



# A New Database of Giant Impacts over a Wide Range of Masses and with Material Strength: A First Analysis of Outcomes

Alexandre Emsenhuber<sup>1,2,3</sup> , Erik Asphaug<sup>2</sup> , Saverio Cambioni<sup>4</sup> , Travis S. J. Gabriel<sup>5</sup> , Stephen R. Schwartz<sup>2,6</sup> , Robert E. Melikyan<sup>2</sup> , and C. Adeene Denton<sup>2</sup>

<sup>1</sup> Universitäts-Sternwarte, Ludwig-Maximilians-Universität München, Scheinerstraße 1, 81679 München, Germany; [emsenhuber@usm.lmu.de](mailto:emsenhuber@usm.lmu.de)

<sup>2</sup> Lunar and Planetary Laboratory, University of Arizona, 1629 E. University Boulevard, Tucson, AZ 85721, USA

<sup>3</sup> Space Research and Planetary Science, University of Bern, Gesellschaftsstrasse 6, 3012 Bern, Switzerland

<sup>4</sup> Department of Earth, Atmospheric & Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

<sup>5</sup> U.S. Geological Survey, Astrogeology Science Center, 2255 N. Gemini Drive, Flagstaff, AZ 86001, USA

<sup>6</sup> Planetary Science Institute, 1700 E. Fort Lowell, Suite 106, Tucson, AZ 85719-2395, USA

Received 2023 March 20; revised 2023 December 24; accepted 2023 December 31; published 2024 March 4

## Abstract

In the late stage of terrestrial planet formation, planets are predicted to undergo pairwise collisions known as giant impacts. Here, we present a high-resolution database of giant impacts for differentiated colliding bodies of iron–silicate composition, with target masses ranging from  $1 \times 10^{-4} M_{\oplus}$  up to super-Earths ( $5 M_{\oplus}$ ). We vary the impactor-to-target mass ratio, core–mantle (iron–silicate) fraction, impact velocity, and impact angle. Strength in the form of friction is included in all simulations. We find that, due to strength, the collisions with bodies smaller than about  $2 \times 10^{-3} M_{\oplus}$  can result in irregular shapes, compound-core structures, and captured binaries. We observe that the characteristic escaping velocity of smaller remnants (debris) is approximately half of the impact velocity, significantly faster than currently assumed in  $N$ -body simulations of planet formation. Incorporating these results in  $N$ -body planet formation studies would provide more realistic debris–debris and debris–planet interactions.

*Unified Astronomy Thesaurus concepts:* [Planet formation \(1241\)](#); [Solar system terrestrial planets \(797\)](#); [Extrasolar rocky planets \(511\)](#)

*Supporting material:* machine-readable table

## 1. Introduction

In the last stage of classical terrestrial planet formation, collisions between similar-sized planetary embryos are thought to be the dominant mode of growth (e.g., Wetherill 1985; Kokubo & Ida 2002) where Moon- to Mars-sized bodies accumulate dynamically to form the final planets. Other stages of planet, planetesimal, and satellite formation may also involve giant impacts, or more generally, similar-sized collisions (Asphaug 2010). A correct understanding of how planets accumulate and exchange matter in these numerous giant impacts thus underlies our most basic knowledge of planet formation.

Giant impacts are complicated phenomena. Colliding bodies can be centrally condensed, leading to large mass fractions outside the direct collision zone (e.g., Genda et al. 2012; Movshovitz et al. 2016). Off-axis impacts involve high angular momentum and limited accretion efficiency (Agnor & Asphaug 2004). They result in a complicated postcollision phase (e.g., Cameron 1997). The fraction of the impactor mass  $m_{\text{imp}}$  that gets accreted onto the target of original mass  $m_{\text{tar}}$  and final mass  $m_{\text{L}}$  is given by the accretion efficiency

$$\xi_{\text{L}} = \frac{m_{\text{L}} - m_{\text{tar}}}{m_{\text{imp}}}. \quad (1)$$

A perfect merger has the largest possible accretion efficiency,  $\xi_{\text{L}} = 1$ . A net accretion requires  $\xi_{\text{L}} > 0$ , and  $\xi_{\text{L}} < 0$  describes the mass loss (erosion or disruption).

The slowest possible collisions occur at around the mutual escape velocity at contact, where for colliding spheres

$$v_{\text{esc}} = \sqrt{\frac{2G(m_{\text{tar}} + m_{\text{imp}})}{r_{\text{tar}} + r_{\text{imp}}}}. \quad (2)$$

Here,  $G$  is the gravitational constant, and  $r_{\text{tar}}$  and  $r_{\text{imp}}$  are the target and impactor radii, respectively. The largest bodies of a growing planetary system, under conditions of gravitational self-stirring, collide at relative velocities near their mutual escape velocity (e.g., Wetherill 1985).  $N$ -body simulations find that most solar system giant impacts are faster than  $v_{\text{esc}}$  ( $\sim 1\text{--}4 v_{\text{esc}}$ , e.g., Agnor et al. 1999; Quintana et al. 2016), and the impact angle is often off-axis. Because of this, the impactor and target can undergo “hit-and-run,” where the bodies remain relatively unscathed after a glancing blow, and the accretion efficiency is close to zero (Agnor & Asphaug 2004). This is expected to regulate the pace/velocity of planet formation. In these typical scenarios, the impactor plows through the target mantle and emerges as a deflected, decelerated, and gravitationally intact body called the *runner*. The bodies may then recollide after orbiting the central star.

At more head-on and/or at lower velocities, “graze-and-merge” collisions are possible (e.g., Leinhardt et al. 2010). These are high angular momentum accretions, as represented by the canonical Moon-forming giant impact (Canup & Asphaug 2001).

When averaged over the impact angle, graze-and-merge may be the dominant form of giant impact accretion (Stewart & Leinhardt 2012).

In some graze-and-merge-like scenarios, the gravitationally bound runner can overshoot the target, but the bodies may not entirely escape one another. Tidal friction and transfer of angular momentum around an irregular central mass can then cause material to be captured as a moon, at least temporarily; such “graze-and-capture” collisions include hypotheses for Moon formation (Benz et al. 1987) and Pluto–Charon formation (Canup 2005). This represents an intermediate case between graze-and-merge and hit-and-run.

### 1.1. Dynamical Significance

The simulation of the gravitational interactions of a planetary system ( $N$ -body simulations) depends on how collisions are treated. If the colliding bodies are assumed to be perfectly merging, the mass is conserved, and the new orbit is often placed at the center of mass of the colliding bodies (e.g., Duncan et al. 1998). Perfect merging may be a sufficient assumption for understanding the largest-scale architectures in the solar system (e.g., Kokubo & Genda 2010; Quintana et al. 2016; Walsh & Levison 2019). However, the approximation alters the dynamical evolution and the formation sequence.

Including the realistic collision outcomes increases the formation timescales (Chambers 2013; Quintana et al. 2016), because inefficient accretions—especially hit-and-runs—are common. If the runner returns to the target, or to another nearby accreting body (Emsenhuber et al. 2021), it is a “collision chain,” a hit-and-run followed by another similar-sized collision (merger, hit-and-run, disruption, etc.). During planet formation at  $\sim 1$  au around a solar-mass star, the recollision time for a collision chain can be  $1 \times 10^3$ – $1 \times 10^6$  yr (Emsenhuber & Asphaug 2019a). Such multicollisional pathways could lead to mantle-stripped cores, the “stranded runner” hypothesis for the origin of Mercury (Asphaug & Reufer 2014; Chau et al. 2018), and for metallic asteroids and meteorites (Yang et al. 2007).

Also, a realistic treatment of collisions affects the resulting composition (Carter et al. 2015; Dwyer et al. 2015), and the mixing of material between planetary accretion zones (Burger et al. 2020), and leaves behind a diversity of smaller bodies (e.g., Mars-sized; Emsenhuber et al. 2020). This is because a growing fraction of the remainder end up surviving hit-and-run collisions as the growth of the largest bodies proceeds (Asphaug & Reufer 2014; Asphaug 2017). This could explain the observed increasing diversity of terrestrial planets with decreasing mass (Gabriel & Cambioni 2023).

Implementing realistic treatments of similar-sized collisions can be achieved in  $N$ -body codes in several ways. For example, scaling laws have been developed (e.g., Leinhardt & Stewart 2012; Reinhardt et al. 2022) from 3D giant impact simulations that characterize the outcomes of giant impacts. One application of our new database would be to validate or update the parameters to such scaling laws or develop new ones, as our parameter space includes friction in all simulations, providing more accurate results for impacts at smaller scales. Another approach is to apply machine learning to the outcomes of giant impact simulations (Timpe et al. 2020). Surrogate models (e.g., Cambioni et al. 2019, 2021) can be generated, which relate preimpact conditions to postimpact conditions such as the largest remnant mass  $m_L$ , the second

remnant mass  $m_S$  (runner), dynamical properties, and compositional information. The present database, including an initial assessment of debris, is designed specifically for building machine-learning models and spans a larger range of collisions than an existing giant impact database used for machine-learning models (Cambioni et al. 2019; Emsenhuber et al. 2020). One important limitation, however, is that the impacting bodies are not rotating before impact in our database. Introducing rotation would expand the parameter space considerably, but is a factor to consider given its effect on postimpact outcomes (Timpe et al. 2020), and the profound effects it may have on the internal structure of the resulting bodies (Lock & Stewart 2017).

### 1.2. Geophysical Significance

Certain aspects of giant impacts, such as the mass of remnants, can be modeled by analytical relationships (such as scaling laws). However, obtaining a unified model that can predict the outcomes of impacts across various regimes, such as in small asteroid-scale collisions and in shock-inducing Moon-formation events, is challenging. This is due to numerous physical complexities inherent to collisions at different scales and in different bodies (Gabriel et al. 2020). Transitions in the equations of state (EOSs; Stewart et al. 2020), vapor production (Benz et al. 2007; Carter et al. 2020; Davies et al. 2020), and the presence of shocks in large-scale collisions produce ample vapor and may alter the nature of erosion (Gabriel et al. 2020; Gabriel & Allen-Sutter 2021). At small scales, friction and strength make erosion less likely for a given scaled velocity (e.g., Jutzi 2015), and other forms of dissipation become substantial (Melosh & Ivanov 2018). Even in Mars-scale collisions where rheology has not classically been noted to alter the mass of the largest remnant (e.g., Benz & Asphaug 1999), other aspects of the collision such as heating and debris generation are influenced by the presence of strength (Emsenhuber et al. 2018). To make meaningful progress toward a unifying model across these complex regimes, a set of simulations that spans over a large range of preimpact conditions is required.

Still, giant impact outcomes do have systematic trends across vast ranges. For instance, Jutzi (2019) identified three basic regimes of giant impacts, or similar-sized collisions, for the velocity ranges that can potentially result in accretion. For bodies smaller than  $\sim 100$  km, there is the porosity regime, where the outcome is mainly affected by pore crushing (e.g., Housen & Holsapple 1999; Belton et al. 2007; Jutzi et al. 2015). Intermediate-mass bodies (up to  $\sim 1000$  km) are in the strength regime, where friction is important (for the link between strength and friction, see Section 2.4). The largest bodies are in the gravity regime, where strength and porosity can be ignored. For large enough bodies where mutual escape velocities (and thus impact velocities) exceed the sound speed, the shocks dominate and result in a fourth regime (Gabriel et al. 2020). These regimes are idealizations, and the transitions are not abrupt. Indeed, the transition from strength to gravity dominance may span the entire range of “oligarchic growth,” the canonical late stage of Moon- to Mars-sized bodies (Kokubo & Ida 2002). For this reason, we include a strength model in all of our simulations, even in super-Earth collisions where the effect on the outcome is expected to be negligible. This allows the data to properly capture the smooth transition from strength to gravity dominance.

The end state of the last giant impact represents the beginning of a planet’s long-term geophysical evolution (e.g., Zahnle et al. 2007). Accreted planets can end up with mantles and cores unstable to convection, especially for larger-scale collisions in which the gravitational potential released by the merging material exceeds the energy of melting (Lock et al. 2020). Smooth particle hydrodynamics (SPH), the technique most widely used to model giant impacts, is unable to properly represent long-term convective instability and other effects that happen on a much longer timescale than the collision. To study postimpact geodynamics requires a hybrid numerical approach, as has been applied by Golabek et al. (2018) to model consequences like geothermal overturn, geochemistry, and solidification after giant impacts. For solid final bodies, finite element analysis could be used to predict postcollisional relaxation from these SPH results.

### 1.3. Geochemical Significance

Giant impacts are transformative events, and simulations of giant impacts have become the basis for making quantitative geochemical predictions about late-stage planet formation (e.g., Haghighipour et al. 2018), especially the origin of the Moon (Canup et al. 2021). But colliding planetary bodies in simulations are represented by idealized compositions, and these and other simplifying assumptions need to be recognized. We represent terrestrial planets and their progenitors as differentiated nonporous spheres of forsterite and iron composition, with varying core mass fractions. This simple interior structure allows us to focus on attaining accuracy and reliability of the model predictions, while making a sufficient sweep of the parameter space to obtain results that may be applied generally to models of terrestrial planet formation. Furthermore, there remains a limited availability of shock EOS models for a wider range of mantle materials, and the forsterite EOS (Stewart et al. 2020) is the most up-to-date and widely used for our purposes.

A two-component rock/iron interior is a justifiable approximation for terrestrial bodies in the modeled size range. For instance, 525 km diameter (4) Vesta has a rocky mantle and an iron core (e.g., Russell et al. 2012); so do Mercury, Venus, the Moon, and Earth, to good approximation. So we maintain this assumption from the smallest colliding bodies in our database, about 1000 km diameter, up to  $\sim 5 M_{\oplus}$ , beyond which point gas-free accretion is unlikely (e.g., Rogers 2015).

Numerical studies have revealed the complex thermodynamic evolution that occurs in giant impacts (Carter et al. 2020) and how outcomes can depend on the choice of initial thermal conditions, even for relatively simple two-component planets. For example, a massive magma ocean on the target can enhance the postimpact disk mass and its Earth-isotopic fraction in simulations of Moon formation (Hosono et al. 2019). That said, the material from which the Moon forms represents only  $\sim 1\%$  of the material of the colliding bodies in the Canonical model for Moon formation (Canup 2005). For our database, we thus implement only one thermal state across all colliding bodies: both the core and the mantle start slightly below the solidus, on the basis that convective cooling may be faster than the time between giant impacts.

### 1.4. Scope of Work

We present a significant new database of terrestrial planet-forming giant impacts. To take into account the limitations and bottlenecks of previous works, our new database has the following characteristics:

1. It is applicable to a wide range of impactor-target properties and impact parameters, from the sizes of the largest asteroids and rocky satellites, to terrestrial planets of several Earth-masses ( $M_{\oplus}$ ), which are not expected to be large enough to accrete significant atmospheres (e.g., Rogers 2015).
2. It includes the transitions between collision regimes, especially between graze-and-merge and hit-and-run, i.e., accretion and nonaccretion.
3. It provides the outcomes of collisions for bodies with various core mass fractions, thus enabling self-consistent treatments of core size evolution in collision chains.

Within this scope, we construct a database of 1250 simulations (split in two sets, one comprising 1000 simulations and one 250) suitable for the development of machine-learning models. To compute the final reduced properties of each simulation, we proceed in multiple steps. First, we define the setup of the study (Section 2.1) and determine the initial conditions of hydrodynamical simulations (Section 2.2). Then, we perform simulations using SPH (Section 2.3) and analyze their results (Section 2.3). We also present several results from our calculation: general results (Section 3.1); specific items relevant to low-velocity grazing collisions, satellite capture (Section 3.2), body shapes (Section 3.3), and debris (Section 3.4).

## 2. Methods

### 2.1. Studied Parameters

The first step to generating our database of giant impact simulations is to select the parameters that will be explored. Target mass  $m_{\text{tar}}$  and the impactor-to-target mass ratio  $\gamma = m_{\text{imp}}/m_{\text{tar}}$  are the primary parameters. For the composition, we use a single parameter, the core mass fraction  $Z_{\text{tar}}$  and  $Z_{\text{imp}}$ , which describes the fraction of iron in our two-layered bodies, the rest being forsterite. Core mass fraction is a free parameter in our database as it will allow us to understand the collision outcomes between planets that have been previously eroded, e.g., to know the outcomes of collisions between core-rich bodies. The impact velocity is given in terms of the mutual escape velocity  $v_{\text{coll}}/v_{\text{esc}}$ , which thus increases with total mass. The impact angle  $\theta_{\text{coll}}$  is defined as the angle between the line connecting the centers of mass and the relative velocity vector at impact. We use the convention that  $\theta_{\text{coll}} = 0^\circ$  is a head-on collision, and  $\theta_{\text{coll}} = 90^\circ$  is a grazing collision.

For the present work, we have decided not to include the effect of preimpact rotation. Timpe et al. (2020) found that preimpact rotation had less influence on the giant impact outcomes (apart from final spin, which is strongly correlated) than the colliding mass, mass ratio, core mass fraction, impact angle, and impact velocity. An important advantage of leaving out preimpact rotation is the smaller dimensionality of the parameter space. Accounting for all possibilities of preimpact rotation when the target and impactor are similar in size would require six additional parameters, three for each body, to describe the spin angular momentum vectors.

**Table 1**

Cumulative Distribution of Impact Velocities in the Grid of Hydrodynamical Simulations

$v_\infty/v_{\text{esc}}$	$v_{\text{coll}}/v_{\text{esc}}$	Cumulative Fraction
0	1	0
1	$\sim 1.4$	0.4
2	$\sim 2.2$	$2/3$
3.5	$\sim 3.6$	0.8
6	$\sim 6.1$	0.9
$\sqrt{99}$	10	1

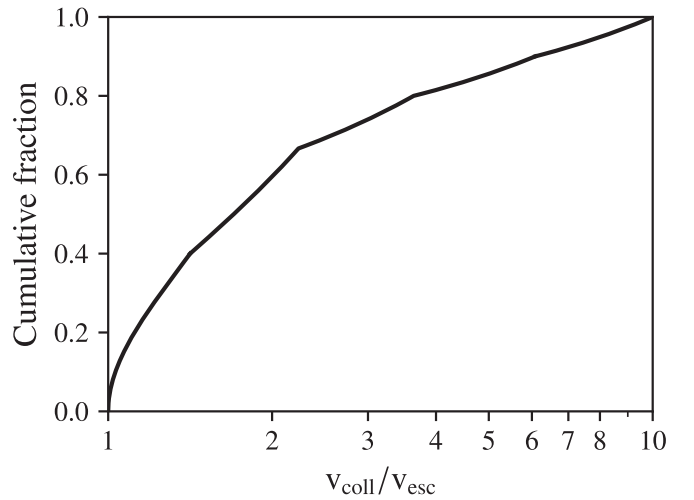
## 2.2. Parameter Range and Initial Conditions

We aim to study the effect of the target’s mass, so we explore a large range from  $1 \times 10^{-4} M_\oplus$  to  $5 M_\oplus$ , which is sampled uniformly in logarithm space. The lower boundary of target mass corresponds to about the mass of (1) Ceres. This range was selected because it incorporates the size range where the effect of friction (included in every simulation in the generated database) is important (e.g., Emsenhuber et al. 2018), yet where porosity is likely unimportant for these massive, differentiated bodies. The upper boundary was selected so that the collisions can be applied to the formation of extrasolar planetary systems that contain super-Earths. We do not include simulations of planets with gas envelopes, so we limit our study to masses below  $5 M_\oplus$  because more massive planets tend to have significant envelopes (e.g., Rogers 2015).

As for the mass ratio  $\gamma$ , our sampling is uniform so that unequal mass ratios are studied as much as more similar mass ratios, as suggested by Valencia et al. (2019). We selected the range of  $0.05 < \gamma < 1$ , covering the  $\sim 1:3$  diameter range defining similar-sized collisions (Asphaug 2010) and as constrained by our capability to computationally resolve the smaller body. The least massive impactor in our database is  $5 \times 10^{-6} M_\oplus$ , somewhat more massive than Main Belt asteroid (16) Psyche.

The core mass fractions of  $Z$  are sampled from a piecewise uniform distribution over the range of 10%–90%, where the range is selected so that the core and mantle are numerically resolved at least several particles across for the chosen resolution. Further, to account for the population of bodies around the average value of Fe/Mg for stars with planets from the Hypatia Catalogue (Hinkel et al. 2014; see also Figure 2 in Scora et al. 2020), and the average chondritic values of the terrestrial planets, we center the distribution so that half the bodies have a core mass fraction below 30%, and the rest are above. For this median value, which is around the value for Earth and Venus, the core radius is about half the body radius.

According to  $N$ -body studies, the majority of giant impacts during terrestrial planet formation occur in the range  $1-2v_{\text{esc}}$  (e.g., Chambers 2013; Emsenhuber et al. 2020), although the median value depends on the damping effects of planetesimals (O’Brien et al. 2006). According to Asphaug (2010), the diversity of planet formation by giant impacts is due to the fact that the transitions in outcomes happen around this accretionary range of velocities. Genda et al. (2012) noted a significant variability at the low-velocity merging and graze-and-merge transition, for minor changes in simulation parameters. Similarly, Cambioni et al. (2019) found the highest rate of misclassification and prediction error to occur at the transition between hit-and-run and merging, in their machine-learning analysis of a prior database of giant impact outcomes.



**Figure 1.** Cumulative distribution of impact velocities in our giant impact database. The parameter distribution was chosen to sample regions where postimpact outcomes transition from one type of impact outcome (e.g., accretion) to another (e.g., hit-and-run).

Therefore, we place emphasis on mapping out this transition by constructing a distribution of impact velocity that favors low-velocity collisions. We note that minor deviations in the transition as a function of different thermal conditions are not covered in this work. We define the cumulative distribution from the relative velocity at infinity, as scaled by the mutual escape velocity  $v_\infty/v_{\text{esc}}$ . This is related to the impact velocity by

$$\left(\frac{v_{\text{coll}}}{v_{\text{esc}}}\right)^2 = 1 + \left(\frac{v_\infty}{v_{\text{esc}}}\right)^2 \quad (3)$$

according to energy conservation. The result is a piecewise uniform distribution as provided in Table 1. The resulting cumulative distribution, given for  $v_{\text{coll}}/v_{\text{esc}}$ , is shown in Figure 1.

As for the impact angle  $\theta_{\text{coll}}$ , we follow the distribution expected for collisions between point impactors onto gravitational targets,  $dP \propto \sin(2\theta_{\text{coll}})d\theta_{\text{coll}}$  (Shoemaker 1962). This ensures that the transitions between hit-and-run and graze-and-merge and accretion, which occur around the nominal range of impact angles at nominal velocities (e.g., Leinhardt & Stewart 2012), are well sampled.

To generate the specific combinations of the parameters, which will be used to perform the hydrodynamical simulations, we use the Latin hypercube sampling (LHS) method (e.g., McKay et al. 1979; Timpe et al. 2020). LHS divides each parameter into  $n$  intervals of equal probability based on the distribution, with  $n$  being the number of samples. Then, one sample is selected randomly from each interval. This ensures that the entire range of possible values for each parameter is sampled. In addition to that, LHS adopts criteria that ensure that the entire parameter space is well sampled. For instance, the algorithm avoids correlations between parameter values in the selected samples, so that the effect of each parameter can be disentangled. In this work, we used the `pyDOE` Python package with the `minmax` setting. This package returns all the samples in the  $[0, 1]$  range with uniform probability. We convert these to the actual collision parameters using the probability distributions discussed above.



For this work, we generated two lists of initial conditions using LHS, one with  $n = 1000$  entries, and one with  $n = 250$ . The intention was to have separate sets for future machine-learning applications. For most of the analysis, we combine the two in a set of 1250 simulations.

### 2.3. Hydrodynamical Simulations

To perform the hydrodynamical simulations, we use the SPH technique (e.g., Monaghan 1992; Rosswog 2009). SPH uses a Lagrangian description with continuum material divided into particles. Quantities are retrieved by performing a kernel interpolation and their spatial derivatives by interpolation of the underlying quantity with the kernel’s gradient. Time evolution is given by the Euler’s equations, except that the density is computed by performing kernel interpolation at each step and corrected for free-surface effects using the method of Reinhardt & Stadel (2017), which works by increasing the density of a particle if there is a spatial imbalance of neighbors around it. An artificial viscosity term inspired by Riemann solvers (Monaghan 1997) with a time-dependent factor (Morris & Monaghan 1997) is included to resolve shocks, the exact form of which is described in Emsenhuber et al. (2021). The interpolation is performed using a cubic spline kernel (Monaghan & Lattanzio 1985) with each particle having about 50 neighbors. We note that SPH tends to spuriously damp the subsonic turbulence (e.g., Cullen & Dehnen 2010; Bauer & Springel 2012; Deng et al. 2019); however, this should affect more the internal mixing within the final bodies rather than their global iron–silicate fraction. Finally, material strength is modeled according to the formulation discussed in the next section.

To close the Euler equations, we need to provide an EOS that provides the pressure as a function of the density and the specific internal energy  $p(\rho, u)$ . The choice of EOS is limited by the requirement that they be thermodynamically reliable in all the energy regimes applicable to a giant impact (Melosh 2007). In this work, we use ANEOS for the iron core (Thompson & Lauson 1972) and a modified version of ANEOS for  $\text{Mg}_2\text{SiO}_4$  (forsterite; Stewart et al. 2020) for the mantle. To improve performance, we precompute tables directly from ANEOS for the anticipated range of values as in previous works (e.g., Benz et al. 1989; Reufer et al. 2012). The code nevertheless retains the capability to call ANEOS for the few cases where the values lie outside of the tabulated range.

The computer code used in this study is SPHLATCH (e.g., Reufer et al. 2012; Asphaug & Reufer 2013, 2014) and includes additional updates and corrections presented in Ballantyne et al. (2023). An earlier version of the same code was used to generate a previous collision database (Reufer 2011) on which the first machine-learning derived surrogate models were based (Cambioni et al. 2019; Emsenhuber et al. 2020), and an updated semiempirical scaling law was developed (Gabriel et al. 2020). Those simulations did not include friction, and the reported database spans  $1 \times 10^{-2} - 1 M_{\oplus}$  planets (for more information, see Gabriel et al. 2020).

### 2.4. Constitutive Strength

The constitutive model we adopt is similar to that from Emsenhuber et al. (2018): elastic perfectly plastic material (Benz & Asphaug 1994, 1995), with the deformation tracked by the deviatoric stress tensor, that is reduced at the Hugoniot

elastic limit with  $\sqrt{J_2}/Y$ , where  $J_2$  is the second invariant of the deviatoric stress tensor (not to be confused with the global gravity coefficient), and  $Y$  is the pressure-dependent yield strength (Collins et al. 2004; Jutzi 2015). We further include a correction tensor to compute the local velocity gradient to achieve angular momentum conservation (Bonet & Lok 1999; Speith 2006).

As a simplification, we assume that cohesion is zero, as its effect is only noticeable in similar-sized collisions in the diameter range  $\lesssim 1$  km (Jutzi 2015). The model thus assumes fully damaged material in the solid state, governed by a friction model, given by

$$Y_d = \mu_p p, \quad (4)$$

where  $\mu_p$  is the coefficient of friction, a material parameter, and  $p$  is pressure. The subscript “d” refers to damaged material. The strength does not increase arbitrarily with pressure; it is limited by the yield strength of intact (nondamaged) material,

$$Y_i = \frac{\mu_i p}{1 + \mu_i p / Y_m}, \quad (5)$$

where  $\mu_i$  is the coefficient of internal friction, and  $Y_m$  is the von Mises plastic limit. The subscript “i” refers to intact material. The *full* form of the yield strength is  $Y_p = \min(Y_d, Y_i)$ , which represents a material supported by friction subject to plastic yielding.

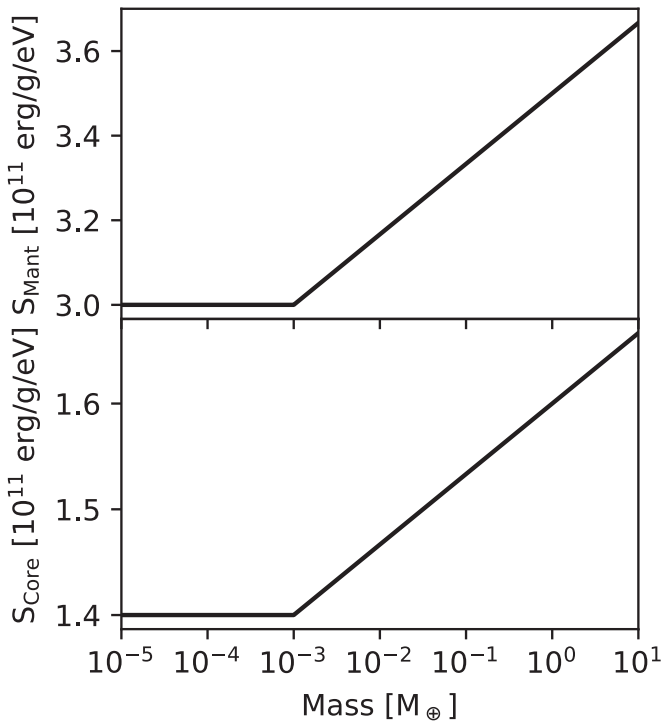
We adopt the constitutive model for “rock materials” in Table A.1 of Collins et al. (2004); here,  $\mu_p = 0.8$ , and  $\mu_i = 2.0$  for all simulations. We assume the same coefficients of friction for the iron cores as well; although it is simplistic, this choice is unlikely to matter because the pressure at the core–mantle boundary in all our bodies exceeds  $Y_m$  for iron, 0.68 GPa. The forsterite mantle is subject to friction, when at lower pressure, and  $Y_m = 3.5$  GPa. We study the effect of our choice for  $\mu_p$  and  $\mu_i$  in Appendix B.

Yield strength is also temperature-dependent. To capture this effect, we further modulate the yield strength as it approaches the melting point with

$$Y_T = Y_p \tanh \left[ \zeta_T \left( \frac{T_M}{T} - 1 \right) \right], \quad (6)$$

where  $\zeta_T = 1.2$  is the thermal softening parameter (Ohnaka 1995; Collins et al. 2004), and  $T_M$  is the melting temperature. The melting temperature is consistently recovered from the EOS by determining the melting temperature for both materials. This is possible because the ANEOS forsterite EOS (Stewart et al. 2020) provides the full phase information, while the older EOS for quartz (Melosh 2007) applied in previous work (e.g., Reufer 2011; Emsenhuber et al. 2018) does not.

In summary, all simulations are performed with the strength of a fully damaged material modulated by yielding beyond each material’s plastic limit, and further by temperature. The strength thus applies to a thin or even unresolved layer in the larger bodies of the database, and to the entire mantle and perhaps some of the core in the smallest bodies in the database. While including strength adds computational cost for planet-scale collisions where strength is not expected to influence outcomes, it ensures constant treatment across the database. Moreover, the simulations involving the most massive bodies are the least computationally expensive, due to the Courant–Friedrichs–Lewy time step criterion being proportional to the



**Figure 2.** Initial specific entropy for the core ( $S_{\text{core}}$ ) and mantle ( $S_{\text{mant}}$ ) as function of the body’s mass. The values of entropy are given in the units of ANEOS,  $\text{erg g}^{-1} \text{eV}^{-1}$ .

spatial resolution (which is itself proportional to the body sizes with a constant number of particles, as in this work). Thus, including friction at the largest scales does not significantly slow down the computation of the database, since most computational effort is on the smaller bodies.

### 2.5. Body Preparation and Initial Thermal State

Colliding bodies need to have their initial thermal state specified. This is particularly important for smaller-scale collisions as our SPH model includes friction. The yield strength in the friction model is modulated by the ratio between the temperature of the material undergoing stresses and that of the melting point (see Equation (6)).

We use isentropic profiles for each of the core and mantle. We leave the initial specific entropy of the core and mantle,  $S_{\text{core}}$  and  $S_{\text{mant}}$  respectively, to be determined. We prescribe thermal profiles such that a large portion of the cores and mantles of the preimpact bodies are in a solid phase, but not far from their melting point. Due to the wide range of body masses that represent our starting conditions, this thermal state is not achievable with a single global value of entropy for the cores or mantles. Hence, we selected the relationship presented in Figure 2. These have two different regimes, with a constant entropy for all bodies whose masses are below  $1 \times 10^{-3} M_{\oplus}$  and a log-linear relationship above. The selected entropies result in core and mantle temperatures of 1651 and 1657 K respectively for the bodies at the low-mass end of the range. For larger bodies, a temperature drop occurs at the core–mantle boundary, which is expected. Thermal conduction is not included in our SPH model, so this temperature jump between materials persists through calculations. To highlight the jump in temperature, we show a profile on an Earth-like body (mass  $M = 1M_{\oplus}$ , and core mass fraction  $Z = 30\%$ ) in Figure 3. Note

that, for the chosen EOS, the outer radius of the body is 6577 km, roughly 3% larger than the Earth.

The body preparation scheme follows that of Emsenhuber et al. (2021): we begin by obtaining a 1D radial profile in hydrostatic equilibrium using the scheme presented in Benz (1991). The profiles are then mapped onto 3D profiles using the methodology described in Reinhardt & Stadel (2017) and further relaxed for 6 hr to damp particle motions. To enforce the prescribed initial value, we use a fixed-entropy SPH formulation (rather than evolving energy) during this relaxation step of problem setup.

We choose the resolution so that there are roughly 100,000 total particles in the simulation (in the target and the impactor). For the smallest mass ratio scenarios ( $\gamma = 0.05$ ), this equates to roughly 5000 particles in the impactor, since the particles are roughly (but not exactly) equal mass in our simulations. Prior studies (e.g., Asphaug 2010) have shown that accretion efficiency is resolved to a few percent for SPH simulations with 20,000 and more particles, and Meier et al. (2021) found that  $Q_{\text{RD}}^*$  (the catastrophic disruption threshold) is converged for  $1 \times 10^5$  particles in SPH simulations using ANEOS, providing additional confidence that our results for lower-energy, hit-and-run scenarios are converged in terms of the largest remnant mass. We also performed simulations with a resolution increased by a factor of 10 (roughly 1 million total particles) that show only small differences compared to the nominal resolution (Appendix C).

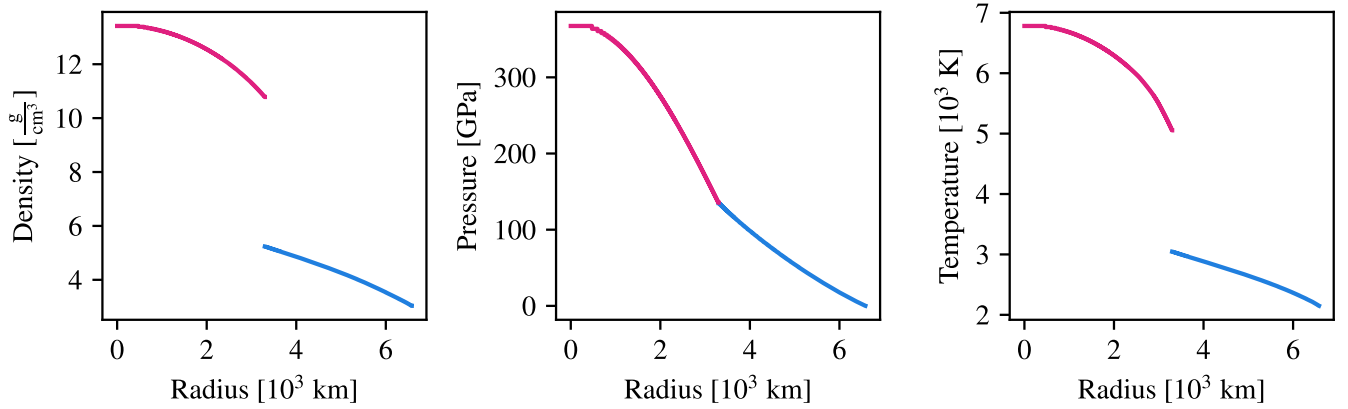
### 2.6. Simulation Evolution

We set up each collision by placing the bodies at a distance equal to 5 times the sum of the radii, well outside the region where tidal influence begins to deform the bodies (more than twice the Roche limit). The collisions are then evolved for  $50 \tau_{\text{coll}}$  past the initial contact, which is defined as

$$\tau_{\text{coll}} = \frac{2(r_{\text{tar}} + r_{\text{imp}})}{v_{\text{coll}}}. \quad (7)$$

$50 \tau_{\text{coll}}$  corresponds to a bit more than 1 day for the low-velocity collisions (that is, at the mutual escape velocity). By inspection, we find that this duration is sufficient in most cases to determine the collision outcome. However, grazing collisions, where the impactor is captured on a bound orbit with a long duration loopback orbit, demand longer integrations. These collisions encompass both graze-and-merge, where the impactor is left on an orbit that intersects with the target, and graze-and-capture regimes. Graze-and-capture refers to collisions where the impactor remains as a bound satellite, akin to the scenario of Canup (2005) for the origin of Pluto–Charon.

For cases that end up near the transition between the hit-and-run and graze-and-merge regimes, the simulations and end-states must be analyzed separately (Emsenhuber & Asphaug 2019a). The runner’s loopback orbit requires days of physical time, typically, so for returning runners with more than about  $\sim 10\%$  of the impactor’s mass, we continue the SPH evolution until it has made at least one more passage of the pericenter. Afterwards, it may be tidally disrupted or partially accreted. We discriminate two cases in these scenarios. If the orbital period of the runner around the target is smaller than about 7 days, the simulation is further evolved until, usually, no such body remains. We checked that heating due to spurious activation of the artificial viscosity (for



**Figure 3.** Initial 1D radial profile of a body with mass  $M = 1 M_{\oplus}$ , and a core mass fraction  $Z = 30\%$ . The red curve corresponds to the iron core, and the blue line corresponds to the forsterite mantle.

instance, due to residual motion inside the bodies) is minimal during the loopback orbit. However, if the orbital period is determined to be longer than 7 days, the simulation is kept in this intermediate state and marked as such. This is to avoid the buildup of numerical imprecision in the gravity solver during such an extended period, which could make the final state less realistic.

### 2.7. Simulation Analysis

We proceed as in Emsenhuber & Asphaug (2019b), Emsenhuber et al. (2020) for the analysis of the SPH simulations, and perform additional analysis of body shapes, interior structures, and debris. We begin with a friends-of-friends search of SPH particles to find contiguous bodies. Each of these bodies is replaced by one *superparticle* with equivalent mass and momentum for the remaining steps. The second step is to find gravitationally bound material and identify them as unique clumps. In our work, we consider a minimum clump mass worthy of analysis to be 10 times the mass of the most massive SPH particle, although in practice our results do not depend on this minimum. The particles not part of any clump are considered *unresolved* debris. Identified clumps can include coorbiting pairs of superparticles and multiple-body systems. A last step is to find physical bodies. When a single contiguous body is found in a clump, all remaining particles in such a clump are attributed to that body. When there are more, the particles not part of a contiguous body are attributed to the body with which they have the lowest relative energy.

These bodies and clumps provide the basic metrics of collision outcomes, which are archived in machine-readable format. The most massive clump (from the second analysis step) is taken as the largest remnant, whose accretion efficiency  $\xi_L$  is given by Equation (1). In a perfect merger,  $\xi_L$  equals 1, and a value of 0 represents a target with a postimpact mass that is equal to its preimpact mass. The second most massive clump is defined as the second remnant, for which we define the accretion efficiency of the second remnant of mass  $m_S$ :

$$\xi_S = \frac{m_S - m_{\text{imp}}}{m_{\text{imp}}} = \frac{m_S}{m_{\text{imp}}} - 1. \quad (8)$$

In a hit-and-run collision, this second remnant is usually made largely of an impactor continuing downrange (i.e., a *runner*). Under most circumstances, mass is eroded from the body, such

that  $\xi_S < 0$ . Material not part of the two most massive remnants is considered debris; either resolved if it is part of a clump or unresolved if not. By analogy to the two largest remnants, debris production is characterized by its accretion efficiency  $\xi_D$ , which is its mass in units of impactor masses, or

$$\xi_D = \frac{m_D}{m_{\text{imp}}}. \quad (9)$$

Mass conservation requires that  $\xi_L + \xi_S + \xi_D = 0$ .

We then derive the orbital–dynamical parameters that are sufficient to set up the postimpact bodies on their new orbits following an identified collision in an  $N$ -body simulation. In this step, we use the radius of an equivalent body with the same mass and core mass fraction obtained from 1D calculations with the initial thermodynamic profile, using the results shown in Appendix A. This is to use the same radii as in the  $N$ -body. Using the postcollision radii would either create additional complexity or lead to the use of inconsistent radii. Having a consistent radii is important to ensure that the relative velocity and impact parameter are correct.

For the collisions with no second remnant, the velocity change (postimpact orbit) of the target is computed, where, to conserve momentum in the center-of-mass frame, it has the equal but opposite momentum of the debris. For the collisions with two major remnants after the initial collision, whether they end up being hit-and-run or graze-and-merge, we determine the final (outgoing) orbits of both the bodies. Since we do not include the preimpact rotation, we can assume the outgoing orbits are in the same plane as the incoming colliding bodies. The true anomaly is not determined or considered herein, as the initial location of the remnants after the collision will be prescribed according to the capabilities and design of  $N$ -body codes that will use our results and simulate the evolution of the remnants.

As in Emsenhuber et al. (2020), there are three additional orbital–dynamical parameters from the postimpact remnants that need to be computed to describe their orbits in sufficient detail for  $N$ -body simulations. We compute the orbital energy as

$$\epsilon' = \frac{v_{\text{coll}}'^2}{v_{\text{esc}}'^2} - 1 = \frac{v_{\infty}'^2}{v_{\text{esc}}'^2}, \quad (10)$$

where the primed ( $'$ ) quantities are computed from the properties of the largest and second remnants instead of the

target and impactor, respectively. The second equality is only valid in the case of an unbound orbit. Next, we compute the impact parameter

$$b' = \frac{h'}{R'_{\text{coll}} v'_{\text{coll}}} \quad (11)$$

where  $h'$  is the norm of the specific relative angular momentum vector. Finally, we have the shift of the orbit's pericenter

$$\Delta\varpi = \varpi' - \varpi. \quad (12)$$

Together, these orbital–dynamical parameters are sufficient to set up the postimpact orbits in an  $N$ -body simulation.

### 3. SPH Results

While there is a great wealth of information in this new, publicly available database to be leveraged, we present a few major conclusions that are worth highlighting. In particular, we find results related to remnant shape and the nature of debris as a function of total mass (reported in Table 2), which provide new insights into planet formation physics.

#### 3.1. General SPH Outcomes

The “regime” column in Table 2 is a flag that is automatically set from the analysis of a collision. It determines which properties of the simulation are returned. It is based on the number of significant remnants that are found, where a significant remnant is taken to be a body whose mass is at least 10% that of the impactor; considering the resolution of the simulations and the range of impactor-to-target mass ratios covered, this means that the significant remnants' classification is restricted to bodies of at least  $\sim 500$  SPH particles. The core mass fractions are only determined on significant remnants to avoid computing the statistics on bodies that have low numbers of SPH particles in the simulation and thus may be under-resolved. To accommodate graze-and-merge collisions, we provide two sets of outcomes for the simulations results: the first for the end state of the simulation and the other after only a single encounter. Depending on the type of collision, these may or may not be at the same time.

The different regime labels are as follows:

1. 727 entries or 58.2% are hit-and-run collisions where two unbound significant remnants exist. Per our definition of a significant remnant, this includes erosive hit-and-runs. Here, all the values are computed at the end of the simulation.
2. 218 entries or 17.4% are finished graze-and-merge collisions, where two significant remnants were temporarily gravitationally bound after first contact, and collided again thereafter in an accretion. Here, all the properties are determined first at apocenter and then at the end of the simulation.
3. 10 entries or 0.8% are unfinished graze-and-merge collisions, where the loopback orbital period is too long to evolve back to apocenter using SPH. Here, only the properties after the initial grazing encounter are determined, from the end state of the simulation.
4. 281 entries or 22.5% are collisions that result in a single major remnant. Here, the accretion efficiency and core mass fraction are determined only for the largest remnant, at the end of the simulation.

5. 14 entries or 1.1% are fully erosive events where no significant remnant exists. Only the accretion efficiencies (negative) are determined in this case, at the end of the simulation, as there is insufficient resolution to determine the other properties.

We observe that hit-and-runs are the most represented collisions in our database. This is consistent with the fact that these collisions are found for a wide range of impact angles for our favored impact velocities (Cambioni et al. 2019). Only 14 collisions (or 1.1% of the total) are supercatastrophic. This is unsurprising as they only occur at large velocities and low impact angles (Leinhardt & Stewart 2012), which our selected velocity distribution does not favor.

For unfinished grazing collisions—those labeled 3—we refrain from providing the final values. The final outcome is less relevant in the context of planetary formation, where the loopback orbit would be perturbed by third bodies, such as the Sun (Emsenhuber & Asphaug 2019b).

Due to the sensitivity on the impact angle and velocity, grazing collisions can have a diversity of specific outcomes even for similar initial conditions. Example end states include the capture of a substantial part of the runner as a satellite, or producing bodies that are far from their original hydrostatic equilibrium. We provide a selection of these outcomes in Figure 4 that will be discussed in more detail in the following sections.

#### 3.2. Capture as Satellite

Some of the peculiar outcomes of grazing collisions are the situations where most of the impactor (the runner) ends up forming a satellite orbiting the largest remnant. This is akin to a proposed scenario for the formation of the Pluto–Charon binary (Canup 2005) and some scenarios for the formation of our Moon (Benz et al. 1987). We remind the reader that a bound satellite—like any bound material—is considered part of the largest remnant in this study.

We also examine the long-term evolution of the satellite. If the satellite orbit is below the corotation radius, tidal friction will transfer angular momentum over time from the satellite's orbit to the spin of the central body, and it will spiral down on a timescale dictated by tidal dissipation and reimpact the central body. This results in the accretion of the majority of the satellite's mass and ejection of a fraction of the mass to release angular momentum. If the satellite is captured outside the corotation radius, then the opposite occurs, and angular momentum is transferred to the satellite's orbit, and it migrates farther out.

Examples of a satellite capture in our database are shown in Figure 4, left and center panels. One scenario is a collision with  $m_{\text{tar}} = 1.154 \times 10^{-4} M_{\oplus}$ , and  $\gamma = 0.657$ , and whereas the other is a higher-mass target  $m_{\text{tar}} = 1.618 M_{\oplus}$ , with  $\gamma = 0.376$ . Both occur near the mutual escape velocity ( $v_{\text{coll}}/v_{\text{esc}} = 1.02$  and  $1.00$ , respectively) at a grazing incidence ( $75^{\circ}1$  and  $79^{\circ}5$ , respectively), which is characteristic of satellite capture scenarios.

Figure 5 shows how multiple loopback encounters can lead to satellite capture. The blue thin curve plots the relative distance between the largest remnants versus time for the low-mass case in Figure 4 (left panel). The orange curve shows the same information for the high-mass case (center panel). The gaps in the curves indicate the moment the bodies are in



**Table 2**  
Description of the Table Headers for the Outcomes of SPH Simulations

Group	Quantity	Unit	Description
	Set		Name of simulation set
	Num		Simulation number in the set
Collision parameters	$m_{\text{tar}}$	$[M_{\oplus}]$	Mass of the target
	$\gamma$		Impactor-to-target mass ratio
	$Z_{\text{tar}}$		Target's core mass fraction
	$Z_{\text{imp}}$		Impactor's core mass fraction
	$\frac{v_{\text{coll}}}{v_{\text{esc}}}$		Relative velocity at initial contact
	$\theta_{\text{coll}}$	[deg]	Impact angle
	Regime		Automated classification (see main text)
End state	$\xi_L$		Accretion efficiency of largest remnant
	$\xi_S$		Accretion efficiency of second remnant
	$\xi_D$		Accretion efficiency of debris
	$\Omega_L$	[rad hr <sup>-1</sup> ]	Spin rate of the largest remnant
	$Z_L$		Largest remnant's core mass fraction
	$I_{Z,L}$		Largest remnant's moment of inertia factor along spin axis
	$I_{A,L}$		Largest remnant's volume-averaged moment of inertia factor
	$\xi_P$		Accretion efficiency of primary physical body
	$\gamma_P$		Mass ratio of the second to primary physical body
	$n_P$	[rad hr <sup>-1</sup> ]	Mean motion of the second physical body's orbit around the primary
	$N_{\text{res}}$		Number of resolved bodies
	$\frac{v_{\text{RMS}}}{v_{\text{esc}}}$		rms of the velocity of debris
		$\frac{v_{\text{RMS},\infty}}{v_{\text{esc}}}$	
After first enc.	$\xi_L$		Accretion efficiency of largest remnant
	$\xi_S$		Accretion efficiency of second remnant
	$\xi_D$		Accretion efficiency of debris
	$\Omega_L$	[rad hr <sup>-1</sup> ]	Spin rate of the largest remnant
	$Z_L$		Largest remnant's core mass fraction
	$I_{Z,L}$		Largest remnant's moment of inertia factor along spin axis
	$I_{A,L}$		Largest remnant's volume-averaged moment of inertia factor
	$\epsilon'$		Scaled orbital energy of the second remnant's orbit
	$b'$		Impact parameter of the second remnant's orbit
	$\Delta\varpi$	[rad]	Shift of pericenter of the second remnant's orbit
	$\Omega_S$	[rad hr <sup>-1</sup> ]	Spin rate of the second remnant
	$Z_S$		Second remnant's core mass fraction
	$I_{Z,S}$		Second remnant's moment of inertia factor along spin axis
	$I_{A,S}$		Second remnant's volume-averaged moment of inertia factor
	$N_{\text{res}}$		Number of resolved bodies
		$\frac{v_{\text{RMS}}}{v_{\text{esc}}}$	

**Note.** The first eight columns of the table are the initial conditions while the rest summarize the outcomes of the simulations.

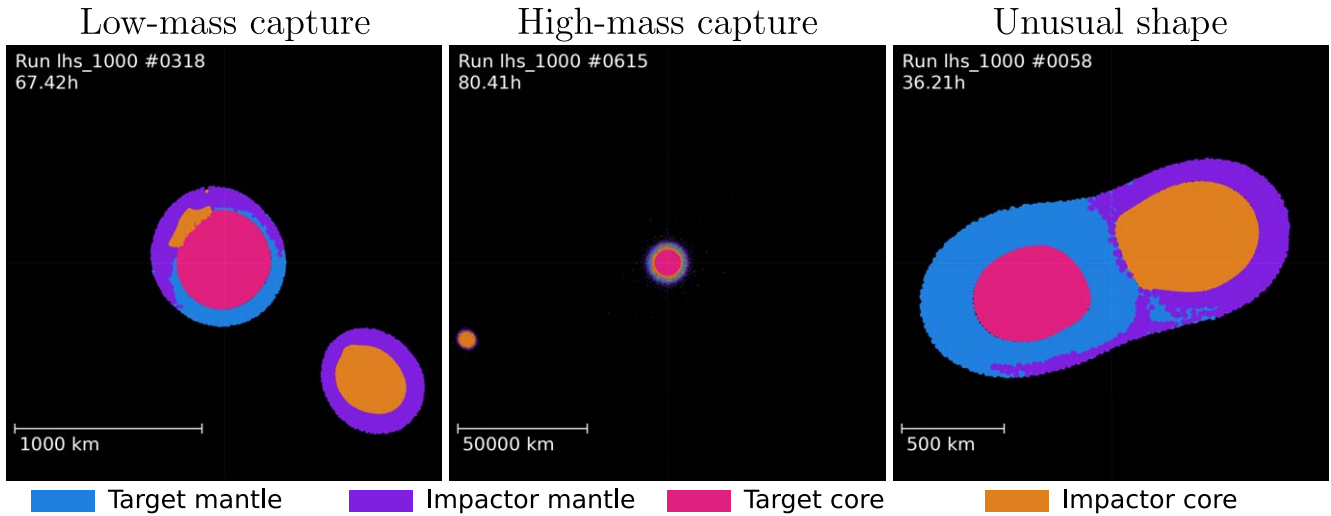
(This table is available in its entirety in machine-readable form.)

physical contact. An *arm* (continuous link of material) usually connects the two bodies until the bodies reach a distance of up to several mutual radii. During this time, the friend-of-friend algorithm cannot identify the central body and satellite as separate objects, and the orbit cannot be determined. The results show that the pathway toward capture differs between the two collisions.

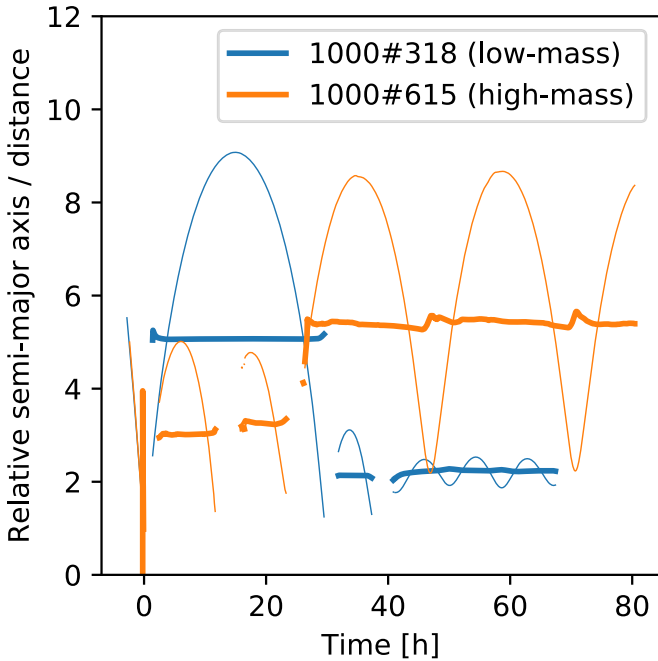
For the low-mass case, each close encounter leads to dissipation in the satellite, which decreases the apocenter. A little orbital angular momentum is transferred onto the target, likely due to the inclusion of material strength in our simulations, which limits its deformation. Consequently, the orbit becomes more circular. For the high-mass case, both the satellite and target are heavily deformed during each loopback encounter, and much of the impactor's material is transferred onto the target during successive encounters. This raises the

pericenter when tidal deformation induces a torque on the remnant of the impactor, which increases its angular momentum.

In the high-mass scenario in Figure 5 (orange lines), it is the third encounter at  $\sim 23$  hr that is the most effective at establishing a relatively stable orbit for the impactor. The satellite's orbit at the end of our simulation is eccentric (see difference between thin curve and thick curve, representing radius and semimajor axis of the orbit respectively), resulting in the satellite spending most of the time many radii away from the target. On its subsequent close approaches, tidal destruction is avoided due to the sufficiently high pericenter. We evolve these simulations long enough for the satellite to make multiple pericenter passages to check that satellites are not destroyed during this phase.



**Figure 4.** Final state of a selection of three, low-velocity grazing simulations, each with different masses of the colliding bodies (left,  $m_{\text{tot}} \approx 1.9 \times 10^{-4} M_{\oplus}$ ; middle,  $m_{\text{tot}} \approx 2.2 M_{\oplus}$ ; right,  $m_{\text{tot}} \approx 4.0 \times 10^{-4} M_{\oplus}$ ); note the scale bars. Left and center panels show capture as satellite; the right panel shows unusual body shape. Colors represent material and origin body, as indicated in the legend: blue for target’s mantle, purple for impactor’s mantle, red for target’s iron core, and yellow for impactor’s core.



**Figure 5.** Time evolution of the relative semimajor axis (thick lines) and distance (thin lines) of the two largest remnants for grazing collisions resulting in satellite capture shown in Figure 4 (left and central panels). The values are given in terms of the sum of the radii of equivalent bodies with the given mass and core mass fraction, according to the results from Appendix A. Gaps indicate physical contacts between the two bodies. The satellite resulting from the collision between the low-mass bodies remains on a low-eccentricity orbit close to the primary, while the satellite in the high-mass collision is left on an eccentric orbit farther out.

For all capture scenarios, we report the mass ratio of the orbiting pairs in Figure 6. For this analysis, we ensure that there is at least one pericenter passage of the loopback orbit. We find that the satellites with more than  $\sim 10\%$  of the mass of the primary occur around the primaries with masses below  $2 \times 10^{-3} M_{\oplus}$ . Friction likely aids in the capture of such massive moons, as exemplified by the low-mass capture scenario shown in Figures 4 and 5 where the satellite remains at about twice the

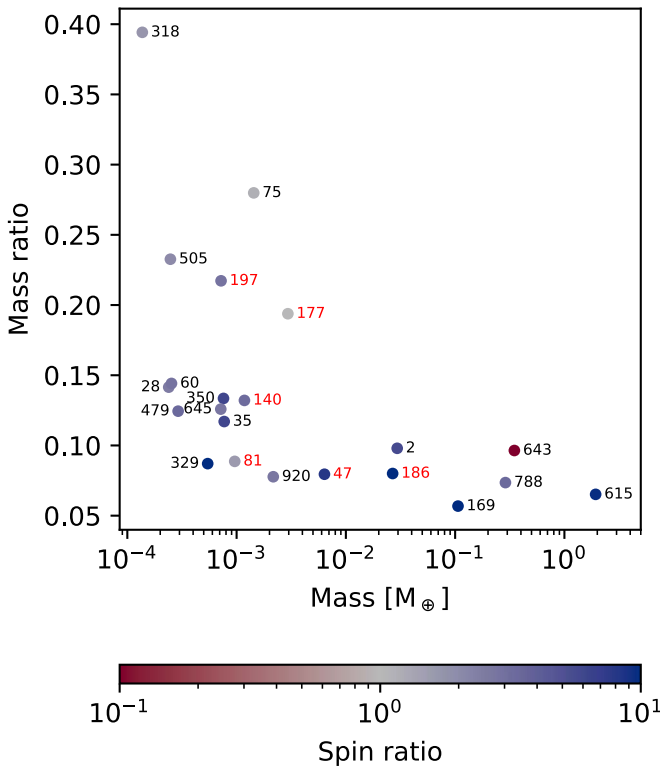
mutual radii. Without friction, such a satellite would likely be disrupted by tidal forces, as it lies inside the Roche limit of a fluid body. Satellites are obtained around more massive bodies, but are of a small mass ratio with respect to the postimpact target. This outcome is consistent with prior work (Nakajima et al. 2022) on the limited ability of larger planets to form fractionally massive moons; however, the result here is for collisional capture and is a direct aspect of the giant impact phase itself. Also, as a general rule, fractionally massive surviving clumps are less common at larger scales of giant impacts because of the greater relative energies involved (Asphaug & Reufer 2014).

### 3.3. Final Body Shapes

An outcome of low-mass, graze-and-merge collisions is the emergence of nonspherical bodies. Friction plays a role in sustaining shapes that are far from hydrostatic equilibrium (e.g., Tanga et al. 2009; Sugiura et al. 2018; Jutzi et al. 2019). For example, in Mars-sized planets and smaller, friction forces can counteract gravity, allowing for the stranding of impact cores in mantle (Emsenhuber et al. 2018). On geologic timescales, these nonhydrostatic shapes and irregular core structures might relax, lowering the moment of inertia and changing the spin state.

An example is shown in the third panel of Figure 4, which shows a cross section of simulation 58 of the 1000 simulation set. These scenarios are analogous to hypotheses for the formation of contact binaries such as 486958 Arrokoth (Stern et al. 2019) and 67P Churyumov-Gerasimenko (Jutzi & Asphaug 2015; Jutzi & Benz 2017), although those scenarios invoke crushable solid bodies while our models are at much larger scales and consider bodies that start out with metallic cores. Here, it can be seen that the two bodies only slightly mix, and the cores remain separated.

The nonspheroidal shape of the final body, in these extreme cases, is maintained not only by material strength but also from the spin induced by the collision. The production of nonspherical shapes generally requires grazing mergers at velocities close to  $v_{\text{esc}}$ . In these cases, the two cores do not



**Figure 6.** Mass ratio of the satellite-to-primary pairs ( $\gamma_p$ ) as function of the primary mass (computed as  $m_{\text{tar}}(1 + \gamma_p)$ ), for all simulations ending in coorbiting secondary remnants. The number in the plots refers to the corresponding run numbers: black for those in the 1000 simulation set, and red for those in the 250 simulation one. The color scale represents the ratio between the orbital frequency of the satellite and the spin rate of the primary ( $n_p/\Omega_L$ ), which is 1 for synchronous orbit and must be  $<1$  for stable, outward tidal migration.

immediately intersect during the initial collision, yet there is enough dissipation in the initial encounter that the bodies remain bound, rotate quickly, and maintain a spin-supported shape. A caveat of our analysis, however, is that bodies are not spinning before the collision, so preimpact spin is not taken into account.

We detected bodies with nonspherical shapes by computing the moment of inertia of all massive remnants and comparing the values of an equivalent nonrotating spherical body of the same composition under a hydrostatic equilibrium. Figure 7 is a scatter plot of the ratio between these values as a function of mass. Ratios close to 1 indicate bodies whose internal structure is comparable to that of initial bodies.

The analysis reveals several features. The remnants from hit-and-run collisions, where  $\xi_L \approx 0$  (in gray in the figure), in the end have their global structures barely affected by the collision. Accretionary or erosive collision usually leads to an increased moment of inertia, due to spin-induced deformation (equatorial bulge), for accretionary events, or pressure release, for erosive events. Finally, we see two regions that exhibit moment of inertia factors larger than 2: one for small bodies (below  $\sim 2 \times 10^{-3} M_{\oplus}$ ) and another for large bodies (around  $1 M_{\oplus}$ ).

For small bodies below  $\sim 2 \times 10^{-3} M_{\oplus}$  (or 2000 km diameter) for the modeled terrestrial compositions, the high moment of inertia factor reflects the peculiar shapes. In Figure 4 (right panel), the contact binary would plot toward the top left of Figure 7. It is not a unique case in our database; however, this compound-core scenario only happens for small

bodies. The specific compound-core outcome and other aspects of shape depend on the influence of friction, especially in the silicate material that supports the denser core material, where for consistency with past and ongoing research we have adopted the coefficients of friction in Table A1 of Collins et al. (2004). The upper limit of  $\sim 2 \times 10^{-3} M_{\oplus}$ , which coincides with the limit for satellites with large mass ratios (Section 3.2), suggests that material strength has global-scale effects up to that mass. We note that this upper limit depends on the constitutive strength model and its parameters. As a consequence, the bodies with significantly higher initial temperature or lower von Mises plastic limit  $Y_m$  would exhibit the transition to more spherical shapes at smaller sizes.

In the case of the accretionary events in Figure 7, the increased moment of inertia factor is due to a combination of two effects. The predominant one is the spin, which causes a rotational bulge. Unlike the peculiar shapes we just discussed, however, the shape of these bodies is nearly symmetric about their spin axis. Other effects, limited to the masses above  $\sim 1 \times 10^{-2} M_{\oplus}$ , are thermal in nature. This is because mergers, especially those between similar-mass bodies (that is, for  $\gamma \approx 1$ ), release the most binding energy (Asphaug & Reufer 2013; Carter et al. 2020), increasing temperature. These bodies will also preferentially form distended vapor-rich outer structures due to shock heating in the largest remnant (Lock & Stewart 2017). This in turn expands the bodies relative to a starting configuration (Lock & Stewart 2017) and increases their moment of inertia.

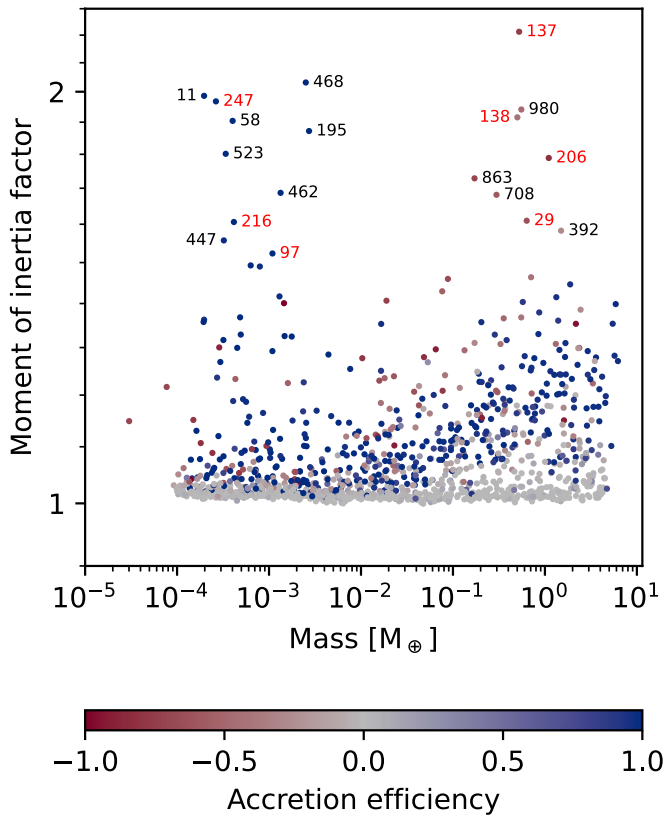
Strongly erosive collisions ( $\xi_L \approx -1$ ) also exhibit similar expanded states because of pressure release, as in these cases usually only the most central part of the body remains. Even-lower values of  $\xi_L$  are obtained in supercatastrophic disruptions, and show the same effect, but they are not plotted because the bodies are comprised of relatively few particles.

### 3.4. Debris Properties

One of the main objectives of our new simulations is to better characterize the debris to represent it with known accuracy in planet formation studies. For this, we need to understand not only debris production efficiency  $\xi_D$  but also its velocity and the size distribution.

#### 3.4.1. Fragment Sizes and Number

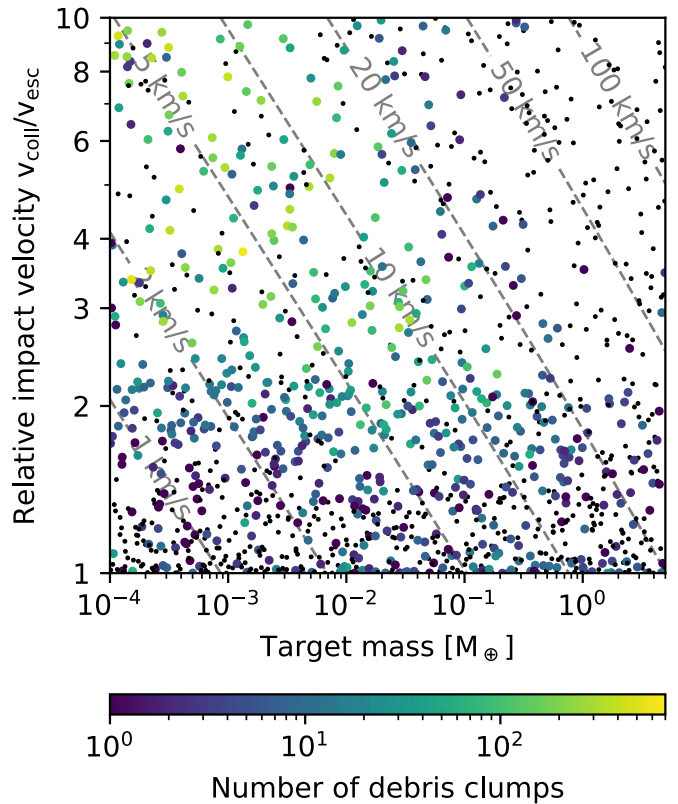
Resolvable fragments are clumps that are obtained in the debris field, identified in the same way as we identify the largest and second remnants, by performing friend-of-friend and gravity searches in that order. The smallest mass that can be resolved depends on the numerical resolution of the simulations; however, the size–frequency distribution as a whole is much less sensitive to numerical resolution (Appendix C.2). We assume that a fragment requires at least ten particles to be meaningful, for our database with roughly  $1 \times 10^5$  particles per simulation, and thus, only the clumps larger than  $\sim 1 \times 10^{-4}$  the total mass are regarded as significant. The particles in the smallest clumps have less than a full neighbor-list (approximately 50 particles), and while this means that SPH forces are not computed accurately, the self-gravity that causes clumping is accurate. We have therefore verified that the results discussed here remain if we change the fragment resolution limit from 10 to 50 to 100 SPH particles.



**Figure 7.** The moment of inertia factor of the postimpact largest remnant (y-axis) as a function of its mass, with accretion efficiency of the largest remnant ( $\xi_L$ ) in color. The moment of inertia factor is the moment of inertia divided by that of a hydrostatic body with the same mass and core mass fraction, as outlined in Appendix A. Both axes use logarithmic scales. To avoid catastrophic events, only collisions where  $\xi_L > -1$  are shown. The simulation number is shown when the moment of inertia factor is larger than 1.5. A black label is used for runs in the 1000 simulation set, and red for those in the 250 simulation set.

Fewer than half of our modeled collisions produce fragments 10 SPH particles in size or larger, besides the two largest remnants. No debris fragment is larger than 10% of the total mass, and only a few collisions produce any fragment larger than 1% of the total mass. The only exception is the second remnant, which we do not include in the debris size–frequency distribution. Most massive fragments are found in either graze-and-merge or disruptive hit-and-run collisions, where the impactor is broken in multiple fragments.

It is extremely difficult for  $N$ -body simulations to track debris directly, as the number of bodies ( $N$ ) would increase dramatically whenever there is a collision. Thus, debris need to be treated differently from the two largest remnants, with a more statistical approach, such as tracer particles (e.g., Levison et al. 2012; Walsh & Levison 2019), bodies of a fixed mass (e.g., Chambers 2013), or as a background surface density (e.g., Carter et al. 2015). Possible statistical approaches include looking for correlations between debris size and velocity (Section 3.4.2) or fitting a size–frequency distribution to the bodies in the debris field (as in Appendix C). Both require that at least a few clumps are reliably determined, and we thus investigated the conditions that are favorable for the production of many resolved debris clumps. We find that a well-populated size distribution of debris is generated in graze-and-merge and shallow angle, erosive collisions. Graze-and-merge ejects clumps due to the large angular momentum involved



**Figure 8.** Scatter plots of the number of resolved ( $>10$  SPH particle) fragments, excluding the largest and second remnants, found in our simulation database as function of target mass and relative velocity  $v_{\text{coll}}/v_{\text{esc}}$ . The color represents number of fragments in a single simulation. Collisions not producing any fragment are shown by smaller black circles. Although these points include simulations with various impactor-to-target mass ratios, we show absolute impact velocities for the target mass alone, assuming a core mass fraction of 30%. It can be seen that the number of resolved clumps decreases with increasing target mass.

(Asphaug & Reufer 2013), and the shallow-angle erosive collisions ( $v_{\text{imp}}/v_{\text{esc}} \geq 2$ ) naturally produce ample amounts of debris.

We further find that low-mass bodies ( $m_{\text{tar}} \lesssim 0.1 M_{\oplus}$ ) are favored for the production of a large number of resolved debris clumps. We illustrate this in Figure 8, where the number of fragments for each collision is shown as functions of the target’s mass  $m_{\text{tar}}$  and the relative collision velocity  $v_{\text{coll}}/v_{\text{esc}}$ . Here, we observe that there is a strong transition at  $v_{\text{coll}} \approx 20 \text{ km s}^{-1}$ , which corresponds to  $m_{\text{tar}} = 0.1 M_{\oplus}$  for  $v_{\text{coll}}/v_{\text{esc}} \approx 4$ . While the maximum number of fragments identified in any collision is larger than 400, none of the collisions above  $20 \text{ km s}^{-1}$  has more than 40 resolvable fragments. At even larger velocities ( $v_{\text{coll}} \approx 50 \text{ km s}^{-1}$ ), only few collisions report any resolvable fragment; nearly all the debris are found in unresolved particles. This effect could be due to stronger shocks or vaporization being more common at a larger absolute velocity, both of which can prevent the formation of debris clumps.

Our debris results should be taken with caution. First, we use a constant number of particles for our simulations, which means that the minimum mass of resolved fragments scales with the target mass. Therefore, fragments of a given absolute size may be resolved in collisions involving low-mass bodies and not with high-mass bodies. Second, high-velocity



collisions involve extreme amounts of heating, where numerical effects in SPH may be exacerbated.

### 3.4.2. Debris Dynamics

Collision after collision, the fraction of the total mass that becomes debris may become significant with time (Emsenhuber et al. 2020). These debris are reaccreted by the large bodies on timescales that are millions of years (Jackson & Wyatt 2012), although some works suggest that the mass fraction is never significant at any given time (Carter et al. 2015). As such, debris orbital evolution must be modeled consistently during  $N$ -body calculations. For this purpose, we report postimpact velocities and their correlation with remnant sizes.

First, we compute the rms of the debris velocity distribution relative to the center of mass, reported as  $v_{\text{rms}}$  in Table 2. This value is computed from their specific kinetic energy. Because the debris are in the gravitational potential of the largest remnants, this value is time-dependent. To account for this, we also report the velocity of the debris calculated using the total orbital energy (kinetic and potential) of the debris field ( $v_{\text{RMS},\infty}$ ).

We then check for any correlation between debris size and their velocity using the Spearman’s rank correlation coefficient. We find that, over the 1250 simulations presented here, 591 have a sufficient number of resolved debris such that a correlation can be computed. Of those, 62 have  $p$ -values consistent with there being no correlation lower than 1%, indicating that the correlation is statistically significant. Of those, 10 have a positive correlation with debris size and velocity, and 52 have a negative correlation. Thus, only a minority (about 10%) of our simulated collisions with resolved debris fields show correlations between debris size and velocity. In sum, we find that debris velocity and mass are not correlated in a systematic way.

To broadly describe debris velocity fields in our simulations, we compute the mean velocity of the debris at infinity in terms of the relative velocity of the precursor bodies at initial contact (see Figure 9). The left panel shows that the collisions producing the most debris (shown toward the right of the panel) have velocities that trend near a value of about 0.45, with a significant dispersion around the mean due to the other parameters (mainly the mass ratio  $\gamma$ ). The collisions that produce a smaller fraction of debris exhibit the greatest dispersion and a lower mean velocity; however, capturing the debris velocity accurately is less important for those events, from the point of view of planet formation, as their debris constitutes a much smaller fraction of the total mass. The right panel of Figure 9 shows that the debris velocity in terms of  $v_{\text{coll}}$  is largely independent of  $v_{\text{coll}}/v_{\text{esc}}$  of the incoming bodies. This indicates that  $v_{\text{coll}}$  is a better reference for the debris velocity rather than, for instance,  $v_{\text{esc}}$ .

As a rule of thumb, we find that the mean debris velocity at infinity  $v_{\text{RMS},\infty}$  is about half that of the collision velocity  $v_{\text{coll}}$ . This can be significantly greater than  $v_{\text{esc}}$  of the largest remnant. Thus, the assumption that the debris always leave at near-escape velocities, which is made by Chambers (2013) and other works based on their prescription (e.g., Clement et al. 2019), severely underestimates debris velocity. The consequence is an artificial damping of the velocities of a late-stage accreting planetary system and, as a potential consequence, an artificial limitation to the exchanges of materials between

formation regions. The debris velocities we obtain differ significantly from the ejecta velocity distributions of collisions reported in Leinhardt & Stewart (2012); this is understandable for two reasons. First, their results rely on a discrete element model instead of solving the Euler equations, and thus are idealized representations of the collisional physics at the high velocities of planet formation. Second, and more fundamentally, their results are computed for the catastrophic disruption regime (their Figure 5), in which the escaping velocities are governed by impact ejection mechanics rather than gravitational accretion dynamics. Our focus is on the regimes relevant to planetary accretion, so our database barely extends into their regime of catastrophically disruptive collisions (Table 1). Debris are not blasted from the target so much as they are flung out by the complex gravitational interactions of two similar-mass bodies, which explains why the debris velocities in our database are governed by  $v_{\text{esc}}$ .

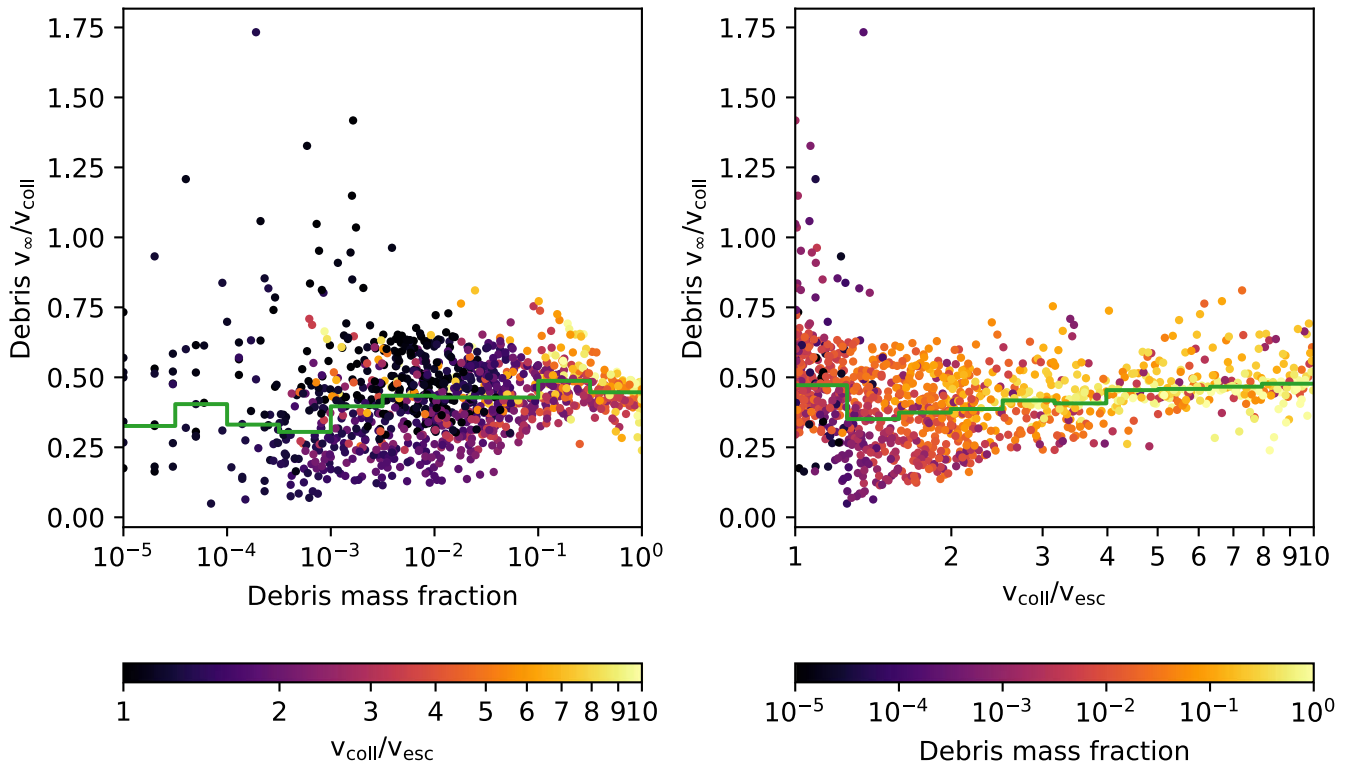
If the relative velocity at infinity  $v_{\text{RMS},\infty}$  is half of  $v_{\text{coll}}$ , then for collisions where  $v_{\text{coll}}/v_{\text{esc}} < 2\sqrt{3}/3 \simeq 1.15$ , the debris velocity at infinity  $v_{\text{RMS},\infty}$  can be larger than the relative velocity of the incoming bodies at infinity  $v_{\infty}$ . This is not paradoxical, as it happens in the case of strong gravitational torques by the more massive bodies in a close approach, and the subsequent ejection of arms of material. Collisions with such low impact velocities usually produce relatively little debris in any case, as can be seen by the color scale on the right panel of Figure 9.

Several simulations exhibit debris with negative total energy and are not represented in Figure 9. This is confined to weakly interacting hit-and-runs resulting in two large remnants with debris that resides between them in space. These debris experience the gravitational potential of both remnants, which are in opposite directions. Thus, it is not a contradiction that the debris are not bound to either remnant while they have negative energy. These cases result in debris with low relative velocity (lower than that of the runner). This peculiar behavior should not be affecting the overall analysis, as they only occur in simulations with a low-mass fraction of debris. While the major characteristics of the largest remnants are well determined in this study, a more informed assessment of debris will require higher-resolution simulations running to a much later time.

## 4. Summary and Conclusion

We present a new database of 1250 giant impact simulations relevant to terrestrial planet formation, spanning  $1 \times 10^{-6} M_{\oplus}$  to  $5 M_{\oplus}$  bodies, performed using SPH including friction. The data set has six free parameters: the target mass, the impactor-to-target mass ratio, the core mass fractions of target and impactor, and the impact velocity and angle. The preimpact rotation was not included. Some simulations were performed with a resolution increase by a factor of 10. Only small differences were found between our nominal resolution (100k SPH particles) and the high-resolution case (1M SPH particles).

In previous work, Cambioni et al. (2019), Gabriel et al. (2020), and Emsenhuber et al. (2020) reported on the accretion efficiencies, the change in composition, and the relative orbits of the largest remnants for giant impacts, based on a previous database (Reufer 2011). As part of this new and more comprehensive database, we additionally calculate the moments of inertia of the final bodies, and determine the presence of major satellites.



**Figure 9.** Debris velocity at infinity given in terms of the relative velocity between the initial bodies (target and impactor) at initial contact for all the simulations of this study. Left panel shows the values as function of the mass fraction of debris at the end and color coded by the initial  $v_{\text{coll}}/v_{\text{esc}}$ . Right panel shows the same data with the horizontal axis and color code inverted. The step-wise line shows the median of the values in the given range.

All simulations apply a solid friction model modulated by temperature-dependent yielding. As expected, the effects of strength are the most significant at smaller sizes. Friction acts to reduce differential flow velocities, and this can substantially change the gravitational and angular momentum dynamics. Because of this, large satellites, several percents of the target mass, are found more commonly in the aftermath of collisions up to  $\sim 2 \times 10^{-3} M_{\oplus}$ . Earth-mass and larger collisions do not tend to feature captured impactors as satellites. Low-mass bodies also have a diversity of extended, nonspherical postimpact shapes and interior structures. These observed strength-supported features are consistent with previous studies on similar-sized collisions at smaller scales than those examined here (e.g., Richardson et al. 2009; Jutzi & Benz 2017). One possible example of such a strength supported structure is asteroid Pallas, over 500 km diameter, whose shape substantially deviates from equilibrium considering its current rotation (Marsset et al. 2020).

We also present a systematic characterization of debris produced in giant impacts. We find that debris fragments are very small, in most cases, compared to the total mass of colliding bodies (less than 1%). Also, contrary to the common assumption (e.g., Chambers 2013) that debris escape at a low velocity, we find that the the mean debris velocity at infinity from giant impacts is about half of the relative velocity of the incoming objects, with a dispersion of  $\sim 50\%$  from the mean. We did not find any systematic correlation between debris size and velocity, so an assumption that the two are independent is sufficient. Such simple scaling could help future  $N$ -body-based evolution models to incorporate realistic debris dynamics.

## Acknowledgments

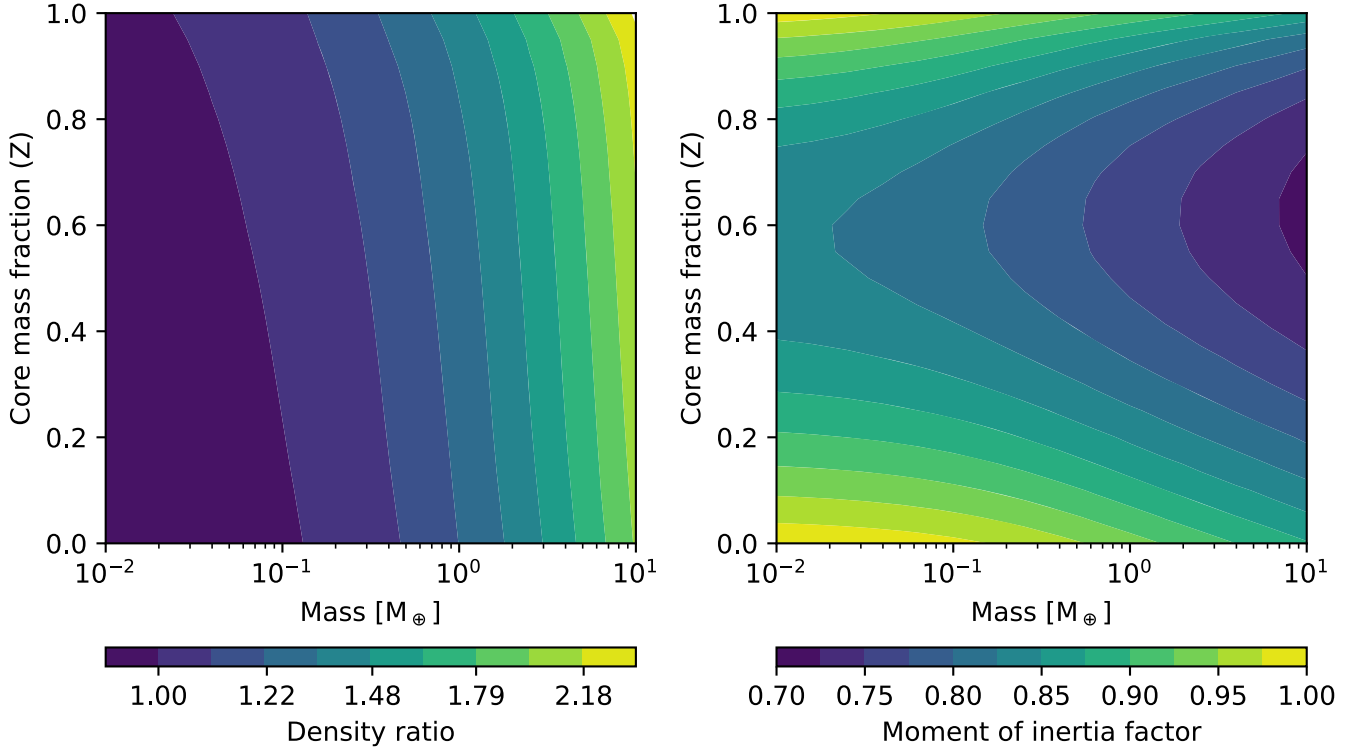
We thank the two anonymous reviewers, whose comments and suggestions helped improve this manuscript. A.E., E.A., S. R.S., and R.M. acknowledge support from NASA under grant 80NSSC19K0817 and the University of Arizona. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 362051796. S.C. acknowledges support from the Crosby Distinguished Postdoctoral Fellowship of the Department of Earth, Atmospheric and Planetary Sciences of the Massachusetts Institute of Technology. T.S.J.G. acknowledges support from the U.S. Geological Survey Astrogeology Science Center. An allocation of computer time from the UA Research Computing High Performance Computing (HPC) is gratefully acknowledged. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

*Software:* SPHLATCH (Reufer et al. 2012; Asphaug & Reufer 2013, 2014; Emsenhuber et al. 2018; Ballantyne et al. 2023), matplotlib (Hunter 2007), SciPy (Virtanen et al. 2020), NumPy (Harris et al. 2020).

## Appendix A

### 1D Calculations: Bulk Density and the Moment of Inertia

To ensure that the remnants can be treated consistently from preimpact to postimpact in an  $N$ -body simulation, we use the value for the radius that an initial body with the same mass and core mass fraction would have in our simulations. For this purpose, we compute a grid of 1D radial structures using the same approach as for the initialization of SPH collisions. Once these models are computed, we



**Figure 10.** Density ratio with respect to the reference value from the EOS (left) and moment of inertia factor  $I/I_{\text{ref}}$  (such that 1 is a uniform density sphere; right) as function of body mass and core mass fraction.

compute not only the bulk density (which is needed for the radius) but also other related properties, such as the moment of inertia.

The results of these calculations are provided in Figure 10. To leave out the changes due to variable composition (the core mass fraction) and focus on the internal compression due to self-gravity, we show the ratio of the bulk density to the reference density from the EOS we are using for the SPH simulations: ANEOS for the iron core (with a reference density of  $7.65 \text{ g cm}^{-3}$ ) and a modified version of the same (Stewart et al. 2020) for forsterite in the mantle (with a reference density of  $3.32 \text{ g cm}^{-3}$ ). For a given core mass fraction  $Z$ , the reference density  $\rho_{\text{ref}}$  is computed as

$$(\rho_{\text{ref}}/\text{g cm}^{-3})^{-1} = (1/7.65)Z + (1/3.32)(1 - Z). \quad (\text{A1})$$

The calculations were performed for bodies down to  $1 \times 10^{-5} M_{\oplus}$ . Figure 10 shows only the results for bodies larger than  $1 \times 10^{-2} M_{\oplus}$  because the smaller objects exhibit negligible compression. Their density is slightly below the reference value of the EOS because of the thermal expansion due to the internal energy. One outcome of the calculations is that the iron core is more compressible than the mantle in the EOSs that we use. This results in slightly different core radii-to-total radius ratios in bodies with different masses but otherwise the same core mass fraction.

The right panel of Figure 10 depicts the ratio of the moment of inertia of the computed bodies with respect to the moment of inertia of an equivalent body with the same mass and radius, but assuming a constant density,

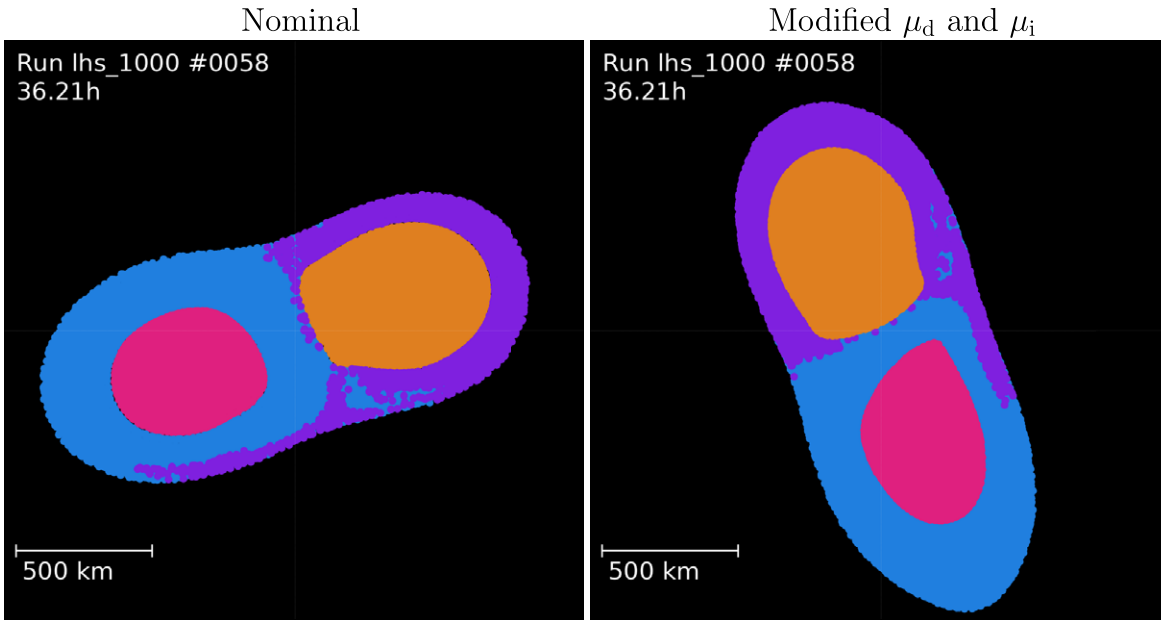
$$I_{\text{ref}} = \frac{2}{5} m r^2. \quad (\text{A2})$$

The bodies with intermediate core mass fractions have the smallest values of the moment of inertia ratio, as expected. Nevertheless, the effect of the different compressibility of the core and mantle is also noticeable here, as the ratio is smaller for a given mass at  $Z = 1$ , than at  $Z = 0$ . We also computed the ratio of the potential energies, but since the results are very similar to the moment of inertia, they are not shown here.

## Appendix B Effect of the Solid Model Parameters

To check how our choice for the values of  $\mu_d$  and  $\mu_i$  affects our results, we perform one simulation where we altered the values of  $\mu_d$  and  $\mu_i$ . We selected simulation #0058 of `1hs_1000`, as it is the one with the smallest bodies, lowest velocity, and an oblique impact angle, which are those most affected by the inclusion of strength in our model. The nominal values for the parameters are  $\mu_d = 0.8$ , and  $\mu_i = 2$ , as discussed in Section 2.4. For the altered values, we selected  $\mu_d = 0.63$ , and  $\mu_i = 1.58$  following Johnson et al. (2016) and references therein.

The final state of the two simulations is provided in Figure 11. The two show essentially the same outcome with an elongated body and separate cores. Differences at smaller scales nevertheless appear, such as an even more elongated shape and thinner separation between the cores in the simulation with altered parameters. Thus, we find that the precise choice of the values for  $\mu_d$  and  $\mu_i$  parameters of the solid model does not affect the global outcome of the simulations. The results presented in this work are therefore robust in this regard.



**Figure 11.** Final state of simulation #0058 with different values for parameters of the solid model. The left panel shows the nominal case and is the same as the right panel of Figure 4. The right panel shows the same initial conditions, but with  $\mu_d = 0.63$ , and  $\mu_i = 1.58$ .

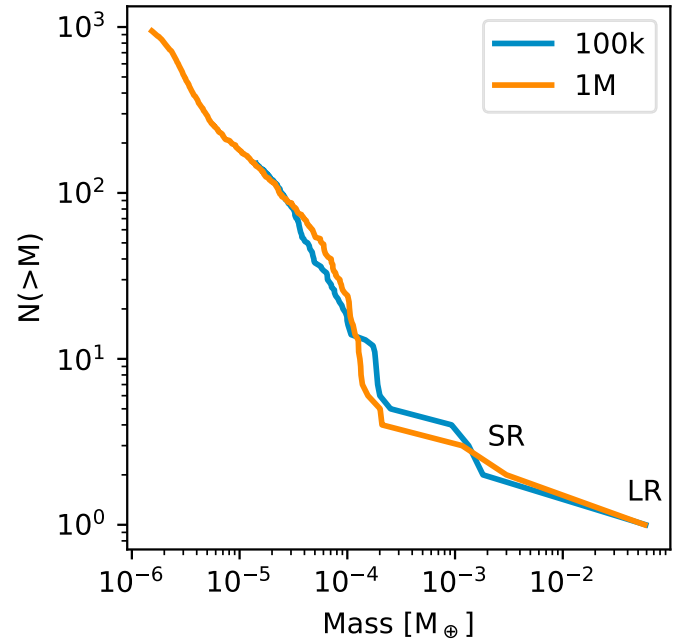
### Appendix C Effect of SPH Resolution

The resolution of SPH simulations is set by the number of particles (resolution) in the simulation. To check whether our results are affected by resolution, we performed five additional runs with a different number of particles but otherwise identical SPH parameters. For this analysis, we predominantly focus on metrics related to debris, namely the clumps or fragments that are made of relatively few SPH particles.

#### C.1. Global Results

We provide a comparison of the global outcomes of five simulations in Tables 3 and 4. The former represents the outcomes after the first contact (i.e., the first phase of a graze-and-merge collision), and the latter represents the final outcome. The columns provided are the same as those in Table 2, except they include a new parameter,  $N$ , for the SPH resolution (particle count), and we removed the outcome  $n_p$  since satellites are not produced in these collisions. Furthermore, we did not include the number of resolved clumps  $N_{\text{res}}$  in these tables, as we analyze this later in this section.

The simulations for this resolution study were selected to represent a variety of outcomes from the overall database. Most of the metrics that we are studying are similar whether the number of particles  $N$  is  $1 \times 10^5$  or  $1 \times 10^6$ . The perfect merger collision, shown in the first two rows in each table, shows nearly identical results in all outcomes as a function of resolution. For the others, the accretion efficiencies show differences in the range of 0.01–0.03, which can equate roughly to less than a few percent of total mass depending on the impactor mass. The final rotation rate and core mass fraction of the largest remnant are also very similar across resolution in the final state. For the intermediate state, the rotation rate of the largest remnant seems to be somewhat variable, showing  $\sim 20\%$  relative deviation in some cases. The debris velocities are also more uncertain, in two cases with



**Figure 12.** Comparison of the size–frequency distribution of the remnants for collision 566 of the 1000 simulation set for two different SPH resolutions, as indicated in the legends. The two texts “LR” and “SR” denote the largest and second remnants, respectively.

differences of up to  $0.2v_{\text{esc}}$ . These differences, however, do not generally exceed the statistical variations of these quantities in the database under small variations in the impact angle or velocity. Additionally, as noted, in cases where the core mass fractions of the initial bodies are very large, the mantle materials are less well resolved, and thus may be subject to SPH challenges resolving material discontinuity across the core–mantle boundary. While these cases might be expected to show greater variations as a function of  $N$ , in our resolution comparisons with the largest core mass fractions, we do not observe large deviations in the composition or other properties



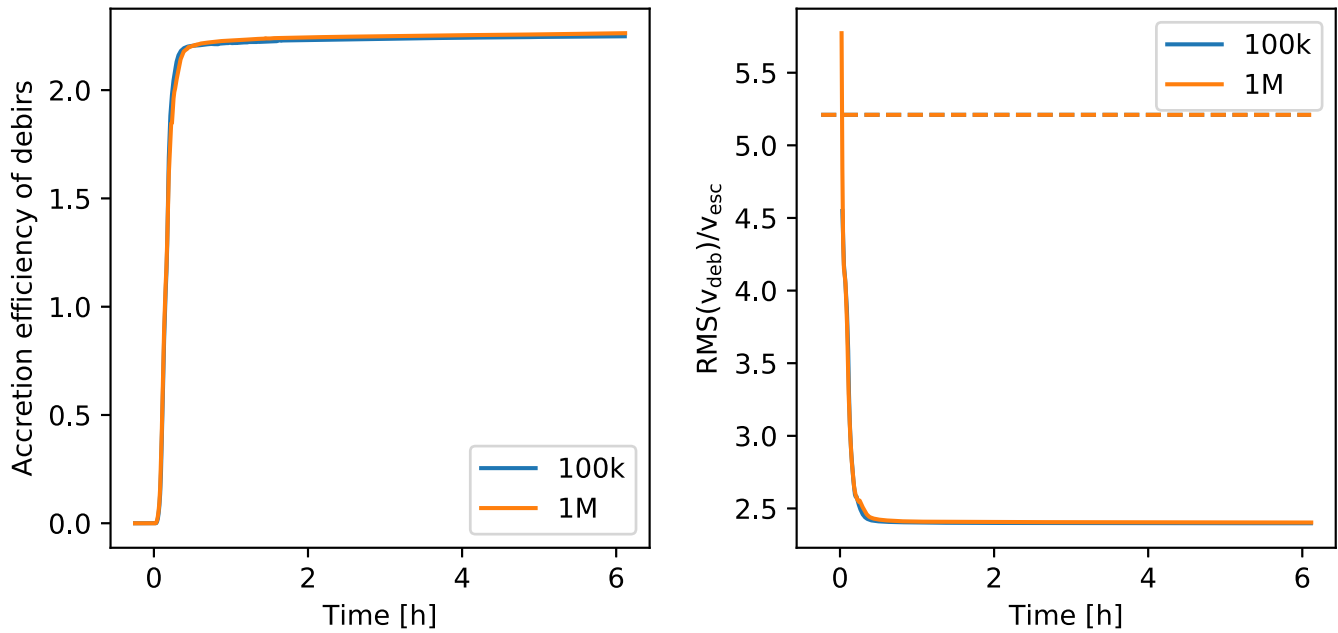
**Table 3**  
Intermediate Outcomes of Five Representative Collision Scenarios Computed at Two Different SPH Resolutions,  $N = 1 \times 10^5$  (from the Database) and  $N = 1 \times 10^6$ , Shown after First Contact

Collision Parameters							After First Enc.													
$m_{\text{tar}}$ ( $M_{\oplus}$ )	$\gamma$	$Z_{\text{tar}}$	$Z_{\text{imp}}$	$\frac{v_{\text{coll}}}{v_{\text{esc}}}$	$\theta_{\text{coll}}$ (deg)	$N$	$\xi_L$	$\xi_S$	$\xi_D$	$\Omega_L$ (rad hr $^{-1}$ )	$Z_L$	$I_{ZL}$	$I_{AL}$	$\epsilon'$	$b'$	$\Delta\varpi$ (rad)	$\Omega_S$ (rad hr $^{-1}$ )	$Z_S$	$I_{ZS}$	$I_{AS}$
$1.72 \times 10^{-4}$	0.961	0.371	0.284	1.169	51.2	$1 \times 10^5$	0.004	-0.004	0.000	0.840	0.369	1.127	1.043	-0.196	0.892	-0.687	0.834	0.281	1.135	1.047
$1.72 \times 10^{-4}$	0.961	0.371	0.284	1.169	51.2	$1 \times 10^6$	0.004	-0.004	0.000	0.734	0.368	1.114	1.036	-0.147	0.881	-0.640	0.717	0.281	1.111	1.037
$7.14 \times 10^{-4}$	0.319	0.192	0.294	1.837	30.3	$1 \times 10^5$	0.209	-0.321	0.112	0.844	0.200	1.107	1.039	0.334	0.719	-0.979	1.579	0.312	1.174	1.063
$7.14 \times 10^{-4}$	0.319	0.192	0.294	1.837	30.3	$1 \times 10^6$	0.208	-0.354	0.146	0.764	0.201	1.108	1.032	0.350	0.699	-0.994	1.239	0.324	1.120	1.045
$1.01 \times 10^{-3}$	0.052	0.398	0.130	1.026	51.7	$1 \times 10^5$	0.464	-0.469	0.005	0.318	0.393	1.024	1.016	-0.261	0.940	-1.030	6.720	0.160	11.162	7.585
$1.01 \times 10^{-3}$	0.052	0.398	0.130	1.026	51.7	$1 \times 10^6$	0.412	-0.414	0.002	0.273	0.394	1.013	1.009	-0.230	0.882	-1.035	10.048	0.158	20.187	13.636
$4.80 \times 10^{-2}$	0.812	0.731	0.270	2.225	12.5	$1 \times 10^5$	-0.081	-0.764	0.844	0.713	0.725	1.049	1.214	-0.392	0.245	-2.556	1.023	0.623	7.755	5.539
$4.80 \times 10^{-2}$	0.812	0.731	0.270	2.225	12.5	$1 \times 10^6$	-0.145	-0.725	0.870	0.591	0.724	0.954	1.241	-0.423	0.262	-2.538	0.632	0.626	1.488	1.273
$3.42 \times 10^{-1}$	0.458	0.542	0.520	5.210	23.5	$1 \times 10^5$	-1.249	-1.000	2.249	0.018	...	1.853	1.733	...	...	...	...	...	...	...
$3.42 \times 10^{-1}$	0.458	0.542	0.520	5.210	23.5	$1 \times 10^6$	-1.263	-1.000	2.263	0.040	...	1.686	1.636	...	...	...	...	...	...	...

**Note.** Intermediate outcomes are especially useful for comparing the graze-and-merge cases where the bodies recollide later in the simulation. Each group of two rows represents one scenario simulated at both resolutions: graze-and-merge, hit-and-run, another graze-and-merge, near-head-on erosive, and slightly off-axis highly erosive/disruptive.

**Table 4**  
The Final Outcomes of the Same Five Collision Scenarios in Table 3, Again as a Function of SPH Resolution

Collision Parameters							End State										
$m_{\text{tar}}$ ( $M_{\oplus}$ )	$\gamma$	$Z_{\text{tar}}$	$Z_{\text{imp}}$	$\frac{v_{\text{coll}}}{v_{\text{esc}}}$	$\theta_{\text{coll}}$ (deg)	$N$	$\xi_{\text{L}}$	$\xi_{\text{S}}$	$\xi_{\text{D}}$	$\Omega_{\text{L}}$ ( $\text{rad hr}^{-1}$ )	$Z_{\text{L}}$	$I_{\text{Z,L}}$	$I_{\text{A,L}}$	$\xi_{\text{P}}$	$\gamma_{\text{P}}$	$\frac{v_{\text{RMS}}}{v_{\text{esc}}}$	$\frac{v_{\text{RMS},\infty}}{v_{\text{esc}}}$
$1.72 \times 10^{-4}$	0.961	0.371	0.284	1.169	51.2	$1 \times 10^5$	1.000	-1.000	0.000	4.366	0.326	2.474	1.801	1.000	0.000	...	...
$1.72 \times 10^{-4}$	0.961	0.371	0.284	1.169	51.2	$1 \times 10^6$	1.000	-1.000	0.000	4.367	0.326	2.394	1.746	1.000	0.000	...	...
$7.14 \times 10^{-4}$	0.319	0.192	0.294	1.837	30.3	$1 \times 10^5$	0.209	-0.321	0.112	0.844	0.200	1.107	1.039	0.209	0.203	0.838	0.828
$7.14 \times 10^{-4}$	0.319	0.192	0.294	1.837	30.3	$1 \times 10^6$	0.208	-0.354	0.146	0.764	0.201	1.108	1.032	0.165	0.179	0.956	0.947
$1.01 \times 10^{-3}$	0.052	0.398	0.130	1.026	51.7	$1 \times 10^5$	0.917	-0.928	0.011	0.534	0.389	1.040	1.018	0.875	0.004	0.561	0.543
$1.01 \times 10^{-3}$	0.052	0.398	0.130	1.026	51.7	$1 \times 10^6$	0.894	-0.942	0.048	0.506	0.388	1.032	1.011	0.836	0.003	0.357	0.320
$4.80 \times 10^{-2}$	0.812	0.731	0.270	2.225	12.5	$1 \times 10^5$	0.278	-0.953	0.675	1.559	0.696	1.410	1.268	0.206	0.032	1.050	1.078
$4.80 \times 10^{-2}$	0.812	0.731	0.270	2.225	12.5	$1 \times 10^6$	0.248	-0.923	0.675	1.570	0.697	1.236	1.176	0.170	0.052	1.134	1.127
$3.42 \times 10^{-1}$	0.458	0.542	0.520	5.210	23.5	$1 \times 10^5$	-1.249	-1.000	2.249	0.018	0.868	1.853	1.733	-1.249	0.000	2.398	2.395
$3.42 \times 10^{-1}$	0.458	0.542	0.520	5.210	23.5	$1 \times 10^6$	-1.263	-1.000	2.263	0.040	0.868	1.686	1.636	-1.263	0.000	2.403	2.400



**Figure 13.** Comparison of the accretion efficiency of debris  $\xi_D$  and the rms of the debris velocity for collision 561 of the 1000 simulation set for two different SPH resolutions, as indicated in the legends. The dashed horizontal line in the right panel shows the impact velocity.

of the remnants. Thus, while the higher-resolution studies produce some differences, which are important to explore further as resources permit, overall, we conclude that our findings are robust at the  $N = 1 \times 10^5$  resolution of the database.

### C.2. Size–Frequency Distribution

To check for the robustness of the size–frequency distribution with resolution, we selected one simulation where several hundreds of clumps are produced as debris: simulation 566, with  $m_{\text{tar}} = 4.80 \times 10^{-2} M_{\oplus}$ ,  $\gamma = 0.81$ ,  $v_{\text{coll}}/v_{\text{esc}} = 2.23$ , and  $\theta_{\text{coll}} = 12^\circ.5$ . The size–frequency distributions at the end of the evolution for two simulations, one from the database (labeled “100k”) and a similar one but with 1 million SPH particles instead (labeled “1M”), are shown in Figure 12. Overall, the two distributions appear to match quite well; the mass of the largest remnant is similar in both simulations, and the total number of clumps at the resolution limit of the nominal simulation is also similar. There are nevertheless a few differences at intermediate remnant masses. The high-resolution simulation produces a more massive second remnant, as expected (e.g., Genda et al. 2015). The high-resolution run also only has two remnants comparable in mass to the second-largest, whereas the nominal-resolution simulation has three.

### C.3. Debris Velocity

To understand whether debris masses and velocities are influenced by common choices in SPH resolution used in the giant impact studies, we perform a simulation at  $10^6$  particles and compare it to the nominal resolution in the data set ( $10^5$  particles). We select a run from the database where debris represents roughly 70% of the total mass, and where no clumps are found (simulation 561,  $m_{\text{tar}} = 3.42 \times 10^{-1} M_{\oplus}$ ,  $\gamma = 0.46$ ,  $v_{\text{coll}}/v_{\text{esc}} = 5.21$ ,  $\theta_{\text{coll}} = 23^\circ.5$ ). As shown in Figure 13, the amount of debris produced and their velocities are very similar in both simulations; the debris are traveling at roughly 2.5 times the mutual escape velocity of the colliding bodies, or

nearly half of their relative velocity at initial contact. In addition, no clumps (minimum of 10 particles) are found within the debris even when using 1 million particles, so the debris size remains unresolved.

### ORCID iDs

Alexandre Emsenhuber <https://orcid.org/0000-0002-8811-1914>  
 Erik Asphaug <https://orcid.org/0000-0003-1002-2038>  
 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>  
 Robert E. Melikyan <https://orcid.org/0000-0003-2018-3273>  
 C. Adeene Denton <https://orcid.org/0000-0002-6696-2961>

### References

- Agnor, C., & Asphaug, E. 2004, *ApJL*, 613, L157  
 Agnor, C. B., Canup, R. M., & Levison, H. F. 1999, *Icar*, 142, 219  
 Asphaug, E. 2010, *ChEG*, 70, 199  
 Asphaug, E. 2017, in *Planetesimals: Early Differentiation and Consequences for Planets*, ed. L. T. Elkins-Tanton & B. P. Weiss (Cambridge: Cambridge Univ. Press)  
 Asphaug, E., & Reufer, A. 2013, *Icar*, 223, 544  
 Asphaug, E., & Reufer, A. 2014, *NatGe*, 7, 564  
 Ballantyne, H. A., Jutzi, M., Golabek, G. J., et al. 2023, *Icar*, 392, 115395  
 Bauer, A., & Springel, V. 2012, *MNRAS*, 423, 2558  
 Belton, M. J. S., Thomas, P., Veverka, J., et al. 2007, *Icar*, 187, 332  
 Benz, W. 1991, in *Late Stages of Stellar Evolution. Computational Methods in Astrophysical Hydrodynamics (Lecture Notes in Physics, Vol. 373)*, ed. C. B. De Loore (Berlin: Springer), 258  
 Benz, W., Anic, A., Homer, J., & Whitby, J. A. 2007, *SSRv*, 132, 189  
 Benz, W., & Asphaug, E. 1994, *Icar*, 107, 98  
 Benz, W., & Asphaug, E. 1995, *CoPhC*, 87, 253  
 Benz, W., & Asphaug, E. 1999, *Icar*, 142, 5  
 Benz, W., Cameron, A. G. W., & Melosh, H. J. 1989, *Icar*, 81, 113  
 Benz, W., Slattery, W. L., & Cameron, A. G. W. 1987, *Icar*, 71, 30  
 Bonet, J., & Lok, T.-S. L. 1999, *CMAME*, 180, 97  
 Burger, C., Bazsó, Á., & Schäfer, C. M. 2020, *A&A*, 634, A76  
 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
- Canup, R. M. 2005, *Sci*, **307**, 546
- Canup, R. M., & Asphaug, E. 2001, *Natur*, **412**, 708
- Canup, R. M., Richter, K., Dauphas, N., et al. 2021, arXiv:2103.02045
- 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. 2013, *Icar*, **224**, 43
- Chau, A., Reinhardt, C., Helled, R., & Stadel, J. 2018, *ApJ*, **865**, 35
- Clement, M. S., Kaib, N. A., Raymond, S. N., Chambers, J. E., & Walsh, K. J. 2019, *Icar*, **321**, 778
- Collins, G. S., Melosh, H. J., & Ivanov, B. A. 2004, *M&PS*, **39**, 217
- Cullen, L., & Dehnen, W. 2010, *MNRAS*, **408**, 669
- Davies, E. J., Carter, P. J., Root, S., et al. 2020, *JGRE*, **125**, e06227
- Deng, H., Reinhardt, C., Benitez, F., et al. 2019, *ApJ*, **870**, 127
- Duncan, M. J., Levison, H. F., & Lee, M. H. 1998, *AJ*, **116**, 2067
- Dwyer, C. A., Nimmo, F., & Chambers, J. E. 2015, *Icar*, **245**, 145
- Emsenhuber, A., & Asphaug, E. 2019a, *ApJ*, **875**, 95
- Emsenhuber, A., & Asphaug, E. 2019b, *ApJ*, **881**, 102
- Emsenhuber, A., Asphaug, E., Cambioni, S., Gabriel, T. S. J., & Schwartz, S. R. 2021, *PSJ*, **2**, 199
- Emsenhuber, A., Cambioni, S., Asphaug, E., et al. 2020, *ApJ*, **891**, 6
- Emsenhuber, A., Jutzi, M., & Benz, W. 2018, *Icar*, **301**, 247
- Gabriel, T. S. J., & Allen-Sutter, H. 2021, *ApJL*, **915**, L32
- Gabriel, T. S. J., & Cambioni, S. 2023, *AREPS*, **51**, 671
- Gabriel, T. S. J., Jackson, A. P., Asphaug, E., et al. 2020, *ApJ*, **892**, 40
- Genda, H., Fujita, T., Kobayashi, H., Tanaka, H., & Abe, Y. 2015, *Icar*, **262**, 58
- Genda, H., Kokubo, E., & Ida, S. 2012, *ApJ*, **744**, 137
- Golabek, G. J., Emsenhuber, A., Jutzi, M., Asphaug, E. I., & Gerya, T. V. 2018, *Icar*, **301**, 235
- Haghighipour, N., Maindl, T. I., Schäfer, C. M., & Wandel, O. J. 2018, *ApJ*, **855**, 60
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
- Hinkel, N. R., Timmes, F. X., Young, P. A., Pagano, M. D., & Turnbull, M. C. 2014, *AJ*, **148**, 54
- Hosono, N., Karato, S.-i., Makino, J., & Saitoh, T. R. 2019, *NatGe*, **12**, 418
- Housen, K. R., & Holsapple, K. A. 1999, *Icar*, **142**, 21
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Jackson, A. P., & Wyatt, M. C. 2012, *MNRAS*, **425**, 657
- Johnson, B. C., Blair, D. M., Collins, G. S., et al. 2016, *Sci*, **354**, 441
- Jutzi, M. 2015, *P&SS*, **107**, 3
- Jutzi, M. 2019, *P&SS*, **177**, 104695
- Jutzi, M., & Asphaug, E. 2015, *Sci*, **348**, 1355
- Jutzi, M., & Benz, W. 2017, *A&A*, **597**, A62
- Jutzi, M., Holsapple, K., Wünneman, K., & Michel, P. 2015, in *Asteroids IV*, ed. P. Michel, F. E. DeMeo, & W. F. Bottke (Tucson, AZ: Univ. Arizona Press), 679
- Jutzi, M., Michel, P., & Richardson, D. C. 2019, *Icar*, **317**, 215
- Kokubo, E., & Genda, H. 2010, *ApJL*, **714**, L21
- Kokubo, E., & Ida, S. 2002, *ApJ*, **581**, 666
- Leinhardt, Z. M., Marcus, R. A., & Stewart, S. T. 2010, *ApJ*, **714**, 1789
- Leinhardt, Z. M., & Stewart, S. T. 2012, *ApJ*, **745**, 79
- Levison, H. F., Duncan, M. J., & Thommes, E. 2012, *AJ*, **144**, 119
- Lock, S. J., & Stewart, S. T. 2017, *JGRE*, **122**, 950
- Lock, S. J., Stewart, S. T., & Čuk, M. 2020, *E&PSL*, **530**, 115885
- Marsset, M., Brož, M., Vernazza, P., et al. 2020, *NatAs*, **4**, 569
- McKay, M. D., Beckman, R. J., & Conover, W. J. 1979, *Technometrics*, **21**, 239
- Meier, T., Reinhardt, C., & Stadel, J. G. 2021, *MNRAS*, **505**, 1806
- Melosh, H. J. 2007, *M&PS*, **42**, 2079
- Melosh, H. J., & Ivanov, B. A. 2018, *GeoRL*, **45**, 2597
- Monaghan, J. J. 1992, *ARA&A*, **30**, 543
- Monaghan, J. J. 1997, *JCoPh*, **136**, 298
- Monaghan, J. J., & Lattanzio, J. C. 1985, *A&A*, **149**, 135
- Morris, J. P., & Monaghan, J. J. 1997, *JCoPh*, **136**, 41
- Movshovitz, N., Nimmo, F., Korycansky, D. G., Asphaug, E., & Owen, J. M. 2016, *Icar*, **275**, 85
- Nakajima, M., Genda, H., Asphaug, E., & Ida, S. 2022, *NatCo*, **13**, 568
- O'Brien, D. P., Morbidelli, A., & Levison, H. F. 2006, *Icar*, **184**, 39
- Ohnaka, M. 1995, *GeoRL*, **22**, 25
- Quintana, E. V., Barclay, T., Borucki, W. J., Rowe, J. F., & Chambers, J. E. 2016, *ApJ*, **821**, 126
- Reinhardt, C., Meier, T., Stadel, J. G., Otegi, J. F., & Helled, R. 2022, *MNRAS*, **517**, 3132
- 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
- Richardson, D. C., Michel, P., Walsh, K. J., & Flynn, K. W. 2009, *P&SS*, **57**, 183
- Rogers, L. A. 2015, *ApJ*, **801**, 41
- Rosswog, S. 2009, *NewAR*, **53**, 78
- Russell, C. T., Raymond, C. A., Coradini, A., et al. 2012, *Sci*, **336**, 684
- Scora, J., Valencia, D., Morbidelli, A., & Jacobson, S. 2020, *MNRAS*, **493**, 4910
- Shoemaker, E. M. 1962, in *Physics and Astronomy of the Moon*, ed. Z. Kopal (New York: Academic Press), 283
- Speith, R. 2006, *Improvements of the Numerical Method Smoothed Particle Hydrodynamics*, Habilitationsschrift
- Stern, S. A., Weaver, H. A., Spencer, J. R., et al. 2019, *SCI*, **364**, aaw9771
- Stewart, S., Davies, E., Duncan, M., et al. 2020, in *AIP Conf. Proc.* 2272, ed. J. M. D. Lane et al. (Melville, NY: AIP Publishing), 080003
- Stewart, S. T., & Leinhardt, Z. M. 2012, *ApJ*, **751**, 32
- Sugiura, K., Kobayashi, H., & Inutsuka, S. 2018, *A&A*, **620**, A167
- Tanga, P., Comito, C., Paolicchi, P., et al. 2009, *ApJL*, **706**, L197
- Thompson, S. L., & Lauson, H. S. 1972, *Improvements in the CHART-D Radiation-Hydrodynamic Code III: Revised Analytic Equations of State*, Tech. Rep., Sandia Natl. Laboratory SC-RR-71 0714, <https://www.osti.gov/biblio/4208328>
- Timpe, M. L., Han Veiga, M., Knabenhans, M., Stadel, J., & Marelli, S. 2020, *ComAC*, **7**, 2
- Valencia, D., Paracha, E., & Jackson, A. P. 2019, *ApJ*, **882**, 35
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261
- Walsh, K. J., & Levison, H. F. 2019, *Icar*, **329**, 88
- Wetherill, G. W. 1985, *SCI*, **228**, 877
- Yang, J., Goldstein, J. I., & Scott, E. R. D. 2007, *Natur*, **446**, 888
- Zahnle, K., Arndt, N., Cockell, C., et al. 2007, *SSRv*, **129**, 35