fast, disrupted runners, this can involve a complex chain of subsequent collisions, or no return at all. For lower-energy hit-and-runs, as expected in an accreting population and thus the subject of our study, the runner emerges relatively intact and escapes more slowly, leading to a higher return probability. These “hit-and-run return” collisions are a common mechanism for the late-stage growth of planets (Paper II).

2.2. Return to Middle-Earth

For nominal late-stage velocities $v_{\text{coll}} \approx 1.2 - 1.4v_{\text{esc}}$, most giant impacts are hit-and-run collisions (Figure 2) that produce relatively intact escaping runners, stripped of a fraction of their atmosphere, hydrosphere, crust, and mantle. Paper II showed that runners escaping from slow hit-and-runs with proto-Earth return to collide with post-hit-and-run proto-Earth (“middle-Earth”) about half the time, and that an equal fraction go on to collide with (and ultimately accrete with) proto-Venus. Venus loses far fewer of its runners—an asymmetry of formation that is the subject of that paper.

For runners that return to proto-Earth, the interlude between the hit-and-run and the return can be a few thousand years, or longer than 30 Myr. The average, of order 0.1–1 Myr, is much shorter than the error bars in geochronology, so a hit-and-run return origin of the Moon would likely appear geochemically as a single, complex event.

Even if proto-Earth is initially not rotating, the hit-and-run causes it to rotate, although it seldom makes a massive disk. The induced rotation period ($P_{\text{rot}} \sim 10$ – 11 hr in the simulations considered below) establishes a preimpact vector for the next collision. A key finding of Paper I is that, with the exception of the very early returns ($\lesssim 1000$ yr), there is no memory of the orientation from one collision to the next. The distribution of φ is indistinguishable from the expected random distribution $dp = \frac{1}{2} \sin \varphi d\varphi$, where φ is the offset angle between the two collisions. This is true whether or not other planets are included in the integration. Prograde and retrograde returns (parallel or antiparallel, $\varphi = 0^\circ$ or 180°) are least probable, while the mean and most probable returns are orthogonal, $\varphi = 90^\circ$, the sort that knock a planet on its side.

Because of the transfer of momentum to the target and the loss of kinetic energy to shocks and escaping debris, for nominal impact angles the runner’s egress velocity, escaping the postimpact target, is considerably slower than its inbound velocity v_{coll} (Emsenhuber et al. 2020). Another key finding of Paper I is that slower egress velocities lead to comparably slow return velocities. Given the high likelihood of hit-and-run collisions over the expected velocity range of giant impacts, the slowing down of runners is vital to late-stage planet formation because it leads to the low velocities that are required for accretion. Efficient accretion is not limited to the lower left blue region of Figure 2, but includes whatever collisions lead to runners that end up there.

3. Methods

Moon formation by hit-and-run return has three acts that we model independently: the hit-and-run, which slows down proto-Theia and causes initial mixing; the dynamical evolution of Theia's returning orbit; and the canonical-like graze-and-merge collision that forms the Moon.

There are many possibilities for the hit-and-run, and we focus on lower-velocity scenarios that were arguably most common. We sample existing hit-and-run simulations from Paper I that start with nonrotating differentiated planets 0.15 and $0.95 M_{\oplus}$ and choose two examples that lie marginally above the hit-and-run transition. For an impact angle of $\sim 45^{\circ}$ this corresponds to a velocity of 1.15 – $1.20 v_{\text{esc}}$. For these, we calculate the egress velocity of the runner escaping from the target, as well as their postimpact masses and the target rotation. The tendency in this velocity range is for the projectile to be accreted, or to escape relatively intact.

For the second act or interlude, we follow the procedure of Paper II and transfer the target and the escaping runner at their new velocity into the N -body code `mercury` (Chambers 2012), to track their dynamical evolution. We clone the outcome of each hit-and-run (middle-Earth + runner, ignoring debris) into 1000 random orientations and evolve each cloned set (including the major planets) until there is a follow-on collision, or for 20 Myr, at which time most giant impact chains at 1 au are complete. Longer evolutions up to 400 Myr are explored systematically in Paper II. From this we obtain the expected distribution of parameters (impact angles and velocities) that define candidate returning-runner scenarios for Moon formation.

The third act is the return collision, which we model using our updated SPH code for several cases within the distribution of return velocities and angles. As there are at least three major parameters to consider, even neglecting rotation and composition, times two collisions, it is not yet possible to explore all cases. For now we focus on the most probable ($\varphi \sim 90^{\circ}$, $\theta_{\text{coll}} \sim 45^{\circ}$) and fiducial cases, with the intention of hypothesis validation.

We now present the hydrocode methodology, along with the setup of the collisions and the postimpact analysis. We consider various geometries and velocities and include preimpact rotation of the target. We describe the computing of material exchange and the analysis of isotopic mixing in successive collisions. We evaluate the disk mass and the associated angular momentum and composition, applying two different methods in the literature for evaluating the postimpact disk.

3.1. Giant Impact Simulations

Our updated SPH scheme is fully described in Paper II; it is designed and extensively tested for modeling giant impacts (Reufer et al. 2012; Emsenhuber et al. 2018; Paper I). General reviews of SPH include Monaghan (1992) and Rosswog (2009).

SPH is a Lagrangian technique with material distributed into particles. A kernel interpolation is used to compute the quantities at any location, and spatial derivatives are computed using an interpolation with the derivatives of the kernel. To retrieve the neighbor particles for these derivatives, a hierarchical spatial tree is used (Barnes & Hut 1986) that also gives a rapid solution to the self-gravitational potential. In our code, density is retrieved using the kernel interpolation with a correction term for particles close to the surface (Reinhardt & Stadel 2017).

Entropy that arises as a result of shocks is evolved through artificial viscosity, and pressure is computed from the entropy and density using the (M-)ANEOS EOS (Thompson & Lauson 1972; Melosh 2007). For the terrestrial bodies in this set of papers, we use iron for the core and forsterite Mg_2SiO_4 (olivine) for the mantle. Friction is ignored for giant impacts of this scale (Emsenhuber et al. 2018), although it is likely to be important with regard to the structure of debris.

The initialization of pressure, density, and entropy inside each planet is performed in the same way as in Paper II. For spinning planets we also initialize a uniform rotation. First, we retrieve a 1D radial hydrostatic profile using the algorithm of Benz (1991) and the iron/forsterite EOSs. From the 1D profile, an initial SPH 3D spherical body is obtained using the methodology of Reinhardt & Stadel (2017). To establish preimpact rotation, each particle is given a radius-dependent velocity

$$\mathbf{v}_{\text{ini}} = \boldsymbol{\Omega} \times \mathbf{r}_{\text{ini}}, \quad \boldsymbol{\Omega} = \begin{pmatrix} 0 \\ 0 \\ \Omega \end{pmatrix}, \quad (5)$$

where $\Omega = 2\pi/P$ and P is the desired rotation period. The rotating body is then evolved by itself in SPH with a damping term that uses \mathbf{v}_{ini} as the reference velocity. This forces the body to spin at the prescribed uniform rate and allows the equatorial bulge to form, which changes not only the dynamics but also the collision cross section.

In all simulated collisions, to allow the tidal bulge to form, and for torque to accumulate before the impact, the bodies begin their evolution at a distance of 5 times the sum of the radii, about twice the Roche limit.

3.2. Return Collision Setup and Rotations

Bodies from the hit-and-runs are evolved dynamically using the `mercury` code (Chambers 2012) for 1000 clones representing the velocity $\sim 1.01 v_{\text{esc}}$ in a distribution of directions. As we shall see, the return collision ends up being $v_{\text{coll}} \lesssim 1.05 v_{\text{esc}}$ in most cases, so we model return collisions as either 1.00 or $1.05 v_{\text{esc}}$, fixing the impact angle at $\sim 45^{\circ}$ in each case.

The targets must be set up with a rotation period $P_{\text{rot}} = 10$ – 11 hr, the result of the hit-and-runs described. This requires specification of the orientation of the return collision (Figure 3). The alignment angle φ is between the rotation axis of the target and the orbital angular momentum vector in the two-body frame of the collision, while the offset angle ω is between the periape (as if there were no collision) and target spin vector. For now we apply zero rotation to the reimpacting runner, although barely escaping runners tend to rotate rapidly (Asphaug et al. 2006), which might be important to disk composition.

For random return collisions it was shown in Paper I that the offset angle ω is uniformly distributed, while φ follows the probability distribution $dp = \frac{1}{2} \sin \varphi d\varphi$ with a maximum at $\varphi = 90^{\circ}$. For the limited set of studies presented here, we therefore focus on impact angles of around 45° and alignment angles of around 90° , with other alignment angles (prograde, retrograde, and nonrotating) for comparison. As with the first collision, we start the bodies at a distance of 5 times the sum of the radii.

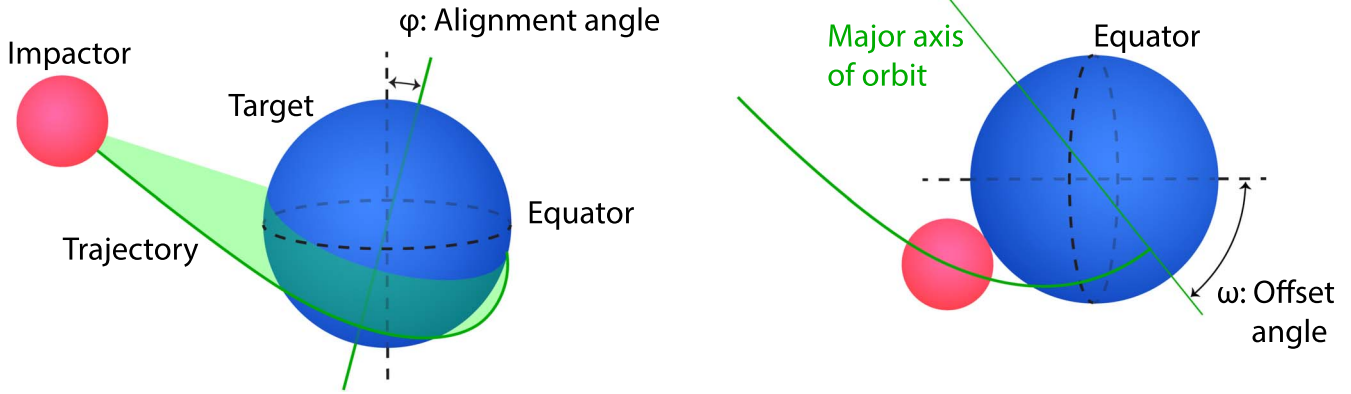


Figure 3. A rotating target has two additional parameters: the alignment angle φ between the target equator and the plane of the collision in the two-body frame (left), and the offset angle ω between the orbit periapease and the spin vector (right). φ ranges from prograde (0°) to retrograde (180°), the most probable being a perpendicular (90°) collision. The offset angle (right) is shown for the case $\varphi = 90^\circ$ and an impact pericenter at the pole ($\omega = 0^\circ$). The angular momentum offset φ matters most.

3.3. Determination of the Protolunar Disk

In our proposed scenario the protolunar disk is a consequence of the terminal collision, when the runner’s mass and angular momentum are finally accreted. A slow hit-and-run usually does not produce a massive disk (Paper II), but it has three primary effects: it spins up the target ahead of the next collision, it exchanges an initial mass fraction between the colliding bodies (middle panel of Figure 1), and it causes significant heating of the target that might influence the terminal collision.

After the giant impact has evolved in SPH to the point that there are no further major collisional interactions, we evaluate the disk. (By “disk” we mean any postimpact orbiting material, despite the possibility of the mass being dominated by huge clumps.) We first identify all particles that are gravitationally bound to the single largest remnant; escaping particles are removed from the analysis. The surface of the largest remnant is then computed as the layer whose density is approximately three-quarters of the reference density of the EOS. The nonescaping particles outside that surface are potential disk particles that are then evaluated according to two approaches.

The first approach (Reufer et al. 2012, as adapted by Emsenhuber & Asphaug 2019a) is to calculate the pericenter of the two-body problem for each potential disk particle about the central body. If the pericenter is below the identified surface, the particle is added to Earth; otherwise, it is added to the disk. The second approach (e.g., Canup et al. 2001; Canup 2004) is to calculate the radius of a circular orbit with the equivalent angular momentum to each potential disk particle and then proceed similarly, adding the particle to the disk if this equivalent circular orbit is outside the identified surface. The latter is more inclusive, as all particles having their pericenter above the surface also have their equivalent radius above the surface. We include both approaches in our analysis to compare with published results for postimpact disks.

For the case of a fluid-particle postimpact disk, the final coagulation to form the Moon is at best $\sim 50\%$ efficient according to Salmon & Canup (2012), consistent with the results of Ida et al. (1997). If such disk approximations are correct, then to be successful a giant impact needs to place at least 2 lunar masses into orbit, although this depends on disk angular momentum (Kokubo et al. 2000b).

3.4. Material Exchange in Two Collisions

Two collisions in a row provide two stages of material exchange, as illustrated in Figure 1. The fractions of proto-Theia and proto-Earth in the final Earth’s mantle and the protolunar disk are computed by convolving the material exchanges/contributions calculated in both collisions individually. We assume that complete mixing occurs within each silicate reservoir during the interlude and that the bodies are fully differentiated. Core material is not exchanged between colliding bodies in these relatively low energy events. We have not attempted to evaluate any metal-silicate equilibration that may occur inside each body in response to either collision. While equilibration could have significant implications, for example, in the resetting of ^{182}W (see Dwyer et al. 2015), such a study is outside the scope of our analysis, which focuses on silicate compositions.

The first stage of silicate mass exchange is defined by the equilibration factor (Reufer et al. 2012)

$$\delta f_T^l = \frac{f_{\text{mant}}^{s \leftarrow t}}{f_{\text{mant}}^{l \leftarrow t}} - 1, \quad (6)$$

where t and i stand for the original target (proto-Earth) and impactor (proto-Theia) of the hit-and-run, respectively, and l and s stand for the largest remnant (middle-Earth) and second-largest remnant (Theia, the runner) of that collision, respectively. Thus, $f_{\text{mant}}^{s \leftarrow t}$ is the mantle mass fraction of Theia that comes from proto-Earth, $f_{\text{mant}}^{l \leftarrow t} = 1 - f_{\text{mant}}^{l \leftarrow i}$ is the mantle mass fraction of middle-Earth that comes from proto-Earth, and $f_{\text{mant}}^{l \leftarrow i}$ is the mantle mass fraction of middle-Earth that comes from the impactor. These mixing values are given in Table 1 of Paper II for a range of giant impacts, from which we select our starting hit-and-run collisions.

The combined silicate equilibration factor is computed as the product of both collisions,

$$\delta f_T^c = \frac{f_{\text{mant}}^{d \leftarrow t}}{f_{\text{mant}}^{e \leftarrow t}} - 1, \quad (7)$$

where superscript “ $d \leftarrow t$ ” indicates the fraction of the protolunar disk that originates from the proto-Earth, and superscript “ $e \leftarrow t$ ” indicates the fraction of Earth that does.

These give the cumulative relative contributions in the mantles of the Moon and Earth. Under the assumption of complete mixing within each silicate reservoir during the interlude, we can then write

$$f_{\text{mant}}^{\text{d} \leftarrow \text{t}} = f_{\text{mant}}^{\text{d} \leftarrow \text{l}} f_{\text{mant}}^{\text{l} \leftarrow \text{t}} + (1 - f_{\text{mant}}^{\text{d} \leftarrow \text{l}}) f_{\text{mant}}^{\text{s} \leftarrow \text{t}} \quad \text{and} \quad (8)$$

$$f_{\text{mant}}^{\text{e} \leftarrow \text{t}} = f_{\text{mant}}^{\text{e} \leftarrow \text{l}} f_{\text{mant}}^{\text{l} \leftarrow \text{t}} + (1 - f_{\text{mant}}^{\text{e} \leftarrow \text{l}}) f_{\text{mant}}^{\text{s} \leftarrow \text{t}}. \quad (9)$$

Putting everything together results in a final combined mantle equilibration ratio

$$\delta f_{\text{T}}^{\text{c}} = \frac{1 + \delta f_{\text{T}}^{\text{l}} (1 - f_{\text{mant}}^{\text{d} \leftarrow \text{l}})}{1 + \delta f_{\text{T}}^{\text{l}} (1 - f_{\text{mant}}^{\text{e} \leftarrow \text{l}})} - 1. \quad (10)$$

If $f_{\text{T}}^{\text{l}} = -1$, that is, no mixing in the hit-and-run, then $\delta f_{\text{T}}^{\text{c}}$ is equivalent to

$$\delta f_{\text{T}}^{\text{c}2} = \frac{f_{\text{mant}}^{\text{d} \leftarrow \text{l}}}{f_{\text{mant}}^{\text{e} \leftarrow \text{l}}} - 1, \quad (11)$$

which, as expected, is the equilibration of the second collision alone. A further intuitive understanding of Equation (10) is obtained by assuming $f_{\text{mant}}^{\text{e} \leftarrow \text{l}} = 1$ and using Equation (11) to replace $f_{\text{mant}}^{\text{d} \leftarrow \text{l}}$ so that all of middle-Earth's mantle is made only of proto-Earth mantle. Then,

$$\delta f_{\text{T}}^{\text{c}} \approx -\delta f_{\text{T}}^{\text{l}} \delta f_{\text{T}}^{\text{c}2}. \quad (12)$$

Across our simulations, we obtain $f_{\text{mant}}^{\text{e} \leftarrow \text{l}} \approx 0.96$, so this assumption results in reasonable values for the estimation of $\delta f_{\text{T}}^{\text{c}}$.

4. Results

4.1. Constraining the First Collision

The first giant impact is defined by our hypothesis to be a hit-and-run that meets three basic requirements: the runner is approximately $0.1 M_{\oplus}$ per the canonical model; the runner retains a massive silicate mantle, i.e., is more or less chondritic; and the egress velocity is slow so that it is likely to be followed by a low-velocity return. This implies a proto-Theia with a mass of order $0.15 M_{\oplus}$ and a hit-and-run collision velocity $v_{\text{coll}} \lesssim 1.2 v_{\text{esc}}$, depending on angle.

We do not consider impacts that are close to head-on, in this velocity range, because these end up being accretions that are incompatible with Moon formation, having insufficient angular momentum. As for highly grazing initial impacts, these have much less direct mass intersection, so the runner is not sufficiently slowed down and emerges largely intact, although not without significant geophysical transformations (Asphaug et al. 2006). From the point of view of Moon formation a highly grazing collision may serve primarily as a dynamical scattering event, without significant mass exchange or deceleration, so we do not consider these explicitly either in our current study.

We have also not explicitly considered giant impacts faster than $1.2 v_{\text{esc}}$. In close to head-on cases these can also be accretions, though not efficient, and incompatible with Moon formation. At higher velocity they can be disruptive to the projectile and erosive to the target. At more typical impact angles ($\sim 30^{\circ}$ – 60°) these faster impacts are hit-and-runs causing significant mass loss from the projectile ($\xi_{\text{imp}} \ll 0$), or projectile disruption, and erosion of the target. We have not

Table 1

The Hit-and-run Slows Down Proto-Theia from a Starting Value 1.15 – $1.20 v_{\text{esc}}$ to a Runner Egress Velocity $\sim 1.01 v_{\text{esc}}$ (Figure 4)

$\frac{v_{\text{coll}}}{v_{\text{esc}}}$	θ_{coll} (deg)	m_{lr} (M_{\oplus})	m_{sr} (M_{\oplus})	P_{lr} (hr)	Z_{lr}	Z_{sr}	$\delta f_{\text{T}}^{\text{l}}$
1.15	47.5	0.94	0.11	11	30%	32%	−83.6%
1.20	43.0	0.95	0.081	10	30%	36%	−75.4%

Note. It removes some mantle from proto-Theia, reducing its mass and increasing its iron fraction (Z_{sr}). It does not make a disk, but spins up the largest remnant, that is, P_{lr} which is important to the follow-on collision.

attempted to model the latter, simply because the egress velocity from a fast hit-and-run is usually also fast, so that the return collision is not as likely. When it occurs, it may itself be a hit-and-run, if not a head-on merger. A faster hit-and-run could therefore be the start of a longer chain, explored in our previous papers, in which case our scenario could be the end of a series of collisions, in a system more laden with debris.

Focusing on impact angles $\sim 45^{\circ}$, target masses $0.9 M_{\oplus}$, impactor masses $0.15 M_{\oplus}$ ($\gamma = 1/6$), and chondritic compositions, two cases from Paper II are just above the hit-and-run threshold, one at $1.15 v_{\text{esc}}$ at 47.5° , and another at $1.20 v_{\text{esc}}$ collision at 43° . The parameters and outcomes of these collisions are given in Table 1 and are indicated by the ‘‘Prior’’ column in Table 2. The slower, more grazing hit-and-run produces a somewhat more massive runner that is ultimately more successful at protolunar disk production. A larger runner could of course be obtained in the faster case by starting with a larger proto-Theia; these two specified cases are intended as examples. The core mass fraction Z_{sr} of the runner following each collision is somewhat greater than the starting value $Z_{\text{imp}} = 30\%$, due to mantle stripping, which is more significant in the faster, somewhat more head-on collision.

All simulations start with nonrotating proto-Earths and proto-Theias. The largest remnant (middle-Earth, subscript ‘‘I’’ in Equation (11)) ends with a rotation period ~ 11 hr following the slower collision and ~ 10 hr after the faster, more head-on collision. According to Canup (2008), this is where preimpact spin begins to play an important role in giant impacts, so we include it in setting up the return collision.

The deflection angles in these hit-and-runs, relative to the incoming vector, are 52° and 57° , measured between the approach vector and the egress vector, where the larger deflection is for the egress that just barely escapes, the faster impact. If the approach angle is uniformly distributed in the center-of-mass frame of the collision, then the deflection angle does not play an important role in this analysis.

So the four principal outcomes of the hit-and-run, in order of significance to Moon formation, are deceleration of the runner, spin-up of the target, heating of the target, and mantle (silicate) loss from the projectile. A slow hit-and-run usually does not produce a massive disk.

4.2. Dynamical Evolution of the Runner

A slow egress velocity from a hit-and-run collision will usually lead to a correspondingly slow return collision, which is likely to be a graze-and-merge collision for most impact angles, thus compatible with Moon formation. However, this depends on the dynamical perturbations the runner has along the way. The longer the interlude between the collisions, the

Table 2
Results from Our SPH Simulations of the Terminal Accretion Collisions

Prior	$\frac{v_{\text{coll}}}{v_{\text{esc}}}$	θ_{coll} (deg)	φ (deg)	ω (deg)	m_{esc} (M_{\oplus})	L_{tot} (L_{EM})	Pericenter above Surface						Equivalent Radius above Surface					
							m_{disk} (M_{\oplus})	L_{disk} (L_{EM})	$\frac{m_{\text{Fe}}}{m_{\text{disk}}}$	ϕ_{disk} (deg)	δf_{T}^2	$\delta f_{\text{T}}^{\text{c}}$	m_{disk} (M_{\oplus})	L_{disk} (L_{EM})	$\frac{m_{\text{Fe}}}{m_{\text{disk}}}$	ϕ_{disk} (deg)	δf_{T}^2	$\delta f_{\text{T}}^{\text{c}}$
c. case*	1.00	45	0.175	0.950	0.887	0.168	1.0%	0.3	-58.5%	...	0.964	0.173	3.6%	0.3	-57.2%	...
1.15*	1.00	45	0.079	1.100	1.120	0.224	1.6%	0.3	-56.8%	-47.2%	1.218	0.229	4.1%	0.2	-56.1%	-46.1%
1.15	1.00	45	0	...	0.113	1.430	1.433	0.295	1.4%	0.3	-62.4%	-51.8%	1.547	0.298	3.6%	0.4	-62.2%	-51.2%
1.15	1.00	45	180	...	0.160	0.744	0.745	0.145	0.5%	0.9	-56.1%	-46.6%	1.003	0.154	7.9%	0.9	-57.4%	-47.2%
1.15	1.00	45	90	0	0.081	1.152	1.101	0.219	0.7%	17.5	-58.4%	-48.5%	1.224	0.224	4.8%	17.5	-57.9%	-47.6%
1.15	1.00	45	90	90	0.194	1.123	1.159	0.225	2.1%	20.1	-57.0%	-47.4%	1.307	0.230	8.2%	20.0	-56.2%	-46.3%
1.15	1.00	48	90	0	0.062	1.208	1.403	0.331	0.1%	19.8	-73.1%	-60.8%	2.131	0.338	27.8%	19.6	-73.8%	-60.8%
1.15	1.05	45	90	0	0.994	1.049	0.834	0.173	2.3%	19.4	-52.0%	-43.2%	1.569	0.178	42.4%	18.9	-52.3%	-43.1%
1.15	1.05	48	90	0	0.356	1.214	1.283	0.308	0.1%	16.6	-75.5%	-62.9%	2.870	0.325	44.0%	15.7	-77.5%	-63.9%
1.20	1.00	45	0	...	0.092	1.172	0.812	0.146	1.4%	1.6	-67.3%	-50.5%	0.858	0.148	1.8%	1.6	-66.3%	-49.2%
1.20	1.00	45	90	0	0.506	0.828	0.292	0.054	0.1%	23.2	-50.5%	-37.9%	0.352	0.055	6.8%	22.9	-50.0%	-37.7%

Note. The first five columns are the initial conditions, with “prior” defining the body properties from the corresponding entries in Table 1. The first row is our simulation of the canonical case ($0.9 M_{\oplus}$ target, $0.1 M_{\oplus}$ projectile, 30% core mass fraction), and the asterisk is for nonrotating targets. The total escaping mass m_{esc} and total angular momentum L_{tot} are shown, and then the disk properties computed using both approaches (see text): disk mass m_{disk} , disk angular momentum L_{disk} , disk iron mass fraction $\frac{m_{\text{Fe}}}{m_{\text{disk}}}$, disk inclination ϕ_{disk} , and final disk–planet compositional imbalance $\delta f_{\text{T}}^{\text{c}}$.

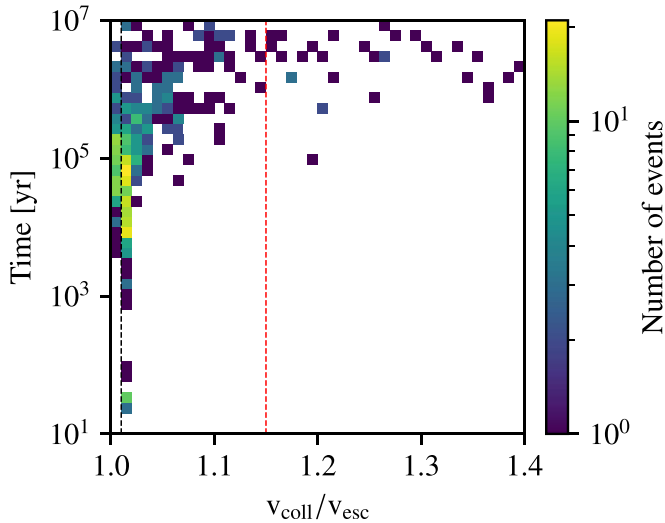


Figure 4. Two-dimensional histogram of impact velocity (normalized to v_{esc}) vs. time between the collisions, for hit-and-run returns with proto-Earth at 1 au (from Table 1 in Paper II). Dynamics are computed including the major solar system planets starting on their current orbits. Plotted is the slower hit-and-run, $v_{\text{coll}}/v_{\text{esc}} = 1.15$ (red dashed line), for parameters $\theta = 47^\circ 5$, $m_{\text{imp}} = 0.15 M_{\oplus}$, and $m_{\text{tar}} = 0.90 M_{\oplus}$. The escaping runner velocity is $v_{\text{run}}/v_{\text{esc}} = 1.01$ (black dashed line).

longer the opportunity for encounters with proto-Earth, proto-Venus, or other bodies that would energize the subsequent encounters.

This tendency is shown in Figure 4, which summarizes the dynamical evolution of 1000 clones of middle-Earth and Theia (remnants of each hit-and-run, integrated with the present major planets) for egress velocities $1.01v_{\text{esc}}$ in the center-of-mass frame starting at 1 au. Plotted is a 2D histogram of the return velocity ($v_{\text{imp}}/v_{\text{esc}}$) versus number of years between the two giant impacts, for the distribution of clones launched in random directions. The hit-and-run velocity $v_{\text{imp}}/v_{\text{esc}} = 1.15$ is shown by the red dashed line, and the egress velocity $v_{\text{run}}/v_{\text{esc}} = 1.01$ is shown by the black dashed line. Histogram values are logarithmic, with yellow indicating the majority of returns.

Return collisions happen over a range of times, mostly within $\sim 10^4$ – 10^6 yr. Early returns are close to the egress velocity, which is close to v_{esc} ; later returns happen at a wider range of velocities owing to gravitational encounters. Non-impacting runners are not plotted in Figure 4 but are studied in Paper II; these either move on to another planet (usually Venus) or become part of the background population.

4.3. Final Accretion and Moon Formation

We simulate the return collisions using the SPH methods described in Section 3 and Paper II. It is not feasible to span the possibilities of mass ratio, velocity, core size, collision angle, and rotation vector, so we limit our current analysis to 10 cases that seem representative of the predicted terminal collisions, based on largest remnants from two of the hit-and-run collisions in Paper II, plus a simulation of the canonical case for comparison.

We use the largest remnants of the hit-and-runs in Table 1 to prescribe targets of the appropriate mass and rotation period and projectiles (runners) of the appropriate mass. The target of each return collision is set at $\sim 0.95 M_{\oplus}$ in each case. The projectile is $0.11 M_{\oplus}$ in the slower case and $0.08 M_{\oplus}$ in the

faster case. The alignment angle (Figure 3) ranges from prograde ($\varphi = 0^\circ$) to retrograde ($\varphi = 180^\circ$), the most likely being 90° , so we focus on orthogonal collisions. We also vary the offset angle ω , although this is less significant. The same composition and entropy profiles are applied as in the original collision, and target rotations are applied as described in Section 3. As discussed further below, a much hotter middle-Earth may lead to a more massive, more Earth-composition disk (Hosono et al. 2019), but we do not model this possibility.

Each collision in Table 2 was evolved to at least 48 hr after contact, in order for graze-and-merge collisions to complete and for the reported disk quantities to converge. In two cases at slightly more grazing incidence (48° , discussed further below) the integration time was doubled to follow their extended graze-and-merge orbits. One of the challenges of giant impact simulations in the accretion regime is that many scenarios can require days to evolve the point that an outcome (merger versus hit-and-run) is determined. This can require significant machine time, plus greater long-range dynamical accuracy than SPH provides. It also makes it difficult to compare outcomes for disk mass and angular momentum, when disk states are evaluated at different times.

Results are given in Table 2 for the total escaping mass not bound to the final system, the total bound angular momentum, and the mass, angular momentum, inclination, and composition of the disk. Two sets of results are given corresponding to the two approaches (see Section 3) for determining what particles belong to the disk at the end of a simulation. Overall, for a number of our modeled return collisions the final properties are comparable to those obtained in canonical simulations, with an improvement in most cases, with isotopic mixing improving by 10%–15%. We have not performed any tuning to obtain a favorable result, other than to consider hit-and-runs that end with a low-velocity runner and a return impact angle close to the expected value.

4.3.1. Protolunar Disk Mass, Angular Momentum, and Composition

SPH simulations of giant impacts, and especially their disks, become inaccurate after tens of hours for the reasons described above, yet the postimpact state continues to evolve dynamically. The disk quantities must therefore be estimated, which is done by treating each nonescaping SPH particle as a test body that orbits the largest remnant and applying a criterion, as was described in Section 3. The more conservative criterion (e.g., Kokubo et al. 2000b; Reufer et al. 2012) is that the pericenter of each particle’s orbit must be outside the computed target surface (“pericenter above surface” in Table 2). The other approach (e.g., Canup et al. 2001; Canup 2004) is to convert each SPH particle’s orbit into a circular orbit with the equivalent angular momentum. Here we define the equivalent radius as a circular orbit of semimajor axis $a(1 - e^2)$, where e is the eccentricity of the particle orbit. If this circular orbit is outside the target, the particle is added to the disk (“equivalent radius above surface” in Table 2).

The mass and composition of the postimpact orbiting material (not always disk-like) are found in simulations (e.g., Canup et al. 2013) to evolve rapidly owing to the reimpact of massive projectile-remnant clumps. This is apparent when comparing some of the pericenter and equivalent-radius disk evaluations applied to our simulations. Three of the cases end up with several- M_{\oplus} disks in the equivalent-radius evaluation (1.6 – $2.9 M_{\oplus}$), but their iron mass fractions are large, 28%–44%,

respectively. In the pericenter evaluation of the same SPH outcomes, these cases end up with less than 2% iron in the disk, and at most $1.3 M_{\oplus}$ disk mass. This is because iron-rich clumps come crashing down in the stricter pericenter evaluation.

Using the pericenter evaluation, we find that hit-and-run return collisions can lead to silicate disks significantly greater than one lunar mass, in a final system with the right L_{EM} , and with a low iron mass fraction. The angular momentum in the postimpact disk, expressed as L_{disk}/L_{EM} in Table 2, equals or exceeds the proposed minimum value of ~ 0.18 (e.g., Canup et al. 2001) in most of the simulated cases. The disk mass and disk angular momentum are both greater, in most of our return collisions, than in our simulation of the canonical model.

When using the equivalent-radius evaluation for these same simulations, the disk masses are substantially greater, although the disk angular momentum only slightly increases. This is because this criterion includes higher-eccentricity particles that are rejected by the pericenter evaluation; the equivalent radius of these particles is only slightly above the surface. Using the scaling of Kokubo et al. (2000a), a disk with a small specific angular momentum is unable to coalesce into a large satellite; material with a specific angular momentum lower than that of an equivalent orbit at $\approx 0.6 R_{\text{Roche}} = 1.7 R_{\oplus}$ does not contribute to the final satellite. Disk mass estimates obtained with the equivalent-radius approach are thus not representative of the final mass of the satellite.

When comparing equivalent methods of evaluating the postimpact disk, we see that hit-and-run return collisions can lead to solutions comparable to, or improving upon, simulations of the canonical case. This is also true for disk isotopic composition. The final protolunar disk is composed of proto-Theia and proto-Earth materials, with proportions obtained over the course of two collisions using the conventions of Reufer et al. (2012) and Paper II (Equation (10)). This is compared to the mantle composition of the final target, also derived from both components in two collisions, to compute the final difference, Moon versus Earth composition, f_T^c in Table 2, which would be 0 for equally derived composition and -100% for lunar silicates derived entirely from proto-Theia. Compositional similarity is improved by $\sim 10\%$ – 15% compared to the canonical case, depending on the use of the pericenter or equivalent-radius criterion for which particles end up in the disk. The equivalent-radius approach also tends to leave more of Theia’s iron in the disk.

The improvement in isotopic equilibration, although significant, is not sufficient to erase any major differences in proto-Earth/proto-Theia composition. As discussed above, their similarity might not have to be so exact (e.g., Cano et al. 2020). But we consider this aspect of our results further, noting that our computed f_T^c is likely to underestimate the final homogenization for several reasons.

For one thing, SPH tends to suppress mixing at interacting boundaries (e.g., Deng et al. 2019). This might artificially limit the material exchange in the hit-and-run collision when the planets shear past and through one another. Hit-and-run is not a “bounce” but an intense thermophysical and dynamical contact lasting about half an hour. Suppression of mixing might also limit the entrainment of Earth materials into the disk, which means that canonical scenarios might improve as well, if mixing is greater than computed. But the effect would be of compounded significance in collision-chain scenarios.

For another, Theia and middle-Earth are initialized, in our setup for the terminal collision, with the same subsolidus entropy profile as proto-Theia and proto-Earth in the first collision (see Section 3). Although hit-and-runs do not impart the same degree of heating as accretions, for the reasons noted above, they are nevertheless violently energetic and can produce a global magma ocean (Nakajima et al. 2021). If so, and the second collision happened before it cooled, then according to Hosono et al. (2019) this would enhance protolunar disk production and increase the fractional contribution from proto-Earth.

Furthermore, our simplified representation of hit-and-run as two bodies coming in and two bodies going out, like a bounce, ignores the fact that a collision chain generates substantial heliocentric debris. Much of this will return within a few million years (i.e., Jackson & Wyatt 2012) while the Moon is in close orbit around the Earth, potentially leading to ballistic material exchanges and other modifications. Debris in a hit-and-run can be tens of percent of the total colliding mass, depending on starting velocity (Figure 2), but for slow collision chains debris production is dominated by the final merger (m_{esc} in Table 2). Faster, more debris-laden giant impacts are relatively common in terrestrial planet formation but have yet to be modeled in the context of our proposed scenario.

4.3.2. Sensitivity to Parameters

There is considerable sensitivity to the return impact angle and velocity, even within our very limited range of collisional parameters. Four of the modeled return collisions are set up with $\varphi = 90^\circ$ and $P_{\text{rot}} = 11$ hr and the same colliding bodies, only varying the impact angle, $\theta_{\text{coll}} = 45^\circ$ or 48° , and velocity, $v_{\text{coll}} = 1.00$ or $1.05 v_{\text{esc}}$. The $\theta_{\text{coll}} = 45^\circ$ collisions are graze-and-merge events ending in accretion, resembling the canonical case but with pronounced disk inclination. The $\theta_{\text{coll}} = 48^\circ$ collisions are also graze-and-merge, but the bound runner misses the target when it loops back, passing inside the threshold for tidal mass loss (Sridhar & Tremaine 1992) so that the runner’s energy is further reduced.

In the $1.00 v_{\text{esc}}$ case at 48° , this leads to subsequent accretion and the spin-out of a massive disk. For the $1.05 v_{\text{esc}}$ case we end the simulation after $t = 96$ hr, by which time dissipation has begun to circularize the runner’s orbit with a period of ~ 4 hr around the target. It is unclear whether the bodies will merge, as in the previous case, or end in a direct capture reminiscent of early models of lunar formation (Benz et al. 1986) and scenarios for Pluto–Charon formation (Canup 2011). Because the runner orbits just inside the corotation radius of Earth, it is expected to eventually collide, and could then make a delayed protolunar disk.

As for prograde collisions ($\varphi = 0^\circ$) into spinning targets, these are found to produce more massive disks, and greater system angular momentum overall, but at the expense of resulting in somewhat less mixed isotopic compositions (e.g., $L_{\text{tot}} = 1.433 L_{EM}$ and $\delta f_T^c = -51.8\%$ for the case with prior $v_{\text{coll}}/v_{\text{esc}} = 1.15$ and $\varphi = 0^\circ$).

Retrograde collisions lead to insufficient disk mass and lower combined angular momentum in our study. For example, $m_{\text{disk}} = 0.745 M_{\oplus}$ and $L_{\text{tot}} = 0.744 L_{EM}$ for the case with prior $v_{\text{coll}}/v_{\text{esc}} = 1.15$ and $\varphi = 180^\circ$. Moon formation might still be possible in retrograde collisions, but with a more massive, higher-velocity projectile providing larger orbital angular momentum to compensate for the opposing spin.

For the retrograde collisions we have studied, the angular momentum acquired is sufficient to reverse the rotation of the counterrotating proto-Earth, so the disk (although less than a lunar mass in these cases) and target end up rotating in the same direction. Earth’s rotation ends up being slower by about a factor of two than for the prograde but otherwise-identical collision, 8.9 hr versus 4.5 hr. We do not recover the trend identified by Canup (2008), where the escaping mass m_{esc} is larger for retrograde than for prograde impacts.

4.3.3. Target Rotation and Disk Inclination

To first order, we find that protolunar disks produced in orthogonal (that is, expected) cases are similar to nonrotating canonical cases: silicate disks of $\sim 1\text{--}2$ lunar masses with very little iron. But there are two important differences: impacts into rotating targets liberate more material into heliocentric orbit, likely due to the lower preimpact binding energy of the rotating target, and the disk ends up highly inclined to the rotating postimpact Earth in the case of an orthogonal collision.

In prograde collisions we find that the disk ends up aligned within about 1° to the equator of Earth. But in orthogonal collisions the postimpact disks end up inclined by up to $\sim 20^\circ$ in the cases we have modeled. This is because the disk is formed almost entirely in the return collision and thus retains Theia’s angular momentum, while Earth’s final spin is the combination of the hit-and-run and the terminal accretion.

A strongly inclined protolunar disk is an intriguing outcome of our model, given the Moon’s pronounced and unexplained orbital inclination with respect to Earth’s equator (Touma & Wisdom 1998). But the connection is not straightforward. A gas-particle disk would damp to the equator; in this case the disk angular momentum reported in Table 2 would be multiplied by the cosine of the inclination to obtain the parallel component, although this is close to 1 in all our cases. For an inclined postimpact state to result in an inclined Moon, rapid coagulation is required, or the presence of nearly lunar-mass postimpact clumps to begin with. Afterward the inclination evolution would depend on the tidal dissipation that drove the Moon’s migration (e.g., Chen & Nimmo 2016).

4.3.4. Remnants and Debris

Most material remains gravitationally bound to one of the two largest remnants, for the relatively gentle hit-and-runs considered here, with some exchange across the collisional interface. The total debris mass m_{esc} not bound to either the target or the runner is $\lesssim 10^{-3} M_{\oplus}$ (Paper II), less than 1/10 of a lunar mass, in these cases. Although the total energy is lower, more debris is produced in escape-velocity mergers than in slow hit-and-runs. Higher-velocity hit-and-runs and accretions are increasingly dispersive and debris producing, as seen in the right panel of Figure 2.

The structure and size distribution of giant impact debris are uncertain, for reasons similar to why we do not yet know, with any confidence, the structure of the postimpact disk. Like the disk, the escaping debris is an energetic minor fraction of the total colliding mass, whose production is sensitive to geometry, velocity, temperature, EOS, and numerical treatment. Even if the physics and dynamics of debris production are accurately modeled, resolving m_{esc} explicitly requires quite high numerical resolution. For Earth-mass colliding systems simulated with 500,000 SPH particles, as in Paper II, clumps smaller than

several 100 km diameter are not well resolved (a few smoothing lengths). Extensive structures like plumes and spiral arms can be resolved, but not their components.

The total debris mass produced in these hit-and-runs is much smaller than the runner mass, m_{sr} , and is therefore ignored for now in the major bodies’ dynamical evolution. It would exert some dynamical friction, increasing the tendency toward final accretion. Geologically the production and size distribution of debris are much more significant. One-tenth of a lunar mass is twice the mass of the current main belt, fluxing through the Earth–Moon system on a timescale that overlaps the early solidification and tidal migration timescale of the Moon. It therefore makes a distinguishable difference to early lunar geology whether the returning debris of Moon formation took the form of millions of 10 km cratering projectiles, or hundreds of basin-forming bodies, or tens of 1000 km bodies, or dust. For example, Perera et al. (2018) showed that if the returning projectiles were large enough to punch through the insulating crust while the Moon was solidifying, the effect would be to greatly accelerate early lunar cooling. Conversely, as noted, one 1000 km impactor could form a remelted hemispheric magma ocean after the Moon had largely solidified.

Returning collisions would most likely happen before the Moon had tidally evolved far away from Earth. Proximity during the bombardment could lead to further mass exchanges and mixing, in addition to the diffusion and equilibration processes described above. Heliocentric impact velocities at the Moon are greater than the escape velocity from Earth at lunar orbit. Earth escape velocity at R_{Roche} is $\sim 6 \text{ km s}^{-1}$, almost 3 times the lunar surface escape velocity, so in the case of slow tidal evolution the return bombardment could lead to substantial erosion of the Moon.

Despite its potential significance to the problem of Moon formation, we have not attempted to track the evolution and reimpact of debris in our simulations because it remains a poorly constrained aspect of the problem. Debris production depends sensitively on the details of the collision chain, we are unable to numerically resolve the debris sizes in any case, and including debris adds new dimensions to an already wide-ranging parameter space.

5. Conclusions

Moon formation was the final major episode of Earth’s accretion, a calamity at the end of the late stage when planetary embryos strayed from their respective feeding zones and collided. Terrestrial planets may have grown pairwise in a series of violent mergers, building up Venus and Earth in perhaps dozens of events, leaving the smaller planets as unaccreted remnants or survivors.

But pairwise accretion does not generally proceed through effective mergers; far from it, hit-and-run collisions are expected to happen about half the time. This means that pairwise accretion proceeds about half the time through hit-and-run return, the process of sequential giant impacts.

5.1. Pathways to the Canonical Model

The goal of this paper is to demonstrate that hit-and-run collisions often lead to terminal mergers that can form a Moon-sized silicate satellite. It builds on prior work. Paper I tracked the dynamical fate of “runners” after hit-and-run collisions at 1 au, focusing on velocities $v_{\text{coll}} \lesssim 1.2v_{\text{esc}}$ that were common in

the late stage. A barely escaping runner was shown to recollide with proto-Earth in about $\sim 0.1\text{--}1$ Myr, in most cases, with a return velocity close to v_{esc} , a terminal merger that is the basis for our hypothesis.

Paper II focused on Earth and Venus, examining the demographics of collision chains. It treated hit-and-run collisions explicitly by applying the surrogate model of Cambioni et al. (2019) and Emsenhuber et al. (2020). This enabled N -body computations to explicitly track the dynamics of targets and runners through sequential collisions. For intermediate-velocity hit-and-runs, with the other major planets included, it was shown that proto-Earth loses a significant fraction of its runners, roughly half, compared to Venus.

Paper II also showed that faster runners, from faster giant impacts into proto-Earth, are most likely to be ultimately accreted by Venus. Also, longer, more complex, more debris-producing chains are relatively common, as are stranded runners. It was seen that proto-Earth could serve as a “vanguard” during late-stage accretion, slowing down interloping major bodies by colliding with them, and delivering their remnants mostly to Venus. Earth, having no such vanguard, would end up comparatively depleted in outer solar system materials.

In an epoch where accretion is occurring overall, $v_{\text{coll}} \lesssim 1.2v_{\text{esc}}$, about half of giant impacts are hit-and-run. Most of those have barely escaping runners that are likely to come back for a terminal merger, hence our proposed scenario where “proto-Theia” (mass $\gtrsim m_{\text{imp}} \gtrsim 0.15 M_{\oplus}$) collides with proto-Earth at $v_{\text{coll}} \sim 1.15 - 1.20v_{\text{esc}}$, where for now we have considered only the most likely impact angle $\theta_{\text{coll}} \sim 45^\circ$. We evaluate the material exchange in this hit-and-run collision and obtain the target’s final spin rate and the runner’s mass and egress velocity. The target and runner evolve dynamically until there is a second collision, which can resemble the canonical scenario.

A faster initial hit-and-run with proto-Earth is possible, but as noted, these runners are more likely to terminate at Venus (Paper II). A hit-and-run return starting with a faster projectile, for example, from the outer solar system, is therefore a less likely scenario for Moon formation around Earth than it would be for Venus. While our scenario allows for higher (and, we argue, more realistic) initial relative velocities than the canonical model, it favors a barely escaping runner and hence a proto-Theia deriving from the inner solar system.

Just as $\theta_{\text{imp}} \sim 45^\circ$ is the most probable impact angle, $\varphi \sim 90^\circ$ is the most probable alignment angle. Orthogonal return collisions are the norm. We find that these lead to disk masses and compositions that are consistent with Moon formation. Moreover, two collisions in a row lead to improved isotopic equilibration, by about 10% in simulations, and a substantially inclined postimpact disk.

5.2. Further Reaccretion and Mixing

The return timescale of runners and debris, of order $0.1\text{--}1$ Myr, is shorter than most error bars on geochronometry applied to the Moon-formation era. A hit-and-run return and its consequences would be resolvable in relative time but not absolutely. It is therefore important to understand the sequential events of follow-on collisions and material exchanges, which altogether constitute the giant impact.

Our nominal scenario improves Earth–Moon isotopic similarity but is not a sufficient explanation. It is not yet clear whether proto-Earth and proto-Theia must be so perfectly equilibrated (e.g., Cano et al. 2020), and $\delta f_{\text{T}}^c = 0$ is too high a

bar to attain. For now we note that our scenario is a significant improvement and also that our methodology underestimates the degree of compositional equilibration for several reasons.

Numerical effects in SPH could limit the particle exchanges during collisional shearing in the simulated hit-and-runs. These could likewise limit proto-Earth and Theia exchanges in simulations of the canonical model, so that accounting for more mixing would improve canonical models as well. If mixing is suppressed generally by SPH in giant impacts, this error would accumulate more significantly in our two-collision model.

Also, we have ignored that the runner is stripped of a fraction of its mantle by the hit-and-run. Theia ends up $\sim 10\text{--}20\%$ more metal-rich than proto-Theia in our starting cases (Table 1), yet with no knowledge of proto-Theia we use a chondritic runner in every case to represent the terminal collision. Because the two major bodies remain on close orbits, the target ends up accreting the majority of the runner’s stripped silicates, according to its greater mass (Asphaug & Reufer 2014). Consequently, Theia’s overall contribution to the Moon is less than we compute.

Another important consideration is that the postimpact middle-Earth might retain a magma ocean from the hit-and-run, given that the interlude timescale may be much shorter than the cooling timescale. From SPH simulations of proto-Earth targets with deep magma oceans, Hosono et al. (2019) obtain more massive protolunar disks with significantly greater proto-Earth mass fraction. Slow hit-and-runs are not as effective as mergers at producing magma oceans, however (Nakajima et al. 2021).

Additionally, we neglect any blending of material between the newly formed Earth and Moon, caused by the reimpact of escaping materials ranging in size from dust to diverse large bodies (e.g., Genda et al. 2017). An era of strong dynamical exchanges will result (e.g., Gladman 1993; Jackson & Wyatt 2012), and a fraction of returning debris would strike the Moon, potentially to strip away, overturn, or remelt its outer layers. Most would strike Earth, potentially to layer the lunar crust with its mega-ejecta. More violent hit-and-runs than we have studied, at higher velocity and smaller impact angle (Reufer et al. 2012), produce much more debris and would increase the significance of these effects.

Catastrophic disruption of the projectile, at velocities faster than about $1.3v_{\text{esc}}$ (leaving the target mostly intact), would produce a deeply mantle-stripped metallic runner, or multiple runners (right panel of Figure 2), with half or more of the projectile escaping as variegated remnants and debris. Longer, more energetic collision chains would lead to more thorough mixing, although at some point perhaps not a Moon. A higher-velocity proto-Theia would originate farther from proto-Earth, counteracting the desired tendency toward isotopic similarity. Also, as noted, according to Paper II Earth is more likely to lose its faster runners to Venus than to retain them, in which case Moon formation would not occur at Earth.

Longer chains, multiple runners, and masses of lesser remnants leave us with many scenarios to consider, as well as the dynamics, chronology, and geochemistry those entail—scenarios for protolunar disk formation and, afterward, for the survival and evolution of the disk and early Moon in the face of returning remnants. And each scenario provides physical and geochemical consequences that might be confirmed or refuted by the earliest lunar geology.

Here we have only analyzed the slow and simple cases without much debris. Considering the other extreme, the most obvious upper limit on debris production is that the Moon exists. Debris-laden chains may be dominated by a Mars-mass Theia whose return collision produces the Moon, but if there are several lunar masses of accompanying debris, their return bombardment over the next ~ 1 Ma might destabilize or destroy the new-formed Moon.

5.3. Why Earth?

Our hypothesis for Moon formation begins with two terrestrial planets colliding at $\lesssim 1.2v_{\text{esc}}$, which for expected impact angles and mass ratios is a hit-and-run collision. A return collision often happens some 0.1–1 Myr later, much slower, and unaligned to the first. Proto-Earth finally acquires the runner and its angular momentum, producing a lunar-mass disk of mixed silicate composition, in the common case of an impact angle $\sim 45^\circ$. This provides a dynamical pathway to the canonical scenario and has the additional advantage of causing greater isotopic mixing.

But it is not without conundrums of its own. It was found in Paper II that proto-Venus is more effective than proto-Earth at ultimately accreting the runners of hit-and-run collisions. In fact, Venus is the terminus for more than half of proto-Earth's faster runners. Earth mainly holds on to its slower runners, which motivates our present focus on low-velocity hit-and-run events. For faster hit-and-runs, a disk-forming terminal graze-and-merge collision would have been more probable at Venus. This shifts the question from *Why does Earth have a moon?* to *Why doesn't Venus have a moon?*

One approach is to invoke the small number statistics of giant impacts. Maybe Venus did not suffer any giant impacts (Jacobson et al. 2017). Or maybe the final giant impact into Venus was head-on, wiping out any satellite from before, instead of the requisite graze-and-merge collision. If the terrestrial planets formed directly by pebble accretion, maybe there was only one giant impact, Theia and proto-Earth (Johansen et al. 2021). But it would have to be slow, nearly v_{esc} , to form the Moon by the demonstrated mechanism of a graze-and-merge collision. A faster collision, especially between bodies of comparable mass, would likely be a hit-and-run collision. While this could set the stage for a Moon-forming return collision such as we propose, a fast runner from Earth is more likely to end up at Venus. So this is not an immediate solution to our conundrum.

While Venus may have been more likely than Earth to have acquired a massive satellite by our hypothesis, it may also have been more likely to have lost one. For the same reason that Venus reaccretes a greater fraction of its runners, compared to Earth, it also reaccretes a greater fraction of its giant impact debris, of which it produces more for a given projectile (the mass ratio being larger, and the denominator v_{esc} being smaller in Figure 2). The orbit being smaller, 0.7 au, the debris will recollide with Venus sooner, with greater efficiency, and at higher velocity. So the argument, that returning debris would erode or even destroy a massive satellite after its formation, is more relevant to Venus than Earth.






The interconnectedness of terrestrial planet formation demonstrated in this set of papers points to the significance of understanding the solidification, layering, and early excavation geology of the Moon, with its profound record of the aftermath of its formation, and of obtaining surface samples

from Venus, which, if there was a late stage, should have a commonality with Earth.

E.A., A.E., S.C., and S.R.S. were supported by NASA grant 80NSSC19K0817, “Application of Machine Learning to Giant Impact Studies of Planet Formation.” Allocation of computer time from the University of Arizona center of High Performance Computing (HPC) is gratefully acknowledged. The authors thank Kevin Zanhle for his insights, as well as two anonymous referees, whose detailed efforts helped us clarify our research and improve its context. And like so many, we are especially indebted to Jay Melosh (1947–2020) for his insightful teachings about planetary geology and formation and his singular enthusiasm for giant impacts and the physics of collisions.

Software: matplotlib (Hunter 2007), collresolve (Emsenhuber & Cambioni 2019).

ORCID iDs

Erik Asphaug  <https://orcid.org/0000-0003-1002-2038>
 Alexandre Emsenhuber  <https://orcid.org/0000-0002-8811-1914>
 Saverio Cambioni  <https://orcid.org/0000-0001-6294-4523>
 Travis S. J. Gabriel  <https://orcid.org/0000-0002-9767-4153>
 Stephen R. Schwartz  <https://orcid.org/0000-0001-5475-9379>

References

- Agnor, C., & Asphaug, E. 2004, *ApJL*, **613**, L157
 Agnor, C. B., Canup, R. M., & Levison, H. F. 1999, *Icar*, **142**, 219
 Armitage, R. M. G., Georg, R. B., Williams, H. M., & Halliday, A. N. 2012, *GeCoA*, **77**, 504
 Asphaug, E. 2010, *ChEG*, **70**, 199
 Asphaug, E. 2014, *AREPS*, **42**, 551
 Asphaug, E., Agnor, C. B., & Williams, Q. 2006, *Natur*, **439**, 155
 Asphaug, E., Jutzi, M., & Movshovitz, N. 2011, *E&PSL*, **308**, 369
 Asphaug, E., & Reufer, A. 2013, *Icar*, **223**, 544
 Asphaug, E., & Reufer, A. 2014, *NatGe*, **7**, 564
 Asphaug, E. 2017, in *Planetesimals: Early Differentiation and Consequences for Planets*, ed. L. T. Elkins-Tanton & B. P. Weiss (Cambridge: Cambridge Univ. Press), 37, doi:[10.1017/9781316339794.002](https://doi.org/10.1017/9781316339794.002)
 Barboni, M., Boehnke, P., Keller, B., et al. 2017, *SciA*, **3**, e1602365
 Barnes, J., & Hut, P. 1986, *Natur*, **324**, 446
 Barr, A. C. 2016, *JGRE*, **121**, 1573
 Belbruno, E., Gott, J., & Richard, I. 2005, *AJ*, **129**, 1724
 Benz, W. 1991, in *Late Stages of Stellar Evolution. Computational Methods in Astrophysical Hydrodynamics*, ed. C. B. De Loore (Berlin: Springer), 258
 Benz, W., Slattery, W. L., & Cameron, A. G. W. 1986, *Icar*, **66**, 515
 Benz, W., Slattery, W. L., & Cameron, A. G. W. 1988, *Icar*, **74**, 516
 Bitsch, B., Lambrechts, M., & Johansen, A. 2015, *A&A*, **582**, A112
 Borg, L. E., Connelly, J. N., Boyet, M., & Carlson, R. W. 2011, *Natur*, **477**, 70
 Budde, G., Burkhardt, C., & Kleine, T. 2019, *NatAs*, **3**, 736
 Burkhardt, C., Kleine, T., Dauphas, N., & Wieler, R. 2012, *E&PSL*, **357**, 298
 Cambioni, S., Asphaug, E., Emsenhuber, A., et al. 2019, *ApJ*, **875**, 40
 Cambioni, S., Jacobson, S. A., Emsenhuber, A., et al. 2021, *PSJ*, **2**, 93
 Cameron, A. G. W. 1997, *Icar*, **126**, 126
 Cameron, A. G. W., & Benz, W. 1991, *Icar*, **92**, 204
 Cameron, A. G. W., & Ward, W. R. 1976, *LPSC*, **7**, 120
 Cano, E. J., Sharp, Z. D., & Shearer, C. K. 2020, *NatGe*, **13**, 270
 Canup, R. M. 2004, *Icar*, **168**, 433
 Canup, R. M. 2008, *Icar*, **196**, 518
 Canup, R. M. 2011, *AJ*, **141**, 35
 Canup, R. M. 2012, *Sci*, **338**, 1052
 Canup, R. M., & Asphaug, E. 2001, *Natur*, **412**, 708
 Canup, R. M., Barr, A. C., & Crawford, D. A. 2013, *Icar*, **222**, 200
 Canup, R. M., Righter, K., Dauphas, N., et al. 2021, arXiv:2103.02045
 Canup, R. M., Ward, W. R., & Cameron, A. G. W. 2001, *Icar*, **150**, 288
 Carlson, R. W., Borg, L. E., Gaffney, A. M., & Boyet, M. 2014, *RSPTA*, **372**, 20130246

- Carter, P. J., Leinhardt, Z. M., Elliott, T., Walter, M. J., & Stewart, S. T. 2015, *ApJ*, **813**, 72
- Carter, P. J., Lock, S. J., & Stewart, S. T. 2020, *JGRE*, **125**, e06042
- Chambers, J. E. 2012, Mercury: A Software Package for Orbital Dynamics, Astrophysics Source Code Library, ascl:1201.008
- Chambers, J. E. 2013, *Icar*, **224**, 43
- Chambers, J. E. 2016, *ApJ*, **825**, 63
- Chambers, J. E., & Wetherill, G. W. 1998, *Icar*, **136**, 304
- Chandrasekhar, S. 1969, Ellipsoidal Figures of Equilibrium (New Haven, CT: Yale Univ. Press)
- Charnoz, S., Sossi, P. A., Lee, Y.-N., et al. 2021, *Icar*, **364**, 114451
- Chau, A., Reinhardt, C., Helled, R., & Stadel, J. 2018, *ApJ*, **865**, 35
- Chen, E. M. A., & Nimmo, F. 2016, *Icar*, **275**, 132
- Connelly, J. N., & Bizzarro, M. 2016, *E&PSL*, **452**, 36
- Čuk, M., Hamilton, D. P., Lock, S. J., & Stewart, S. T. 2016, *Natur*, **539**, 402
- Čuk, M., & Stewart, S. T. 2012, *Sci*, **338**, 1047
- Cummings, W. D. 2019, Evolving Theories on the Origin of the Moon (Cham: Springer Nature), doi:10.1007/978-3-030-29119-8
- Darwin, G. H. 1879, *RSPT*, **170**, 447
- Dauphas, N. 2017, *Natur*, **541**, 521
- Dauphas, N., Burkhardt, C., Warren, P. H., & Fang-Zhen, T. 2014, *RSPTA*, **372**, 20130244
- Deng, H., Ballmer, M. D., Reinhardt, C., et al. 2019, *ApJ*, **887**, 211
- Dwyer, C. A., Nimmo, F., & Chambers, J. E. 2015, *Icar*, **245**, 145
- Dwyer, C. A., Stevenson, D. J., & Nimmo, F. 2011, *Natur*, **479**, 212
- Elkins-Tanton, L. T., Burgess, S., & Yin, Q.-Z. 2011, *E&PSL*, **304**, 326
- Emsenhuber, A., & Asphaug, E. 2019a, *ApJ*, **875**, 95
- Emsenhuber, A., & Asphaug, E. 2019b, *ApJ*, **881**, 102
- Emsenhuber, A., & Cambioni, S. 2019, collresolve, v1.1, Zenodo, doi:10.5281/zenodo.3560892
- Emsenhuber, A., Cambioni, S., Asphaug, E., et al. 2020, *ApJ*, **891**, 6
- Emsenhuber, A., Cambioni, S., Asphaug, E., & Gabriel, T. S. J. 2021, *PSJ*, **2**, 199
- Emsenhuber, A., Jutzi, M., & Benz, W. 2018, *Icar*, **301**, 247
- Frost, D. J., & McCammon, C. A. 2008, *AREPS*, **36**, 389
- Gabriel, T. S. J., & Allen-Sutter, H. 2021, *ApJL*, **915**, L32
- Gabriel, T. S. J., Jackson, A. P., Asphaug, E., et al. 2020, *ApJ*, **892**, 40
- Genda, H., & Abe, Y. 2005, *Natur*, **433**, 842
- Genda, H., Iizuka, T., Sasaki, T., Ueno, Y., & Ikoma, M. 2017, *E&PSL*, **470**, 87
- Genda, H., Kokubo, E., & Ida, S. 2012, *ApJ*, **744**, 137
- Gladman, B. 1993, *Icar*, **106**, 247
- Goldstein, B. E., Phillips, R. J., & Russell, C. T. 1976, *LPSC*, **3**, 3321
- Gross, J., Treiman, A. H., & Mercer, C. N. 2014, *E&PSL*, **388**, 318
- Haghighipour, N., Maindl, T. I., Schäfer, C. M., & Wandel, O. J. 2018, *ApJ*, **855**, 60
- Halliday, A. N. 2000, *E&PSL*, **176**, 17
- Hallis, L. J., Anand, M., Greenwood, R. C., et al. 2010, *GeCoA*, **74**, 6885
- Hartmann, W. K. 2014, *RSPTA*, **372**, 20130249
- Hartmann, W. K., & Davis, D. R. 1975, *Icar*, **24**, 504
- Hauri, E. H., Weinreich, T., Saal, A. E., Rutherford, M. C., & Van Orman, J. A. 2011, *Sci*, **333**, 213
- Hosono, N., Iwasawa, M., Tanikawa, A., et al. 2017, *PASJ*, **69**, 26
- Hosono, N., Karato, S.-i., Makino, J., & Saitoh, T. R. 2019, *NatGe*, **12**, 418
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Ida, S., Canup, R. M., & Stewart, G. R. 1997, *Natur*, **389**, 353
- Jackson, A. P., Gabriel, T. S. J., & Asphaug, E. I. 2018, *MNRAS*, **474**, 2924
- Jackson, A. P., & Wyatt, M. C. 2012, *MNRAS*, **425**, 657
- Jacobsen, S. B. 2005, *AREPS*, **33**, 531
- Jacobson, S. A., Morbidelli, A., Raymond, S. N., et al. 2014, *Natur*, **508**, 84
- Jacobson, S. A., Rubie, D. C., Hernlund, J., Morbidelli, A., & Nakajima, M. 2017, *E&PSL*, **474**, 375
- Johansen, A., Ronnet, T., Bizzarro, M., et al. 2021, arXiv:2102.08611
- Jutzi, M. 2015, *P&SS*, **107**, 3
- Kaib, N. A., & Cowan, N. B. 2015, *Icar*, **252**, 161
- Kenyon, S. J., & Bromley, B. C. 2006, *AJ*, **131**, 1837
- Kokubo, E., Canup, R. M., & Ida, S. 2000a, in Origin of the Earth and Moon, ed. R. M. Canup et al. (Tucson, AZ: Univ. Arizona Press), 145
- Kokubo, E., & Genda, H. 2010, *ApJL*, **714**, L21
- Kokubo, E., & Ida, S. 2002, *ApJ*, **581**, 666
- Kokubo, E., Ida, S., & Makino, J. 2000b, *Icar*, **148**, 419
- Kortenkamp, S. J., & Hartmann, W. K. 2016, *Icar*, **275**, 239
- Kraus, R. G., Stewart, S. T., Swift, D. C., et al. 2012, *JGRE*, **117**, E09009
- Kruijjer, T. S., & Kleine, T. 2017, *E&PSL*, **475**, 15
- Lambrechts, M., & Johansen, A. 2012, *A&A*, **544**, A32
- Lambrechts, M., Morbidelli, A., Jacobson, S. A., et al. 2019, *A&A*, **627**, A83
- Leinhardt, Z. M., Marcus, R. A., & Stewart, S. T. 2010, *ApJ*, **714**, 1789
- Leinhardt, Z. M., & Stewart, S. T. 2012, *ApJ*, **745**, 79
- Lock, S. J., Stewart, S. T., & Čuk, M. 2020, *E&PSL*, **530**, 115885
- Lock, S. J., Stewart, S. T., Petaev, M. I., et al. 2018, *JGRE*, **123**, 910
- Marcus, R. A., Stewart, S. T., Sasselov, D., & Hernquist, L. 2009, *ApJL*, **700**, L118
- Mastrobuono-Battisti, A., Perets, H. B., & Raymond, S. N. 2015, *Natur*, **520**, 212
- Matsumura, S., Brasser, R., & Ida, S. 2017, *A&A*, **607**, A67
- Maurice, M., Tosi, N., Schwinger, S., Breuer, D., & Kleine, T. 2020, *SciA*, **6**, eaba8949
- Melosh, H. J. 1989, Impact Cratering: A Geologic Process (Oxford: Oxford Univ. Press)
- Melosh, H. J. 2007, *M&PS*, **42**, 2079
- Melosh, H. J. 2014, *RSPTA*, **372**, 20130168
- Meyer, J., Elkins-Tanton, L., & Wisdom, J. 2010, *Icar*, **208**, 1
- Monaghan, J. J. 1992, *ARA&A*, **30**, 543
- Morbidelli, A., Chambers, J., Lunine, J. I., et al. 2000, *M&PS*, **35**, 1309
- Movshovitz, N., Nimmo, F., Korycansky, D. G., Asphaug, E., & Owen, J. M. 2016, *Icar*, **275**, 85
- Nakajima, M., Golabek, G. J., Wünnemann, K., et al. 2021, *E&PSL*, **568**, 116983
- Nakajima, M., & Stevenson, D. J. 2014, *Icar*, **233**, 259
- Nakajima, M., & Stevenson, D. J. 2015, *E&PSL*, **427**, 286
- Nakajima, M., & Stevenson, D. J. 2018, *E&PSL*, **487**, 117
- Nemchin, A., Timms, N., Pidgeon, R., et al. 2009, *NatGe*, **2**, 133
- Nie, N. X., & Dauphas, N. 2019, *ApJL*, **884**, L48
- O'Brien, D. P., Morbidelli, A., & Levison, H. F. 2006, *Icar*, **184**, 39
- Pahlevan, K., & Stevenson, D. J. 2007, *E&PSL*, **262**, 438
- Peplowski, P. N., Evans, L. G., Hauck, S. A., et al. 2011, *Sci*, **333**, 1850
- Perera, V., Jackson, A. P., Elkins-Tanton, L. T., & Asphaug, E. 2018, *JGRE*, **123**, 1168
- Pierazzo, E., & Melosh, H. J. 2000, *AREPS*, **28**, 141
- Qin, L., Alexander, C. M. O. D., Carlson, R. W., Horan, M. F., & Yokoyama, T. 2010, *GeCoA*, **74**, 1122
- Quarles, B. L., & Lissauer, J. J. 2015, *Icar*, **248**, 318
- Raskin, C., & Owen, J. M. 2016, *ApJ*, **831**, 26
- Raymond, S. N., Quinn, T., & Lunine, J. I. 2004, *Icar*, **168**, 1
- Reinhardt, C., & Stadel, J. 2017, *MNRAS*, **467**, 4252
- Reufer, A. 2011, PhD thesis, Univ. Bern
- Reufer, A., Meier, M. M. M., Benz, W., & Wieler, R. 2012, *Icar*, **221**, 296
- Rosswog, S. 2009, *NewAR*, **53**, 78
- Rufu, R., Aharonson, O., & Perets, H. B. 2017, *NatGe*, **10**, 89
- Rufu, R., & Canup, R. M. 2020, *JGRE*, **125**, e06312
- Safronov, V. S. 1969, Evolution of the Protoplanetary Cloud and Formation of the Earth and the Planets (Moscow: Nauka)
- Salmon, J., & Canup, R. M. 2012, *ApJ*, **760**, 83
- Sekine, Y., & Genda, H. 2012, *P&SS*, **63**, 133
- Shoemaker, E. M. 1962, in Physics and Astronomy of the Moon, ed. Z. Kopal (New York: Academic), 283
- Solomon, S. C. 1977, *PEPI*, **15**, 135
- Spicuzza, M. J., Day, J. M. D., Taylor, L. A., & Valley, J. W. 2007, *E&PSL*, **253**, 254
- Sridhar, S., & Tremaine, S. 1992, *Icar*, **95**, 86
- Stevenson, D. J. 1987, *AREPS*, **15**, 271
- Stewart, S., Davies, E., Duncan, M., et al. 2020, in AIP Conf. Ser. 2272, ed. J. M. D. Lane et al. (Melville, NY: AIP), 080003
- Stewart, S. T., & Leinhardt, Z. M. 2012, *ApJ*, **751**, 32
- Thompson, S. L., & Lauson, H. S. 1972, Improvements in the CHART-D Radiation-hydrodynamic Code III: Revised Analytic Equations of State, Tech. Rep. SC-RR-71 0714, Albuquerque, NM: Sandia National Laboratories, <https://www.osti.gov/biblio/4208328-improvements-chart-radiation-hydrodynamic-code-iii-revised-analytic-equations-state>
- Tian, Z., & Wisdom, J. 2020, *PNAS*, **117**, 15460
- Tian, Z., Wisdom, J., & Elkins-Tanton, L. 2017, *Icar*, **281**, 90
- Timpe, M. L., Han Veiga, M., Knabenhans, M., Stadel, J., & Marelli, S. 2020, *ComAC*, **7**, 2
- Tonks, W. B., & Melosh, H. J. 1993, *JGR*, **98**, 5319
- Touboul, M., Kleine, T., Bourdon, B., Palme, H., & Wieler, R. 2007, *Natur*, **450**, 1206
- Touma, J., & Wisdom, J. 1998, *AJ*, **115**, 1653
- Visser, R. G., Ormel, C. W., Dominik, C., & Ida, S. 2020, *Icar*, **335**, 113380
- Wada, K., Kokubo, E., & Makino, J. 2006, *ApJ*, **638**, 1180
- Ward, W. R., Canup, R. M., & Rufu, R. 2020, *JGRE*, **125**, e06266
- Warren, P. H. 1985, *AREPS*, **13**, 201
- Wetherill, G. W. 1980, *ARA&A*, **18**, 77

- Wiechert, U., Halliday, A. N., Lee, D. C., et al. 2001, *Sci*, 294, 345
- Wisdom, J., & Tian, Z. 2015, *Icar*, 256, 138
- Wissing, R., & Hobbs, D. 2020, *A&A*, 643, A40
- Wolf, R., & Anders, E. 1980, *GeCoA*, 44, 2111
- Wood, J. A., Dickey, J. S. J., Marvin, U. B., & Powell, B. N. 1970, *Sci*, 167, 602
- Yang, J., Goldstein, J. I., & Scott, E. R. D. 2007, *Natur*, 446, 888
- Youdin, A. N. 2011, *ApJ*, 742, 38
- Young, E. D., Kohl, I. E., Warren, P. H., et al. 2016, *Sci*, 351, 493
- Zahnle, K., Arndt, N., Cockell, C., et al. 2007, *SSRv*, 129, 35
- Zahnle, K. J., Lupu, R., Dobrovolskis, A., & Sleep, N. H. 2015, *E&PSL*, 427, 74
- Zhang, J., Dauphas, N., Davis, A. M., Leya, I., & Fedkin, A. 2012, *NatGe*, 5, 251