Dynamics of Planets, Stars & Compact Objects

Research Thesis in Partial Fulfilment of the Requirements for the Degree of
Doctor of Philosophy

Mor Rozner

Submitted to the Senate of the Technion – Israel Institute of Technology
Nisan 5784, Haifa, May 2024

Related Publications

The Research thesis was done in the faculty of physics under the supervision of Prof. Hagai B. Perets. The author of this thesis states that the research, including the collection, processing and presentation of data, addressing and comparing to previous research, etc., was done entirely in an honest way, as expected from scientific research that is conducted according to the ethical standards of the academic world. Also, reporting the research and its results in this thesis was done in an honest and complete manner, according to the same standards.
This thesis is based on original research, conducted during my PhD studies. Several papers were published/accepted and hopefully will be published soon, based on the results of this research. I led all the papers below, apart from the papers signed with *, which I co-led. See below the papers discussed in the thesis

Published/accepted

1. **Rozner M.** & Perets H. B., Soft no more: gas shielding protects soft binaries from disruption in gas-rich environments, accepted for publication in APJ
2. **Rozner M**. & Perets H. B. , *Born to be wide: the distribution of wide binaries in the field and soft binaries in clusters*, APJ, 955, 2 (2023)
3. **Rozner, M.**, Generozov A. & Perets H.B., *Binary formation through gas-assisted capture and the implications for stellar, planetary and compact-object evolution*, MNRAS, 521,1 (2023)
4. **Rozner, M.** & Perets H.B., *Binary evolution, gravitational-wave mergers and explosive transients in multiple-populations gas-enriched globular-clusters*, APJ, 931,2
5. **Rozner, M.**, Glanz H. , Perets, H. B. and Grishin, E., *Inflated Eccentric Migration of Evolving Gas-Giants I: Accelerated Formation and Destruction of Hot and Warm Jupiters*, APJ, 931, 1 (2022)
6. *Glanz H., **Rozner, M.** , Perets, H. B. and Grishin, E., *Inflated Eccentric Migration of evolving gas giants II: Numerical methodology and basic concepts*, APJ, 931, 1 (2022)
7. **Rozner, M.**, Veras, D. & Perets, H. B., Rapid destruction of planetary debris around WDs through wind erosion, MNRAS, 502, 4 (2021)
8. **Rozner, M.**, Grishin, E., Perets, H. B., *The Wide-Binary Origin of The Pluto-Charon System*, MNRAS, 497, 4 (2020)
9. *Grishin, E., **Rozner, M.** , Perets, H. B., *Erosion-driven Size Redistribution of Protoplanetary Disk Solids and the Onset of Streaming Instability and Pebble Accretion*, APJL, 898,1 (2020)
10. **Rozner, M.**, Grishin, E., Perets, H. B., *The aeolian-erosion barrier for the growth of metre-size objects in protoplanetary discs*, MNRAS, 496, 4 (2020)

Contents

1 Introduction **5**

1.1 Planet Formation & Dynamics 5

1.1.1 Planet Formation 6

1.1.2 The Hierarchical Three-Body Problem 11

1.1.3 Tidal Forces 16

1.1.4 Hot & warm Jupiters 20

1.2 Gravitational waves 22

1.3 Dynamics in gas-rich environments 25

1.3.1 Gas dynamical friction (GDF) 26

1.3.2 Globular clusters as gas-rich environments 29

1.4 Binary formation, evolution & distribution 31

1.4.1 Binary formation 31

1.4.2 Soft & hard binaries 32

2 Aeolian erosion in protoplanetary disks **34**

2.1 Aeolian-Erosion in Protoplanetary-Disks I 35

2.2 Aeolian-Erosion in Protoplanetary-Disks II – Implications 49

2.3 Aeolian erosion III - white dwarf disks 58

3 The Wide-Binary Origin of The Pluto-Charon System **71**

4 Inflated eccentric migration of hot & warm Jupiters **83**

5 Dynamics in gas-rich environments **118**

5.1 Binary evolution & gravitational-wave mergers in multiple-population gas-enriched globular Clusters 120

5.2 Gas-assisted binary formation 140

5.3 Gas shielding of soft binaries 163

6 The distribution of soft binaries in clusters and wide binaries in the field 182

7 Discussion and Summary 201

Bibliography 203

List of Figures

1.1 Effect of gas density on particle radial velocity. Adopted from Weidenschilling (1977). 7

1.2 Schematic description of a hierarchic triple system. Adopted from `http://www.astro.ucla.edu/\protect\unhbox\voidb@x\protect\penalty\@M\{}snaoz/NEKL.html`, see also an animation there. 12

1.3 An example of LK evolution, for a system with $m_1 = 1.57 \times 10^24$g, $m_2 = 1.35 \times 10^{24}$g and a normalized inner separation $\alpha = a_{in}/R_H = 0.1$ (where R_H is the Hill Radius) for the inner binary, $m_3 = 2 \times 10^{33}$g, $e_{out} = 0.2488$. 14

1.4 An illustration for the bulges raise due to tidal forces, in the weak tide model. Ω and $\dot{\theta}$ are the frequencies of the planet and star correspondingly. Adopted from Hut (1981). 17

1.5 Masses and radii of known hot Jupiters, relative to Jupiter and colored by equilibrium temperature. Adopted from Komacek and Youdin (2017). 21

1.6 The detected gravitational-wave strain amplitude as a function of time for GW150914, the first signal detected nearly simultaneously by the LIGO Hanford and Livingston observatories on September 14, 2015. Figure adapted from `https://creativecommons.org/licenses/by/3.0/` 22

1.7 The gravitational-wave spectrum probed by strain-sensitive gravitational-wave detectors. Adopted from Bailes et al., 2021. 23

1.8 Schematic diagram illustrating the mechanisms affecting the BH population and driving binary formation and evolution. Taken from Tagawa et al. (2020). 25

1.9 An illustration of dynamical friction (taken from lecture notes by Prof. Frank van den Bosch) 27

1.10 Examples of the wake induced in gas dynamical friction for subsonic velocity ($\mathcal{M} = v/c_s < 1$) and supersonic velocity ($\mathcal{M} = v/c_s > 1$). Taken from Ostriker (1999). 28

1.11 Evidence for different chemical abundances. Taken from Bastian and Lardo (2018) 30

2.1 The characteristic timescales for the aeolian-erosion as a function of its initial radius. Each line represents different starting separation. For smaller sizes the time is truncated by the separation-dependent limit where there is no erosion, for dust size of 0.1cm. The dots correspond to timescales from numerical simulation. Adopted from Rozner et al. (2020b). 38

2.2 The dependence of the evolution on the initial radii of the bodies at a fixed distance of $a = 1$ AU from the star. Adopted from Rozner et al. (2020b). 39

2.3 The dependence of the evolution on the initial radii of the bodies at a fixed distance of $a = 1$ AU from the star. Adopted from Rozner et al. (2020b). 39

2.4 Critical stokes number $\tau_\star$ as a function of the orbital separation a. Solid lines are solutions of the laminar velocity only. Dashed lines are the solution with both laminar and turbulent velocity, with the α-viscosity equal to $\alpha = 0.01$. Red (top), green (middle) and blue (bottom) lines correspond to detaching grain sizes of $1, 0.1, 0.01$ cm, respectively. Black circles indicate numerical integration of the erosion equation with laminar velocities for $a = 0.3, 1, 2$AU and $d = 0.01, 0.1, 1$cm , respectively. Black squared indicate the same numerical integration but with both laminar and turbulent velocities. Adopted from Grishin et al. (2020b). 51

2.5 The evolution of triple-layered objects, with each layer comprising 1/3 of the radius, of the same density 3.45g cm^{-3}, but with different size of inner grains that build each layer. The object is embedded at a white dwarf disk at a distance of 0.005AU from the white dwarf. 60

3.1 Parameter-diagram of the different possible regimes of behavior of Pluto-Charon system. The colored regimes correspond to the different dynamical regimes of the system: no collision, extra precession no collision, secular and non-secular evolution. The lines correspond to illustrations of systems with different hierarchies – encapsulated in α, and different eccentricities. Adopted from Rozner et al. (2020c). . . . 74
3.2 Cumulative distributions of the impact properties. Upper left: Normalized pericenter q/R_{tot}; Upper right: Time of collision; Lower left: Final inclination at impact; Lower right: Velocity at impact. The vertical dashed line is the escape velocity. Adopted from Rozner et al. (2020c). 75

4.1 The thermal and orbital evolution of a HJ-progenitor migrating due to weak tides, with photosphere heating due to irradiation. The thermal (**a,c**- effective temperature and radius) and orbital (**b,d** - semimajor axis and eccentricity) evolution of 1 M_J HJ progenitors with different initial radii (1 R_J – blue, 1.3 R_J – orange and 1.5 R_J – green). Tides are modeled through a weak-tide model, irradiation of the planetary outer layer is included, but with no additional efficient heat conduction to the core. Shown are gas-giants with an initial semimajor axis of 1 AU and initial eccentricity of 0.98. The solid lines correspond to the semi-analytical calculation and the dashed to the numerical. 86
4.2 2-dimensional histogram of the eccentricity and orbital period of hot and warm Jupiters after a Hubble time, as obtained from the population-synthesis models initialized with $\sigma_1 R_3 a_1 df_{0.1}$. The probability density is normalized according to the total fraction of successful formation of HJs and WJs among all initial conditions sampled. 88
4.3 A diagram of the final fractions concluded from the Monte-Carlo simulation and a comparison inflated ($\sigma_1 R_3 a_1 df_{0.1}$) and non-inflated ($\sigma_1 R_1 a_1 df_{0.1}$) initial radii. The fractions could be summed over to 1 – we present just the results of the simulation without further normalization. The results are based on 10^4 runs of the semi-analytical model per each case. 89

4.4 The period distribution as found by the Monte-Carlo simulation, based on the semi-analytical model, for different initial radius distribution. The rest of the parameters are drawn according to $\sigma_1 R_3 a_1 df_{0.1}$. In the inset figure we introduce the probability distribution function of WJs only (with the same color code) 90

5.1 The effects of gas hardening, GWs, and three-body hardening. The blue dashed line represents the maximal SMA in which GW emission catalyzes a binary merger within Hubble time. We consider the evolution of a binary with masses $m_1 = m_2 = 10\ M_\odot$ and the initial separation of $a_0 = 1$ AU. We consider an exponential decaying background gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\rm gas})$ with $\rho_{g,0} = 1.74 \times 10^6\ M_\odot {\rm pc}^{-3}$ and $\tau_{\rm gas} = 50$ Myr. 123

5.2 The combined effect of gas hardening, three-body hardening and GWs on a binary, for different initial separations. The blue dashed line represents the maximal SMA in which GW emission catalyzes a binary merger within Hubble time. The purple dashed line corresponds to the widest binary allowed by evaporation considerations. The solid lines correspond to the evolution of the SMA, starting from an initial separation of $a_0 = 1, 10, 100, 200$ AU, and given an exponential decaying gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\rm gas})$ with $\tau_{\rm gas} = 50$ Myr. 124

5.3 The effects of gas hardening and GWs on eccentric orbit. We consider the evolution of a binary with masses $m_1 = m_2 = 10\ M_\odot$ and an initial separation of $a_0 = 1$ AU. We consider an exponential decaying background gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\rm gas})$ with $\rho_{g,0} = 1.74 \times 10^6\ M_\odot {\rm pc}^{-3}$ and $\tau_{\rm gas} = 50$ Myr. The solid lines correspond to semimajor axis evolution and the dashed lines to pericenter evolution. 125

5.4 Evolution of the stellar velocities and binary orbital elements from numerical simulations with different gas density and sound speeds. . . 142

5.5 Left: The capture rate per object for different masses, in a SF environment, for our fiducial model specified above. Right: The capture rates for equal masses, given different gas number densities. 144

5.6 Left: The capture rate per object for different masses, in a second (or later) star generation environment in a GC, for our fiducial model specified above. Right: The capture rates for equal masses, given different gas masses. 145

5.7 Total BH-BH capture rate for artificially inflated AGN disks with $\dot{M}/M_{\rm Edd} = 0.1$. The x-axis shows the aspect ratio at 3 pc after rescaling. If the disk aspect ratio is increased by an order of magnitude, the rate of stable captures goes to 0. The vertical red line shows the aspect ratio for the Sirko and Goodman (2003) model for fiducial parameters 146

5.8 The shielding radius – the maximal radius for which a binary is considered hard when accounting for the effect of both gas hardening and stellar encounters, as a function of different gas densities, for different binaries of equal masses $m = m_1 = m_2$. Taken from Rozner and Perets (2024). 166

5.9 Comparison between the semimajor axes evolution with and without the contribution of gas-shielding, for a binary with $m = m_1 = m_2 = 1\ M_\odot$. 167

5.10 Upper panel: the gas density evolution over time, for initial gas density of $\rho_{g,0} = 10^3\ M_\odot\ {\rm pc}^{-3}$. Lower panel: The evolution of different initial semimajor axes for a background density with $\rho_{g,0} = 10^3\ M_\odot\ {\rm pc}^{-3}$, for a binary with masses $m = m_1 = m_2 = 1\ M_\odot$. The dashed line corresponds to the soft-hard limit boundary as derived based on stellar interactions only. 168

5.11 The results of a Monte Carlo simulation, with two equal masses $m = m_1 = m_2 = 1\ M_\odot$ averaged over 1000 iterations. The red dashed line corresponds to the critical semimajor axis between soft and hard binaries when considering stellar interactions only (without the effect of gas), and the green dashed line corresponds to the shielding radius as calculated relative to the *initial* gas density. The different solid lines correspond to different times. Upper panel: Initial background density of $\rho_{g,0} = 10^3\ M_\odot\ {\rm pc}^{-3}$. Lower panel: $\rho_{g,0} = 10^4\ M_\odot\ {\rm pc}^{-3}$. Note the different scales. 169

5.12 The normalized difference between the number of binaries in different separations, based on the results of our Monte Carlo simulation, with two equal masses $m = m_1 = m_2 = 1\ M_\odot$ averaged over 1000 iterations. The red dashed line corresponds to the critical semimajor axis between soft and hard binaries when considering stellar interactions only (without the effect of gas), and the green dashed line corresponds to the shielding radius as calculated relative to the *initial* gas density. The different solid lines correspond to different times. Upper panel: Initial background density of $\rho_{g,0} = 10^3\ M_\odot\ \mathrm{pc}^{-3}$. Lower panel: $\rho_{g,0} = 10^4\ M_\odot\ \mathrm{pc}^{-3}$. Note the different scales. 170

6.1 The fraction of soft binaries in clusters/wide binaries in the field for different choices of mass distribution, as derived from the Monte-Carlo simulation, is based on our analytical derivation. (averaged over 100 Monte-Carlo realizations). 186

6.2 The semi-major axis distribution, as derived from the Monte-Carlo simulation, based on our analytical derivation, averaged over 1000 realizations, for a cluster containing both stars and planets ($N_\star = N_p$), in comparison to results from N-body simulations. 187

6.3 The binary fraction as a function of the primary mass, for different masses, as derived from an averaging over 1000 runs of the Monte-Carlo simulation, for a cluster containing 100 stars, drawn from a Kroupa mass function, and an equal number of planets, in comparison to N-body simulations. 188

List of Tables

5.1 Maximum velocities for capture in different regimes of GDF, with no headwind. 141

Abstract

The research in this dissertation comprises studies in three main directions, connected by similar ideas, tools and methods. I will discuss the formation and dynamics of planets, dynamics in gas-rich environments and distributions of soft binaries in clusters and wide binaries in the field.

Planet formation is a long and challenging journey, that takes along several orders of magnitudes, from small grains to full-sized planets. I studied a novel barrier in planet formation – aeolian erosion – which is inspired by geological models. I showed that this process could lead to some significant effects also on other processes during the initial stages of planet formation, and shaping the size distribution of objects in protoplanetary disks. I then discussed later stages of planet evolution including the formation of Pluto-Charon from a wide binary, in a general mechanism that could explain the formation of other Kuiper-belt binaries. Then I discussed the coupled thermal-dynamical evolution of hot & warm Jupiters. I treated high eccentricity migration of initially inflated gas giants, and followed their contraction using a thermal model, and then analyzed its effect on final fractions and properties. I showed that thermal evolution substantially affects the dynamical evolution and final properties of formed Jupiters.

Binaries are abundant in many astrophysical scales, and while they were studied extensively in gas-free environments, there are some unexplored phenomena in gas-rich environments. Such mergers took place at the early stages of the clusters formation. I investigated the evolution of compact binaries in the gas-rich environment of the second-generation star formation and discussed a novel gravitational wave channel in these environments. Gas could also assist in binary formation. I studied the conditions for binary formation in several gas-rich media including star-forming regions, AGN disks and the gaseous regions of second-generation star formation in globular clusters, and described the possible following dynamical processes. Dynamical binary formation requires an external dissipation force, and gas induced energy decrease could serve as such. If these binaries survive after their creation, they could contribute significantly to the binaries population. I then discussed the process of shielding soft binaries from disruptions and a novel hard binary formation channel

from soft binaries in gas-rich environments. Usually, soft binaries are prone to disruptions due to interactions with other stars. However, in gas-rich environments, all the interaction with the gas extracts energy from all the binaries, soft or hard, and by that harden also soft binaries. If the gas density is high enough, by the end of the gas phase, substantial fraction of soft binaries become hard and avoid disruption. This process might affect the total number of hard binaries and by that contribute to various phenomena such as supernovae and tidal interactions. Since several binary formation channels favor the creation of soft binaries, gas shielding could play an important role in creating hard binaries by converting soft binaries into hard binaries via efficient gas hardening. Modeling the processes described above could in principle revise the dynamics of gas-rich clusters and gas-rich environments in general, and lead to a better understanding of the properties of the populations in these environments.

I also derived the distribution of soft binaries in clusters and wide binaries in the field using tools from statistical mechanics. Thanks to the good agreement between our analytical results and N-body simulations in most cases, our results could be used to evade expensive numerical calculations and give us a better understanding of the properties of these binaries.

List of Symbols

Below I introduce the main symbols used in this dissertation, as sorted by chapter.

Symbol	Meaning
Chapter 2: Aeolian erosion in protoplanetary disks	
$v_\star$	threshold velocity for aeolian-erosion
Σ_g	gas surface density
ρ_g	gas density
ρ_p	density of the eroded object
d	typical size of the grains that compose the eroded object
R	radius of the eroded object
m	mass of the eroded object
$v_{\rm rel}$	relative velocity between the eroded object and the gas
$a_{\rm coh}$	the cohesion acceleration
η	gas pressure support
St	Stokes number
Re	Reynolds number
$v_{\rm th}$	thermal velocity
c_s	sound of speed
λ	mean free path
n_g	gas number density
σ	Neutral collision cross-section
T	temperature
α	Shakura Sunyaev parameter
h/r	aspect ratio of the disk
Chapter 3: The wide binary origin of the Pluto-Charon system	
m_1, m_2	masses of the companions of the inner binary
m_3	the mass of the outer tertiary
$P_{\rm in}$ ($P_{\rm out}$)	period of the inner (outer) orbit
$a_{\rm in}$ ($a_{\rm out}$)	inner (outer) binary separation
$e_{\rm out}$	eccentricity of the outer binary
i_0	initial mutual inclination
e_0	initial eccentricity
R_H	Hill radius
$\alpha = a_{\rm in}/R_H$	normalized inner separation

R_L	Laplace radius
Chapter 4: Inflated eccentric migration of hot & warm Jupiters	
$M_\star$	mass of the host star
M_p	mass of the planet
R_p	radius of the planet
T_c	central temperature of the planet
T_{eff}	effective temperature of the planet
L_{cool}	energy rate of cooling
L_{ext}	energy rate of heating
$L_\star$	stellar luminosity
τ_p	planetary tidal lag
k_{AM}	planetary apsidal motion constant
γ	polytropic index
P_c	central pressure
P_{RCB}	pressure at the boundary layer between the radiative and convective regions
Chapter 5: Dynamics in gas-rich environments	
ρ_g	gas density
τ_{gas}	gas lifetime
c_s	speed of sound
$n_\star$	stellar number density
$\rho_\star$	stellar density
σ	velocity dispersion
α	Shakura Sunyaev parameter
$\mathcal{M}$	Mach number
h/r	scale height
Ω_{bin}	angular frequency of the binary
v_{rel}	relative velocity between the object and the gas
R_{Hill}	Hill radius
q	mass ratio
v_{esc}	escape velocity
a_{SH}	critical separation between soft and hard binaries
$\bar{m}$	mean mass of a star in the cluster
Chapter 6: The distribution of soft binaries in clusters and wide binaries in the field	
ρ_i	density of the ith mass species
σ	velocity dispersion
R_{cut}	cutoff radius
R_{Hill}	Hill radius
Q	Virial ratio
$N_\star$	number of stars
N_i	the number of stars from the ith species

Chapter 1

Introduction

In my PhD I studied the dynamics and evolution of planets, compact objects and stars. My research comprises three main directions: formation and dynamics of planetary systems, dynamics in gas-rich environments and binary formation, evolution and distribution. These directions frequently overlap and give unique perspectives on each other. I will start with a short introduction to each one of these directions and their components (note that some chapters are related to several introduction sections). Then, in the following chapters, I will describe each project, the tools I used, and key results, followed by the full paper(s) that this project yielded.

1.1 Planet Formation & Dynamics

In this chapter, I will discuss the formation and dynamics of planetary systems. I will present a new barrier in the early stages of planet formation, and then discuss the dynamics and architecture of already-formed planets. This chapter serves as an introduction for chapters 2,3 and 4.

1.1.1 Planet Formation

Planet formation could be divided crudely into three stages. While the early stages and the last ones are well understood, the growth and dynamics in the intermediate regime – which contains pebbles and boulders up to planetesimals –reveal many unsolved challenges. My proposed research will focus on the intermediate regime and present new challenges, as well as the architecture and dynamics of planets after their initial formation.

Barriers and Suggested Solutions

The formation of planets is thought to be a bottom-up process (Safronov, 1972), which takes place in protoplanetary, and includes the growth of planets from small grains to their final size – this process strays over a large number of orders of magnitudes and comprises many physical processes. The growth of dust aggregates and sub-cm size pebbles in protoplanetary disks is understood theoretically and experimentally and is explained by sticking (Wurm and Blum, 1998). The growth of large objects ($\gtrsim$ 1km) to fully formed planets is also well understood and is explained by pebble-accretion (Ormel and Klahr, 2010; Perets and Murray-Clay, 2011; Lambrechts and Johansen, 2012), in which large objects grow via accretion of smaller objects ($\sim$ cm – m).

There is a gap in our physical understanding of growth in the intermediate regime – around meter-size (and actually a bit below, especially for some of the barriers we discuss here, but the barrier is still coined as the 'meter-size' barrier). Several processes potentially quench the growth in this stage and state potential barriers that objects should overcome. Among them are the bouncing barrier – objects rebound off one from each other instead of sticking (Zsom et al., 2010; Kruss et al., 2016; Booth et al., 2018); fragmentation barrier – violent collisions between objects might break them apart (e.g. Blum and Wurm, 2008) and the famous radial-drift barrier (Adachi et al., 1976; Weidenschilling, 1977) – the radial velocity is maximized for objects around meter-size, which drifts them rapidly towards the star, before their expected growth (see Fig. 1.1).

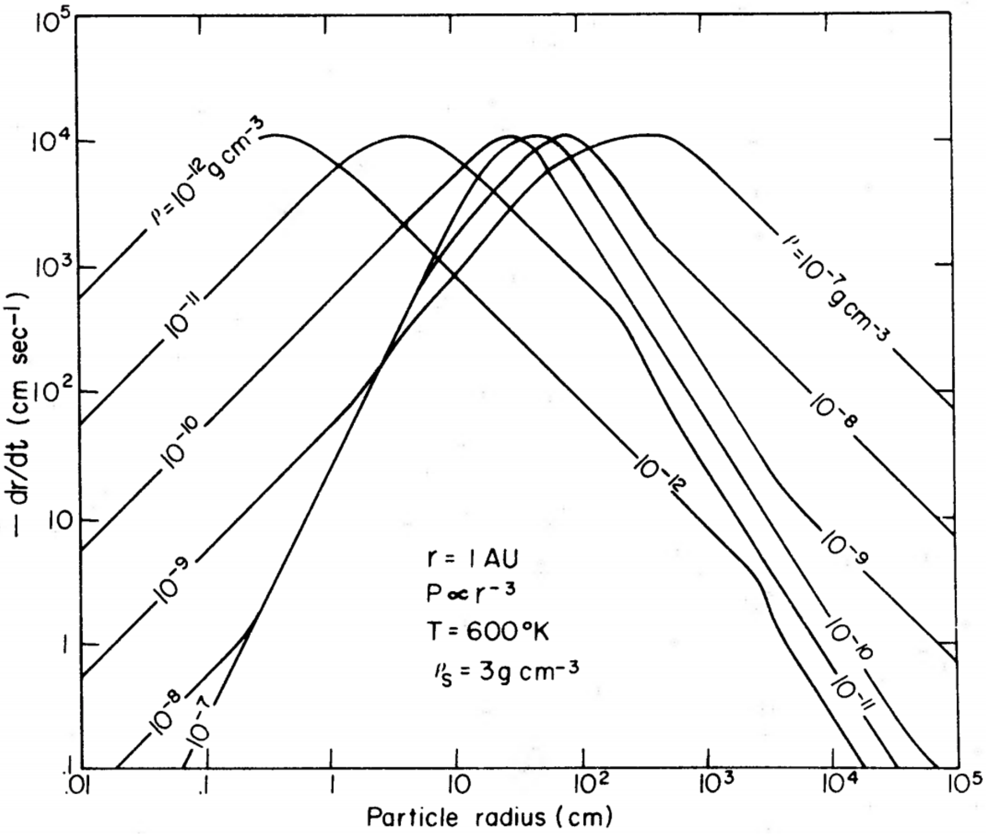

Figure 1.1: Effect of gas density on particle radial velocity. Adopted from Weidenschilling (1977).

There were many suggested solutions to the meter-size barrier, see Blum (2018) for a review and discussion about some of them. Streaming instability clumping induces overdensity of dust and pebbles that finally leads to gravitational collapse of the overdense regions into large planetesimals (Youdin and Goodman, 2005; Johansen et al., 2007). Isolated particles, that drift faster, join the newly formed clump, fueling positive feedback – an exponential growth that bypasses the meter-size barrier. Unfortunately, streaming instability requires fine-tuned conditions in order to maintain the conditions for gravitational collapse, such as pre-concentration of pebbles and large initial metallicity. Another suggested solution is successful collisional growth in relatively small velocities from the possible velocity distribution (Windmark et al., 2012a,b; Garaud et al., 2013; Booth et al., 2018). However, in this channel the growth timescales are too long, and objects larger than 0.1cm are not likely to survive (e.g. Blum, 2018). Recently, Grishin et al. (2019) suggested that planetesimals could be exchanged and captured between protoplanetary disks, such that just a small fraction of km-size objects – which could be explained by the fine-tuned mechanisms that were discussed above – should be produced in order to seed the entire birth cluster with planetesimals.

Disk Structure and Drag Laws

Primordial protoplanetary disks consist mostly of gas – the dust-to-gas ratio (Chiang and Youdin, 2010). Henceforth, objects in the disk experience a significant gas drag, which is determined by the ratio of the object's radius R and the mean free path of the gas, and for objects with a diameter larger than the mean free path λ – also on the Reynolds numbers. When $R \lesssim \lambda$, the drag force is modeled by considering individual and independent particle collisions; for $R \gtrsim \lambda$, the gas behaves as a fluid. Following Perets and Murray-Clay (2011), we take the boundary between these two regimes at $R = 9\lambda/4$. The Reynolds number of a particle in the protoplanetary disk is defined by $Re = 2Rv_{rel}/(0.5\bar{v}_{th}\lambda)$ where R is the radius of the object, v_{rel} is its velocity relative to the gas, $\bar{v}_{th} = (8/\pi)^{1/2}c_s$ is the thermal velocity of the gas and λ is the mean free path of the gas. Weidenschilling (1977) (who followed Whipple (1972, 1973)) introduced three regimes of drag-force, which are defined by the Reynolds numbers: Epstein ($Re < 1$), Stokes ($1 < Re < 800$) and Ram pressure ($Re > 800$). The full range of Reynolds numbers could be fitted with an empirical formula such that we could write a single unified expression for the drag-force that includes all the regimes (Perets and Murray-Clay, 2011 and references therein),

$$\mathbf{F}_D = \frac{1}{2} C_D(Re) \pi R^2 \rho_g v_{rel}^2 \hat{\mathbf{v}}_{rel}, \tag{1.1}$$

$$C_D(Re) = \frac{24}{Re}(1+0.27Re)^{0.43} + 0.47\left[1-\exp\left(-0.04Re^{0.38}\right)\right] \tag{1.2}$$

where ρ_g is the gas density in the protoplanetary disk.

The relative velocity of objects in disks could be decomposed into two components, $v_{rel,\phi}$ and $v_{rel,r}$ such that $v_{rel}^2 = v_{rel,\phi}^2 + v_{rel,r}^2$.

The pressure support presents in gaseous disks leads to a difference between the Keplerian velocity $\Omega_k = (GM_\star/a^3)^{1/2}$ and the actual angular velocity Ω_g: $\Omega_g - \Omega_k \approx (2\Omega_k a \rho_g)^{-1} \partial P/\partial r$ where $\partial P/\partial r$ is the pressure gradient and a is the radial distance in the disk. Furthermore, the pressure support leads to the radial-drift, which states the radial-drift meter-size barrier (Weidenschilling, 1977) discussed above. The components of the relative velocity are given by (e.g. Armitage, 2010; Perets and Murray-Clay, 2011)

$$v_{\mathrm{rel,r}} = -\frac{2\eta v_k \mathrm{St}}{1+\mathrm{St}^2}, \; v_{\mathrm{rel},\phi} = -\eta v_k \left(\frac{1}{1+\mathrm{St}^2} - 1\right) \tag{1.3}$$

where $v_k = \sqrt{GM_s tar/a}$ is the Keplerian velocity and the Stokes number is defined by

$$\mathrm{St} = \Omega_K t_{\mathrm{stop}}, \; t_{\mathrm{stop}} = \frac{m v_{\mathrm{rel}}}{F_D} \tag{1.4}$$

In order to model the protoplanetary-disk structure, we use the canonical Minimal-Mass-Solar-Nebula (MMSN) model (Hayashi, 1981), we will give here just a brief review of the parameters (taken from Perets and Murray-Clay (2011)), a more comprehensive review could be found in Armitage (2010). The surface density of the disk is taken to be $\Sigma_{g,0} = 2 \times 10^3 (a/AU)^{-3/2} g \times cm^{-2}$ with an exponential decay in a typical disk lifetime of 3Myr (Haisch et al., 2001), $\Sigma_g = \Sigma_{g,0} \exp(-t/\tau_{disk})$. The scale height varies with the distance from the center of the disk and is given by

$$\frac{H}{a} \sim \frac{c_s}{\Omega_k a} \sim 0.022 \left(\frac{a}{AU}\right)^{2/7} \tag{1.5}$$

where c_s is the speed of sound. The gas density profile decreases with a and is given by

$$\rho_g(a) \sim \frac{\Sigma_g}{2H} \sim 3 \times 10^{-9} \left(\frac{a}{\mathrm{AU}}\right)^{-16/7} \mathrm{g/cm^3} \tag{1.6}$$

the mean free path of the gas is given by $\lambda \sim (a/\mathrm{AU})^{16/7}\,\mathrm{cm}$ and the radial temperature profile by $T(a) = 120(a/\mathrm{AU})^{-3/7}\mathrm{K}$.

The disk parameters play an important role in the dynamics of objects in protoplanetary disks. These parameters dictate the growth and destruction rates, and different disk profiles might lead finally to different size distributions and dynamics.

1.1.2 The Hierarchical Three-Body Problem

Planetary systems give rise naturally to hierarchical systems, from moons to binaries in the Kuiper belt with the Sun as a third perturber, and many other examples – and henceforth are a great motivation to the studies of hierarchical systems.

Binaries, triples and higher multiplicities systems are ubiquitous in various scales (e.g. Raghavan et al., 2010; Perets, 2011; Kulkarni and Loeb, 2012; Tokovinin, 2014a,b and a plethora of others). Moreover, there are a lot of systems that are usually considered to be binaries, but actually are overlooked triples; for example, binaries near supermassive black holes (e.g. Antonini and Perets, 2012) and so does every binary in the Solar system – with the Sun as a third companion (Perets and Naoz, 2009; Grishin et al., 2020a). Furthermore, most of the stars are born and raised in clusters (Lada and Lada, 2003), which expose stars to potential frequent encounters with others. Henceforth, the dynamics of binaries, triples and higher multiplicity systems play a key role in astrophysical systems in general, and in planetary dynamics in particular.

The three-body problem is one of the most famous non-integrable problems, tracing back to the pioneering work of Newton and Poincare. These systems/interactions could be divided crudely into two classes – hierarchical and non-hierarchical. Here we will focus on hierarchical systems/interactions, and in this chapter – introduce a brief review (see a detailed review in Toonen et al., 2016).

Most of the stable astrophysical systems are hierarchical since non-hierarchical tend to decompose. One essential building block of hierarchical systems is the hierarchical triple system. Hierarchical triples are systems composed of an inner binary and a distant tertiary. These systems could be described as two binaries – the inner binary and an outer binary which includes the triary and the inner binary as companions (see Fig. 1.2. When the periods of the two binaries are well separated, i.e. the period of the outer binary is much larger than the period of the inner one, we can average the orbits of the binaries and consider them as two ellipse-shaped mass wires that interact weakly with each other. This method of averaging long-term phase evolution over timescales longer than the orbital period is called secular evolution and was developed by Von Zeipel (1916); Lidov (1962); Kozai (1962), see a detailed review in Naoz (2016). From now on, this mechanism will be related to the LK mechanism.

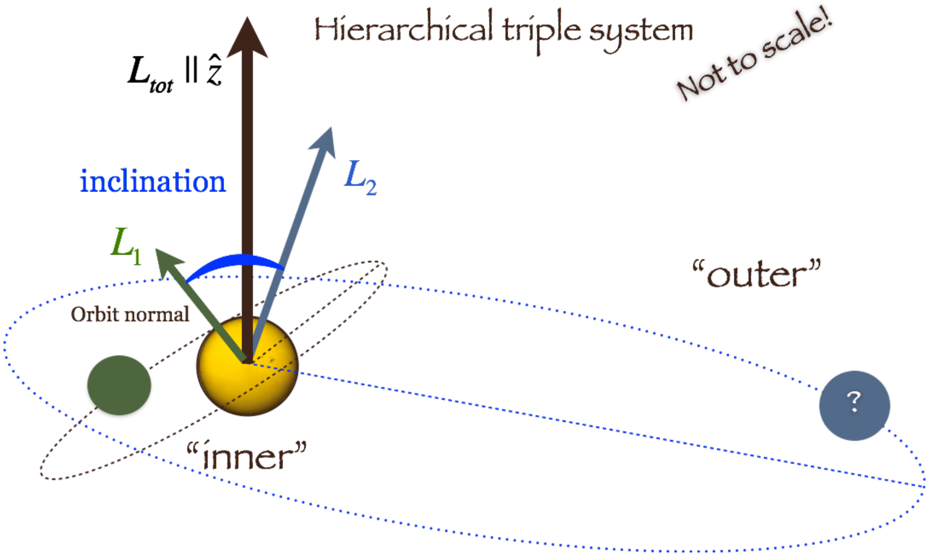

Figure 1.2: Schematic description of a hierarchic triple system. Adopted from `http://www.astro.ucla.edu/~snaoz/NEKL.html`, see also an animation there.

The LK formalism ignores and averages over timescales shorter than the secular timescale given by(Kinoshita and Nakai, 1999; Antognini, 2015).

$$\tau_{\rm sec} \approx \frac{8}{15\pi} \frac{m_1 + m_2 + m_3}{m_3} \frac{P_{\rm out}^2}{P_{\rm in}} \left(1 - e_{\rm out}^2\right)^{3/2} \tag{1.7}$$

where m_1 and m_2 are the masses of the companions of the inner binary, m_3 is the mass of the outer tertiary $P_{\rm out}$, $P_{\rm in}$ are the periods of the outer and inner binaries correspondingly and $e_{\rm out}$ is the eccentricity of the outer binary.

When the periods of the inner and outer binaries are well separated, the Hamiltonian of the problem could be decomposed into two Keplerian Hamiltonians and a weak interaction term between the two orbits (Lidov, 1962; Kozai, 1962). Over long timescales, the orbits exchange angular momentum, while the energy exchange is negligible and the inner and outer semi-major axes remain roughly constant. These conditions induce periodic variations of eccentricity and inclination.

The motion is governed by the Hamiltonian with the perturbation expanded in multipole expansion (Harrington, 1968)

$$\begin{aligned}
\mathcal{H} &= \frac{Gm_1m_2}{2a_{\text{in}}} + \frac{Gm_3(m_1+m_2)}{2a_{\text{out}}} + \mathcal{H}_{\text{pert}}; \\
\mathcal{H}_{\text{pert}} &= \frac{G}{a_{\text{out}}} \sum_{j=2}^{\infty} \left(\frac{a_{\text{in}}}{a_{\text{out}}}\right)^j \left(\frac{r_1}{a_{\text{in}}}\right)^j \left(\frac{a_{\text{out}}}{r_2}\right)^{j+1} \mathcal{M}_j P_j\left(\cos\Phi\right), \\
\mathcal{M}_j &= m_1m_2m_3 \frac{m_1^{j-1} - (-m_2)^{j-1}}{(m_1+m_2)^j} \qquad (1.8)
\end{aligned}$$

where r_i is the distance between the two companions of the i-th binary, P_i is the i-th Legendre polynomial and Φ is the angle between r_2 and r_1.

The standard LK formalism considers double-averaging, i.e. averaging over both inner and outer mean anomalies. This averaging enables an extraction of the secular changes in the system – changes in the orbital elements along timescales much longer than the orbital period. The lowest order, and most significant, is the quadruple order. In the quadruple order, the maximal eccentricity, and the corresponding minimal inclination, could be written as (Innanen et al., 1997)

$$e_{\text{max,LK}} = \sqrt{1 - \frac{5}{3}\cos^2 i_0}, \;\; i_{\text{min,LK}} = \arccos\left(\pm\sqrt{\frac{3}{5}}\right) \qquad (1.9)$$

The possible i_{min}, 39.23° and 140.77°, set the boundaries where LK evolution is active. Due to the secular evolution of the distant perturber, the inner binary experiences extreme oscillations of its eccentricity and mutual inclination, where the highest eccentricity is obtained with the lowest mutual inclination and vice versa. See an example of LK evolution in Fig. 1.3.

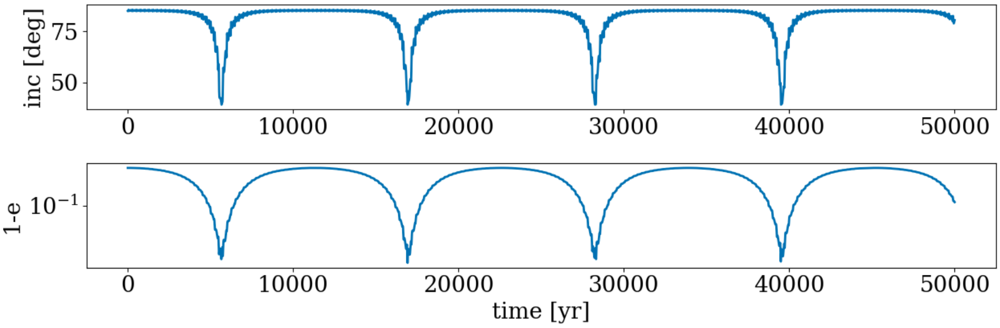

Figure 1.3: An example of LK evolution, for a system with $m_1 = 1.57 \times 10^24$g, $m_2 = 1.35 \times 10^{24}$g and a normalized inner separation $\alpha = a_{in}/R_H = 0.1$ (where R_H is the Hill Radius) for the inner binary, $m_3 = 2 \times 10^{33}$g, $e_{out} = 0.2488$.

One of the core assumptions of the LK formalism is $P_{out} \gg P_{in}$, and when the discussed systems are just mildly hierarchical, short-term effects gain significance, the LK assumptions and results lose accuracy and the double-averaging approximation breaks (Antonini and Perets, 2012; Antognini, 2015; Luo et al., 2016; Grishin et al., 2018). The breakdown leads to corrections in the evolution and in particular, corrections of the maximal eccentricity and critical inclinations for onset, which will be indexed by QS (stands for quasi-secular) (Grishin et al., 2018).

$$e_{\max,\mathrm{QS}} = \sqrt{1 - \frac{5}{3}\cos^2 i_0 \frac{1 + \frac{9}{8}\epsilon_{\mathrm{SA}} \cos i_0}{1 - \frac{9}{8}\epsilon_{\mathrm{SA}} \cos i_0}}, \tag{1.10}$$

$$i_{\min,\mathrm{QS}} = \arccos\left(\pm\sqrt{\frac{3}{5} - \frac{27}{40}\epsilon_{\mathrm{SA}}}\right) \tag{1.11}$$

were $\epsilon_{\mathrm{SA}} = P_{\mathrm{out}}/2\pi\tau_{\mathrm{sec}}$ is the strength of the single averaging (Luo et al., 2016; Grishin et al., 2018) where τ_{sec} is defined in eq. 1.7.

The LK mechanism gives rise to many interesting physical phenomena and important implications(see a detailed review in Naoz, 2016). LK oscillations might enhance the formation of short-period and contact-binaries (e.g. Fabrycky and Tremaine, 2007; Perets and Naoz, 2009; Naoz and Fabrycky, 2014; Grishin and Perets, 2016; Grishin et al., 2020a; Lyra et al., 2020), play an important role in the formation of hot-Jupiters which includes some unique features (e.g. Wu and Murray, 2003; Fabrycky and Tremaine, 2007; Naoz et al., 2011b, 2012; Petrovich, 2015a; Anderson et al., 2016; Hamers and Tremaine, 2017) and enhance the merger of compact objects and trigger supernovae (e.g. Blaes et al., 2002; Thompson, 2011; Antonini et al., 2016a, 2017; Hamers and Thompson, 2019) and produce blue-stragglers (e.g. Perets and Fabrycky, 2009; Antonini et al., 2016b; Fragione and Antonini, 2019).LK formalism is also useful for system more hierarchical then triples (e.g. Hamers et al., 2015; Hamers and Portegies Zwart, 2016; Hamers, 2018, 2020 and references therein).

1.1.3 Tidal Forces

Planets, planetesimals and smaller objects, are exposed to dissipative forces during their motion, which originates from a variety of physical processes. The motion of objects in a medium leads to dissipation, whether the medium is collisional – e.g. swarm of stars/planets, or non-collisional – e.g. gaseous disk. The interaction between the medium and the object tightly depends on the size of the object, while large objects are affected mostly by type I migration (Armitage, 2010) and gas dynamical friction (Grishin and Perets, 2015, 2016), the dissipative force that dominates for smaller objects is aerodynamic gas drag, see a discussion on drag laws in subsection 1.1.1.

Simplified stellar and planetary models treat objects as point masses, but when objects are close enough, their finite size gives rise to a gradient in the gravitational force – tide force – in which they pull each other – the gravitational pull from the closest side is the strongest and one of the furthest points is the weakest, and all the rest of the points are somewhere in between correspondingly. Since planets and stars are not completely solids, tide forces might induce the deformation of objects.

Weak Tides

The most common description for tidal evolution is given by the weak tide approximation (Darwin, 1879; Alexander, 1973; Hut, 1981). Generally, tides could act on both of the objects since both of the objects torque one on another, but here for simplicity, we will assume that only one object – a planet – feels the tides from the second object – a star. See equations 2-4 in Miller et al. (2009). The bulges that rise due to the gradient in the gravitational force are modeled by two point masses hanging in the sides of the object with the tides, in correspondence to the lag time between the two objects, as seen in figure 1.4.

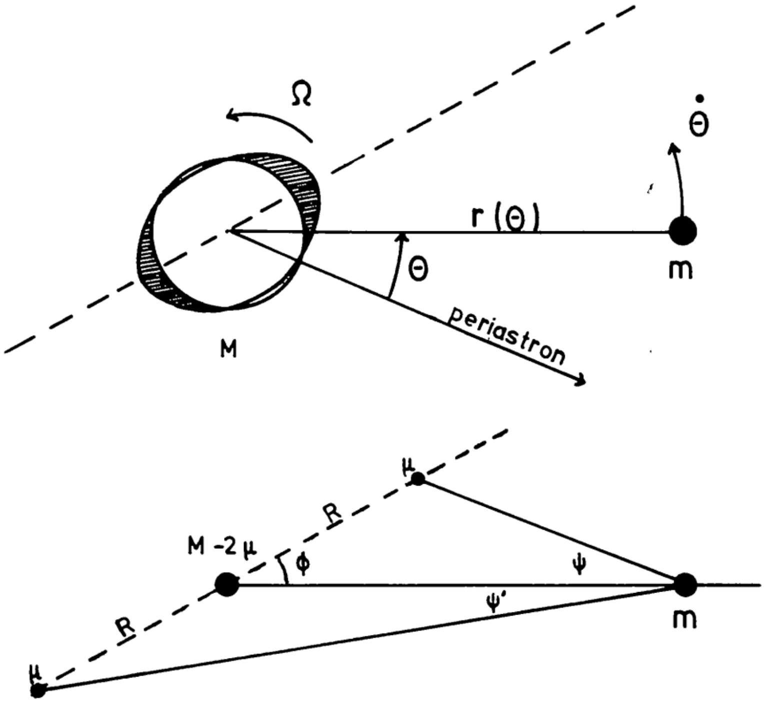

Figure 1.4: An illustration for the bulges raise due to tidal forces, in the weak tide model. Ω and $\dot{\theta}$ are the frequencies of the planet and star correspondingly. Adopted from Hut (1981).

The orbital parameters of the planet change due to the tidal forces, and their change is governed by the following equations, which are averaged over a period

$$\frac{da}{dt} = -21k_{AM,p}n^2\tau_p\frac{M_\star}{M}\left(\frac{R}{a}\right)^5 ae^2\frac{f(e)}{(1-e^2)^{15/2}}, \tag{1.12}$$

$$\frac{de}{dt} = -\frac{21}{2}k_{AM}n^2\tau_p\frac{M_\star}{M}\left(\frac{R}{a}\right)^5 e\frac{f(e)}{(1-e^2)^{13/2}} \tag{1.13}$$

where $M_\star$ is the mass of the host star, M, R, e, a and n are the mass, radius, orbital eccentricity, orbital semimajor axis and mean motion of the Jupiter correspondingly, τ_p is the planetary tidal lag time and k_{AM} is the planetary apsidal motion constant.

Further models parametrize the efficiency of tidal-dissipation by a quality factor Q, which is an inverse measure of the dissipation, which is proportional to the ratio between the maximum energy stored in the tidal oscillation, E_0, and the energy loss rate (Munk and MacDonald, 1960; Goldreich and Soter, 1966), or explicitly by

$$Q = 2\pi E_0 \left(- \oint \dot{E} dt \right)^{-1} \tag{1.14}$$

Estimating Q from first principles is not trivial, since there are many uncertainties in the mechanisms of tidal dissipation.

Dynamical Tides

One of the core assumptions of the weak tide model is a static tide in hydrostatic equilibrium; this assumption breaks for eccentric orbits or small pericenters, and excitation/damping of internal waves/modes should be taken into consideration (Zahn, 1977; Mardling, 1995a,b; Lai, 1997; Ivanov and Papaloizou, 2004, 2007; Ogilvie, 2014). Non-hydrostatic motions produced by tidal gravitational field give rise to waves that propagate through interior, with frequencies equal to the tidal forcing frequencies (which are typically smaller than the host star's dynamical frequency). Here we will present a combined description of equilibrium and dynamical tides, the dynamical tides become important for large frequencies.

The energy deposition due to dynamical tides in the limit of high eccentricities is given by (see eq. 26 in Press and Teukolsky (1977))

$$\Delta E = \frac{GM^2}{R} \left(\frac{M_\star}{M} \right)^2 \sum_{\ell=2}^{\infty} \left(\frac{R}{a(1-e)} \right)^{2(\ell+1)} T_\ell(\eta), \tag{1.15}$$

$$\eta = \left(\frac{M}{M_\star + M} \right)^{1/2} \left(\frac{a(1-e)}{R} \right)^{3/2} \tag{1.16}$$

where $T_\ell(\eta)$ are dimensionless functions that express the efficiency of energy deposition.

Moe and Kratter (2018) introduced an analytic prescription of ΔE up to quadruple order (i.e., $\ell = 2$) and approximated $T_2(\eta)$ by a powerlaw

$$\Delta E = f_{dyn} \frac{M_\star + M}{M} \frac{G M_\star^2}{R} \left(\frac{a(1-e)}{R} \right)^{-9} \tag{1.17}$$

where f_{dyn} is a numerical factor, which depends on the polytrope of the planet. For main-sequence stars, with $n = 3$, $f_{dyn} = 0.3$.

The corresponding migration is given by (Moe and Kratter, 2018)

$$\frac{da}{dt} = \frac{a}{P} \frac{\Delta E}{E_{orb}} \tag{1.18}$$

where P is the period of the Jupiter and E_{orb} is its orbital energy.

Once the eccentricity is below a certain value, in Moe and Kratter (2018) taken to be 0.8, equilibrium tides are the more dominant.

For eccentric enough orbits, oscillatory modes might grow chaotically and lead to more efficient migration, circularization and heat injection/extraction (Ivanov and Papaloizou, 2004; Vick and Lai, 2018; Teyssandier et al., 2019; Vick et al., 2019) – chaotic tides. The conditions of the model are that energy transfer is roughly constant over many pericenter passages, occurs instantaneously at pericentre and the mode energy is much smaller than the binding energy.

1.1.4 Hot & warm Jupiters

The discovery of 51 Pegasi b opened a new era in planetary science (Mayor and Queloz, 1995). While it was expected that exoplanets would be similar to the Solar system planets in their properties, 51 Pegasi b was unique in several aspects and opened a new class of planets – hot Jupiters. Hot Jupiters are giant planets with orbital periods $\lesssim 10$ days; since their first discovery, they have been widely observed, partially because of observational biases. However, there is no consensual model of their formation. The suggested formation models of hot Jupiters could be classified roughly into three channels: in-situ formation, disk-migration and high-eccentricity tidal-migration; see a detailed review in Dawson and Johnson (2018). In order to have in-situ formation, gravitational instability or core accretion should take place. The proximity of the current location of hot Jupiters to their host star rules out both of the mechanisms, due to the high temperatures and low disk mass that characterize this environment (Bodenheimer et al., 2000; Rafikov, 2005). In the gas disk-migration scenario (Goldreich and Tremaine, 1980; Lin and Papaloizou, 1986) (see a detailed review in Baruteau et al., 2014), torques from the protoplanetary disk lead to shrinkage of the semimajor axis from several AUs to ~ 0.01AU while maintaining low eccentricity during the migration. High-eccentricity tidal migration requests high eccentricity as an initial condition. There are several mechanisms that lead to high-eccentricity, among them are scattering (Rasio and Ford, 1996; Weidenschilling and Marzari, 1996), Lidov-Kozai oscillations (Lidov, 1962; Kozai, 1962) or combinations of the two. Many models of high-eccentricity tidal migration as a formation channel of hot Jupiters were suggested (e.g Wu and Murray, 2003; Fabrycky and Tremaine, 2007; Naoz et al., 2011a; Petrovich, 2015a,b; Anderson et al., 2016; Hamers et al., 2017).

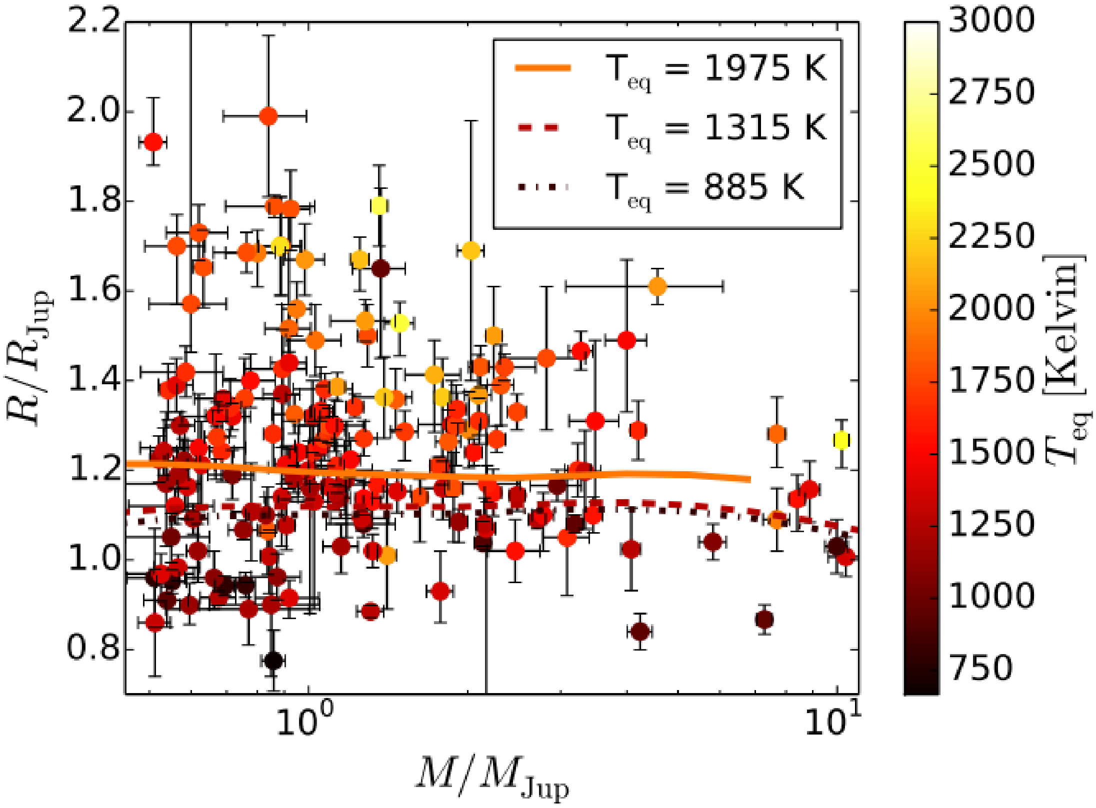

Figure 1.5: Masses and radii of known hot Jupiters, relative to Jupiter and colored by equilibrium temperature. Adopted from Komacek and Youdin (2017).

The size distribution of formed warm and hot Jupiters is wide and includes some radii around $2R_J$. The existence of hot Jupiters with large radii, reaching $\gtrsim 2R_J$, is a smoking gun for thermal inflation. Previous analytical and numerical studies considered various heat sources, including Ohmic heat (Ginzburg and Sari, 2015; Komacek and Youdin, 2017; Komacek et al., 2020).

1.2 Gravitational waves

This section serves as an introduction for chapter 5.1. Gravitational waves (GWs) were predicted first theoretically predicted more than a century ago, by Albert Einstein (Einstein, 1916), who himself doubted the capability of detecting them. Another revolution in Astronomy took place with the direct detection of gravitational waves emitted from a binary black hole merger (GW150914), by the Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) detector in 2016 (Abbott et al., 2016), see figure 1.6.

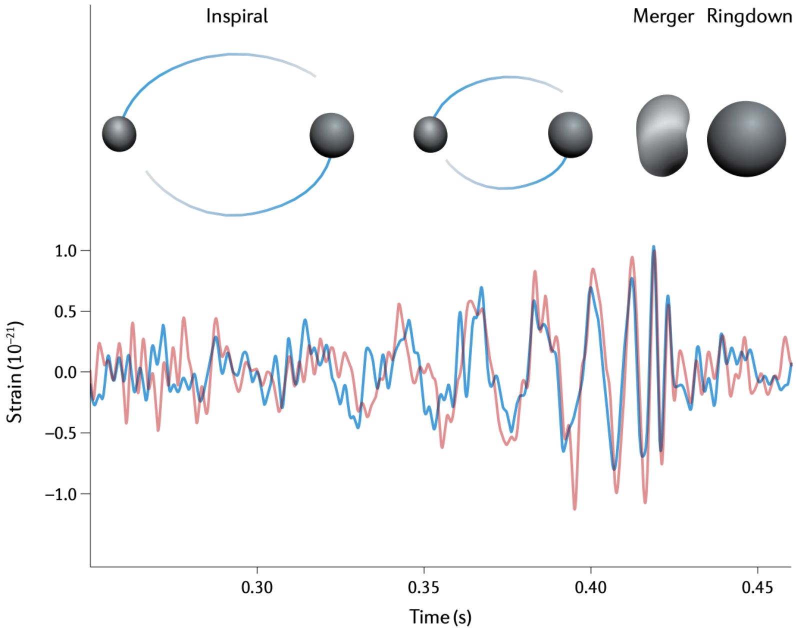

Figure 1.6: The detected gravitational-wave strain amplitude as a function of time for GW150914, the first signal detected nearly simultaneously by the LIGO Hanford and Livingston observatories on September 14, 2015. Figure adapted from `https://creativecommons.org/licenses/by/3.0/`

The detection not only approved our theoretical knowledge but also opened a new window to the Universe, enabling the study of exotic objects from a new point of view. Although now we are shifting from the detection of few events to data-driven studies, the data we have so far is only the tip of the iceberg, and many peculiarities are about to come in other observation epochs, and other instruments that are about to operate in the future.

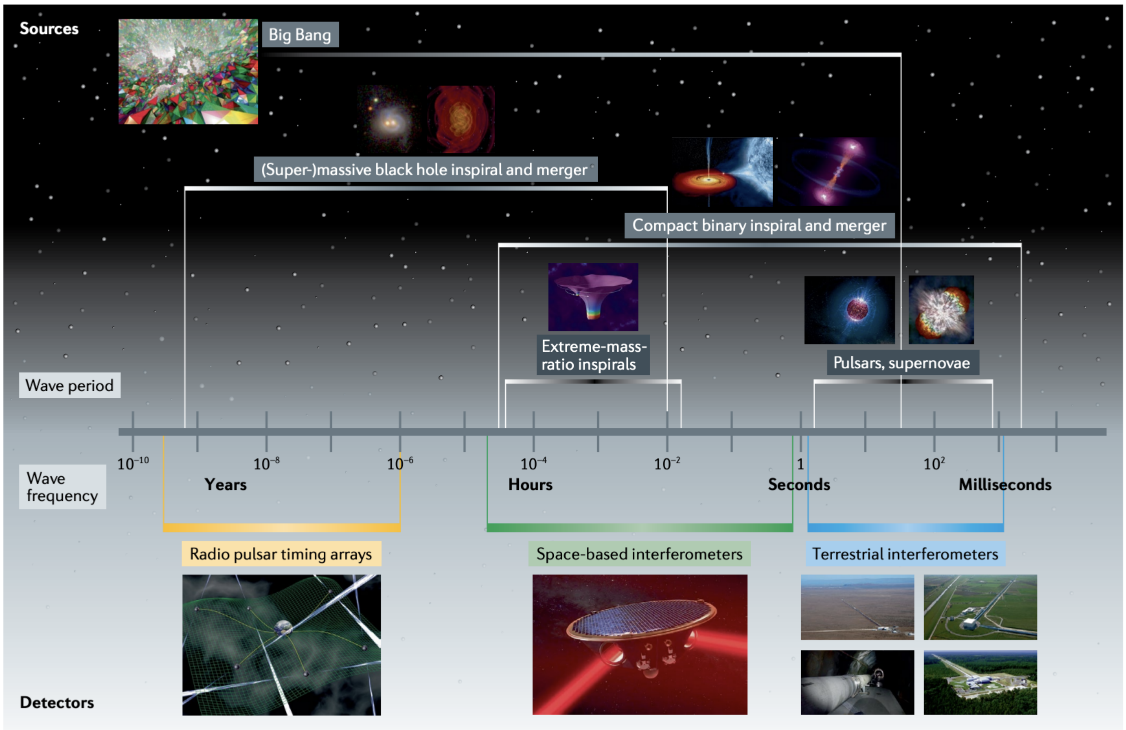

Figure 1.7: The gravitational-wave spectrum probed by strain-sensitive gravitational-wave detectors. Adopted from Bailes et al., 2021.

Gravitational waves emerge from accelerating objects, with a non-symmetric motion.

Various gravitational channels were suggested, including common-envelope assisted mergers (e.g. Dominik et al., 2012; Ginat et al., 2020), chemically homogeneous evolution (de Mink and Mandel, 2016), evolution of triples and quadruples (e.g. Antonini et al., 2017; Silsbee and Tremaine, 2017; Fragione and Antonini, 2019; Michaely and Perets, 2019), gravitational captures (e.g. O'Leary et al., 2009; Rasskazov and Kocsis, 2019; Samsing et al., 2020), and dense star clusters (e.g. Portegies Zwart and

McMillan, 2000; O'Leary et al., 2006; Samsing et al., 2014). Another class of mergers occurs in dissipative environments, such as gaseous environments of AGN disks (e.g. McKernan et al., 2012; Stone et al., 2017; Tagawa et al., 2020). Later on, we will discuss a novel merger channel in gas-rich globular clusters (Rozner and Perets, 2022b).

The merger channels differ from each other by the physical processes that lead to mergers and the environments in which they take place, as well as the different merging objects that leave different signatures.

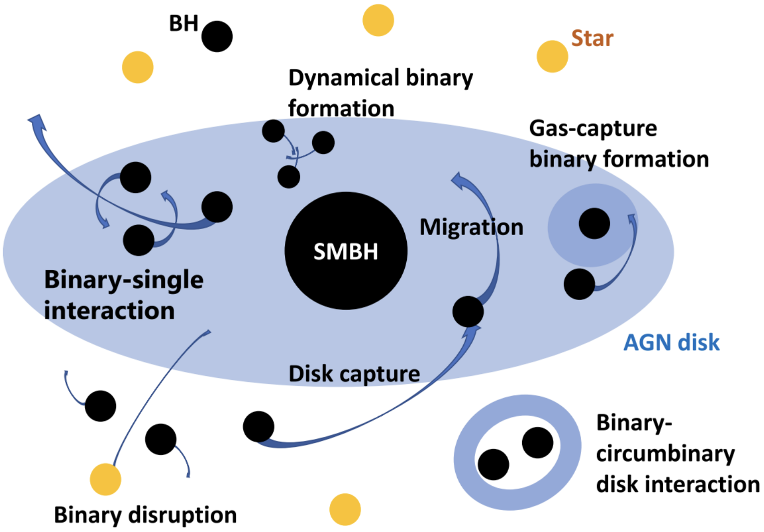

Figure 1.8: Schematic diagram illustrating the mechanisms affecting the BH population and driving binary formation and evolution. Taken from Tagawa et al. (2020).

1.3 Dynamics in gas-rich environments

This section serves as an introduction to chapters 5.1,5.2 and 5.3. Gas-rich environments are abundant in the Universe through different scales and include AGN disks, star-forming regions, protoplanetary disks and regions of late star formation in globular clusters. While these environments are overall very different, the dynamics of objects in gaseous media could in principle be explained using similar physical laws. Dynamics in gas is essentially different than the dynamics in gas-dilute environments, and hence should be treated with different, unique tools.

Gas-rich environments were suggested to serve as a major source of gravitational waves, as energy dissipation is enhanced there (McKernan et al., 2012; Stone et al., 2017; Tagawa et al., 2020; Rozner and Perets, 2022c). Moreover, various dynamical processes are expected to take place in these regions, as the population of stars is affected by the perturbing forces. See a schematic illustration of some of the mechanisms that are expected to take place in the gaseous environment of AGN disks in Fig. 1.8.

There are several suggested approaches to modeling the evolution in gaseous environments. Since I discuss here mostly binaries embedded in gas, I will focus on the dynamics of binaries. Binaries embedded in gas interact with it, exchange angular momentum and energy, and possibly accrete gas. Due to the complexity of these processes, there is still no fully closed unified model for their motion, and there are several approaches used to treat them. One approach is considering their motion under the torques of accretion circumbinary disks, formed due to the accretion to the Hill sphere. In such disks, it was suggested that torques similar to the ones described by type I/II migration of planets in protoplanetary disks could lead to the shrinkage of the binary separation (e.g. Artymowicz et al., 1993; McKernan et al., 2012; Stone et al., 2017; Tagawa et al., 2020). The migration is still highly debated, and some models showed that under some conditions, even outward migration is plausible (e.g. Duffell et al., 2020). Another approach is discussed in Antoni et al. (2019), in which they simulated Bondi–Hoyle–Lyttelton (BHL) supersonic flows and derived the corresponding energy dissipation, fitted to an analytical theory. Here we will focus on a third approach – gas dynamical friction (GDF) (Ostriker, 1999). In the following subsection, we will elaborate further on this approach.

1.3.1 Gas dynamical friction (GDF)

Gas dynamical friction (GDF) is induced by the motion of a massive object through a collisional gas-rich background, and leads to a deceleration of the moving object due to gravitational interactions with the environment. The model is based on dynamical friction, with the proper modifications from collisionless discrete (stellar background) to collisional continuous (gas) medium and different physical cutoffs. In stellar dynamical friction, motion perturbs the distribution of field stars and produces wakes in their densities (Chandrasekhar, 1944; Ostriker, 1999). It could also be thought of as a series of encounters between objects that evolve the systems towards a thermal equilibrium and extract energy and momentum from the massive object by doing that.

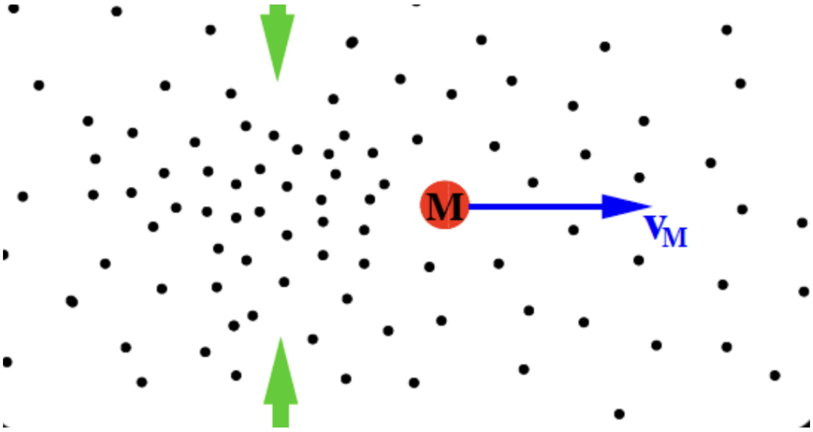

Figure 1.9: An illustration of dynamical friction (taken from lecture notes by Prof. Frank van den Bosch)

The GDF force on an object with mass m is (Ostriker, 1999),

$$\mathbf{F}_{\rm GDF} = -\frac{4\pi G^2 m^2 \rho_g}{v_{\rm rel}^3}\mathbf{v}_{\rm rel} I(v/c_s) \tag{1.19}$$

where G is the gravitational constant, ρ_g is the gas density, c_s is the sound speed, and $\mathbf{v}_{\rm rel}$ is the relative velocity between the object and the gas. The function I is given by

$$I(\mathcal{M}) = \begin{cases} \frac{1}{2}\log(1-\mathcal{M}^{-2}) + \ln\Lambda, & \mathcal{M} > 1 \\ \frac{1}{2}\log\left(\frac{1+\mathcal{M}}{1-\mathcal{M}}\right) - \mathcal{M}, & \mathcal{M} < 1 \end{cases} \tag{1.20}$$

where $\ln\Lambda$ is the Coulomb logarithm.

There is a qualitative difference between the behavior of supersonic and subsonic motions. For subsonic motion, the wake is confined in a sphere behind the perturber, while for the supersonic case, there is an enhanced density wake trailing after the perturber. See a demonstration in Fig. 1.10.

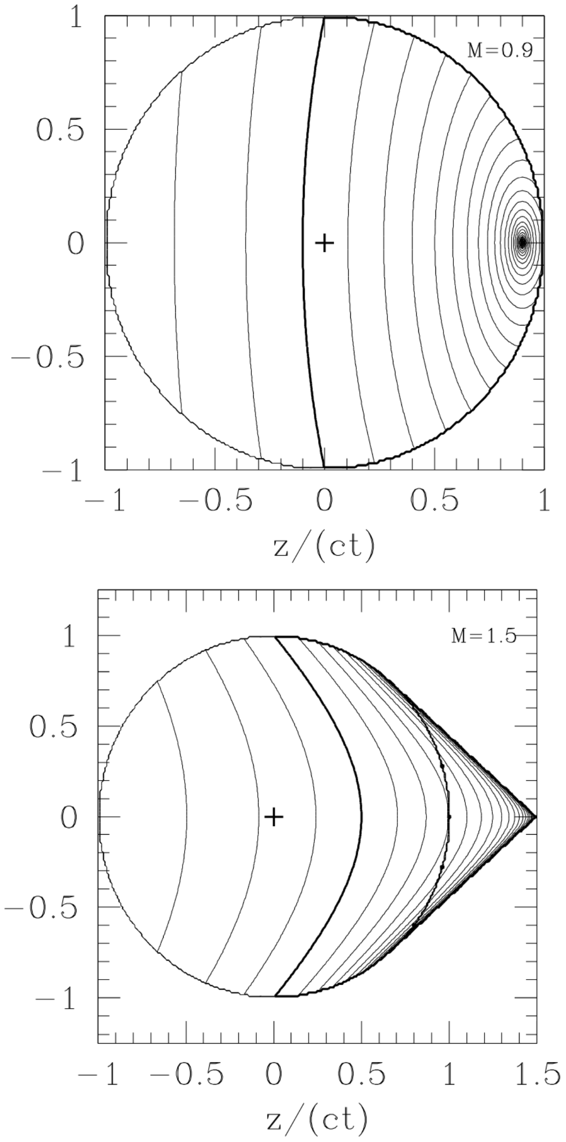

Figure 1.10: Examples of the wake induced in gas dynamical friction for subsonic velocity ($\mathcal{M} = v/c_s < 1$) and supersonic velocity ($\mathcal{M} = v/c_s > 1$). Taken from Ostriker (1999).

1.3.2 Globular clusters as gas-rich environments

Globular clusters were thought for many years to host a simple/single-age population – i.e. to contain stars that were born during a single burst of star formation. However, observations that studied chemical abundances showed that the vast majority of clusters host at least two populations (e.g. Carretta et al. (2009)). Evidence for that includes different chemical abundances of light elements, which is not expected from standard stellar evolution processes. Second-generation stars have enhancements in He, N, and Na abundances, and depletions in C and O with respect to the field stars of the host galaxy (e.g. Carretta et al., 2009 and references therein). The typical age difference between the generations is The existence of a second population might indicate the second significant gas-rich epoch. The origin of the second population is still debated, but in most of them it is suggested that the cluster has a second significant gas-rich epoch. Several models were suggested, including the asymptotic giant branch (AGB) scenario, fast-rotating massive stars and interacting binaries, the early disk accretion scenario the extended cluster formation even and more (see a detailed review in Bastian and Lardo, 2018 and references therein). A different proposed model is that there is only one single generation of star formation, later accompanied by the accretion/depletion of elements that explain the chemical abundance anomalies (Bastian et al., 2013).

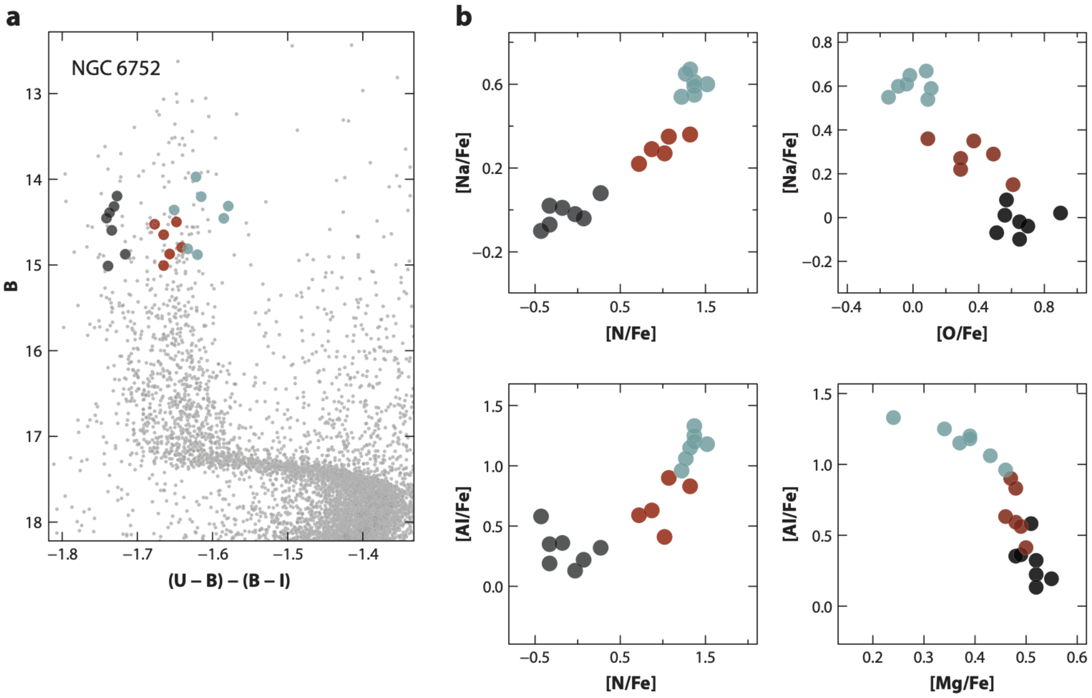

Figure 1.11: Evidence for different chemical abundances. Taken from Bastian and Lardo (2018)

Even from an agnostic point of view regarding the formation channel of the second population, most of the formation channels require high concentrations of gas that allow us to treat clusters at these epochs as gas-rich clusters. This allows us to treat the stellar dynamics in clusters during these epochs as in gas-rich environments.

1.4 Binary formation, evolution & distribution

This section serves as an introduction to chapters 5.2, 5.3 and 6. Binary systems are ubiquitous in a wide range of scales and astrophysical systems, from planetary scales to black holes. Indeed, the majority of stars reside in binaries, or systems of even higher multiplicities (e.g. Raghavan et al., 2010). Hence, they play a crucial role in the dynamics and evolution of stellar and planetary systems. For example, binary interactions could lead to mass transfers, collisions and even mergers between binary components, the formation of exotic stars and compact objects. Compact binaries could give rise also to explosive transient events such as supernovae, gamma-ray bursts and gravitational waves. Higher hierarchies are abundant as well, as discussed in the previous chapters, and for example, every binary in the Kuiper belt could be thought of as a hierarchical triple with the Sun as a third perturber (Perets and Naoz, 2009; Rozner et al., 2020c). Understanding the formation and dynamics of binaries is therefore essential for a complete understanding of stellar and planetary systems.

1.4.1 Binary formation

There are several suggested channels of binary formation, roughly divided into two categories: primordial formation, in which the components of the binary are formed together as a bound system, and dynamical formation in which the components are formed separately, and become a binary later on (see detailed reviews in e.g. Bonnell, 2001; Kratter, 2011; Lee et al., 2020).

Primordial binary formation channels include fragmentation of gas/dust blobs (e.g. Bate et al., 1995, 2002; Clarke, 2009; Offner et al., 2010; Kratter and Lodato, 2016), these produce close binaries either by forming them initially with small separations or with larger separations and later migration induced by energy dissipation from the environment (e.g. by interaction with circumstellar disks Clarke and Pringle, 1991).

The dynamical formation of a binary from an unbound pair requires a dissipative force. Several suggestions for that include tidal force (Press and Teukolsky, 1977), three-body encounters (Aarseth and Heggie, 1976) and dynamical friction Goldreich et al. (2002b). Later in this dissertation, we will discuss a novel formation channel of binaries using gas dissipation (Tagawa et al., 2020; Rozner et al., 2023).

1.4.2 Soft & hard binaries

Binaries in dense environments could be categorized roughly into two categories, based on their energy relative to the mean energy of the environment. Binaries with energies higher than the mean energy are called 'hard binaries', and binaries with smaller energies are called 'soft binaries'. They differ from each other by their evolution even in a qualitative manner – while soft binaries tend to get softer statistically (and even get disrupted) as they get kicks from background stars, hard binaries get statistically harder (Heggie, 1975). Collisions of hard binaries with other stars harden them statistically, and by that pump kinetic energy into the cluster, i.e. cluster heating. This kinetic energy could lead to cluster expansion/halting the collapse of the cluster (Hut, 1983).

While hard binaries got a lot of interest, wide binaries ($a \gtrsim 10^3$ AU) were usually overlooked. The era of GAIA is a renaissance for their studies, as more and more observational data gathers to reveal their unique properties (e.g. El-Badry et al., 2019, 2021).

Wide binaries play an important role in probing large scales, shedding light on the Galactic potential, MACHOS and potential signatures of axion dark matter (e.g. Bahcall et al., 1985; Chanamé and Gould, 2004; Blas et al., 2017; Rozner et al., 2020a). Moreover, external perturbations such as flybys or galactic tidal perturbations could drive them into high eccentricities, leading to the formation of compact binaries/collisional/merger events (e.g. Kaib and Raymond, 2014; Michaely and Perets, 2016; Grishin and Perets, 2022; Michaely and Naoz, 2022). Given their important role in dynamics, studying their origins and properties is essential to our full understanding of cluster dynamics.

Chapter 2

Aeolian erosion in protoplanetary disks

2.1 Aeolian-Erosion in Protoplanetary-Disks I

Based on Rozner et al. (2020b)

As we discussed in section 1.1.1, there are many barriers that objects in protoplanetary disks should overcome in order to become planets. In this study, we introduced another barrier – the aeolian-erosion barrier.

Aeolian-erosion is a purely-mechanical process, discussed usually in geophysics or in context of dunes of terrestrial planets in general (Bagnold, 1941; Kruss et al., 2019), but also was discussed in the context of objects in protoplanetary-discs, from an experimental point of view – especially in recent papers (Paraskov et al., 2006; Schräpler and Blum, 2011; Demirci et al., 2020; Schaffer et al., 2020; Demirci and Wurm, 2020). In our work, we adopted the terminology and the analytical approach from the geophysical literature on wind-erosion in dunes – mostly from the pioneering work of Bagnold (1941); our analog to wind is the gas-drag discussed in section 1.1.1.

There are three main types of wind erosion: suspension, saltation and creeping. Suspension describes the process of wind that swipes particles, and takes them away from the surface far enough such that they won't come back; saltation describes heavier particles that are lifted and then their fallback induces an avalanche of small particles that swiped from the surface; creeping is a rolling of particles that are too heavy to be lifted on the surface; see a detailed discussion in Shao (2008). We claim that the main aeolian-erosion process to take place in protoplanetary disks is the analog for suspension – the scenario we suggest is that aggregates lose their outer layer, which is made of small grains due to the effect of gas-drag. Saltation requires much more significant self-gravity – which is absent in this scenario for objects $\lesssim$ 50km; creeping requires more massive grains than we discuss here.

The aeolian-erosion strongly depends on the velocity of the objects relative to the gas. If this relative velocity is low enough, aeolian-erosion won't be effective. The threshold velocity from which aeolian-erosion induced by gas drag is effective, is determined by the balance between drag force, cohesion and self-gravity. For small objects, where self-gravity is negligible, only the effect of the cohesion of the object should be compared with the drag force. A full derivation of the threshold is described in Shao and Lu (2000). When self-gravity is neglected, the threshold velocity is given by

$$v_\star \approx \sqrt{A_N \frac{\gamma}{\rho_g d}} = 2600 \left(\frac{\rho_g}{3 \times 10^{-9}\ \mathrm{g\ cm^{-3}}}\right)^{-1/2} \left(\frac{d}{0.1\ \mathrm{cm}}\right)^{-1/2} \frac{\mathrm{cm}}{\mathrm{sec}} \tag{2.1}$$

where $A_N = 1.23 \times 10^{-2}$, and $\gamma = 0.165\ \mathrm{g\ s^{-2}}$ are determined empirically from Shao and Lu (2000), the gas-density ρ_g in normalized at 1 au and d is the typical size of the grains that composed the outer (eroded layer) – normalized to 0.1cm. When self-gravity is taken into consideration as well, the expression becomes

$$v_\star = \sqrt{A_N \left(\sigma_p g d + \frac{\gamma}{\rho_g d}\right)}, \tag{2.2}$$

where $\sigma_p = \rho_p/\rho_g$, ρ_p is the density of the object, ρ_p is the density of the gas and the gravitational acceleration is $g = Gm/R^2$.

rtional to d^{-2}). Assuming that the relative velocity is larger than the threshold velocity v_{th}, the typical sweeping time for individual grain is $t_{\mathrm{sw}} \sim v_{\mathrm{rel}}/a_{\mathrm{coh}}$. The work done on the body is $W \sim p \cdot v_{\mathrm{rel}} t_{\mathrm{sw}} A$ where $p = \rho_g v_{\mathrm{rel}}^2/2$ is the dynamic pressure, A is the effective shear surface. The work is equal to the energy loss $\Delta E \sim -\Delta m \cdot v_{\mathrm{rel}}^2/2$. Therefore, the net mass change is $\Delta m \sim -A\rho_g \cdot v_{\mathrm{rel}}^2/a_{\mathrm{coh}}$. The effective shear surface is only linearly proportional to the size R, since only a thin layer of width $v_{\mathrm{rel}}\Delta t$ is affected by the wind for small enough time $\Delta t \lesssim t_{\mathrm{sw}}$, thus $A \sim R v_{\mathrm{rel}} \Delta t$. In the limit of $\Delta t \to 0$, the differential equation for the mass loss rate is then

$$\frac{dm}{dt} = -\frac{\rho_g}{a_{\rm coh}} v_{\rm rel}^3 R \propto -\rho_g \rho_p v_{\rm rel}^3 d^2 R \tag{2.3}$$

or in terms of radius,

$$\frac{dR}{dt} = -\frac{\rho_g v_{\rm rel}^3}{4\pi R \rho_p a_{\rm coh}}. \tag{2.4}$$

Here, m_d is the mass of the released grain of size d, where we assume that the densities of both the grain and the entire eroding body are the same.

The timescale for the aeolian-erosion of an object to half its size can be approximated by

$$t_{\rm ero} = \frac{R}{|\dot{R}|} = \frac{4\pi R^2 \rho_p a_{\rm coh}}{\rho_g v_{\rm rel}^3} \tag{2.5}$$

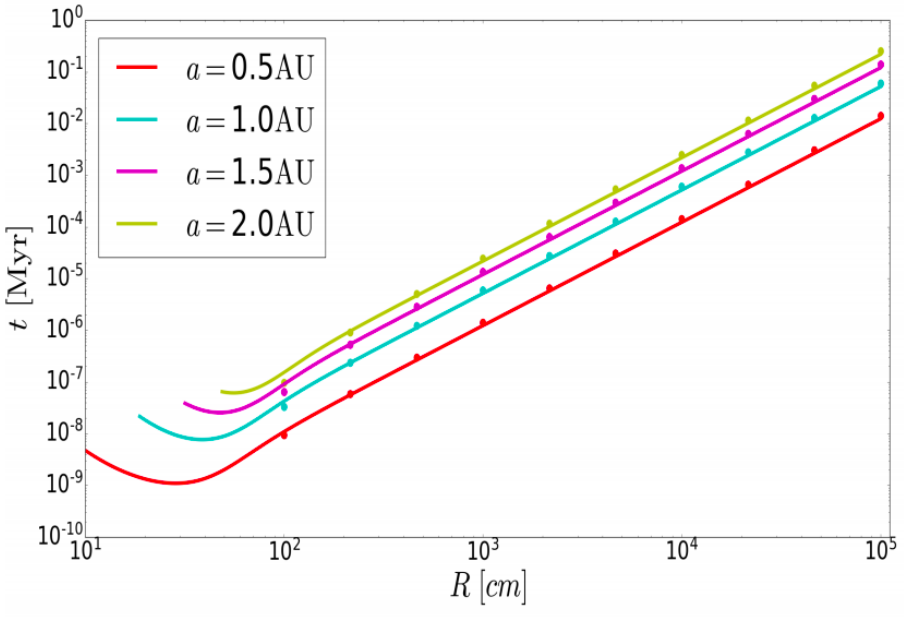

Figure 2.1: The characteristic timescales for the aeolian-erosion as a function of its initial radius. Each line represents different starting separation. For smaller sizes the time is truncated by the separation-dependent limit where there is no erosion, for dust size of 0.1cm. The dots correspond to timescales from numerical simulation. Adopted from Rozner et al. (2020b).

Aeolian-erosion in protoplanetary disks turns out to be very rapid and efficient. Moreover, its effect reaches up to $\sim 10^{5}$cm objects, and these objects are grinded down to ~ 10cm within the protoplanetary-disk lifetime.

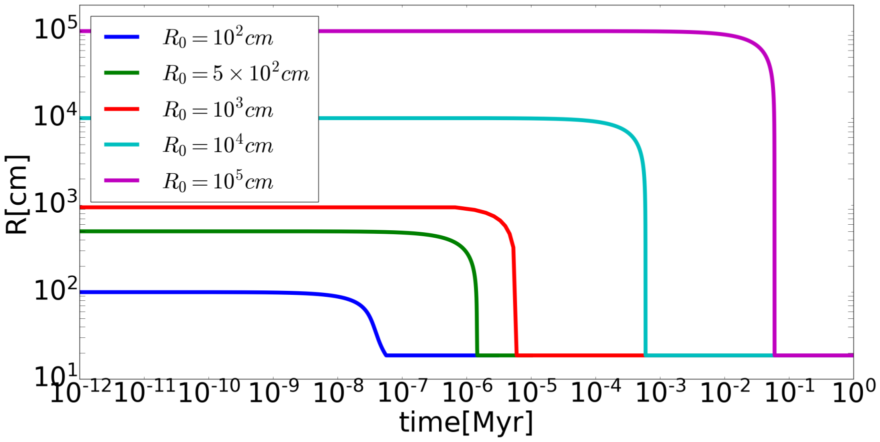

Figure 2.2: The dependence of the evolution on the initial radii of the bodies at a fixed distance of $a = 1$ AU from the star. Adopted from Rozner et al. (2020b).

As can be seen from eq. 2.4, the rate of aeolian-erosion is proportional to the density of the gas; since the density of the gas decreases with the distance from the center of the disk, aeolian-erosion is most effective in short distances.

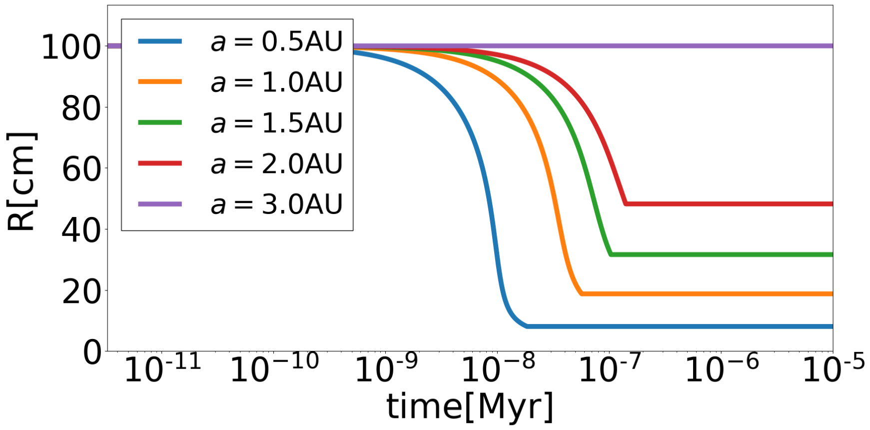

Figure 2.3: The dependence of the evolution on the initial radii of the bodies at a fixed distance of $a = 1$ AU from the star. Adopted from Rozner et al. (2020b).

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **496,** 4827–4835 (2020)
Advance Access publication 2020 June 27

doi:10.1093/mnras/staa1864

The aeolian-erosion barrier for the growth of metre-size objects in protoplanetary discs

Mor Rozner,[⋆] Evgeni Grishin and Hagai B. Perets
Technion – Israel Institute of Technology, Haifa 3200003, Israel

Accepted 2020 June 24. Received 2020 June 24; in original form 2019 October 5

ABSTRACT

Aeolian erosion is a destructive process that can erode small-size planetary objects through their interaction with a gaseous environment. Aeolian erosion operates in a wide range of environments and under various conditions. Aeolian erosion has been extensively explored in the context of geophysics in terrestrial planets. Here we show that aeolian erosion of cobbles, boulders, and small planetesimals in protoplanetary discs can constitute a significant barrier for the early stages of planet formation. We use analytic calculations to show that under the conditions prevailing in protoplanetary discs small bodies (10–10^4 m) are highly susceptible to gas-drag aeolian erosion. At this size-range aeolian erosion can efficiently erode the planetesimals down to tens-cm size and quench any further growth of such small bodies. It thereby raises potential difficulties for channels suggested to alleviate the metre-size barrier. Nevertheless, the population of ∼decimetre-size cobbles resulting from aeolian erosion might boost the growth of larger (>km size) planetesimals and planetary embryos through increasing the efficiency of pebble-accretion, once/if such large planetesimals and planetary embryos exist in the disc.

Key words: comets: general – minor planets, asteroids: general – planets and satellites: formation.

1 INTRODUCTION

The growth of dust aggregates and sub-cm size pebbles in protoplanetary discs can be understood theoretically and experimentally (Wurm & Blum 1998). The growth of km-size objects or larger planetary embryos to fully formed planets could also be efficient, and possibly proceed through mechanisms such as a pebble-accretion (Ormel & Klahr 2010; Perets & Murray-Clay 2011; Lambrechts & Johansen 2012). However, the growth of pebbles, cobbles (up to 25 cm), and boulders in the intermediate regime from ∼cm to metre up to km-size planetesimals is not well understood. Several physical processes potentially quench planetesimal growth in this size range. These growth barriers include fast radial drift on to the host star of (typically) cm–metre-size bodies at few-au scales (Adachi, Hayashi & Nakazawa 1976; Weidenschilling 1977), and inefficient growth of dust-aggregates, pebbles, cobbles, and boulders due to collisional fragmentation and erosion (Blum & Wurm 2000; Brauer, Henning & Dullemond 2008; Birnstiel, Dullemond & Brauer 2010; Güttler et al. 2010; Krijt et al. 2015), leading to the so-called metre-size barrier.

Several solutions to the metre-size barrier had been proposed, including particle trapping eddies (Klahr & Henning 1997), instabilities in turbulent discs near the snow line (Brauer et al. 2008), and collisional growth (Windmark et al. 2012). Recently, Grishin, Perets & Avni (2019) have suggested that planetesimals can be exchanged and captured between protoplanetary discs, some of them already on ∼km-size scale. Only a tiny fraction of protoplanetary discs are required to form planetesimals in situ in order to ‘seed’ the entire birth cluster with planetesimals. Thus, the formation of the first planetesimals can be an exponentially rare event, consistent with the various fine-tuned models for planetesimal formation. Pfalzner & Bannister (2019) had suggested to take the seeding model one step further and start with a population of planetesimals already at the stage of star formation and collapse of giant molecular clouds.

The streaming instability (Youdin & Goodman 2005; Johansen & Youdin 2007), where the coupled dust-gas evolution clumps dust at localized regions and eventually leads to direct gravitational collapse, is a promising route to planetesimal formation, although it requires fine-tuned conditions, such as large initial metallicity (see recent review by Blum 2018, and references therein). Yang, Johansen & Carrera (2017) showed that a slightly above Solar metallicity of only $Z \approx 0.03$ is required for streaming instablity for optimal range Stokes numbers St ∼ 0.1. However, recently Krapp et al. (2019) showed that some range of a mass distribution of particles slows down the growth of unstable modes, and does not converge with the number of species sampled, thus poses severe limitations for the onset of streaming instability.

Here we identify an additional, erosion-induced barrier for planetesimal growth. This physical process efficiently erodes bodies in the size range of 1 m to 1 km embedded in the gaseous protoplanetary disc (depends on the parameters of the disc and of the object). The erosion-induced barrier effectively makes the metre-size barrier into ∼100 m size barrier, thereby challenges the collisional growth models and support the direct collapse models into km-sized planetesimals, such as the streaming instability. Aeolian erosion in protoplanetary discs, currently not included in dust and planetary

---

[⋆] E-mail: morozner@campus.technion.ac.il

Published by Oxford University Press on behalf of the Royal Astronomical Society

standard growth models, significantly affects the evolution and growth of sub-km bodies embedded in the discs and their size distribution.

Aeolian erosion is a completely mechanical process, discussed usually in the context of dunes of terrestrial planets (Bagnold 1941; Kruss et al. 2019), but also was discussed in the context of objects in protoplanetary discs, from an experimental point of view (Paraskov, Wurm & Krauss 2006; Schräpler & Blum 2011). There are three main types of erosion: suspension, saltation, and creeping. Suspension describes the process of wind that swipes particles, and takes them away from the surface; saltation describes heavier particles that are lifted and then their fallback induces an avalanche of small particles that swiped from the surface; creeping is a rolling of particles that are too heavy to be lifted on the surface; see a detailed discussion in Shao (2008). The main erosion type in protoplanetary discs is suspension, since saltation requires significant self-gravity and creeping involves relatively massive grains (Paraskov et al. 2006).

Here we study aeolian erosion in protoplanetary discs and show that it may have far-reaching implications for planet formation. It gives rise to significant mass-loss from cobbles and boulders rock-size bodies, up to the level of quenching their growth and critically grinding them down to decimetre-size cobbles. In the following, we analyse the effects of aeolian erosion in discs on such small-sized objects, and explore the symbiotic relations between aeolian erosion and other dominant physical processes that take place in discs.

The paper is organized as follows: In Section 2, we introduce the settings and the phenomena of aeolian erosion in protoplanetary discs. We discuss the characteristic time-scales of aeolian erosion and dynamical evolution. In Section 3, we discuss the symbiotic relations between aeolian erosion and other dominant physical processes that take place in discs. Finally, in Section 5, we summarize our results and discuss future implications.

2 GAS DRAG AND AEOLIAN EROSION IN PROTOPLANETARY DISCS

2.1 Drag laws

Objects in gaseous protoplanetary disc with density $\rho_{\rm gD}$ experience aerodynamic drag force, expressed by the drag law

$$\boldsymbol{F}_{\rm D} = \frac{1}{2} C_{\rm D}({\rm Re})\pi R^2 \rho_{\rm g} v_{\rm rel}^2 \hat{\boldsymbol{v}}_{\rm rel}, \tag{1}$$

where R is the radius of the object and $v_{\rm rel}$ is the object's velocity relative to the gas. The drag coefficient, $C_{\rm D}$, depends on the geometry of the particle and the relative velocity. For spherical bodies, the drag coefficient depends only on the Reynolds number Re.

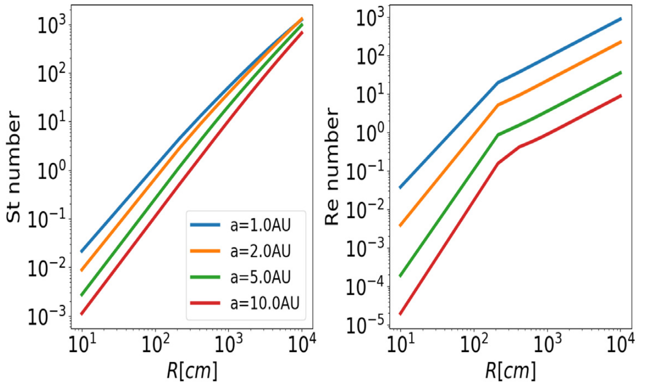

Figure 1. Stokes and Reynolds numbers dependence on the radius of the object. Note the different *y*-axis.

The motion of an object in gaseous disc can be determined by the relative velocity $v_{\rm rel}$, the particle size R, and the distance to the star. Due to radial pressure gradient in the disc, gas in the disc revolves in sub-Keplerian velocities, $v_{\rm gas} - v_{\rm k} \approx \eta v_{\rm k}$ where $v_{\rm gas}$ is the azimuthal velocity of the gas, $v_{\rm k}$ is the Keplerian velocity, and $\eta \sim (c_{\rm s}/v_{\rm k})^2$ is the small correction due to pressure gradients and $c_{\rm s}$ is the speed of sound (Perets & Murray-Clay 2011, see Table A1 for exact expressions). Very small objects strongly coupled to the gas and move with it, while very large objects are little affected by the gas. In the intermediate regime, cobbles/boulders orbiting at sub-Keplerian velocities experience 'headwind' from the gas in the disc slowing them down. Such objects could therefore lose angular momentum and inspiral to the inner parts of the disc (Weidenschilling 1977). Using polar coordinates, the components of the relative velocity between the object and gas are given by (e.g. Perets & Murray-Clay 2011)

$$v_{\rm rel,r} = -\frac{2\eta v_{\rm k}{\rm St}}{1+{\rm St}^2},\ v_{\rm rel,\phi} = -\eta v_{\rm k}\left(\frac{1}{1+{\rm St}^2}-1\right), \tag{2}$$

where the Stokes number is defined by

$${\rm St} = \Omega t_{\rm stop};\ t_{\rm stop} = \frac{m v_{\rm rel}}{F_{\rm D}}, \tag{3}$$

where Ω is the angular Keplerian velocity. $F_{\rm D}$ is the drag force (see equation 1). For the drag coefficient, we adopted the fitting used in Perets & Murray-Clay (2011)

$$C_{\rm D}({\rm Re}) = \frac{24}{\rm Re}(1+0.27{\rm Re})^{0.43} + 0.47\left[1-\exp\left(-0.04{\rm Re}^{0.38}\right)\right]. \tag{4}$$

The fitting formula is valid for a wide range of Reynolds numbers, $10^{-3} < {\rm Re} < 10^5$ (Brown & Lawyer 2003), which covers most of the drag regimes. In particular, equation (4) covers the ram-pressure and Stokes regimes (Weidenschilling 1977). In the ram-pressure regime, ${\rm Re} \gg 1$, $C_{\rm D} \approx 0.47$, while in the Stokes regime, ${\rm Re} < 1$, $C_{\rm D} \rightarrow 24/{\rm Re}$. In the intermediate regime, where $1 < {\rm Re} < 800$, $C_{\rm D} \propto {\rm Re}^{-3/5}$.

Fig. 1 shows the Stokes and Reynolds number are an increasing function of the size of the object. Stokes and Reynolds numbers correspond to the coupling to the gas in the disc, which becomes weaker for larger objects.

2.2 Aeolian erosion threshold and time-scales

The balance between drag force, cohesion, and self-gravity dictates a lower velocity threshold from which the drag-force can resist the self-gravity and cohesion and discharge particles from the surface of objects. For small objects, where self-gravity is negligible, only the effect of the cohesion of the object should be compared with the drag force. A full derivation of the threshold is described in Shao & Lu (2000). The threshold velocity in the limit of small objects is then

$$v_\star \approx \sqrt{A_N \frac{\gamma}{\rho_{\rm g} d}} = 2600 \left(\frac{\rho_{\rm g}}{3\times10^{-9}\ {\rm g\ cm^{-3}}}\right)^{-1/2}\left(\frac{d}{0.1\ {\rm cm}}\right)^{-1/2}\frac{\rm cm}{\rm s} \tag{5}$$

where $A_N = 1.23 \times 10^{-2}$, and $\gamma = 0.165$ g s^{-2} are determined empirically from Shao & Lu (2000), and the gas density in normalized

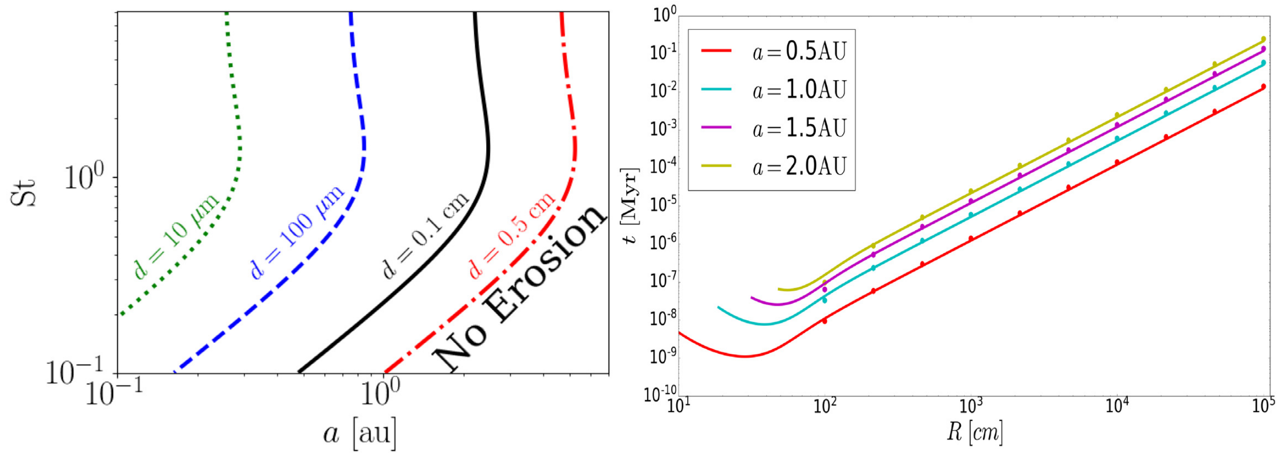

Figure 2. Left: Areas on the separation–Stokes number plane where erosion is effective. Transitional lines for dust size of 10, 100 μm, 0.1 and 0.5 cm are represented by the dotted green, dashed blue, solid black, and dash–dotted red, respectively. Areas to the right with large a are where erosion is ineffective. Right: The characteristic time-scales for the aeolian erosion as a function of its initial radius. Each line represents different starting separation. For smaller sizes the time is truncated by the separation-dependent limit where there is no erosion, for dust size of 0.1 cm. The dots correspond to time-scales from numerical simulation.

at 1 au. A_N is a function of the Reynolds number, and includes the friction force, which is significantly lower than the cohesion force and scales as d/R (Zimon 1982).

The size d can range from micron sized dust to larger grains of ~ 1 mm. Radio and infra-red observations on protoplanetary discs show abundance of mm-sized grains with a typical power-law size distribution (Andrews 2009, 2015). Moreover, the wind-tunnel experiments of dust aeolian erosion that we compared our results to were performed for grains of size of 0.5 mm. We therefore choose a canonical size of $d = 0.1$ cm $= 1$ mm, although the results are generic and explored for a range of grain sizes (Figs 2 and 7).

The threshold velocity sets a regime of typical velocities in which objects in protoplanetary discs could be significantly affected by aeolian erosion. Note that the threshold strongly depends on the gas density and the typical size of the swept particles.

The general expression for the threshold velocity contains a self-gravity term as well (Shao & Lu 2000)

$$v_\star = \sqrt{A_{\rm N}\left(\sigma_{\rm p} g d + \frac{\gamma}{\rho_{\rm g} d}\right)}, \tag{6}$$

where $\sigma_{\rm p} = \rho_{\rm p}/\rho_{\rm g}$ and the gravitational acceleration is $g = Gm/R^2$. The contribution from self-gravity becomes comparable to the contribution from cohesion just for objects with sized $R \gtrsim 50$ km.

$$\frac{\gamma/(\rho_{\rm g} d)}{\sigma_{\rm p} g d} = \frac{\gamma}{\rho_{\rm p} g d^2} = \frac{3\gamma}{4\pi\rho_{\rm p}^2 R G d^2} \approx \frac{5\times 10^4}{R d^2/\mathrm{cm}^3}, \tag{7}$$

which means that for $d = 0.1$ cm, the critical radius in which the self-gravity becomes important is $\sim 5\times10^6$ cm $= 50$ km.

The question whether self-gravity can play a role and keep a binary stable (even for extreme mass ratios of a boulder and a dust grain) had been addressed in Perets & Murray-Clay (2011). Their fig. 2 shows the regions where a binary is stable against the shearing from the wind (i.e. wind-shearing radius – WISH). For small micron-sized grains they are tightly coupled to the gas, thus their WISH radius is smaller than the physical radius of the boulder. Thus, once liberated from the boulder, they are immediately sheared apart. The white zones in fig. 2 of Perets & Murray-Clay (2011) are where the WISH radius larger than the physical boulder size, but smaller than the Hill radius (Grishin et al. 2017), where the Solar tide shears the binary apart. For 1 au, only planetsimals above 1 km can have 0.1 cm bound dust grains, while smaller boulders cannot retain the liberated grains. The minimal size of the planetesimal will decrease to ~ 1 km for larger dust sizes of 1 cm or for larger location of ~ 5 au. Further than that, aeolian erosion is ineffective and we do not deal with larger distances and grains. To summarize, only planetesimals of size $\gtrsim 1$ km could keep the dust grains bound to them, otherwise the grains are essentially lost to the wind once they overcome the cohesion forces.

The velocity profile changes for different streamlines in the flow. Far from the surface of the body, the velocity is the free streaming velocity $v_{\rm rel}$, dictated by the size of the object (equation 2). Close to the surface, boundary layer effects might change the relative velocity, which may even vanish if the no-slip condition is applied. Nevertheless, the flow around the object is well approximated by a shear flow, and erosion occurs when the shear stress overcomes the cohesion forces. It is possible to define an effective friction velocity, $u^\star$ that measures the strength of the shear stress (e.g. Demirci et al. 2020). While the friction velocity is somewhat lower than the free streaming velocity for large Reynolds numbers (Greeley et al. 1980), they are practically indistinguishable for lower Reynolds numbers. Indeed, we follow Shao & Lu (2000), where they use the friction velocity both for the erosion threshold and for the typical relative velocity for the drag forces.

Above the threshold velocity, the shear pressure induces a mass-loss in rate. The mass-loss rate was derived in Bagnold (1941) for dust saltation or erosion on planetary bodies. We modify the Bagnold (1941) and replace the retaining force from self-gravity to cohesion. The heuristic derivation is as follows.

Consider a wind of density $\rho_{\rm g}$ and velocity $v_{\rm rel}$ blows upon a body of size R. The cohesive acceleration that holds the grains together is $a_{\rm coh}$ (which is proportional to d^{-2}). Assuming that the relative velocity is larger than the threshold velocity $v_{\rm th}$, the typical sweeping time for individual grain is $t_{\rm sw} \sim v_{\rm rel}/a_{\rm coh}$. The work done on the body is $W \sim p v_{\rm rel} t_{\rm sw} A$ where $p = \rho_{\rm g} v_{\rm rel}^2/2$ is the dynamic pressure, A is the effective shear surface. The work is equal to the energy loss $\Delta E \sim -\Delta m v_{\rm rel}^2/2$. Therefore, the net mass change is $\Delta m \sim -A \rho_{\rm g} v_{\rm rel}^2/a_{\rm coh}$. The effective shear surface is only linearly proportional to the size R, since only a thin layer of width $v_{\rm rel}\Delta t$ is affected by the wind for small enough time $\Delta t \lesssim t_{\rm sw}$, thus $A \sim R v_{\rm rel} \Delta t$. In the limit of $\Delta t \to 0$, the differential equation for the mass-loss rate is then

$$\frac{{\rm d}m}{{\rm d}t} = -\frac{\rho_{\rm g}}{a_{\rm coh}} v_{\rm rel}^3 R \propto -\rho_{\rm g} \rho_{\rm p} v_{\rm rel}^3 {\rm d}^2 R \tag{8}$$

or in terms of radius,

$$\frac{{\rm d}R}{{\rm d}t} = -\frac{\rho_{\rm g} v_{\rm rel}^3}{4\pi R \rho_{\rm p} a_{\rm coh}}. \tag{9}$$

Here, m_d is the mass of the released grain of size d, where we assume that the densities of both the grain and the entire eroding body are the same. Unless stated otherwise, we consider the aeolian erosion of cobbles and boulders with $\rho_{\rm p} = 3.45 \ {\rm g\,cm^{-3}}$, that correspond to rocky objects as described in Pollack et al. (1996) . Shao & Lu (2000) note that the cohesion force is linearly proportional to the grain size, $F_{\rm coh} = m_{\rm d} a_{\rm coh} = \beta d$. The numerical value of β is uncertain. Shao & Lu (2000) have investigated the strength of the cohesive acceleration in wind tunnel experiments. They relied on early experiments of Phillips (1980) for powder particles with relatively weak cohesion and $\beta \approx 10^{-2}$ g s^{-2} (10^{-5} N m^{-1}). On the other hand, Paraskov et al. (2006) refers to stronger cohesion, which was measured by atomic force microscopy by Heim et al. (1999) where the force to separate $\sim\mu$-sized grains was around 10^{-7} N which leads to $\beta \approx 10^2$ g s^{-2} (10^{-1} N m^{-1}). We continue the expected linear scaling and adopt the value of stronger $\beta = 10^2$ g s^{-2} here. Here we only consider objects which composition behaves like loose soil. Objects of more complex compositions, such as, e.g. ice-coated objects, might behave chemically/physically different, and not allow for wind-driven erosion.

Given the strong dependence of the aeolian erosion rate on the relative velocity, the density profile of the disc, the appropriate (size dependent) Stokes and Reynolds numbers as well as η play a significant role in modelling aeolian erosion. The peak of relative velocity is $\sim \eta v_{\rm k}$, and henceforth even a small difference in η can introduce significant changes in the aeolian-erosion rate. The aeolian-erosion dependence can be non-trivial, due to the mutual dependence of the relative velocity and the Stokes number (see Appendix A for further details).

The time-scale for the aeolian erosion of an object to half its size can be approximated by

$$t_{\rm ero} = \frac{R}{|\dot{R}|} = \frac{4\pi R^2 \rho_{\rm p} a_{\rm coh}}{\rho_{\rm g} v_{\rm rel}^3}. \tag{10}$$

The relative velocities and gas density depend on the model of the gaseous disc. Our disc models follow those used in our previous papers (Perets & Murray-Clay 2011; Grishin & Perets 2015) where the gas density surface profile is the minimal mass solar nebula (MMSN), $\Sigma_{\rm g} = 2 \times 10^3 (a/{\rm au})^{-3/2}$ g cm^{-2}. The aspect ratio is $h/a = 0.022(a/{\rm au})^{2/7}$. The relation of the surface density and the aspect ratio lead to the gas density of $\rho_{\rm g} \approx 3 \times 10^{-9} (a/{\rm au})^{-16/7}$. The gas-pressure support parameter is $\eta \approx 2 \times 10^{-3} (a/{\rm au})^{4/7}$ and the temperature profile is $T = 120(a/{\rm au})^{-3/7}$.

Given our disc model, for a metre-sized object, this time-scale is about 1 yr. Note that this expression is a crude estimate for the time-scale, as it does not take into account the dynamics of the problem.

The left-hand panel of Fig. 2 shows the area in on the a–St plane where the relative velocity is larger than the threshold velocity, and erosion can take plane. More distant objects have higher threshold velocity due to the lower local gas density, hence the erosion is mostly effective in the inner disc; moreover, beyond the ice line the behaviour of the cohesion law might change, in our scope we assume that the only change is the density of the object. Smaller dust grains also increase the threshold velocity. The maximal separation where erosion is allowed occurs when the relative velocity is maximal, which occurs at ${\rm St} = \sqrt{2}$. For typical size of $d = 0.1$ cm the erosion takes place only within the snow line, $a \lesssim 2.7$ au.

In the right-hand panel of Fig. 2, we show the characteristic time-scales determined by equation (10). At a fixed distance from the centre of 1 au, aeolian erosion is most effective around $\sim 10^2$ cm, and for these sizes, embedded objects can be eroded down to half their size in $\sim 10^{-8}$ Myr. More generally, objects in the size range of ~ 1–10^4 cm can be eroded down to half their initial radii over a typical lifetime of a protoplanetary disc or less. Note that here we present just a rough estimate for the time-scales; a more detailed discussion on these issues follows below.

2.3 Dynamical evolution

In order to study the evolution of objects under the influence of aeolian erosion, we integrate equation (9) numerically using a Runge–Kutta integrator. We use our disc model and dust size of $d = 0.1$ cm, unless stated otherwise. Fig. 3 presents the time evolution of objects with various initial radii and distances from the star.

For a fixed distance of $a = 1$ au, objects of sizes ~ 1–10^4 cm on circular orbits are eroded significantly down to a size of ~ 15 cm. Aeolian erosion of larger objects takes more time, objects of 10^3 cm will be eroded in ~ 1 yr. We find that bodies with initial metre size are eroded significantly up to $\lesssim 2.7$ au from the star, during the typical lifetimes of gaseous protoplanetary discs.

As can be seen in Fig. 3, given sufficient time, the embedded bodies are eventually eroded to a typical final size, at which point the bodies could not be eroded because their relative velocity is smaller than the threshold velocity for erosion (see Fig. 2). Note that objects on eccentric orbits could experience higher headwind velocities even at these small sizes, and therefore dynamical excitation of planetesimals could strengthen the effects of protoplanetary disc erosion.

The self-gravity sets an upper limit for the erosion of objects in protoplanetary discs. As can be seen from Fig. 4, objects of size $\gtrsim 10^4$ cm do not erode efficiently during the disc lifetime.

3 EFFECTS OF RADIAL DRIFT AND TURBULENCE

Other physical processes occur in the young protoplanetary disc and potentially couple to the effects of the aeolian erosion, in particular turbulence and radial drift. In the following, we discuss some of these aspects.

Radial drift due to gas drag is thought to be one of the most dominant processes in the disc. It could potentially lead to in-spiral of the cobbles, boulders, and planetesimals towards the star and their possible destruction over short time-scales of about $10^3 - 10^4$ yr.

In a steady state, the equations of motion in the presence of gas drag can be solved self-consistently (e.g. Perets & Murray-Clay 2011). The radial drift is given by steady-state solution of the radial

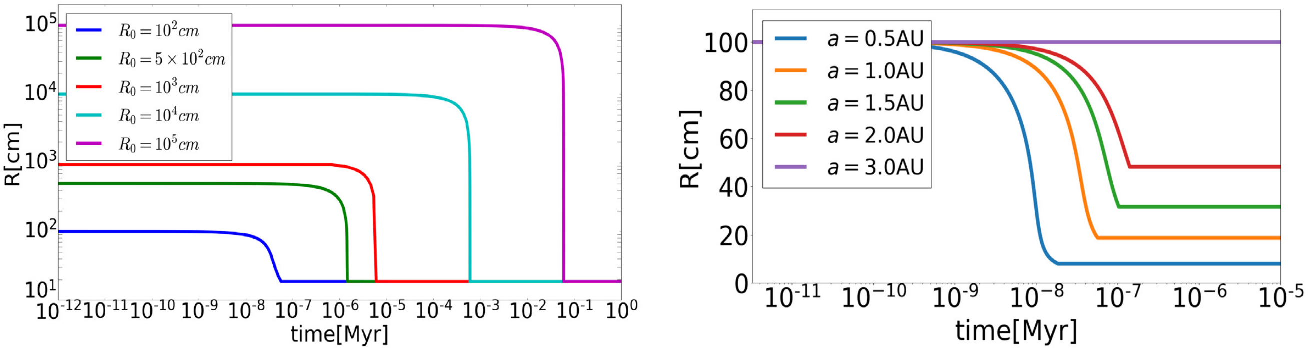

Figure 3. The evolution of the size of objects embedded in the protoplanetary disc due to aeolian erosion. Left: Dependence of the evolution on the initial radii of the bodies at a fixed distance of $a = 1$ au from the star. Right: Dependence of the evolution at different distances from the star for objects with a fixed initial radius of 10^2 cm.

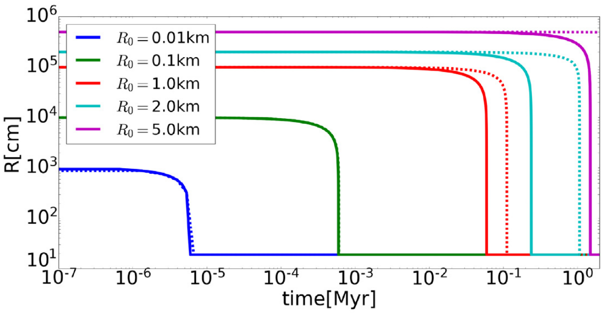

Figure 4. The evolution of the size of objects embedded in the protoplanetary disc due to aeolian erosion including self-gravity. The dependence of the evolution on the initial radii of the bodies at a fixed distance of $a = 1$ au from the star. Solid lines do not take into account self-gravity, and dashed lines do.

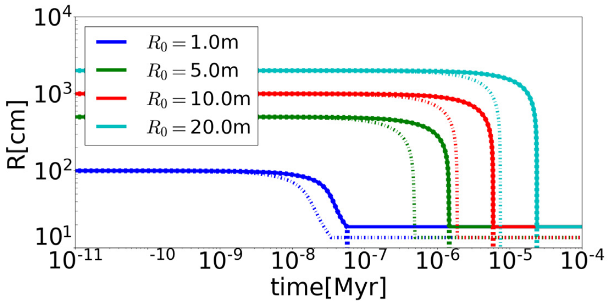

Figure 5. Effects of radial drift and turbulence. Solid lines are the same as in Fig. 3. Dot–dashed lines represent the evolution including radial drift, which is indistinguishable from the solid lines. Dashed line represent the evolution including turbulent velocities.

velocity:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -v_r = -\frac{2\eta v_k \mathrm{St}}{1+\mathrm{St}^2}, \tag{11}$$

where v_r is the radial component of the relative velocity.

Fast radial drift, which peaks at St ≈ 1 (decimetre to metre-size objects for distances of 1 au), can in principle enhance the aeolian erosion.

Prima facie, inspiral in the disc transfers objects to the inner regions of the disc, where the radial gas density increases and the aeolian erosion is more effective. However, the time-scales of aeolian erosion are shorter/comparable to these of radial drift.

In order to study the importance of turbulent velocities for aeolian erosion, we parametrize the strength of turbulence in the disc using the standard Shakura–Sunyaev α prescription describing the effective kinematic viscosity, here taken to be $\alpha = 0.01$ and constant during the evolution. The effective kinematic viscosity of the turbulent gas is then given by Shakura & Sunyaev (1973), $\nu = \alpha c_\mathrm{s} H_\mathrm{g}$, where $H_\mathrm{g} = a(c_\mathrm{s}/v_\mathrm{k})$ is the scale height of the gas. The turbulent velocity of the largest scale eddies is $v_\mathrm{t} = \sqrt{\alpha} c_s$. The turbulence adds a non-zero root mean square velocity, i.e. $\langle \delta v^2 \rangle = \langle \delta v_\mathrm{rel}^2 \rangle + \langle v_\mathrm{turb}^2 \rangle$. Ormel & Cuzzi (2007) derived an analytical expression for the relative velocity between particle and gas in turbulence; although this derivation, which is rooted in the work of Cuzzi & Hogan (2003) uses the assumption of St ≪ 1, it turns out to work over a wider range of Stokes numbers, and gives

$$v_\mathrm{p,t}^2 = v_\mathrm{t}^2 \left(\frac{\mathrm{St}^2 (1 - \mathrm{Re_t}^{-1/2})}{(\mathrm{St}+1)(\mathrm{St} + \mathrm{Re_t}^{-1/2})} \right), \tag{12}$$

$$\mathrm{Re_t} = 4.07 \times 10^{10} \alpha \left(\frac{a}{\mathrm{au}} \right)^{-1}, \tag{13}$$

where $v_\mathrm{p,t}$ is the magnitude of the relative turbulent velocity between the particle and the gas. $\mathrm{Re_t}$ is the turbulent Reynolds number, defined as $\mathrm{Re_t} = \alpha c_\mathrm{s} H_\mathrm{g}/(v_\mathrm{th}\lambda)$. The turbulent Reynolds number characterize the interaction of the object with the turbulence and sets an eddie scale.

Fig. 5 shows the dynamical evolution due to additional effects of radial drift and turbulence. The solid lines are essentially the same as in Fig. 3. The dot–dashed lines are with radial drift, and are indistinguishable from the solid lines. The time-scales for the radial drift are much longer than the dynamical time-scales that occurs in the aeolian erosion process. Thus, radial drift is not important for the aeolian erosion in these regimes; its only effect is below the characteristic final size of the object. The addition of turbulent velocities increases the relative velocities involved and strengthens the aeolian erosion. The dot–dashed lines show the evolution of the eroding bodies with turbulent velocities included. The erosion is faster due to the higher velocities involved, and also stops at lower size of the eroding body.

Fig. 6 shows the evolution of a 1 m particle with turbulent velocities for different values of α. Larger values of α increase the erosion rate and also result in lower final size. Weak turbulence levels of $\alpha \lesssim 10^{-3}$ at $a = 1$ au do not change the evolution, and the erosion is dominated by the laminar velocities. Since $v_\mathrm{t} \propto c_\mathrm{s} H_\mathrm{g} \propto a^{3/2}$, at larger separations weak turbulence could be more effective.

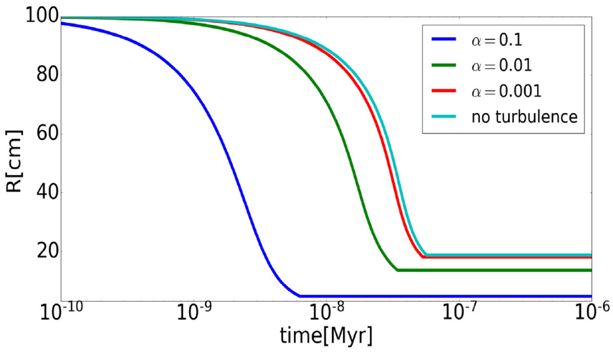

Figure 6. The dependence of aeolian erosion on turbulent velocities, for metre size objects at a fixed a of 1 au distance from the star. We added an artificial lower cut-off at a radius of 1 cm.

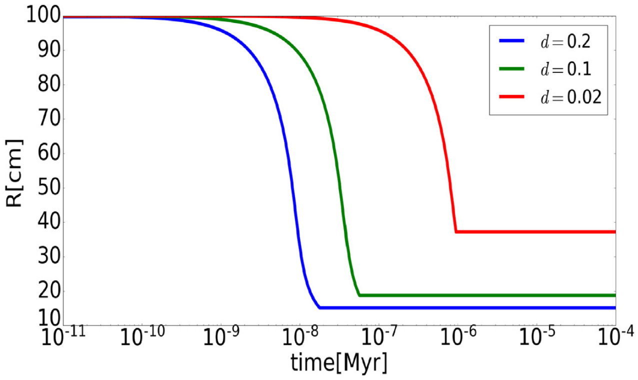

Figure 7. The dependence of the evolution of object of initial metre size in a constant distance from the centre of 1 au on the size of swept grains.

4 DISCUSSION

4.1 Dependence on dust, boulders, and disc properties

The aeolian erosion rate sensitively depends on the size d of the dust grains undergoing suspension. Larger grain size leads to lower threshold velocity, since the cohesive acceleration scales as d^{-2}. These trends are depicted in Fig. 7 where smaller grains indeed lead to larger final size and slower evolution, and vice versa. For small enough size of grains, erosion will not take place due to the strong cohesion acceleration. An important caveat is that for large grains (larger than few centimetres), the cohesion force is not the only force that hold the grains together – Van der Waals is a microscopic force – and henceforth they are out of the scope of our paper.

We used the composition of Rocky material with density $\rho_{\rm p} = 3.45\ {\rm g\ cm^{-3}}$ (Pollack et al. 1996) throughout. Compositions may vary from mostly ice with $\rho_{\rm p} = 1.4\ {\rm g\ cm^{-3}}$ (and lesser densities for porous ice Krijt et al. 2015) up to purely metallic composition with $\rho_{\rm p} = 7.8\ {\rm g\ cm^{-3}}$ (Pinhas, Madhusudhan & Clarke 2016). Since $a_{\rm coh} \propto \rho_{\rm p}$, the time-scale actually does not depend on the density. However, particles with larger density will be less coupled to the gas, therefore denser objects will by eroded to lower size, and vice versa.

We have presented the aeolian erosion barrier under a certain protoplanetary disc model. In reality, there is large spread and uncertainty in the properties of the observed discs Chiang & Laughlin (2013) and Raymond & Cossou (2014). Here we briefly discuss how varying the disc parameters affect the evolution. Changing the gas density at 1 au will have similar effect as the dust grain size for the threshold velocity, with smaller density leads to slower evolution and larger final size. Similar behaviour is expected for varying the gas pressure gradient η, since it affects the relative velocity. Changing the disc temperature changes the sound speed and the scale height, therefore changing the relative velocity and the final size. Lower temperatures will result in smaller final sizes. All of the varied parameters affect the evolution only by a factor of a few and to not change much the overall dynamics.

We have considered the effects of aeolian erosion on circular orbits. Even a small eccentricity, $e > c_{\rm s}/v_{\rm k} \approx 0.022$ could lead to large, supersonic relative velocities, which in turn makes aeolian erosion much more efficient, giving rise to effective erosion of even small bodies, which would otherwise not be susceptible to erosion due to their strong coupling to the gas. For large enough velocities, the pressure can cause significant heating of the outer layers and lead to thermal ablation of the object D'Angelo & Podolak (2015). Although aeolian erosion is just a mechanical processes, it could be important also in cases where just ablation is considered until now. Nevertheless, gas drag rapidly circularize any eccentricity. The erosion/ablation time-scale could be shorter if the typical sizes are small enough. Studying the coupled effects of erosion/ablation and circularization is beyond the scope of this paper.

One possibility for long-term eccentric evolution might the case where some process keeps the bodies eccentric for a long amount of time (e.g. resonances, circumbinary discs and/or external perturbations), the subsequent aeolian erosion in such cases could be much more efficient. The relative velocity could also be altered if binary planetesimals are present (Perets 2011; Grishin & Perets 2016). Finally, similar processes of planetesimal aeolian erosion could be important for planetesimals in scaled-down discs, such as circumplanetary discs (Fujita et al. 2013) or discs around white dwarfs (Grishin & Veras 2019).

4.2 Caveats and comparison to experiments

Paraskov et al. (2006) studied the aeolian erosion of dust aggregates in wind tunnel experiments. They used dust piles, cuboids, and hemispheres. The erosion rate measurement were only possible for cuboids. For 10 cm cuboids at relative velocity of 63 m $\rm s^{-1}$ it is $\dot{m} \sim 10^{-1}\ {\rm g\ h^{-1}}$. The erosion time-scale can thus be estimated as $m/\dot{m} \sim 40\ {\rm h} = 4.5 \times 10^{-3}$ yr. The $\propto R^2$ scaling of the erosion time (or the linear $\propto R$ scaling of the erosion rate) leads to erosion rate estimate of ~ 0.4 yr, which is comparable to our numerical erosion rate of ~ 0.1 yr for 1 m boulder.

Our aeolian erosion model excludes chemically complicated objects such as ice-coated objects. We assume that erosion acts just on outer shells of objects, where we can assume that the chemical interactions are controlled mainly by cohesion forces that behave like loose soil. More complex erosion models are required, but they are beyond the scope of this study.

Given our current data, our results are not completely comparable to the experiments, since there many uncertainties and uncontrolled conditions. The ambient temperature and composition of the gas flow is different. The targets themselves are dust piles, and the erosion rate of spherical piles is not determined, since they start cracking and break apart due to the shear pressure in the experiment. The strength of the cohesive forces in our model is a wide extrapolation of the values obtained for μ-sized dust piles and are uncertain. Spherical configurations may have stronger cohesive forces. On the other hand, the impact velocities of the streamlines that hit the eroding dust grain could be lower depending on geometry, so the empirical rates could be also a lower estimate. Nevertheless, the empirical and our modelled erosion rates are relatively comparable, given the vast uncertainty involved. Future experiments could determine the validity of our model more precisely, in particular the scaling of the erosion rate with the relative velocity, target size, and ambient density.

The erosion is less efficient for smaller dust grains. If the dust grains are on μ-sized, the threshold velocity is too large so that the erosion will be quenched. On the other hand, the larger grains could be eroded easier and faster. In the limit of a rubble pile ~ 100 m body consists of ~ 10 cm cobbles, erosion should be efficient and expend to larger disc separations. However, the forces that bind together cobbles and boulders are probably stronger than purely cohesive forces, and the extrapolation of the linear force dependence on the grain size from measurements of μ-sized grains to 10 cm cobbles is not entirely justified. We therefore caution to draw conclusions on the erosion of a larger body on to ~ 10 cm cobbles. The binding forces of the cobbles and boulders should be studies in more detail in the future.

Large objects that are composed of compactified large rocks cobbles and boulders (tens of centimetres or large) are safe from aeolian erosion in the short time, since the forces that hold them are not only mainly Van der Waals forces. However, with time the small dust particles filling will erode and the overall large body can still erode and the components fragmented away.

Finally, some of the parameter space in Fig. 2 may be inaccessible, since each grain size *d* imposes a *minimal* Stokes number, which varies with the disc location. For a fixed grain size, the Stokes number is proportional to $\mathrm{St} \propto \Sigma_{\mathrm{g}}^{-1} \propto a^{3/2}$, thus for larger separations smaller grains will have larger Stokes numbers. The erosion will stop once the larger body is eroded into its fundamental grains.

4.3 Relation to other growth mechanisms

Aeolian erosion operates even under conditions a priori more favourable to planetesimal growth such as migration traps, if the disc is turbulent. Aeolian erosion significantly affects the evolution of small bodies and their size distribution and therefore has important implications for the evolution of protoplanetary discs and their constituent dust aggregates, cobbles, boulders, and planetesimals.

Another aspect of aeolian erosion in discs is its contribution to growth via pebble accretion. Observations show that mm–cm sized particles are present throughout most of protoplanetary disc's lifetime, including transitional discs with gaps carved by growing protoplanets (Lommen et al. 2009; Banzatti et al. 2011; Jin et al. 2019). Planetesimals and protoplanets must therefore co-exist. The gas-pebble coupling in the presence of a planetary core changes the trajectories of the pebbles and leads to accretion with the core. This pebble-accretion scenario is efficient for optimal size of the pebble reservoir Grishin, Rozner & Perets (2020). In terms of the Stokes number, bodies with $10^{-2} \lesssim \mathrm{St} \lesssim 1$ are efficiently deflected and accreted on to the core (see e.g. fig. 7 of Lambrechts & Johansen 2012). Numerical simulations of particles up to 10 m size show that the accretion rate is stronger for the larger bodies in the presence of very massive cores ($>1\,\mathrm{M}_\oplus$), which alter the trajectories of the gas itself (Morbidelli & Nesvorny 2012). The destruction of 10 m bodies could therefore damage the efficiency of pebble accretion in this case. However, we focus on the first stages of pebble accretion in the presence of smaller cores of sizes below $\lesssim 10^{25}$ g, such that gaseous streamlines will remain intact. Otherwise, different relative velocities should be considered, and not the ones involved in generating Fig. 2 left.

As discussed above, aeolian erosion leads to the erosion of planetesimals into a typical cobble-size range, which are then relatively unaffected by aeolian erosion. Henceforth, aeolian erosion can assist in the growth of planetary embryos at later stages through the provision of small similar-size cobbles, which can be more efficiently accreted on existing embryos through pebble accretion.

A possible extension of our study would be to consider different initial shapes of objects and the effect of aeolian erosion on them – it might structure them into more aerodynamic shapes. Aeolian erosion might then potentially explain the unique elongated shape of the interstellar object such as 'Oumuamua (1I/2017 U1) (Meech et al. 2017). Another direction is to study in details the erosion process of larger objects/larger grains and more complicated shapes of grains and the forces between them, e.g. geometric ways to hold grains together (Goldreich & Sari 2009). Finally, the different size distribution of objects in the disc produced by the erosion process driving them to similar sizes could be important for processes such as streaming instability, which depend on the size distribution (Krapp et al. 2019).

5 SUMMARY

In this paper, we presented an analytic model for aeolian erosion of cobbles, boulders, and planetesimals in protoplanetary discs. The time-scales for erosion are fast and roughly comparable to laboratory experiments (Paraskov et al. 2006). The aeolian erosion is robust and is effective for a wide range of disc structures, dust and planetesimal properties, and turbulence levels. Only small dust grain below 10^{-2} cm are generally safe against aeolian erosion, while larger portions of the disc are susceptible to aeolian erosion for larger grain sizes.

The aeolian erosion is essentially a barrier to planetesimal formation, even at sizes of ~ 100 m. This favours direct gravitational collapse and disfavours coagulation models. On the other hand, the grinding down of larger objects on to dust with typical sizes could be beneficial for planet formation. Small dust grains, cobbles, and boulders of a preferable size are vital for pebble accretion and streaming instability, while other grain sizes prevent growth. The recycled grains can participate in subsequent growth processes, pending on their size and location on the disc.

ACKNOWLEDGEMENTS

We thank to Ruth Murray-Clay, Mickey Rosenthal, Tyler Takaro, Richard Booth, Ynaqin Wu, and Gerhard Wurm for useful discussions and the anonymous referee for useful comments. HBP acknowledges support from the Minerva Center for Life under Extreme Planetary Conditions.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

Adachi I., Hayashi C., Nakazawa K., 1976, Prog. Theor. Phys., 56, 1756
Andrews S. M., 2015, PASP, 127, 961
Andrews S. M., Wilner D. J., Hughes A. M., Qi C., Dullemond C. P., 2009, ApJ, 700, 1502

Armitage P. J., 2010, Astrophysics of Planet Formation, Cambridge University Press, Cambridge, UK
Bagnold R. A., 1941, The Physics of Blown Sand and Desert Dunes. Methuen, London
Banzatti A., Testi L., Isella A., Natta A., Neri R. , Wilner D.J., 2011, A&A, 525, A12
Birnstiel T., Dullemond C. P., Brauer F., 2010, A&A, 513, A79
Blum J., 2018, Space Sci. Rev., 214, 52
Blum J., Wurm G., 2000, Icarus, 143, 138
Brauer F., Henning T., Dullemond C. P., 2008, A&A, 487, L1
Brown P. P., Lawyer D. S., 2000, J. Environ. Eng., 129, 222
Chiang E., Laughlin G., 2013, MNRAS, 431, 3444
Cuzzi J. N., Hogan R. C., 2003, Icarus, 164, 127
D'Angelo G., Podolak M., 2015, ApJ, 806, 203
Demirci T., Schneider N., Steinpilz T., Bogdan T., Teiser J., Wurm G., 2020, MNRAS, 493, 5456
Fujita T., Ohtsuki K., Tanigawa T., Suetsugu R., 2013, AJ, 146, 140
Goldreich P., Sari R., 2009, ApJ, 691, 54
Greeley R., Leach R., White B., Iversen J., Pollack J. B., 1980, Geophys. Res. Lett., 7, 121
Grishin E., Perets H. B., 2015, ApJ, 811, 54
Grishin E., Perets H. B., 2016, ApJ, 820, 106
Grishin E., Veras D., 2019, MNRAS, 489, 168
Grishin E., Perets H. B., Zenati Y., Michaely E., 2017, MNRAS, 466, 276
Grishin E., Perets H. B., Avni Y., 2019, MNRAS, 487, 3324
Grishin E., Rozner M., Perets H. B., 2020, preprint (arXiv:2004.03600)
Güttler C., Blum J., Zsom A., Ormel C. W., Dullemond C. P., 2010, A&A, 513, A56
Heim L.-O., Blum J., Preuss M., Butt H.-J., 1999, Phys. Rev. Lett., 83, 3328
Jin S., Isella A., Huang P., Li S., Li H., Ji J., 2019, ApJ, 881, 108
Johansen A., Youdin A., 2007, ApJ, 662, 627
Klahr H. H., Henning T., 1997, Icarus, 128, 213
Krapp L., Benítez-Llambay P., Gressel O., Pessah M. E., 2019, ApJ, 878, L30
Krijt S., Ormel C. W., Dominik C., Tielens A. G. G. M., 2015, A&A, 574, A83
Kruss M., Musiolik G., Demirci T., Wurm G., Teiser J., 2019, preprint (arXiv:1911.01692)
Lambrechts M., Johansen A., 2012, A&A, 544, A32
Lommen D., Maddison S. T., Wright C. M., van Dishoeck E. F., Wilner D. J., Bourke T. L., 2009, A&A, 495, 869
Meech K. J. et al., 2017, Nature, 552, 378
Morbidelli A., Nesvorny D., 2012, A&A, 546, A18
Ormel C. W., Cuzzi J. N., 2007, A&A, 466, 413
Ormel C. W., Klahr H. H., 2010, A&A, 520, A43
Paraskov G. B., Wurm G., Krauss O., 2006, ApJ, 648, 1219
Perets H. B., 2011, ApJ, 727, L3
Perets H. B., Murray-Clay R. A., 2011, ApJ, 733, 56
Pfalzner S., Bannister M. T., 2019, ApJ, 874, L34
Phillips M., 1980, J. Phys. D: Appl. Phys., 13, 221
Pinhas A., Madhusudhan N., Clarke C., 2016, MNRAS, 463, 4516
Pollack J. B., Hubickyj O., Bodenheimer P., Lissauer J. J., Podolak M., Greenzweig Y., 1996, Icarus, 124, 62
Raymond S. N., Cossou C., 2014, MNRAS, 440, L11
Rosenthal M. M., Murray-Clay R. A., Perets H. B., Wolansky N., 2018, ApJ, 861, 74
Schräpler R., Blum J., 2011, ApJ, 734, 108
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shao Y., 2008 , in Atmospheric and Oceanographic Sciences Library, Vol. 37, Physics and Modelling of Wind Erosion, 2nd edn. Springer Netherlands, Dordrecht
Shao Y., Lu H., 2000, J. Geophys. Res., 105, 22,437
Weidenschilling S. J., 1977, MNRAS, 180, 57
Windmark F., Birnstiel T., Ormel C. W., Dullemond C. P., 2012, A&A, 544, L16
Wurm G., Blum J., 1998, Icarus, 132, 125
Yang C. C., Johansen A., Carrera D., 2017, A&A, 606, A80
Youdin A. N., Goodman J., 2005, ApJ, 620, 459
Zimon A. D., 1982, Adhesion of Dust, Powder, 2nd edn. Springer, New York, NY

APPENDIX A: DISC PARAMETERS

In the following, we consider the dependence of aeolian-erosion efficiency on the properties of the protoplanetary disc. In Fig. A1, we present these dependences. In Table A1, we present the parameters we used through the paper.

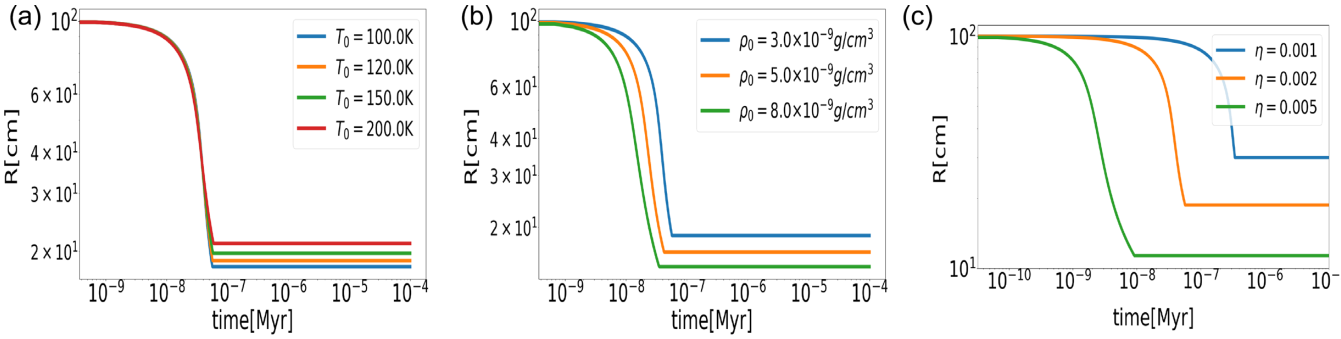

Figure A1. Time evolution of metre-size objects at a fixed distance of 1 au from the star. Part (a) shows the dependence on temperature in units of *K*. Part (b) shows the dependence on the overall central density in units of g cm^{-3}. Part (c) shows the dependence on the gas-pressure support parametres.

Table A1. Supplementary parametres.

Symbol	Definition	Expression	Reference
γ		$0.165\,\mathrm{g\,s^{2}}$	Kruss et al. (2019)
A_N		1.23×10^{-2}	Shao & Lu (2000)
β		$10^{2}\,\mathrm{g\,s^{-1}}$	scaled from Paraskov et al. (2006) and references therein
ρ_{g}	Radial gas density	$3 \times 10^{-9}\left(\frac{a}{\mathrm{au}}\right)^{-16/7}\frac{\mathrm{g}}{\mathrm{cm}^{3}}$	Perets & Murray-Clay (2011)
ρ_{p}	Planetesimals' density	$3.45\frac{\mathrm{g}}{\mathrm{cm}^{3}}$	Pollack et al. (1996)
η	Gas-pressure support parametre	$2 \times 10^{-3}\left(\frac{a}{\mathrm{au}}\right)^{4/7}$	
St	Stokes number	Ωt_{stop}	Perets & Murray-Clay (2011), Armitage (2010)
Re	Reynolds number	$\frac{4Rv_{\mathrm{rel}}}{v_{\mathrm{th}}\lambda}$	Perets & Murray-Clay (2011)
v_{th}	Thermal velocity	$\sqrt{\frac{8}{\pi}}c_{\mathrm{s}}$	Perets & Murray-Clay (2011)
c_{s}	Speed of sound	$\sqrt{\frac{k_B T}{\mu}}$	Perets & Murray-Clay (2011)
μ	Mean molecular weight	$3.93 \times 10^{-24}\,\mathrm{g}$	Rosenthal et al. (2018)
λ	Mean-free path	$\frac{1}{n_{\mathrm{g}}\sigma}$	Perets & Murray-Clay (2011)
n_{g}	Gas number density	$\frac{\rho_{\mathrm{g}}}{\mu}$	Perets & Murray-Clay (2011)
σ	Neutral collision cross-section	$10^{-15}\,\mathrm{cm}$	Perets & Murray-Clay (2011)
T	Temperature	$120\left(\frac{a}{\mathrm{au}}\right)^{-3/7}\,\mathrm{K}$	Perets & Murray-Clay (2011)
α	Shakura–Sunyaev constant	10^{-2}	Rosenthal et al. (2018)

This paper has been typeset from a TEX/LATEX file prepared by the author.

2.2 Aeolian-Erosion in Protoplanetary-Disks II – Implications

Based on Grishin et al. (2020b)

In this study, we discussed the implications of aeolian-erosion in protoplanetary disks on streaming instability and pebble-accretion.

As discussed in section 1.1.1, streaming-instability (Youdin and Goodman, 2005; Johansen et al., 2007) is a suggested mechanism to bypass the meter-size barrier. The main idea behind streaming instability is an initially small overdensity of pebbles that triggers gravitational collapse into planetesimals via positive feedback. Although this direction of solution gains popularity (e.g. Morbidelli et al., 2009; Li et al., 2019; Nesvorný et al., 2019), streaming-instability is efficient just under certain special conditions, for example, high metallicity, local dust-to-ratio above unity and optimal size of pebbles and pressure gradients are required. Another assumption in the 'classical' theory of streaming instability is the mono-size distribution of particles. Recently, this assumption was re-examined by Krapp et al. (2019) and unfortunately, it turned out that streaming-instability is significantly less efficient when the multi-species population is considered – in the extreme case of a wide distribution of pebble sizes, the growth timescale of the streaming-instability unstable mode doesn't converge. As we stated in Rozner et al. (2020b), aeolian-erosion in protoplanetary disks grinds down pebbles in the disk rapidly and efficiently to a final typical size which depends on the distance from the center of the disk – aeolian-erosion induces size-distribution in the disk, on relatively short timescales. Henceforth, aeolian-erosion could assist in setting the right conditions for streaming instability and limit the effect of multi-size distribution.

Planetesimals growth is explained by accreting pebbles (mostly cm-m size, or equivalently Stokes number in the range $0.1 - 1$) – by mechanism which is called pebble-accretion (Ormel and Klahr, 2010; Perets and Murray-Clay, 2011; Lambrechts and Johansen, 2012). The relative velocity of pebbles is reduced as they pass by larger objects, because of the gas drag, and they remain trapped by the large object's gravity. Finally, they spiral onto the surface of the object and accrete onto it. The typical final size produced by aeolian-erosion in disks lies within the size regime of pebbles that accrete efficiently onto planetesimals and henceforth aeolian-erosion enhances and feeds growth by pebble-accretion.

In Grishin et al. (2020b), we derived numerically the critical final size induced by aeolian-erosion (which depends on the distance from the center and the rest of the parameters of the disk). This critical size changes when we take into consideration also turbulence in the disk, by creating a reach reservoir of pebbles in the optimal size for pebble-accretion.

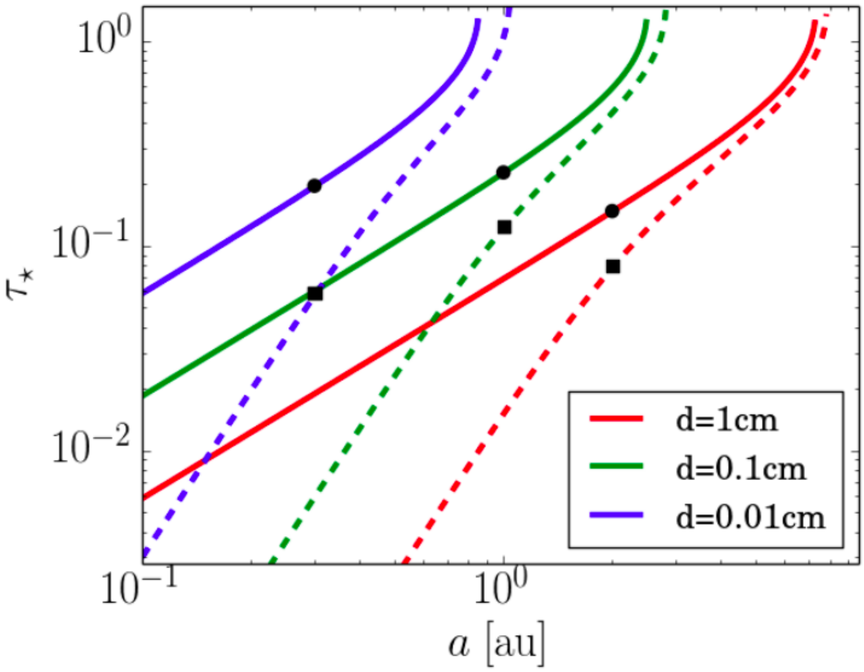

Figure 2.4: Critical stokes number $\tau_\star$ as a function of the orbital separation a. Solid lines are solutions of the laminar velocity only. Dashed lines are the solution with both laminar and turbulent velocity, with the α-viscosity equal to $\alpha = 0.01$. Red (top), green (middle) and blue (bottom) lines correspond to detaching grain sizes of $1, 0.1, 0.01$ cm, respectively. Black circles indicate numerical integration of the erosion equation with laminar velocities for $a = 0.3, 1, 2$AU and $d = 0.01, 0.1, 1$cm , respectively. Black squared indicate the same numerical integration but with both laminar and turbulent velocities. Adopted from Grishin et al. (2020b).

The Astrophysical Journal Letters, 898:L13 (6pp), 2020 July 20 https://doi.org/10.3847/2041-8213/aba266

Erosion-driven Size Redistribution of Protoplanetary Disk Solids and the Onset of Streaming Instability and Pebble Accretion

Evgeni Grishin, Mor Rozner, and Hagai B. Perets
Technion, Israel Institute of Technology, Haifa 3200003, Israel; eugeneg@campus.technion.ac.il

Received 2020 April 7; revised 2020 June 24; accepted 2020 July 2; published 2020 July 20

Abstract

The formation of the first planetesimals and the final growth of planetary cores relies on the abundance of small pebbles. The efficiencies of both the streaming instability (SI) process, suggested to catalyze the early growth of planetesimals, and the pebble-accretion process, suggested to accelerate the growth of planetary cores, depend on the sizes of solids residing in the disk. In particular, these processes were found to be sensitive to size distribution of solids, and efficient planetesimal formation and growth through these channels require a limited pebble size distribution. Here we show that aeolian erosion, a process that efficiently grinds down boulders into a mono-sized distribution of pebbles, provides a natural upper limit for the maximal pebble sizes (in terms of their Stokes number). We find the dependence of this upper limit on the radial separation, disk age, turbulence strength, and the grain-size composition of the boulders in the disk. SI is favorable in areas with a Stokes number less than 0.1, which is found in the inner sub-astronomical-unit regions of the disk. This upper limit shapes the size distribution of small pebbles and thereby catalyzes the early onset of planetesimal formation due to SI, and the later core accretion growth through pebble accretion.

Unified Astronomy Thesaurus concepts: Planet formation (1241); Protoplanetary disks (1300); Planetesimals (1259)

1. Introduction

The early stages of planet formation occur in protoplanetary disks around young stars, which initially contain mostly gas and roughly 1% of dust. Planet formation takes place over many orders of magnitude, beginning with micron-sized dust grains, which collisionally grow to centimeter-sized pebbles and later grow into kilometer-sized planetesimals, and eventually form planetary embryos and planets (Chiang & Youdin 2010).

Although the early growth of dust grains can be understood through collisional processes, the formation of the first planetesimals proves to be a major challenge. Small grains are tightly coupled to the gas flow and can efficiently grow to millimeter–centimeter pebbles. The larger meter-sized boulders are partially decoupled from the gas flow and experience various growth barriers (Blum & Wurm 2008, and references therein). In particular, the radial-drift barrier prevents particles from growing beyond centimeter–meter scales, since such boulders are effectively lost to the main star (Adachi et al. 1976; Weidenschilling 1977), and collisional fragmentation limits rapid growth of ~meter-size boulders (Blum & Wurm 2008). Interstellar planetesimal seeding (Grishin et al. 2019) could provide large enough planetesimals to young systems, thus liberating them from their initial growth barriers. The generation of the first planetesimals, however, is still debated.

Recently, we suggested that aeolian-erosion gives rise to an additional potential growth barrier for pebble/boulder/rock growth, where beyond a certain threshold velocity, the headwind from the gas flow erodes material from the surface of the boulder, as it overcomes the cohesive forces holding its material together (Rozner et al. 2020). The erosion can either grind down larger boulders into smaller pebbles, or set an additional growth barrier for the growing pebbles, even if the other barriers are circumvented.

The streaming instability (SI; Youdin & Goodman 2005) is a potentially promising mechanism to overcome the radial drift (and other barriers) to form planetesimals. SI catalyzes the localized concentration of solids in the disk to the point where gravitational collapse can operate and directly form large planetesimals. Possible observations and simulations that support this channel rely on studies of asteroid size distributions (Morbidelli et al. 2009; Li et al. 2019) and binary Kuiper Belt object binary masses, compositions (Nesvorný et al. 2010), and orientations (Nesvorný et al. 2019). However, the robustness of SI is debated. In particular, SI that leads to the production of strong clumping and successful planetesimal formation requires large metallicity in the protoplanetary disk, a local dust to gas ratio above unity, and an optimal size of the pebbles and pressure gradients (Johansen et al. 2009b; Yang et al. 2017; Sekiya & Onishi 2018).

Early SI studies assumed simple mono-size distribution of solids in the disk. However, recently, Krapp et al. (2019) showed that SI proves to be far less efficient when multisize solid distribution is considered. They find that for a sufficiently wide distribution of pebble sizes, the timescale for the growth of the SI unstable mode is linearly decreasing with the number of species and does not converge (see Figures 2 and 4 of Krapp et al. 2019). Interferometric and scattered light observations of young disks suggest the coexistence of both small μ-sized grains and ~centimeter-sized pebbles (Menu et al. 2014; van Boekel et al. 2017). Thus, the existence of a wide size distribution, typically expected in planet formation models (Bai & Stone 2010; Schaffer et al. 2018) could severely limit the applicability of the SI scenario.

At later stages, the formation of gas/ice giants requires the growth of planetary cores in the standard core accretion scenario (Pollack et al. 1996). The source of the accreted solids was first attributed to planetesimals, but the accretion rate was found to be too slow to efficiently grow planetary cores at large separations. However, it was later suggested that wind-assisted accretion of pebbles could provide a more efficient channel for planetary accretion and growth (Ormel & Klahr 2010; Perets & Murray-Clay 2011; Lambrechts & Johansen 2012). The growth rate and hence the final embryo/planet mass depend on many parameters, including the pebble sizes and abundance, the location in the disk, core formation times (Bitsch et al. 2015; Visser & Ormel 2016; Ormel & Liu 2018), turbulence levels (Rosenthal & Murray-Clay 2018, 2019), and planetary envelope structure and evolution (Lambrechts & Lega 2017; Brouwers et al. 2018).

Both the early planetesimal formation via SI and later subsequent formation of planets due to core accretion rely on the flow of pebbles. Only pebbles of a certain size range, pending the disk model and radial location, can significantly contribute. Thus, the concentration of pebbles of similar sizes in a localized region in the disk could be beneficial for the formation and growth of planetesimals/planets (Liu et al. 2019).

Various mechanisms for concentration of particles have been suggested, including vortices (Barge & Sommeria 1995; Raettig et al. 2015), zonal flows (Johansen et al. 2009a), pressure bumps (Pinilla et al. 2012; Zhu et al. 2012) or planetary torques (Benítez-Llambay & Pessah 2018; Chen & Lin 2018). These mechanisms involve either complex turbulent magnetohydrodynamical effects and/or preexisting planets and have been studied mostly numerically. Here we present a simple, analytic model for the redistribution of disk solid sizes due to a different mechanism, namely aeolian-erosion.

In this Letter we utilize the aeolian-erosion barrier as a natural source of size-segregation and concentration. We focus on the first stages of planet formation assuming no planets or pressure bumps are present. We consider laminar disk flow, and later discuss turbulent disks. In Section 2 we review the aeolian-erosion mechanism and derive the upper limit for the critical Stokes number of surviving solids as a function of the radial location on the disk and the size of the detached grains (i.e., assuming pebbles/boulders are composed of small grains of some typical size, which are removed by the head winds) for laminar and turbulent flows. This, in turn, effectively determines the maximal size of eroded pebbles that survive in the disk. We discuss the implications of the aeolian-erosion pebble size-limit for the SI and pebble accretion processes in Section 3 and summarize in Section 4.

2. Critical Stokes Number from Aeolian Erosion

2.1. Aeolian Erosion

In Rozner et al. (2020), we introduced and discussed the concept of the aeolian-erosion barrier. As small pebbles grow into boulders they are held by cohesive forces. The wind from the gas flow can detach dust grains and pebbles from the surface of the growing boulder. The threshold relative wind velocity at the point when the shear pressure overcomes the cohesion and detaches the particle from the boulder surface is derived from Shao & Lu (2000)

$$v_{\rm th} = \sqrt{\frac{A_N \gamma}{\rho_g d}}, \tag{1}$$

where ρ_g is the local gas density and d is the typical size of the grains composing the pebble. A_N is a dimensionless number that depends on the Reynolds number, and γ is the surface energy. Wind tunnel experiments found a good fit with a constant value of $A_N = 1.23 \times 10^{-2}$ and γ in the range of $(1.65\text{–}5) \times 10^{-1}$ g s^{-2} for grain sizes in the range of 50–1800 μm (Iversen & White 1982). Recent microgravity experiments of silicate glass spheres measured the surface energy in the range of $\gamma = 7.8 \pm 3.8 \times 10^{-2}$ g s^{-2} (Demirci et al. 2020). We choose $\gamma = 1.65 \times 10^{-1}$ g s^{-2} to be compatible with both experiments.

When the relative velocity exceeds the threshold velocity, grains from the outer layer of the pebble/boulder are removed and the mass loss rate is fast. The erosion timescale is (Rozner et al. 2020).

$$t_{\rm ero} = \frac{R}{|\dot{R}|} = \frac{4\pi R^2 \rho_p F_{\rm coh}}{\rho_g v_{\rm rel}^3 m_d}, \tag{2}$$

where R is the size of the body, $v_{\rm rel}$ is the relative velocity, $F_{\rm coh}$ is the strength of the cohesive forces and m_d is the mass of the released grains. The cohesive force scales as $F_{\rm coh} \propto d$, the grain size, with a proportionality constant around 10^2 g s^{-2}, determined from experiments (see Rozner et al. 2020 and Shao & Lu 2000 for details and references.) Generally, the erosion will be very fast, comparable to dynamical timescales for particles less than $\lesssim$10 m (see, e.g., Figure 2 of Rozner et al. 2020), which is comparable to the rapid erosion rates determined in wind tunnel experiments of Paraskov et al. (2006), and more recent microgravity experiments of Demirci et al. (2020). The mass loss continues until the relative wind velocity (which changes due to the continuous decrease in the size of the eroding pebble) becomes smaller than the threshold velocity.

The gas flows in a sub-Keplerian velocity determined by the pressure gradient profile and the location in the disk. The deviation from Keplerian velocity is measured by $\eta \propto (h/a)^2$, where h is the scale height and a is the distance from the star. Using polar coordinates, the components of relative velocity between the object and gas are (we generally follow the same disk model as assumed in Perets & Murray-Clay 2011 and references therein)

$$v_{\rm rel,r} = -\frac{2\eta v_k \tau_s}{1+\tau_s^2},\ v_{\rm rel,\phi} = -\eta v_k\left(\frac{1}{1+\tau_s^2} - 1\right), \tag{3}$$

where the Stokes number is defined by

$$\tau_s = \Omega t_{\rm stop};\ t_{\rm stop} = \frac{m v_{\rm rel}}{F_D}, \tag{4}$$

where Ω is the angular Keplerian velocity, v_k is the Keplerian velocity, m is the object's mass, and $t_{\rm stop}$ is the stopping time. F_D is the aerodynamic drag force.

In Figure 1 we show the aeolian-erosion time evolution of bodies of various initial sizes, but using the Stokes number as a measure. In obtaining Figure 1 we used the flared Chiang–Goldreich disk model (Chiang & Goldreich 1997, see also Perets & Murray-Clay 2011 and Grishin & Perets 2015), with $\eta \approx 2 \times 10^{-3}(a/\mathrm{au})^{4/7}$ and $\rho_g = 3 \times 10^{-9}(a/\mathrm{au})^{-16/7}$ g cm^2. We used $a = 1$ au and $d = 0.1$ cm, similarly to our default assumption in Rozner et al. (2020). The final Stokes number

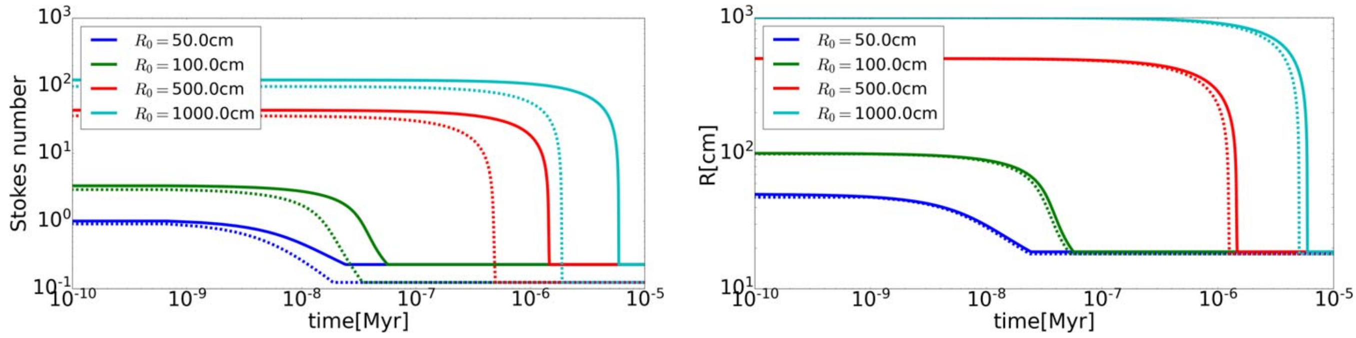

Figure 1. Time-evolution of the Stokes number (left) and of the body size (right) on the initial size of the body at a fixed distance of $a = 1$ au from the star, and dust grains of size $d = 0.1$ cm. Solid lines correspond to integration with the laminar relative velocity. Dashed lines depict integration of both laminar and turbulent velocity.

will lower for the case of turbulent velocities, as explained below.

2.2. Critical Stokes Number

The conclusion from Figure 1 is that the size distribution of particles is limited to a critical Stokes number, $\tau_\star$, which depends on the properties of the composing grains, the sizes of the eroding pebbles and the properties of the disk. For only laminar relative velocities, we present an analytic solution for $\tau_\star$ as a function of the grain and disk properties. For the turbulent case, we arrive at a fifth-order polynomial and find its roots numerically. We discuss the implications for SI later in Section 3.

2.2.1. Laminar Case

The scalar relative velocity from Equation (3) is

$$v_{\rm rel} = \sqrt{v_r^2 + v_\phi^2} = \eta v_K \frac{\tau_s\sqrt{4 + \tau_s^2}}{1 + \tau_s^2} \equiv \eta v_k g(\tau_s). \quad (5)$$

The erosion is quenched once $v_{\rm rel}(\tau_s) \leqslant v_{\rm th}$, which defines the critical Stokes number $\tau_\star$ as a function of the radial location on the disk. By setting the dimensionless laminar relative velocity $\kappa_{\rm l} \equiv \eta v_k / v_{\rm th}$ the condition becomes $g(\tau_\star) = \kappa_{\rm l}^{-1}$. Inverting the equation leads to a second degree polynomial, solved via the standard quadratic formula to yield

$$\tau_\star = \left[\frac{1 - 2\kappa_{\rm l}^2 + \kappa_{\rm l}\sqrt{4\kappa_{\rm l}^2 - 3}}{\kappa_{\rm l}^2 - 1}\right]^{1/2}. \quad (6)$$

The existence of a real solution requires $\kappa_{\rm l} \geqslant \sqrt{3}/2$. Note that the case $a = 1$ au and $d = 0.1$ cm leads to $\tau_\star \approx 0.22$, which is the critical Stokes number in our example in Figure 1.

2.2.2. Turbulent Case

The disk could also be turbulent. The strength of the turbulence is parameterized by the standard Shakura–Sunyaev parameter α. The relative turbulent velocity depends on α and on the dimensionless Stokes and turbulent Reynolds numbers. The turbulent Reynolds number is the ratio of the turbulent to molecular viscosity or the ratio of the largest eddy to the mean free path (Rosenthal et al. 2018). In any case, the turbulent Reynolds number is of the order $\sim\alpha \times 10^{10}$ and much larger than any typical Stokes number.

In the limit of infinite turbulent Reynolds number, the turbulent velocity component is given by $v_{\rm turb} = \sqrt{\alpha}\, c_s \sqrt{\tau_s/(1 + \tau_s)}$, where the sound speed is $c_s \approx 6.6 \times 10^4 (a/{\rm au})^{-3/14}$ cm s^{-1}. Note that the ratio $c_s/v_k \approx 0.022(a/{\rm au})^{2/7}$ is the aspect ratio of the disk, as set from the disk profile. The total relative velocity is the sum of the squares of the laminar and turbulent velocities, $v_{\rm tot}^2 = v_{\rm rel}^2 + v_{\rm turb}^2$.

The erosion stops once $v_{\rm tot} \leqslant v_{\rm th}$. Similarly to the laminar case, we can define the dimensionless turbulent velocity $\kappa_{\rm turb} \equiv \sqrt{\alpha}\, c_s / v_{\rm th}$, and the condition for the critical Stokes number becomes:

$$\kappa_{\rm l}^2 \frac{\tau_\star^2(4 + \tau_\star^2)}{(1 + \tau_\star^2)^2} + \kappa_{\rm turb}^2 \frac{\tau_\star}{1 + \tau_\star} = 1. \quad (7)$$

After some algebra, Equation (7) can be rewritten as a fifth-order polynomial in $\tau_\star$:

$$\begin{aligned} p(\tau_\star) = & (\kappa_{\rm l}^2 + \kappa_{\rm turb}^2 - 1)\tau_\star^5 + (4\kappa_{\rm l}^2 + 2\kappa_{\rm turb}^2 - 2)\tau_\star^3 \\ & + (\kappa_{\rm l}^2 - 1)\tau_\star^4 + (4\kappa_{\rm l}^2 - 2)\tau_\star^2 + (\kappa_{\rm turb}^2 - 1)\tau_\star - 1 \\ = & \ 0. \end{aligned} \quad (8)$$

Unfortunately, there is no explicit expression for the roots of a fifth-degree polynomial,[1] but the roots can be found numerically.

Figure 2 shows the critical Stokes number τ_s as a function of the orbital separation. Solid lines are the solution of Equation (6) with only laminar disk considered, while dashed lines are the solution to the turbulent disk, Equation (8) with $\alpha = 10^{-2}$. Equation (8) has been solved numerically using the `numpy.polynomial` module. The solution is the smallest positive real solution. Each line represents different grain size that determines the threshold velocity in Equation (1). Generally the critical Stokes number is a decreasing function on the radial separation. The larger the size of the pebbles, the farther in disk will erosion take place.

---

[1] Solutions to third- and forth-order polynomials by radicals were known already in the 16th century. The first attempts of a proof of no analytic formula for the fifth degree was presented by Paolo Ruffini (Ruffini 1799). His proof was incomplete and corrected by Niels Henrik Abel (Abel 1824). This is known as the Abel–Ruffini theorem. Later, it was superseded by what is known today as Galois' theory (Galois 1846), which was published postmortem only in 1846, 14 yr after the tragic death of Évariste Galois at 1832.

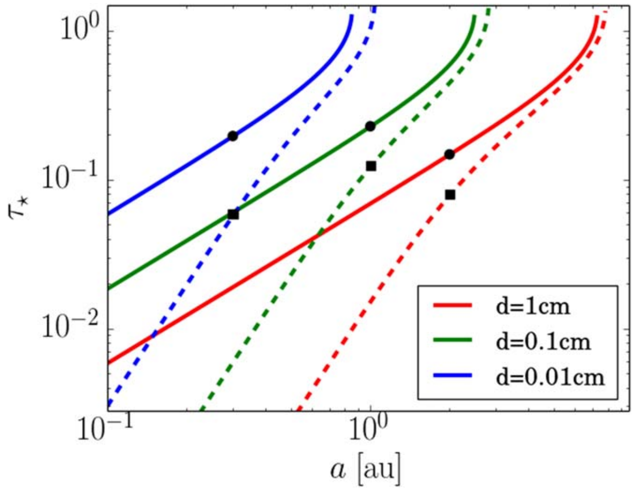

Figure 2. Critical Stokes number $\tau_\star$ as a function of the orbital separation a. Each line is the solution to Equation (8). Solid lines are solutions of the laminar velocity only ($\kappa_t \equiv 0$ in Equation (8)). Dashed lines are the solution with both laminar and turbulent velocity, with the α-viscosity equals to $\alpha = 0.01$. Red (top), green (middle), and blue (bottom) lines correspond to detaching grain sizes of 1, 0.1, 0.01 cm, respectively. Black circles indicate numerical integration of the erosion equation with laminar velocities for $a = 0.3$, 1, 2 au and $d = 0.01$, 0.1, 1 cm, respectively. Black squares indicate the same numerical integration, but with both laminar and turbulent velocities.

2.3. Disk Structure and Evolution

The calculation of the critical Stokes number was done under the assumption of a minimal mass solar nebula (MMSN; Hayashi 1981; Perets & Murray-Clay 2011) background gas density. In reality, disk profiles could vary in shape and slope (Raymond & Cossou 2014), and the gas density may vary due to various global and local effects. Transitional disks of depleted gas density are favorable for the formation of super-Earths (Lee & Chiang 2016), and the formation of ice giants requires the core to form relatively late in order to avoid runaway gas accretion (Bitsch et al. 2015). SI was considered and found to be more efficient following disk evolution in depleted disks, where the metallicity is artificially enhanced (Carrera et al. 2017).

Here we focus on the global disk dissipation and do not discuss local and/or transient effects, which could potentially be important, but are beyond the scope of the current study. We demonstrate the dependence on the results on the different gas densities.

Observations of young clusters show that protoplanetary disks live only a few megayears and could be fitted with exponential time dependence (Mamajek 2009). We assume for simplicity that the gas density follows an exponential decay law, $\rho_g(t) = \rho_g(0)\exp(-t/\tau_{\rm disk})$, where $\tau_{\rm disk} \approx 3$ Myr. Since $v_{\rm th} \propto \rho_g^{-1/2}$, the dimensionless parameter $\kappa_{\rm l}$ will decrease until the erosion stops. For laminar velocity, the critical Stokes number depends on time via $\kappa_{\rm l}$, which will approach $\sqrt{3}/2$ at a finite erosion-stopping time

$$t_{\rm es} = \tau_{\rm disk} \ln \frac{4 d \rho_0 \eta^2 v_K^2}{3 A_N \gamma}. \qquad (9)$$

At this time, the critical Stokes number will increase up to a limiting value of $\tau_\star(\kappa_{\rm l} \to \sqrt{3}/2^+) = \sqrt{2} \approx 1.414$. For our fiducial values of $d = 0.1$ cm at 1 au, the erosion-stopping time is $t_{\rm es} \approx 1.96\tau_{\rm disk} \approx 5.9$ Myr. The result sensitively depends on the location in the disk. At larger radial locations $t_{\rm es}$ is reached faster since $\kappa_{\rm l}$ is smaller there, and vice versa.

For turbulent velocities, both $\kappa_{\rm l}$ and $\kappa_{\rm turb}$ will decrease with decreasing gas density. The critical Stokes number will increase, and generally larger Stokes numbers are possible. The erosion-stopping time is hard to compute analytically, but we expect it to be similar to the time obtained for the laminar case.

To summarize, the critical Stokes number increases with time as the disk is depleted. Therefore, assuming some supply rate $\dot{M}_{\rm supp}(t)$ of larger boulders and planetesimals (e.g., from pebbles drifting from larger separations where erosion was inefficient), the time-dependent erosion will leave traces of eroded material with a time-dependent critical Stokes number $\tau_\star(t)$. The rate of erosion of larger boulders leading to production of grains/pebbles by the aeolian-erosion is $dN(\tau_\star)/dt = \dot{M}_{\rm supp}(t)/m_d(\tau_\star) \propto \tau_\star^{-3}$, where $m_d(\tau_\star)$ is the mass of the grain at Stokes number $\tau_\star$. In principle, the production rate can be integrated to obtain the total number of new grains at a given time, but the integration is not trivial since both $\tau_\star$ and $\dot{M}_{\rm supp}(t)$ could have a complicated dependence on time. The number of new grains should decrease as the Stokes number increases.

3. Discussion and Implications

Size distributions: The initial size distribution of disk solids is usually considered to be following a power law with index q ($n(r) \propto r^{-q}$). Observations of interstellar dust indicate that $q = 3.3$–3.6 (Mathis et al. 1977). Evidence of multiple grain-size populations have also been detected in molecular clouds (Pagani et al. 2010; Andersen et al. 2013) and in protoplanetary disks (Banzatti et al. 2011; Jin et al. 2019). The actual formation channels for boulders beyond the drift and fragmentation barriers are debated. Various mechanisms have been suggested to overcome the growth barriers, such as local pressure maxima, particle pile-ups, rapid coagulation, etc. (see Section 4.3 in Armitage 2010 and references therein). Nevertheless, a large reservoir of $\tau_s \sim 1$ pebbles is the starting point of the pebble accretion paradigm, and the numerical SI study of Krapp et al. (2019) uses a wide range of sizes up to $\tau_s \sim 1$. The interstellar pebble and planetesimal reservoir could have been captured in most stages of the protoplanetary disk lifetime (Grishin et al. 2019), or at an earlier stage during the molecular cloud phase (Pfalzner & Bannister 2019), which would enrich the protoplanetary nebula with an abundance of pebbles and boulders. We remain agnostic to the exact mechanism that forms these boulders and assume that a large reservoir exists, similarly to the standard pebble accretion scenario and other studies that assume an initial size distribution (e.g., Krapp et al. 2019).

Regardless of the theoretical and observational uncertainties, the power law is expected to be steep. At $t = t_0$ the distribution is strictly a power law. As time progresses, dust will grow and the minimal size will increase. In addition, particles with $\tau_s > \tau_\star$ will be eroded to smaller pebbles with $\tau_\star$. If the growth is slow or inefficient, there will be little effect on the underlying distribution, since the total mass is dominated by the lighter dust particles. The only changes in the underlying power-law distribution are the boundaries of the minimal and maximal sizes, shaped by growth and erosion (and other barriers), respectively.

Nevertheless, the shaping of the dust size distribution could have an effect on a local scale. Since each radial separation a determines a typical Stokes number $\tau_\star(a)$, different locations will have different typical dust sizes, which could in turn serve as a ubiquitous mono-dispersed local population. This population can be important for the onset of other growth mechanisms as described below.

Streaming instability: Particles with $\tau_s > \tau_\star$ rapidly erode to $\tau_\star$ on dynamical timescales, much faster that the growth of SI. Thus, $\tau_\star$ is a natural upper limit for the allowed Stokes numbers for the initial multispecies size distribution. The inner parts of the disk will have lower $\tau_\star$. Although this natural upper limit is considered a barrier, it could actually help catalyze planet formation via SI.

Recently, Krapp et al. (2019) have provided the first systematic study of the linear growth of the multispecies SI. They varied the minimal and maximal ranges of the Stokes number, the number of species N and the local dust-to-gas density, ϵ, and studied the timescale for the growth of the most unstable mode in each case. The most striking conclusion is that the convergence was not achieved with increasing number of species. In particular, even for favorable conditions with $\epsilon = 1$, convergence was achieved for $\max(\tau_s) = 0.1$ after $N \sim 100$ species, but for $\max(\tau_s) = 1$ the timescale for the growth of the unstable mode is linearly increasing with the number of species, and does not seem to converge (see Figures 2 and 4 of Krapp et al. 2019).

By truncating the maximal range of τ_s to $\tau_\star$, the SI mechanism can achieve convergence. Convergence is typically achieved for $\max \tau_s \lesssim 0.1$. Thus, the SI is favorable in areas in the disk for which $\tau_\star \lesssim 0.1$, which we find to be the regions inward to $\sim$1 au, pending on dust size, disk model, and turbulence levels. The boundaries of these areas, where $\tau_\star \approx 0.1$, could therefore be the most favorable areas for SI, since this is the optimal Stokes number at which SI is effective with the lowest possible metallicity $Z \approx 0.03$, as shown in Yang et al. (2017).

Pebble accretion: SI is a growth mechanism for the first planetesimals. Once planetary cores of $\gtrsim 10^{2-3}$ km are formed, further growth is proceeded by accretion of pebbles until a critical core mass is reached, where runaway gas accretion begins leading to gas/ice giant formation. The efficiency of pebble accretion depends on their coupling to the gas, i.e., their Stokes number. Pebbles with $\tau_s \lesssim 10^{-3}$ are well coupled to the gas flow and unaffected by the core. Pebbles with $10^{-3} \leqslant \tau_s \leqslant 0.1$ are affected by the core's gravity, but contribute less to the overall collisions and accretion rates (Lambrechts & Johansen 2012). Pebbles with $\tau_s \gtrsim 0.1$ are accreted onto the protoplanet when the impact parameter is within the Hill-sphere. Pebbles with $\tau_s \approx 1$ are attracted from wider distances, but the horseshoe orbits with small impact parameters are lost (see Figure 7 of Lambrechts & Johansen 2012). The overall accretion rates are faster for Stokes numbers in the range of $0.1 \lesssim \tau_s \lesssim 1$, for large enough protoplanetary core of 10^3 km as seen, e.g., in Figure 10 of Ormel & Klahr (2010).

The radial erosion-induced stratification of dust sizes plays a similar role in the efficiency of pebble accretion. Similarly to SI, there are favorable regions in the disk where the critical Stokes number is around $\tau_\star \approx 0.1$–1, where pebble accretion is most probable. Since these are generally regions close to 1 au and inwards, the accreting cores are unlikely to form gas giants. Only for boulders composed of relatively large dust grains of $d \approx 1$ cm, could erosion be effective up to larger distances of $\sim$7 au, which is compatible with the formation locations of ice/giant planets. Evolved disks have lower densities (therefore, less erosive), and even larger grain composition (or closer separation) is required to be effective.

Caveats: In the derivation of Equations (6) and (8) we used dimensionless quantities. In reality, there are limitations to the smallest Stokes number available. The Stokes number is defined as $\tau_s \equiv (\pi/2)\rho_p d/\Sigma_g$. For our disk models, it is roughly $\tau_s \sim 10^{-3}(d/\mathrm{cm})(a/\mathrm{au})^{3/2}$. Thus, for the size of $d = 1$ cm the minimal Stokes is $\sim 10^{-3}$, which increases to $\sim$0.02 at $a \approx 7$ au. Obviously, the erosion cannot grind down boulders to sizes smaller than the fundamental composing-grain size, d; therefore, there is a physical limitation on the minimal Stokes number in our formalism.

Growing boulders and planetesimals can be porous and have various sizes and different densities and cohesive forces. From Equation (1) it is evident that detaching larger grains is easier than smaller ones. Thus, if an eroding object is composed of grains of various sizes, only grains above some threshold can be detached, which will affect the structure of the growing boulders and requires further study.

The erosion timescales are usually shorter than the radial drift times, but the drift itself is much faster than the disk's lifetime. As the particle will drift inward, its critical Stokes number will keep decreasing due to the decrease of the threshold velocity. Obviously, with no drift stopping mechanism, the body will be lost. Nevertheless, even if the body is lost, some of the fractions of the detached grains during the erosion process may survive and serve as reservoirs for the later growth mechanisms.

It is also tempting to apply our formalism for large pebbles of sizes $\sim$10 cm, since they are more favorable to efficient erosion. However, the aeolian-erosion formalism is relying on the assumption that the cohesive forces are linearly proportional to the dust grain size d. The proportionality constant was derived experimentally for small grain sizes of μ-size. We extrapolated the linear behavior up to $\lesssim$0.1 cm pebbles in Rozner et al. (2020), largely based on laboratory experiments of Paraskov et al. (2006) for 0.05 cm size grains, which seem to be consistent with our derived erosion rates. It is unclear if the linear proportionality could be extended beyond 1 cm scale. On the other hand, erosion of smaller grains from the surface may destabilize and weaken the cohesion of larger grains, possibly attached through contact with smaller grains. In this case erosion might be even more efficient. More generally, the nature of the forces that bring together the planetesimals that are composed of pebbles could be different and depend on the composition, porosity, and equation of state, as well as self-gravity for the larger objects. We therefore caution using our model to larger dust/pebble sized and defer it to future studies.

4. Summary

In this Letter we showcased that aeolian-erosion can efficiently grind down solids in protoplanetary disks into smaller grains/pebbles down to the point where they are coupled to the gas flow. The strength of the coupling is measured by the critical Stokes number $\tau_\star$ (Equations (6), (8)), which in turn depends on the ratio of the threshold velocity v_{th} (Equation (1)) and the typical relative laminar and turbulent velocities, and on the size of the detaching grains/pebbles d.

The dependence can be related to the radial location on the disk (Figure 2), and the general trend is that $\tau_\star$ is decreasing with decreasing radial location, until some critical separation where aeolian-erosion becomes inefficient.

Growth of planetesimals due to the streaming instability and the growth of planetary cores due to pebble accretion rely on large numbers of pebbles with "optimal" Stokes numbers with nontrivial coupling with the gas. A wide size-distribution of small particles slows down the growth, since fewer particles participate, and complex coupling between different sizes may play a role and hinder the growth; therefore, simplified assumptions in modeling of these processes through the use of ubiquitous, mono-sized pebbles is heavily criticized. However, as we show here, aeolian-erosion processes naturally produce particle sizes of typical Stokes number, depending on the radial separation. Erosion may therefore allow for a realistic, naturally produced limited pebble-size range. Optimal Stokes numbers are a natural consequence and are expected to then be present at preferred locations. The critical Stokes numbers depend not only on locations but also on time. Evaporating disks with lower gas density increase the critical Stokes number with time. Therefore, depleted disks (at later times or with local cavities) are better sites for planet/planetesimal formation mechanisms that require nontrivial coupling (e.g., $\tau_\star \approx 0.1$–1) of gas and dust, such as the streaming instability or pebble accretion.

We thank Jake Simon, Allona Vazan, Andrew Youdin, and Yanqin Wu for useful discussions. E.G. and M.R. acknowledge useful discussions at the Rocky Worlds workshop at the KICC, which initiated this work. H.B.P. acknowledges support from the MINERVA center for "Life under extreme planetary conditions" and the Kingsley fellowship at Caltech.

ORCID iDs

Evgeni Grishin https://orcid.org/0000-0001-7113-723X
Hagai B. Perets https://orcid.org/0000-0002-5004-199X

References

Abel, N. H. 1824, Œuvres Complètes de Niels Henrik Abel (in French) (2nd ed.; Oslo: Grøndahl & Søn), 28
Adachi, I., Hayashi, C., & Nakazawa, K. 1976, PThPh, 56, 1756
Andersen, M., Steinacker, J., Thi, W. F., et al. 2013, A&A, 559, A60
Armitage, P. J. 2010, Astrophysics of Planet Formation (Cambridge: Cambridge Univ. Press)
Bai, X.-N., & Stone, J. M. 2010, ApJL, 722, L220
Banzatti, A., Testi, L., Isella, A., et al. 2011, A&A, 525, A12
Barge, P., & Sommeria, J. 1995, A&A, 295, L1
Benítez-Llambay, P., & Pessah, M. E. 2018, ApJL, 855, L28
Bitsch, B., Lambrechts, M., & Johansen, A. 2015, A&A, 582, A112
Blum, J., & Wurm, G. 2008, ARA&A, 46, 21
Brouwers, M. G., Vazan, A., & Ormel, C. W. 2018, A&A, 611, A65
Carrera, D., Gorti, U., Johansen, A., & Davies, M. B. 2017, ApJ, 839, 16
Chen, J.-W., & Lin, M.-K. 2018, MNRAS, 478, 2737
Chiang, E., & Youdin, A. N. 2010, AREPS, 38, 493
Chiang, E. I., & Goldreich, P. 1997, ApJ, 490, 368
Demirci, T., Schneider, N., Steinpilz, T., et al. 2020, MNRAS, 493, 5456
Galois, E. 1846, J. Math. Pures Appl., 11, 417
Grishin, E., & Perets, H. B. 2015, ApJ, 811, 54
Grishin, E., Perets, H. B., & Avni, Y. 2019, MNRAS, 487, 3324
Hayashi, C. 1981, PThPS, 70, 35
Iversen, J. D., & White, B. R. 1982, Sedim, 29, 111
Jin, S., Isella, A., Huang, P., et al. 2019, ApJ, 881, 108
Johansen, A., Youdin, A., & Klahr, H. 2009a, ApJ, 697, 1269
Johansen, A., Youdin, A., & Mac Low, M.-M. 2009b, ApJL, 704, L75
Krapp, L., Benítez-Llambay, P., Gressel, O., & Pessah, M. E. 2019, ApJL, 878, L30
Lambrechts, M., & Johansen, A. 2012, A&A, 544, A32
Lambrechts, M., & Lega, E. 2017, A&A, 606, A146
Lee, E. J., & Chiang, E. 2016, ApJ, 817, 90
Li, R., Youdin, A. N., & Simon, J. B. 2019, ApJ, 885, 69
Liu, B., Ormel, C. W., & Johansen, A. 2019, A&A, 624, A114
Mamajek, E. E. 2009, in AIP Conf. Ser. 1158, Exoplanets and Disks: Their Formation and Diversity, ed. T. Usuda, M. Tamura, & M. Ishii (Melville, NY: AIP), 3
Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, ApJ, 217, 425
Menu, J., van Boekel, R., Henning, T., et al. 2014, A&A, 564, A93
Morbidelli, A., Bottke, W. F., Nesvorný, D., & Levison, H. F. 2009, Icar, 204, 558
Nesvorný, D., Li, R., Youdin, A. N., Simon, J. B., & Grundy, W. M. 2019, NatAs, 3, 808
Nesvorný, D., Youdin, A. N., & Richardson, D. C. 2010, AJ, 140, 785
Ormel, C. W., & Klahr, H. H. 2010, A&A, 520, A43
Ormel, C. W., & Liu, B. 2018, A&A, 615, A178
Pagani, L., Steinacker, J., Bacmann, A., Stutz, A., & Henning, T. 2010, Sci, 329, 1622
Paraskov, G. B., Wurm, G., & Krauss, O. 2006, ApJ, 648, 1219
Perets, H. B., & Murray-Clay, R. A. 2011, ApJ, 733, 56
Pfalzner, S., & Bannister, M. T. 2019, ApJL, 874, L34
Pinilla, P., Birnstiel, T., Ricci, L., et al. 2012, A&A, 538, A114
Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, Icar, 124, 62
Raettig, N., Klahr, H., & Lyra, W. 2015, ApJ, 804, 35
Raymond, S. N., & Cossou, C. 2014, MNRAS, 440, L11
Rosenthal, M. M., & Murray-Clay, R. A. 2018, ApJ, 864, 66
Rosenthal, M. M., & Murray-Clay, R. A. 2019, arXiv:1908.06991
Rosenthal, M. M., Murray-Clay, R. A., Perets, H. B., & Wolansky, N. 2018, ApJ, 861, 74
Rozner, M., Grishin, E., & Perets, H. B. 2020, MNRAS, 496, 4827
Ruffini, P. 1799, Teoria Generale delle Equazioni, Parte Prima (in Italian) (Bologna: Nella Stamperia di S. Tommaso D'Aquino)
Schaffer, N., Yang, C.-C., & Johansen, A. 2018, A&A, 618, A75
Sekiya, M., & Onishi, I. K. 2018, ApJ, 860, 140
Shao, Y., & Lu, H. 2000, JGR, 105, 437
van Boekel, R., Henning, T., Menu, J., et al. 2017, ApJ, 837, 132
Visser, R. G., & Ormel, C. W. 2016, A&A, 586, A66
Weidenschilling, S. J. 1977, MNRAS, 180, 57
Yang, C. C., Johansen, A., & Carrera, D. 2017, A&A, 606, A80
Youdin, A. N., & Goodman, J. 2005, ApJ, 620, 459
Zhu, Z., Nelson, R. P., Dong, R., Espaillat, C., & Hartmann, L. 2012, ApJ, 755, 6

2.3 Aeolian erosion III - white dwarf disks

Based on Rozner et al. (2021)

The mechanism of aeolian-erosion in protoplanetary disks, which we presented in Rozner et al. (2020b); Grishin et al. (2020b) might take place also in white-dwarf disks, with some necessary modifications. Here we studied the effect of aeolian erosion on white-dwarf disks and its implications.

White dwarfs are the final evolutionary stage of the vast majority of all-stars. A few percent of the white dwarfs are surrounded by planetary debris disks (Manser et al., 2020), and over fifty white-dwarf disks have been already discovered (Zuckerman et al., 2010; Farihi, 2016). Most of the observed disks are dusty, although over one dozen disks with gaseous components were discovered as well (Dennihy et al., 2020; Melis et al., 2020). The abundance of objects with different sizes in white-dwarf disks – from grain size to minor planets, reveals rich and interesting dynamics (see a detailed review of post-main sequence evolution in Veras, 2016), which includes many physical processes, among them are replenishment and accretion. Frequent collisions in white-dwarf disks (Jura, 2008; Metzger et al., 2012; Kenyon and Bromley, 2017a,b) lead to replenishment of grains in the disk – large objects break into smaller ones, and gradually are removed by radiation pressure. The debris disk might give rise to dynamical excitation and perturbations of mass that will eventually drive matter onto the white dwarf within the disk lifetime (Girven et al., 2012; Veras and Heng, 2020). Poynting-Robertson drag exploits radiation force and causes loss of angular-momentum for small pebbles, by that – carries them to the white-dwarf (Burns et al., 1979; Rafikov, 2011b,a). Polluted white dwarfs, which are $25 - 50\%$ of the whole white-dwarf population (Zuckerman et al., 2010; Koester et al., 2014), constitute an observational signature to the accretion of heavy elements.

The dynamics and architecture of white-dwarf disks are similar to protoplanetary disks in some aspects, and different in others. (see Rafikov, 2011a; Metzger et al., 2012;

Veras, 2016 for detailed reviews of white-dwarf disks' known range of parameters). The scales discussed in white-dwarf disks are much smaller, and here we can discuss objects in a distance of $\lesssim 1R_\odot$. Moreover, the density profile, the mass of the disk, the temperature profile changes – which might enable thermal ablation as well and the lifetime of the disk is shorter and ranges between $10^4 - 10^6$ year (Veras and Heng, 2020).

Aeolian-erosion in white-dwarf disks might add another piece to the evolution puzzle in white-dwarf disks and explain the accretion rates onto white-dwarfs. The scenario we describe is the following: objects in white-dwarf disks are eroded down to a small size, where Poynting-Robertson drag is efficient, and is able to carry the particles to the white-dwarf. Objects in eccentric disks might be eroded in shorter timescales since the relative velocity in the pericenter is larger. Similarly to protoplanetary disks, aeolian-erosion might reshape the size distribution of objects in the disk. In white-dwarf disks, these constraints will be important, since the lack of experimental/theoretical knowledge on the size distribution. In this project, we will use the analytical and numerical tools we developed in Rozner et al. (2020b); Grishin et al. (2020b) in order to study the role of aeolian-erosion in white dwarf disks.

From our preliminary results, large objects ($\sim 10^4$ year) in white-dwarf disks eroded down to a typical size of ~ 0.1cm in short timescales of $\sim$ year; km-size objects are expected to be eroded in typical timescales of $\sim 10^2$ years to the same final size.

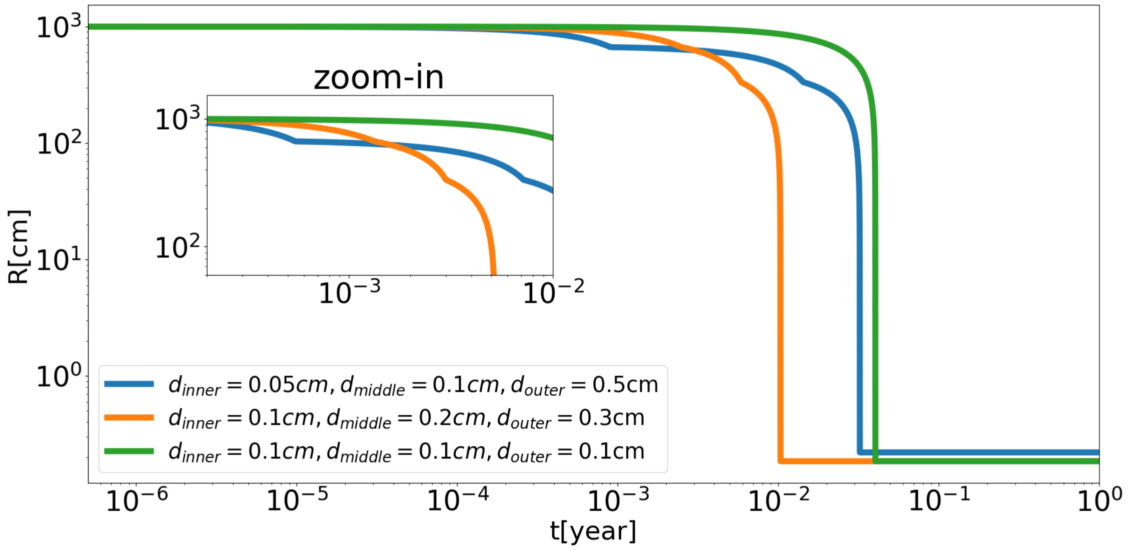

Figure 2.5: The evolution of triple-layered objects, with each layer comprising 1/3 of the radius, of the same density 3.45g cm^{-3}, but with different size of inner grains that build each layer. The object is embedded at a white dwarf disk at a distance of 0.005AU from the white dwarf.

In Fig. 2.5 we present an example of the time evolution of a differentiated object embedded in white-dwarf disks under the effect of aeolian-erosion. First and foremost, the aeolian-erosion timescales are extremely short, which points out that as far as the conditions for initiation of aeolian-erosion are implied, this mechanism is robust and might affect significantly the population of grains in white-dwarf disks. Erosion enables us to decompose the layers of an object (as far as they are in the correct regimes in which erosion is effective) and to reveal the inner layers in short timescales. Since different sizes of grains impose different rates of aeolian-erosion, the total aeolian-erosion timescales change such that considering larger grains shortens the timescales, and smaller grains lengthen them. Moreover, it can be seen that the final size is determined by the innermost layer, and larger grains lead to smaller final sizes.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **00,** 1 (2021)
Advance Access publication 2021 February 9

doi:10.1093/mnras/stab329

Rapid destruction of planetary debris around white dwarfs through aeolian erosion

Mor Rozner,[1★] Dimitri Veras[2,3†] and Hagai B. Perets[1]
[1] *Physics Department, Technion - Israel Institute of Technology, Haifa 320004, Israel*
[2] *Centre for Exoplanets and Habitability, University of Warwick, Coventry CV4 7AL, UK*
[3] *Department of Physics, University of Warwick, Coventry CV4 7AL, UK*

Accepted 2021 February 2. Received 2021 January 14; in original form 2020 November 24

ABSTRACT

The discovery of numerous debris discs around white dwarfs (WDs) gave rise to extensive study of such discs and their role in polluting WDs, but the formation and evolution of these discs is not yet well understood. Here, we study the role of aeolian (wind) erosion in the evolution of solids in WD debris discs. Aeolian erosion is a destructive process that plays a key role in shaping the properties and size distribution of planetesimals, boulders, and pebbles in gaseous protoplanetary discs. Our analysis of aeolian erosion in WD debris discs shows that it can also play an important role in these environments. We study the effects of aeolian erosion under different conditions of the disc and its erosive effect on planetesimals and boulders of different sizes. We find that solid bodies smaller than ~ 5 km will be eroded within the short disc lifetime. We compare the role of aeolian erosion in respect to other destructive processes such as collisional fragmentation and thermal ablation. We find that aeolian erosion is the dominant destructive process for objects with radius $\lesssim 10^3$ cm and at distances $\lesssim 0.6\,\mathrm{R}_\odot$ from the WD. Thereby, aeolian erosion constitutes the main destructive pathway linking fragmentational collisions operating on large objects with sublimation of the smallest objects and Poynting–Robertson drag, which leads to the accretion of the smallest particles on to the photosphere of WDs, and the production of polluted WDs.

Key words: minor planets, asteroids: general – planets and satellites: dynamical evolution and stability – planets and satellites: formation – planets and satellites: physical evolution – white dwarfs.

1 INTRODUCTION

White dwarfs (WDs) are the final evolutionary stage of the vast majority of all stars. Metal-polluted WDs, which represent 25–50 per cent of the whole WD population (Zuckerman et al. 2010; Koester, Gänsicke & Farihi 2014), constitute an observational signature of the accretion of heavy elements. A few per cent of the WDs are surrounded by planetary debris discs (Manser et al. 2020), and over 60 WD discs have already been discovered (Zuckerman et al. 2010; Farihi 2016). Most of the observed discs are dusty, although about 20 discs with gaseous components were discovered as well (Dennihy et al. 2020; Gentile Fusillo et al. 2020; Melis et al. 2020)

The quickly growing population of these discs motivates theoretical explorations of their structure and evolution, which can reveal both how they are formed (Veras et al. 2014, 2015; Malamud & Perets 2020a,b) and how they accrete on to WDs photospheres, creating observable metal pollution (Jura & Young 2014; Harrison, Bonsor & Madhusudhan 2018; Hollands, Gänsicke & Koester 2018; Doyle et al. 2019; Swan et al. 2019; Bonsor et al. 2020). An interesting observed feature of WD discs is that nearly all the discs are now thought to showcase flux variability on time-scales of weeks to decades (Farihi et al. 2018; Xu et al. 2018; Swan et al. 2020).

The abundance of objects with different sizes in WD discs – from grain size to minor planets, reveals rich and interesting dynamics (see a detailed review of post-main-sequence evolution in Veras 2016), which includes many physical processes, among them are replenishment and accretion. Frequent collisions in WD discs (Jura 2008; Metzger, Rafikov & Bochkarev 2012; Kenyon & Bromley 2016, 2017a,c) lead to replenishment of grains in the disc – where large objects break into smaller ones, and gradually are removed by radiation pressure and accretion. The debris discs might give rise to dynamical excitations and perturbations of mass that will eventually drive matter on to the WD within the disc lifetime (Girven et al. 2012; Veras & Heng 2020). Poynting–Robertson drag results from the radiation force and causes loss of angular momentum for small pebbles, carrying them to the WD (Burns, Lamy & Soter 1979; Rafikov 2011a,b).

The dynamics and architecture of WD discs are similar to protoplanetary discs in some aspects, and different in others. While there are some processes that take place in both of them, such as fragmentation, their parameters might differ significantly. The scales of WD discs are much smaller, the density and temperature profiles are different, and the typical velocities could be much higher.

One important process in protoplanetary discs is aeolian (wind) erosion. Aeolian erosion is a purely mechanical destructive process, which is very common in many occasions in nature, mainly discussed in the context of sand dunes (Bagnold 1941). Recently, we showed that aeolian erosion can play an important role in planet

---

★ E-mail: morozner@campus.technion.ac.il
† STFC Ernest Rutherford Fellow.

Published by Oxford University Press on behalf of Royal Astronomical Society

formation by setting a new growth-barrier for pebbles/boulders in protoplanetary discs, and affecting pebble accretion and streaming instability (Grishin, Rozner & Perets 2020; Rozner, Grishin & Perets 2020). Aeolian erosion in protoplanetary discs is rapid and efficient, as also verified in lab experiments and numerical simulations of the conditions of protoplanetary discs (Paraskov, Wurm & Krauss 2006; Demirci & Wurm 2020; Demirci et al. 2020a,b; Schaffer et al. 2020). Aeolian erosion leaves signatures on the dynamics of objects in the disc, e.g. it fuels pebble-accretion and induces a redistribution of sizes in the disc, and might potentially lead to reshaping or complete destruction of objects (Grishin et al. 2020; Rozner et al. 2020). In contrast with protoplanetary discs, the temperatures in WD discs are high enough to maintain thermal ablation (Podolak, Pollack & Reynolds 1988; Pollack et al. 1996; D'Angelo & Podolak 2015) along with aeolian erosion, and accelerate the destruction of small objects.

In this paper, we suggest that aeolian erosion could take place in WD discs and interact symbiotically with other dynamical processes in the disc. In fact, we show that for certain regimes, for objects with radius $\lesssim 10^3$ cm and at distances $\lesssim 0.6\,{\rm R}_\odot$ from the WD, in a disc with $M_{\rm disc} = 10^{24}$ g for aeolian erosion is the dominant destruction process. The dominance regime changes with the disc and object's parameters. As we will discuss later, the mass of the disc should be $\gtrsim 10^{21}$ g to enable aeolian erosion.

The efficiency of aeolian erosion, as well as of thermal ablation, depends on the composition of the object and the disc temperature, and acts to grind down large objects into small pebbles. For small pebbles, Poynting–Robertson drag becomes efficient and carries the pebbles on to the WD. Hence, aeolian erosion might set a lower limit on the accretion flux on to the WD and enhance the disc gas replenishment, by assuring repeatedly regenerating the large abundance of small objects. Moreover, as we showed in Grishin et al. (2020), aeolian erosion induces redistribution of the particle sizes and their abundance in the disc. The size distribution of objects in WD discs as well as the sizes of objects which accrete on WDs are not well constrained observationally, although some theoretical constraints have been established (Kenyon & Bromley 2017b, c). Aeolian erosion might shed light on in this direction as well.

The paper is organized as follows: in Section 2 we briefly review the parameters space and various models of WD discs. In Section 3 we review the models of aeolian erosion and thermal ablation in WDs discs. In Section 4 we present our results and include eccentric orbits, multilayer objects and the relationship with thermal effects. In Section 5 we discuss our results and suggest possible implications: we discuss the dependence of our model on the disc parameters, the symbiotic relations with other processes in the disc including collisional cascade, external seeding and further disc generations. In Section 6, we discuss the caveats and limitations of our study. In Section 7 we summarize the paper and suggest future directions.

2 WHITE DWARF DISCS

In contrast with protoplanetary discs which properties are better-constrained, WD disc parameters are uncertain by orders of magnitude. Here we briefly review the ranges of these parameters that will be used in the rest of the paper.

The total mass of a WD disc ranges between 10^{12} and 10^{25} g (see a detailed discussion in Metzger et al. 2012; Veras & Heng 2020 and references therein). The mass of the gas in the disc, parametrized by the gas-to-dust ratio, is highly unconstrained and ranges between 10^{-5} and unity (Veras 2016). Observations set the lower limit of the inner radius to be $\lesssim 0.2\,{\rm R}_\odot$ (e.g. Rafikov 2011a) and the outer radius to be $\gtrsim 1.2\,{\rm R}_\odot$ (e.g. Gänsicke et al. 2006).

The surface density profile is given as (Metzger et al. 2012)

$$\Sigma_{\rm g}(a) = \Sigma_{\rm g,0}\left(\frac{4.72a}{0.6R_\odot}\right)^{-\beta}\left(\frac{M_{\rm disc}}{10^{24}\,{\rm g}}\right), \qquad (1)$$

where a is the distance from the centre of the disc, $\Sigma_{\rm g,\,0}$ is the fiducial surface density and β is an arbitrary exponent which is parametrized as $\beta = n + 1/2$, where n describes the viscosity power law $\nu(a) \propto a^n$. The gas density $\rho_{\rm g}$ is determined by the surface density,

$$\rho_{\rm g} = \frac{\Sigma_{\rm g}}{2h(a)}, \qquad (2)$$

where h is the height of the disc.

This scaling of β also induces a temperature scaling (Metzger et al. 2012),

$$T(a) \propto \begin{cases} \text{constant}, & n = 3/2; \\ a^{-1/2}, & n = 1 \end{cases} \qquad (3)$$

where $n = 3/2$ corresponds to an optically thick disc, and $n = 1$ corresponds to an optically thin disc. The aspect ratio of the disc is loosely constrained and strays over some orders of magnitude. We will use an aspect ratio of 10^{-2} unless stated otherwise. The aforementioned choice of parameters yields typical values of $\Sigma_{\rm g} = 5.1 \times 10^3\,{\rm g\,cm}^3$, $\rho_{\rm g} = 2.55 \times 10^{-5}\,{\rm g\,cm}^{-3}$, and $T = 10^3$ K at $a = 0.6\,{\rm R}_\odot$.

Unless stated otherwise, we will use $n = 3/2$ for all our profiles.

The lifetime of WD discs (at least the ones within a couple of Solar radii from the WD) ranges between $\sim 10^4$ and 10^6 yr (Girven et al. 2012; Veras & Heng 2020). Girven et al. (2012) estimated the age of WD discs assuming a constant accretion rate to the WD. Recently, Veras & Heng (2020) introduced a different estimation method, arising from the dynamical processes that the disc should go through and their typical time-scales.

3 AEOLIAN EROSION AND THERMAL DESTRUCTIVE PROCESSES IN DISCS

3.1 Aeolian erosion

Consider a spherical object with radius R that resides at a constant distance a from the WD and is built from grains with a typical size d, which move in a gaseous medium with a density $\rho_{\rm g}$. The pressure support present in gaseous discs leads to a difference between the Keplerian velocity $\Omega_{\rm K}$ around the WD and the actual angular velocity $\Omega_{\rm g}$, $\Omega_{\rm g} - \Omega_{\rm k} \approx (2\Omega_{\rm K} a \rho_{\rm g})^{-1}\partial P/\partial r$ where $\partial P/\partial r$ is the pressure gradient. Furthermore, the pressure support leads to the radial drift, which presents one of the fundamental problems in planet formation in protoplanetary discs – the metre-size barrier (Weidenschilling 1977).

The radial velocity in protoplanetary discs obtains a maximum for ~ 1 m size objects, and these objects inspiral to the inner parts of the disc on time-scales shorter than the expected growth time-scale. The drift velocity is lower for smaller objects – they are better coupled to the gas and hence have slower relative velocities; larger objects are loosely coupled to the gas and experience slower drift velocities. For objects in WD discs, the maximal drift velocity is obtained for smaller objects, in size $\lesssim 1$ cm (Kenyon & Bromley 2017c):

$$v_{\rm rel,r} = -\frac{2\eta v_{\rm k} St}{1 + St^2}, \qquad (4)$$

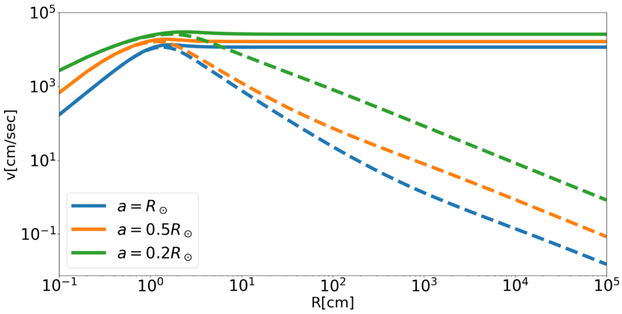

Figure 1. The relative velocity between objects and the gas, in different constant distances from the WD. The solid lines are for $v_{\rm rel}$ and the dashed ones are for the radial component, $v_{\rm r}$.

$$v_{{\rm rel},\phi} = -\eta v_{\rm k} \left(\frac{1}{1 + St^2} - 1 \right) \quad (5)$$

where the Stokes number is defined by

$$St = \Omega_{\rm K} t_{\rm stop}, \quad (6)$$

$$t_{\rm stop} = \frac{m v_{\rm rel}}{F_{\rm D}} \quad (7)$$

where m is the mass of the object, $F_{\rm D}$ is the drag force, which is given by

$$\mathbf{F}_{\rm D} = \frac{1}{2} C_{\rm D}(Re) \pi R^2 \rho_{\rm g} v_{\rm rel}^2 \hat{\mathbf{v}}_{\rm rel}. \quad (8)$$

For the drag coefficient, we adopted an empirical fitted formula, based on experimental data in the regime $10^{-3} \leq Re \leq 10^5$ (Perets & Murray-Clay 2011 and references therein)

$$C_D(Re) = \frac{24}{Re} (1 + 0.27 Re)^{0.43} + 0.47 \left[1 - \exp\left(-0.04 Re^{0.38} \right) \right] \quad (9)$$

where Re is the Reynolds number, defined by

$$Re = \frac{4 R v_{\rm rel}}{v_{\rm th} \lambda}, \quad (10)$$

where $v_{\rm th} = (8/\pi)^{1/2} c_{\rm s}$ is the mean thermal velocity, $c_{\rm s} = \sqrt{k_{\rm B} T / \mu}$ is the speed of sound, k_B is Boltzmann constant, T is the temperature of the disc, μ is the mean molecular weight, taken to be $2.3 m_{\rm H}$ (following Perets & Murray-Clay 2011) with $m_{\rm H}$ the mass of a hydrogen atom and $\lambda = \mu / (\rho_{\rm g} \sigma)$ is the mean free path of the gas.

In Fig. 1, we present the relative velocities between the gas and objects for different sizes. The radial component of the velocity peaks for $\sim$1cm size objects, and is about $\sim 10^5$ cm s^{-1}. Much smaller objects are well-coupled to the gas, which leads to smaller velocities, while much larger objects are weakly coupled to the gas and hence are not affected significantly by it. The relative velocity could change for different parameters of the disc, and as we have mentioned, the parameter space is currently wide.

Objects in gaseous discs experience gas drag, which depends strongly on their velocity relative to the gas, $v_{\rm rel}$. The gas-drag can play the role that is usually played by the wind in aeolian erosion and trigger loss of the outer layer of objects, as long as they reach the threshold conditions.

The threshold velocity is dictated by the balance between cohesion forces, self-gravity and gas-drag, i.e. the headwind should be stronger than the attraction force between grains in order to initiate aeolian erosion.

The threshold velocity is given by

$$v_\star = \sqrt{\frac{A_{\rm N}}{\rho_{\rm g}} \left(\rho_{\rm p} g d + \frac{\gamma}{d} \right)}, \quad (11)$$

where $A_{\rm N} = 1.23 \times 10^{-2}$, and $\gamma = 0.165$ g s^{-2} are determined empirically from Shao & Lu (2000). Both of them are intrinsic characteristic of the materials that rise from the cohesion forces that hold the particles together – mostly electrostatic forces and van der Waals forces. The gravitational acceleration is $g = Gm/R^2$. Above the threshold velocity, i.e. $v_{\rm rel} > v_\star$, the aeolian erosion rate is given by (Rozner et al. 2020),

$$\frac{{\rm d}R}{{\rm d}t} = -\frac{\rho_{\rm g} v_{\rm rel}^3}{4\pi R \rho_{\rm p} a_{\rm coh}} \quad (12)$$

where $a_{\rm coh}$ is the cohesion acceleration. The derivation rises from the work done on the eroded object by the shear pressure; see a detailed derivation from equations (8) and (9) of Rozner et al. (2020). The derivation is based on estimation of the typical sweeping rate of grains from the outer layer and calculating the work done by them.

3.2 Thermal destructive processes

Along with aeolian erosion, the high temperatures that are usually found in WD discs might give rise to destructive thermal processes such as thermal ablation (Podolak et al. 1988; Pollack et al. 1996; D'Angelo & Podolak 2015) and sublimation (e.g. Metzger et al. 2012; Shestakova, Demchenko & Serebryanskiy 2019). The heat that the outer layer absorbs might lead to phase transitions and then mass loss. There are two regimes, separated by a critical temperature $T_{\rm cr}$, in which the latent heat required for vaporization is zero, and varies according to the material (e.g. Opik 1958; Podolak et al. 1988; D'Angelo & Podolak 2015). See the fiducial parameters for ablation in Table A1.

Below the critical temperature, the rate in which vaporization removes mass is dictated by Hertz–Knudsen–Langmuir equation; see a discussion in D'Angelo & Podolak (2015). In this regime, aeolian erosion significantly dominates for our choice of parameters, and the time-scale for thermal ablation below the critical temperature is $\gtrsim 95$ yr – much longer than typical aeolian erosion time-scale, which enables us to neglect the thermal effect and focus on the mechanical processes.

Above the critical temperature, the contribution from thermal processes might add a significant contribution to aeolian erosion and should be added to equation 12; assuming blackbody emission, the thermal ablation term is given by

$$\left. \frac{{\rm d}R}{{\rm d}t} \right|_{\rm ablation} = \frac{1}{L_{\rm s} \rho_{\rm p}} \epsilon_{\rm s} \sigma_{\rm SB} \left(T_{\rm cr}^4 - T_{\rm g}^4 \right), \quad (13)$$

where $\epsilon_{\rm s}$ is the thermal emissivity of the object ($\epsilon_{\rm s} = 1$ for a perfect blackbody), $L_{\rm s}$ is the particle specific vaporization energy, $T_{\rm cr}$ is the critical temperature – which depends on composition of the material – and $T_{\rm g}$ is the gas temperature. At high enough temperatures, close to the WD, small grains sublimate to gas (e.g. Metzger et al. 2012; Shestakova et al. 2019).

4 RESULTS

In this section we present the evolution of objects in WD discs due to aeolian erosion and thermal ablation, starting from a fiducial set of parameters and then vary them.

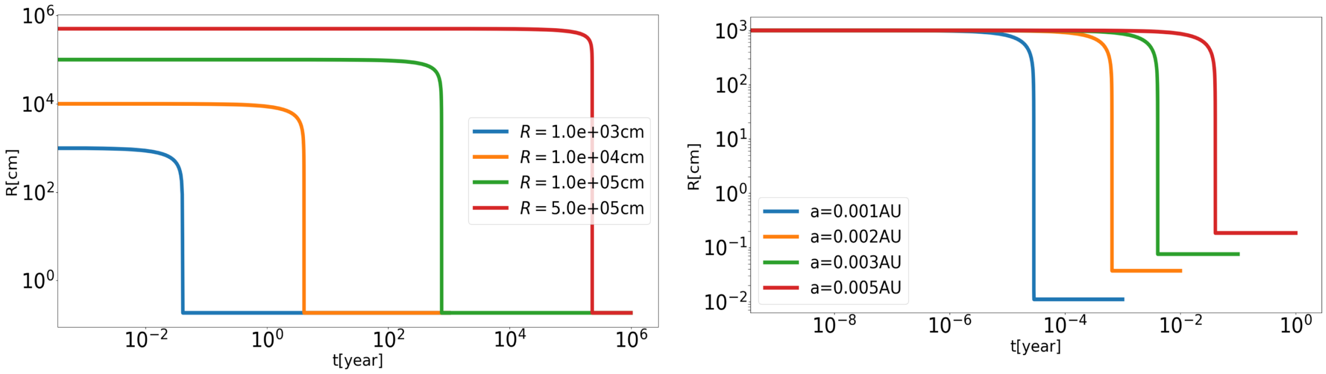

Figure 2. The time evolution of objects embedded in a WD disc due to aeolian erosion. Left: Evolution at a constant distance $a_0 = 0.005$ AU $\approx 1\,\mathrm{R}_{\odot}$ from the WD. Right: Different distances from the WD, with a constant initial size of embedded objects $R_0 = 10^3$ cm.

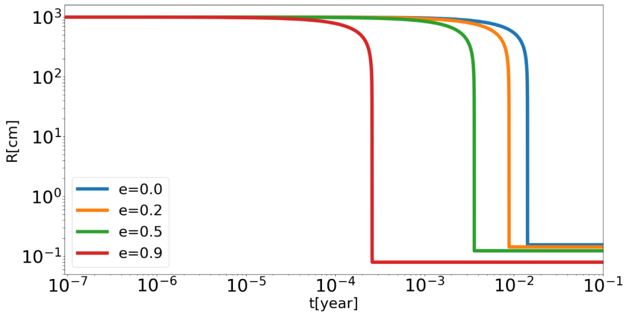

Figure 3. The effect of aeolian erosion on an 10^3 cm object embedded in a WD disc at a distance of $1\,\mathrm{R}_{\odot}$ from the centre of the WD. The relative velocity of the object is considered as the maximal velocity in the orbit.

Aeolian erosion in WD discs is quite efficient, as manifested in Fig. 2. The time-scales for aeolian erosion are extremely short, and for a constant distance $a = 0.005$ AU $\approx 1\,\mathrm{R}_{\odot}$ from the WD disc, objects with radii as large as $\sim 5 \times 10^5$ cm are eroded within the expected disc lifetime. The chosen fixed distance from the WD, along with the rest of the parameters that we do not vary currently, dictates the final size to which objects are ground down – from ~ 0.01 cm for 0.001 AU to ~ 0.18 cm for 0.005 AU.

The final size is determined by the initial conditions of the eroded object, and induces a size re-distribution according to distance from the WD. The outer parts of the disc are more dilute then the inner ones, and since the rate of aeolian erosion is proportional to the gas density, as can be seen in equation (12), aeolian erosion is more efficient in the inner part of the discs, which leads to shorter time-scales and smaller final sizes. Aeolian erosion has a Goldilocks region of sizes in which it attains its maximal efficiency, since the relative velocity, which plays a significant role in this process, varies with the coupling of objects to the gas.

4.1 Eccentric orbits

Aeolian erosion depends strongly on the relative velocity (to the third power), as can be seen from equation (12). Hence, its effects on objects on eccentric orbits may differ substantially from the effects in the circular case, and lead to stronger more significant erosion due to the higher velocities involved, that might even be supersonic and under some conditions lead to the prompt disruption of objects (Demirci et al. 2020a).

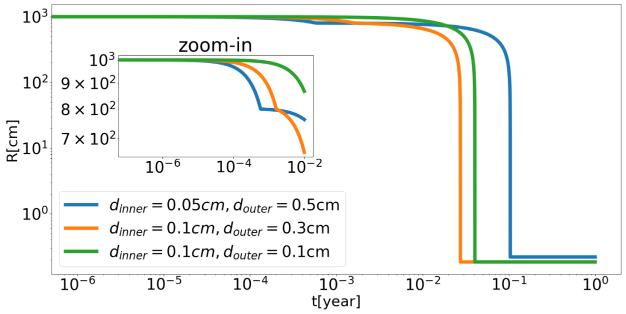

Figure 4. The evolution of two-layered differentiated objects (see inset legend), with an outer layer comprising 20 per cent of the radius and an inner core of the remaining 80 per cent of the radius, under the effects of aeolian erosion; all of the layers have the same density of $3.45\,\mathrm{g\,cm^{-3}}$. The object is embedded in a WD disc and resides at a distance of 0.005 AU from the host WD.

The velocity of a planetesimal in an eccentric orbit is given by

$$\mathbf{v}_{\mathrm{p}} = v_{\mathrm{k}}\sqrt{2 - \frac{r}{a}}\hat{\mathbf{v}}_{\mathrm{p}} = \sqrt{GM\left(\frac{2}{r} - \frac{1}{a}\right)}\hat{\mathbf{v}}_{\mathrm{p}} \tag{14}$$

where M is the mass of the WD, v_k is the Keplerian velocity, r is the distance and a is the semimajor axis.

Assuming a circular disc, the gas moves with a velocity

$$\mathbf{v}_{\mathrm{g}} = v_{\mathrm{k}}\sqrt{1-\eta}\hat{\mathbf{v}}_{\mathrm{g}}. \tag{15}$$

The magnitude of the relative velocity between the gas and the planetesimal is given by

$$v_{\mathrm{rel}} = |\hat{\mathbf{v}}_{\mathrm{p}} - \hat{\mathbf{v}}_{\mathrm{g}}| = \sqrt{v_{\mathrm{p}}^2 + v_{\mathrm{g}}^2 - 2v_{\mathrm{p}}v_{\mathrm{g}}\hat{\mathbf{v}}_{\mathrm{g}}\cdot\hat{\mathbf{v}}_{\mathrm{p}}}. \tag{16}$$

For simplicity, we will assume that the evolution of objects experiencing aeolian erosion is dominated by the maximal relative velocity in the orbit, when the phase between the gas and the objects is maximal – $\pi/2$, and also assume that the object is at the pericentre, i.e. $r = r_{\mathrm{p}} = a(1 - e)$. See Mai et al. (2020) for more detailed results.

As can be seen in Figs 3, motion in eccentric orbits shortens the time-scales of aeolian erosion and the final size of eroded objects is smaller and could attain ~ 0.08 cm for $e = 0.9$.

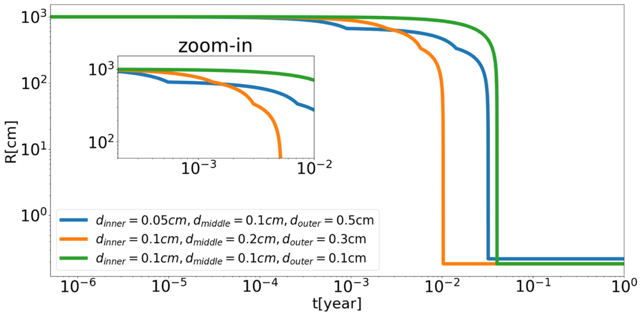

Figure 5. The evolution of triple-layered objects, with three equal layers, with each layer comprising 1/3 of the radius, of the same density 3.45 g cm^{-3}, but with different size of inner grains that build each layer (see inset legend). The object is embedded in a WD disc and resides at a distance of 0.005 AU from the WD.

4.2 Multilayer objects

Previously, we discussed objects which are composed from a single-size grain distribution. However, the physical reality might be more complicated and could give rise to a non-trivial internal size distribution. In the following we relax homogeneous composition assumption, and consider the effects of aeolian erosion on inhomogeneous objects. To the best of our knowledge no previous study considered the internal structure of grains within WD discs, and we adopt internal size distribution models usually considered for asteroids.

The Brazil nut effect in asteroids (Matsumura et al. 2014) suggests that in a mixture of particles, the larger ones tend to end up on the surface of objects. Therefore, when this process acts, the inner structure of an asteroid is such that the larger grains are in the outer layers.

In Figs 4 and 5, we present how differentiated objects react to aeolian erosion. Erosion enables us to decompose the layers of an object – as far as they are in the correct regimes in which erosion is effective – and to reveal the inner layers in short time-scales. Since different sizes of grains impose different rates of aeolian erosion, the total aeolian erosion time-scales change such that considering larger grains shortens the time-scales, and smaller grains lengthen them. Moreover, it can be seen that the final size is determined by the innermost layer, and larger grains lead to smaller final sizes.

4.3 Thermal effects

At high temperatures, thermal processes become more significant and might strengthen the effect of aeolian erosion and give rise to further destruction and shorten the time-scales.

As can be seen in Fig. 6, thermal ablation shortens the time-scales in which objects are ground down to their final scales – here manifested for icy objects. Rocky objects will require higher disc temperatures for thermal ablation to be significant, since the critical temperature for rock is 4000 K; due to the large range of possibilities for WD discs in general – and for temperatures in particular – the temperature in the disc might even exceed the critical temperature for rocks and other materials, at least for the hottest youngest WDs. For coated objects, there could be a combined process of ablating the outer icy layer and then mechanically eroding inner layers.

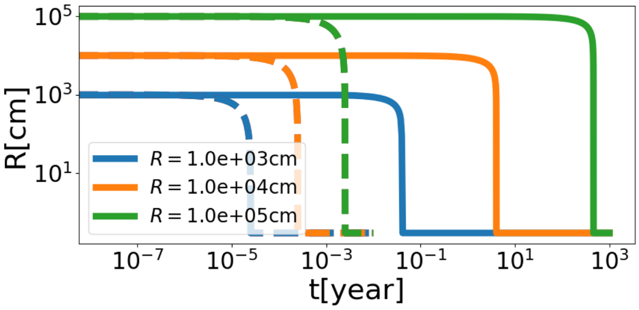

Figure 6. The relationship between aeolian erosion and ablation for an icy object with initial size of 10^3 cm, embedded at a constant distance of 0.005 AU from the WD. Solid lines describes aeolian erosion only, and dashed lines describe the combined effect of aeolian erosion and thermal ablation (above the threshold temperature for ice).

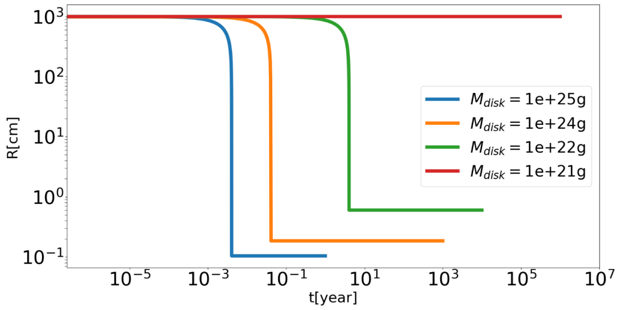

Figure 7. The dependence on disc mass for an object with initial size of 10^3 cm, and constant distance from the centre of 0.005 AU.

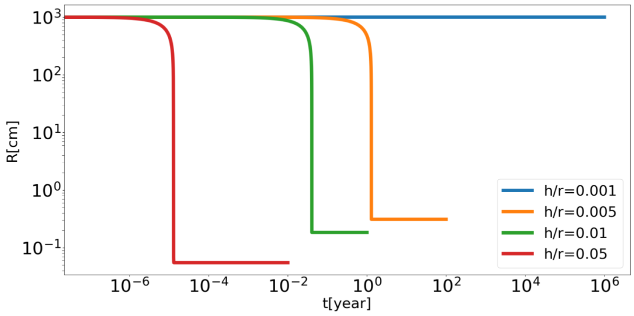

Figure 8. The dependence on aeolian erosion on the aspect ratio of the disc, $h/r = c_s^2/r\Omega$ where Ω is the angular velocity, for objects of initial size 10^3cm, embedded at the disc in a distance of 0.005 AU.

5 DISCUSSION AND IMPLICATIONS

5.1 Parameters dependence

The possible parameter range for WD discs is wide and enables us to study a multitude of combinations of parameters. Here we will present a parameter space exploration for solid bodies embedded in WD discs and subjected to the effects of aeolian erosion.

In Fig. 7 we examine the dependence of aeolian erosion on the disc mass. As can be seen in equation (1), the surface density of the disc, and hence the gas density, grow linearly with the disc mass. Since the aeolian erosion rate is proportional to the gas density, the rate becomes stronger for larger disc masses.

In Fig. 8, we present the dependence of aeolian erosion on the aspect ratio. Higher aspect ratios lead to stronger aeolian erosion.

And in our default choice of parameters, the aspect ratio should be $\gtrsim 10^{-3}$ in order to maintain significant aeolian erosion.

5.2 Symbiotic relations with other processes in the disc

5.2.1 *Collisional cascade*

Fragmentational collisional erosion is another destructive process that gradually diminishes the mass of objects in WD discs. This destructive process generates a collisional cascade which grinds down $10^5 - 10^7$ cm objects to 10^{-6} cm objects within $10^2 - 10^6$ years (e.g. Kenyon & Bromley 2017b, c). Collisional cascades and aeolian erosion both cause objects to lose mass, such that in the presence of gas, these two processes are symbiotic.

The aeolian erosion time-scale is given by

$$t_{\rm erosion} = \frac{R}{|\dot{R}|} = \frac{4\pi R^2 \rho_p a_{\rm coh}}{\rho_g v_{\rm rel}^3}. \tag{17}$$

The fragmentation time-scale is estimated from the collisional time-scale. The collision time-scale for a mono-disperse swarm is given by (Kenyon & Bromley 2017a, b, c and references therein)

$$t_0 = \frac{r_0 \rho P}{12\pi \Sigma_g}, \tag{18}$$

where r_0 is the radius of all the objects in the swarm, ρ_p is their density, P is their orbital period and Σ_0 is the initial surface density. When a multi-disperse swarm is considered, the modification of the time-scale is parametrized by a collision parameter $\alpha_c \propto (v^2/Q_D^*)^{-1}$, where v is the collision velocity and Q_D^* is the binding energy of the object (see Leinhardt & Stewart 2012 and references therein), such that the time-scale from multi-species is given by $t_c = \alpha t_0$, which can be shorter compared with the mono-disperse collision time-scales. When the replenishment is efficient enough, the background distribution of objects remains roughly constant. However, generally, the background density changes with time as well, adding complications to the analysis.

The ratio between the time-scales of aeolian erosion and fragmentational collisions is given by

$$\frac{t_{\rm erosion}}{t_{\rm fragment}} = \frac{96\pi^2 h a_{\rm coh}}{\alpha_c P v_{\rm rel}^3} R \tag{19}$$

such that the time-scales are comparable for a ratio around unity. From equating the ratio to unity, one can set the transition radius between the aeolian erosion dominated regime and the collisional fragmentation dominated regime, which depends on the rest of the parameters.

In Fig. 9, we compare the typical time-scales of aeolian erosion and fragmentational collisions at which objects are destroyed. For the fragmentation collision time-scales, we consider a swarm with a maximal size given by $R_{\rm max}$, and assume for simplicity that the fragmentation of intermediate objects (i.e. not the largest or the smallest object in the swarm) starts just when they become the largest of the swarm, which is justified by the hierarchical character of the fragmentational cascade. For the background distribution, we assume a power-law distribution of $N(r) \sim r^{-3.5}$ (Kenyon & Bromley 2017a), and a value of v^2/Q_D^* that yields $\alpha_c \sim 10^2$. Aeolian erosion dominates at small sizes, and fragmentational collisions at larger sizes. The transition point between the regimes varies with the distance from the WD, such that the regime of aeolian erosion grows with smaller distances. The aeolian erosion time-scales are very short in these regimes, but might be comparable to or longer than the orbital period time-scales of the objects.

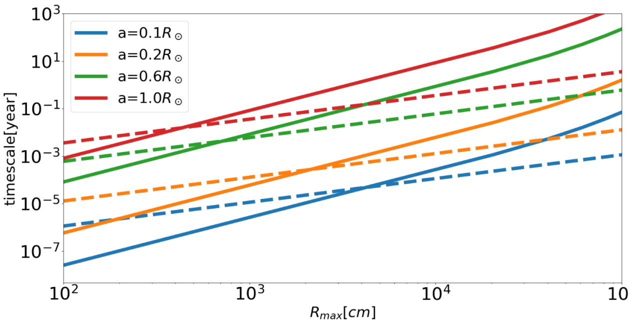

Figure 9. The typical time-scales of aeolian erosion (solid lines) and fragmentational collision (dashed lines) as a function of the destructed object's radius, at different constant distances from the WD. $R_{\rm max}$ is the radius of the largest object in the swarm.

We would like to stress that although the destructive character of both aeolian erosion and collisional cascades leads to shrinkage of objects in the discs, they differ by their intrinsic physical mechanisms. Aeolian erosion originates from shear pressure that is induced by gas drag. Hence, the presence of gas in the disc is a necessary condition to initiate this process. The collisional cascade arises from collisions between particles in the disc.

5.2.2 *External seeding of objects into the disc*

The abundance of large objects in WD discs might be replenished as a result of seeding (Grishin & Veras 2019; Grishin, Perets & Avni 2019). Exo-planetesimals from external sources could enrich the abundance of objects in the disc that will eventually be eroded. External capture especially contributes to the number of large objects, although the total captured mass is small compared to the dust initial abundance. Furthermore, seeding could bring into the disc new materials that might eventually end as pollutants on the WD. The capture rate is dictated by the supply rate and the capture probability; both vary with the origin and the size of the captured objects.

The captured objects change the size distribution in the disc, and might contribute to the steady-state distribution. The injection of external objects might be at a sufficient rate to cancel out the destructive processes and maintain a distribution with larger objects than expected. The revised size distribution should be derived from a full swarm simulation which includes collisional cascade, aeolian erosion, growth and seeding.

5.2.3 *Further disc generations*

The parameters and characteristics of the disc might vary from one generation to another. Wide-binary evolution could lead to an evolved donor that transfers mass to its companion, such that the captured material forms a disc (Perets 2010; Perets & Kenyon 2013; Schleicher & Dreizler 2014; van Lieshout et al. 2018). A binary main-sequence system evolves such that the more massive star sheds material which is accreted on the secondary and forms a protoplanetary disc; the mass that is lost from the system, which is much greater than the mass transferred from one star to the other, leads to expansion of the binary orbit (Veras et al. 2011; Kratter & Perets 2012; Veras & Tout 2012). Afterwards, a second generation of debris and planets form in the pre-evolved system, such that the secondary evolves off the main sequence and sheds material to its WD companion and a protoplanetary disc formed – similarly to the previous stages. Finally, the binary orbit expands and a third-generation debris disc is formed.

Higher generation discs might give rise to different disc composition, since metallicity varies from one generation to another, such that the formed accretion discs are metal and dust rich. Higher metallicity environments are better for planet formation (Fischer & Valenti 2005), such that the formed discs are likely to have planetary and planetesimal structures.

5.3 Possibilities for wind erosion of SDSS 1288+1040

One notable potential application of aeolian erosion theory is the disc orbiting the WD SDSS J1228+1040 (Manser et al. 2019). This disc contains a planetesimal on an orbit with $a = 0.73R_{\odot}$ and $e \approx 0.54$. This orbit is fully embedded in a dusty and gaseous planetary debris disc, which extends beyond $1R_{\odot}$.

Because of this compact planetesimal orbit, and because the planetesimal has been observed over at least 4,000 orbits, it cannot represent a typical rubble-pile Solar system asteroid. Instead, it likely represents an iron-rich remnant core of a planet, and harbours non-zero internal tensile strength. The internal structure and size are poorly constrained: the estimated radius range is $R = 1.2$–120 km.

Hence, for only the lowest end of this size range could aeolian erosion could be effectual. Fig. 2 illustrates that km-sized objects may be eroded in time-scales under $\sim10^6$ yr. However, the structures of the objects in that figure are likely to be very different than the strong and dense planetesimal around SDSS 1228+1040. Furthermore, a potentially competing effect are bodily gravitational tides. Veras et al. (2019) show that these tides could instigate a WD to engulf objects containing just $10^{-3}M_{\oplus}$, but only for objects with sufficiently low internal viscosities.

6 CAVEATS

Aeolian erosion is a very effective process in WD discs, as long as the amount of gas in the disc is non-negligible. The current observations of WD discs with a gaseous component cannot yet well constrain several of the disc parameters, such as the disc mass and its scale height. Also, the exact origins of the gas in such disc are not well understood, with suggested origins include sublimation (e.g. Metzger et al. 2012), grain–grain collisional vaporization and sputtering. Recently, Malamud, Grishin & Brouwers (2021) presented another channel of gas production in WD discs, via interactions between an eccentric tidal stream and a pre-existing dusty compact disc.

Along with the destructive processes, there could be mass influx that we did not take into consideration in this paper. As manifested in Kenyon & Bromley (2016), Kenyon & Bromley (2017c) in the context of collisional cascade, the rate of mass input might equalize the mass loss such that objects that are a priori expected to be pulverized and will maintain their mass for long time-scales. However, here the combined effect of fragmentational erosion and aeolian erosion might play a role, and these two processes together will lead to mass depletion of objects.

We have neglected any asphericity in the object, which is assured due to the lack of perfect packing efficiency of its constituent grains. Furthermore, we have neglected any asphericity that develops as a result of an aeolian erosion, which acts in the direction of the headwind, and might reshape the eroded objects.

We focus on aggregates, with a weak outer layer. Some of the objects that could be potentially affected from aeolian erosion are destroyed already in tidal shredding or collisions. The cohesion forces hold for loosely bound objects, i.e. they describe aggregates and other forms of cohesion laws should be taken into account in case of different internal physics. Once the cohesion law is dictated, the suggested prescription of aeolian erosion will be very similar to the one we sketched in this paper.

For aeolian erosion to be effective, the dominant stripping force on the weak outer layers of the aggregates would need to be erosive rather than tidal. The strength of the tidal force can vary significantly depending on physical properties; for large homogeneous rubble-piles, this value can vary by a factor of about 2 (around $1\,\mathrm{R}_{\odot}$) depending on spin and fluidity, and can cause stripping on an intermittent, yearly time-scale (Veras et al. 2017). Such intermittency perhaps suggests that erosive and tidal forces may act in concert in certain cases, particularly as the aggregate changes shape.

Another aspect of the physics which we did not model is the asteroid spin barrier, which refers to the minimum spin rate at which an asteroid breaks itself apart. This barrier is well established at about 2.2 h in the gravity-dominated regime, for spherical rubble piles larger than about 300m (Pravec, Harris & Michalowski 2002; Hu et al. 2020). However, because this spin limit is a general function of both the asphericity of the object and its internal cohesion (see the appendix of Veras, McDonald & Makarov 2020), incorporating this limit into our modelling would not be trivial. Our aggregates change shape through time, and spin variations are not necessarily monotonic.

7 SUMMARY

The growing number of WD discs (both gaseous and non-gaseous) that have been observationally detected and characterized leave open the possibility for constraining theoretical models for the origins and evolution of such discs. However, even at the theoretical level, there are important gaps in our physical understanding of the dynamics and processes that take place in WD discs.

In this paper we focused on the processes of aeolian erosion, which, to date, were not considered in the context of WD discs. We made use of an analytical model for aeolian erosion in WD discs, based on our studies of such processes in protoplanetary discs as presented in Rozner et al. (2020). We find that the typical time-scales of aeolian erosion in WD discs are extremely short, with aeolian erosion grinding down even km-size objects within the disc lifetime. Consequently, such processes are likely to play an important role in the evolution of small solid bodies in the disc. We also studied the relationship between aeolian erosion and other physical processes in WD discs and its amplification due to the combined effect (see Sections 3.2 and 5.2). Along with collisional cascade and thermal ablation, aeolian erosion grinds down efficiently large objects into small ones with a characteristic final size. The eroded objects experience dynamical processes that finally grind down planetesimals/rocks/pebbles/boulders into sufficiently small particles such that these could drift towards the WD via Poynting–Robertson drag and contribute to its pollution.

Aeolian erosion is the most efficient and becomes the dominant destruction process for small objects, and the critical radius for its dominance is determined by the parameters of the disc, the physical characteristics of the eroded objects, the distance from the WD and the parameters of the collisional cascade (see equation 19).

Similarly to protoplanetary discs, aeolian erosion in WD discs induces a re-distribution of particles size, according to the distance of the particles from the WD. Hence, aeolian erosion sets constraints on the parameters of WD discs that might narrow down the current parameter space. Due to the extremely short time-scales of aeolian erosion, it is not likely that observable variations in WD discs would be explained by replenishment from aeolian erosion.

ACKNOWLEDGEMENTS

We thank the anonymous reviewer for useful comments which have improved the manuscript. MR and HBP acknowledge support from the Minerva Center for Life Under Extreme Planetary Conditions, the Lower Saxony-Israel Niedersachsisches Vorab research cooperation fund, and the European Union's Horizon 2020 Framework Programme under grant agreement No 865932-ERC-SNeX. DV gratefully acknowledges the support of the STFC via an Ernest Rutherford Fellowship (grant ST/P003850/1).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

Bagnold R. A., 1941, The Physics of Blown Sand and Desert Dunes. Methuen, London
Bonsor A., Carter P. J., Hollands M., Gänsicke B. T., Leinhardt Z., Harrison J. H. D., 2020, MNRAS, 492, 2683
Burns J. A., Lamy P. L., Soter S., 1979, Icarus, 40, 1
D'Angelo G., Podolak M., 2015, ApJ, 806, 203
Demirci T., Wurm G., 2020, A&A, 641, A99
Demirci T., Schneider N., Teiser J., Wurm G., 2020a, A&A, 644, A20
Demirci T., Schneider N., Steinpilz T., Bogdan T., Teiser J., Wurm G., 2020b, MNRAS, 493, 5456
Dennihy E. et al., 2020, ApJ, 901, 5
Doyle A. E., Young E. D., Klein B., Zuckerman B., Schlichting H. E., 2019, Science, 366, 356
Farihi J., 2016, New Astron. Rev., 71, 9
Farihi J. et al., 2018, MNRAS, 481, 2601
Fischer D. A., Valenti J., 2005, ApJ, 622, 1102
Gänsicke B. T., Marsh T. R., Southworth J., Rebassa-Mansergas A., 2006, Science, 314, 1908
Gentile Fusillo N. P. et al., 2020, preprint (arXiv:2010.13807)
Girven J., Brinkworth C. S., Farihi J., Gänsicke B. T., Hoard D. W., Marsh T. R., Koester D., 2012, ApJ, 749, 154
Grishin E., Veras D., 2019, MNRAS, 489, 168
Grishin E., Perets H. B., Avni Y., 2019, MNRAS, 487, 3324
Grishin E., Rozner M., Perets H. B., 2020, ApJ, 898, L13
Harrison J. H. D., Bonsor A., Madhusudhan N., 2018, MNRAS, 479, 3814
Hollands M. A., Gänsicke B. T., Koester D., 2018, MNRAS, 477, 93
Hu S., Richardson D. C., Zhang Y., Ji J., 2020, preprint (arXiv:2009.12605)
Jura M., 2008, AJ, 135, 1785
Jura M., Young E. D., 2014, Annu. Rev. Earth Planet. Sci., 42, 45
Kenyon S. J., Bromley B. C., 2016, ApJ, 817, 51
Kenyon S. J., Bromley B. C., 2017a, ApJ, 839, 38
Kenyon S. J., Bromley B. C., 2017b, ApJ, 844, 116
Kenyon S. J., Bromley B. C., 2017c, ApJ, 850, 50
Koester D., Gänsicke B. T., Farihi J., 2014, A&A, 566, A34
Kratter K. M., Perets H. B., 2012, ApJ, 753, 91
Kruss M., Musiolik G., Demirci T., Wurm G., Teiser J., 2019, preprint (arXiv:1911.01692)
Leinhardt Z. M., Stewart S. T., 2012, ApJ, 745, 79
Mai C., Desch S. J., Kuiper R., Marleau G.-D., Dullemond C., 2020, ApJ, 899, 54
Malamud U., Perets H. B., 2020a, MNRAS, 492, 5561
Malamud U., Perets H. B., 2020b, MNRAS, 493, 698
Malamud U., Grishin E., Brouwers M., 2021, MNRAS, 501, 3806
Manser C. J. et al., 2019, Science, 364, 66
Manser C. J., Gänsicke B. T., Gentile Fusillo N. P., Ashley R., Breedt E., Hollands M., Izquierdo P., Pelisoli I., 2020, MNRAS, 493, 2127
Matsumura S., Richardson D. C., Michel P., Schwartz S. R., Ballouz R.-L., 2014, MNRAS, 443, 3368
Melis C., Klein B., Doyle A. E., Weinberger A. J., Zuckerman B., Dufour P., 2020, ApJ, 905, 56
Metzger B. D., Rafikov R. R., Bochkarev K. V., 2012, MNRAS, 423, 505
Opik E. J., 1958, Physics of Meteor Flight in the Atmosphere. Interscience, New York.
Paraskov G. B., Wurm G., Krauss O., 2006, ApJ, 648, 1219
Perets H. B., 2010, preprint (arXiv:1001.0581)
Perets H. B., Kenyon S. J., 2013, ApJ, 764, 169
Perets H. B., Murray-Clay R. A., 2011, ApJ, 733, 56
Podolak M., Pollack J. B., Reynolds R. T., 1988, Icarus, 73, 163
Pollack J. B., Hubickyj O., Bodenheimer P., Lissauer J. J., Podolak M., Greenzweig Y., 1996, Icarus, 124, 62
Pravec P., Harris A. W., Michalowski T., 2002, in Bottke W. F., Cellino A., Paolicchi P., Binzel R. P., Asteroids III. Univ. Arizona Press, Tucson, AZ, p. 113,
Rafikov R. R., 2011a, MNRAS, 416, L55
Rafikov R. R., 2011b, ApJ, 732, L3
Rozner M., Grishin E., Perets H. B., 2020, MNRAS, 496, 4827
Schaffer N., Johansen A., Cedenblad L., Mehling B., Mitra D., 2020, A&A, 639, A39
Schleicher D. R. G., Dreizler S., 2014, A&A, 563, A61
Shao Y., Lu H., 2000, J. Geophys. Res., 105, 22,437
Shestakova L. I., Demchenko B. I., Serebryanskiy A. V., 2019, MNRAS, 487, 3935
Swan A., Farihi J., Koester D., Holland s M., Parsons S., Cauley P. W., Redfield S., Gänsicke B. T., 2019, MNRAS, 490, 202
Swan A., Farihi J., Wilson T. G., Parsons S. G., 2020, MNRAS, 496, 5233
van Lieshout R., Kral Q., Charnoz S., Wyatt M. C., Shannon A., 2018, MNRAS, 480, 2784
Veras D., 2016, R. Soc. Open Sci., 3, 150571
Veras D., Heng K., 2020, MNRAS, 496, 2292
Veras D., Tout C. A., 2012, MNRAS, 422, 1648
Veras D., Wyatt M. C., Mustill A. J., Bonsor A., Eldridge J. J., 2011, MNRAS, 417, 2104
Veras D., Leinhardt Z. M., Bonsor A., Gänsicke B. T., 2014, MNRAS, 445, 2244
Veras D., Leinhardt Z. M., Eggl S., Gänsicke B. T., 2015, MNRAS, 451, 3453
Veras D., Carter P. J., Leinhardt Z. M., Gänsicke B. T., 2017, MNRAS, 465, 1008.
Veras D. et al., 2019, MNRAS, 486, 3831
Veras D., McDonald C. H., Makarov V. V., 2020, MNRAS, 492, 5291
Weidenschilling S. J., 1977, MNRAS, 180, 57
Xu S. et al., 2018, ApJ, 866, 108
Zuckerman B., Melis C., Klein B., Koester D., Jura M., 2010, ApJ, 722, 725

APPENDIX A: TABLE OF COMMONLY USED PARAMETERS

In this appendix, we present the default values for the used parameters, unless stated otherwise.

Table A1. Supplementary parametres.

Symbol	Definition	Fiducial Value	Reference
γ		$0.165\,\mathrm{g\,s^{-2}}$	Kruss et al. (2019)
A_N		1.23×10^{-2}	Shao & Lu (2000)
β		$10^2\ \mathrm{g\ s^{-1}}$	scaled from Paraskov et al. (2006) and refs. therein
ρ_p	Planetesimals' density	rock $3.45\,\mathrm{g\,cm^3}$, ice $1.4\,\mathrm{g\,cm^3}$	Pollack et al. (1996)
μ	Mean molecular weight	$3.85 \times 10^{-24} g$	Perets & Murray-Clay (2011)
σ	Neutral collision cross-section	10^{-15} cm	Perets & Murray-Clay (2011)
T	Temperature (optically thick)	1000 K	
M_disc	Disc mass	10^{24} g	
Σ_g	Surface density profile (optically thick)	5.1×10^3 g cm^2	Grishin & Veras (2019)
h/r	Aspect ratio	10^{-2}	
d	Typical 'building-block' grain size	0.1 cm	
T_cr	Critical temperature	ice 648K, rock 4000 K	Podolak et al. (1988)
L_s	Particle specific vaporization energy (solid)	ice 2.83×10^{10} erg g^{-1}, rock 8.08×10^{10} erg g^{-1}	D'Angelo & Podolak (2015)
	Size dist. in the disc (fragmentation only)	$N(r) \propto r^{-3.5}$	Kenyon & Bromley (2017b)

This paper has been typeset from a TEX/LATEX file prepared by the author.

Chapter 3

The Wide-Binary Origin of The Pluto-Charon System

Based on Rozner et al. (2020c)

The dwarf planet Pluto is the first and largest object in the Kuiper-belt to be discovered (Tombaugh, 1946), and its most massive moon is Charon. Together they constitute the first Kuiper-belt binary (KBB) to be discovered (Christy and Harrington, 1978). The origin of Pluto-Charon binary is currently unknown, though several models were suggested. In this study, we used the LK mechanism, presented in section 1.1.2, and suggested that the Pluto-Charon binary originated from a wide-binary that finally got closer. We used analytical calculations and N-body simulations in order to study this formation channel.

The leading model to the formation of Pluto-Charon is the giant impact model (McKinnon, 1984, 1989; Canup, 2005, 2011; Desch, 2015; Sekine et al., 2017). According to this model, the progenitor of Charon hit proto-Pluto with a velocity comparable to the escape velocity and either merged (immediately/after a rebound Leinhardt and Stewart, 2012) with Pluto and ejected a massive disk of debris which then formed Charon, or grazed Pluto and was then directly captured to be the currently observed Charon. However, there are some problems in the giant impact model, and the major one is that probability of low-velocity collision between unrelated (unbound) most massive KBOs in the Solar system (proto-Pluto and proto-Charon) is a potentially low-probability event (Canup, 2005). In Rozner et al. (2020c) we showed that the formation channel from wide-binaries presents a large parameter space and doesn't require potential fine-tune as in the giant impact model.

In fact, all the binaries in the Solar system – including Pluto-Charon – are a part of a hierarchical triple system with the third companion being the Sun; this provides a fertile ground for significant secular and quasi-secular effects Perets and Naoz (2009); Grishin et al. (2020a). The significant hierarchy allows to treat the triple system Pluto-Charon – Sun as an inner binary – the Pluto-Charon binary – orbited by the outer binary – the Sun and analyze the evolution in secular/quasi-secular tools that could naturally give rise to a low-velocity grazing impact at high inclination which is a prerequisite for the formation of the Pluto-Charon system. It provides an alternative and possibly more robust channel for the origin of this system, consistent with the likely origin of other contact KBBs.

Consider an inner wide-binary composed from the progenitors of Pluto and Charon, with separation $a_{\rm in}$, eccentricity $e_{\rm in}$ and total mass $m_{\rm in}$, and the Sun as a distant perturber, forming together a hierarchical triple where the separation of the outer binary is given by $a_{\rm out}$ and the mass by $m_{\rm out}$. The dynamical evolution of the system could be described by four main regimes: non-collisional, precession dominated (without collision), collisional secular (or quasi-secular) evolution and collisional non-secular evolution. The initial conditions of the system, i.e. initial mutual inclination i_0, eccentricity e_0 and the separation between the inner binary companions $a_{\rm in}$, determine the regime; or equivalently $j_z = \sqrt{1-e^2}\cos i$ and $a_{\rm in}$, where i is the mutual inclination of the inner binary. Note that from the features of the secular and quasi-secular evolution, $a_{\rm in}$ and $a_{\rm out}$ are approximately conserved through the evolution. We normalize the inner semi-major axis by the Hill radius, $\alpha = a_{\rm in}/R_H$. The Hill radius is defined by $R_H = a_{\rm out}(1-e_{\rm out})\left((m_{\rm Pluto}+m_{\rm Charon})/3M_\odot\right)^{1/3} \approx 6\times10^6$ km. In Fig. 3.1 we present a roadmap for the regimes of the problem

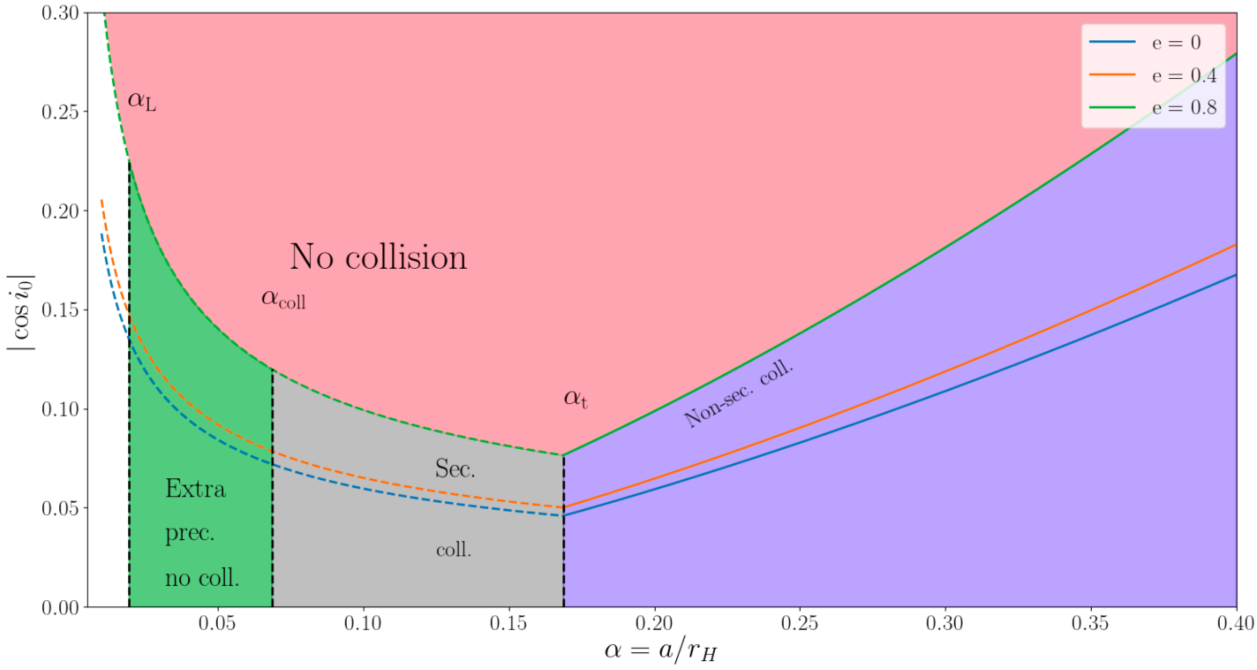

Figure 3.1: Parameter-diagram of the different possible regimes of behavior of Pluto-Charon system. The colored regimes correspond to the different dynamical regimes of the system: no collision, extra precession no collision, secular and non-secular evolution. The lines correspond to illustrations of systems with different hierarchies – encapsulated in α, and different eccentricities. Adopted from Rozner et al. (2020c).

The critical α-s where calculated analytically, by comparing the 'strength' of each regime and comparing between them, see a detailed derivation in Rozner et al. (2020c).

In order to gather statistics on the system, and verify our analytic results, we simulate numerically the dynamics of the progenitors of Pluto-Charon and the Sun, See Fig. 3.2 for the cumulative distribution function. We used the publicly available N-body code `REBOUND` (Rein and Liu, 2012a); we chose `IAS15`, a fast, adaptive, high-order integrator for gravitational dynamics, accurate to machine precision over a billion orbits (Rein and Spiegel, 2015).

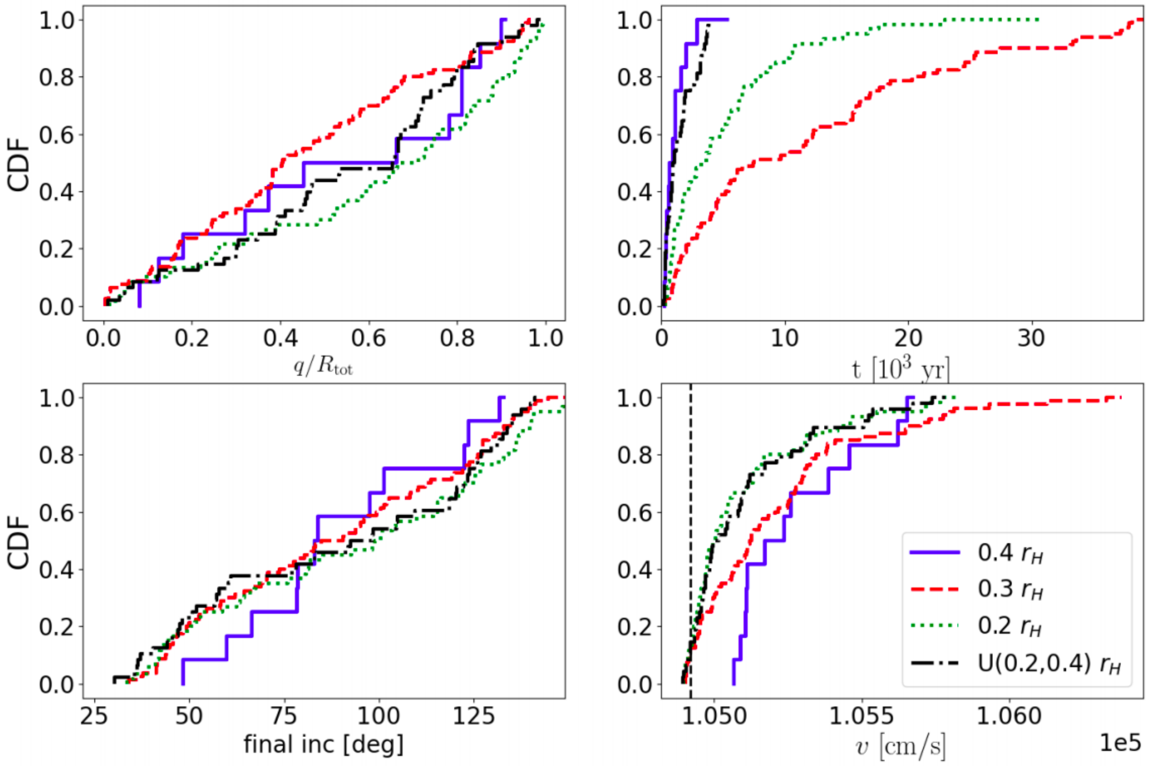

Figure 3.2: Cumulative distributions of the impact properties. Upper left: Normalized pericenter $q/R_{\rm tot}$; Upper right: Time of collision; Lower left: Final inclination at impact; Lower right: Velocity at impact. The vertical dashed line is the escape velocity. Adopted from Rozner et al. (2020c).

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **497,** 5264–5270 (2020)
Advance Access publication 2020 August 18

doi:10.1093/mnras/staa2446

The wide-binary origin of the Pluto–Charon system

Mor Rozner ,★ Evgeni Grishin and Hagai B. Perets
Technion, Israel Institute of Technology, Haifa, 3200003, Israel

Accepted 2020 August 12. Received 2020 August 10; in original form 2020 July 20

ABSTRACT

The Pluto–Charon binary system is the best studied representative of the binary Kuiper-belt population. Its origins are vital to understanding the formation of other Kuiper-belt objects (KBO) and binaries and the evolution of the outer Solar system. The Pluto–Charon system is believed to form following a giant impact between two massive KBOs at relatively low velocities. However, the likelihood of a random direct collision between two of the most massive KBOs is low and is further constrained by the requirement of a low-velocity collision, making this a potentially fine-tuned scenario. Here, we expand our previous studies and suggest that the proto-Pluto–Charon system was formed as a highly inclined wide-binary, which was then driven through secular/quasi-secular evolution into a direct impact. Since wide-binaries are ubiquitous in the Kuiper belt with many expected to be highly inclined, our scenario is expected to be robust. We use analytic tools and few-body simulations of the triple Sun–(proto-)Pluto–Charon system to show that a large parameter space of initial conditions leads to such collisions. The velocity of such an impact is the escape velocity of a bound system, which naturally explains the low-velocity impact. The dynamical evolution and the origins of the Pluto–Charon system could therefore be traced to similar secular origins as those of other binaries and contact-binaries (e.g. Arrokoth) and suggest that they play a key role in the evolution of KBOs.

Key words: Kuiper belt: general – Kuiper belt objects: Pluto – Kuiper belt objects: Charon – planets and satellites: dynamical evolution and stability – planets and satellites: formation.

1 INTRODUCTION

The Kuiper belt hosting numerous Kuiper-belt objects (KBOs) is a relic of ancient era of the Solar system and henceforth preserves valuable clues regarding the dynamics that led the Solar system to its current state. KBO-binaries (KBBs) are ubiquitous among KBOs, in particular massive ones, and tens of per cents of the current large KBOs are found to be part of bound binary (and satellite) systems (Goldreich, Lithwick & Sari 2002; Noll et al. 2008; Fraser et al. 2017).

The dwarf-planet Pluto was the first and largest KBO to be discovered (Tombaugh 1946), as well as the first KBB to be found, with the discovery of its most massive companion Charon (Christy & Harrington 1978). Recent data from the New Horizons spacecraft set firmly the size and density of the current Pluto and Charon to be $R_{\rm Pluto} = 1188.3 \pm 1.6$ km, $\mathrm{R}_{\rm Charon} = 606 \pm 1$ km, $\rho_{\rm Pluto} = 1854 \pm 11$ kg/m^3, and $\rho_{\rm Charon} = 1701 \pm 33$ kg/m^3 Nimmo et al. (2017). The mass ratio between Charon and Pluto, which is given by 0.1218:1, the relatively close distance between them of $\sim 2 \times 10^7$ m (Stern et al. 2015), and the large mutual inclination of $\sim 119^\circ$ to Pluto's orbit (Naoz, Perets & Ragozzine 2010) are unique among moons in the Solar system.

There are three major models suggested for the formation of KBBs: gravitational collapse, a giant impact, and dynamical capture. The gravitational collapse *in situ* formation (Nesvorný, Youdin & Richardson 2010) requires a large gravitationally unstable pebble cloud – at least as massive as the Pluto–Charon system – since mass is lost during the formation. However, simulations showed that the formation of such a particle cloud is unlikely (Johansen et al. 2015). In the giant impact scenario (McKinnon 1984, 1989; Canup 2005, 2011; Desch 2015; Sekine et al. 2017), the progenitor of Charon hit proto-Pluto with a velocity comparable to the escape velocity and either merged (immediately/after a rebound, Leinhardt & Stewart 2012) with Pluto and ejected a massive disc of debris, which then formed Charon, or grazed Pluto and was then directly captured to be the currently observed Charon. In the dynamical capture scenario of Goldreich et al. (2002), two (typically massive) unbound KBOs become bound through a close passage during which the relative velocities are dissipated through dynamical friction by the planetesimals in their surrounding environment, forming initially very wide-binaries, close to the Hill radius of the system. Further dissipation could drive the orbit into shorter period.

The low-velocity collision between unrelated (unbound) and most massive KBOs in the Solar system (proto-Pluto and proto-Charon) is a potentially low-probability event (Canup 2005), given the rarity of such objects, but it depends on the timing and location (distance from the Sun) of the impact event. Moreover, it is likely to explain neither the formation of the rest of Pluto's moons nor the absence of a fossil bulge (McKinnon et al. 2017; Nimmo et al. 2017). Although Kenyon & Bromley (2014) claim that the number of the collisions might be high, one should note that the abundance of objects in the relevant sizes and velocities in the ancient Kuiper belt is no more than a few tens (Canup 2005). The time of the Pluto–Charon formation is unknown and is restricted to occur only in the 'pre-installation phase', which took place in the first 500 Myr of the Solar system

---

★ E-mail: morozner@campus.technion.ac.il

Published by Oxford University Press on behalf of the Royal Astronomical Society

history (Greenstreet, Gladman & McKinnon 2015); this uncertainty may give rise to corrections of orders of magnitudes in the estimate presented in Canup (2005).

In practice, all binaries in the Solar system are a part of a triple system – with the third companion being the Sun; this provides a fertile ground for significant secular and quasi-secular effects (Perets & Naoz 2009; Grishin et al. 2020). In this view, the Pluto–Charon system is in fact a part of a hierarchical triple, with the Sun as a third distant perturber, which could be important to the dynamics of the Pluto–Charon binary. The significant hierarchy allows to treat the triple system as an inner binary – the Pluto–Charon binary – orbited by the outer binary – the Sun. Here, we show that the secular or quasi-secular evolution of the Sun and the proto-Pluto–Charon triple system could naturally give rise to a low-velocity grazing impact at high inclination, which is a prerequisite for the formation of the Pluto–Charon system. It provides an alternative and possibly more robust channel for the origin of this system, consistent with the likely origin of other contact KBBs (Perets & Naoz 2009; Grishin et al. 2020).

We begin with a brief introduction of the role of secular/quasi-secular evolution in the evolution of KBBs (2), followed by an analytical description of the dynamics of Pluto–Charon binary in the different regimes (3), we then describe our numerical results (4), discuss caveats (5), and summarize (6).

2 SECULAR AND QUASI-SECULAR EVOLUTION OF KBO BINARIES

The three-body problem is one of the most famous non-integrable problems, tracing back to the pioneering work of Poincare (Poincaré 1892). Fortunately, under certain conditions, some cases could be analysed using perturbative methods. Hierarchical triples are systems that contain an inner binary and a distant tertiary. The system could be described as two binaries – the inner one and the outer one. When the period of the outer binary is much larger than the period of the inner one, we can average their orbits and consider them as two ellipse-shaped mass-wires that interact weakly with each other. This averaging method in the analysis of the secular behaviour is the Lidov–Kozai (LK) mechanism (Kozai 1962; Lidov 1962), which gives rise to the exchange of the inner inclination and eccentricity that could be significant and may lead to extreme orbital evolution, and even flips from prograde to retrograde orbits and vice versa (Naoz 2016). This formalism ignores and averages over time-scales shorter than the secular time-scale given by (Kinoshita & Nakai 1999; Antognini 2015)

$$\tau_s \approx \frac{8}{15\pi}\frac{m_1+m_2+m_3}{m_3}\frac{P_{\rm out}^2}{P_{\rm in}}\left(1-e_{\rm out}^2\right)^{3/2}, \quad (1)$$

where m_1 and m_2 are the masses of the companions of the inner binary, m_3 is the mass of the outer tertiary $P_{\rm out}, P_{\rm in}$ are the periods of the outer and inner binaries correspondingly, and $e_{\rm out}$ is the eccentricity of the outer binary.

The standard LK mechanism relies on significantly large separation between the time-scales of the inner and outer binaries, $P_{\rm in}/P_{\rm out}$. When the systems are mildly hierarchical, short-term effects become important as well, and LK assumptions and results lose accuracy. The evolution is in the quasi-secular regime rather in the 'standard' LK regime (Antonini & Perets 2012; Antognini 2015; Luo, Katz & Dong 2016; Grishin, Perets & Fragione 2018).

The LK mechanism and its quasi-secular generalization (Antonini & Perets 2012) enhance the mergers and collisions of a variety of stellar objects (see e.g. for a review Naoz 2016). In particular, it could play a key role in the evolution of KBO binaries and formation of short period and contact-binaries (Perets & Naoz 2009; Naoz et al. 2010; Porter & Grundy 2012; Grishin & Perets 2016; Grishin et al. 2017; Michaely, Perets & Grishin 2017; Grishin et al. 2020; Lyra, Youdin & Johansen 2020).

The collision velocity that was predicted is comparable to the escape velocity or slightly above (Canup 2005), as well as the high inclination of Pluto is suggestive of a possible collision of two-bound objects at high inclination. These are natural outcomes from a secular/quasi-secular evolution and could therefore point to the possible involvement of quasi-secular evolution that drives the formation of the Pluto–Charon binary, similarly to the process that likely formed the contact binary KBO Arrokoth.

The KBO contact binary (2014) MU_{69} (Arrokoth) was discovered by New Horizons search team using the *Hubble Space Telescope* and was chosen as a main target of an extended exploration mission of New Horizons (Stern et al. 2018b, 2019). One of the leading models of the formation of Arrokoth – which is supported by new measurements (Stern et al. 2019; McKinnon et al. 2020) – argues that the formation of Arrokoth arose from a gentle collision between two perturbed wide companions (Grishin et al. 2020). Under certain conditions, wide-binaries could be perturbed significantly such that they could form later contact binaries during secular and quasi-secular evolution.

In the following, we propose and analyse a similar formation channel for the Pluto–Charon system via secular/quasi-secular evolution that leads ultimately to a collision between the inner binary companions – Pluto and Charon, due to perturbations from the distant perturber – the Sun.

3 ANALYTICAL DESCRIPTION

We propose that the current Pluto–Charon short-period binary system originated from a much wider KBO-binary system possibly formed through the gravitational instability or the dynamical capture scenario and later evolved through secular or quasi-secular evolution.

Consider an inner wide-binary composed from the progenitors of Pluto and Charon, with separation $a_{\rm in}$, eccentricity $e_{\rm in}$ and total mass $m_{\rm in}$, and the Sun as a distant perturber, forming together a hierarchical triple where the separation of the outer binary is given by $a_{\rm out}$ and the mass by $m_{\rm out}$. The dynamical evolution of the system could be described by four main regimes: non-collisional, precession dominated (without collision), collisional secular (or quasi-secular) evolution, and collisional non-secular evolution. The initial conditions of the system, i.e. initial mutual inclination i_0, eccentricity e_0, and the separation between the inner binary companions $a_{\rm in}$, determine the regime, or equivalently $j_z = \sqrt{1-e^2}\cos i$ and $a_{\rm in}$, where i is the mutual inclination of the inner binary. Note that from the features of the secular and quasi-secular evolution, $a_{\rm in}$ and $a_{\rm out}$ are approximately conserved through the evolution. We normalize the inner semi-major axis by the Hill radius, $\alpha = a_{\rm in}/R_{\rm H}$. The Hill radius is defined by $R_{\rm H} = a_{\rm out}(1-e_{\rm out})\left((m_{\rm Pluto}+m_{\rm Charon})/3M_\odot\right)^{1/3} \approx 6\times 10^6$ km.

Hereafter, we briefly review the behaviour in the different dynamical regimes.

3.1 Standard LK oscillations

When the periods of the inner and outer binaries are well separated, the Hamiltonian of the problem could be decomposed into two Keplerian Hamiltonians and a weak interaction term between the two orbits (Kozai 1962; Lidov 1962). Over long time-scales, the orbits exchange angular momentum, while the energy exchange is negligible and the inner and outer semi-major axes remain roughly constant. These conditions induce periodic variation of eccentricity and inclination.

The motion is governed by the Hamiltonian with the perturbation expanded in multipole expansion (Harrington 1968)

$$\mathcal{H} = \frac{Gm_1m_2}{2a_{\rm in}} + \frac{Gm_3(m_1+m_2)}{2a_{\rm out}} + \mathcal{H}_{\rm pert};$$
$$\mathcal{H}_{\rm pert} = \frac{G}{a_{\rm out}}\sum_{j=2}^{\infty}\left(\frac{a_{\rm in}}{a_{\rm out}}\right)^j\left(\frac{r_1}{a_{\rm in}}\right)^j\left(\frac{a_{\rm out}}{r_2}\right)^{j+1}\mathcal{M}_jP_j\left(\cos\Phi\right),$$
$$\mathcal{M}_j = m_1m_2m_3\frac{m_1^{j-1}-(-m_2)^{j-1}}{(m_1+m_2)^j} \qquad (2)$$

where r_i is the distance between the two companions of the i-th binary, P_i is the i-th Legendre polynomial, and Φ is the angle between r_2 and r_1.

The standard LK formalism considers double-averaging, i.e. averaging over both inner and outer mean anomalies. The averaging is done in von Zeipel technique (Von Zeipel 1916) and enables an extraction of the secular changes in the system – changes in the orbital elements along time-scales much longer than the orbital period. The lowest order, and most significant, is the quadruple order. Since the inner binary cannot exceed the Hill radius, the next octupole order is weaker by at least $a_{\rm in}/a_{\rm out} \le (m_{\rm in}/3m_{\rm out})^{1/3} \approx 10^{-3}$ and the octupole evolution can be safely neglected. We therefore focus only on quadrupole evolution.

In the quadruple order, the maximal eccentricity, and the corresponding minimal inclination, could be written as (Innanen et al. 1997)

$$e_{\rm max,LK} = \sqrt{1-\frac{5}{3}\cos^2 i_0},\ \ i_{\rm min,LK} = \arccos\left(\pm\sqrt{\frac{3}{5}}\right). \qquad (3)$$

The possible $i_{\rm min}$, 39.23° and 140.77° set the boundaries where LK evolution is active.

Due to the secular evolution of the distant perturber – the Sun – the inner binary, which contains the progenitors of Pluto and Charon, experiences extreme oscillations of its eccentricity and mutual inclination, where the highest eccentricity obtained with the lowest mutual inclination and vice versa. Under the constraints introduced above, the LK mechanism might lead to collisions (Perets & Naoz 2009; Grishin et al. 2020).

3.2 Quasi-secular and non-secular regime

The double-averaging approximation breaks down when the system is mildly hierarchical, i.e. when the inner and outer periods become comparable (e.g. Antonini & Perets 2012; Luo et al. 2016; Grishin et al. 2018). The breakdown leads to corrections in the evolution and, in particular, corrections of the maximal eccentricity and critical inclinations for onset, which will be indexed by QS (Grishin et al. 2018).

$$e_{\rm max,QS} = \sqrt{1-\frac{5}{3}\cos^2 i_0\frac{1+\frac{9}{8}\epsilon_{\rm SA}\cos i_0}{1-\frac{9}{8}\epsilon_{\rm SA}\cos i_0}}, \qquad (4)$$

$$i_{\rm min,QS} = \arccos\left(\pm\sqrt{\frac{3}{5}-\frac{27}{40}\epsilon_{\rm SA}}\right) \qquad (5)$$

were $\epsilon_{\rm SA}$ is the strength of the single averaging, given by equation (12) (Luo et al. 2016; Grishin et al. 2018), $\epsilon_{\rm SA} = P_{\rm out}/2\pi\tau_{\rm s}$ where $P_{\rm out}$ is the period of the outer binary and $\tau_{\rm s}$ is defined in equation (1).

Quasi-secular analysis enables us to treat collisions of less hierarchical systems, i.e. wider inner binaries.

3.3 Precession due to oblateness

Non-spherical shapes of objects lead to corrections in the gravitational potential, which generate an extra precession that might quench the LK mechanism. The leading term that encapsulates the dynamics induced by oblateness is the J_2 coefficient (Murray & Dermott 1999). Nimmo et al. (2017) introduced a detailed study of the observational properties of Pluto and Charon. The measured upper bounds for the oblateness of Pluto and Charon were 0.006 and 0.005 correspondingly. For Pluto, the oblateness is proportional to $R_{\rm Pluto}\omega^2/2g$ up to a constant factor of order unity where $R_{\rm Pluto}$ is Pluto's radius, ω is its rotation angular frequency, and g is the surface gravity.

3.4 Roadmap

Here, we will review the roadmap that describes the transitions between the different regimes, as shown in Fig. 1.

Since the objects are not completely spherical, using a point-mass gravitational potential of spherical objects neglects a potentially important aspect of the evolution. Oblateness induces extra apsidal precession on the Keplerian ellipse. When oblateness-induced precession becomes comparable or larger than that induced by the LK evolution (here taken up to quadruple order), the oscillations are quenched and the behaviour is dictated by the precession. For small enough values of α, LK oscillations are completely quenched due to oblateness-induced precession. Above a critical value of α, signed by α_L, LK oscillations become more important and the oblateness-induced precession is non-negligible and prevents the significant eccentricity growth, thereby avoiding collisions; the parameter space corresponding to this scenario is coloured with green – extra precession no collision. The oblateness manifests itself by a dimensionless constant, J_2, which is the second Laplace coefficient in the expansion of the potential. The ratio between the LK-induced and oblateness-induced precession is given by (Liu, Muñoz & Lai 2015; Grishin et al. 2020)

$$\epsilon_{\rm rot} = \frac{3}{2}J_2\frac{m_{\rm in}}{m_{\rm out}}\frac{a_{\rm out}^3(1-e_{\rm out}^2)^{3/2}R_{\rm Charon}}{\alpha^5R_{\rm H}^5}, \qquad (6)$$

where $J_2 = 0.005$ is the upper bound of the measured oblateness of Charon (Nimmo et al. 2017). Oblateness effects become important when $\epsilon_{\rm rot} \gtrsim 1$, and for $\epsilon_{\rm rot} = 3/2$, we define Laplace radius as $R_{\rm L} = \alpha_{\rm L}R_{\rm H}$ (Tremaine, Touma & Namouni 2009).

The quenching of LK oscillations becomes less significant when $\alpha > \alpha_{\rm L}$, where $\alpha_{\rm L}$ is given by

$$\alpha_{\rm L} = \left(\frac{J_2m_{\rm in}a_{\rm out}^3(1-e_{\rm out}^2)^{3/2}R_{\rm tot}^2}{m_{\rm out}R_{\rm H}^5}\right)^{1/5} \approx 0.02. \qquad (7)$$

The transition between the precession-dominated regime and the secular-collision regime could be derived from the characteristics of the LK oscillations. The possible eccentricity regime for a collision to occur is restricted to the range $e_{\rm min} = e_{\rm coll}$ and $e_{\rm max}$; the minimal and maximal eccentricities depend on the geometric configuration of the system, which affects the LK evolution (e.g. Kinoshita & Nakai 1999; Perets & Naoz 2009; Naoz 2016; Grishin et al. 2017)

$$e_{\rm coll} = 1-\frac{R_{\rm tot}}{\alpha R_{\rm H}},\ \ e_{\rm max} = \sqrt{1-\frac{5}{3}(1-e_{\rm out}^2)\cos^2 i_{\rm out}}. \qquad (8)$$

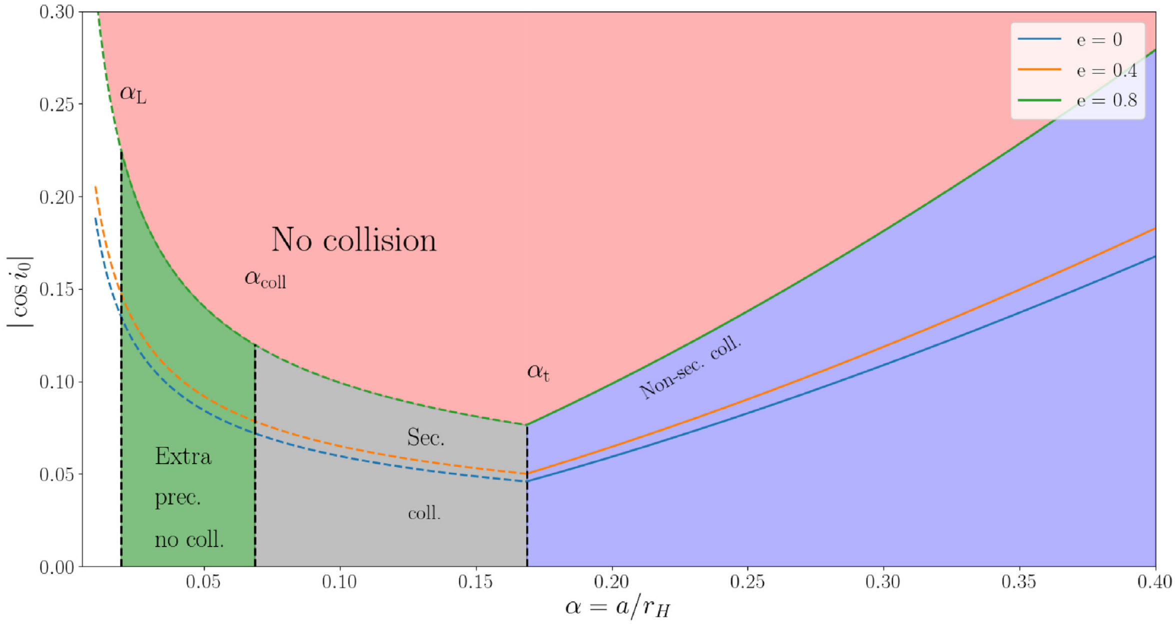

Figure 1. Parameter diagram of the different possible regimes of behaviour of Pluto–Charon system. The coloured regimes correspond to the different dynamical regimes of the system: no collision, extra precession no collision, secular and non-secular evolution. The lines correspond to illustrations of systems with different hierarchies – encapsulated in α and different eccentricities.

The lower bound on the eccentricity enables us to derive a critical inclination,

$$\cos i_0 = \sqrt{\frac{6R_{\rm tot}}{5(1-e_0^2)\alpha R_{\rm H}}}. \tag{9}$$

Significant LK oscillations (of an initially circular orbit) occur for inclinations in the range 40° $\lesssim i_0 \lesssim$ 140°, where large inclinations correspond to small eccentricities and vice versa.

In the presence of oblate bodies and under the assumption of maximal initial inclination of $\cos i_0 = 90°$, the maximal eccentricity could be estimated by the implicit expression (Liu et al. 2015; Grishin et al. 2020)

$$\frac{\epsilon_{\rm rot}}{3}\left(\frac{1}{(1-e_{\rm max}^2)^{3/2}}-1\right) = \frac{9}{8}e_{\rm max}^2, \tag{10}$$

which could be approximated by $e_{\rm max} \approx 1 - \frac{2}{9}\epsilon_{\rm rot}^{3/2}$ under the assumption of large eccentricity ($e_{\rm max}^2 \approx 1$) and weak effect of the rotation term ($\epsilon_{\rm rot} \ll 1$).

Henceforth, the minimal α value for collision (in the secular regime) is given by

$$\alpha_{\rm coll} = \left(\frac{2R_{\rm H}^3\alpha_{\rm L}^{10}}{81R_{\rm tot}^3}\right)^{1/7} \approx 0.07. \tag{11}$$

The regime that corresponds to ’pure’ secular collision is coloured with grey – secular collision.

The transition between secular and quasi-secular regime can be derived from the strength of the quasi-secular corrections over the strength of the ’standard’ LK ones, here taken to quadrupole order. The strength of the perturbations from single averaging – averaging over the inner orbit only – is given by (Luo et al. 2016; Grishin et al. 2018)

$$\epsilon_{\rm SA} = \frac{P_{\rm out}}{2\pi\tau_{\rm s}} = \left(\frac{a_{\rm in}}{a_{\rm out}(1-e_{\rm out}^2)}\right)^{3/2}\frac{M_\odot}{(m_{\rm tot}m_{\rm in})^{1/2}} \approx$$
$$\approx \frac{\alpha^{3/2}}{\sqrt{3}(1+e_{\rm out})^{3/2}}, \tag{12}$$

where $P_{\rm out}$ is the period of the outer binary and $m_{\rm tot}$ is the total mass of the triple system. Due to large eccentricity of Pluto’s orbit around the Sun, the effective quasi-secular corrections are encapsulated by $\tilde{\epsilon}_{\rm SA} = \epsilon_{\rm SA}(1+2\sqrt{2}e_{\rm out}/3)$. For $\cos i_0\sqrt{1-e^2} \lesssim 9\tilde{\epsilon}_{\rm SA}/8$, the fluctuations in the angular momentum are larger than its initial value. The evolution experiences orbital flips, and the eccentricity is unbound, which hallmarks the transition to the non-secular regime, given by

$$\alpha_t = 3^{1/3}\left[\frac{128}{135}\frac{(1+e_{\rm out})^3}{\left(1+\frac{2\sqrt{2}}{3}e_{\rm out}\right)^2}\left(\frac{M_\odot}{m_{\rm in}}\right)^{1/3}\frac{R_{\rm tot}}{a_{\rm out}}\right]^{1/4} \approx 0.17. \tag{13}$$

The regime that describes non-secular collisions is coloured with blue – non-secular collisions.

4 NUMERICAL RESULTS

In order to verify the analytic result and simulate the dynamics of the progenitors of Pluto–Charon and the Sun, we used the publicly available *N*-body code `REBOUND` (Rein & Liu 2012). We use `IAS15`, a fast, adaptive, high-order integrator for gravitational dynamics, accurate to machine precision over a billion orbits (Rein & Spiegel 2015). We integrate different sets of initial conditions; in all of them, we set outer semi-major $a_{\rm out} = 39.482$AU, outer eccentricity $e_{\rm out} = 0.2488$, and changing mutual inclinations and outer semi-major axes. Canup (2005) suggests a range of possible parameters for Pluto and Charon progenitors; we use the average masses from the constrained mass ranges, i.e. $m_{\rm Pluto} \approx 1.57 \times 10^{24}$g and $m_{\rm Charon} \approx 1.35 \times 10^{24}$g.

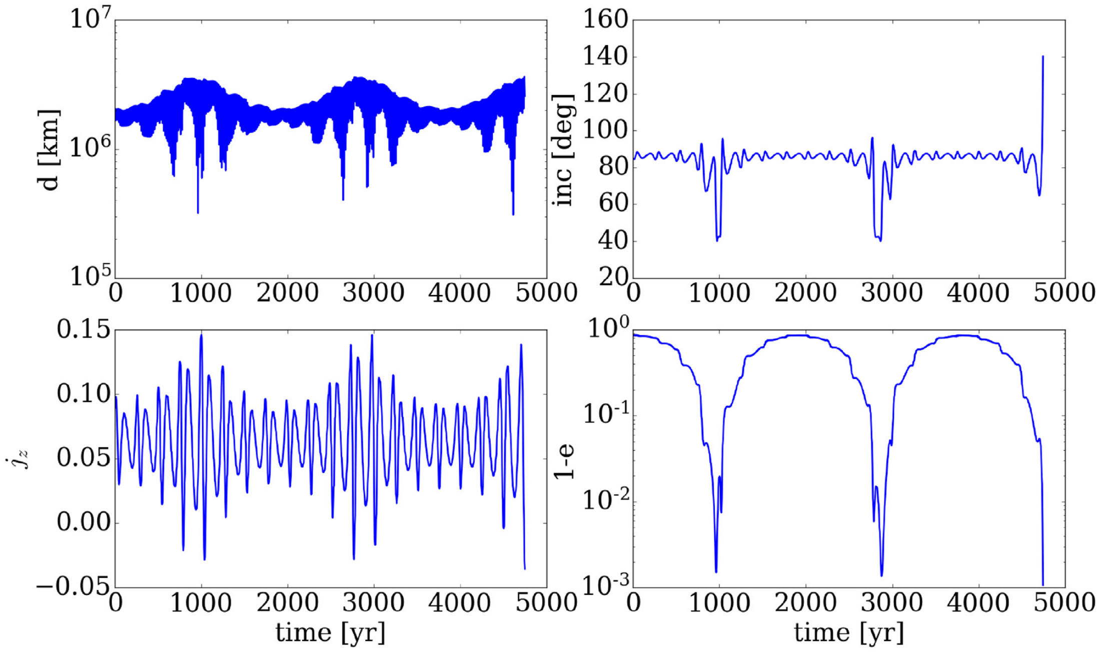

Figure 2. Evolution of the system for the following initial conditions: scaled initial inner semi-major axis $\alpha = 0.3$, initial inner eccentricity $e_{\mathrm{in}} = 0.15$, initial mutual inclination $i_{\mathrm{mutual}} = 85^{\circ}$, argument of periapsis $\omega_{\mathrm{out}} = 0$, longitude of the ascending node $\Omega_{\mathrm{out}} = \pi/4$, and mean anomalies $\mathcal{M}_1 = 0$ and $\mathcal{M}_2 = -\pi/4$.

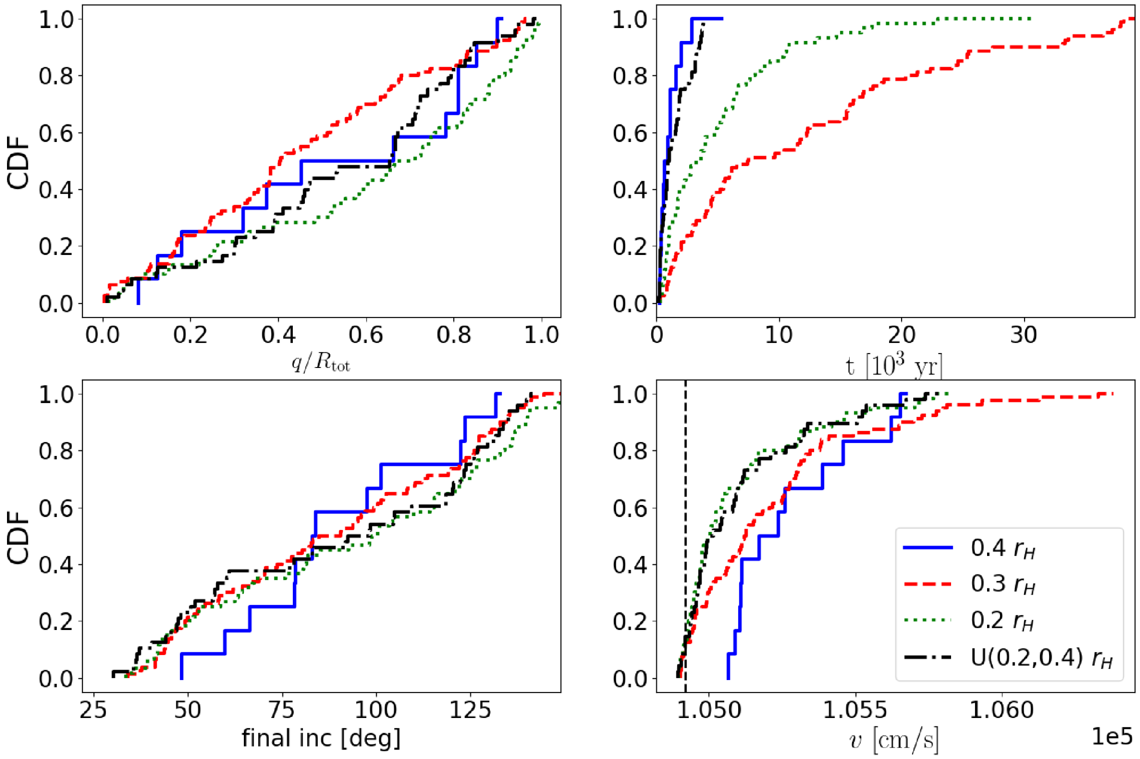

Figure 3. Cumulative distributions of the impact properties. Upper left: Normalized pericentre q/R_{tot}; Upper right: Time of collision; Lower left: Final inclination at impact; Lower right: Velocity at impact. The vertical dashed line is the escape velocity.

The total radius of the object is taken to be $R_{\mathrm{tot}} = R_{\mathrm{Pluto}} + R_{\mathrm{Charon}} \approx$ 1794km .

Fig. 2 presents an example of the behaviour in the quasi-secular regime. Here, we study the evolution of the system with initial normalized inner separation of $\alpha = 0.3$ in which case, the evolution is in the non-secular regime. The inclination flips, the eccentricity is excited significantly during the evolution, and the inclination reaches a high value at the time of the collision. The angular momentum in the z direction, defined by $j_z = \cos i\sqrt{1-e^2}$, oscillates, and its envelope structure can be derived analytically (Luo et al. 2016). Collision occurs after ~ 4500 yr.

In order to study the statistics of the collisional behaviour of the Pluto–Charon binary system, we follow the approach of Grishin et al. (2020) and study a sample of cases where α is either 0.2, 0.3, 0.4, or chosen randomly from a uniform distribution in the range [0.2, 0.4] and run the simulation for 5×10^4 yr. In Fig. 3, we show the cumulative distribution function of collision parameters. Fig. 3 presents the cumulative distribution function (CDF) of the collision parameters: $q/R_{\rm tot}$ where $q = a(1-e)$ is the closest approach, time in units of thousand years, final inclination in degrees, and velocity in units of cm/s.

The successful collisions fractions after 5×10^4 yr are: for $\alpha = 0.2$ is ≈ 30 per cent, for $\alpha = 0.3$ is ≈ 40 per cent, for $\alpha = 0.4$ is ≈ 5 per cent, and for the uniformly sampled $\alpha \in [0.2, 0.4]$ is ≈ 20 per cent.

The lower left-hand panel of Fig. 3 show the consistency with the uniform distribution in $\cos i_0$ where i_0 ranges between 40° and 140°. As can be seen from the upper right-hand panel, the typical collision time-scale is a few thousands of years and the collision velocity is close to the escape velocity – lower right-hand panel, which is given by $v_{\rm esc} = \sqrt{2G(m_{\rm Pluto} + m_{\rm Charon})/R_{\rm tot}} \approx 10.49 \times 10^4 {\rm cm/s}$, where $R_{\rm tot}$ is the minimal possible distance between the binary companions Pluto and Charon, i.e. the sum of their radii. As expected, since all the sampled values of α are predicted analytically to be in the regime of quasi-secular/non-secular or no collision at all, the behaviour of the system is chaotic and spans over a wide range in the parameter space, as can be seen in the upper left-hand panel. Collisions in large values of α, i.e. $\alpha \geq 0.4$, become more rare, since the system is less stable (Grishin et al. 2017).

5 CAVEATS

In wider and less hierarchical systems, the external perturbations by the Sun could lead to instabilities and consequently lead to physical collisions or the escape of objects in the system through a chaotic evolution. While collisions are the major consequence discussed in this paper, escape is an unwanted byproduct. Very wide systems can also become unstable due to the perturbations flyby encounters with other KBOs (Heggie's law, Heggie 1975), but here we neglect such encounters and analyse only isolated systems.

Analysis of the simulated systems shows signatures of instability around $\alpha = 0.4$, as can be seen in Fig. 3. Close to and beyond this limit, the number of collisions decreases significantly. We find (see Fig. 3) that ≈ 94 per cent of the systems with $\alpha = 0.4$ break up during the simulation. For initially circular systems, Grishin et al. (2017) showed that the stability of systems could be sustained for even slightly wider systems; the somewhat lower limit we observed results can rise from considering systems with significant eccentricities.

Another caveat arises from uncertainties in the mass estimates, resulting from uncertainties in the chemical differentiation of the progenitors (see a detailed discussion in Stern et al. (2018a) and references therein), which we didn't take into consideration in our paper. The chosen masses are taken to be the average masses of the range given by Canup (2005), which might lead to some small uncertainties in our results. Furthermore, some larger uncertainties arise from the unknown initial separations between the progenitors of Pluto and Charon. We sampled some possible separations in order to explore a range of possibilities, but different initial separations could change the time-scales and evolution of the binary.

We treated the Pluto–Charon – Sun system in isolation. In principle, other planets, in particular, Neptune, might affect the evolution of the system and add non-trivial corrections.

6 SUMMARY

In this paper, we proposed a novel formation channel for the origin of the Pluto–Charon system from a wide-binary, via secular and quasi-secular evolution that might lead to the collision between the components of the originally wide-binary progenitor.

We use analytic criteria to set the different regimes of evolution: secular (and quasi-secular), non-secular, and non-collisional, and made use of N-body simulations to verify and study these regimes. Our results indicate that collisions, consistent with the impact required to explain Pluto–Charon properties, are a natural byproduct of secular and quasi-secular evolution of wide-binary progenitors. The required impact parameters can be reproduced from a wide range of initial conditions, suggesting this scenario as a robust formation channel for the origins of the Pluto–Charon system, and alleviating potential fine-tuned conditions required for the currently suggested origin from a random low-velocity collision between two of the most massive and relatively rare KBOs in the Kuiper belt. The model could be used in the future for similar systems in the Solar system and beyond and shed light on the formation process of highly inclined contact or tidally synchronized binaries where regular LK oscillations that some of them could not have produced such configurations include (139775) 2001 QG$_{298}$ (Lacerda 2011) and potentially 67P/Churyumov–Gerasimenko (Marsden 1969), the latter one might be produced by a regular LK mechanism, since its obliquity is high but not extreme. Very large obliquities are a signature of quasi-secular evolution, but even large but not extreme obliquities might belong to the quasi-secular regime.

ACKNOWLEDGEMENTS

HBP acknowledges support from the MINERVA centre for 'Life under extreme planetary conditions' and the Kingsley fellowship at Caltech.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

Antognini J. M. O., 2015, MNRAS, 452, 3610
Antonini F., Perets H. B., 2012, ApJ, 757, 27
Canup R. M., 2011, AJ, 141, 35
Canup R. M., 2005, Science, 307, 546
Christy J. W., Harrington R. S., 1978, AJ, 83, 1005
Desch S. J., 2015, Icarus, 246, 37
Fraser W. C. et al., 2017, Nature Astron., 1, 0088
Goldreich P., Lithwick Y., Sari R., 2002, Nature, 420, 643
Greenstreet S., Gladman B., McKinnon W. B., 2015, Icarus, 258, 267
Grishin E., Perets H. B., 2016, ApJ, 820, 106
Grishin E., Perets H. B., Zenati Y., Michaely E., 2017, MNRAS, 466, 276
Grishin E., Perets H. B., Fragione G., 2018, MNRAS, 481, 4907
Grishin E., Malamud U., Perets H. B., Wand el O., Schaefer C. M., 2020, Nature, 580, 463
Harrington R. S., 1968, AJ, 73, 190

Heggie D. C., 1975, MNRAS, 173, 729
Innanen K. A., Zheng J. Q., Mikkola S., Valtonen M. J., 1997, AJ, 113, 1915
Johansen A., Mac Low M.-M., Lacerda P., Bizzarro M., 2015, Sci. Adv., 1, 1500109
Kenyon S. J., Bromley B. C., 2014, AJ, 147, 8
Kinoshita H., Nakai H., 1999, Celest. Mech. Dyn. Astron., 75, 125
Kozai Y., 1962, AJ, 67, 591
Lacerda P., 2011, AJ, 142, 90
Leinhardt Z. M., Stewart S. T., 2012, ApJ, 745, 79
Lidov M. L., 1962, Planet. Space Sci., 9, 719
Liu B., Muñoz D. J., Lai D., 2015, MNRAS, 447, 747
Luo L., Katz B., Dong S., 2016, MNRAS, 458, 3060
Lyra W., Youdin A. N., Johansen A., 2020, preprint (arXiv:2003.00670)
Marsden B. G., 1969, IAU Circ., 2191
McKinnon W. B., 1984, Nature, 311, 355
McKinnon W. B., 1989, ApJ, 344, L41
McKinnon W. B. et al., 2017, Icarus, 287, 2
McKinnon W. B. et al., 2020, Science, 367, aay6620
Michaely E., Perets H. B., Grishin E., 2017, ApJ, 836, 27
Murray C. D., Dermott S. F., 1999, Solar system dynamics. Cambridge University Press, Cambridge, UK
Naoz S., 2016, ARA&A, 54, 441
Naoz S., Perets H. B., Ragozzine D., 2010, ApJ, 719, 1775
Nesvorný D., Youdin A. N., Richardson D. C., 2010, AJ, 140, 785
Nimmo F. et al., 2017, Icarus, 287, 12
Noll K. S., Grundy W. M., Chiang E. I., Margot J. L., Kern S. D., 2008, in Barucci M. A., Boehnhardt H., Cruikshank D. P., Morbidelli A., Dotson R., eds, Binaries in the Kuiper Belt, University of Arizona Press, Tucson, p. 345
Perets H. B., Naoz S., 2009, ApJ, 699, L17
Poincaré H., 1892, Les méthodes nouvelles de la mécanique céleste
Porter S. B., Grundy W. M., 2012, Icarus, 220, 947
Rein H., Liu S. F., 2012, A&A, 537, A128
Rein H., Spiegel D. S., 2015, MNRAS, 446, 1424
Sekine Y., Genda H., Kamata S., Funatsu T., 2017, Nature Astron., 1, 0031
Stern S. A. et al., 2015, Science, 350, aad1815
Stern S. A., Grundy W. M., McKinnon W. B., Weaver H. A., Young L. A., 2018a, ARA&A, 56, 357
Stern S. A., Weaver H. A., Spencer J. R., Elliott H. A., 2018b, Space Sci. Rev., 214, 77
Stern S. A. et al., 2019, Science, 364, aaw9771
Tombaugh C. W., 1946, Leaflet Astron. Soc. Pac., 5, 73
Tremaine S., Touma J., Namouni F., 2009, AJ, 137, 3706
Von Zeipel H., 1916, Astron. Fys., 11

This paper has been typeset from a TEX/LATEX file prepared by the author.

Chapter 4

Inflated eccentric migration of hot & warm Jupiters

Based on Rozner et al. (2022); Glanz et al. (2022)

By the end of core accretion, gas giants are born inflated with radii that can reach 10 R_J, and then cool down and contract rapidly to a typical size of $\sim 4\ R_J$ (Guillot et al., 1996; Ginzburg and Chiang, 2019). Then, the gas giant goes over a slower phase of contraction within Kelvin-Helmholtz timescales (tens on Myrs), until it reaches asymptotically its (effective) final size. The contraction process depends on different parameters, such as the mass of the planet, and the external heating sources injected into the planet by irradiation/tidal heating. and also dust composition (e.g. (Ginzburg and Sari, 2015, 2017; Ginzburg and Chiang, 2019)).

While this process is well known, the thermal evolution of gas giants is usually overlooked when eccentric tidal migration is studied, although they could take place on similar timescales. Since tides depend strongly on the radius of the planet, integrating the thermal evolution which is connected directly to the radius evolution changes significantly the population of formed hot and warm Jupiters (HJs & WJs correspondingly). In this project, we studied the coupled thermal-evolution, using two methods: semi-analytical and numerical. In the semi-analytical model we characetrized the gas giant using its equation of state and connected between the radius evolution and the energy evolution, then we solved the coupled differential equations that couple between the tidal and thermal evolutions.

The tidal evolution equation for weak tides are given by (Hut, 1981; Hamers and Tremaine, 2017)

$$\frac{da}{dt} = -21 k_{\mathrm{AM}} n^2 \tau_p \frac{M_\star}{M_p} \left(\frac{R_p}{a}\right)^5 a e^2 \frac{f(e)}{(1-e^2)^{15/2}}, \tag{4.1}$$

$$\frac{de}{dt} = -\frac{21}{2} k_{\mathrm{AM}} n^2 \tau_p \frac{M_\star}{M_p} \left(\frac{R_p}{a}\right)^5 e \frac{f(e)}{(1-e^2)^{13/2}} \tag{4.2}$$

where $M_\star$ is the mass of the host star, $M_p,\ R_p,\ e,\ a$ and n are the mass, radius, orbital eccentricity, orbital semimajor axis and mean motion of the gas-giant correspondingly, $\tau_p = 0.66$ sec is the planetary tidal lag time and $k_{\rm AM} = 0.25$ is the planetary apsidal motion constant ($k_{\rm AM}$ and τ_p are taken from Hamers and Tremaine, 2017). $f(e)$ is defined by

$$f(e) = \frac{1 + \frac{45}{14}e^2 + 8e^4 + \frac{685}{224}e^6 + \frac{255}{448}e^8 + \frac{25}{1792}e^{10}}{1 + 3e^2 + \frac{3}{8}e^4} \tag{4.3}$$

The energy extracted per period and hence the migration and circularization rate, scales as R_p^5. Consequently, the migration timescales of initially inflated gas-giants, where the planetary radius, $R_p > R_J$, are shortened relative to migration of non-inflated gas-giants, i.e. with a constant radius of R_J.

The radius evolution, derived by using the Virial theorem that states a relation between the total and potential energies $E = -(3\gamma - 4)U/(3\gamma - 3)$, such that $U \propto 3GM_p^2/R(5 - \tilde{n})$ where $\tilde{n} = 1/(\gamma - 1)$ is the polytropic index, γ is the heat capacity ratio, taken as $\gamma = 5/3$ and the proportion constant is determined by numerical gauge. The change in the planet luminosity is mostly determined by the change in the thermal energy of ions, which are not degenerate, and their equation of state is given by the ideal gas equation $E = (M_p/\mu)k_B T_c$ where μ is the molecular weight, taken as the proton mass and T_c is central temperature of the planet. Following this relation, we derive the following equation for the radius change due to heating and cooling,

$$\frac{dR_p}{dt} \propto \frac{5\gamma - 6}{3\gamma - 4} \frac{R_p^2}{GM_p^2} \left(L_{\rm ext} - L_{\rm cool}\right) \tag{4.4}$$

Then, in the numerical method, we used a `MESA` (Paxton et al., 2011) model that evolve the gas giants thermally, and then after every thermal step we solved the very same tidal equation presented in the semi-analytical model, coupled together with the thermal evolution via the `AMUSE` framework (Portegies Zwart et al., 2009).

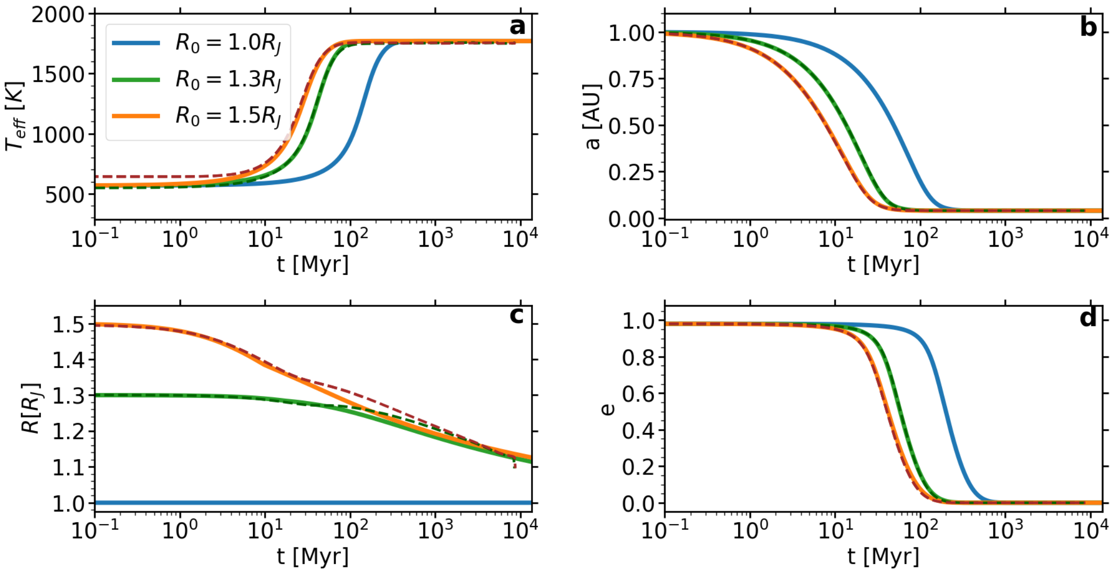

Figure 4.1: The thermal and orbital evolution of a HJ-progenitor migrating due to weak tides, with photosphere heating due to irradiation. The thermal (**a,c**- effective temperature and radius) and orbital (**b,d** - semimajor axis and eccentricity) evolution of 1 M_J HJ progenitors with different initial radii (1 R_J – blue, 1.3 R_J – orange and 1.5 R_J – green). Tides are modeled through a weak-tide model, irradiation of the planetary outer layer is included, but with no additional efficient heat conduction to the core. Shown are gas-giants with an initial semimajor axis of 1 AU and initial eccentricity of 0.98. The solid lines correspond to the semi-analytical calculation and the dashed to the numerical.

In Fig. 4.1 we present an example for the coupled thermal-dynamical evolution of a hot Jupiter. The effective temperature of the gas giant is roughly flat until it reaches the final temperature, dictated by the irradiation from the host star. The radius contracts within a relatively long timescale, such that for the whole migration process the gas giant is inflated to some degree. On the right panles, we present the dynamical evolution of the hot Jupiter candidate and show that the migration timescale shortens significantly, as expected. It could be seen that the analytical and numerical models are within well agreement.

We then proceeded with the semi-analytical approach only, which is computationally cheaper, and was proved to agree with the results of the numerical simulations, to gain a better understanding of the behaviour of hot and warm Jupiters populations. The total fraction of hot/warm Jupiters was calculated using the following equation

$$\mathcal{F}_{\mathrm{HJ/WJ}} = f_J \times f_{2J} \times f_{\mathrm{unstable}} \times f_{ecc,J} \times f_{\mathrm{HJ/WJ}} \tag{4.5}$$

where f_J is the fraction of stellar systems hosting gas-giants, f_{2J} is the probability to find at least two gas-giants (in most of the cases, in order to scatter a gas-giant into a highly eccentric orbit leading to eccentric migration, another planet as massive as Jupiter is needed), f_{unstable} is the fraction of unstable systems, in which planet-planet scattering is likely to occur and $f_{\mathrm{ecc,J}}$ is the fraction of sufficiently eccentric gas-giants that would evolve through eccentric migration to become HJs&WJs.

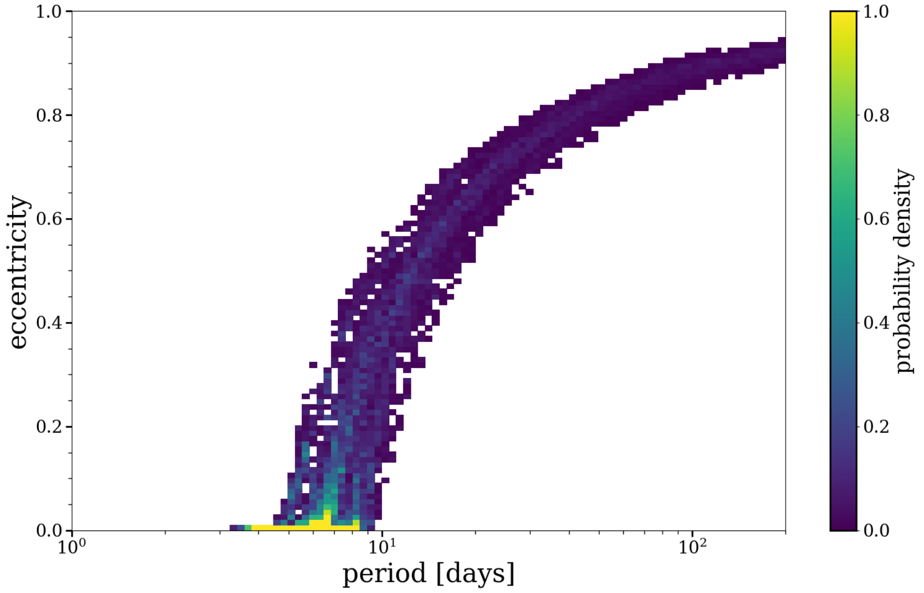

Figure 4.2: 2-dimensional histogram of the eccentricity and orbital period of hot and warm Jupiters after a Hubble time, as obtained from the population-synthesis models initialized with $\sigma_1 R_3 a_1 df_{0.1}$. The probability density is normalized according to the total fraction of successful formation of HJs and WJs among all initial conditions sampled.

In Fig. 4.2, we present the 2d period-eccentricity probability density. It could be seen that the majority of hot Jupiters formed from our model are circular, where the warm Jupiters have a non-zero eccentricity. This plot could be thought as the 'HR diagram' of hot Jupiters evolution – they start their migration as cold Jupiters, with large periods and high eccentricity, and the tidal dissipation leads to circularization and period shrinkage. The empty regions are determined by the tidal disruption and conservation of angular momentum during the evolution.

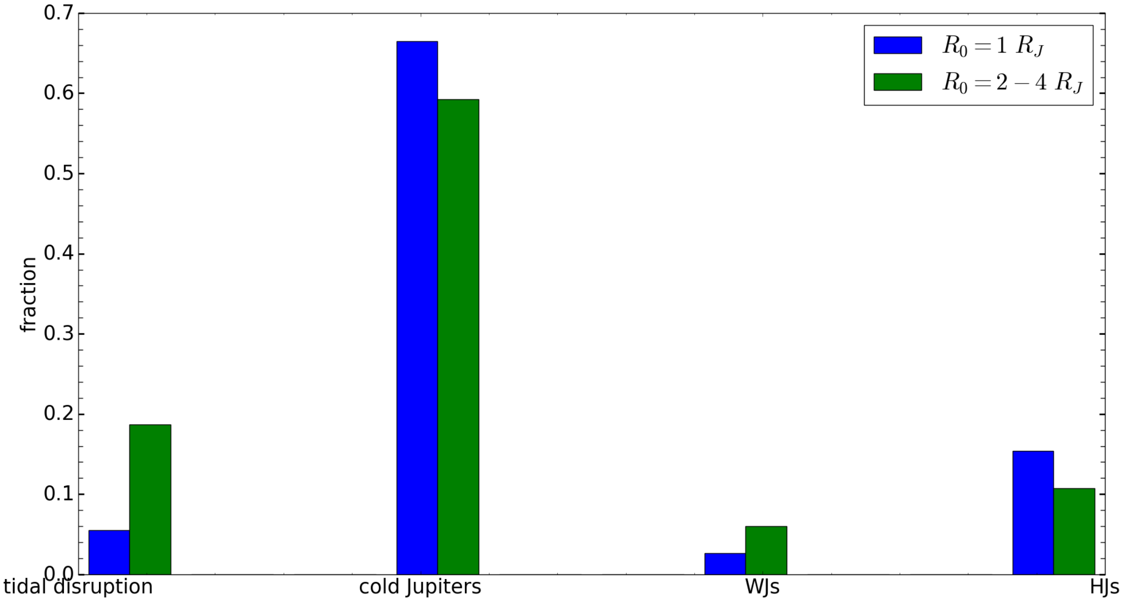

Figure 4.3: A diagram of the final fractions concluded from the Monte-Carlo simulation and a comparison inflated ($\sigma_1 R_3 a_1 df_{0.1}$) and non-inflated ($\sigma_1 R_1 a_1 df_{0.1}$) initial radii. The fractions could be summed over to 1 – we present just the results of the simulation without further normalization. The results are based on 10^4 runs of the semi-analytical model per each case.

In Fig. 4.3, we present an histogram of the different population of Jupiters, comparing between initially-inflated and non-inflated gas giants. It could be seen that the inflated ones are more dynamically active, i.e. we produce less cold Jupiters. However, while the fraction of warm Jupiters increases, the fraction of hot Jupiters decreases, as inflated gas-giants are more prone to tidal disruption.

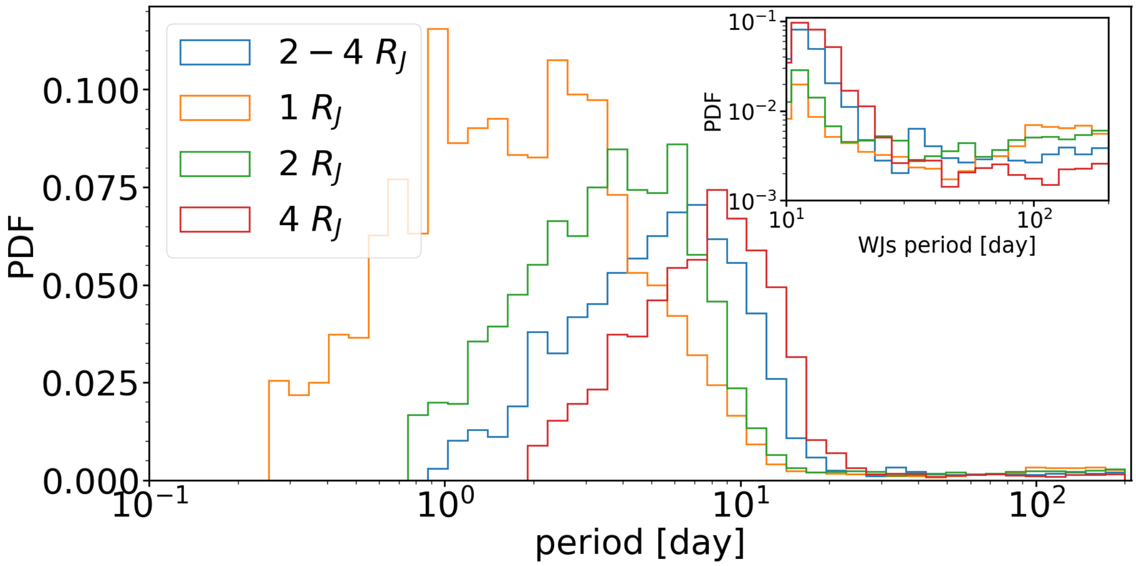

Figure 4.4: The period distribution as found by the Monte-Carlo simulation, based on the semi-analytical model, for different initial radius distribution. The rest of the parameters are drawn according to $\sigma_1 R_3 a_1 df_{0.1}$. In the inset figure we introduce the probability distribution function of WJs only (with the same color code)

.

Initial inflation affects also on the final distribution of orbital parameters. As can be seen in Fig. 4.4, the population of initially-inflated gas giant tend to reproduce a population of hot Jupiters that peak on longer periods than the non-inflated ones.

To summarize, in this project we investigated the coupling of thermal and dynamical evolution of hot and warm Jupiters, during their tidal migration. We found that the effect is significant, and increase substantially the disruption rates of hot Jupiter candidates, together with the rate of warm Jupiters. The formalism we introduced here is general and could be used in principle to treat other physical phenomena.

The Astrophysical Journal, 931:10 (15pp), 2022 May 20 https://doi.org/10.3847/1538-4357/ac6808

Inflated Eccentric Migration of Evolving Gas Giants I – Accelerated Formation and Destruction of Hot and Warm Jupiters

Mor Rozner[1], Hila Glanz[1], Hagai B. Perets[1], and Evgeni Grishin[1,2]
[1] Technion—Israel Institute of Technology, Haifa, 3200002, Israel; morozner@campus.technion.ac.il
[2] School of Physics and Astronomy, Monash University, Clayton 3800, VIC, Australia

Received 2021 November 28; revised 2022 April 4; accepted 2022 April 17; published 2022 May 19

Abstract

Hot and warm Jupiters (HJs and WJs, correspondingly) are gas giants orbiting their host stars at very short orbital periods ($P_{\mathrm{HJ}} < 10$ days; $10 < P_{\mathrm{WJ}} < 200$ days). HJs and a significant fraction of WJs are thought to have migrated from initially farther-out birth locations. While such migration processes have been extensively studied, the thermal evolution of gas giants and its coupling with migration processes are usually overlooked. In particular, gas giants end their core accretion phase with large radii, then contract slowly to their final radii. Moreover, intensive heating can slow the contraction at various evolutionary stages. The initial large inflated radii lead to faster tidal migration, due to the strong dependence of tides on the radius. Here, we explore this accelerated migration channel, which we term inflated eccentric migration, using a semi-analytical, self-consistent model of the thermal–dynamical evolution of the migrating gas giants, later validated by our numerical model (see the companion paper, paper II). We demonstrate our model for specific examples and carry out a population synthesis study. Our results provide a general picture of the properties of the formed HJs and WJs via inflated migration, and their dependence on the initial parameters/distributions. We show that the tidal migration of gas giants could occur much more rapidly then previously thought, and could lead to the accelerated destruction and formation of HJs and an enhanced formation rate for WJs. Accounting for the coupled thermal–dynamical evolution is therefore critical to understanding the formation of HJs/WJs, and the evolution and final properties of the population, and it plays a key role in their migration processes.

Unified Astronomy Thesaurus concepts: Hot Jupiters (753); Exoplanets (498); Exoplanet formation (492); Exoplanet dynamics (490); Exoplanet tides (497); Exoplanet migration (2205); Exoplanet evolution (491); Exoplanet structure (495); Exoplanet astronomy (486); Extrasolar gaseous planets (2172)

1. Introduction

The formation of gas-giant planets is thought to occur either through the gradual bottom-up growth of planetary embryos, followed by rapid gas accretion (core accretion), or through the direct formation of gas giants, following gravitational instability (Mizuno 1980; Boss 1997; Armitage 2010). However, the observed proximity of hot and warm Jupiters (HJs and WJs, correspondingly) to their host star sets severe constraints on both these mechanisms. The high temperatures, high velocities, low disk masses, and solids that characterize these environments constrict the predicted formation rates from the in situ channel (Bodenheimer et al. 2000; Rafikov 2005). Though in situ formation may still potentially explain a non-negligible fraction of the population of WJs residing farther out (Huang et al. 2016; Anderson et al. 2020), the formation of HJs and eccentric WJs is not likely to occur in situ. These have therefore been suggested to form at larger separations from the star, before migrating inward to their current locations (see Dawson & Johnson 2018 for a review).

The currently suggested migration channels can generally be divided between disk migration channels and high-eccentricity tidal migration channels (Dawson & Johnson 2018). While such migration scenarios could potentially explain the origins of HJs and WJs, detailed studies of these migration scenarios have struggled to reproduce the rates and properties of the formed HJs and WJs (Dawson & Johnson 2018; Zhu & Dong 2021). However, such models usually did not self-consistently account for role of the thermal evolution of the gas giants, from their initially hot inflated state to their later contraction (and possible heating), during their eccentric migration.

In high-eccentricity migration mechanisms, the semimajor axis shrinks due to energy dissipation induced by tidal forces. This process requires initial high eccentricity, as tidal coupling strongly depends on the distance from the star, which becomes small at pericenter approaches of highly eccentric orbits. The planet can initially be excited to such high eccentricities by planet–planet scattering (Rasio & Ford 1996; Chatterjee et al. 2008; Jurić & Tremaine 2008), through the von-Ziepel–Lidov–Kozai (ZLK) mechanism (von Zeipel 1910; Kozai 1962; Lidov 1962) coupled to tidal evolution (e.g., Kiseleva et al. 1998; Wu & Murray 2003; Fabrycky & Tremaine 2007; Naoz et al. 2011; Petrovich 2015a; Grishin et al. 2018; Vick et al. 2019), or coupled multiplanet secular evolution (e.g., Wu & Lithwick 2011; Beaugé & Nesvorný 2012; Hamers et al. 2017), or their combination. In this paper, we focus on excitation by planet–planet scattering, and we will discuss the secular channel in a future study. The current tidal migration models show that the typical migration timescales could be long, and the production rates of HJs and WJs (or both) are too low in comparison with observations (particularly for WJs). However, as we discuss below, the initial hot inflated state of a gas giant at birth, and its later radiative and tidal heating, could change this picture.

Core accretion formation of gas giants proceeds through the runaway accretion of gas from the protoplanetary disk typically onto a few Earth-mass solid cores (Perri & Cameron 1974; Bodenheimer & Pollack 1986). At the end of the core accretion process, the recently formed gas giants reach large radii, possibly up to $10\,R_J$ for dusty planets (Ginzburg & Chiang 2019). After a very rapid contraction phase, when the planet's radius contracts to $4\,R_J$ (Guillot et al. 1996), the contraction slows down, and takes place at typical Kelvin–Helmholtz timescales (tens of Myrs), as the planet contracts asymptotically to reach its (effective) final size. The contraction depends on the mass and, as we discuss in depth, on the externally injected energies, due to tides and/or irradiation, with only weak dependence on the exact initial conditions at birth (leaving aside the uncertainties/degeneracy in the models of the early evolution of gas giants; e.g., Marley et al. 2007). It should be noted that dustier planets could lead to even slower contraction (Ginzburg & Chiang 2019).

Typically, eccentric migration scenarios assume that the radii of the migrating gas giants are constant throughout their evolution, and are taken to be their asymptotic radii at late times, $\sim 1R_J$, neglecting the initial inflated radii at birth and the later contraction. Since tidal forces strongly depend on the radius of the affected object, initially inflated planets should migrate faster than planets that have already contracted to their final radii. The dynamical and thermal evolutions cannot be decoupled, due to their mutual strong dependence on each other. To date, these issues have only been partially studied by Wu et al. (2007), Miller et al. (2009), and Petrovich (2015a), and even these studies have only addressed limited aspects of coupled evolution models, using simplified approaches.

In this paper and Glanz et al. (2021; hereafter, paper II), we study the coupled dynamical–thermal evolutions of HJs and WJs, which evolve via inflated eccentric migration, after the end of the core accretion phase, and compare them with corresponding cases of initially noninflated planets and the inefficient heat conduction that is typically studied in the literature. We present a semi-analytical model, described in this paper, and compare it with a numerical model, discussed in paper II. Both approaches couple the dynamical and tidal evolutions of the planets with their thermal evolutions. The former makes use of a simple analytic approach for the thermal evolution and its effect on the radius evolution, and the latter follows the evolution through a numerical model, using the stellar and planetary evolution code `MESA` (Paxton et al. 2011, 2013), coupled with the `AMUSE` framework (Portegies Zwart et al. 2009). We find excellent agreement between the model results, in terms of the overall evolutions and properties of the systems, as we discuss below and in paper II. Consequently, and due to the simplicity of the semi-analytical approach relative to the computationally costly numerical one, we discuss the results from the numerical evolution only for specific cases, and show comparisons to the analytic model (see below and in paper II). The efficient semi-analytical approach allows us to study both the large parameter space of the possible initial conditions and the formed population of HJs/WJs, which are inaccessible to the numerical approach, since it is computationally expensive and has limitations in terms of long timescales/small radii (see paper II for further details).

In the following, we describe our semi-analytical approach, and present the various aspects of inflated eccentric migration and its outcomes. We then consider the overall distribution of the initial conditions, the resulting rates and properties of the formed HJs and WJs, and their dependence on these initial distributions.

In order to present our approach, we first discuss the conditions for inflated eccentric migration arising from planet–planet scattering. We then discuss tidal migration in general in Section 3, and elaborate the weak and dynamical tidal models and their use in our models. In Section 4, we describe the semi-analytical model for inflated migration and demonstrate the formation of HJs and WJs in specific examples. In Section 5, we demonstrate the use of the semi-analytical model in several examples, and compare to the numerical model. In Section 6, we discuss the population synthesis of the formation of HJs and WJs. In Section 7, we present the results of the population synthesis, the choice of the parameters, and the role played by them. In Section 8, we discuss our findings and the further implications. In Section 9, we summarize our results.

2. Planet–Planet Scattering

In this study, we focus on planetary systems in which the initial conditions for high-eccentricity migration are dictated by planet–planet scattering. There are other channels for eccentricity excitation, such as secular Lidov–Kozai (LK) evolution in triple systems (e.g., Wu & Murray 2003; Naoz et al. 2011; Petrovich 2015a, 2015b; Grishin et al. 2018; Vick & Lai 2018) and secular resonances (e.g. Wu & Lithwick 2011; Hamers et al. 2017). Such processes typically operate on longer timescales, lead to intermittent/quasiperiodic high eccentricities, and are also sensitive to precession induced by tidal interactions, which can partially quench the level of eccentricity excitation. In this work, we consider only planet–planet scattering, but inflated eccentric migration is likely to be important for those other channels, such as the ones involve secular evolution (see Wu et al. 2007; Petrovich 2015a). It should be noted that in general the timescales of these channels could be longer, suggesting a priori that the effect of the initial inflation may be smaller. However, initially inflated planets will still leave a signature on the dynamical evolution. First, a larger fraction of the initial planets will be prone to tidal disruption. Moreover, the final distribution of the semimajor axes is expected to change accordingly (Petrovich 2015a). Another potential effect of initial inflation is a tidal quenching of the LK mechanism, which passes with the contraction of the planet. Although several studies have been done on the secular channel, to our knowledge the coupling there between the thermal and dynamical evolution is not self-consistent; however, self-consistent modeling could easily be implied in our semi-analytical model. This channel is out of the scope of this paper, and is left for follow-up studies.

In multiplanet systems, mutual gravitational interactions between the planets perturb their orbits and may destabilize the system, leading to the ejection of planets, mutual collisions, collisions with the star, and more general excitation of the planets' eccentricities and inclinations (Rasio & Ford 1996; Weidenschilling & Marzari 1996; Chatterjee et al. 2008). The gravitational encounters could also involve tidal circularization during the scattering process, which might enhance the fractions of shorter-period and high-eccentricity gas giants (Nagasawa et al. 2008). More generally, it has been found that strong planet–planet scatterings eventually give rise to a Rayleigh distribution of the eccentricities of the surviving planets (e.g., Jurić & Tremaine 2008). It should be noted that

Table 1

Normalized Formation Fractions of HJs and WJs—$f_{\rm HJ}$ and $f_{\rm WJ}$, Correspondingly—as Derived from the Monte Carlo Simulation, Based on the Semi-analytical Model

Model	$f_{\rm HJ}$	$f_{\rm WJ}$	Model	$f_{\rm HJ}$	$f_{\rm WJ}$
$\sigma_1 R_1 a_1 w$	1%	0.3%	$\sigma_1 R_1 a_1 df_{0.1}$	2%	0.4%
$\sigma_1 R_2 a_1 w$	0.5%	0.4%	$\sigma_1 R_3 a_1 df_{0.1}$	1.6%	1%
$\sigma_1 R_3 a_1 w$	0.4%	0.4%	$\sigma_1 R_1 a_1 df_{0.01}$	1.5%	0.2%
$\sigma_1 R_4 a_1 w$	0.2%	0.4%	$\sigma_1 R_2 a_1 df_{0.01}$	1.2%	0.3%
$\sigma_2 R_1 a_1 w$	0.5%	0.2%	$\sigma_1 R_3 a_1 df_{0.01}$	1%	0.4%
$\sigma_2 R_2 a_1 w$	0.3%	0.2%	$\sigma_1 R_4 a_1 df_{0.01}$	0.9%	0.5%
$\sigma_2 R_3 a_1 w$	0.2%	0.2%	$\sigma_2 R_1 a_1 df_{0.01}$	0.8%	0.1%
$\sigma_2 R_4 a_1 w$	0.1%	0.3%	$\sigma_2 R_2 a_1 df_{0.01}$	0.7%	0.2%
$\sigma_1 R_1 a_1 df_1$	2.5%	0.8%	$\sigma_2 R_3 a_1 df_{0.01}$	0.6%	0.3%
$\sigma_1 R_3 a_1 df_1$	1.8%	2%	$\sigma_2 R_4 a_1 df_{0.01}$	0.6%	0.4%
$\sigma_1 R_1 a_1 dc_1 f_{0.1}$	0.1%	0.2%	$\sigma_1 R_3 a_1 dc_1 f_{0.01}$	0.2%	0.6%
$\sigma_1 R_3 a_1 dc_{10} f_{0.01}$	0.1%	0.6%	$\sigma_1 R_1 a_1 dc_{10} f_{0.01}$	0.1%	0.2%
$\sigma_1 R_3 a_1 dc_1 f_{0.1}$	0.06%	0.7%			

Note. $\sigma_1 = 0.5$; $\sigma_2 = 0.4$; $R_1 = 1\,R_J$; $R_2 = 2\,R_J$; $R_3 \sim U[2, 4]\,R_J$; $R_4 = 4\,R_J$; $a_1 \sim LU[0.4, 5]$au; d relates to the dynamical tides model; w relates to the weak tides model; c_x relates to the deposition of x% of the irradiation at the center and tidal heating (corresponding to the tides model), such that the absence of c_x stands for 0%; and f_x relates to a variation in $f_{\rm dyn}$, i.e., f_x stands for $f_{\rm dyn} = x$. The results are based on the statistics from $\geqslant 10^4$ Monte Carlo simulations per each set of initial parameters.

although good fits exist, the exact final eccentricity distribution is still unknown in analytic terms, and usually a preliminary N-body simulation is needed to determine the distribution at the end of the planet–planet scattering. In our calculations, we assume that destabilized systems evolve through planet–planet scattering, leading to the eccentricity distribution given by

$$\frac{dN}{de} \propto e \exp\left[-\frac{1}{2}\left(\frac{e}{2\sigma_e^2}\right)^2\right], \qquad (1)$$

where we adopt a value of $\sigma_e = 0.5$ in our fiducial model (see Table 1 and Section 6). We note that some studies have considered somewhat different distributions, with even higher fractions of highly eccentric orbits, which might improve our results (e.g., Nagasawa et al. 2008; Carrera et al. 2019).

The Rayleigh distribution of the eccentricities arising from the planet–planet scattering determines the fractions of the planets that would evolve through inflated eccentric migration to become HJs/WJs, and would not be tidally disrupted by the host star.

The tidal disruption radius is given by $r_{\rm dis} = \eta R_p (M_\star/M_p)^{1/3}$, where $\eta = 2.7$ (Guillochon et al. 2011), and planets with pericenters approaching these values are assumed to be disrupted and, naturally, are not considered to be HJs/WJs. Planets with too high pericenter approaches, which are not affected by tides, or those that do migrate through inflated eccentric migration, but do not become HJs or WJs in the relevant time considered, i.e., the age considered for the HJs/WJs, are similarly not considered to be HJs and WJs.

The timescale for the planet–planet scattering sets the initial radius of the planet in the migration stage. Gas giants usually form with a typical radius that can exceed even $4\,R_J$, from which there will be a rapid cooling phase up to $4\,R_J$ (Guillot et al. 1996). The following contraction phase is slower, along the Hayashi track. This timescale could take less than a Myr, to reach a typical radius of $1.5\,R_J$. The decoupling timescale from the planet–planet scattering could stray from less then a Myr to a few Myrs, and even more (e.g., Dawson & Johnson 2018, Figure 3), such that the initial radius for the migration phase could be large (corresponding to a short decoupling time) or small (corresponding to a long decoupling time), depending on the planet–planet scattering conditions. A more detailed calculation, which sets out more accurately the initial distribution for our semi-analytical-based population synthesis, is out of the scope of this paper, and is left for a future study. However, we do take this into account by considering several possible initial radii distributions.

3. High-eccentricity Tidal Migration

Tidal migration is a dissipative process, where the tides raised on the planet by the host star extract energy from the planet's orbit, typically leading to its inward migration into shorter-period orbits. Tides raised on the star by the planet can also contribute to tidal migration, but these are typically negligible compared to the effects of the tides raised on the planet, though they might become important under some circumstances (Ginzburg & Sari 2017). In the following, we only consider the tides raised on the planet, and postpone the consideration of the tides on the star to later studies. These tides on the star will lead to corrections in Equation (3), which may be included (e.g., Miller et al. 2009).

In order for the tidal dissipation to be effective, and lead to a significant migration, a small pericenter approach should be considered. This dictates the basic initial conditions for eccentric migration. High-eccentricity tidal migration could therefore be roughly divided into two separate steps: reducing the planet's angular momentum and reducing the planet's energy. In the first step, the HJ/WJ progenitor is excited into an eccentric orbit via planet–planet scattering (Rasio & Ford 1996; Chatterjee et al. 2008; Jurić & Tremaine 2008), as we discuss here, or through other channels for eccentricity excitation (e.g., via the ZLK mechanism; von Zeipel 1910; Kozai 1962; Lidov 1962; Wu & Murray 2003; Fabrycky & Tremaine 2007; Naoz et al. 2011; Petrovich 2015a). In the second step, energy extraction via tides leads to the migration and circularization of the planet's orbit. The energy extracted from the orbit (calculated over an orbital period) per period is dissipated in the planet, and therefore the injected energy heats the planet. The injected energy per unit time is given by

$$L_{\rm tide} = -\frac{E}{a}\frac{da}{dt}, \qquad (2)$$

where E is the orbital energy and a is the semimajor axis. The angular momentum is approximately conserved. For HJs, the final orbit is usually circular, and given the angular momentum conservation one can estimate the final semimajor axis of the HJ, finding $a_{\rm final} = a_0(1 - e_0^2)$, where a_0 and e_0 are the initial semimajor axis and the eccentricity, correspondingly. The same consideration could be applied to any other given final eccentricity (or any eccentricity during the evolution).

Modeling tides is not trivial; in particular, their strong dependence on the internal structure of the planet and other physical aspects of the problem raise many complications. Here, we consider two tide models: one is the widely used tidal model of weak/equilibrium tides (Darwin 1879; Goldreich & Soter 1966; Alexander 1973; Hut 1981), the second one is a model for dynamical tides (Zahn 1977; Mardling 1995a, 1995b). The latter could be especially important, and more efficient during the early migration phases when the planet's orbit is still highly eccentric; and, in that sense, considering only the weak tides model is potentially conservative in terms of the efficiency of the eccentric migration and the long timescales that result (Lai 1997). Our approach is general, and any other tide model could potentially be incorporated, such as the chaotic dynamical tides model (Ivanov & Papaloizou 2004, 2007; Vick & Lai 2018; Wu 2018; Vick et al. 2019), which can potentially further shorten the migration timescales, due to its more efficient extraction of energy. The tide models are discussed below. Here, we consider only nonchaotic dynamical tides. It should be noted that the very same equations are used in the detailed numerical model in paper II. To keep this paper as succinct as possible, and to minimize the overlap with paper II, we refer the reader to paper II for further discussion of tidal evolution, which is also relevant to the semi-analytical model.

3.1. Equilibrium Tide Model

The equilibrium tide model is used in many physical scenarios and has been widely discussed (e.g., Darwin 1879; Goldreich & Soter 1966; Alexander 1973; Hut 1981).

Let us consider an orbit-averaged time evolution of the eccentricity and semimajor axis. If we assume that pseudosynchronization of the planetary spin and the orbit occurs on a short timescale, and that the angular momentum is conserved during the migration, one finds that (e.g., Hut 1981; Hamers & Tremaine 2017)

$$\frac{da}{dt} = -21 k_{\rm AM} n^2 \tau_p \frac{M_\star}{M_p}\left(\frac{R_p}{a}\right)^5 a e^2 \frac{f(e)}{(1-e^2)^{15/2}}, \tag{3}$$

$$\frac{de}{dt} = -\frac{21}{2} k_{\rm AM} n^2 \tau_p \frac{M_\star}{M_p}\left(\frac{R_p}{a}\right)^5 e \frac{f(e)}{(1-e^2)^{13/2}}, \tag{4}$$

where $M_\star$ is the mass of the host star, M_p, R_p, e, a, and n are the mass, radius, orbital eccentricity, orbital semimajor axis, and mean motion of the gas giant, correspondingly, $\tau_p = 0.66$ sec is the planetary tidal lag time, and $k_{\rm AM} = 0.25$ is the planetary apsidal motion constant ($k_{\rm AM}$ and τ_p are taken from Hamers & Tremaine 2017). $f(e)$ is defined by

$$f(e) = \frac{1 + \frac{45}{14}e^2 + 8e^4 + \frac{685}{224}e^6 + \frac{255}{448}e^8 + \frac{25}{1792}e^{10}}{1 + 3e^2 + \frac{3}{8}e^4}. \tag{5}$$

The energy extracted per period, and hence the migration and circularization rate, scales as R_p^5. Consequently, the migration timescales of initially inflated gas giants, where the planetary radius $R_p > R_J$, are shortened relative to the migration timescales of noninflated gas giants, i.e., with a constant radius of R_J. The contraction timescales are long enough to maintain inflated gas giants along a significant part of their dynamical evolution, such that the initial radius of an HJ/WJ will leave a signature on its expected final parameters, which could be also observed.

3.2. Dynamical Tides

Tidal forcing from the star might excite the internal energy modes of the planet (mainly the fundamental f-mode), which might induce an enhanced response (Mardling 1995a, 1995b; Lai 1997; Ogilvie 2014). The energy is mostly extracted during the pericenter approach, and the extraction is more efficient compared with the equilibrium tide model, potentially leading to the even more rapid circularization and migration of the planet. The eccentricity decay is accompanied by pseudosynchronization with the angular frequency of the star, and the excitation of the oscillations in the planet becomes less pronounced as the orbital eccentricity decreases. The energy dissipation by the various modes is gradually suppressed, until the transition to the regime in which the equilibrium tides become more dominant. The quadrupole order of the energy dissipation can be written as follows (Press & Teukolsky 1977; Moe & Kratter 2018):

$$\Delta E = f_{\rm dyn} \frac{M_\star + M_p}{M_p} \frac{G M_\star^2}{R_p} \left(\frac{a(1-e)}{R_p}\right)^{-9}, \tag{6}$$

with $f_{\rm dyn} = 0.1$, unless stated otherwise (Moe & Kratter 2018), being taken for our case of the tidal response of a gas giant. Combining this prescription with the equations of orbital energy and angular momentum, and assuming a constant pericenter, leads to the following equations of the orbital semimajor axis and eccentricity along the migration (Moe & Kratter 2018):

$$\frac{da}{dt} = \frac{a}{P}\frac{\Delta E}{E}, \quad \frac{de}{dt} = \frac{1-e}{a}\frac{da}{dt}. \tag{7}$$

While dynamical tides dominate for large eccentricities, weak tides constitute a more physical description for low ones. The ratio of the migration rate due to dynamical tides to the migration rate due to weak tides is given by:

$$\beta(R_p, a, e) \coloneqq \frac{da/dt|_{\rm dyn}}{da/dt|_{\rm weak}} = \frac{2 f_{\rm dyn} R_p^3 (1-e^2)^{15/2}}{21 G M_p k_{\rm AM} \tau_p (1-e)^9 P e^2 f(e)}, \tag{8}$$

where P is the period of the planet. The transition between the dynamical and weak tides occurs roughly at $\beta \sim 1$, and we set a lower artificial cutoff at $e = 0.2$, such that the transition occurs at $\max\{0.2, e|_{\beta=1}\}$, to avoid the divergence of dynamical tides at $e = 0$.

4. Combined Thermal–Dynamical Semi-analytical Model

We model inflated eccentric migration by coupling the orbital equations, governed by the tidal migration model, with the thermal evolution of the planet, dictated by heating due to tides, irradiation, and thermal cooling. The thermal evolution leads to a planetary contraction due to cooling, which can be slowed by external heating sources and, in extreme cases, even be stopped/partially reversed. The tidal evolution is strongly affected by the radius of the planet, such that the thermal evolution changes the migration rate and timescale.

The virial theorem states a relation between the total and potential energies $E = -(3\gamma - 4)U/(3\gamma - 3)$, such that $U \propto 3GM_p^2/R(5 - \tilde{n})$, where $\tilde{n} = 1/(\gamma - 1)$ is the polytropic index, γ is the heat capacity ratio, taken as $\gamma = 5/3$, and the proportion constant is determined by the numerical gauge. The change in the planet's luminosity is mostly determined by the change in the thermal energy of the ions, which are not degenerate, and their equation of state is given by the ideal gas equation $E = (M_p/\mu)k_B T_c$, where μ is the molecular weight, taken as the proton mass, and T_c is the central temperature of the planet. Following this relation, we derive the following equation for the radius change due to heating and cooling:

$$\frac{dR_p}{dt} \propto \frac{5\gamma - 6}{3\gamma - 4}\frac{R_p^2}{GM_p^2}(L_{\rm ext} - L_{\rm cool}). \tag{9}$$

It should be noted that the energy deposition and loss are generally not constant, and could vary with time and the changes of the orbital parameters—the semi-analytical approach takes this into account. In a similar manner, one can calculate the evolution of the central temperature of the planet:

$$\frac{dT_c}{dt} = \frac{m_p}{M_p k_B}(L_{\rm ext} - L_{\rm cool}), \tag{10}$$

where m_p is the proton mass, $L_{\rm cool}$ describes the cooling of the planet, and $L_{\rm extra}$ describes the further external sources of energies injected, such as tidal heating and irradiation, or any other general source of heating. The external luminosity could generally be a function of the optical depth $\tau_{\rm dep}$ in which it is deposited, or, equivalently, of the pressure $P_{\rm dep}$. The effect of the deposition is strongest for $\tau_{\rm dep} = \tau_c$, i.e., at the center, and it decreases as $\tau_{\rm dep}$ decreases. Although deposition over layers has different physical consequences than deposition at the center (Youdin & Mitchell 2010; Spiegel & Burrows 2013; Komacek & Youdin 2017), for our purposes, the energy injected at outer layers could be translated to a reduced deposition at the center, as shown by Ginzburg & Sari (2016) for the case of a power-law distribution.

To reinflate the gas giant, it can be seen directly from Equation (9) that the minimal external luminosity added should be comparable to the radiating luminosity. For a tidal disruption, a larger energy is required, of the order of magnitude of the binding energy of the planet.

The reduction can be considered by an overall multiplication factor, depending on the depth of the deposition (Ginzburg & Sari 2015, 2016), and the effective temperature could be inferred from the effect of the irradiation of the host star. The exact multiplication factor depends on the not-yet-well-understood heat transfer processes in the planet (which we discuss in more detail below). The planet cools through a blackbody emission, which is given by

$$L_{\rm cool} \propto 4\pi R_p^2 \sigma_{\rm SB} T_{\rm eff}^4, \tag{11}$$

where a correction for blackbody radiation arises from the radiative–convective boundary (RCB) and the transition from isolation to insolation (Ginzburg & Sari 2015). While deposition at the photosphere does not significantly change the cooling rate of the planet, central deposition could lead to more prominent effects. A weak external energy source, such as deposition at the photosphere, could be thought of as "standard" Kelvin–Helmholtz cooling, up to a small correction arising from the external source, which becomes important when $L_{\rm dep}/L_{\rm cool} \gtrsim 1$. The effective temperature at late times is solely determined by the stellar irradiation, and it reaches a steady state, corresponding with the typical observationally inferred values for the effective temperatures of HJs and WJs. At early times, the temperature of the gas giant could exceed this final temperature, before cooling to a quasi-steady temperature, with its temperature then rising again to the one expected from stellar irradiation (see, for example, Figures 1(a) and 2(a)).

While the thermal evolution, and hence also the evolution of the planetary radius, are governed by the central temperature of the planet, the effective temperature, which also appears in the cooling equation, could be changed significantly, without changing the central temperature.

The relation between the effective and central temperatures depends on the pressure gradient inside the planet:

$$\frac{T_{\rm eff}}{T_c} \propto \left(\frac{P_{\rm RCB}}{P_c}\right)^{\gamma/(\gamma-1)}; \tag{12}$$

$$P_c \sim \frac{GM_p^2}{R_p^4},\ P_{\rm RCB} \sim \frac{GM_p}{\bar{\kappa}R_p^2}, \tag{13}$$

where P_c is the central pressure of the planet, $P_{\rm RCB}$ is the pressure at the boundary layer between the radiative and convective regions, and $\bar{\kappa}$ is the mean opacity, averaged over the planet's atmosphere. We consider $\bar{\kappa} = 5 \times 10^{-2}\ {\rm cm^2\ g^{-1}}$, following the typical values in the literature (Ginzburg & Chiang 2019). The proportion factors could be larger than the order of unity, and they could be found more precisely from numerical simulations. The photosphere luminosity is determined by the effective boundary condition for $L_{\rm cool}$. In general, the RCB layer is dictated by the internal adiabat of pressures. It is not located at the photosphere, and the photosphere is in fact an underestimation of it (see the further discussions in Thorngren et al. 2019 and Sarkis et al. 2021). It should be noted that the pressure of the RCB is dictated by both the internal adiabat and the external irradiation, as shown in Equation (12).

The opacity might vary strongly with temperature changes. For high-temperature planets, i.e., $T \gtrsim 10^4$ K, the opacity is given by the Kramers bound-free opacity law $\kappa \propto \rho T^{-3.5}$ (H^- opacity; Kippenhahn et al. 2012; Ginzburg & Sari 2015). Moreover, dustier planets could have higher initial opacity, and hence finish their core accretion with larger radii, which remain inflated for longer timescales, enhancing the effect discussed in this paper.

Coupling Equations (9) and (10) with orbit-averaged evolution equations defined by the tidal migration model (i.e., for weak tides, this is given by Equation (3), and for dynamical tides, by Equation (7)), provides a complete,

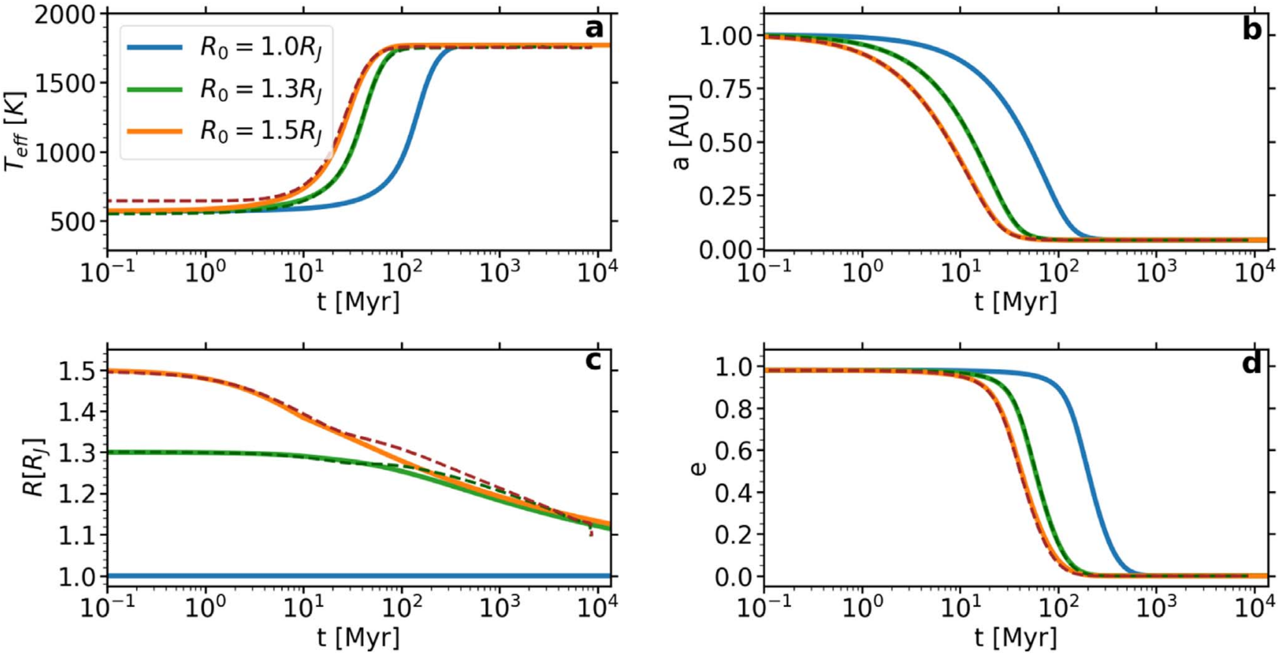

Figure 1. The thermal and orbital evolutions of an HJ progenitor migrating due to weak tides, with photosphere heating due to irradiation. The thermal—panels (a) and (c): effective temperature and radius—and orbital—panels (b) and (d): semimajor axis and eccentricity—evolutions of 1 M_J HJ progenitors with different initial radii are shown (1 R_J—blue; 1.3 R_J—orange; and 1.5 R_J—green). Tides are modeled through a weak tide model; irradiation of the planetary outer layer is included, but with no additional efficient heat conduction to the core. Gas giants with an initial semimajor axis of 1 au and an initial eccentricity of 0.98 are shown. The solid lines correspond to the semi-analytical calculation and the dashed to the numerical calculation.

consistent semi-analytical description of the coupled dynamical–thermal model.

In the following, we discuss the addition of external heating sources, which are discussed in more detail in paper II.

4.1. External Heating

Our semi-analytical model is general, and could effectively take account of any coupling between the thermal and dynamical evolution, as well as include external heating sources, such as tidal heating and irradiation. The efficiency of the heat deposition due to such sources depends both on the heating rate and on the depth of the deposition. Heat deposition in the central parts of a planet has a more significant effect compared with the limited effect of deposition at the photosphere, where radiative cooling effectively disposes of much of the heating.

External heating plays a dual role, keeping the planet inflated for longer time (and, in extreme cases, inflating it) and modifying the effective temperature. At late stages, the observed effective temperature of a gas giant is determined by the the irradiation/other external energy sources applied on the planet, such that the external heating, even if it does not play a role in inflation, should be taken into consideration in the effective temperature calculation. The relative role of the external heating can change during the planet's thermal–dynamical evolution. When the planet is sufficiently far away from the star, the heating due to irradiation is less than the internal heat, and hence it negligibly affects the thermal evolution. When the planet migrates closer to the star, the radiative heating (and tidal heating, when applicable) increases, and one then needs to consider irradiation when determining the planet's temperature. This can be determined by the planetary equilibrium temperature:

$$T_{\mathrm{eff,irr}}(t) \propto \left(\frac{L_\star}{16\pi\sigma_{\mathrm{SB}}r^2(t)}\right)^{1/4}, \qquad (14)$$

where $r(t)$ is the instantaneous distance of the planet from the star. In order to include it in our model self-consistently, and in an efficient way, we consider the averaged effective temperature over an orbit (given that the heating rate is sufficiently small during dynamical times). Similar considerations would apply to the effects of other heating sources on the effective temperature.

5. Case Study Examples

In the following, we present the results of the semi-analytical model, and compare them with the results of the numerical model—we refer the reader to paper II for further examples of this comparison and an extensive description of the numerical approach, its limitations, and its applications.

The semi-analytical model could potentially take into account any general distribution of external energies, by the integration of the contributions from different optical depths (see Ginzburg & Sari 2016 for the case of a constant energy source with a power-law distribution in the optical depth). Currently, for simplicity, the injection of heat is simply given as a point source in the center, while in the numerical modeling one needs to distribute the energy injection over some finite region, so as not to lead to an unstable or divergent result from the nonphysical injection at a singular point. The current chosen distribution in the numerical prescription is a semi-Gaussian (with the dispersion changing with the location of the deposition), which injects heat with a deviation of 15%–50% from the intended value (see paper II for the technical details

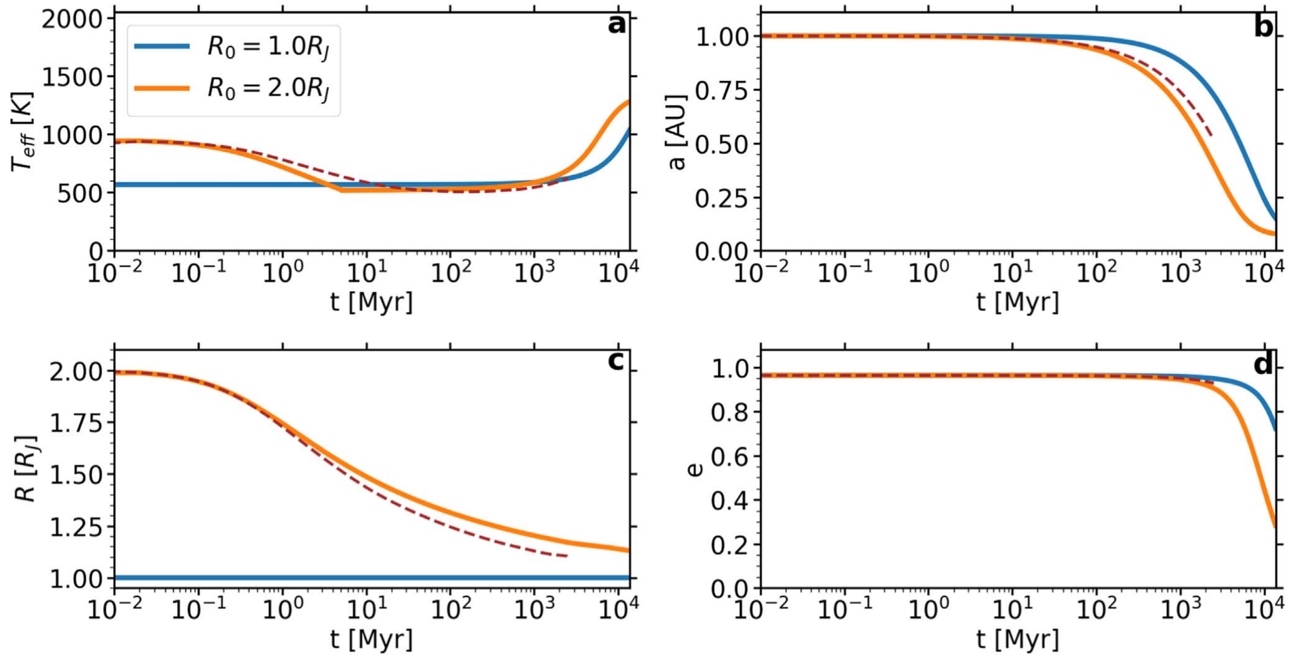

Figure 2. The thermal and orbital evolutions of HJ and WJ candidates migrating due to weak tides, including irradiation, for an initial semimajor axis of 1 au, initial radii of 1 R_J and 2 R_J, and an initial eccentricity of 0.963 (blue—$R_0 = 1\ R_J$; orange—$R_0 = 2\ R_J$). The initial 2 R_J model finalizes with an orbital period of ~ 9.8 days, and the constant 1R_J with 24 days. The solid lines correspond to the semi-analytical calculation and the dashed to the numerical calculation. Panels (a)–(d): time evolutions of the effective temperature, semimajor axis, radius, and eccentricity.

regarding the distribution). It should be noted that the actual physical distribution is not yet understood, since the external heating sources are likely to lead to a distribution with a preference toward the outer areas of the planet, and its extension depends on the exact nature of the heat transfer. Nevertheless, since the effect of a generally distributed external energy on the dynamical evolution could be translated into some appropriate amount of energy injected at the center, we are practically able to treat the effect of any external energy distribution, through an effective heat injection in the center, with an order-unity constant prefactor to calibrate between the distributed injection in the numerical model and in the semi-analytical one.

In Figure 1, we present an example of the coupled dynamical–thermal evolutions of eccentric $1\,M_J$ gas giants with different initial radii, an initial semimajor axis of 1 au, and an initial eccentricity of 0.98. We compare the evolution of a constant $1\,R_J$ with that of initially inflated planets. All the planets in our model experience photospheric heating induced by stellar irradiation.

We find that the semi-analytical and numerical approaches are in excellent agreement, and yield similar results (see the additional examples in paper II).

The typical migration and circularization timescales are correspondingly given by (for the weak equilibrium tide model):

$$\tau_{\rm mig,weak} = \frac{a}{|\dot{a}_{\rm weak}|} = \frac{(1-e^2)^{15/2} M_p}{21 k_{\rm AM} n^2 \tau_p M_\star e^2 f(e)} \left(\frac{a}{R_p}\right)^5, \qquad (15)$$

$$\tau_{\rm circ,weak} = \frac{e}{|\dot{e}_{\rm weak}|} = \frac{2}{21} \frac{M_p (1-e^2)^{13/2}}{k_{\rm AM} M_\star n^2 \tau_p f(e)} \left(\frac{a}{R_p}\right)^5. \qquad (16)$$

These timescales are somewhat optimistic, since the initial radius shrinks with time. However, it could be seen that initial larger radii lead to shorter migration/circularization timescales.

A larger initial radius shortens the migration and circularization timescales, which become 10 times shorter in this case, as expected from the strong dependence of the tidal forces on the radius. The radius contracts to a final radius of $\gtrsim 1\,R_J$ within a Hubble time. A significant part of the migration takes place with a radius larger than R_J, since the migration timescale is shorter than the contraction timescale, manifesting the key role played by the initial inflated radius on the evolution. It should be noted that the radius of the initially noninflated planet changes as well, but negligibly. In addition, the effective temperature is roughly constant at the early stages of the evolution, but increases as the planet gets closer to the star, since the irradiation from the star determines the effective equilibrium temperature of the planet.

In Figure 2, we present the coupled thermal–dynamical evolution of a formed WJ and HJ. Similar to Figure 1, the migration timescales are significantly shortened for a planet with the same initial separation, eccentricity, and mass, but a different initial radius.

We can see a good agreement between the semi-analytical model and the numerical one, throughout the evolution. Note that the numerical model terminates due to the limitations of the (older) `MESA` version used in this regime; for more details regarding the termination criteria of the numerical model, see paper II. In such cases, the numerical model could be extrapolated using the semi-analytical model.

We find that all the WJs produced through inflated eccentric migration are all still eccentric, though somewhat circularized, and that their final states are dictated by the initial angular momentum and the energy dissipation rate. Inflated eccentric migration enhances the migration rate, such that planets that would never otherwise have become HJs or WJs, when considering an initial and constant $1\,R_J$ radii, migrate more efficiently, enabling WJs to become HJs and little-/nonmigrating planets to become WJs. Furthermore, inflated WJs, given the same initial conditions, would be less eccentric, since they proceed faster in their migration; some of the expected WJs from the $1\,R_J$ case become HJs once inflated migration is accounted for. These transitions between regimes, the effective accelerated changes in the migration flows of gas giants, are discussed in more detail in Section 6.

The semi-analytical approach provides an efficient, simple, and computationally inexpensive approach to modeling the evolution, allowing us not only to consider a large phase space of initial conditions, but also to consider a detailed population synthesis study that describes the evolution and formation of Jupiters. Nevertheless, a priori, this simple approach cannot describe the detailed internal structure of a planet, and it might give rise to inaccurate modeling of the macroscopic properties of a planet and its evolution. However, we generally find an excellent agreement between the numerical approach and the semi-analytical one, which therefore allows us to use the simple semi-analytical approach robustly.

6. Population Synthesis Study and Occurrence Rate Estimates

In order to calculate the occurrence rate of HJs and WJs and their properties, we use a population synthesis study. Given the excellent agreement that we found between the semi-analytical model and the numerical model (see Figures 1, 2, and the further discussion in paper II), we rely on the semi-analytical model to provide a fast and efficient evolution model, enabling us to study and analyze a wide range of initial conditions, and to explore the formed population of HJs and WJs through a population synthesis study, without the need of a costly (computational) numerical simulation.

We consider various plausible choices for the initial conditions of potential HJ/WJ progenitors (see Table 1), and randomly sample the parameter space of the initial conditions.

In order to characterize the population of HJs and WJs formed through regular eccentric migration and through inflated eccentric migration, we make several assumptions regarding the initial conditions. We also reiterate that in this study we are only considering eccentric migration following planet–planet scattering, while other secular evolution models for eccentricity excitation will be explored elsewhere.

We assume a continuous star formation rate in the Galactic disk, since the planet formation rate is proportional to the star formation rate (e.g., Behroozi & Peeples 2015), and we therefore sample the age (i.e., the evolution time) for each planet from a uniform distribution in the range 1–12 Gyr. We then evolve each planet in our sample using our semi-analytical approach, and examine its properties after a given time. We define an HJ as a gas giant with a final period shorter than 10 days, and we define a (migrating) WJ as a gas giant with a final period of 10–200 days, only considering cases where the semimajor axis has shrunk by at least a factor of 2 relative to the initial one, so as not to consider possible contributions from WJs that have potentially been formed in situ. A gas giant is assumed to be disrupted if its pericenter is smaller than the Roche radius. We define $f_{\mathrm{HJ/WJ}}$ as the fraction of formed HJs/WJs, as described above, and use it to derive the total expected frequency of HJ and WJ systems in the galaxy. The fraction of formed HJs and WJs for any given time t is given by

$$f_{\mathrm{HJ/WJ}}(t) = \int \frac{dN}{dm_p}\frac{dN}{de_0}\frac{dN}{da_0}\frac{dN}{dR_{p,0}} \times \chi(m_p, e_0, R_{p,0}, a_0, t)\, de_0 da_0 dR_{p,0} dm_p, \tag{17}$$

where dN/de_0, dN/da_0, and $dN/dR_{p,0}$ are the differential distributions of the initial eccentricities, the semimajor axes, and the initial planetary radii, respectively. χ is an indicator function determining whether a given system has evolved to become an HJ/WJ after a time t, based on the semi-analytical coupled orbital–thermal evolution. Calculating the fraction of formed HJs/WJs, we also take into consideration the possibility of tidal disruption, and exclude the disrupted population from the population of formed HJs/WJs. The tidal disruption radius is given by $r_{\mathrm{dis}} = \eta R_p (M_\star/M_p)^{1/3}$, where $\eta = 2.7$ (Guillochon et al. 2011), and planets with a pericenter approaching these values are assumed to be disrupted, so, naturally, are not considered to be HJ/WJs. Planets with a sufficiently large pericenter approach, which are negligibly affected by tides, and do not migrate, are also not considered as migration-formed HJs/WJs, nor are those that do migrate due to inflated eccentric migration, but do not migrate close enough to the star to be considered as HJs or WJs in the relevant time considered.

In order to calculate the total occurrence rate, the frequency of HJs and WJs among the stellar systems is given by

$$\mathcal{F}_{\mathrm{HJ/WJ}} = f_J \times f_{2J} \times f_{\mathrm{unstable}} \times f_{ecc,J} \times f_{\mathrm{HJ/WJ}}, \tag{18}$$

where f_J is the fraction of stellar systems hosting gas giants; f_{2J} is the probability of finding at least two gas giants (in most cases, in order to scatter a gas giant into a highly eccentric orbit that leads to eccentric migration, another planet as massive as Jupiter is needed); f_{unstable} is the fraction of unstable systems, in which planet–planet scattering is likely to occur; and $f_{\mathrm{ecc,J}}$ is the fraction of sufficiently eccentric gas giants that would evolve through eccentric migration to become HJs and WJs.

We study each choice of parameters (see the parameter sampling description below) for the initial distribution of $\gtrsim 10^4$ semi-analytical simulations to determine the fractions of cases that successfully evolve to become an HJ/WJ. Calculating this at at a given time since birth (up to a specific chosen time) provides us with the delay-time distribution, i.e., the fractions of HJ or WJ as a function of time since planet formation. As time goes by, planets that were observed to be WJs could further migrate to become HJs, HJs could be disrupted, and planets that were too distant to be WJs could become them. The rate at which the different areas in the parameter space are filled is determined by the initial conditions of the planet (a_0, e_0, R_0, and m_p), the tide model (weak, dynamical, etc.), and the external energy sources that could slow the contraction.

We consider a specific star formation history, and integrate the delay-time distribution (our Green function), weighted by the star formation rate, to obtain a realistic estimate of the current fractions of HJs/WJs in the galaxy at the current time, containing both young and old planetary systems. For disk stars, most relevant for the currently observed exoplanet hosts, we consider a continuous, uniform rate of star formation for the Galactic disk, as mentioned above. This is generally consistent with the inferred local star formation history of the Galactic disk stellar populations where most of the exoplanets have been observed to date. Note, however, that the current exoplanet samples are dominated by those identified by the Kepler mission. The age distribution for Kepler stars peaks at around 2.5 Gyr, and gradually falls off to larger ages (Berger et al. 2020).

In the following, we describe the parameters characterizing the initial conditions, the specific ranges of these parameters, and their motivations. In order to evolve a gas giant, we need to ascribe a planet with both physical and orbital properties, including mass, initial radius, initial separation, and initial eccentricity.

6.1. Parameter Space of the Initial Conditions

Taking the observationally inferred occurrence rate and the distributions of the gas giants as initial conditions is not fully self-consistent, since the observed systems have potentially already been affected by evolution. However, given the overall very low fractions of HJs and WJs among the entire exoplanet population, and given the lack of direct data on the initial state of the systems, it is likely that most systems did not evolve significantly after their initial formation, following the dissipation of the disk. Thus, overall, the currently observed period and mass distribution are assumed to still reflect the postformation initial conditions of gas giants.

In the following, we first discuss our choices for the initial distributions, before presenting the resulting populations in the next section.

Planetary radii. Planet formation models suggest that gas giants form with large radii and rapidly contract to a radius of $\sim 4\,R_J$, regardless of their mass, where a phase transition occurs and the gas giants contract and cool at a slower rate (Guillot et al. 1996). Therefore, we assume that the initial planet radii are uniformly distributed between some lower and maximal radii, generally extending between $1\,R_J$ and $4\,R_J$, where we consider four possible subranges, defined by $R_1 - R_4$ (see Table 1). It should be noted that a more self-consistent choice of radii distribution should be taken from a planet–planet scattering simulation coupled to the radius evolution. This is out of the scope of the current paper, and will be left for future studies.

Stellar and planetary masses. In this study, we only consider Sun-like stellar hosts, all having the same (Solar) mass.

The planetary masses are chosen from a power-law distribution with an exponent of −1.1, in the range 0.1–10 M_J, consistent with observations (Butler et al. 2006).

Semimajor axes. The semimajor axes of planets are assumed follow a log-uniform distribution, and we consider the range between 0.4 au and 5 au (the planets could be scattered into highly eccentric orbits, with eccentricities close to 1 orbit for $\sim 1\,M_J$ planets for separations greater than 0.4 au; Dawson & Johnson 2018). We also considered other distributions extending to larger separations, as mentioned below.

Eccentricities. The eccentricities are assumed to follow a Rayleigh distribution, i.e., $dN/de = (e/\sigma_e^2)\exp(-e^2/(2\sigma_e^2))$, consistent with planet–planet scattering models (e.g., Jurić & Tremaine 2008 and references therein). In order to shorten the running times, we sample the eccentricities starting from 0.8, and then properly normalize the results according to the Rayleigh distribution.

6.2. Normalization Factors

In the following, we consider the various factors used to calculate the overall predicted occurrence rates of HJs and WJs from inflated migration.

Gas-giant occurrence. The occurrence rate of planetary systems hosting a gas giant is of the order of 25% (Wang et al. 2015), but this depends on metallicity, and could be as low as 5% for low-metallicity hosts. Here, we adopt a fiducial fraction of the stellar systems hosting such planets as $f_J = 0.17$ for our model, to account for a nonextreme average case. This fraction is decreased by only considering gas giants that were not destroyed during the migration.

Occurrence of planetary systems with at least two gas giants. We will assume that every system that has one gas giant has at least two initially, i.e., $f_{2J} = 1$ for a given f_J, generally consistent with observations (Bryan et al. 2016), finding that $> 50\%$ of all planets residing between 1 au and 5 au have additional gas-giant companions, and considering that planet scattering typically ejects most planets from the systems, leaving behind two to three planets.

Fraction of (initially) unstable systems. We take the fraction of unstable systems (given that they host at least two gas giants) as $f_{\rm unstable} = 0.75$. The actual number is unknown, but the overall consistency of the observed eccentricity distribution with a Rayleigh distribution suggests that planet–planet scattering due to unstable systems is ubiquitous (Ford & Rasio 2008), given that planets are generally thought to form initially on circular orbits.

Overall normalization prefactor. Taken together, following Equation (18), we get a normalization factor, $f_{\rm norm} = 0.1257$, of $F_{\rm HJ/WJ} = f_{\rm norm} f_{\rm HJ/WJ}$, with $f_{\rm HJ/WJ}$ calculated from our models.

7. Results

Our main results for the HJ/WJ occurrence rates from the various models considered are summarized in Table 1, in which we explore a large parameter space. We will note briefly the different parameters taken into consideration there: $\sigma_1 = 0.5$; $\sigma_2 = 0.4$; $R_1 = 1\,R_J$; $R_2 = 2\,R_J$; $R_3 \sim U[2,\ 4]\,R_J$; $R_4 = 4\,R_J$; $a_1 \sim LU[0.4,\ 5]$au; d relates to the dynamical tides model; w relates to the weak tides model; c_y relates to the deposition of y % of the irradiation at the center and tidal heating (corresponding to the tides model), such that the absence of c_x stands for 0%; and f_x relates to a variation in $f_{\rm dyn}$, i.e., f_x stands for $f_{\rm dyn} = x$. U stands for a uniform distribution and LU for a log-uniform distribution.

The distributions of the orbital properties are only shown for our fiducial model, i.e., $\sigma_1 R_2 a_1 d f_{0.1}$.

We considered a variety of models, which differ in terms of the initial radii distribution of the planets, the initial eccentricity distribution, and the initial semimajor axes distribution. In addition, we also consider different types of tidal evolution, either weak or dynamical, and different efficiencies of external heating (the fraction of heat injected into the center of the planet, which depends on the not-yet-understood heat transfer processes).

7.1. Occurrence Rates

As can be seen in Table 1, the different models and physical processes included give rise to large differences in the fractions of HJs and WJs, up to factors of 10–20 between the most extreme cases. The models we consider cannot robustly

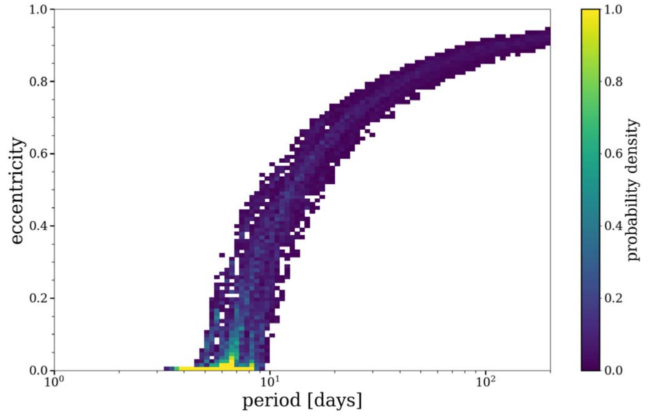

Figure 3. Two-dimensional histogram of the eccentricities and orbital periods of HJs and WJs after a Hubble time, as obtained from the population synthesis models initialized with $\sigma_1 R_3 a_1 df_{0.1}$. The probability density is normalized according to the total fraction of successful formations of HJs and WJs among all of the initial conditions sampled.

reproduce the observationally inferred occurrence rates of HJs ($F_{\rm HJ}^{\rm Obs} = 0.3\%–1.5\%$). However, our strong dynamical tides models (df1 and df0.1), without efficient central heating, give occurrence rates of $1.6–2.5 \times f_{\rm norm} = 0.2\%–0.32\%$, i.e., marginally reproduce the lower estimates for the observed HJ occurrence rate. The same models can robustly reproduce the occurrence rates of eccentric WJs ($F_{\rm WJ}^{\rm Obs} = 0.04 - 1.37\%$, given that 30%–90% of all WJs are assumed to have formed in situ, or through disk migration, producing low-eccentricity WJs; Huang et al. 2016). We find occurrence rates of $0.4–2 \times f_{\rm norm} = 0.05\%–0.26\%$ in our strong dynamical tides models. In general, we find that the consideration of initially inflated gas giants gives rise to occurrence rates of WJs that are two to three times higher compared with models of constant noninflated gas giants, but that it decreases the occurrence rate of HJs by 20%–30%, and up to a factor of two in some cases.

The occurrence rate of HJs could be enhanced by more rapid inflated eccentric migration, bringing them in from larger distances, compared with noninflated planets. However, we find that the overall occurrence rate of HJs decreases with inflated migration. This is due to the enhanced tidal disruption of the now larger gas giants (with correspondingly larger Roche radius) when they are first scattered into high eccentricities; see Figure 4. The tidal disruption is determined merely by the initial conditions, apart from cases of extremely efficient heating, which leads to reinflation. Hence, the decoupling in the latter stages from the planet–planet scattering phase, which is expected to lead to initial lower radii, also leads to a higher formation fraction of HJs. In contrast, the occurrence rate of WJs increases, due to the flow in the parameter space enabling gas giants that would otherwise (if they were not initially inflated) not have migrated (or little migrated) to migrate more rapidly, so as to attain sufficiently small semimajor axes to become WJs, while the migration of WJs to become HJs does not increase at a similar level.

7.2. Parameter-space Evolution

Figure 3 shows the resulting distribution of the eccentricities and orbital periods of our population synthesis models initialized with $\sigma_1 R_3 a_1 df_{0.1}$, where we find the majority of HJs formed via eccentric migration to have been circularized, whereas the WJs are eccentric, suggesting that their migration process has not yet terminated. The empty region on the left reflects the tidal disruption of planets with pericenters below the Roche radius. The empty region on the right reflects the conservation of angular momentum along the migration, such that an initial high eccentricity sets a lower bound on the semimajor axis; planets with too high pericenters are not affected by tides and do not migrate.

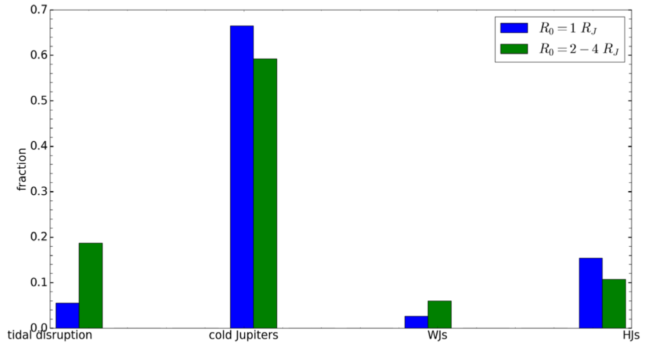

Figure 4. A diagram of the final fractions concluded from the Monte Carlo simulation and a comparison of the inflated ($\sigma_1 R_3 a_1 df_{0.1}$) and noninflated ($\sigma_1 R_1 a_1 df_{0.1}$) initial radii. The fractions could be summed over to 1—we just present the results of the simulation, without further normalization. The results are based on 10^4 runs of the semi-analytical model per each case.

In Figure 4, we present a histogram describing the final fractions of the possible outcomes of eccentric inflated migration: cold Jupiters, WJs, HJs, or tidal disruption. The population is dominated by cold Jupiters, i.e., gas giants with periods larger than 200 days or gas giants that did not migrate at least half of their initial semimajor axis. Inflated eccentric migration reduces the percentage of cold Jupiters, since the migration is more efficient in this model, leading to a more significant flow in the parameter space from the cold Jupiter regime to the WJs. The fraction of WJs increases, due to inflated eccentric migration, from similar considerations. However, the fraction of HJs also reduces, due to tidal disruption.

In Figure 6, we present the eccentricity distribution of HJs and WJs. It can be seen that within this time, HJs tend to circularize, while WJs tend toward higher eccentricities, although they could obtain lower eccentricities, with the lowest at $\lesssim 0.1$.

7.3. HJ and WJ Populations

HJs and WJs migrate faster via the inflated eccentric migration channel, as also manifested for a specific case in Figure 1. The HJ population formed via inflated eccentric migration can be distinguished from noninflated migrating giants. The distribution of the formed inflated HJs is shifted toward larger periods, as can be seen in Figure 5, in agreement with the findings of Petrovich (2015b), who conducted a population study of the cooling of initially inflated giants (where no external heating was considered in the study) for a secular evolution channel. The larger sizes of the inflated HJs make them susceptible to tidal disruption at larger pericenter approaches, as can be seen directly from the expression for the Roche limit $r_{\rm dis} = \eta R_p (M_\star / M_p)^{1/3}$, due to the increased radius and small pericenter.

Indeed, the fraction of tidally disrupted planets increases significantly for models with inflated giants (e.g., see Figure 4). These disruptions arise from the initial conditions, and not the

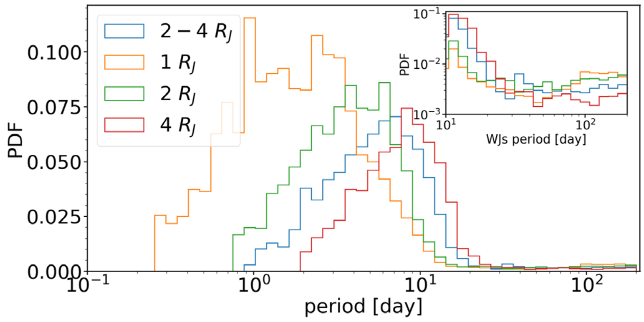

Figure 5. The period distribution as found by the Monte Carlo simulation, based on the semi-analytical model, for different initial radius distributions. The rest of the parameters are drawn according to $\sigma_1 R_3 a_1 d f_{0.1}$. In the inset figure, we introduce the probability distribution function of WJs only (with the same color codes).

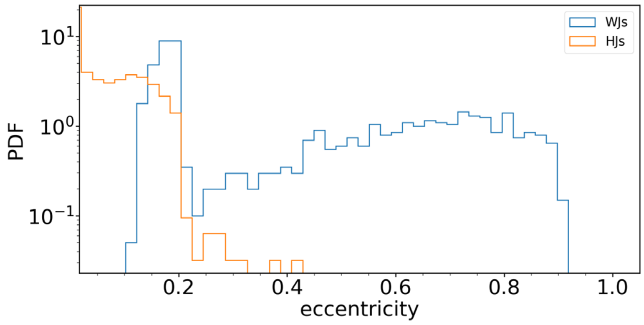

Figure 6. The eccentricity distribution of HJs and WJs as found by the Monte Carlo simulation, based on the semi-analytical model, after 1 Gyr. The the parameters are drawn according to $\sigma_1 R_3 a_1 d f_{0.1}$.

later evolution. Planets are initiated at their largest sizes following their formation, and they can therefore be disrupted at higher pericenter approaches at these times, but then the Roche radius increases as the planets contract.

Only in cases where significant external tidal/radiative heating is able to reinflate planets do tidal disruptions occur during the migration. Indeed, models with significant heating of the central parts (c1, c10) show even higher tidal disruption rates, and, moreover, such disruptions occur during the migration of the planets, following their tidal and radiative inflation, and not immediately after their initial scattering to high eccentricities.

Star formation modifies both the populations of HJs and WJs. As a convolution of single-time star formation events, it gives rise to the further formation of WJs, together with a smaller fraction of HJs, and it indicates that WJs might be younger than HJs, since gas giants that are currently observed as WJs could migrate in and become HJs. We therefore predict that, on average, WJs should reside in younger systems than HJs, if eccentric migration plays a significant role in their production.

7.4. Effects of External Heating on the Population

External heating leads to slowed cooling and, hence, increases in both the production of WJs and the tidal disruption rate (see, e.g., Table 1 and Figure 4). In terms of the WJ–HJ parameter space, external heating speeds up the flow from cold Jupiters to WJs, and so on; taken together, it gives rise to an increased total number of WJs, compared with HJs and tidal disruptions.

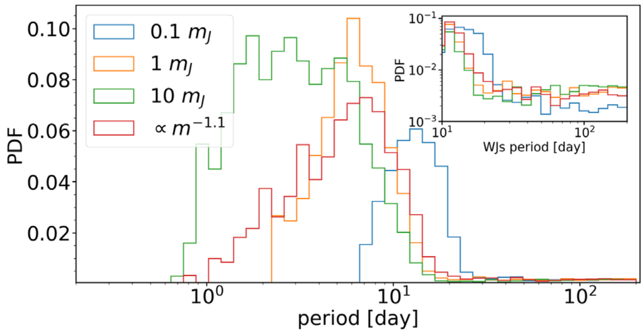

Figure 7. The period distribution as derived from the Monte Carlo simulation, based on the semi-analytical model, for different masses, normalized to 1. The rest of the parameters are drawn according to $\sigma_1 R_3 a_1 d f_{0.1}$. The red line corresponds to our standard mass distribution—$\propto m^{-1.1}$, within the range [0.1, 10] m_J. In the inset figure, we introduce the probability distribution function of WJs only (with the same color codes).

The efficiency of the external heating deposition depends on the depth of the deposition, its duration, and its amplitude. For example, irradiation deposited at the outer layers of the gas giants mainly contributes to the effective temperature, but makes a negligible contribution to the heating of the central parts of the planet, and consequently little affects the radius of the planet, nor the dynamical tidal evolution, which strongly depends on the radius.

This is not the case if a fraction of the irradiation is deposited at the center. In this case, there could be a significant effect that might even lead to inflation, if the fraction of centrally deposited irradiation is sufficiently large. The exact process of energy transfer from the outer planets to the interior is still unknown, and several mechanisms have been suggested (e.g., Arras & Socrates 2010; Batygin & Stevenson 2010; Youdin & Mitchell 2010), where the main motivations for these have been the observations of old inflated HJs, which likely require such efficient heat transfer mechanisms. We encapsulate the uncertainties in the external heating source and the depth of the deposition in an efficiency of deposition at the center, similar to other studies that have focused on thermal evolution, rather than on coupled thermal–dynamical evolution (e.g., Bodenheimer et al. 2001; Komacek et al. 2020). Our central heat deposition models therefore do not correspond specifically to any of the suggested models, but rather bracket the potential effects of the potentially efficient heat transfer.

7.5. Dependence on Parameters

The final population and its properties depend strongly on the choice of the initial distributions and their parameters. While some of the distributions are well constrained from observations, others suffer from large uncertainties, which we account for by considering several choices of parameters. In the following, we discuss several possible choices of parameters and their effects on the final distributions.

In Figure 5, we present the dependence of the final periods on the initial radii distribution. The population formed via the inflated eccentric migration channel peaks in a larger period and gives rise to enhanced filling of the available parameter space of WJs.

As shown in Figure 7, the initial population of more massive gas giants gives rise to a postmigration population residing in smaller periods. This might be expected, given that migration is more efficient for massive planets, as can be seen directly from

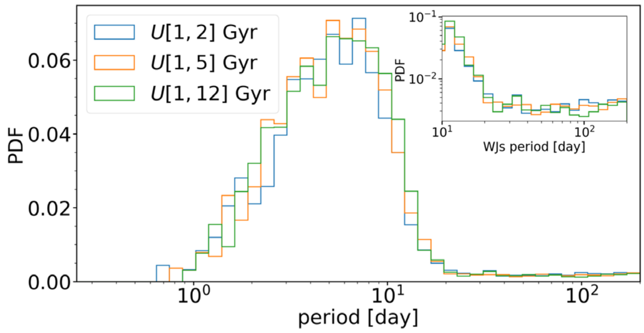

Figure 8. The period distribution as derived from the Monte Carlo simulation, based on the semi-analytical model, after different times, normalized to 1, considering continuous star formation. The rest of the parameters are drawn according to $\sigma_1 R_3 a_1 df_{0.1}$. In the inset figure, we introduce the probability distribution function of WJs only (with the same color codes).

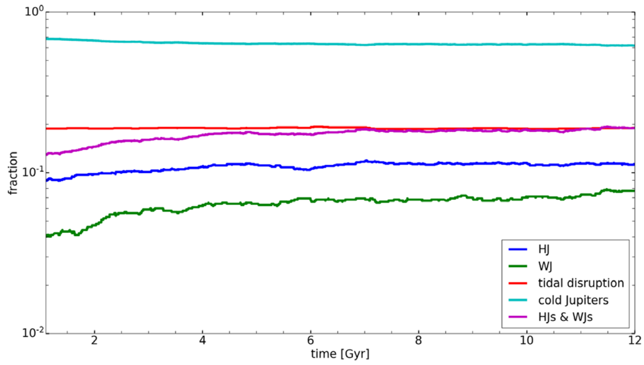

Figure 9. Delay-time distribution, as derived from the Monte Carlo simulation, for $\sigma_1 R_3 a_1 df_{0.1}$, after one star formation event. The plot is normalized to 1 (each time) and is based on 10^4 runs of the semi-analytical-based Monte Carlo simulation per case.

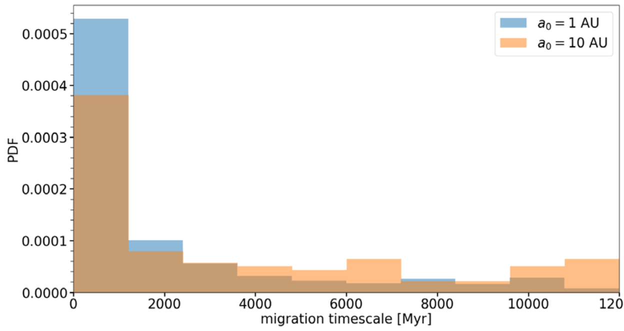

Figure 10. Histogram of the migration timescales of HJs, for different initial semimajor axes, after 12 Gyr from a single star formation event, as derived from our population synthesis based on the semi-analytical model (the rest of the parameters are sampled according to our fiducial model).

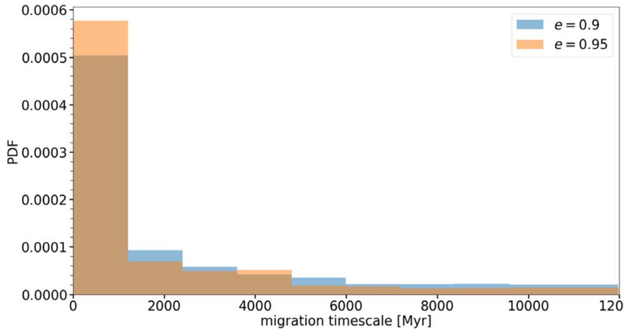

Figure 11. Histogram of the migration timescales of HJs, for different initial eccentricities, after 12 Gyr from a single star formation event, as derived from our population synthesis based on the semi-analytical model (the rest of the parameters are sampled according to our fiducial model).

the tidal evolution equations (e.g., Equations (3) and (7), for weak and dynamical tides, correspondingly), although this is not trivial, given the transitions and flows between cold Jupiters, WJs, HJs, and tidally disrupted planets, which change in each model. It should be noted that the migration timescale of lower-mass planets is shorter than that for more massive ones (see also paper II). However, they are more vulnerable to tidal disruption, such that the overall effect is that lower masses are more efficient in the production of WJs, rather than HJs.

The observational mass distribution sets more weight on the less massive gas giants, such that the total period distribution flattens to include a larger range of periods, from HJs to WJs.

In Figures 8 and 9, we show the time evolutions of the gas-giant population. As time goes by, more and more cold Jupiters migrate inward to become WJs, some of them migrate to become HJs, and some will migrate further and be disrupted. This can be seen by following the peak of the distribution. At early times, there is a quick rise in HJs, which form more rapidly; but at later times, WJs form, and the peak gradually moves toward larger periods. The vast majority of the HJs migrate via timescales shorter than 1 Gyr. Considering star formation leads to a continuous flow of formed HJs, and could basically be understood as a time convolution of the distribution derived from a single star formation event. In models without central heating, planets do not reinflate, and all tidal disruptions occur promptly, following the scattering of the planets, while tidal and radiative heating do not inflate the planets, which attain their maximal radii following their formation. When central heating is efficient, planets can reinflate and become larger than their original radii, and be more prone to tidal disruptions at later times. The exact amount of heat needed for reinflation can be roughly estimated by setting the total luminosity to be larger than 0, as can be seen in Equation (9). The effect of reinflation can also be seen from the decreased fraction obtained for gas giants with large amounts of energy injected in their centers (see Table 1).

In Figures 10, 11, and 12, we present the dependence of the migration timescale on the initial semimajor axis and eccentricity, correspondingly. As expected, larger initial semimajor axes lead to larger migration timescales of HJs, and higher initial eccentricities lead to more efficient migration that finally leads to smaller migration timescales. Larger initial radii lead to extremely short migration timescales, while long timescales are cut out of the histogram, due to the elevated disruption rates.

In Figure 13, we present the dependence of the eccentricity distribution on the dispersion. Lower-eccentricity distributions give rise to larger rates of WJs, but smaller rates of HJs, as expected. Eccentric tidal migration is more efficient, i.e., it extracts more energy from the orbit, when larger eccentricities are included. Since the WJs produced via inflated migration could be thought of as transient HJs that did not manage to end their migration after a given time, their fraction increases when a lower-eccentricity dispersion is taken into consideration.

In Figure 14, we present the dependence of the dynamical tides model on $f_{\rm dyn}$. Larger $f_{\rm dyn}$ corresponds to the more

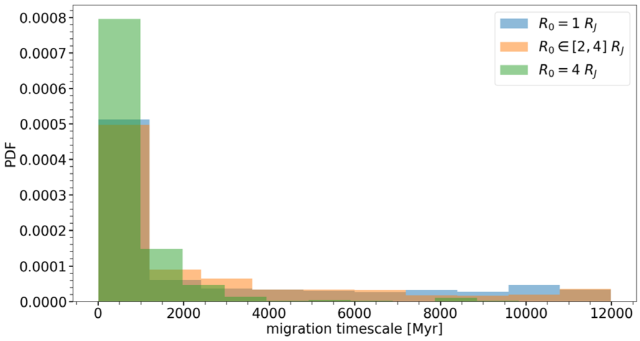

Figure 12. Histogram of the migration timescales of HJs, for different initial radii, after 12 Gyr from a single star formation event, as derived from our population synthesis based on the semi-analytical model (the rest of the parameters are sampled according to our fiducial model).

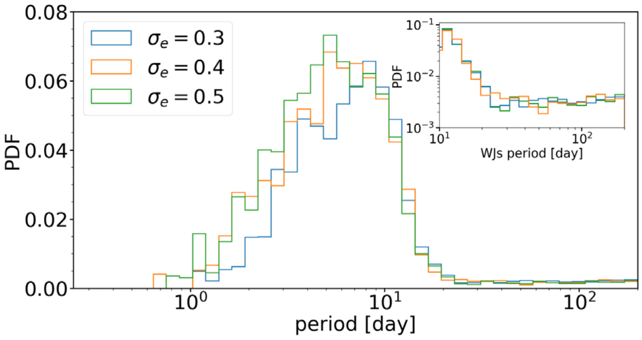

Figure 13. The period distribution as derived from the Monte Carlo simulation, based on the semi-analytical model, for different masses, normalized to 1. The rest of the parameters are drawn according to $\sigma_1 R_3 a_1 d f_{0.1}$, which corresponds to $\sigma_e = 0.5$. In the inset figure, we introduce the probability distribution function of WJs only (with the same color codes).

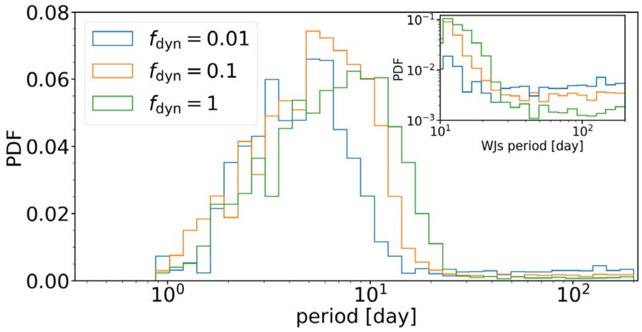

Figure 14. The period distribution as derived from the Monte Carlo simulation, based on the semi-analytical model, for different choices of $f_{\rm dyn}$, normalized to 1. The rest of the parameters are drawn according $\sigma_1 R_3 a_1 d$, which corresponds to $\sigma_e = 0.5$. In the inset figure, we introduce the probability distribution function of WJs only (with the same color codes).

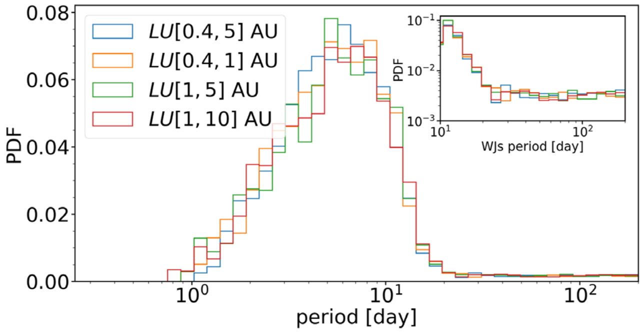

Figure 15. The period distribution as derived from the Monte Carlo simulation, based on the semi-analytical model, for different choices of semimajor axis distribution, normalized to 1, considering continuous star formation. The rest of the parameters are drawn according $\sigma_1 R_3 d f_{0.1}$. *LU* stands for logarithmic uniform distribution. In the inset figure, we introduce the probability distribution function of WJs only (with the same color codes).

efficient extraction of energy via tidal force, which yields faster tidal migration. In terms of the population, the period distribution peak moves toward larger periods as the efficiency of the tides rises, and there is enhanced tidal disruption, together with enhanced production of HJs and WJs.

In Figure 15, we present the dependence on the choice of the boundaries in the semimajor axis distribution, and consider the contributions from different semimajor axes. The overall distribution suggests a preference for HJ production from large initial semimajor axes, but the results seem robust under the choices of distributions.

8. Discussion and Implications

Key findings. As discussed and shown above, inflated eccentric migration following planet–planet scattering in planetary systems could potentially explain the whole population of eccentric WJs and a significant fraction of HJs (or even all of them, given the lowest inferred estimates). It also leads to a high disruption rate of systems, as it accelerates the parameter-space flow from HJs to tidally disrupted gas giants. Inflated eccentric migration could also play an important role in other systems, where high eccentricities are excited through secular processes in general and ZLK oscillations in particular—these will be discussed in future studies (see also Petrovich 2015b for a study of contracting planets in this context). Our results suggest that any modeling of planetary systems, and in particular of young planetary systems, should self-consistently account for thermal evolution. It should consider the evolving size of the planets and the coupling between the thermal and dynamical evolutions. These could play key roles in the system's dynamics, and in the final sculpting of its architecture.

The longer overall timescales and shorter time spent at eccentric orbits would potentially allow for more significant contraction of the gas giants before significant migration occurs, giving rise to a weaker, though still important, effect of inflated eccentric migration. It should be noted that it is important to model the ZLK–Octupole order, in this context, given the conditions and timescales involved (Naoz et al. 2011).

Inflated eccentric migration is most pronounced at the earliest times, when planets are still in their infancy/at a young age and are still far more inflated than at later times, after contraction. Moreover, when eccentricity excitation occurs through secular processes, tidal effects can induce tidal precession, which can quench the eccentricity excitation, giving rise to lower eccentricities that would be expected without the effects of tides, thereby leading to larger pericenter approaches and less effective tidal dissipation and migration.

Implications for giant-planet formation. Our results could have a wide range of implications with respect to various aspects of giant-planet formation and evolution. They shed light on and point to the critical role played by the physical evolution of the planets, and its coupling with the dynamical evolution, which significantly changes the behavior of eccentric migration processes. It gives rise to the efficient formation of both HJs and eccentric WJs, where the latter, in particular, are more difficult to form efficiently through previously studied eccentric migration. We provide their relative fractions overall, and as a function of the age of the systems, as well as detailed predictions of their physical properties. Furthermore, the shorter migration timescales due to inflated eccentric migration could give rise to very young HJs, Myrs old, which are typically suggested to form through disk migration (see the detailed review of disk migration in Baruteau et al. 2014) or via other channels that focus on the stages after gas dissipation (e.g., Wu et al. 2007). The HJs and WJs formed in the proposed channel could have a range of inclinations, even retrograde ones, but would generally have a preference for prograde orbits, given that their initial inclinations were excited by planet–planet scattering (Beaugé & Nesvorný 2012).

Potential caveats and challenges. In our models, all the WJs formed via eccentric tidal migration are effectively transient HJs that did not complete their migrations within a given time, i.e., eccentric WJs rather than circular ones (although they could reach relatively small eccentricities, and even $\lesssim 0.1$), as can be seen in the eccentricity distributions in Figures 3 and 13. Even with the inflated radii, allowing for the formation of lower-eccentricity WJs, low-eccentricity (<0.6) WJs can hardly be formed through inflated migration, and were likely formed in situ and/or through disk migration, generally consistent with the analysis by Anderson et al. (2020). We note that the paucity of >0.9 eccentricity suggested by Socrates et al. (2012), and ruled out by Dawson et al. (2015), does not constrain our models, which indeed show >0.9 WJs to be very rare. That being said, the apparent low frequency of 0.6–0.9 eccentricity WJs is a challenge to inflated eccentric migration, and any other eccentric migration model. In fact, this is a potentially more general difficulty—any successful model of HJ/WJ production should also be able to suppress HJ/WJ formation through the various types of eccentric migration.

On the theoretical front, neither the tidal interactions of gas giants nor the heat transfer to the core are understood, giving rise to large uncertainties in the evolutionary models. Here, we tried to bracket these potential caveats, but, naturally, better understanding of these processes is critical for the assessment and modeling of any eccentric migration model.

Future work on eccentric migration. Though we have focused on the role of inflated migration for specific types of eccentric migration (initially by planet–planet scattering, and considering some specific models for weak and dynamical tides), the same coupled evolution is important for any suggested eccentric migration model. Follow-up papers may consider other such models and their variants. In addition, accounting for the inflated sizes of gas giants in their early phases is also important for the increased likelihood of their physical collisions with other planets (due to their larger cross sections), which, in turn, can give rise to the heating and further inflation of the collision product (Lin & Ida 1997), which could then affect the migration and further collisions. Eccentric inflated migration could also be coupled to other processes that play a role in planet formation and dynamics, such as photoevaporation (e.g., Tripathi et al. 2015). Since the geometric cross sections of initially inflated planets are larger than the cross sections of the noninflated ones, the role of photoevaporation might change accordingly. Furthermore, although we have discussed the eccentric tidal migration channel in this paper, initially inflated gas giants could also be discussed in the context of disk migration, where the radius plays a role in the evolution, too.

Central heating. We can use our model to set constraints on the effective amount of energy penetrating to the center, since it will affect the contraction timescale and hence the migration timescale. The exact amount of the energy deposit and its distribution are still unknown, but by using our population synthesis—assuming a given distribution, e.g., for simplicity, that all the energy is deposited at the center—we can estimate the energy amount needed to explain the observational results. It should be noted that the parameter space is large and includes many degeneracies. For example, the choice of the initial eccentricity distribution might shift the distribution in a similar direction as external heating would do.

9. Summary

In this paper, we have proposed the formation channel of hot and warm Jupiters via inflated eccentric migration, and discussed the implications on the population. Here, we have focused on the semi-analytical approach, and discussed specific examples and a detailed population synthesis. We have compared the semi-analytical approach with the numerical planet evolution models described in paper II, and found that the semi-analytical model is in good agreement with the numerical model in all the regimes available for the numerical model. This allows us to make use of the more efficient semi-analytical model to explore the dynamics of large populations of planets as well as the resulting populations of HJs/WJs for a wide range of initial conditions.

Our models are general and are able to include in principle any kind of external heating/dynamical evolution. For brevity, we have demonstrated the use of the semi-analytical model for several examples only. We have presented specific examples for inflated eccentric migration and provided a detailed population synthesis study, based on the semi-analytical model. We have studied the dependence of the resulting HJ/WJ populations on the assumptions made regarding the tidal models and heat transfer processes, as well as the initial properties of the progenitor planets and their initial orbits.

Using the detailed semi-analytical model and population synthesis results, we have showed that inflated eccentric migration could significantly shorten the migration timescales of HJs and WJs, generally form WJs more efficiently, and give rise to enhanced rates of tidal disruptions of gas giants. We have also considered the effect of the external energy injected into the migrating gas giants on their final formed population, and found that it leads to enhanced tidal disruption, together with a smaller fraction of HJs and a larger fraction of WJs compared with models without efficient central heating.

Inflated eccentric migration leads to significant differences in the final distribution of the parameters as compared with noninflated models, and suggests that inflated migration plays an important role in migration and should generally be accounted for in any eccentric migration models (and possibly also in disk migration models).

We gratefully acknowledge fruitful discussions with Sivan Ginzburg, Thaddeus D. Komacek, Michelle Vick, Nicholas C.

Stone, and Eden Saig. M.R. acknowledges the generous support of the Azrieli fellowship.

ORCID iDs

Mor Rozner https://orcid.org/0000-0002-2728-0132
Hila Glanz https://orcid.org/0000-0002-6012-2136
Hagai B. Perets https://orcid.org/0000-0002-5004-199X
Evgeni Grishin https://orcid.org/0000-0001-7113-723X

References

Alexander, M. E. 1973, Ap&SS, 23, 459
Anderson, K. R., Lai, D., & Pu, B. 2020, MNRAS, 491, 1369
Armitage, P. J. 2010, Astrophysics of Planet Formation (Cambridge: Cambridge Univ. Press)
Arras, P., & Socrates, A. 2010, ApJ, 714, 1
Baruteau, C., Crida, A., Paardekooper, S. J., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al., 667 (Tuscon, AZ: Univ. Arizona Press)
Batygin, K., & Stevenson, D. J. 2010, ApJL, 714, L238
Beaugé, C., & Nesvorný, D. 2012, ApJ, 751, 119
Behroozi, P., & Peeples, M. S. 2015, MNRAS, 454, 1811
Berger, T. A., Huber, D., van Saders, J. L., et al. 2020, AJ, 159, 280
Bodenheimer, P., Hubickyj, O., & Lissauer, J. J. 2000, Icar, 143, 2
Bodenheimer, P., Lin, D. N. C., & Mardling, R. A. 2001, ApJ, 548, 466
Bodenheimer, P., & Pollack, J. B. 1986, Icar, 67, 391
Boss, A. P. 1997, Sci, 276, 1836
Bryan, M. L., Knutson, H. A., Howard, A. W., et al. 2016, ApJ, 821, 89
Butler, R. P., Wright, J. T., Marcy, G. W., et al. 2006, ApJ, 646, 505
Carrera, D., Raymond, S. N., & Davies, M. B. 2019, A&A, 629, L7
Chatterjee, S., Ford, E. B., Matsumura, S., & Rasio, F. A. 2008, ApJ, 686, 580
Darwin, G. H. 1879, RSPT, 170, 1
Dawson, R. I., & Johnson, J. A. 2018, ARA&A, 56, 175
Dawson, R. I., Murray-Clay, R. A., & Johnson, J. A. 2015, ApJ, 798, 66
Fabrycky, D., & Tremaine, S. 2007, ApJ, 669, 1298
Ford, E. B., & Rasio, F. A. 2008, ApJ, 686, 621
Ginzburg, S., & Chiang, E. 2019, MNRAS, 490, 4334
Ginzburg, S., & Sari, R. 2015, ApJ, 803, 111
Ginzburg, S., & Sari, R. 2016, ApJ, 819, 116
Ginzburg, S., & Sari, R. 2017, MNRAS, 469, 278
Glanz, H., Rozner, R., Perets, H. B., & Grishin, E. 2021, arXiv:2111.12714
Goldreich, P., & Soter, S. 1966, Icar, 5, 375
Grishin, E., Perets, H. B., & Fragione, G. 2018, MNRAS, 481, 4907
Guillochon, J., Ramirez-Ruiz, E., & Lin, D. 2011, ApJ, 732, 74
Guillot, T., Burrows, A., Hubbard, W. B., Lunine, J. I., & Saumon, D. 1996, ApJL, 459, L35
Hamers, A. S., Antonini, F., Lithwick, Y., Perets, H. B., & Portegies Zwart, S. F. 2017, MNRAS, 464, 688
Hamers, A. S., & Tremaine, S. 2017, AJ, 154, 272
Huang, C., Wu, Y., & Triaud, A. H. M. J. 2016, ApJ, 825, 98
Hut, P. 1981, A&A, 99, 126
Ivanov, P. B., & Papaloizou, J. C. B. 2004, MNRAS, 347, 437
Ivanov, P. B., & Papaloizou, J. C. B. 2007, MNRAS, 376, 682
Jurić, M., & Tremaine, S. 2008, ApJ, 686, 603
Kippenhahn, R., Weigert, A., & Weiss, A. 2012, Stellar Structure and Evolution (Berlin: Springer), 2012
Kiseleva, L. G., Eggleton, P. P., & Mikkola, S. 1998, MNRAS, 300, 292
Komacek, T. D., Thorngren, D. P., Lopez, E. D., & Ginzburg, S. 2020, ApJ, 893, 36
Komacek, T. D., & Youdin, A. N. 2017, ApJ, 844, 94
Kozai, Y. 1962, AJ, 67, 591
Lai, D. 1997, ApJ, 490, 847
Lidov, M. L. 1962, P&SS, 9, 719
Lin, D. N. C., & Ida, S. 1997, ApJ, 477, 781
Mardling, R. A. 1995a, ApJ, 450, 722
Mardling, R. A. 1995b, ApJ, 450, 732
Marley, M. S., Fortney, J. J., Hubickyj, O., Bodenheimer, P., & Lissauer, J. J. 2007, ApJ, 655, 541
Miller, N., Fortney, J. J., & Jackson, B. 2009, ApJ, 702, 1413
Mizuno, H. 1980, PThPh, 64, 544
Moe, M., & Kratter, K. M. 2018, ApJ, 854, 44
Nagasawa, M., Ida, S., & Bessho, T. 2008, ApJ, 678, 498
Naoz, S., Farr, W. M., Lithwick, Y., Rasio, F. A., & Teyssandier, J. 2011, Natur, 473, 187
Ogilvie, G. I. 2014, ARA&A, 52, 171
Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3
Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4
Perri, F., & Cameron, A. G. W. 1974, Icar, 22, 416
Petrovich, C. 2015a, ApJ, 799, 27
Petrovich, C. 2015b, ApJ, 805, 75
Portegies Zwart, S., McMillan, S., Harfst, S., et al. 2009, NewA, 14, 369
Press, W. H., & Teukolsky, S. A. 1977, ApJ, 213, 183
Rafikov, R. R. 2005, ApJL, 621, L69
Rasio, F. A., & Ford, E. B. 1996, Sci, 274, 954
Sarkis, P., Mordasini, C., Henning, T., Marleau, G. D., & Mollière, P. 2021, A&A, 645, A79
Socrates, A., Katz, B., Dong, S., & Tremaine, S. 2012, ApJ, 750, 106
Spiegel, D. S., & Burrows, A. 2013, ApJ, 772, 76
Thorngren, D., Gao, P., & Fortney, J. J. 2019, ApJL, 884, L6
Tripathi, A., Kratter, K. M., Murray-Clay, R. A., & Krumholz, M. R. 2015, ApJ, 808, 173
Vick, M., & Lai, D. 2018, MNRAS, 476, 482
Vick, M., Lai, D., & Anderson, K. R. 2019, MNRAS, 484, 5645
von Zeipel, H. 1910, AN, 183, 345
Wang, J., Fischer, D. A., Horch, E. P., & Huang, X. 2015, ApJ, 799, 229
Weidenschilling, S. J., & Marzari, F. 1996, Natur, 384, 619
Wu, Y. 2018, AJ, 155, 118
Wu, Y., & Lithwick, Y. 2011, ApJ, 735, 109
Wu, Y., & Murray, N. 2003, ApJ, 589, 605
Wu, Y., Murray, N. W., & Ramsahai, J. M. 2007, ApJ, 670, 820
Youdin, A., & Mitchell, J. 2010, DPS Meeting, 42, 27.15
Zahn, J. P. 1977, A&A, 500, 121
Zhu, W., & Dong, S. 2021, ARA&A, 59, 42

The Astrophysical Journal, 931:11 (12pp), 2022 May 20 https://doi.org/10.3847/1538-4357/ac6807

Inflated Eccentric Migration of Evolving Gas Giants II – Numerical Methodology and Basic Concepts

Hila Glanz[1], Mor Rozner[1], Hagai B. Perets[1], and Evgeni Grishin[1,2]

[1] Technion—Israel Institute of Technology, Haifa, 3200002, Israel; Glanz@tx.technion.ac.il

[2] School of Physics and Astronomy, Monash University, Clayton 3800, VIC, Australia

Received 2021 November 28; revised 2022 April 4; accepted 2022 April 17; published 2022 May 19

Abstract

Hot and warm Jupiters (HJs&WJs) are gas-giant planets orbiting their host stars at short orbital periods, posing a challenge to their efficient in situ formation. Therefore, most HJs&WJs are thought to have migrated from an initially farther-out birth location. Current migration models, i.e., disk migration (gas-dissipation driven) and eccentric migration (tidal evolution driven), fail to produce the occurrence rate and orbital properties of HJs&WJs. Here we study the role of thermal evolution and its coupling to tidal evolution. We use `AMUSE`, a numerical environment, and `MESA`, planetary evolution modeling, to model in detail the coupled internal and orbital evolution of gas giants during their eccentric migration. In a companion paper, we use a simple semianalytic model, validated by our numerical model, and run a population-synthesis study. We consider the initially inflated radii of gas giants (expected following their formation), as well study the effects of the potentially slowed contraction and even reinflation of gas giants (due to tidal and radiative heating) on the eccentric migration. Tidal forces that drive eccentric migration are highly sensitive to the planetary structure and radius. Consequently, we find that this form of inflated eccentric migration operates on significantly (up to an order of magnitude) shorter timescales than previously studied eccentric-migration models. Therefore, inflated eccentric migration gives rise to the more rapid formation of HJs&WJs, higher occurrence rates of WJs, and higher rates of tidal disruptions, compared with previous eccentric-migration models that consider constant ∼Jupiter radii for HJ and WJ progenitors. Coupled thermal–dynamical evolution of eccentric gas giants can therefore play a key role in their evolution.

Unified Astronomy Thesaurus concepts: Exoplanet formation (492); Hot Jupiters (753); Exoplanet migration (2205); Exoplanet evolution (491)

1. Introduction

Gas-giant planets are thought to have formed from either core accretion, in which runaway gas accretion takes place onto the massive, $\sim 10 M_{\oplus}$, core (Perri & Cameron 1974; Bodenheimer & Pollack 1986), or from a direct collapse from the gas disk (Mizuno 1980; Boss 1997; Armitage 2010). However, as the efficiencies of both channels are greatly affected by the local environment properties, such as the temperature, density, composition, and velocities, they cannot solely describe the formation of gas giants that have extremely short-period orbits around their host stars. These include the population of Hot Jupiters (HJs), with orbits of a few days (Bodenheimer et al. 2000; Rafikov 2005), as well as some warm Jupiters (WJs) with small pericenters. The remaining nonnegligible fraction of more distant WJs might still be formed in situ, as discussed in Huang et al. (2016) and Anderson et al. (2020). Consequently, HJs&WJs are thought to have formed at larger separations from their stars and migrated inward due to dynamical interactions, either with other bodies leading to high-eccentricity migration or with the gas from the protoplanetary disk, producing a drag force (see Dawson & Johnson 2018 for a review). Nevertheless, past studies on these migration models could not reproduce the observed formation rates and properties of the current population of HJs&WJs (Dawson & Johnson 2018; Zhu & Dong 2021), as the typical migration timescales are potentially too long to produce the inferred numbers of HJs/WJs and their appropriate timescales.

A planet orbiting its host star at close separation experiences significant tidal forces raised by the host star. The gravitational interaction between the star and the bulge raised due to tides on the planet (and to a much lesser degree the tides raised on the star by the planet) eventually gives rise to the dissipation of orbital energy in the planetary atmosphere. This, in turn, leads to the orbital decay of the planet into shorter periods and more circular orbits. Consequently, planets on highly eccentric orbits with a close pericenter approach to the host star may experience tidal migration, generally termed eccentric migration.

The strength of the tides strongly depends on the planetary radius, which is typically considered as some constant ∼Jupiter radius, R_J, in eccentric-migration models. However, gas giants are thought to form with far larger inflated radii and then cool and contract, where external heating by radiation and/or tidal heating may slow down their cooling and possibly even reinflate them. After reaching large radii of up to $10R_J$ by the end of the core accretion (Ginzburg & Chiang 2019), the gas giants contract to smaller radii, initially in a rapid process to $4R_J$ (Guillot et al. 1996), followed by a slower thermal contraction within a Kelvin–Helmholtz timescale ($\sim 10^8$ yr), reaching radii of ~ 1.5–$2.5R_J$ that continuously shrink in an even slower rate, depending on their mass and external energies. Because tidal migration depends strongly on the planetary radius, inflated planets could give rise to far faster eccentric migration compared with nonevolving constant-Jupiter-radius gas giants typically considered in eccentric-migration models. As the radius of such planets might decrease in a comparable timescale to the high-eccentricity migration timescale, considering the internal evolution of the planet (initially thermal contraction and cooling) can therefore play a key role in their dynamical evolution.

Here and in a companion paper (Rozner et al. 2021, hereafter Paper I) we explore for the first time a self-consistent thermal–dynamical evolution of migrating planets over a wide parameter space and throughout their evolution beginning at very high eccentricities (but see Wu et al. 2007; Miller et al. 2009; Petrovich 2015, where some of these issues were partially studied). We couple the thermal evolution of gas giants and their dynamical evolution through eccentric tidal migration, as well as consider possible reinflation and slow contraction of the planets due to external heating sources. We find that the eccentric migration of such inflated Jovian planets, which we term inflated eccentric migration, significantly alters their dynamical evolution and could play a key role in any type of eccentric migration and, in particular, give rise to much (up to an order of magnitude) faster eccentric migration.

Here we present our numerical method, where we use `MESA` and `AMUSE` to accurately simulate the internal evolution of these planets during their migration. Our numerical results can be used to study the effect of other types of dynamical evolution and external energy sources. In Paper I, we present a semianalytical approach to simulate such a migration, where we use the same equations of motion but simple modeling of the internal/thermal evolution, which therefore requires less computational resources in order to be used. Here we present some comparisons between the results of both methods and find a good agreement. This also validates our use of the semianalytical approach in the study and characterization of a large population of HJ and WJ progenitors, which we present in Paper I.

In the next section, we describe our calculation method. We first discuss the considered external energy sources affecting the evolution of the giant planets (Section 2.1), then we explain the mechanism of high-eccentricity migration with different tide models in Section 2.2. Later (Section 2.3), we describe our numerical simulation method to couple the dynamical and tidal evolution of the planets with their thermal evolution. In Section 3, we present our results and their implications on the formation of HJs&WJs, followed by discussions in Section 4, and finally, we summarize in Section 5.

2. Methods

2.1. External Energy Sources

In the absence of any internal heating sources, following its formation and final runaway accretion stages, a newly born gas giant begins to continuously cool down and contract. However, a variety of external heating sources can affect the planet during its life. These can include heating the planetary surface through irradiation by its host star, tidal heating induced by the star when the planet migrates, or any other potential heating sources resulting from other interactions and dynamical processes (e.g., collisions with other planets, Lin & Ida 1997, which can affect the early stages of planetary evolution and growth). Here, we consider the evolution of fully formed planets after they had been excited to high eccentricity, such that they experience strong tidal interactions with the host star. Besides the initial excitation to high eccentricity, the definition of our initial conditions, and the tidal interaction with the host star, we assume that no further interaction with other stellar or planetary bodies occurs. Figure 2 demonstrates the fast contraction from the initially inflated radii to about $2R_J$, in less than a Myr, such that a scattering prior to this stage is less probable, and even in such cases, the binary would more likely be disrupted rather than rapidly migrate to produce an HJ/WJ (see the discussion on flow in parameter space in Paper I). Therefore, we begin our models after a gas giant has already finished the core accretion stages and any planetary-scattering epoch and reached the initial eccentricity for its migration. Generally, these processes are thought to have been finalized by the first few Myr of evolution. As we describe in Section 2.3, we examine different initial radii at the time of coupling, such that the external energies are included both during the rapid contraction shown in Figure 2 and after all initial models have already converged and continued on the same cooling timescale.

Hereafter we study the effects of two sources of external energy: tidal heating and irradiation flux from the host star, both of them taken into consideration self-consistently together with the migration of the planet toward the host star and the thermal cooling of the migrating planet.

The distribution of the heat from the different sources inside the planet depends on the specific mechanism and the internal structure of the planet. Irradiation flux heats the surface of the planet and dissipates to deeper layers, but tidal heating may cause a deformation of the internal structure and therefore can potentially heat deeper layers more efficiently. We define $r_{\rm ext}$ as the radial distance inside the planet in which most of the external energy source is deposited.

The irradiation luminosity (averaged over an orbital period) is deposited in the photosphere of the planet (i.e., $r_{\rm ext} = r_{\rm irr} = R_p$ where R_p is the radius of the planet), is given by:

$$\bar{L}_{\rm irr} = \frac{1}{\mathcal{T}}\int_{\mathcal{T}} L_{\rm irr}(r(t))dt = \left(\frac{R_p}{a}\right)^2 \frac{L_\star}{\sqrt{1-e^2}} \tag{1}$$

where $\mathcal{T}$ is the orbital period and $r(t)$ is the distance between the planet and its host star.

The energy from tidal heating is given by the tidal model, which determines its internal distribution (i.e., $r_{\rm ext} = r_{\rm tides}$). We discuss the different heat distributions in Sections 2.2.1 and 2.2.2 for the equilibrium and dynamical tides models. We explain our numerical method of the internal heat distribution in Section 2.3.

We find that due to the planet's own radiation and cooling of the planet, the effect of deposition of irradiation and/or tidal heating on the dynamical evolution is mostly negligible when the energy is deposited in the planetary photosphere. In this case, most of the deposited energy is quickly irradiated away and does not heat the planetary interior. Consequently, the planetary radius is not affected by the heating processes in this case nor does it affect the migration timescale. However, some processes, such as ohmic dissipation (Batygin & Stevenson 2010), can provide a channel for heat conduction into internal regions. Deeper deposition at the inner layers could lead to a much more significant effect, such that even 1% of the external energy deposited at the center of the planet could induce larger radii than R_J even after Gyr when the planet is already very close to its star (see Bodenheimer et al. 2001; Guillot & Showman 2002; Komacek et al. 2020 and references therein). When using $r_{\rm ext} = 0$ to deposit the energy around the center of the planet, and multiply the right side of Equation (14) by an efficiency parameter, we find that very high energy deposition in the core can indeed give rise to planetary inflation, as can be seen in Figures 5, 6, and 11, which in the case of strong inflation can lead to disruption. We further discussed this in Section 4.2; see also the semianalytic study in Paper I.

2.2. High-eccentricity Tidal Migration

Tidal migration occurs when strong tidal forces from the star act to exchange energy and angular momentum between the orbit and the planet, leading to the growth of tidal bulges and thus an orbital decrease. Given the strong dependence of tides on the distance from the host star (Equations (3) and (8)), efficient tidal migration requires a close approach of the planet to the star. If the planet is born far from the star, as expected for gas giants, a close approach can occur only if the planet resides in a highly eccentric orbit, for which the pericenter approach is close to the star for tidal effects to become significant. Therefore, one can divide high-eccentricity tidal migration into two separate stages: reducing the planet's angular momentum and reducing the planet's energy. In the first stage, the HJ/WJ progenitor, which is likely formed on a relatively circular orbit, is excited into an eccentric orbit via planet–planet scattering (Rasio & Ford 1996; Chatterjee et al. 2008; Jurić & Tremaine 2008), as we discuss here, or through other channels for eccentricity excitation such as via the Von-Ziepel–Lidov–Kozai (ZLK) mechanism and secular chaos (e.g von Zeipel 1910; Kozai 1962; Lidov 1962; Wu & Murray 2003; Fabrycky & Tremaine 2007; Nagasawa et al. 2008; Naoz et al. 2011; Wu & Lithwick 2011; Petrovich 2015; Hamers et al. 2017; Wu 2018). In the second stage, energy extraction via tides leads to migration and circularization of the planet's orbit. The energy extracted from the orbit during an orbital period is dissipated in the planet, affecting its overall luminosity, which can affect the internal structure as a result. Assuming a complete transfer from the orbital energy to the planet, the injected/incoming luminosity can be described as follows:

$$L_{\rm tide} = -\frac{E}{a}\frac{da}{dt}, \tag{2}$$

where E is the orbital energy and a is the semimajor axis.

Modeling tides in giant planets is not trivial, and its strong dependence on the internal structure of the planet, turbulent viscosity processes dissipating energy, and other physical aspects of the problem has induced some long-standing debates on the nature and specific properties of tidal dissipation. Here we adapt the widely used tidal model of weak/equilibrium tides (Darwin 1879; Goldreich & Soter 1966; Alexander 1973; Hut 1981) but also consider more briefly the importance of dynamical tides (e.g., Zahn 1977; Mardling 1995a, 1995b). The latter could be especially important and more efficient during the early migration phases when the planet's orbit is still highly eccentric, and in that sense, considering only the weak-tides model is potentially conservative in terms of the efficiency of eccentric migration (Lai 1997).

Here we present a general approach, which can account for any tide model and is demonstrated here using both equilibrium tides and dynamical tides. We note that other models, such as chaotic-dynamical tides (Vick & Lai 2018; Vick et al. 2019), are likely to further shorten the migration timescales; these are to be left for future works.

In the next subsections, we explain how we model the migration of a planet due to equilibrium and dynamical tides, where we describe the equations of motion and the corresponding heat that should be transferred to the planet.

2.2.1. Equilibrium Tide Model

In this tidal model, the gravity from the star raises tides on the planet, leading to the formation of an equilibrium bulge on the planet, which is treated as an external point mass along the calculation (Hut 1981). Due to the timescale involved in raising the bulge, and the spin of the planet, the bulge position lags with respect to the position of the star, and the mutual interaction of the stellar gravity and the bulge torques the planet. When the lag time between the objects is much smaller than the spin or orbital period of the planet, one can invoke the weak-tide approximation (Darwin 1879; Goldreich & Soter 1966; Alexander 1973; Hut 1981). Under the assumption of pseudo-synchronization (of the planetary spin and the orbit) and conservation of the angular momentum, the orbital-averaged time evolution of the eccentricity and semimajor axis is given by Hut (1981) and Hamers & Tremaine (2017):

$$\frac{da}{dt} = -21 k_{\rm AM} n^2 \tau_p \frac{M_\star}{M_p}\left(\frac{R_p}{a}\right)^5 a e^2 \frac{f(e)}{(1-e^2)^{15/2}}, \tag{3}$$

$$\frac{de}{dt} = -\frac{21}{2} k_{\rm AM} n^2 \tau_p \frac{M_\star}{M_p}\left(\frac{R_p}{a}\right)^5 e \frac{f(e)}{(1-e^2)^{13/2}}, \tag{4}$$

where $M_\star$ is the mass of the host star, and M_p, R_p, e, a, n, and Ω_p are the mass, radius, orbital eccentricity, orbital semimajor axis mean motion, and spin frequency of Jupiter correspondingly; $\tau_p = 0.66$ s is the planetary tidal-lag time, $k_{\rm AM} = 0.25$ is the planetary apsidal motion constant (Hamers & Tremaine 2017), and

$$f(e) := \frac{1 + \frac{45}{14}e^2 + 8e^4 + \frac{685}{224}e^6 + \frac{255}{448}e^8 + \frac{25}{1792}e^{10}}{1 + 3e^2 + \frac{3}{8}e^4}. \tag{5}$$

Here we ignore the influence of the tides on the host star, as these are typically negligible in comparison with the tides on the planet. The energy associated with the tides according to Equations (2) and (3) scales as R_p^5, leading to a very strong dependence of the migration timescale on the planet's radius. Consequently, the migration timescales of initially inflated gas giants should be shorter than the timescales of non-inflated gas giants with a constant R_J radius. We note that the contraction timescales are sufficiently long to maintain inflated gas giants throughout a significant part of their dynamical evolution, such that the initial radius of an HJ/WJ will leave a signature on its expected final parameters, which could be also observed.

We consider the location of the tidal bulge, given by Murray & Dermott (1999),

$$h_{\rm weak} = \frac{M_\star}{M_p} R_p \left(\frac{R_p}{a}\right)^3, \tag{6}$$

implying a peak of the external heat from tides at $r_{\rm tides} = R_p - h_{\rm weak}$ from the center of the planet.

2.2.2. Dynamical Tides

At very large eccentricities, tidal energy mostly dissipates near periastron, raising a large tidal bulge on the primary (the giant planet in our case). Consequently, such tidal evolution cannot be parameterized by its average over the entire orbit, as done in the equilibrium tide model (Moe & Kratter 2018). The energy associated with this tidal deformation might excite internal energy modes of the planet (mainly the fundamental f-mode), which might induce an enhanced response (Mardling 1995a, 1995b; Lai 1997; Ogilvie 2014), potentially leading to even more rapid circularization and migration of the planet. The eccentricity decay is accompanied by pseudo-synchronization with the angular frequency of the host star and the excitation of oscillations in the planet becomes less pronounced as the orbital eccentricity decreases. As a result, the energy dissipation by the various modes is gradually suppressed, until a transition to the regime in which equilibrium tides are more dominant (Mardling 1995b). The quadrupole order of the energy dissipation can be written as follows (Press & Teukolsky 1977; Moe & Kratter 2018):

$$\Delta E = f_{\rm dyn}\frac{M_\star + M_p}{M_p}\frac{GM_\star^2}{R_p}\left[\frac{a(1-e)}{R_p}\right]^{-9}, \tag{7}$$

with $f_{\rm dyn} = 0.1$, as Moe & Kratter (2018) (following the calculation of McMillan 1986) found in good agreement with observations of pre-Main-Sequence (pre-MS), which are approximated by the same polytropic index $n = 3/2$ assumed for gas giants. We note that these are estimates containing a large uncertainty, and the exact displacement of this energy is still not well understood. We test the implication of such a choice in the Results section.

At very high eccentricities, the orbital angular momentum can be very low, and one might not ignore the spin angular momentum as done in the above prescription of the equilibrium tides. Therefore, during the migration stage dominated by dynamical tides, one can assume conservation of the pericenter instead of the angular momentum (Moe & Kratter 2018). Combining this prescription with the equations of the orbital energy and angular momentum, and assuming a constant pericenter, leads to the following equations of the orbital semimajor axis and eccentricity along the migration (Moe & Kratter 2018):

$$\frac{da}{dt} = -\frac{a}{\mathcal{T}}\frac{\Delta E}{E}, \tag{8}$$

$$\frac{de}{dt} = \frac{1-e}{a}\frac{da}{dt}, \tag{9}$$

where $\mathcal{T} = 2\pi\sqrt{\frac{a^3}{G(M_\star + m_p)}}$ is the orbital period of the planet around its host star.

While dynamical tides dominate for large eccentricities, weak tides will be a more physical description for low ones (Mardling 1995b). The ratio of the migration rate due to dynamical tides to the migration rate due to weak tides is given by

$$\beta(R_p, a, e) \equiv \frac{da/dt|_{\rm dyn}}{da/dt|_{\rm weak}} = \frac{2f_{\rm dyn}R_p^3 A(e)}{21GM_p k_{AM}\tau_p\mathcal{T}}, \tag{10}$$

where

$$A(e) \equiv \frac{(1-e^2)^{15/2}}{(1-e)^9 e^2 f(e)}. \tag{11}$$

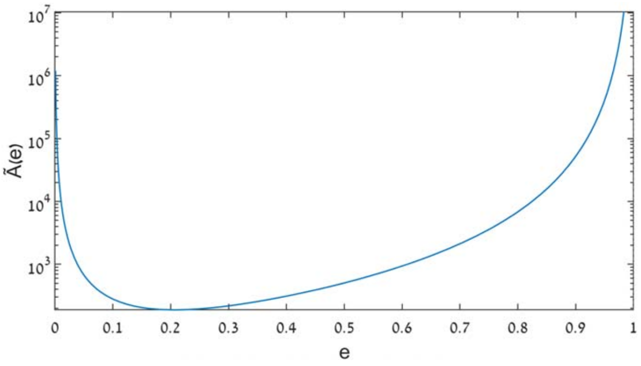

Figure 1. $\tilde{A}(e)$ from Equation (12) showing the dependence of the dynamical-to-weak-tide migration strengths ratio (β) on the eccentricity.

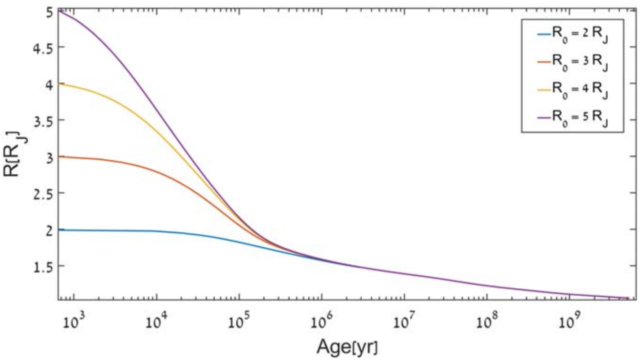

Figure 2. Contraction of a gas giant, modeled with `MESA` version 15140 (Paxton et al. 2013, 2018, 2019), with different initial radii, without deposition of external heat.

Because the pericenter is assumed to remain constant during the migration with dynamical tides, one can write

$$\beta(R_p, a, e) = \tilde{A}(e)\cdot R_p^3\cdot B(m_p,\ldots), \tag{12}$$

where $\tilde{A}(e) \equiv A(e)\cdot(1-e)^{-3/2}$. The transition between the dynamical and weak tides occurs roughly at $\beta \sim 1$, and we set a lower artificial cutoff at $e = 0.2$, at approximately the point where $\tilde{A}(e)$ gets its minimum value (see Figure 1). In this way, we avoid the divergence of dynamical tides at $e = 0$, and the transition occurs at $\max\{0.2, e|_{\beta=1}\}$. We note that considering migration due to dynamical tides with a very small pericenter such that B in Equation (12) is greater than the maximum value of $\frac{1}{\tilde{A}(e)\cdot R_p^3}$, leads to $\beta > 1$ for the entire migration until $e = e_{\rm trans}$. We further discuss the transition point between the different tide models in Section 3.2.

We note that at a sufficiently large eccentricity and low pericenter, the oscillatory modes inside the planet due to tides could grow chaotically (Ivanov & Papaloizou 2004, 2007; Vick & Lai 2018; Wu 2018) and can potentially increase the energy exchange and hence lead to faster migration and circularization. However, modeling such a scenario will be left for future work.

In our modeling, the energy from the dynamical tides is deposited into the planet's photosphere, i.e., $r_{\rm tides} = R_p$. However, more accurate future models might include different internal distributions, as the deposition heat from dynamical tides inside a planet is not yet understood (Sun et al. 2018) and is beyond the scope of this work. In Section 4.2, we briefly

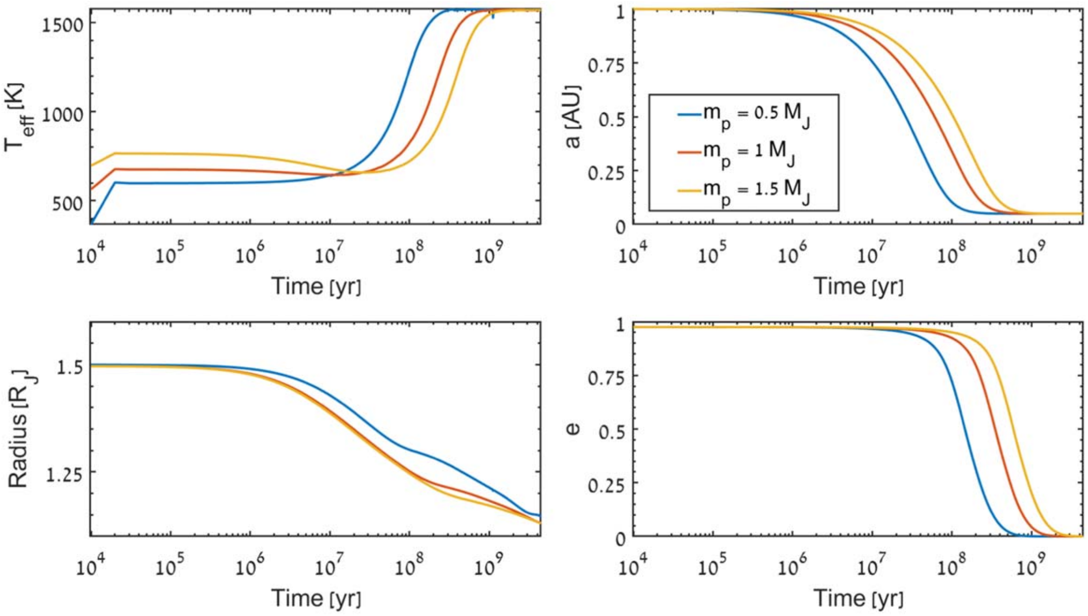

Figure 3. The numerical models' results of the thermal and orbital evolution of an HJ progenitor with a different initial mass, migrating due to weak tides, including irradiation and tidal heating. The initial radius is 1.5 R_J, the initial semimajor axis is 1 au, and the initial eccentricity is 0.975.

discuss the effect of a deeper deposition of the tides, which may arise from efficient ohmic dissipation.

We compare the evolution of inflated eccentric migrating planets due to two different tide models—weak and dynamical tides in Section 3.2—where one can notice the stronger effect of the dynamical tides not only on the migration process but on the planet's structure as well. In the same subsection, one can find a further discussion on the differences between the different tide models, in addition to the influence of different parameter choices for the dynamical tides model.

2.3. Numerical Coupling of the Thermal–Dynamical Evolution

In our numerical approach, we couple the orbital-averaged equations, derived from the tidal migration model, with numerical modeling of the internal evolution of the planet, affected by the deposition of heat and the cooling due to its own irradiation. We use the `AMUSE` framework (Portegies Zwart et al. 2009) (version 13.2.1 including self-contributions that should be available in future versions) to combine between the different codes. The internal evolution of our planets is modeled with the stellar evolution code `MESA` (Paxton et al. 2011, 2013) version 2208, which is a one-dimensional code that solves the stellar equations (Kippenhahn et al. 2012) assuming hydrostatic equilibrium and spherical symmetry. We use the OPAL/SCVH equations of state (Rogers & Nayfonov 2002) and opacity tables corresponding to the existence of molecules at low temperatures at the outer layers of the planets (Grevesse & Sauval 1998; Freedman et al. 2008). In order to simulate such planets, our code can combine any given opacity table by specifying the transition temperatures. In this way, one can consider available dust opacities as well as any future opacity tables that are relevant for planets and were not available in the original version of `MESA` within `AMUSE`. We begin by creating the initial planet model as a pre-main-sequence low-mass star, which has no nuclear burning and hence mimics the evolution of a planet. The planet model then contracts according to the equilibrium between its own gravity and thermal cooling, as can be seen in Figure 2. We thus evolve this model in isolation until it reaches the initial radius at which the migration process is assumed to begin, and from this point, we couple its further internal evolution (i.e., thermal cooling) to the dynamical evolution. When simulating the coupled migration process, after each orbital evolution step, we calculate the corresponding external heat source (extra heat as termed in `MESA` Paxton et al. 2013), $L_{\rm ext}$, such that the energy equation of the planet at each radial distance from its core becomes

$$\frac{dL}{dm} = -T\frac{ds}{dt} + \frac{dL_{\rm ext}}{dm}, \tag{13}$$

where T is the temperature of the model and s is the specific entropy, and $L_{\rm ext}$ is calculated according to the orbital parameters, through the dependency of tides and irradiation, described in the previous sections. In order to include the external heat term via `AMUSE`, we updated the current interface to support the inclusion of any external heat distribution during the evolution with `MESA`.

We consider a heating source that deposits its energy at some typical region inside the planet, at a distance of $r_{\rm ext}$ from the planet's center. Owing to the lack of known distribution of the heat, we adapt a Gaussian heat distribution similar to Spiegel & Burrows (2013) and Komacek & Youdin (2017):

$$\frac{dL_{\rm ext}}{dm} = \frac{L_{\rm ext}}{\sqrt{2\pi\sigma_{\rm ext}^2}} \exp\left[-\frac{1}{2}\left(\frac{r - r_{\rm ext}}{\sigma_{\rm ext}}\right)^2\right]\frac{dr}{dm}, \tag{14}$$

where $dr/dm = (4\pi\rho r^2)^{-1}$ and $\sigma_{\rm ext} = 0.5H_{p,\rm ext}$ is half the scale height, computed at $r_{\rm ext}$ according to Paxton et al. (2011):

$$H_p(r) = \begin{cases} \min\left\{\frac{P}{\rho g}, \sqrt{\frac{P}{G\rho^2}}\right\} & , r \neq 0 \\ R_p & , r = 0 \end{cases}, \tag{15}$$

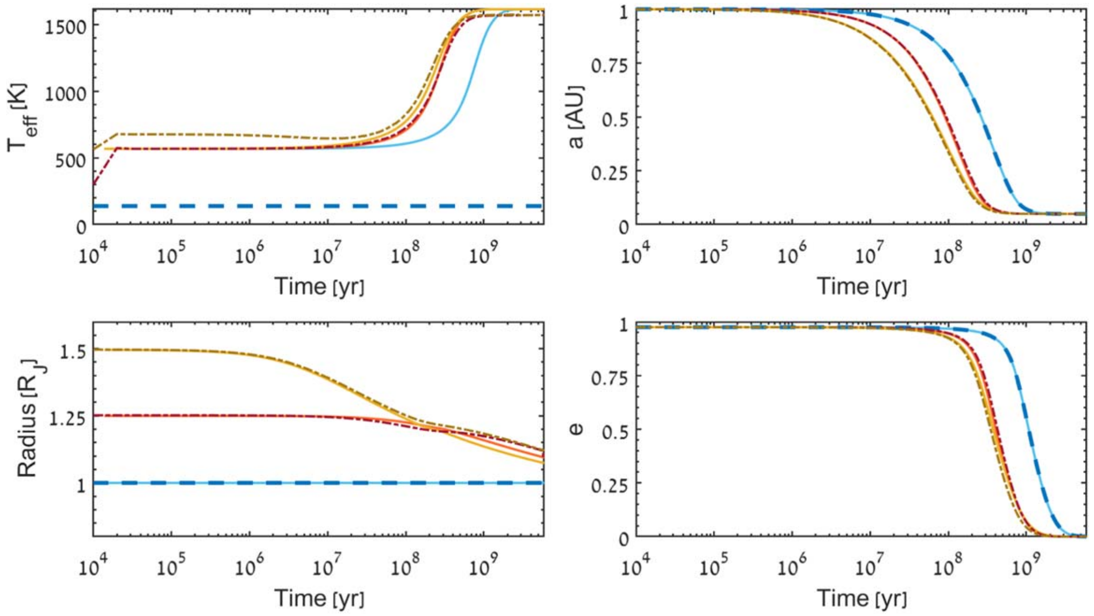

Figure 4. Same as Figure 3, but all with same mass 1 M_J and different initial radii. The blue line corresponds to the migration of a gas giant assumed to have a constant 1 R_J throughout its evolution, without consideration of any thermal evolution, as typically done in eccentric-migration models. The solid lines correspond to the results from the semianalytical model and the dashed lines to results from the numerical model. In light blue is the semianalytical simulation of a gas giant with an initial 1 R_J radius but now considering its thermal evolution.

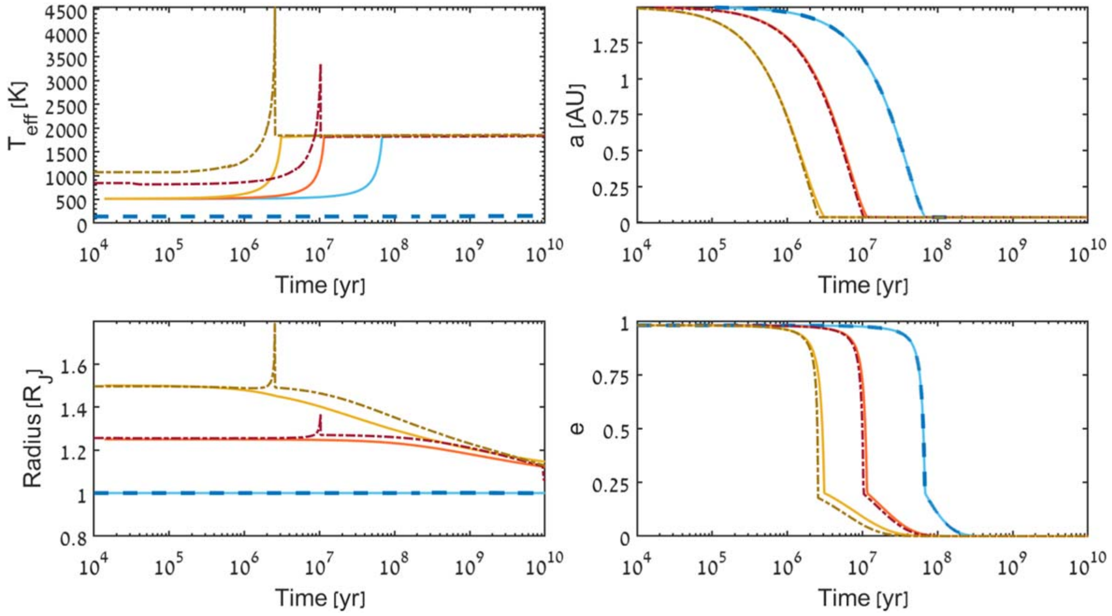

Figure 5. The thermal and orbital evolution of an HJ progenitor migrating through dynamical tides. The initial semimajor axis is $a_0 = 1.5$ au, and the initial eccentricity is $e_0 = 0.98$. Blue lines show the eccentric migration of a constant $1R_J$ planet, without including its internal evolution; orange and yellow correspond to models with $R_0 = 1.25R_J$ and $R_0 = 1.5R_J$, in which we include the thermal evolution affected by both irradiation from the star and tidal heating. Dashed lines are the results of our numerical model compared with the straight lines derived with the semianalytical approach presented in Paper I, where light blue corresponds to the semianalytical model of a 1 R_J planet when including thermal evolution.

where P and ρ are the pressure and density distributions of the planet. Because this distribution depends on the pressure and density profiles, a deposition at a higher location in the atmosphere, at low-pressure regions, is likely to have only a little influence on the interior of the planet, whereas deposition of energy directly into the core significantly affects the evolution of the planetary structure. Integrating over Equation (14) gives approximately the same amount of heat as L_{ext}. We note that the sum of this discrete numerical distribution over all shells might be different (lower) than the total heat calculated from Equations (1) and (2) when using a Gaussian distribution around some radial distance inside a spherical model with a discrete mass distribution. The change we find is by a factor of 2 at most, i.e., this can effectively be

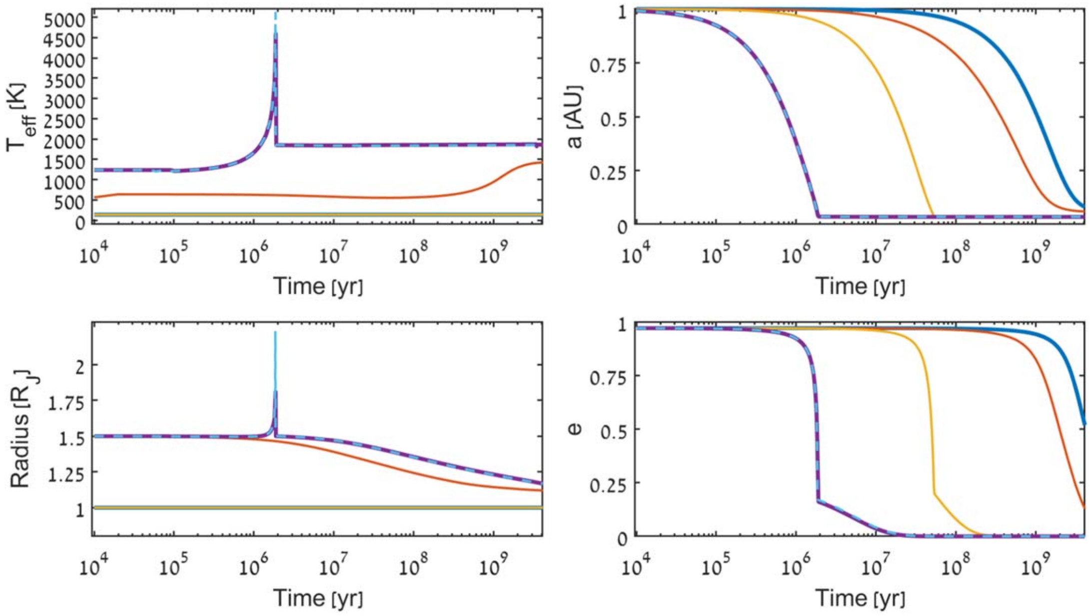

Figure 6. Different models of migrating planets with 1 M_J and initial orbital parameters of $a_0 = 1$ au and $e_0 = 0.97$. Both models with initial $R_0 = 1$ R_J (blue and yellow) are evolved only dynamically, with a constant radius; other models are evolved according to the different tide models when including tidal heating and irradiation from the host star during their migration. Purple and yellow lines correspond to models evolved with dynamical tides; orange and blue correspond to equilibrium tide models. The light-blue dashed line also evolves under the dynamical tides, but when the tidal heating is injected at a depth, it equals the size of the bulge for both tide models.

translated to a lower efficiency of the heat conductance to the central parts. Other distributions can be chosen and easily used by changing the heat distribution.

The time of each step in our simulation is chosen to be much shorter than the typical timescales for the orbital/thermal changes. At each step, we use the current properties of the planet and evolve the orbital parameters according to the specific tides model (i.e., Equation (3) or (8)).

The amount of deposited energy by tides is derived from the tides model and is deposited in the planet; see Equation (2). We assume the energy is deposited at a typical radius of the planet, as discussed above, and smooth it as a Gaussian distribution of the corresponding heat (Komacek & Youdin 2017) $L_{\rm ext} = L_{\rm tides}$, with a pick at $r_{\rm ext} = r_{\rm tides}$, inside the `MESA` model, as described in Equation (14), where the exact $r_{\rm tides}$ depends on the tides model, as described in Section 2.2. The irradiation flux from the host star, which changes due to the orbital evolution throughout the migration process, is described in Equation (1). The corresponding heat is distributed in the photosphere of the planet, using $r_{\rm ext} = R_p$ in Equation (14).

After injecting both tidal and radiation energies, i.e., $L_{\rm ext} = L_{\rm tides} + L_{\rm irr}$, we evolve the planet model with `MESA` for the same duration as was done for the orbital evolution. Our simulations terminate when one of the following conditions is fulfilled: (1) the planet has passed its Roche limit, defined as $r_t = \eta R_p \left(\frac{M_\star}{M_p}\right)^{1/3}$ (Guillochon et al. 2011), and cannot survive in its current condition, (2) the evolution time has passed the Hubble time, (3) the planet has cooled and contracted to levels such that currently used finite opacity and equations of states (EOS) tables, as well as other parameters in the version of `MESA` used, are no longer adequate. Satisfying the last criterion (3) means that the contraction timescale of the current model (including external heating) is much shorter than the migration timescale. During the migration, this termination condition was achieved only in some of our simulations that produced WJs. In such cases, the numerical model cannot be compared with the semianalytic model throughout the evolution. Modeling of these regimes can, however, be followed in the semianalytical model (see Paper I).

Our numerical model has been developed such that one can choose different orbital evolution models and different internal evolution codes. Using the current `MESA` module, one can follow the evolution of an externally built `MESA` model and include any external heat sources distributed around any desired location according to Equation (14).

3. Results

In the next section we present the results of our numerical simulations of the inflated eccentric migration of gas giants. We simulated different candidate models producing HJs&WJs, where we tested the effect of the different heat sources on their migration under the different tide models.

In Paper I we present a simpler semianalytical model that uses the same equations of motion for the orbital change but is coupled with equations approximating the thermal/radius change, instead of coupled to the much more detailed, yet computationally expensive, numerical model of the internal evolution as done here. We compare the results of the semianalytical model to those of our numerical model described here and find excellent agreement (see also Paper I).

3.1. *Hot and Warm Jupiter Candidates with Different Initial Properties*

The thermal–dynamical evolution is affected by the different properties of the migrating planet: its mass, radius, and the internal distribution of heat from external energy sources.

Observations show that the majority of the giant-planet population have mass in the range of 0.1–10 M_J (Butler et al. 2006).

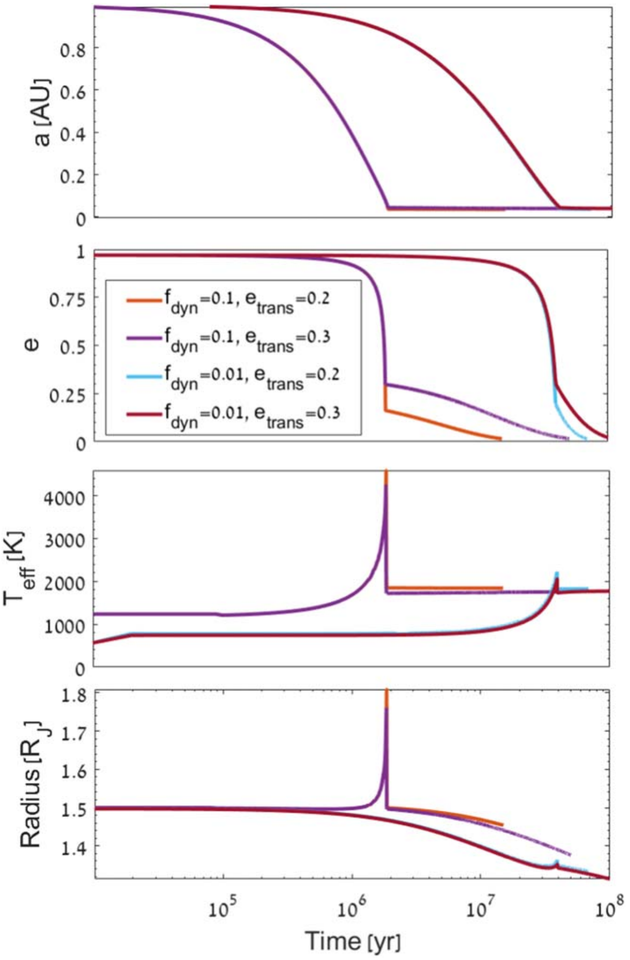

Figure 7. Migrating planets with 1 M_J and initial orbital parameters of $a_0 = 1$ au, $e_0 = 0.97$, and $R_0 = 1.5R_J$, affected by dynamical tides and irradiation from a Sun-like host star.

In Figure 3 we present a comparison between the evolution of planets of different masses according to the equilibrium tides model, showing that planets with lower masses migrate faster. The slower contraction of the radii shows inflated eccentric migration to have an even larger impact in these cases when considering the formation of low-mass HJs&WJs. This can be explained by the opposite dependence of the weak-tides EOS on the planetary mass (Equation (3)). In this case, the resulting giants have migrated to become HJs with a final orbital period of ∼4 days.

The strong dependence of tides on the radius of the migrating object implies a faster migration for larger initial radii. In Figure 4 we compare the dynamical and thermal evolution of gas planets, considering weak tides, which are initialized with different radii. We notice that an initially more inflated planet can migrate an order of magnitude faster, or even more, than a planet with a constant 1 R_J. We discuss the possible implications of the different assumptions regarding the initial radii on the HJ and WJ population in Paper I.

3.2. Different Tide Models

The equations of motion of a migrating planet due to equilibrium tides are derived with the assumption of a small eccentricity. As was described in Section 2.2.2, high eccentricities are likely to excite additional modes inside the planet that can lead to even larger effect on the migration process. In Figure 5 we show the migration of gas-giant planets with different initial radii but now affected by dynamical tides, compared with the migration of a constant 1 R_J radius gas giant. The migration of the initially inflated planets is indeed shorter by more than an order of magnitude compared with the migration of the constant radius planet.

In Figure 6 we compare the evolution of migrating planets when considering the two different tidal models, where one can notice the larger effect of the dynamical tides, with a greater dependence on the planetary radius.

As described in Section 2.2.2, the dynamical tides model still has many uncertainties, among them are the value of $f_{\rm dyn}$ in Equation (8) (McMillan 1986; Mardling 1995b; Lai 1997) and the transition point to equilibrium tides (Moe & Kratter 2018; Grishin & Perets 2022). When using a value of $f_{\rm dyn} = 0.1$, and if $\beta > 1$ all the way to $e_{\rm trans}$, the energy associated with the dynamical tides increases rapidly as the semimajor axis decreases. In this case, even when the heat is deposited only at the planet's photosphere, sufficient heat is transported to the central part as to give rise to a radial expansion (reinflation) of the gas giant (see Figures 5, 6, and 7). This can be seen in Figure 1, showing the dependence of β on the eccentricity, which has a minimum at 0.2, and goes to infinity for $e \to 0$ and $e \to 1$. Because Equations (7) and (8) are no longer valid at $e = 0$, one must use another condition to cease the dynamical tides prior to this point; in correspondence with Figure 1, we choose a lower limit of the eccentricity with dynamical tides between 0.2 and 0.3. However, as there must be a smooth transition into equilibrium tides prior to circularization, the large jump in the tidal energy during the transition is probably not physical. In our model, we consider the heat associated with the equilibrium tides to be deposited at a depth of the bulge height, while the heat from the dynamical tides, which might be larger, is deposited around the surface, such that the depth of energy deposition changes after the transition to a slightly deeper layer and a lower luminosity for the weak tides, as can be seen in Figure 10. In Figure 6 we also examine the evolution when $r_{\rm ext}$ always equals the difference between R_p and the height of the bulge, also during the influence of the dynamical tides (light-blue dashed line), and because the bulge is much smaller than the size of the planet, and therefore very close to the surface, the difference, compared with the evolution when dynamical tides are deposited around the surface (purple line), is relatively negligible. The same behavior can be seen in Figure 11, when the external heat is deposited around the core for the entire evolution, which does not affect the dynamical evolution. In addition, as can be seen in Figure 7, a lower value of $f_{\rm dyn}$ leads to a smoother transition, with a lower impact of the dynamical tides on the migration. Figure 7 shows only a minor difference in the eccentricity evolution between $e_{\rm trans} = 0.2$ and $e_{\rm trans} = 0.3$, but a significant difference in the migrations with the different efficiency parameters ($f_{\rm dyn}$). We compare the effect of the efficiency parameter on the overall formation rate of the HJ and WJ population in Paper I, showing an increase in both populations when using larger values of $f_{\rm dyn}$. We note that very large dynamical tides can lead to planet disruption, even when still not reaching the tidal radius.

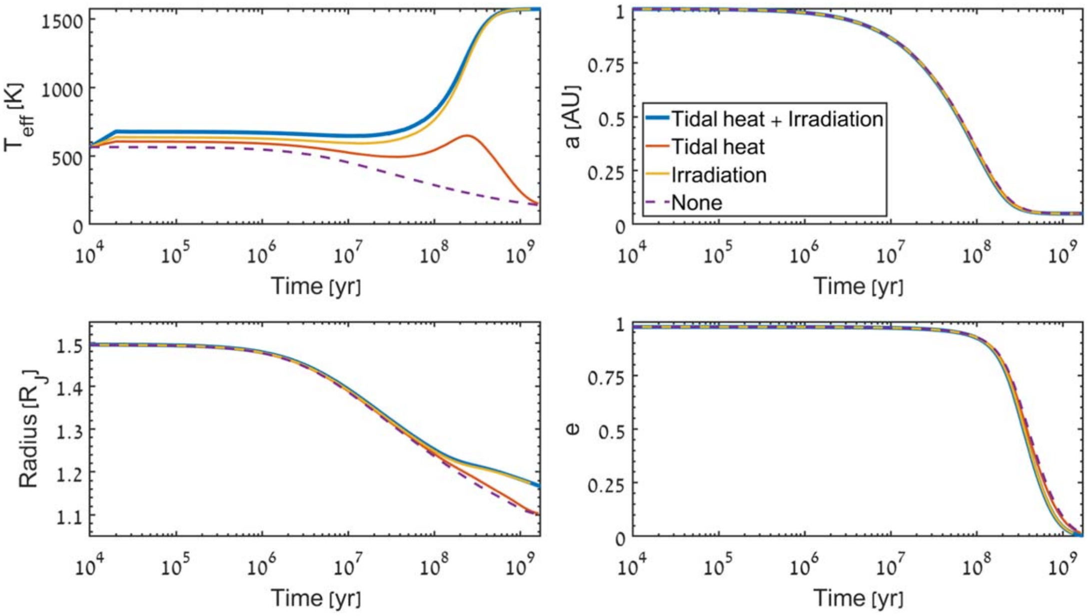

Figure 8. Comparison of the thermal and orbital evolution of an HJ candidate migrating through weak tides. We consider the following initial conditions: initial radius of 1.5 R_J, initial semimajor axis of 1 au, and initial eccentricity of 0.975 (partially presented in Figure 4), and we consider different external energy sources. Red lines correspond to evolution without any external energies. The orange line corresponds to evolution with (weak) tidal heating, the blue line to evolution with irradiation from the star, and the green line model includes both irradiation from the star and (weak) tidal heating. All the lines were produced with the numerical approach.

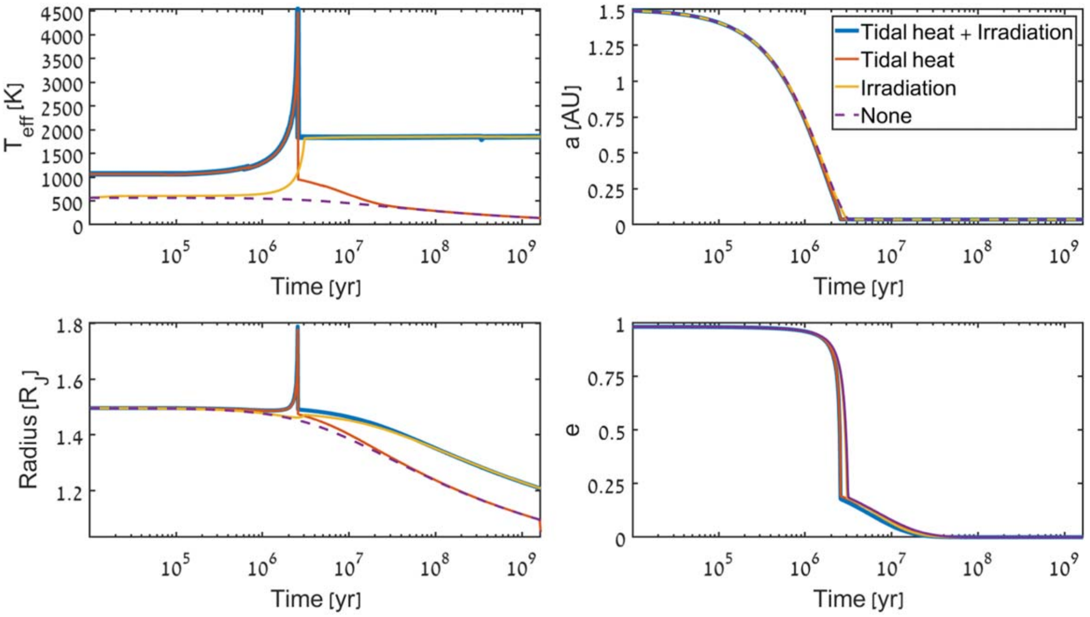

Figure 9. Comparison of the thermal and orbital evolution of an HJ candidate due to dynamical tides. We consider the following initial conditions: initial radius of 1.5 R_J, initial semimajor axis of 1.5 au, and initial eccentricity of 0.98 (partially presented in Figure 5). Different external energy sources are considered. The red line corresponds to evolution without any external energies; the orange one includes (dynamical) tidal heating, the blue line includes irradiation from the star, and the green line includes both irradiation from the star and (dynamical) tidal heating. All the lines were produced with the numerical approach.

3.3. The Effect of Different Energy Sources

Here and in Paper I we considered the influence of two different external heat sources on the migration of HJ and WJ candidates—irradiation from the host star and the energy from the tides acting on the planet. Our approach allows the inclusion of any external energy source and couples its effects to both the thermal and dynamical processes. In Figure 8 and Figure 9, we demonstrate the differences in the migration of a gas giant with both tide models described in Section 2.2 when including the different combinations of heat sources in the thermal evolution of the planet. Both figures show the importance of the irradiation energy to achieve the observed effective temperature range of such planets. The effect of tidal heating in these two cases is very minor in terms of the final properties of the planet, when the migration terminates. However, as was stated in Section 2.2.2, the exact deposition of the dynamical tides inside the planet is unknown, in edition

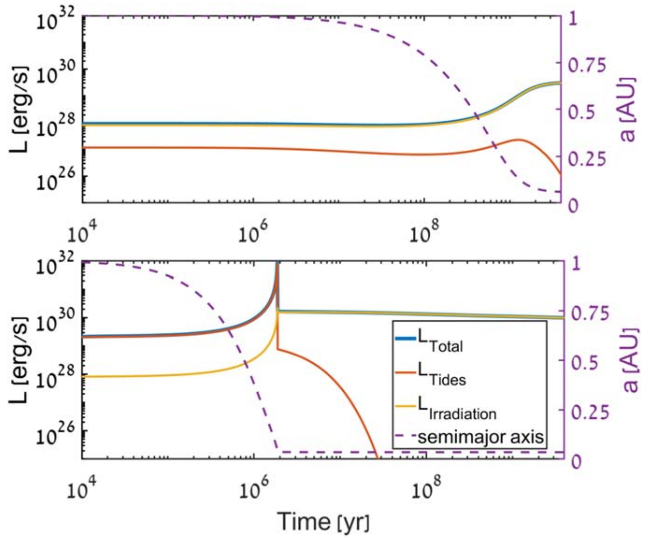

Figure 10. The calculated external heat sources along the migration of a 1 M_J planet with an initial radius of 1.5 R_J, initial semimajor axis of 1 au, and initial eccentricity of 0.97, under the influence of equilibrium tides (upper panel) and dynamical tides (lower panel); simulations are presented in Figure 6.

to its efficiency (i.e., $f_{\rm dyn}$) and the exact transition to equilibrium tides ($e_{\rm trans}$). Therefore, the effect of the dynamical tides can be greater if the efficiency parameter is larger, as well as when the transition to weak tides occurs at a lower eccentricity. In the upper panel of Figure 10 we see that the radiation is indeed the most dominant along the entire evolution. On the other hand, the lower panel of the same Figure 10 shows an example where the dynamical tides play the dominant role down to the transition point at $e_{\rm trans} = 0.2$. In this case, we used a value of $f_{\rm dyn} = 0.1$, while a larger value would increase this effect, as can be seen in Figure 7.

We emphasize that there are many uncertainties regarding the deposition of heat by dynamical tides, which generally affect any eccentric-migration model and the structure of tidal-migrating planets. We discuss the possible implications of the different assumptions on the HJ&WJ population in Paper I. However, an in-depth study of the exact behavior of dynamical tides and their workings is beyond the scope of this paper, while here we consider several models and bracket their general implications.

4. Discussion

4.1. Inflated Hot Jupiters and Heat Transfer

Although our study focuses on the early evolution of migrating Jupiters, at which time they still retain large inflated radii following their initial formation, observations show the existence of at least some older inflated HJs, even at Gyr timescales. The abundance of such inflated HJs was suggested to indicate that an external source of deposited energy is required in order to keep HJs at an inflated phase or to reinflate them after they already contracted (Guillot & Showman 2002; Baraffe et al. 2010; Thorngren & Fortney 2018). Several external energy sources and/or processes that conduct heat from the outer layers of the planet to the interior part (hence keeping the planets hotter) were suggested as a solution to the inflation (Ginzburg & Sari 2015). These include tidal heating (Bodenheimer et al. 2001), ohmic heating (Batygin & Stevenson 2010), and irradiation from the star (Burrows et al. 2007). However, there is still no consensus on the origin of the population of such old inflated HJs. Nonetheless, because observed WJs are usually not inflated (Miller & Fortney 2011), this energy should potentially relate to the orbital separation from the host star or to the migration timescales, which are correlated with the orbital energy and angular momenta of the migrating planet. Depending on the energy source, its duration, and its strength, it could potentially affect the migration process and shorten it. More generally, if other processes exist that keep planets inflated, i.e., leading to even longer contraction timescales, our suggested inflated eccentric migration should be even more efficient than already suggested by our results.

4.2. Internal Distribution of Energy

When considering tidal heating and irradiation flux around the migrating planet's photosphere, we find that due to the efficient radiation of this energy, the effect on the migration is mostly negligible (though it does determine the planet's effective surface temperature).

However, as can be seen in Figure 11, if the energy is deposited at a deeper region (when using a smaller $r_{\rm ext}$ in Equation (14)), the planet may slow its contraction or even reinflate, thus its migration will be further accelerated. One can see that even a low efficiency of heat conductance to the center of the planet of only 1% of the energy is distributed around the center of the planet. When multiplying the expression of $dL_{\rm ext}/dm$ in Equation (14) by 0.01, the planet's radius can increase and affect the migration time. We note that the strength of the dissipation of irradiation energy on the planet slightly varies due to the change in the mass distribution of the model and therefore the change in $\sigma_{\rm ext}$ and the result of the integral in Equation (14).

4.3. Formation of Warm Jupiters

Inflated eccentric migration enhances the migration rate such that planets that could not become HJs/WJs when considering an initial and constant 1 R_J radii, migrate more efficiently and now become WJs. Furthermore, inflated WJs, given the same initial conditions, would be less eccentric because they proceed faster in their migration; some of the expected WJs from the 1 R_J case will turn out to be HJs because their inflated migration sped up to enable that. Figure 12 shows the evolution of two models of migrating WJ candidates during a Hubble time, where one outcome can be considered as a WJ (initialized with $R = 1R_J$) and the other already as an HJ (initialized with $R = 3R_J$). As both migration cases have not yet terminated at a Hubble time, one can deduce that ongoing star formation will enlarge the fraction of eccentric WJs, which are effectively on their way to become HJs on longer timescales. The fractions of WJs decay with time and the fraction of HJs increases, as WJ candidates end as HJs if their migration is efficient enough. Therefore, star formation gives rise to an increment in the fraction of WJs on account of the fraction of HJs.

5. Summary

In this paper, together with Paper I, we proposed a new efficient model for the formation of HJs and WJs by considering the radial/thermal evolution of the originally inflated planet along its migration. Here we used `AMUSE` (Portegies Zwart et al. 2009) to couple numerically the dynamical evolution of such planets according to different

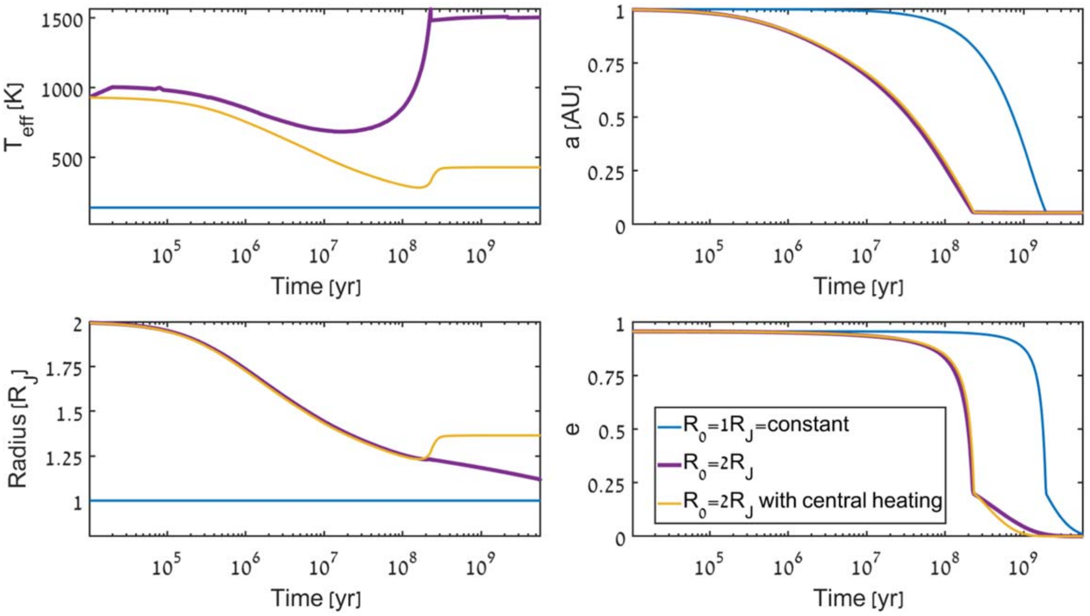

Figure 11. Comparison of the migration of a constant ($R_0 = 1R_J$) radius Jupiter-mass planet (blue) and the migration of a similar planet that is also affected by dynamical tides and irradiation (purple) with a constant $R_0 = 1\ R_J$ (blue). The injection of external heat is distributed according to Equation (14). The yellow line corresponds to the case where only 1% of the external heat is injected, but it is now distributed around the center of the planet, i.e., at $r_{\rm ext} = 0$. In all simulations, $a_0 = 1$ au and $e_0 = 0.955$.

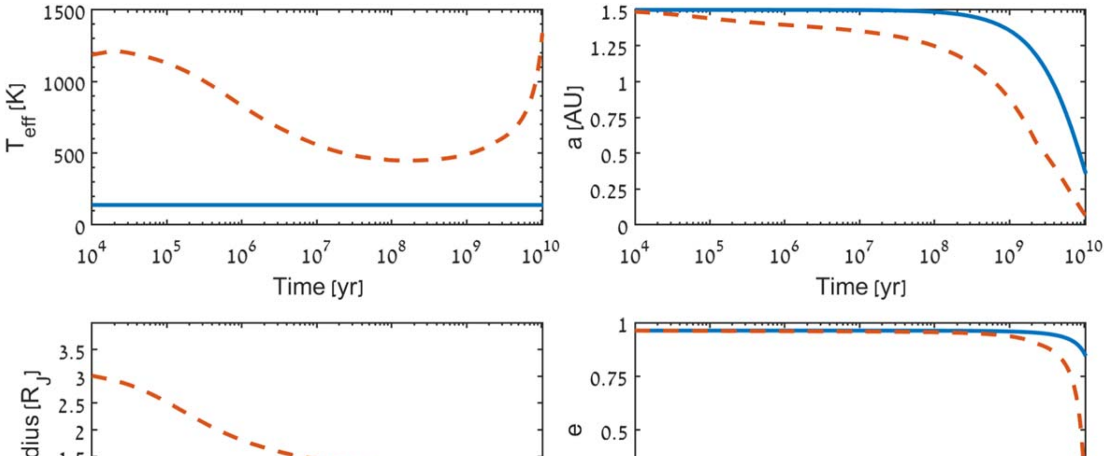

Figure 12. The thermal and orbital evolution of HJ and WJ candidates migrating due to dynamical tides, initialized with a semimajor axis of 1.5 au and initial eccentricity of 0.963. The blue line shows the dynamical evolution of a constant $R_0 = 1\ R_J$, whereas in orange is the evolution (thermal and dynamical) of $R_0 = 3\ R_J$ affected by both irradiation and tidal energies. The initial 3 $R_{\rm J}$ model reached an orbital period of ~ 6.4 days after a Hubble time and became an HJ, and the constant $1R_{\rm J}$ finalized at an orbit of 24.2 days, in the WJ regime.

tidal models with their internal evolution along the migration process (using `MESA`; Paxton et al. 2011).

Here and in Paper I, we showed that the inflated eccentric-migration process efficiently accelerates the migration of such gas planets, compared with eccentric-migration models where the thermal evolution of the planets is not considered. Initially inflated planets and planets that reinflated due to tidal and/or radiative heating experience stronger tides, allowing for planets initialized at larger separations to migrate inwards and inducing higher rates of tidal disruptions of gas giants.

We find that the energy deposited by tides is mostly negligible in the equilibrium tides regime (weak tides) when deposited close to the planet's surface. Tidal heating can be important and even lead to planetary inflation if highly efficient dynamical tides are considered ($f_{\rm dyn} > 0.1$). In addition, efficient heat transfer from the outer regions of the planets where radiative and/or tidal heating is deposited to the central parts also gives rise to significant thermal evolution and possible inflation of planets during their migration, even when only weak tides or less efficient dynamical tides are considered. As the planets reinflate, the radii of HJs may become larger, but the number of disruptions may increase (see Paper I for further discussion). Identifying the exact processes and efficiencies of heat transfer in gas giants is therefore critical for our understanding of their dynamical evolution and the formation of HJs and WJs. However, this is out of the scope of this paper, and we leave it to future works.

Our numerical and analytical approaches complement each other, and both can account for additional types of dynamical processes and other types of external energies. Our numerical model can be used to simulate the detailed evolution of stellar multiples where one can use the coupled internal evolution part on more than one component. The good agreement between the numerical model presented here and the semianalytic models presented in Paper I support the use of the latter, and the analytic model can well reproduce all of the numerical results.

The more computationally efficient semianalytical could then be used to study the large parameter space of the HJ/WJ populations, as described in Paper I.

6. Software and Third-party Data Repository Citations

Our numerical code can be found under the public repository https://github.com/hilaglanz/InflatedEccenricMigration. Here we used `AMUSE` version 13.2.1 with self-contribution as described, combined with `MESA` version 2208.

We gratefully acknowledge helpful discussions with Sivan Ginzburg, Thaddeus D. Komacek, Michelle Vick, Nicholas C. Stone, and Eden Saig. M.R. acknowledges the generous support of the Azrieli fellowship. H.B.P. acknowledges support for this project from the Minerva center for life under extreme planetary conditions.

Software: AMUSE (Portegies Zwart et al. 2009), MESA (Paxton et al. 2011, 2013).

ORCID iDs

Hila Glanz https://orcid.org/0000-0002-6012-2136
Mor Rozner https://orcid.org/0000-0002-2728-0132
Hagai B. Perets https://orcid.org/0000-0002-5004-199X
Evgeni Grishin https://orcid.org/0000-0001-7113-723X

References

Alexander, M. E. 1973, Ap&SS, 23, 459
Anderson, K. R., Lai, D., & Pu, B. 2020, MNRAS, 491, 1369
Armitage, P. J. 2010, Astrophysics of Planet Formation (Cambridge, UK: Cambridge Univ. Press)
Baraffe, I., Chabrier, G., & Barman, T. 2010, RPPh, 73, 016901
Batygin, K., & Stevenson, D. J. 2010, ApJL, 714, L238
Bodenheimer, P., Hubickyj, O., & Lissauer, J. J. 2000, Icar, 143, 2
Bodenheimer, P., Lin, D. N. C., & Mardling, R. A. 2001, ApJ, 548, 466
Bodenheimer, P., & Pollack, J. B. 1986, Icar, 67, 391
Boss, A. P. 1997, Sci, 276, 1836
Burrows, A., Hubeny, I., Budaj, J., & Hubbard, W. B. 2007, ApJ, 661, 502
Butler, R. P., Wright, J. T., Marcy, G. W., et al. 2006, ApJ, 646, 505
Chatterjee, S., Ford, E. B., Matsumura, S., & Rasio, F. A. 2008, ApJ, 686, 580
Darwin, G. H. 1879, RSPT, 170, 1
Dawson, R. I., & Johnson, J. A. 2018, ARA&A, 56, 175
Fabrycky, D., & Tremaine, S. 2007, ApJ, 669, 1298
Freedman, R. S., Marley, M. S., & Lodders, K. 2008, ApJS, 174, 504
Ginzburg, S., & Chiang, E. 2019, MNRAS, 490, 4334
Ginzburg, S., & Sari, R. 2015, ApJ, 803, 111
Goldreich, P., & Soter, S. 1966, Icar, 5, 375
Grevesse, N., & Sauval, A. J. 1998, SSRv, 85, 161
Grishin, E., & Perets, H. B. 2022, MNRAS, 512, 4993
Guillochon, J., Ramirez-Ruiz, E., & Lin, D. 2011, ApJ, 732, 74
Guillot, T., Burrows, A., Hubbard, W. B., Lunine, J. I., & Saumon, D. 1996, ApJL, 459, L35
Guillot, T., & Showman, A. P. 2002, A&A, 385, 156
Hamers, A. S., Antonini, F., Lithwick, Y., Perets, H. B., & Portegies Zwart, S. F. 2017, MNRAS, 464, 688
Hamers, A. S., & Tremaine, S. 2017, AJ, 154, 272
Huang, C., Wu, Y., & Triaud, A. H. M. J. 2016, ApJ, 825, 98
Hut, P. 1981, A&A, 99, 126
Ivanov, P. B., & Papaloizou, J. C. B. 2004, MNRAS, 347, 437
Ivanov, P. B., & Papaloizou, J. C. B. 2007, MNRAS, 376, 682
Jurić, M., & Tremaine, S. 2008, ApJ, 686, 603
Kippenhahn, R., Weigert, A., & Weiss, A. 2012, Stellar Structure and Evolution (Stellar Structure and Evolution) (Berlin: Springer), 2012
Komacek, T. D., Thorngren, D. P., Lopez, E. D., & Ginzburg, S. 2020, ApJ, 893, 36
Komacek, T. D., & Youdin, A. N. 2017, ApJ, 844, 94
Kozai, Y. 1962, AJ, 67, 591
Lai, D. 1997, ApJ, 490, 847
Lidov, M. L. 1962, P&SS, 9, 719
Lin, D. N. C., & Ida, S. 1997, ApJ, 477, 781
Mardling, R. A. 1995a, ApJ, 450, 722
Mardling, R. A. 1995b, ApJ, 450, 732
McMillan, S. L. W. 1986, ApJ, 306, 552
Miller, N., & Fortney, J. J. 2011, ApJL, 736, L29
Miller, N., Fortney, J. J., & Jackson, B. 2009, ApJ, 702, 1413
Mizuno, H. 1980, PThPh, 64, 544
Moe, M., & Kratter, K. M. 2018, ApJ, 854, 44
Murray, C. D., & Dermott, S. F. 1999, in Solar System Dynamics, ed. C. D. Murray & S. F. McDermott (Cambridge, UK: Cambridge Univ. Press)
Nagasawa, M., Ida, S., & Bessho, T. 2008, ApJ, 678, 498
Naoz, S., Farr, W. M., Lithwick, Y., Rasio, F. A., & Teyssandier, J. 2011, Natur, 473, 187
Ogilvie, G. I. 2014, ARA&A, 52, 171
Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3
Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4
Paxton, B., Schwab, J., Bauer, E. B., et al. 2018, ApJS, 234, 34
Paxton, B., Smolec, R., Schwab, J., et al. 2019, ApJS, 243, 10
Perri, F., & Cameron, A. G. W. 1974, Icar, 22, 416
Petrovich, C. 2015, ApJ, 799, 27
Portegies Zwart, S., McMillan, S., Harfst, S., et al. 2009, NewA, 14, 369
Press, W. H., & Teukolsky, S. A. 1977, ApJ, 213, 183
Rafikov, R. R. 2005, ApJL, 621, L69
Rasio, F. A., & Ford, E. B. 1996, Sci, 274, 954
Rogers, F. J., & Nayfonov, A. 2002, ApJ, 576, 1064
Rozner, M., Glanz, H., Perets, H. B., & Grishin, E. 2021, arXiv:2111.12718
Spiegel, D. S., & Burrows, A. 2013, ApJ, 772, 76
Sun, M., Arras, P., Weinberg, N. N., Troup, N. W., & Majewski, S. R. 2018, MNRAS, 481, 4077
Thorngren, D. P., & Fortney, J. J. 2018, AJ, 155, 214
Vick, M., & Lai, D. 2018, MNRAS, 476, 482
Vick, M., Lai, D., & Anderson, K. R. 2019, MNRAS, 484, 5645
von Zeipel, H. 1910, AN, 183, 345
Wu, Y. 2018, AJ, 155, 118
Wu, Y., & Lithwick, Y. 2011, ApJ, 735, 109
Wu, Y., & Murray, N. 2003, ApJ, 589, 605
Wu, Y., Murray, N. W., & Ramsahai, J. M. 2007, ApJ, 670, 820
Zahn, J. P. 1977, A&A, 500, 121
Zhu, W., & Dong, S. 2021, ARA&A, 59, 42

Chapter 5

Dynamics in gas-rich environments

While the dynamics of binaries in gas-free environments have been studied extensively, gas-rich environments are fertile ground for phenomena that are still largely unexplored. The evolution of binaries in gas-rich environments is essentially unique and leads to various special signatures.

The novelty of the project we studied arises not only from the new dynamics introduced but also from the environments I study. While AGN disks have been widely studied among gaseous environments, globular clusters were thought for many years to host a simple/single-age population – i.e. to contain stars that were born during a single burst of star formation. However, observations that studied chemical abundances showed that the vast majority of clusters host at least two populations (see 1.3.2). Even from an agnostic point of view regarding the formation channel of the second population, most of the formation channels require high concentrations of gas that allow us to treat clusters at these epochs as gas-rich clusters. During my PhD, I started to study these environments as gas-rich (Rozner and Perets, 2022b; Rozner et al., 2023), and we were among the first to consider the dynamics in clusters at these epochs as embedded in gas.

5.1 Binary evolution & gravitational-wave mergers in multiple-population gas-enriched globular Clusters

Based on Rozner and Perets (2022b)

Gas-assisted binary mergers are discussed extensively, mainly in the context of AGN disks (McKernan et al., 2012; Stone et al., 2017; Tagawa et al., 2020). Gas could lead to energy extraction from the binaries, i.e. separation shrinkage.

We introduced a novel gravitational waves channel, originates from gas-rich globular clusters during their gaseous epoch of second (or further) star-formation generation. Here I will briefly review this channel, its settings and unique signatures.

The total mass of the gas is uncertain, but we can assess the minimal required gas density to enable the formation of the second generation. The typical gas density in star-forming regions is usually constrained in the range $10^2 - 10^6$ $M_\odot \mathrm{pc}^{-3}$ (Leigh et al., 2014). Estimates for the 2P gas densities could be obtained from simple order of magnitude calculations, assuming 2P stars were formed from replenished gas. The gas density is then $\rho_g \sim M_g/V_{\mathrm{2P}}$ where M_g is the mass of the gas and V_{2P} is the typical volume in which the 2P stars reside. Following Bekki (2017), $M_{\mathrm{2P}} \sim 10^5 M_\odot$ and $\epsilon_g = 0.3$, then $M_g \sim 3 \times 10^5 M_\odot$, where ϵ is the star-formation efficiency. The infalling replenished gas is likely concentrated in a compact region in the central parts of GCs, probably in a disk structure, such that the typical effective radius that encloses the 2P population is of the order of 1 pc (Bekki, 2017). Taken together, the typical density of the replenished gas is $\sim 3 \times 10^5$ $M_\odot$ pc^{-3}, which lies within the expected range for gas densities in star-forming regions. From this density, we will consider scaling to different gas masses, considering $R_{\mathrm{core}} = 1$ pc and take $\rho_g \sim M_{\mathrm{g}}/R^3_{\mathrm{core}}$ accordingly. In particular, as we discuss below, the 2P gas is likely enclosed in a disk-like configuration, in which case the expected gas densities are higher.

The separation evolution is dictated by the following equation (for the circular case), taking into account the contribution from gas dynamical friction, three-body hardening, and gravitational waves

$$\frac{da_{\rm bin}}{dt} = \left.\frac{da_{\rm bin}}{dt}\right|_{\rm GDF} + \left.\frac{da_{\rm bin}}{dt}\right|_{\rm 3-body} + \left.\frac{da_{\rm bin}}{dt}\right|_{\rm GW} \tag{5.1}$$

where $a_{\rm bin}$ is the binary separation. There is a separation of scales, and for large separations, gas dynamical friction dominates (GDF), for intermediate ones three body hardening (3-body) and for small separations gravitational waves (GW). These components are given by

The binary hardening induced by GDF for the circular case, with binary components with the same mass $m_1 = m_2 = m$ is given by (Rozner and Perets, 2022b),

$$\left.\frac{da_{\rm bin}}{dt}\right|_{\rm GDF} = -\frac{8\pi G^{3/2} a_{\rm bin}^{3/2}}{\sqrt{m_1 + m_2}} \rho_g(t) \frac{m}{v_{\rm rel}^2} \mathcal{I}\left(\frac{v_{\rm rel}}{c_s}\right) \tag{5.2}$$

where $v_{\rm rel}$ is the velocity of the binary relative to the gas, taken as the Keplerian velocity of the binary, i.e. $v_K = \sqrt{G(m_1 + m_2)/a_{\rm bin}}$, which dominates the relative velocity throughout most of the evolution. For very wide binaries, the velocity dispersion dominates over the Keplerian velocity of the binaries and then $v_{\rm rel} \sim \sigma$. For eccentric orbits, under the same assumptions, the velocity relative to the gas is given by

$$\mathbf{v}_{\rm rel} = \frac{\Omega a}{2\sqrt{1 - e^2}} \left[e \sin f \hat{r} + (1 + e \cos f)\hat{\varphi}\right] \tag{5.3}$$

and the binary gas-hardening rate is given by

$$\left.\frac{da}{dt}\right|_{\rm GDF} = \frac{2a^{3/2}}{m_{\rm bin}\sqrt{Gm_{\rm bin}(1-e^2)}}\left[F_r e\sin f + F_\varphi(1+e\cos f)\right], \tag{5.4}$$

$$\left.\frac{de}{dt}\right|_{\rm GDF} = \frac{2}{m}\sqrt{\frac{a(1-e^2)}{Gm_{\rm bin}}}\left[F_r\sin f + F_\varphi(\cos f+\cos E)\right] \tag{5.5}$$

where $\mathbf{F}_{\rm drag} = F_r\hat{r} + F_\varphi\hat{\varphi}$, f is the true anomaly and E is the eccentric anomaly. The orbit-averaged equations are given by

$$\left.\overline{\frac{da}{dt}}\right|_{\rm GDF} = \frac{4F_0(1-e^2)^2}{\pi m_{\rm bin}\Omega^3a^2}\int_0^{2\pi}\frac{Idf}{(1+e\cos f)^2\sqrt{1+2e\cos f+e^2}}, \tag{5.6}$$

$$\left.\overline{\frac{de}{dt}}\right|_{\rm GDF} = \frac{4F_0(1-e^2)^3}{\pi m_{\rm bin}\Omega^3a^3}\int_0^{2\pi}\frac{I(e+\cos f)df}{(1+e\cos f)^2(1+2e\cos f+e^2)^{3/2}} \tag{5.7}$$

where F_0 is given by $\mathbf{F}_{\rm drag} = F_0 I\mathbf{v}_{\rm rel}/v_{\rm rel}^3$. The orbit-averaged equations for GWs are given by

$$\left.\overline{\frac{da}{dt}}\right|_{\rm GW} = -\frac{64G^3m_1m_2(m_1+m_2)}{5c^5a^3(1-e^2)^{7/2}}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right), \tag{5.8}$$

$$\left.\overline{\frac{de}{dt}}\right|_{\rm GW} = -\frac{304G^3em_1m_2(m_1+m_2)}{15c^5a^4(1-e^2)^{5/2}}\left(1+\frac{121}{304}e^2\right) \tag{5.9}$$

For hard binaries, the dynamical hardening rate (up to order unity corrections calibrated usually from numerical simulations) is given by (Spitzer, 1987)

$$\left.\frac{da_{\rm bin}}{dt}\right|_{3-body} = -\frac{2\pi Gn_\star m_{\rm pert}(2m+m_{\rm pert})a_{\rm bin}^2}{mv_\infty} \tag{5.10}$$

where we consider a binary with equal mass components, $m = m_1 = m_2$ and an external perturber with mass $m_{\rm pert}$. For interactions with other massive objects only, $n_\star$ and $m_{\rm pert}$ should be taken as $n_\bullet$ and $\bar{m}_\bullet$ correspondingly.

For a circular binary in the quadruple approximation, the GWs inspiral rate is given by (Peters, 1964),

$$\left.\frac{da}{dt}\right|_{\mathrm{GW}} = -\frac{64G^3 m_1 m_2 (m_1 + m_2)}{5c^5 a^3} \tag{5.11}$$

where G is the gravitational constant and c is the speed of light.

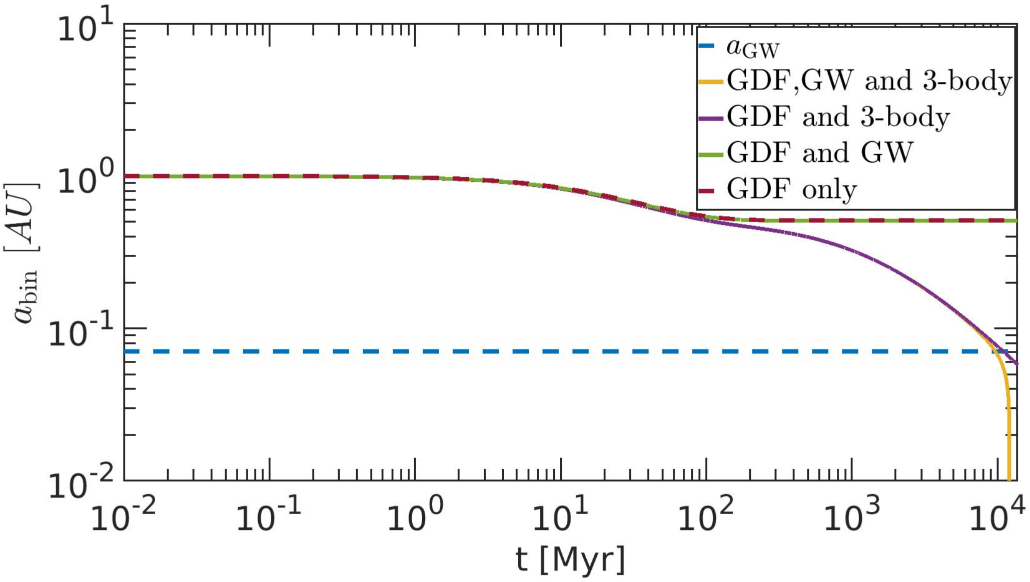

Figure 5.1: The effects of gas hardening, GWs, and three-body hardening. The blue dashed line represents the maximal SMA in which GW emission catalyzes a binary merger within Hubble time. We consider the evolution of a binary with masses $m_1 = m_2 = 10\ M_\odot$ and the initial separation of $a_0 = 1$ AU. We consider an exponential decaying background gas density $\rho_g = \rho_{g,0} \exp(-t/\tau_{\mathrm{gas}})$ with $\rho_{g,0} = 1.74\times10^6\ M_\odot \mathrm{pc}^{-3}$ and $\tau_{\mathrm{gas}} = 50$ Myr.

In Fig. 5.1, we present the evolution of a circular binary, embedded in gas, under the effect of gas dynamical friction, three-body hardening and gravitational waves, and a comparison between the significance of the contributions of each one of these components. It could be seen that the process is efficient and leads to a merger of a binary that won't merge in a gas-free medium.

We then investigated the behavior of all the binaries from the hardest to the transition between soft and hard, and thought that under our fiducial assumptions, all of them are expected to merge (Fig. 5.2).

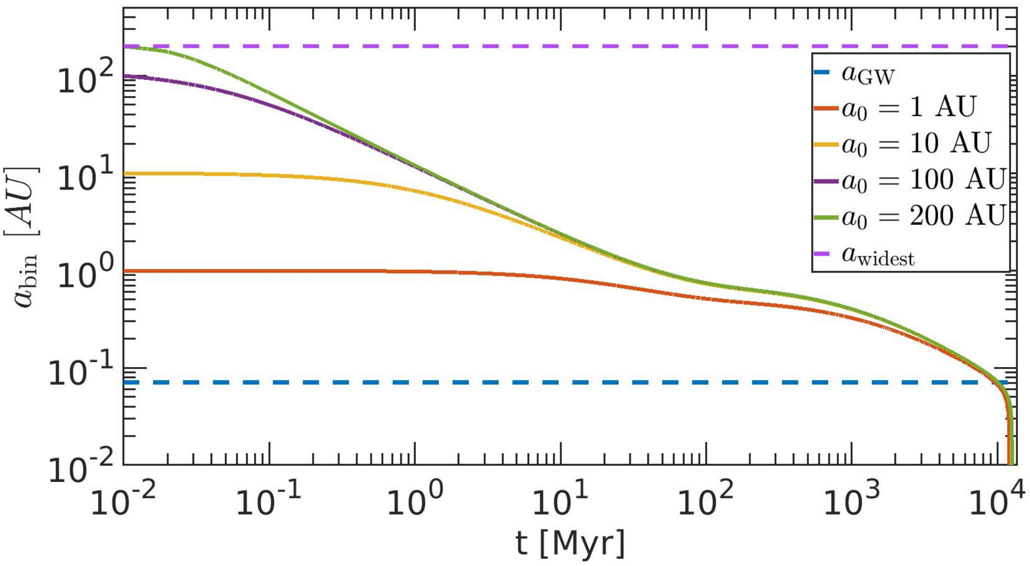

Figure 5.2: The combined effect of gas hardening, three-body hardening and GWs on a binary, for different initial separations. The blue dashed line represents the maximal SMA in which GW emission catalyzes a binary merger within Hubble time. The purple dashed line corresponds to the widest binary allowed by evaporation considerations. The solid lines correspond to the evolution of the SMA, starting from an initial separation of $a_0 = 1, 10, 100, 200$ AU, and given an exponential decaying gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\rm gas})$ with $\tau_{\rm gas} = 50$ Myr.

For eccentric binaries, the hardening time shortens significantly, and could be even no more than few Myr. Moreover, for a flat density profile, gas dynamical friction leads to eccentricity increment, since the force scales as $1/v_{\rm rel}^2$, and the relative velocity could be estimated by the Keplerian velocity, which peaks at the pericenter, and reach the minimal possible value at apocenter. When a small enough pericenter distance is reached, gravitational waves become the dominant dissipation mechanism, leading to fast circularization of the binary, such that only a small residual eccentricity could be potentially observed in the future in the LISA band.

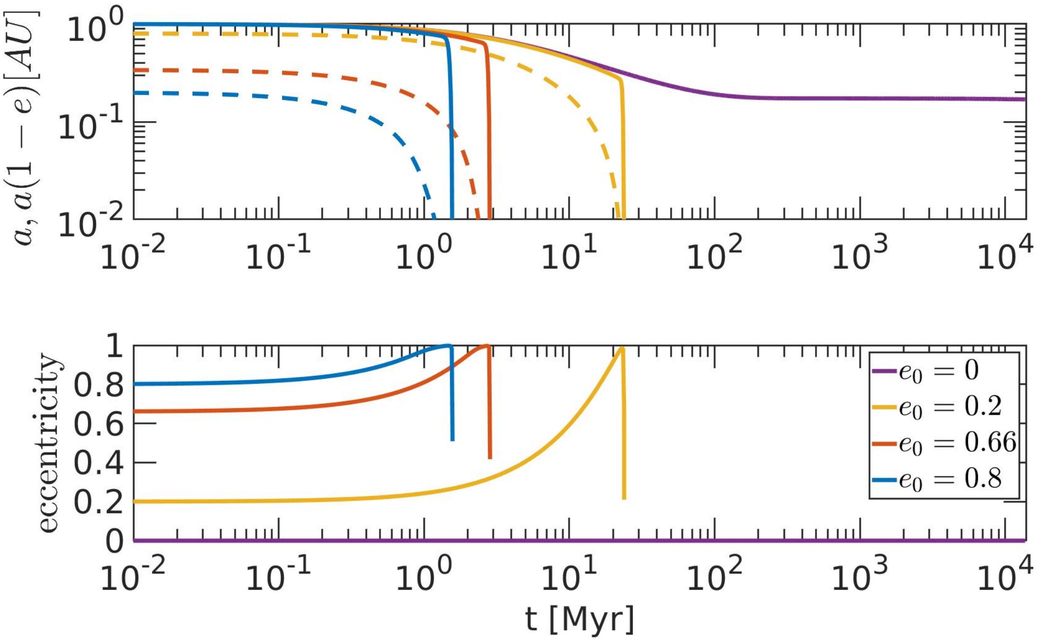

Figure 5.3: The effects of gas hardening and GWs on eccentric orbit. We consider the evolution of a binary with masses $m_1 = m_2 = 10\ M_\odot$ and an initial separation of $a_0 = 1$ AU. We consider an exponential decaying background gas density $\rho_g = \rho_{g,0} \exp(-t/\tau_{\rm gas})$ with $\rho_{g,0} = 1.74 \times 10^6\ M_\odot \rm{pc}^{-3}$ and $\tau_{\rm gas} = 50$ Myr. The solid lines correspond to semimajor axis evolution and the dashed lines to pericenter evolution.

We also calculated the expected merger rates, and found out that they could overlap with the expected rate from observations, although with high uncertainties.

To summarize, we described here a novel gas-assisted gravitational wave channel in globular clusters. We studied the evolution of already-formed binaries under the combined effect of gas dynamical friction, three-body hardening and gravitational waves, and found that this mechanism is highly efficient and could potentially explain a decent fraction of the observed gravitational waves.

The Astrophysical Journal, 931:149 (13pp), 2022 June 1 https://doi.org/10.3847/1538-4357/ac6d55

Binary Evolution, Gravitational-wave Mergers, and Explosive Transients in Multiple-population Gas-enriched Globular Clusters

Mor Rozner and Hagai B. Perets
Technion—Israel Institute of Technology, Haifa, 3200002, Israel; morozner@campus.technion.ac.il

Received 2022 March 1; revised 2022 April 25; accepted 2022 May 4; published 2022 June 6

Abstract

Most globular clusters (GCs) show evidence for multiple stellar populations, suggesting the occurrence of several distinct star formation episodes. The large fraction of second population (2P) stars observed requires a very large 2P gaseous mass to have accumulated in the cluster core to form these stars. Hence, the first population of stars (1P) in the cluster core has had to become embedded in 2P gas, just prior to the formation of later populations. Here we explore the evolution of binaries in ambient 2P gaseous media of multiple-population GCs. We mostly focus on black hole binaries and follow their evolution as they evolve from wide binaries toward short periods through interaction with ambient gas, followed by gravitational-wave (GW) dominated inspiral and merger. We show that this novel GW merger channel could provide a major contribution to the production of GW sources. We consider various assumptions and initial conditions and calculate the resulting gas-mediated change in the population of binaries and the expected merger rates due to gas-catalyzed GW inspirals. For plausible conditions and assumptions, we find an expected GW merger rate observable by aLIGO of the order of up to a few tens of $\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$ and an overall range for our various models of 0.08–25.51 $\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$. Finally, our results suggest that the conditions and binary properties in the early stage of GCs could be critically affected by gas interactions and may require a major revision in the current modeling of the evolution of GCs.

Unified Astronomy Thesaurus concepts: Gravitational waves (678); Gravitational wave sources (677); Globular star clusters (656)

1. Introduction

Stars are thought to form following the collapse of giant molecular clouds (GMCs) and further grow and evolve through accretion from and interaction with the GMC ambient gaseous environment during their early evolution, of up to a few Myr. Following the gas dispersal and depletion, the later long-term evolution of stars and multiple systems is thought to be dominated by their gas-free stellar evolution and their dynamical interactions with other stellar companions and/or stars in the cluster. However, some environments can be replenished with gas, leading to late epochs of stellar and binary evolution of stars embedded in gas. Already decades ago, Bahcall & Ostriker (1976) suggested that stellar compact objects can interact with gaseous disks around massive black holes (BHs; active galactic nuclei (AGNs)), accrete, and give rise to X-ray flarings. Ostriker (1983) suggested that stars and compact objects embedded in AGN disks can accrete gas from the ambient gaseous medium, grow to Chandrasekhar mass, and explode as Type Ia supernovae (SNe), and later Artymowicz et al. (1993) discussed accretion onto stars in AGN disks giving rise to massive stars exploding as core-collapse (CC) SNe and polluting the AGN disks.

The dynamical evolution of binary gravitating objects embedded in a large-scale gaseous environment could be altered through gas dynamical friction (GDF) and accretion that change their orbit and masses and potentially catalyze their merger. We have first discussed binary evolution in gaseous media in the context of catalyzed mergers of binary planetesimals in protoplanetary disks (Perets & Murray-Clay 2011; Grishin & Perets 2016), and later in the context of compact-object binaries in AGN disks (McKernan et al. 2012), where the latter have been extensively studied since then (e.g., Stone et al. 2017; McKernan et al. 2018; Roupas & Kazanas 2019; Tagawa et al. 2020, and references therein). Baruteau et al. (2011) explored the evolution of binary main-sequence (MS) stars in gas disks around massive BHs (MBHs), suggesting that they harden and merge through the interaction with the gas. Various studies followed the evolution of pre-MS/MS binaries embedded in gas just following their formation during the star formation epoch of stars in molecular clouds/young clusters, also suggesting that binaries can shrink and merge through the process (Gorti & Bhatt 1996; Er et al. 2009; Korntreff et al. 2012). It was also suggested that the evolution of embedded binaries could be driven by the formation of a circumbinary disk, which torques the binary. The evolution of binaries in circumbinary disks has been more extensively studied over a wide range of scales from planets, to stars, to MBHs (though typically not in the context of a large-scale gaseous environment), but the exact evolution and even the direction of the binary migration in such circumbinary disks are still debated (e.g., Artymowicz et al. 1991; Artymowicz & Lubow 1994; Bate 2000; Tang et al. 2017; Moody et al. 2019; Muñoz et al. 2019; Duffell et al. 2020; Muñoz et al. 2020, and references therein).

Although the evolution of stars, binaries, and compact objects embedded in gaseous (typically AGN) disks near MBHs has been extensively studied over the past few years, other gas-embedded stellar environments received far less attention. Here and in a companion paper (Perets 2022) we study the evolution of single and binary compact-object binaries in the early gas-rich environments that likely existed in multiple-population globular clusters (GCs) and other young massive clusters (YMCs). We also briefly discuss other (non-compact-object—MS and evolved) stars and binaries in such environments, but we postpone detailed study of the latter to future exploration.

As we discuss below, such gas-rich environments are likely to be far more ubiquitous than AGN disks and potentially play a key role in the production of compact binaries, binary mergers, gravitational-wave sources, and explosive transients.

For decades, GCs were thought to host simple stellar populations formed through a single star formation episode. However, detailed observations over the past decade (see, e.g., Carretta et al. 2009; Bastian & Lardo 2018, and references therein) have shown that the vast majority of galactic GCs host multiple stellar populations showing different light-element content. The origins of multiple populations have been extensively studied, but no clear solution has yet been found (see Renzini et al. 2015; Bastian & Lardo 2018; Gratton et al. 2019, for summaries of the scenarios and their caveats). The current thought is that GCs experienced two or more star formation episodes, in which second-generation/population (2P) stars formed from processed (2P) gas lost from earlier-generation/population (1P) stars and/or accreted external gas. Kinematics show that 2P stars are more centrally concentrated and were likely formed in the inner region of the GC, where the 2P gas is expected to have accumulated.

While the source of the 2P gas is debated, the late formation of 2P stars requires that from tens up to hundreds of Myr after their formation 1P stars had become embedded in a highly gas-rich environment that later produced the 2P stars. The evolution of stars, binaries, and compact objects embedded in gas could therefore be significantly altered in such gaseous environments, following similar processes to those discussed for AGN disks and pre-MS stars embedded in the progenitor GMCs. Such processes were little studied in the context of gas-embedded multiple-population GCs (Vesperini et al. 2010; Maccarone & Zurek 2012; Leigh et al. 2013, 2014; Roupas & Kazanas 2019; Perets 2022, but see works by us and others on some aspects of such evolution), which are the focus of the study below. In particular, in this paper we introduce the effect of gas-catalyzed hardening (shrinkage of the orbit) of binaries in GCs and discuss its implications for GC (and YMC) binary populations and binary mergers, the production of GW sources, and the formation of other merger products, compact binaries, and explosive transient events catalyzed by binary interactions with gas.

In Section 2 we briefly discuss the gas replenishment in multiple-population GCs. In Section 3 we describe the hardening process of binaries in GCs due to GDF and its relation to dynamical hardening by stars and GW inspirals. In Section 4 we introduce our results: in Section 4.1 we focus on the evolution of an individual binary under the effect of gas hardening, and in Section 4.4 we estimate the expected merger rate from the channel we proposed. In Section 5 we discuss our results and additional implications. In Section 6 we summarize and conclude.

2. Multiple Stellar Populations and Early Gas Replenishment in GCs

As discussed above and in Perets (2022), gas could be replenished in GCs (and YMCs) through mass lost from evolved stars and binaries and/or through accretion of external gas onto the clusters (see a detailed review in Bastian & Lardo 2018).

The formation channel sets the amount of gas and hence the dynamics and evolution of embedded stars/binaries. Given the correlation between the fractions of 2P stars and GC properties, it is likely that a large fraction of 2P stars correspond to higher masses of the clusters, larger escape velocities (Mastrobuono-Battisti & Perets 2020), and hence larger mass of replenished gas.

Given the observed kinematics and concentrations of 2P stars and theoretical models for the formation and evolution of 2P stars, it is thought that the replenished gas is concentrated in the central part of GCs, where 2P stars are concentrated. It is likely that the remnant angular momentum of replenished gas gives rise to the formation of 2P in gaseous disks, rather than spherical distribution (Bekki 2010, 2011; Mastrobuono-Battisti & Perets 2013, 2016).

The total mass of 2P gas in GCs is highly uncertain, but given reasonable assumptions on the relation between the gas and the observed populations of 2P stars in GCs, one can provide an estimate of the amount of replenished gas and its density. The typical gas density in star-forming regions is usually constrained in the range 10^2–$10^6\,M_\odot\,\mathrm{pc}^{-3}$ (Leigh et al. 2014). Estimates for the 2P gas densities could be obtained from simple order-of-magnitude calculations, assuming that 2P stars were formed from replenished gas. The gas density is then $\rho_g \sim M_g/V_{\mathrm{2P}}$, where M_g is the mass of the gas and V_{2P} is the typical volume in which the 2P stars reside. Following Bekki (2017), $M_{\mathrm{2P}} \sim 10^5\,M_\odot$ and $\epsilon_g = 0.3$, and then $M_g \sim 3 \times 10^5\,M_\odot$, where ϵ is the star formation efficiency. The infalling replenished gas is likely concentrated in a compact region in the central parts of GCs, such that the typical effective radius that encloses the 2P population is of the order of 1 pc (Bekki 2017). Taken together, the typical density of the replenished gas is $\sim 3 \times 10^5\,M_\odot\,\mathrm{pc}^{-3}$, which lies within the expected range for gas densities in star-forming regions. From this density, we will consider scaling to different gas masses, considering $R_{\mathrm{core}} = 1$ pc, and take $\rho_g \sim M_g/R_{\mathrm{core}}^3$ accordingly. In particular, as we discuss below, the 2P gas is likely enclosed in a disk-like configuration, in which case the expected gas densities are higher. A priori, the binary hardening releases energy that could heat the gas significantly, but from a crude calculation, the cooling rate is high enough to compensate for it (see also Tagawa et al. 2020 for a similar calculation in AGN disks). We also note that the possible production of jets could potentially unbind gas from the disk (Soker 2016; Tagawa et al. 2022), but the study of this possibility is beyond the scope of the current paper.

The total amount of gas is depleted in time, due to formation of stars and/or accretion onto stars, and later gas ejection through possible radiation pressure processes and SNe. For simplicity we assume an exponential decay, i.e., $\rho_g(t) = \rho_{g,0}\exp(-t/\tau_{\mathrm{gas}})$, and consider several possible options for the gas lifetime, to account for uncertainties in the possible gas depletion processes involved.

2.1. Disk Configuration

Gas replenishment leading to the formation of 2P stars in GCs might form a disk-like structure in the cluster nuclei (e.g., Bekki 2010; Mastrobuono-Battisti & Perets 2013).

Following Bekki (2010), we consider a flat disk, i.e., with a constant aspect ratio. We estimate the aspect ratio by

$h/r \sim c_s/v_K$, where $v_K = \sqrt{G(M_{\rm gas} + M_\star)/R_{\rm core}}$ is the typical velocity in the central parsec. The speed of sound $c_s = \sqrt{k_B T_{\rm gas}/\mu m_p}$ ranges between 0.1 and $10\,{\rm km\,s^{-1}}$ (e.g., Bekki 2010; Leigh et al. 2013), in correspondence with the gas temperature $T_{\rm gas}$, such that $c_s \approx 0.6\,{\rm km\,s^{-1}}$ corresponds to a temperature of 100 K, which is the typical temperature in star formation areas, where $\mu = 2.3$, and m_p is the proton mass. Exponential disk models were also considered (Hénault-Brunet et al. 2015), but here we focus on simple models.

In our fiducial model, we consider $c_s = 10\,{\rm km\,s^{-1}}$, unless stated otherwise. Then, the aspect ratio $h/r \approx 0.23$. Following Bekki (2010), we consider a velocity dispersion of $\sigma_{\rm disk} = 10\,{\rm km\,s^{-1}}$ for stars embedded in the disk. As a conservative assumption, we consider the stellar/massive objects' density in the disk to be the same as in the core, i.e., $n_{\star,\rm disk} \approx n_\star = 10^5\,{\rm pc^{-3}}$. However, it should be noted that, due to GDF, stars will migrate and experience inclination damping, and the effective density in the disk is expected to be higher (e.g., Artymowicz et al. 1993; Leigh et al. 2014; Grishin & Perets 2016).

We can estimate the volume ratio between the disk and the core volume by $\pi R_{\rm core}^2 h/(4\pi R_{\rm core}^3/3) \sim 0.75 h\, r^{-1}$. Then, under the assumption that all the second-generation gas is concentrated in the disk, we get a typical gas density of $\rho_{\rm g,disk} \sim 1.74 \times 10^6\, M_\odot\,{\rm pc^{-3}}$. The fraction of stars in the disk will change for thinner/thicker disks correspondingly.

The evolution of binaries in disks differs in several aspects from the evolution in a spherical configuration. For our discussion, the major ones are as follows: the velocity dispersion decreases, the gas density increases, and the total number of stars contained in the disk is only the volumetric fraction of the disk compared with the volume of the spherical core. The fraction might change with time owing to the interaction with gas.

3. Dynamics of Binaries and Their Interaction with Gas: Binary Hardening and Mergers

Binaries embedded in gas interact with it, exchange angular momentum and energy, and possibly accrete gas. These processes are quite complex; here we focus on the interaction through GDF, while other suggested processes for interaction with gas are discussed in Section 4.2.

Besides interaction with gas, binaries in GCs can interact with other stars through dissipative effects such as GW inspirals or tidal evolution and through dynamical encounters with other stars through three-body (or more) encounters (Heggie 1975).

The semimajor axis (SMA) of a given massive binary in a gas-enriched environment evolves through the combined effect of the above-mentioned processes:

$$\frac{da_{\rm bin}}{dt} = \frac{da_{\rm bin}}{dt}\bigg|_{3-{\rm body}} + \frac{da_{\rm bin}}{dt}\bigg|_{\rm GDF} + \frac{da_{\rm bin}}{dt}\bigg|_{\rm GW}, \quad (1)$$

where $a_{\rm bin}$ is the binary SMA.

A priori, all three mechanisms contribute to the evolution of the SMA. However, in practice, each of these processes dominates in a specific regime and can be typically neglected in other regimes. Binaries could shrink to shorter periods (harden), due to the effect of gas interaction or GW inspiral, and get harder or softer (wider), due to three-body interactions with other GC stars. As we discuss in the following, the evolution of hard binaries is dominated by gas interactions at large separations and by GW emission at small separations, while dynamical hardening and softening through three-body encounters (Heggie 1975) can be neglected in these regimes. Nevertheless, binary softening and evaporation before the gas replenishment episode can destroy the widest binaries in the clusters and hence determine the largest possible initial SMAs for binaries in the cluster at the beginning of the gas interaction epoch. Moreover, it could play a role in hardening binaries that did not merge within the gas epoch.

The interaction with gas can also give rise to the formation of new wide binaries through two-body and three-body encounters in gas (Goldreich et al. 2002; Tagawa et al. 2020), allowing for replenishment of binaries in clusters.

In the following we discuss these various processes, while we neglect the effect of direct accretion onto compact objects and their growth, which is beyond the scope of the current paper (though generally such accretion, if effective, likely further accelerates binary hardening; e.g., Roupas & Kazanas 2019).

3.1. Hardening and Softening through Dynamical Encounters with Stars

Due to interactions with other stars, hard binaries tend to get harder, while soft binaries tend to get softer (Heggie 1975); see updated discussion and overview of these issues in Ginat & Perets (2021a, 2021b). Hence, in the absence of a gaseous environment stellar dynamical hardening plays an important role in binary evolution and in catalyzing binary mergers.

3.1.1. Hard Binaries

For hard binaries, the dynamical hardening rate (up to order-unity corrections calibrated usually from numerical simulations) is given by (Spitzer 1987)

$$\frac{da_{\rm bin}}{dt}\bigg|_{3-{\rm body}} = -\frac{2\pi G n_\star m_{\rm pert}(2m + m_{\rm pert})a_{\rm bin}^2}{m v_\infty}, \quad (2)$$

where we consider a binary with equal-mass components, $m = m_1 = m_2$, and an external perturber with mass $m_{\rm pert}$. For interactions with other massive objects only, $n_\star$ and $m_{\rm pert}$ should be taken as $n_\bullet$ and $\bar{m}_\bullet$, respectively.

3.1.2. Soft Binaries

A binary is called a soft binary if its energy is lower than $\bar{m}\sigma^2$. This condition sets a critical SMA,

$$a_{\rm SH} = \frac{2Gm^2}{\bar{m}\sigma^2} \approx 200.53{\rm au}\left(\frac{m}{10\,M_\odot}\right)^2 \left(\frac{43.2\,{\rm km\,s^{-1}}}{\sigma}\right)^2 \left(\frac{0.5\,M_\odot}{\bar{m}}\right). \quad (3)$$

As can be seen, massive stars tend to be hard relative to the background stars in the cluster, due to the scaling $a_{\rm SH} \propto m^2/\bar{m}$. Hence, one should define the hardness of massive binaries relative to both low- and high-mass stars; in particular, the latter will give rise to softer binaries. We then get the following modified expression (Quinlan 1996; Kritos & Cholis 2020):

$$a_{\mathrm{SH},\bullet} \approx \frac{Gm}{4\sigma^2} \approx 1.25\mathrm{au}\left(\frac{m}{10\,M_\odot}\right)\left(\frac{43.2\ \mathrm{km\ s^{-1}}}{\sigma}\right)^2. \quad (4)$$

Soft wide binaries are prone to destruction owing to encounters with other stars. The dynamical evolution of massive binaries is dominated by interactions with other massive stars, and their number density in the core is elevated owing to mass segregation (Sigurdsson & Phinney 1995).

So as to bracket the effect of softening, we consider two possibilities: (1) Softening is dominated by encounters with stellar BHs, where we assume the number density of such objects to be $n_b = n_\bullet = 10^3\ \mathrm{pc}^{-3}$, due to mass segregation to the core, where $\bar{m}_\bullet = 10\,M_\odot$ (see discussion in Miller & Hamilton 2002). (2) Softening is dominated by low-mass ($0.5\,M_\odot$) stars, if the cluster is not well segregated, and $n_b = n_\star = 10^5\ \mathrm{pc}^{-3}$.

Hence, the typical lifetime of a soft massive binary is given by (e.g., Binney & Tremaine 2008)

$$\tau_{\mathrm{evap,massive}} \approx \frac{(m_1+m_2)\sigma}{16\sqrt{\pi}\, n_b \bar{m}_b^2\, Ga\ln\Lambda}, \quad (5)$$

where $\ln\Lambda$ is the Coulomb logarithm and n_b and $\bar{m}_b$ are the number density and the mass of the background stars, respectively, and change according to our choice between (1) and (2). The separation of the widest binaries that survives evaporation until the formation of second-generation stars, signed as τ_{SG}, taken here to be 100 Myr, is then given by

$$a_{\mathrm{widest}} = \max\left\{a_{\mathrm{SH},\bullet}, \frac{(m_1+m_2)\sigma}{16\sqrt{\pi}\, n_b \bar{m}_b^2\, G\tau_{\mathrm{SG}}\ln\Lambda}\right\}. \quad (6)$$

For our fiducial parameters, $a_{\mathrm{widest}} = 24.9$ au for the segregated case and 200.53 au for the non-segregated case. In principle, binaries could soften and be disrupted via encounters during the gas replenishment episode; however, the GDF hardening described in the following is more efficient at this stage. Therefore, binary evaporation due to encounters sets the stage and determines the SMA of the widest binaries at the beginning of the gas enrichment stage, but it can be neglected during the time in which binaries are embedded in gas.

3.2. Gas Dynamical Friction

In gas-rich environments, such as the 2P gas environment of multiple-population GCs/YMCs (and AGN disks), GDF can play a major role in hardening. The evolution of binaries in gaseous media has been studied over a wide range of astrophysical scales from asteroids to MBHs (as discussed in the introduction).

The effect of gas was suggested to be modeled mainly via several approaches. One suggestion is that the accretion of gas onto a binary forms a circumbinary minidisk, due to accretion to the Hill sphere. In such disks, torques similar to the ones described by type I/II migration of planets in protoplanetary disks could lead to the shrinkage of the binary SMA (e.g., Artymowicz et al. 1991; McKernan et al. 2012; Stone et al. 2017; Tagawa et al. 2020). Such migration leads to very efficient mergers, far more efficient than the case of interaction dominated by GDF, as we discuss below. However, these issues are still debated, and some hydrodynamical simulations show that such torques might lead to outward migration (e.g., Moody et al. 2019; Duffell et al. 2020; Muñoz et al. 2020), while other hydrodynamical studies indicate that in thin disks one should have inward migration (Duffell et al. 2020; Tiede et al. 2020). We do note that most studies consider initially circular orbits and generally follow circular orbits, while eccentric orbits could evolve differently, with their orbital eccentricity possibly excited into very high eccentricities, as we discuss below in the context of modeling the evolution through GDF.

Therefore, the approach on which we focus here considers the effects of GDF (Ostriker 1999). When an object has a nonzero velocity relative to the background gas, the interaction with the gas reduces the relative velocity and therefore hardens binaries (e.g., Escala et al. 2004; Baruteau et al. 2011). The binary hardening induced by GDF for the circular case, with binary components with the same mass $m_1 = m_2 = m$, is given by Grishin & Perets (2016),

$$\left.\frac{da_{\mathrm{bin}}}{dt}\right|_{\mathrm{GDF}} = -\frac{8\pi G^{3/2}a_{\mathrm{bin}}^{3/2}}{\sqrt{m_1+m_2}}\rho_g(t)\frac{m}{v_{\mathrm{rel}}^2} f\left(\frac{v_{\mathrm{rel}}}{c_s}\right); \quad (7)$$

$$f(x) = \begin{cases} \frac{1}{2}\log\frac{1+x}{1-x} - x,\ 0 < x < 1, \\ \frac{1}{2}\log(x^2-1) + \log\Lambda_g,\ x > 1 \end{cases}, \quad (8)$$

where f is a dimensionless function derived in Ostriker (1999), and v_{rel} is the velocity of the binary relative to the gas, taken as the Keplerian velocity of the binary, i.e., $v_K = \sqrt{G(m_1+m_2)/a_{\mathrm{bin}}}$, which dominates the relative velocity throughout most of the evolution.

Under this assumption, Equation (7) could be written as

$$\left.\frac{da_{\mathrm{bin}}}{dt}\right|_{\mathrm{GDF}} = -8\pi\sqrt{\frac{Ga_{\mathrm{bin}}^5}{2m}}\rho_g(t) f\left(\frac{v_K}{c_s}\right). \quad (9)$$

For massive binaries, the effect of stellar hardening will be weaker than the effect on less massive stars, as can be seen directly from Equation (2). In contrast, the effect of gas hardening increases with mass (Equation (7)). Comparison of the two shows that hardening is dominated by gas hardening rather than stellar hardening. Moreover, although the effect of GDF decreases as the binary hardens, it decays more slowly than the three-body hardening, as could be seen from the scaling $\dot{a}_{\mathrm{hard},\star} \propto a^2$ and $\dot{a}_{\mathrm{GDF}} \propto a^{3/2}$, and therefore GDF dominates the evolution over stellar hardening throughout the evolution. After gas depletion, three-body hardening becomes the dominant dynamical process for wide binaries, while for sufficiently small separations the evolution is GW dominated.

3.3. Gravitational-wave Inspiral

For stellar-mass objects GW inspiral becomes important only at very small separations and can be neglected with regard to MS (or evolved) stellar binaries that merge before GW emission becomes important. However, GW inspiral plays a key role in the evolution of binaries composed of compact objects.

For a circular binary in the quadruple approximation, the GWs' inspiral rate is given by Peters (1964),

$$\left.\frac{da}{dt}\right|_{\rm GW} = -\frac{64G^3 m_1 m_2 (m_1 + m_2)}{5c^5 a^3}, \tag{10}$$

where G is the gravitational constant and c is the speed of light.

Without gas dissipation, the maximal SMA for GW merger within a Hubble time is given by

$$a_{\rm max,GW} = \left(\frac{64\tau_{\rm Hubble} G^3 m_1 m_2 (m_1 + m_2)}{5c^5}\right)^{1/4}$$
$$\approx 0.07\ {\rm au}\left(\frac{m}{10\,M_\odot}\right)^{3/4}. \tag{11}$$

A compact binary that is driven by GDF to separations below $a_{\rm max,GW}$ would eventually inspiral and merge, even if it survived the gas replenishment stage, and would produce a GW source.

4. Results

Accounting for the effects of the various processes discussed above, we can follow the evolution of binaries in clusters during the gas epoch and assess its outcomes. Overall we find that under plausible conditions all BH binaries initially existing in the cluster inner regions that become embedded in gas during the gas replenishment phase could be driven to short separations and merge within a Hubble time.

These results suggest that gas-catalyzed GW mergers in GCs and YMCs, not considered at all in current modeling of GCs, could serve as an important channel for the production of GW sources and play a key role in the evolution of binaries in such clusters.

Both the GDF and GW inspiral timescales for lower-mass compact objects such as neutron stars (NSs) and white dwarfs (WDs) are longer (as can be seen in Equation (11)), but they are also expected to modify their SMA distribution.

Here we focus on mergers of BHs, and we postpone a detailed discussion of NS and WD mergers to a follow-up paper, but we should already note that potential WD mergers could give rise to the production of explosive events such as Type Ia SNe from mergers of massive WDs (see also Perets 2022) and could produce GW sources observable by planned GW-detection space missions. NS mergers could produce short gamma-ray bursts and aLIGO GW sources. Combined BH−NS or BH−WD binaries with their high mass but lower mass ratio could be driven to mergers at intermediate timescales between the highest and lowest timescales considered here, giving rise to WD/NS disruptions by the BH possibly producing rapid faint SNe (e.g., Zenati et al. 2019, 2020; Bobrick et al. 2022, and references therein) or short GRBs accompanied by a potential GW aLIGO-source. The dynamics of binaries with nonequal masses could, however, be more complicated and is not explored here.

In the following we discuss our results in detail.

4.1. Gas-assisted GW Mergers

In Figure 1 we compare the different hardening processes of binaries in gas-embedded regions. As can be seen, for large separations the evolution is dominated by the gas hardening, while for smaller separations (at late times after the gas depletion) three-body hardening and finally GWs dominate the evolution. The transition between the different regimes is determined by the gas density in the cluster, as well as stellar density. Unless stated otherwise, we consider for our fiducial model a background of stars with typical masses of $\bar{m} = 0.5\,M_\odot$.

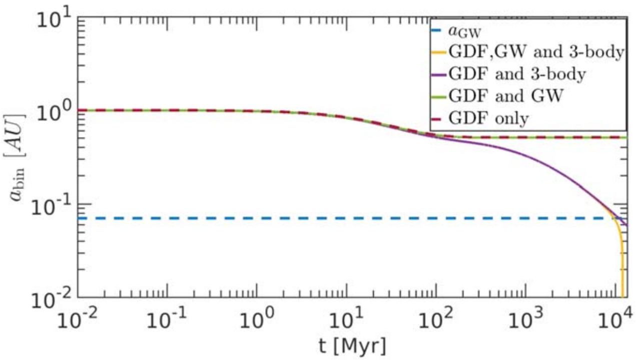

Figure 1. The effects of gas hardening, GWs, and three-body hardening. The blue dashed line represents the maximal SMA in which GW emission catalyzes a binary merger within a Hubble time. We consider the evolution of a binary with masses $m_1 = m_2 = 10\,M_\odot$ and initial separation of $a_0 = 1$ au. We consider an exponential decaying background gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\rm gas})$ with $\rho_{g,0} = 1.74\times10^6\,M_\odot\,{\rm pc}^{-3}$ and $\tau_{\rm gas} = 50$ Myr.

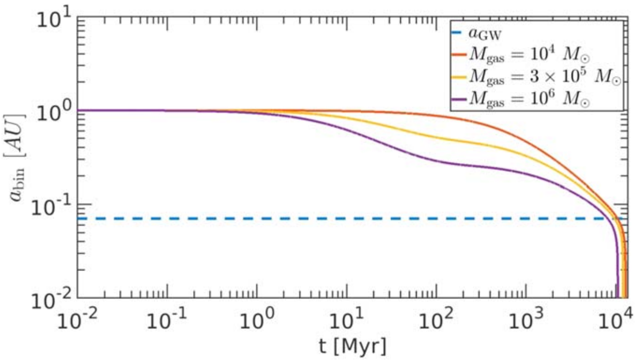

Figure 2. The combined effect of gas hardening, three-body hardening, and GWs on a binary, for different background gas masses (and corresponding gas densities). The blue dashed line represents the maximal SMA in which GW emission catalyzes a binary merger within a Hubble time. The solid lines correspond to the evolution of the SMA, starting from an initial separation of $a_0 = 1$ au, and given different background densities, with an exponential decaying gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\rm gas})$ with $\tau_{\rm gas} = 50$ Myr (which corresponds to $M_{\rm gas,0} = 3\times10^5\,M_\odot$). The velocity dispersions are calculated given the total mass of the gas and stars.

In Figure 2 we present the evolution of binaries with an initial separation of $a_0 = 1$ au, due to GDF, for different ambient gas densities. The gas hardening mechanism is generally very effective and leads to binary migration to small separations within short timescales, given a sufficiently dense gaseous environment. As we discuss below, such gas-assisted evolution would then give rise to high rates of GW mergers of BH binaries, comparable to the BH merger rates inferred from the aLIGO-VIRGO-KAGRA (LVK) collaboration (Abbott et al. 2016, 2021).

It should be noted that the gas could still dominate the evolution even after reaching $a_{\rm GW}$, as long as the gas was not depleted and the timescale for GWs mergers is larger than the GDF-induced merger timescale. In principle, GDF-dominated evolution might even be identified in the GW inspiral (in future space missions) before the merger, under appropriate conditions, if GDF still dominates the evolution in LISA frequencies.

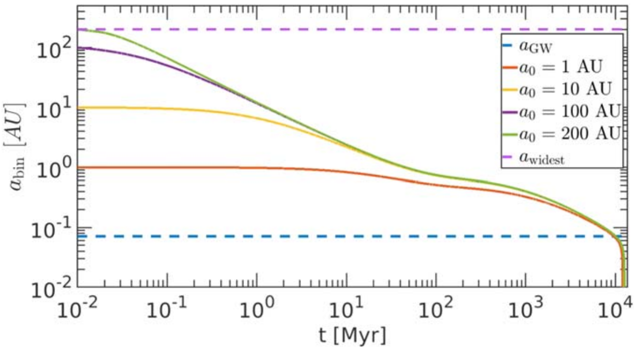

Figure 3. The combined effect of gas hardening, three-body hardening, and GWs on a binary, for different initial separations. The blue dashed line represents the maximal SMA in which GW emission catalyzes a binary merger within a Hubble time. The purple dashed line corresponds to the widest binary allowed by evaporation considerations. The solid lines correspond to the evolution of the SMA, starting from initial separations of $a_0 = 1$, 10, 100, and 200 au, and given an exponential decaying gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\rm gas})$ with $\tau_{\rm gas} = 50$ Myr.

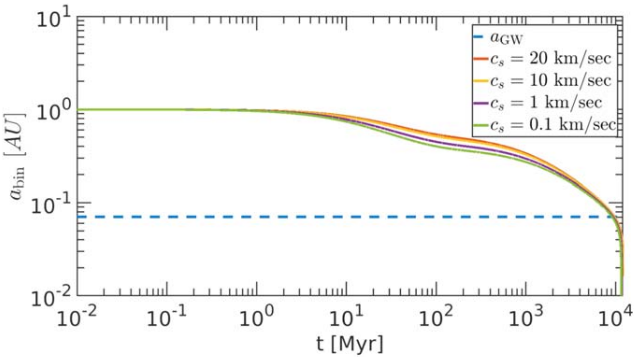

Figure 4. The evolution of the binary separation for different sound speeds. We consider equal-mass binaries with initial separation of $a = 1$ au, masses $m = m_1 = m_2 = 10\,M_\odot$, and an exponential decaying background density with $\rho_{\rm g,0} = 1.74 \times 10^6\,M_\odot$ pc^{-3}. The blue dashed line corresponds to the maximal separation from which a GW merger is expected.

We find that circular binaries shrink and reach final small separations, dictated by the initial conditions, which are not sufficiently small to allow for GW emission alone to drive the binaries to merger even after a Hubble time. Nevertheless, at such a short period, these very hard binaries are more likely to merge owing to dynamical encounters in the long term compared with the primordial population of binary BHs, and they should be appropriately accounted for in simulations of GC stellar populations.

In Figure 3, we introduce the evolution of binaries with different initial separations under the combined effect of GDF, three-body hardening, and GWs. It could be seen that although the merger timescales of wider binaries are slightly larger, all the binaries are expected to merge within a Hubble time. Hence, the effect of the presence of gas in the initial stages is robust across all separations and will modify the binary population. For wide enough binaries, we enter the subsonic range. In order to avoid the discontinuity in Equation (8), we take it as a constant in a small environment around Mach 1—for $\mathcal{M} < 1.01$, we consider $f(\mathcal{M}) \equiv f(1.01)$, where the widest binary we consider corresponds to $\mathcal{M} \approx 0.97$.

In Figure 4 we consider different sound speeds, all of them in the supersonic regime. Higher sound speed leads to larger merger timescales, although the results are robust and do not change steeply between the different choices of sound speed in this regime.

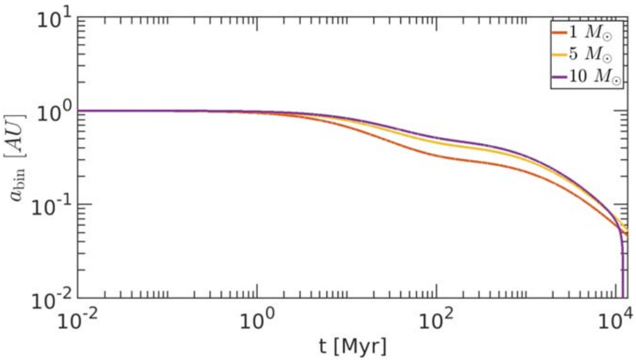

Figure 5. The effect of gas hardening on a binary, as dictated by GDF, for different masses of binaries. The different curves correspond to the evolution of the SMA for different binary masses, starting from an initial separation of $a_0 = 1$ au, given a background density with an exponential decaying gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\rm gas})$ with $\tau_{\rm gas} = 50$ Myr and $\rho_{g,0} = 1.74 \times 10^6\,M_\odot$ pc^{-3}.

In Figure 5, we demonstrate the dependence of gas hardening on different binary masses. As can be seen from Equation (7), lower-mass binaries harden over longer timescales, due to the dependence on the mass that scales as $\propto\sqrt{m}$, for an equal-mass binary with companions $m_1 = m_2 = m$. The final SMA of the binary also depends on the mass of the binary, such that more massive binaries will attain smaller final SMAs.

4.2. Comparison with Other Gas Hardening Models

Heretofore we have considered gas hardening induced by GDF. However, there are other approaches to model gas hardening.

In AGN disks, gas hardening is also modeled using processes similar to migration models in protoplanetary disks (as was suggested in the context of AGN disks; McKernan et al. 2012; Stone et al. 2017; Tagawa et al. 2020). Gas is captured in the Hill sphere of a binary and leads to the formation of a circumbinary minidisk. The disk applies a torque on the binary that leads to separation decay similar to migration type I/II in protoplanetary disks, although there were studies that pointed out that this torque could lead to a softening rather than hardening (Moody et al. 2019). Notwithstanding, we will assume that the formation of a minidisk can take place in GCs and compare the resulting hardening with our GDF model. The typical timescale for hardening due to migration torques is given by (e.g., McKernan et al. 2012), with the parameters relevant to our system,

$$\tau_{\rm type\ II} \sim 46\ {\rm yr}\left(\frac{0.01}{\alpha}\right)\left(\frac{0.23}{h/r}\right)^2\left(\frac{40{\rm yr}^{-1}}{\Omega_{\rm bin}(a_{\rm bin})}\right), \tag{12}$$

where α is the Shakura–Sunyaev parameter, h/r is the aspect ratio, and $\Omega_{\rm bin} = \sqrt{G(m_1 + m_2)/a_{\rm bin}^3}$ is the angular frequency of the binary. We adopt $h/r = 0.23$ and $\alpha = 0.01$ as a conservative value for the viscosity parameter of the disk. We substitute the $\Omega_{\rm bin}$ that corresponds to a binary with a separation of 1 au. Under these assumptions, the migration timescales, which could be used to approximate the hardening timescales, are shorter than the typical migration timescales we derived using the GDF model. These timescales are also shorter than the ones obtained in AGN disks (e.g., Stone et al. 2017; Tagawa et al. 2020), as expected. We therefore expect the merger rates we derived to be similar in this case, and even higher for the lowest gas densities models, where the rates were limited by slower hardening. There were more recent studies that suggested modified migration timescales, here taken for an equal-mass binary

$$\tau_{\text{type II,K}} = \frac{\Sigma_{\text{disk}}}{\Sigma_{\text{disk,min}}}\tau_{\text{type II}},$$

$$\Sigma_{\text{disk,min}} = \frac{\Sigma_{\text{disk}}}{1 + 0.04K},$$

$$K = \left(\frac{m_1}{m_1 + m_2}\right)^2 \left(\frac{h}{r}\right)^{-5} \alpha^{-1}. \tag{13}$$

These factors lengthen significantly the typical migration timescales, such that for our fiducial model we expect $\tau_{\text{typeII,K}} \approx 71{,}515$ yr. This timescale is still much shorter than the expected timescale calculated via the GDF model.

Another approach to modeling gas-induced inspirals is discussed in Antoni et al. (2019). They simulate Bondi–Hoyle–Lyttelton (BHL) supersonic flows and derive the corresponding energy dissipation, fitted to an analytical theory. While the overall gas hardening timescales could be comparable or shorter for the parameters that are in our major interest, there are significant differences in the scaling. The typical inspiral timescale is given by (Equation (52) in Antoni et al. (2019))

$$\tau_{\text{BHL}} = 61\ \text{Myr}\left(\frac{a_0}{\text{au}}\right)^{0.19}\left(\frac{v_{\text{rel}}}{100\ \text{km} \times \text{s}^{-1}}\right)^{3.38} \times \left(\frac{20\ M_\odot}{m_1 + m_2}\right)^{1.19}\left(\frac{7.72 \times 10^7\ \text{cm}^{-3}}{n_{\text{gas}}}\right), \tag{14}$$

where a_0 is the initial separation of the binary and n_{gas} is the number density of the gas, such that $\rho_{\text{gas}} = n_{\text{gas}} m_p$, where m_p is the proton mass.

Each model for gas hardening sets a different critical initial separation from which the binary will merge within a Hubble time. The timescales dictated from both the type II migration and BHL mechanism are even shorter than the ones expected by our fiducial model.

Hence, we will conclude that in all the approaches that we considered to model gas hardening the process is very efficient and leads to a robust rate of mergers, which modifies significantly the binaries' population, while the major difference between them is the time of the merger, dictated by the different gas hardening timescales.

4.3. Eccentric Evolution

The evolution of binaries in a gaseous medium is significantly different for noncircular binaries. Here we derive and solve the equations for an orbit-averaged eccentric evolution of an initially eccentric binary embedded in gas, but we leave a more detailed discussion on the implications for the dynamical three-body hardening of eccentric binaries to future studies.

For simplicity, we will assume that the Keplerian velocity of the binary components dominates the relative velocity to the gas, and that the gas velocity is zero relative to the center of mass of the binary. Hence, the relative velocity between the binary and the gas in the center-of-mass frame is given by

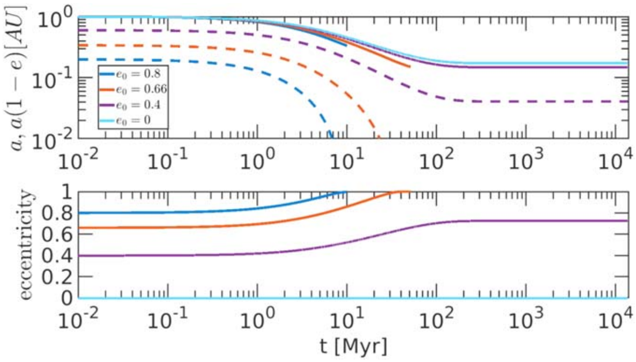

Figure 6. The effects of gas hardening on eccentric orbit. We consider the evolution of a binary with masses $m_1 = m_2 = 10\ M_\odot$ and initial separation of $a_0 = 1$ au. We consider an exponential decaying background gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\text{gas}})$ with $\rho_{g,0} = 1.74 \times 10^6\ M_\odot\ \text{pc}^{-3}$ and $\tau_{\text{gas}} = 50$ Myr. The solid lines correspond to SMA evolution and the dashed lines to pericenter evolution.

$$\boldsymbol{v}_{\text{rel}} = \frac{\Omega a}{2\sqrt{1 - e^2}}[e \sin f \hat{\boldsymbol{r}} + (1 + e\cos f)\hat{\boldsymbol{\varphi}}]. \tag{15}$$

The orbit equations for the GDF for a binary with two equal masses are then given by

$$\left.\frac{da}{dt}\right|_{\text{GDF}} = \frac{4a^{3/2}}{m\sqrt{2Gm(1 - e^2)}}[F_r e \sin f + F_\varphi(1 + e\cos f)], \tag{16}$$

$$\left.\frac{de}{dt}\right|_{\text{GDF}} = \frac{1}{2m}\sqrt{\frac{a(1 - e^2)}{2Gm}}[F_r \sin f + F_\varphi(\cos f + \cos E)], \tag{17}$$

where $\boldsymbol{F}_{\text{drag}} = F_r\hat{\boldsymbol{r}} + F_\varphi\hat{\boldsymbol{\varphi}}$, f is the true anomaly, and E is the eccentric anomaly. The orbit-averaged equations are given by

$$\left.\frac{da}{dt}\right|_{\text{GDF}} = \frac{4F_0(1-e^2)^2}{\pi m_{\text{bin}}\Omega^3 a^2}\int_0^{2\pi} \times \frac{I df}{(1 + e\cos f)^2\sqrt{1 + 2e\cos f + e^2}}, \tag{18}$$

$$\left.\frac{de}{dt}\right|_{\text{GDF}} = \frac{F_0(1-e^2)^3}{\pi m_{\text{bin}}\Omega^3 a^3}\int_0^{2\pi} \times \frac{I(e + \cos f)df}{(1 + e\cos f)^2(1 + 2e\cos f + e^2)^{3/2}}, \tag{19}$$

where F_0 is given by $\boldsymbol{F}_{\text{drag}} = F_0\boldsymbol{v}_{\text{rel}} I / v_{\text{rel}}^3$. The orbit-averaged equations for GWs are given by

$$\left.\frac{da}{dt}\right|_{\text{GW}} = -\frac{64G^3 m_1 m_2(m_1 + m_2)}{5c^5 a^3(1 - e^2)^{7/2}}\left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right), \tag{20}$$

$$\left.\frac{de}{dt}\right|_{\text{GW}} = -\frac{304G^3 e m_1 m_2(m_1 + m_2)}{15c^5 a^4(1 - e^2)^{5/2}}\left(1 + \frac{121}{304}e^2\right). \tag{21}$$

In Figures 6 and 7 we introduce the evolution of eccentric binaries. In Figure 6, we present the evolution due only to GDF, and in Figure 7, we also introduce the effect of GW emission. As can be seen, the eccentricities become extremely

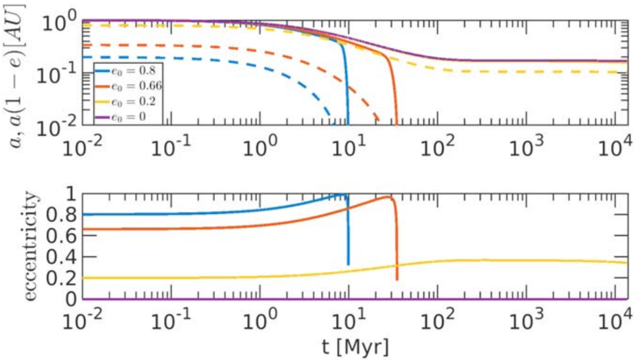

Figure 7. The effects of gas hardening and GWs on eccentric orbit. We consider the evolution of a binary with masses $m_1 = m_2 = 10\,M_\odot$ and initial separation of $a_0 = 1$ au. We consider an exponential decaying background gas density $\rho_g = \rho_{g,0}\exp(-t/\tau_{\rm gas})$ with $\rho_{g,0} = 1.74 \times 10^6\,M_\odot$ pc^{-3} and $\tau_{\rm gas} = 50$ Myr. The solid lines correspond to SMA evolution and the dashed lines to pericenter evolution.

high within short timescales, indicating that the pericenter shrinks significantly. Once the pericenters are sufficiently small, the effect of GWs becomes more significant, the orbit shrinkage is accompanied by eccentricity damping, and the binaries are driven into approximately circular orbit when entering the LIGO-Virgo-KAGRA (LVK) GW bands.

Such eccentric evolution could play a key role in the evolution of the binary populations, as eccentric binaries merge within potentially far shorter timescales than circular binaries. We note, however, that some studies of a circumbinary gas-disk evolution of binaries suggest that they are only excited to moderate eccentricities ∼0.45 (Tiede et al. 2020). Nevertheless, if binary migration occurs through such processes, the overall shrinkage is rapid irrespective of the eccentricity, leading to a fast migration timescale (see previous subsection).

We further discuss these issues, and in particular the implications for the delay time distribution of GW sources from this channel, in Section 4.4. We note that the consideration of eccentric binaries' gas hardening, little studied before, should play a similarly important role in binary evolution in AGN disks, possibly in a different manner than in cases where circumbinary disk evolution is assumed (Samsing et al. 2020; Tagawa et al. 2021).

4.4. Gravitational-wave Merger Rate

In the following we estimate the GW merger rate of binary BHs from the gas-catalyzed channel studied here. We will consider old-formed GCs and YMCs separately, given their different formation history.

In all the models we considered for gas hardening, all the binaries are expected to merge within a Hubble time. However, different gas hardening models suggest different merger timescales. As discussed above, our GDF models suggest that eccentric binaries merge rapidly, and some of the hydrodynamical studies discussed above suggest that even circular binaries merge during the early gas phase. Since most GCs formed very early, such mergers would not be detected by VLK, given the effectively limited look-back time. However, the younger equivalents of GCs, so-called YMCs, continue to form and generally follow the star formation history in the universe. Hence, mergers in such YMCs could occur sufficiently late (and hence closer by) and be detected by VLK, and the contribution of YMCs to the total VLK rate will be the dominant one for the eccentric cases (or for all binaries, according to, e.g., the circumbinary disk migration models). It should be noted that there is observational evidence for gas replenishment also in YMCs (e.g., Li et al. 2016). If, however, gas densities are lower or the binaries are initially circular/in low eccentricity, the final SMA of the binaries could be larger, leading to longer GW merger time catalyzed by three-body hardening (driving the delay time distribution to longer timescales), in which case the contribution from old GCs would be the dominant one.

The rates as a function of the redshift change according to the geometric structure of the 2P stars. Formation of 2P stars in disks is characterized by lower velocity dispersions, which lead to earlier mergers, where for the case of spherical constellation the higher velocity dispersion leads to later mergers.

We will start by estimating the number of mergers per cluster,

$$N_{\rm merge} \sim f_{\rm disk}\, f_{\rm bin,surv}\, f_{\geqslant 20\,M_\odot}\, f_{\rm ret}\, f_{\rm merge}\, N_\star, \tag{22}$$

where $f_{\rm disk}$ is the fraction of stars that reside in the disk, $f_{\rm bin,surv}$ is the fraction of binaries among massive stars that will survive stellar evolution (i.e., SNe), $f_{\geqslant 20\,M_\odot}$ is the fraction of stars with masses that exceed $20\,M_\odot$, $f_{\rm ret}$ is the retention fraction of BHs in the cluster, $f_{\rm merge}$ is the fraction of binaries that merge among the surviving binaries embedded in the disk, and $N_\star$ is the number of stars in the cluster.

Following our geometrical considerations in Section 2.1, we set $f_{\rm disk}$ in the range [2%, 20%]. However, even large fractions could be taken into account if there is a significant capture of objects to the disk.

The binarity fraction of massive BHs is ∼0.7, although even higher values are quite plausible for the massive-star progenitors of BHs (e.g., Sana et al. 2012); stellar evolution may reduce this fraction to a typical value of $f_{\rm bin,surv} = 0.1$ (e.g., Antonini & Perets 2012). We use a Kroupa mass function for the cluster, such that the fraction of stars with masses larger than $20\,M_\odot$ is 2×10^{-3} for a non-segregated environment; for segregated ones we take a fraction of 0.01. The retention fraction from the cluster is taken to be 10% (e.g., Kritos & Cholis 2020 and references therein). Taking into account the initial survival fraction of wide binaries, we consider $f_{\rm merge} \approx 0.49 - 0.61$ for our fiducial model. The lower value corresponds to massive background stars and the upper limit to low-mass background stars ($\bar{m} = 0.5\,M_\odot$); see the discussion below Equation (6).

Following Rodriguez et al. (2016), we consider logarithmically flat distribution of initial SMA in the range [10^{-2}, 10^5] au, where the lower limit is close to the point of stellar contact and the upper one to the Hill radius. It should be noted that although the choice of logarithmically flat is common, there were other choices of distribution considered, based on observational data (see Antonini & Perets 2012 for further discussion).

For our fiducial model, $N_\star = 10^5$ and $M_{\rm cluster} = 10^5\,M_\odot$.

In order to calculate the GW merger from old GCs, we follow the calculation of Rodriguez et al. (2016) and Kritos & Cholis (2020),

$$\mathcal{R}_{\rm old}(z) = \frac{1}{V_c(z)}\int_{z_{\rm min}}^{z} \Gamma_{\rm old}(z')\, n_{\rm old}(z')\, \frac{dV_c}{dz'}(1+z')^{-1}dz', \tag{23}$$

where $\Gamma_{\rm old}$ is the rate of mergers in old GCs; $n_{\rm old}$ is the GC number density, which is taken to be in the range [0.33,

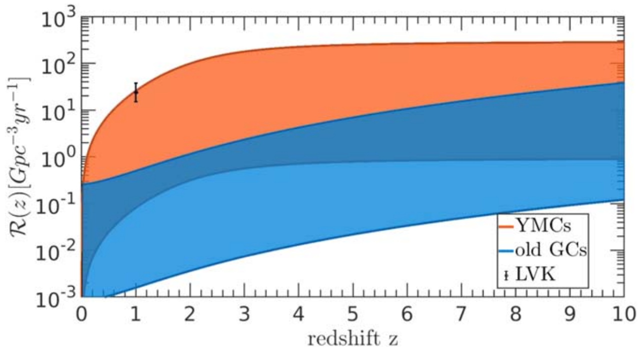

Figure 8. The cumulative contribution to GW rate from YMCs (in red) and old GCs (in blue), from the gas hardening channel, as derived from the GDF. The shaded area relates to the range of parameters. The black line relates to the range of rates inferred by LVK. In the case of circular binaries, the rate will be dominated by old GCs, while for eccentric binaries it will be dominated by YMCs.

2.57]$E^3(z)$ Mpc^{-3} (Portegies Zwart & McMillan 2000; Rodriguez et al. 2016; Kritos & Cholis 2020); dV_c/dz is the comoving volume; and $(1+z)^{-1}$ accounts for the time dilation. The comoving volume is given by Hogg (1999),

$$\frac{dV_c}{dz} = \frac{4\pi c^3}{H_0^3 E(z)}\left(\int_0^z \frac{dz'}{E(z')}\right)^2, \tag{24}$$

$$E(z) = \sqrt{\Omega_M(1+z')^3 + \Omega_\Lambda}, \tag{25}$$

where $\Omega_K = 0$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$ (Planck Collaboration et al. 2016).

As a conservative estimate, we take the merger rate $\Gamma_{\rm old}$ to be $\Gamma_{\rm old} \sim N_{\rm merge}/\tau_{\rm GC}$, where $\tau_{\rm GC}$ is taken to be 10 Gyr. In Figure 8 we present the cumulative rate of expected mergers in old GCs (in blue). There are two types of contributions to the rate: The first are eccentric binaries, such as those with initial eccentricity of 2/3, which corresponds to the mean value of a thermal eccentricity distribution, that will merge within short timescales, i.e., with negligible delay time. These practically follow the star formation rate (SFR). In this case observed contributions are likely to rise from YMCs. The second case corresponds to low-eccentricity/circular binaries, in which there will be a delay time that corresponds to a typical time of $\sim 10^4$ Myr. These contributions will be observed in old GCs.

In this case, the major contribution from our channel to currently observable GW sources would originate not from old GCs but from YMCs. We define a YMC as a cluster formed later than redshift 2 and mass $> 10^4\,M_\odot$ such that we assume for that case that the 2P formation already occurred. The formation rate of YMCs follows the SFR, which enables us to write the merger rate from YMC as (Banerjee 2021)

$$\mathcal{R}_{\rm young}(z) = \frac{N_{\rm mrg}}{N_{\rm samp}}\frac{1}{2\Delta t_{\rm obs}}\frac{\int_{M_{\rm cl,low}}^{M_{\rm cl,high}}\Phi_{\rm CLMF}(M)dM}{\int_{M_{\rm GC,low}}^{M_{\rm GC,high}}\Phi_{\rm CLMF}(M)dM} \times \frac{\int_0^z \Psi_{\rm SFR}(z_f)dz_f}{\int_3^6 \Psi_{\rm SFR}(z_f)dz_f}\rho_{\rm GC}. \tag{26}$$

$N_{\rm mrg}$ is the number of mergers expected in $N_{\rm samp}$ clusters, $\Delta t_{\rm obs} = 0.15$ Gyr (Banerjee 2021) is the uncertainty in the cluster formation epoch, $\Phi_{\rm CLMF} \propto M^{-2}$ (e.g., Portegies Zwart et al. 2010) is the cluster mass function, and we consider $[M_{\rm cl,low}, M_{\rm cl,max}] = [10^4, 10^5]\,M_\odot$ as the available mass range for YMCs and $[M_{\rm GC,low}, M_{\rm GC,high}] = [10^5, 10^6]\,M_\odot$ as the typical present-day masses for GCs. $\rho_{\rm GC}$ is the observed number density of GCs per unit comoving volume. $\Psi_{\rm SFR}(z)$ is the cosmic SFR, which is given by Madau & Dickinson (2014),

$$\Psi_{\rm SFR}(z) = 0.01\frac{(1+z)^{2.6}}{1+[(1+z)/3.2]^{6.2}}M_\odot\ {\rm Mpc}^{-3}\ {\rm yr}^{-1}. \tag{27}$$

We consider $N_{\rm mrg}/N_{\rm samp} = N_{\rm merge}$ and spatial densities in the range [0.33, 2.57] Mpc^{-3}, following Banerjee (2021 and references therein). In Figure 8, we present the cumulative rate of expected mergers in YMCs and GCs. For YMCs, the rate follows the SFR (in general, with a small correction due to the delay time—which is short) and hence peaks in relatively low redshifts. For the eccentric case, the dominant contribution will rise from YMC, while for circular ones the dominant contribution is from GCs.

It should be noted that, in general, there could be a nonnegligible delay time for the binary merger. However, for all the parameters we checked for the disk configuration, the merger timescales are extremely short and are negligible in terms of redshifts.

The total contribution to the GW merger rate from YMCs is in the range $\mathcal{R}_{\rm young} \approx [0.08, 25.51]$ Gpc^{-3}yr^{-1}, which intersects the expected range of LVK, i.e., $23.9^{+14.3}_{-8.6}$ Gpc^{-3} yr^{-1} (Abbott et al. 2021), where the range is bracketed by the models with lowest and highest rates (see Table 1).

In Table 1 we present our calculated rates for different choices of parameters. As expected, higher gas densities lead to larger merger rates and higher sound speeds correspond to thicker disks that host more stars and hence yield more mergers.

4.5. GW Merger Properties

Given the early epoch of gas replenishment, gas-catalyzed mergers operate on primordial binaries in the clusters. The merging components are therefore likely distributed similarly to the primordial distribution of binary components. However, even very wide binaries can merge in this channel compared with only relatively close binaries merging in, e.g., isolated binary evolution channels for GW mergers. This could give rise to significant differences in the expected masses and mass ratios of the merger objects.

Interaction with gaseous media could excite binaries to high eccentricities, due to the dependence of the drag force on the relative velocity between the gas and the binary, which changes along the orbit such that the effect is the strongest at the apocenter. Evolution of eccentric binaries hence shortens significantly the expected merger timescales, as larger separations correspond to small pericenters, in which GWs could dominate the evolution. In this case, eccentric merger could be observed in LVK, but this is not the case for slow mergers in the circular case.

We should remark in passing on the possibility of triples. In triples, the outer component migrates faster than the inner binary, potentially leading to an unstable configuration and effective chaotic three-body interaction (see, e.g., a reversed case of triples expanding owing to mass loss, leading to similar

Table 1
Rates from YMCs for Redshifts $z \leqslant 1$, for Different Choices of Parameters

Model	$\mathcal{R}_{\rm YMC}(z \leqslant 1)$ (Gpc^{-3} yr^{-1})	Model	$\mathcal{R}_{\rm YMC}(z \leqslant 1)$ (Gpc^{-3} yr^{-1})
$\rho_- c_{s-} n_+$	0.32	$\rho_+ c_{s-} n_+$	2.55
$\rho_- c_{s-} n_-$	0.08	$\rho_+ c_{s-} n_-$	0.64
$\rho_- c_{s+} n_+$	3.28	$\rho_+ c_{s+} n_+$	25.51
$\rho_- c_{s+} n_-$	0.82	$\rho_+ c_{s+} n_-$	6.35

Note. $\rho_\pm$ correspond to $\rho_{\rm GC} = 0.33E^3(z)$ Mpc^{-3} and $\rho_{\rm GC} = 2.57E^3(z)$ Mpc^{-3}, $c_{s\pm}$ correspond to $c_{s-} = 1$ km s^{-1} and $c_{s+} = 10$ km s^{-1}, and $n_\pm$ correspond to high density of progenitors and low fraction of hard binaries (n_+, segregated environment) and low density of progenitors and high fraction of hard binaries (n_-, non-segregated environment). These correspond also to different fractions of soft/hard binaries; see Section 3.1.2. Here we present the rates expected for initially eccentric binaries (e.g., $e_0 = 0.66$).

instability, in Perets & Kratter 2012); such chaotic encounters could give rise to eccentric mergers. This possibility and its potential contribution will be discussed elsewhere.

5. Discussion

In the following we discuss our results and implications for the evolution of binaries and singles in gas-enriched GCs.

5.1. Other Aspects of Binary Evolution

As we showed, the presence of gas modifies the binary population in GCs. It leads to an efficient merger of binaries, together with the formation of binaries via the L2 and L3 mechanisms (which were initially used to study the formation of Kuiper-belt binaries (Goldreich et al. 2002) and recently were applied to calculate the formation rate in AGN disks; Tagawa et al. 2020).

After the gas dissipation, the initial properties of the binary, as well as the gas, dictate the final separation, to which all the binaries with initial separations larger than the final separations will converge.

Therefore, gas hardening leaves a significant signature on the binary population and its properties, which sets the ground for further dynamical processes in general and specifically for later dynamical mergers.

In addition to the contribution of the channel to the total rate of GWs, the modification of the properties of binaries (mass, separation, etc.) caused by the gas hardening sets unique initial conditions for the other GW channels. This will induce an indirect signature of the gas hardening on the expected observed mergers. We introduced analytical results that could in principle be plugged in as initial conditions for the later evolution of GCs and the dynamical channels for GW production in such environments. The binary abundance changes owing to the gas hardening, since a significant fraction of binaries could merge, while others form. Furthermore, additional L2/L3-formed binaries could participate and produce GW sources, beyond the primordial binaries considered here. Nevertheless, since stars might be far more abundant than BHs, L2/L3 processes might mostly produce mixed BH—star binaries and may not contribute to the GW merger rate, but they may form other exotic binaries such as X-ray sources, etc., and/or produce microtidal disruption events (Perets et al. 2016; disruption of stars by stellar BHs).

5.2. Implications for Other Gas-rich Environments

The gas-catalyzed dynamics discussed here could take place in any other gas-rich environments, with the proper scaling. While enhanced GW merger rates were discussed in the context of AGN disks (McKernan et al. 2012; Stone et al. 2017; Tagawa et al. 2020, and references therein), they are usually discussed in the context of the evolution of a particular binary or the overall BH merger rate. However, in those cases too, the whole binary populations of both compact objects and stars will change their properties.

A very similar process could take place for young binaries embedded in star formation regions (Korntreff et al. 2012). In this case, the effect is limited to a shorter timescale and compact objects might not yet have formed and are therefore not directly affected (but their progenitor massive stars are).

5.3. YMCs and Very Massive Clusters

YMCs are still relatively little studied in the context of the production of GW sources, although their contribution to the total estimated rate of GWs is potentially not negligible (Portegies Zwart & McMillan 2000; Banerjee 2021). In these clusters, gas can be present up to smaller redshifts, such that the effect from the channel we suggested for GWs could potentially be observed. Hence, their overall contribution to the currently observed merger rates in LVK will be more significant (as can be seen also in Figure 8). Our rate estimates, discussed below, account for both GCs and their younger counterparts, YMCs.

5.4. Dynamics in Gas-enriched Clusters

All the dynamical processes that take place in the early stages of GC evolution might be affected by the presence of gas, e.g., few-body dynamics.

One aspect is that wide binaries that formed during the gas epoch are protected from evaporation by the gas hardening, as they harden within timescales shorter than the typical evaporation/ionization timescales.

GDF could also enhance mass segregation (Indulekha 2013; Leigh et al. 2014). The energy dissipation leads to a change in the velocity dispersion in short timescales, such that massive objects will fall toward the center of the cluster. Moreover, since the more massive objects are prone to merge (as can be seen from Equation (7), or visually from Figure 5), the relaxation will be affected by the modified mass function induced by the gas hardening.

5.4.1. GW Recoils, Spins, and Mass-gap Objects

It is possible that gas accretion onto binaries and not only GDF (e.g., Roupas & Kazanas 2019) could affect their evolution. In particular, sufficient accretion might align the BH spins and orbits, especially if some circumbinary disk forms around the binaries, in which case the GW-recoil velocity following mergers is likely to be small, and allow a larger fraction of merged, now more massive BHs to be retained in the cluster. This in turn would affect the later dynamics in the clusters and the resulting mergers in the dynamical formation channels operating in the clusters. This could then potentially give rise to higher fraction of BHs reaching high (even mass-gap) masses following repeated mergers. The spin evolution and accretion, however, require more detailed study, which is beyond the current scope.

The spin evolution of binaries will be affected by the role played by dynamical encounters, as well as the direction of the gas relative to the binary. In some cases, initially misaligned binaries could be aligned later owing to gas accretion, but when dynamical encounters are dominant, the spins will not be aligned.

5.5. Implications for Neutron Stars and White Dwarfs: Accretion and Explosive Transients

The focus of the current paper is the merger of BHs and the production of GW sources due to gas interactions in multiple-population clusters. However, the evolution of stars and other compact objects such as WDs and NSs could be significantly affected in similar ways. Though some of these aspects are discussed in a companion paper (Perets 2022), we postpone a detailed exploration of these objects to a later stage and only briefly mention qualitatively some potentially interesting implications.

A fraction of the gas could be accreted on objects in the cluster. Gas accretion changes the velocities of the accretors and the overall mass function of objects in the cluster, such that there is a shift toward higher masses (e.g., Leigh et al. 2014), which might affect the dynamical GW channels in clusters that operate after the gas replenishment epoch, since we enrich the abundance of massive objects that are likely to be the progenitors of GWs. Stars that accrete gas could evolve into compact objects that in turn might produce novae. Enhanced accretion in the early stages of the cluster evolution could potentially modify the nova rates and properties (Maccarone & Zurek 2012) and the production of accretion-induced collapse of WDs into NSs (Perets 2022).

We should point out that our scenario suggests a robust merger not only of BHs but also of NSs and WDs. These mergers might leave unique signatures. Besides their contribution to the production of short GRBs and GW sources, binary NS mergers are a promising channel to the production of heavy elements via *r*-process (e.g., Freiburghaus et al. 1999) and would affect the chemical evolution of the clusters.

Thermonuclear explosions of WDs could produce Type Ia SNe, whether via a single-degenerate channel (WD and a nondegenerate companion; Whelan & Iben 1973) or a double-degenerate channel (two WDs; Iben & Tutukov 1984). Both of these channels will be affected by the gas accretion. First, as we mentioned (Leigh et al. 2014 and references therein), the mass function will change. This in turn might change the characteristics of the SNe and their rate. Furthermore, regardless of the mass variation, a large fraction of the compact-object binaries are expected to merge within short timescales, which will also affect the SN rate.

Mergers of WDs could yield a remnant merged object with small or absent natal kick and hence constitute another channel for NS formation. Accretion could potentially change the retention fraction and potentially explain the retention problem in the formation of pulsars (Perets 2022).

5.6. Constraining the Parameters of the Cluster

The amount and origin of gas in GCs during the formation of 2P stars are still uncertain (Bekki 2017). In this channel, we suggest that the amount of gas dictates a final SMA, such that the separation distribution/GW rate could be used to constrain the gas abundance in the cluster and its lifetime.

For sufficiently low gas densities (or lower densities following gas depletion), gas hardening is not efficient enough to lead to a merger. In this case, the terminal SMA of the binary will exceed $a_{\rm GW}$, such that GWs will not be emitted without a further dissipation process. However, if the gas remains for longer timescales, further hardening will occur. For the whole parameter space we considered, the early stages of the hardening process are very efficient, i.e., wide binaries harden and become hard binaries on short timescales.

This channel of production of GW sources could serve as a tracer to later star formation, as it is coupled to the gas that accompanies this formation. The amount of gas and its decay with time are determined by the star formation history. Since these parameters play a role in gas hardening and hence in the final separation distribution at the end of the gas epoch, they could potentially serve to constrain the 2P gas and star formation phase and may help explain some of the differences between 1P and 2P stellar populations.

For example, we might speculate that the inferred difference between the 1P and 2P binary fractions (e.g., Lucatello et al. 2015) could be explained by gas-catalyzed hardening and mergers of MS stars residing in the gaseous region. Such 1P binaries that also accrete a significant mass of 2P gas would appear and be part of the 2P populations, while outside the gas regions binaries are not affected. In this case some of the 2P binaries preferentially merge compared with 1P stars outside the 2P gas region, leading to an overall smaller binary fraction.

That being said, the many uncertainties and degeneracies involved might be challenging in directly connecting current populations with the early conditions directly.

5.7. Caveats and Future Directions

In the following we discuss potential caveats of our model/scenario.

1. The specific scenario for formation of 2P stars is still unknown/debated, and hence there are large uncertainties in the amount of gas in the cluster and its source during the different stages of evolution. Moreover, some explanations for the different chemical composition of the so-called 2P stars might require lower gas masses than the total mass of 2P stars. In these cases, the phenomena we described might be somewhat suppressed, though, as we have shown that even lower gas densities could be highly effective and will not qualitatively change the results.
2. The expected production rates of GW sources depend on the initial parameters of the clusters we consider, including the gas densities, stellar and binary populations, star formation histories, etc. All of these contain many uncertainties, which we did not directly address in this initial study, limited to a small number of models so as to provide an overall estimate to bracket the expected GW rates from this channel. Nevertheless, all of our models show that gas-catalyzed

mergers in multiple-population clusters could produce a significant and even major contribution to the GW merger rate and could play a key role in the general evolution of stars and binaries in such clusters.

3. The interaction of gas with binaries is complex and includes many physical aspects. Here we assumed that the gas density in the cluster, or at least in the region in which the binaries evolve, is spatially constant. Most of the gas should be concentrated in the star-forming region, preferentially toward the inner parts of the cluster. Outer parts of the cluster might be more dilute. Future study could relax the simplified assumption of a constant spatial density and account for a more detailed distribution of gas, stars, and binaries.
4. We assumed that the relative velocity between the objects and the gas is dominated by the Keplerian velocity of the binary dominant. A more realistic approach, but requiring a detailed Monte Carlo or *N*-body simulation, could account for the detailed velocity distribution of binaries in the cluster.
5. As we mentioned in Section 5.5, objects embedded in gas could accrete from it and change their mass over time. As a result, their dynamics will change both in the cluster and as binaries (Roupas & Kazanas 2019). Here we considered constant masses throughout the evolution and neglected the effects of gas accretion. This is a somewhat conservative assumption, in regard to catalysis of mergers, as more massive objects are prone to merge even faster in gas (see Equation (7) and Figure 5).
6. We considered several choices for the gas depletion, assuming an exponential decay, with a fiducial model of 50 Myr and a lifetime of 100 Myr. However, the formation epochs of stars could set different scenarios, e.g., in which gas is abundant in the cluster for longer timescales of ~100 Myr, but only intermittently (Bekki 2017), which will change the picture, or when several wide-scale gas replenishment episodes occur over timescales of even many hundreds of Myr or even Gyr, as might be the case for nuclear clusters.
7. In our analysis we considered for simplicity only equal-mass binaries. Though we do not expect a major change in the results, the generalization to binaries with different masses is more complex and requires more detailed population studies, beyond the scope of the current study.
8. It should be noted that there were studies that suggested more limited efficiency of GDF (e.g., Li et al. 2020; Toyouchi et al. 2020) than considered here. A more detailed comparison is left for further studies.
9. Although the initial parameters of our disk suggest a thick disk, in later stages the disk will be thinner and finally fragment to enable star formation. Hence, for these stages/initial thin disks, the gas hardening epoch should be limited to the regime in which the disk is stable.
10. We restrict ourselves to binaries that are not likely to be disrupted by interactions with other stars. Further disruptions could take place and are encapsulated in $f_{\mathrm{bin,surv}}$ (see Equation (22)).

6. Summary

In this paper we discussed the evolution of binaries in gas-enriched environments that likely existed in the early-stage multiple-population clusters. We showed that the binary interaction with the ambient gas environment significantly affects their evolution and gives rise to major changes in binary population in the cluster and its properties.

Binaries' interaction with gas has been extensively studied over the past few years in the context of AGN disks. Here we show that the environments of multiple-population GCs and YMCs similarly give rise to important effects. In particular, focusing on the production of GW sources from binary BH mergers, we find that gas-enriched multiple-population clusters could provide a significant and possibly major contribution to the production of GW sources of up to a few tens of $\mathrm{Gpc}^{-1}\,\mathrm{yr}^{-1}$, comparable to the GW source production rate inferred by VLK for the local universe. These might even be higher once formation of new binaries due to gas-assisted capture is considered (to be discussed in a follow-up paper).

Moreover, we expect catalyzed mergers of other compact objects, such as NSs and WDs, and of binary MS and evolved stars to give rise to the enhanced rate of a wide range of merger outcomes, producing a range of transient events such as SNe, GRBs, and the formation of X-ray binaries and stellar mergers, which will be discussed elsewhere.

Furthermore, our findings on the overall evolution of binary populations are relevant for other gas-enriched environments such as AGN disks.

Finally, our focus here was on binary BH mergers in multiple-population cluster environments, but we point out that the early gas-enriched phase of such clusters (which in practice is relevant to the vast majority of GCs, given that most GCs show multiple populations) significantly affects all the stellar and binary populations and the overall dynamics inside GCs. Hence, the current modeling of the typical initial conditions in GCs and their evolution might need to be fundamentally revised.

We thank the anonymous referee for important comments and points that significantly improved the manuscript. We would also like to thank Aleksey Generozov, Johan Samsing, Jim Fuller, Kyle Kremer, Noam Soker, and Evgeni Grishin for fruitful discussions. M.R. acknowledges the generous support of Azrieli fellowship. H.B.P. and M.R. acknowledge support from the European Union's Horizon 2020 research and innovation program under grant agreement No. 865932-ERC-SNeX.

Appendix
Fiducial Parameters

Symbol	Definition	Fiducial Value
τ_{gas}	Gas lifetime	50 Myr
τ_{SG}	Formation time of SG	100 Myr
$M_{\star}$	Total mass of stars in cluster	$10^5\ M_{\odot}$
M_{gas}	Gas mass in the cluster	$3 \times 10^5\ M_{\odot}$
$\rho_{\mathrm{g,disk}}$	Initial gas density in disk	$1.74 \times 10^6\ M_{\odot}\ \mathrm{pc}^{-3}$
h/r	Scale height	0.23
σ_{disk}	Disk velocity dispersion	$10\ \mathrm{km\ s}^{-1}$
$\bar{m}$	Average stellar mass	$0.5\ M_{\odot}$
$n_{\star}$	Stellar density	$10^5\ \mathrm{pc}^{-3}$
$n_{\star,\mathrm{disk}}$	Stellar density in disk	$10^5\ \mathrm{pc}^{-3}$
c_s	Sound speed	$10\ \mathrm{km\ s}^{-1}$
$\log\Lambda_g$	Gas Coulomb logarithm	3.1

ORCID iDs

Mor Rozner https://orcid.org/0000-0002-2728-0132
Hagai B. Perets https://orcid.org/0000-0002-5004-199X

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, PhRvL, 116, 061102
Abbott, R., Abbott, T. D., Acernese, F., et al. 2021, arXiv:2111.03606
Abbott, R., Abbott, T. D., Abraham, S., et al. 2021, ApJL, 913, L7
Antoni, A., MacLeod, M., & Ramirez-Ruiz, E. 2019, ApJ, 884, 22
Antonini, F., & Perets, H. B. 2012, ApJ, 757, 27
Artymowicz, P., Clarke, C. J., Lubow, S. H., & Pringle, J. E. 1991, ApJL, 370, L35
Artymowicz, P., Lin, D. N. C., & Wampler, E. J. 1993, ApJ, 409, 592
Artymowicz, P., & Lubow, S. H. 1994, ApJ, 421, 651
Bahcall, J. N., & Ostriker, J. P. 1976, Natur, 262, 37
Banerjee, S. 2021, MNRAS, 503, 3371
Baruteau, C., Cuadra, J., & Lin, D. N. C. 2011, ApJ, 726, 28
Bastian, N., & Lardo, C. 2018, ARA&A, 56, 83
Bate, M. R. 2000, MNRAS, 314, 33
Bekki, K. 2010, ApJL, 724, L99
Bekki, K. 2011, MNRAS, 412, 2241
Bekki, K. 2017, MNRAS, 469, 2933
Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton, NJ: Princeton Univ. Press)
Bobrick, A., Zenati, Y., Perets, H. B., Davies, M. B., & Church, R. 2022, MNRAS, 510, 3758
Carretta, E., Bragaglia, A., Gratton, R. G., et al. 2009, A&A, 505, 117
Duffell, P. C., D'Orazio, D., Derdzinski, A., et al. 2020, ApJ, 901, 25
Er, X.-Y., Jiang, Z.-B., & Fu, Y.-N. 2009, ChA&A, 33, 139
Escala, A., Larson, R. B., Coppi, P. S., & Mardones, D. 2004, ApJ, 607, 765
Freiburghaus, C., Rosswog, S., & Thielemann, F. K. 1999, ApJL, 525, L121
Ginat, Y. B., & Perets, H. B. 2021a, MNRAS, 508, 190
Ginat, Y. B., & Perets, H. B. 2021b, PhRvX, 11, 031020
Goldreich, P., Lithwick, Y., & Sari, R. 2002, Natur, 420, 643
Gorti, U., & Bhatt, H. C. 1996, MNRAS, 283, 566
Gratton, R., Bragaglia, A., Carretta, E., et al. 2019, A&ARv, 27, 8
Grishin, E., & Perets, H. B. 2016, ApJ, 820, 106
Heggie, D. C. 1975, MNRAS, 173, 729
Hénault-Brunet, V., Gieles, M., Agertz, O., & Read, J. I. 2015, MNRAS, 450, 1164
Hogg, D. W. 1999, arXiv:astro-ph/9905116
Iben, I. J., & Tutukov, A. V. 1984, ApJS, 54, 335
Indulekha, K. 2013, JApA, 34, 207
Korntreff, C., Kaczmarek, T., & Pfalzner, S. 2012, A&A, 543, A126
Kritos, K., & Cholis, I. 2020, PhRvD, 102, 083016
Leigh, N. W. C., Böker, T., Maccarone, T. J., & Perets, H. B. 2013, MNRAS, 429, 2997
Leigh, N. W. C., Mastrobuono-Battisti, A., Perets, H. B., & Böker, T. 2014, MNRAS, 441, 919
Li, C., de Grijs, R., Deng, L., et al. 2016, Natur, 529, 502
Li, X., Chang, P., Levin, Y., Matzner, C. D., & Armitage, P. J. 2020, MNRAS, 494, 2327
Lucatello, S., Sollima, A., Gratton, R., et al. 2015, A&A, 584, A52
Maccarone, T. J., & Zurek, D. R. 2012, MNRAS, 423, 2
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Mastrobuono-Battisti, A., & Perets, H. B. 2013, ApJ, 779, 85
Mastrobuono-Battisti, A., & Perets, H. B. 2016, ApJ, 823, 61
Mastrobuono-Battisti, A., & Perets, H. B. 2021, MNRAS, 505, 2548
McKernan, B., Ford, K. E. S., Bellovary, J., et al. 2018, ApJ, 866, 66
McKernan, B., Ford, K. E. S., Lyra, W., & Perets, H. B. 2012, MNRAS, 425, 460
Miller, M. C., & Hamilton, D. P. 2002, MNRAS, 330, 232
Moody, M. S. L., Shi, J.-M., & Stone, J. M. 2019, ApJ, 875, 66
Muñoz, D. J., Lai, D., Kratter, K., & Miranda, R. 2020, ApJ, 889, 114
Muñoz, D. J., Miranda, R., & Lai, D. 2019, ApJ, 871, 84
Ostriker, E. C. 1999, ApJ, 513, 252
Ostriker, J. P. 1983, ApJ, 273, 99
Perets, H. B. 2022, ApJL, 927, L23
Perets, H. B., & Kratter, K. M. 2012, ApJ, 760, 99
Perets, H. B., Li, Z., Lombardi, J. C. J., & James, S. R. J. 2016, ApJ, 823, 113
Perets, H. B., & Murray-Clay, R. A. 2011, ApJ, 733, 56
Peters, P. C. 1964, PhRv, 136, 1224
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Portegies Zwart, S. F., & McMillan, S. L. W. 2000, ApJL, 528, L17
Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, ARA&A, 48, 431
Quinlan, G. D. 1996, NewA, 1, 35
Renzini, A., D'Antona, F., Cassisi, S., et al. 2015, MNRAS, 454, 4197
Rodriguez, C. L., Chatterjee, S., & Rasio, F. A. 2016, PhRvD, 93, 084029
Roupas, Z., & Kazanas, D. 2019, A&A, 621, L1
Samsing, J., Bartos, I., D'Orazio, D. J., et al. 2020, arXiv:2010.09765
Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Sci, 337, 444
Sigurdsson, S., & Phinney, E. S. 1995, ApJS, 99, 609
Soker, N. 2016, NewAR, 75, 1
Spitzer, L. 1987, Dynamical evolution of globular clusters (Princeton, NJ: Princeton Univ. Press)
Stone, N. C., Metzger, B. D., & Haiman, Z. 2017, MNRAS, 464, 946
Tagawa, H., Haiman, Z., & Kocsis, B. 2020, ApJ, 898, 25
Tagawa, H., Kimura, S. S., Haiman, Z., et al. 2022, ApJ, 927, 41
Tagawa, H., Kocsis, B., Haiman, Z., et al. 2021, ApJL, 907, L20
Tang, Y., MacFadyen, A., & Haiman, Z. 2017, MNRAS, 469, 4258
Tiede, C., Zrake, J., MacFadyen, A., & Haiman, Z. 2020, ApJ, 900, 43
Toyouchi, D., Hosokawa, T., Sugimura, K., & Kuiper, R. 2020, MNRAS, 496, 1909
Vesperini, E., McMillan, S. L. W., D'Ercole, A., & D'Antona, F. 2010, ApJL, 713, L41
Whelan, J., & Iben, I. J. 1973, ApJ, 186, 1007
Zenati, Y., Bobrick, A., & Perets, H. B. 2020, MNRAS, 493, 3956
Zenati, Y., Perets, H. B., & Toonen, S. 2019, MNRAS, 486, 1805

5.2 Gas-assisted binary formation

Based on Rozner et al. (2023)

Gas-rich environment revises not only the evolution of already-formed binaries, but could also induce the formation of new ones. Binary formation from an unbound pair requires a dissipation mechanism. Several mechanisms such as tidal dissipation (Press and Teukolsky, 1977) and dynamical friction (Goldreich et al., 2002a) were suggested. Here we suggest gas-induced binary formation (Tagawa et al., 2020; Rozner et al., 2023), which is a general mechanism that acts in various astrophysical systems in different scales, including star-forming regions, AGN disks and the gaseous epoch of second (or further) generation star formation in globular clusters.

There are three major criteria to enable gas-assisted captures: (I) The environment should be dense enough with gas and captured objects (II) The velocity should not exceed a critical velocity, to enable a capture rather than a shear (III) The two objects should initially reside within the same Hill sphere, such that the force between them will dominate over the external potential of the environment.

We derived analytically these criteria and verified our results with N-body simulations. We studied the available phase space for captures in the different environments and discussed possible dynamical implications.

The critical velocity could be calculated by equating the initial energy of an unbound pair to the energy dissipated by gas dynamical friction, taking into consideration the different regimes of gas dynamical friction law (supersonic/subsonic) and focusing (focused/unfocused).

Thus capture occurs for velocities smaller than the critical velocity v_{crit} defined by

$$\frac{1}{2}\mu v_{\rm crit}^2 = \Delta E_{\rm GDF} \approx \mathbf{F}_{\rm GDF}\left(m_1, v_1(v_{\rm crit}), v_g\right) \cdot \boldsymbol{\ell}_1 + \mathbf{F}_{\rm GDF}\left(m_2, v_2(v_{\rm crit}), v_g\right) \cdot \boldsymbol{\ell}_2 \quad (5.12)$$

For the different regimes, the critical velocity is given by

Table 5.1: Maximum velocities for capture in different regimes of GDF, with no headwind.

	Supersonic	Subsonic
Unfocused	$v_x q^{1/4}(1+q)^{3/4}$	$v_s q$
Focused	$\frac{v_x^2}{v_{\rm esc}} \frac{(1+q)^{1/2}}{q}$	$\frac{\sqrt{8 q v_s v_{\rm esc}}}{1+q}$

$v_x = \left(8\pi G^2 \rho_g m_{\rm bin} R_{\rm Hill} \ln \Lambda\right)^{1/4}$

$v_{\rm esc} = \sqrt{\frac{2Gm_{\rm bin}}{R_{\rm Hill}}}$

$v_s = \frac{8\pi G^2 \rho_g m_{\rm bin} R_{\rm Hill}}{3c_s^3}$

To validate our analytic results, we carried out a few-body simulation using `REBOUND` and `REBOUNDx` (Rein and Liu, 2012b; Tamayo et al., 2020). Here, we will show our results for some initial conditions

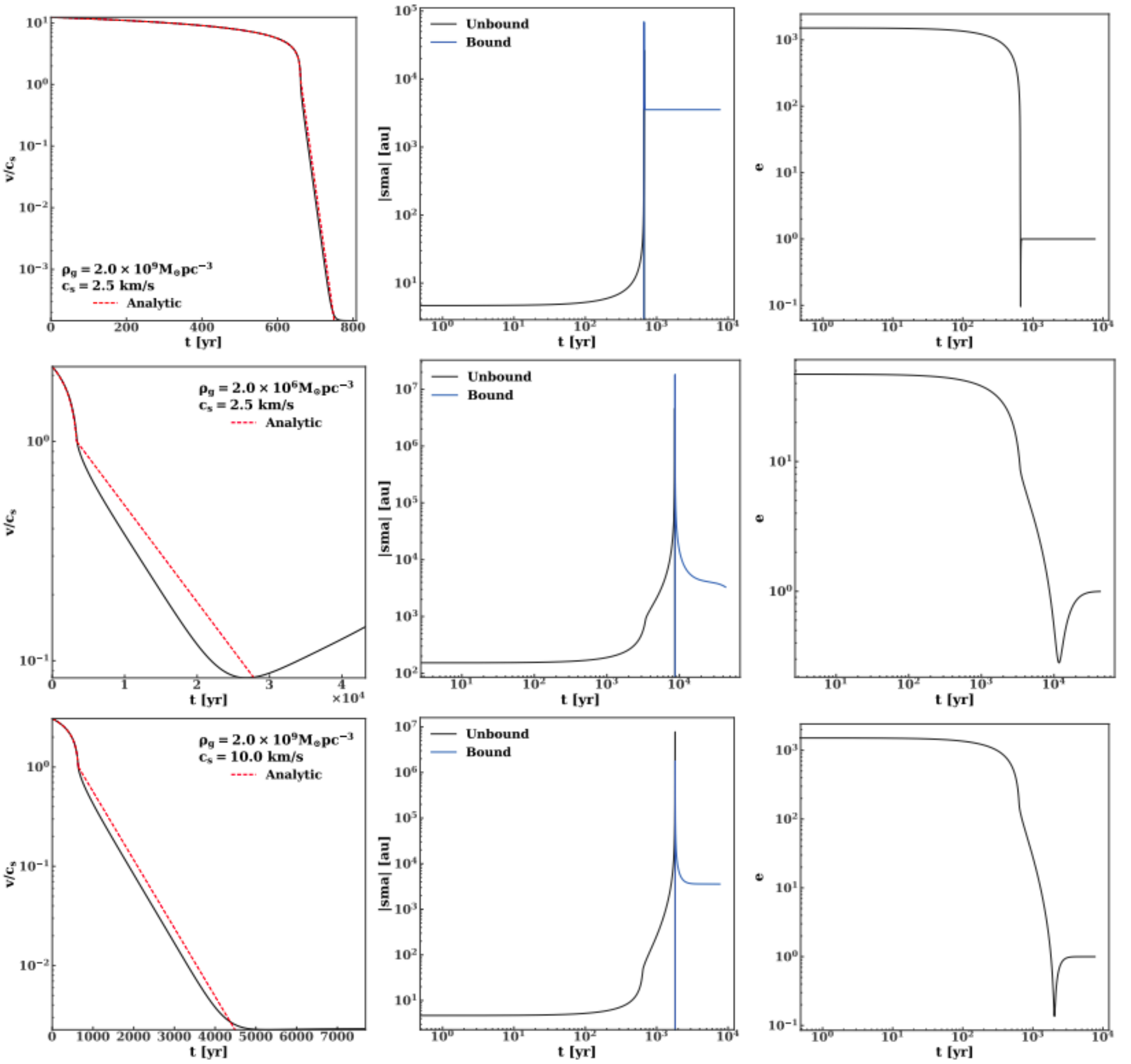

Figure 5.4: Evolution of the stellar velocities and binary orbital elements from numerical simulations with different gas density and sound speeds.

On the left column, we present the velocity evolution, in the middle column the semimajor axis evolution, and on the right column the eccentricity evolution. In the early stages, the velocity decreases, as the objects are decelerating until the capture, when the velocity starts to increase due to the inspiral of the binary. The semimajor axis increases until the capture, in which it theoretically diverges. as $|E| \to 0$ or equivalently, $|a| \to \infty$. Then, during the inspiral, the semimajor axis shrinks as we discussed on the previous chapter. The eccentricity starts from a high value, as the objects are unbound, and slightly after the capture the eccentricity increases steeply. We have a good agreement between the analytical and numerical results.

After setting the conditions for a capture, we calculated the capture rate for every environments by

$$\begin{aligned}
&\Gamma(m_1, m_2) \approx \int_0^{v_{\rm crit}} n_\star(m_2|m_1)\mathcal{A}vp(v)dv, \\
&\mathcal{A} = R_{\rm Hill} z(1+\Theta^2) \\
&z = \min\{R_{\rm Hill}, h_{\rm eff}\}
\end{aligned} \tag{5.13}$$

where m_1 is the mass of the capturer, m_2 is the captured mass, $n_\star(m_2|m_1)$ is the density of candidates for captured masses in the vicinity of m_1, $h_{\rm eff}$ is the effective scale height of the disk (if the environment has a disk-like configuration; e.g. an AGN disk, a gaseous disk in a cluster or a protoplanetary disk), $p(v)$ the velocity distribution and $\Theta = (v_{\rm esc}/v)^2$ is a correction for gravitational focusing. This correction is not valid in systems dominated by a massive central object like AGNs and protoplanetary discs, since it is derived assuming unperturbed two-body trajectories and neglects shearing motion. The critical velocity for capture, $v_{\rm crit}$, is calculated according to the regime (focused/unfocused), as specified in Table 5.1. We assume this is a Maxwellian distribution, such that $p(v) \propto v^2 e^{-v^2/2\sigma^2}$, where σ is the velocity dispersion.

Gas-assisted captures turned out to be an efficient channel in star-forming regions, that scales with the masses of the capturers and the captured objects. Higher masses are more likely to capture objects, and tend to capture smaller masses. The capture rate increases with the gas density.

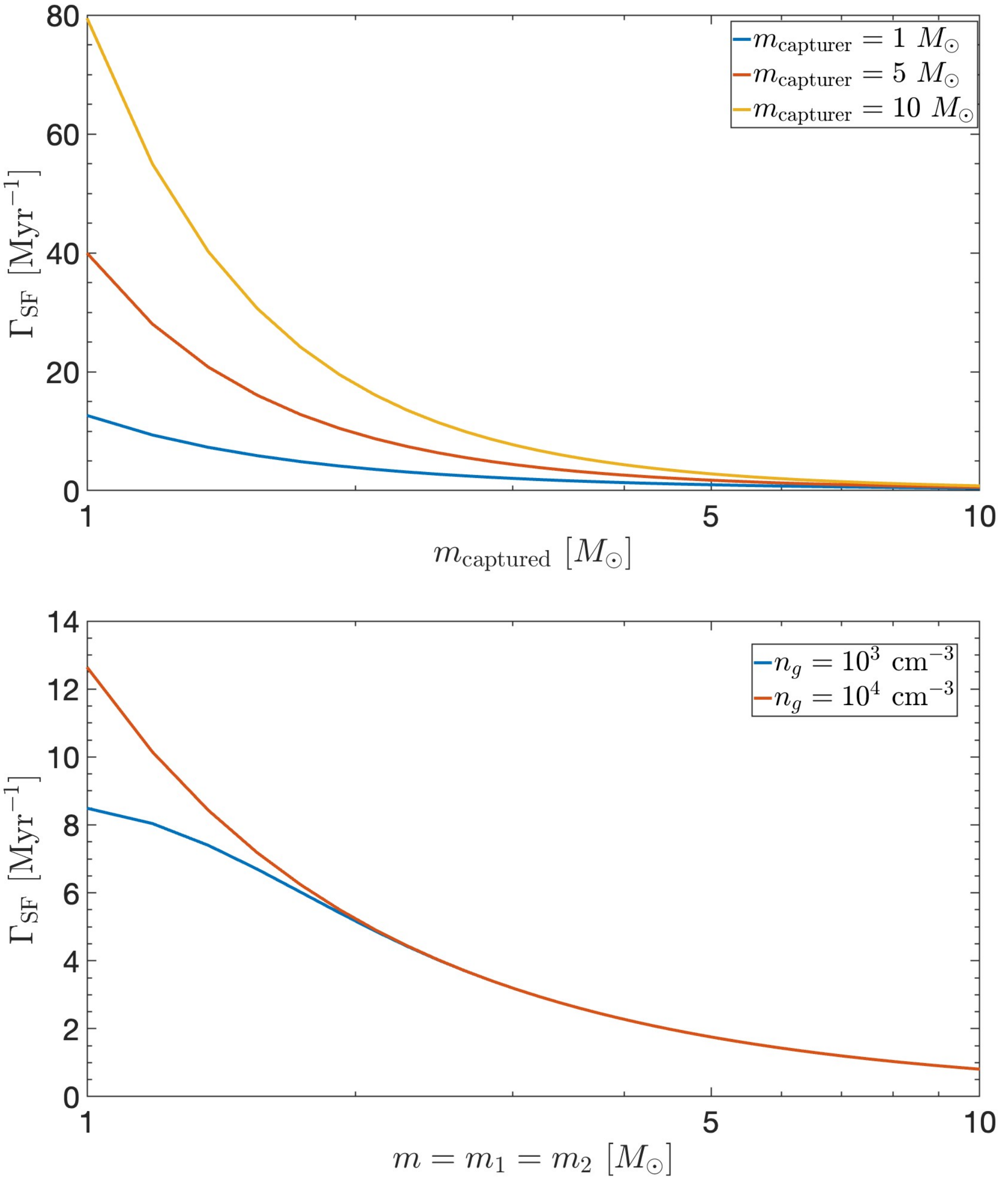

Figure 5.5: Left: The capture rate per object for different masses, in a SF environment, for our fiducial model specified above. Right: The capture rates for equal masses, given different gas number densities.

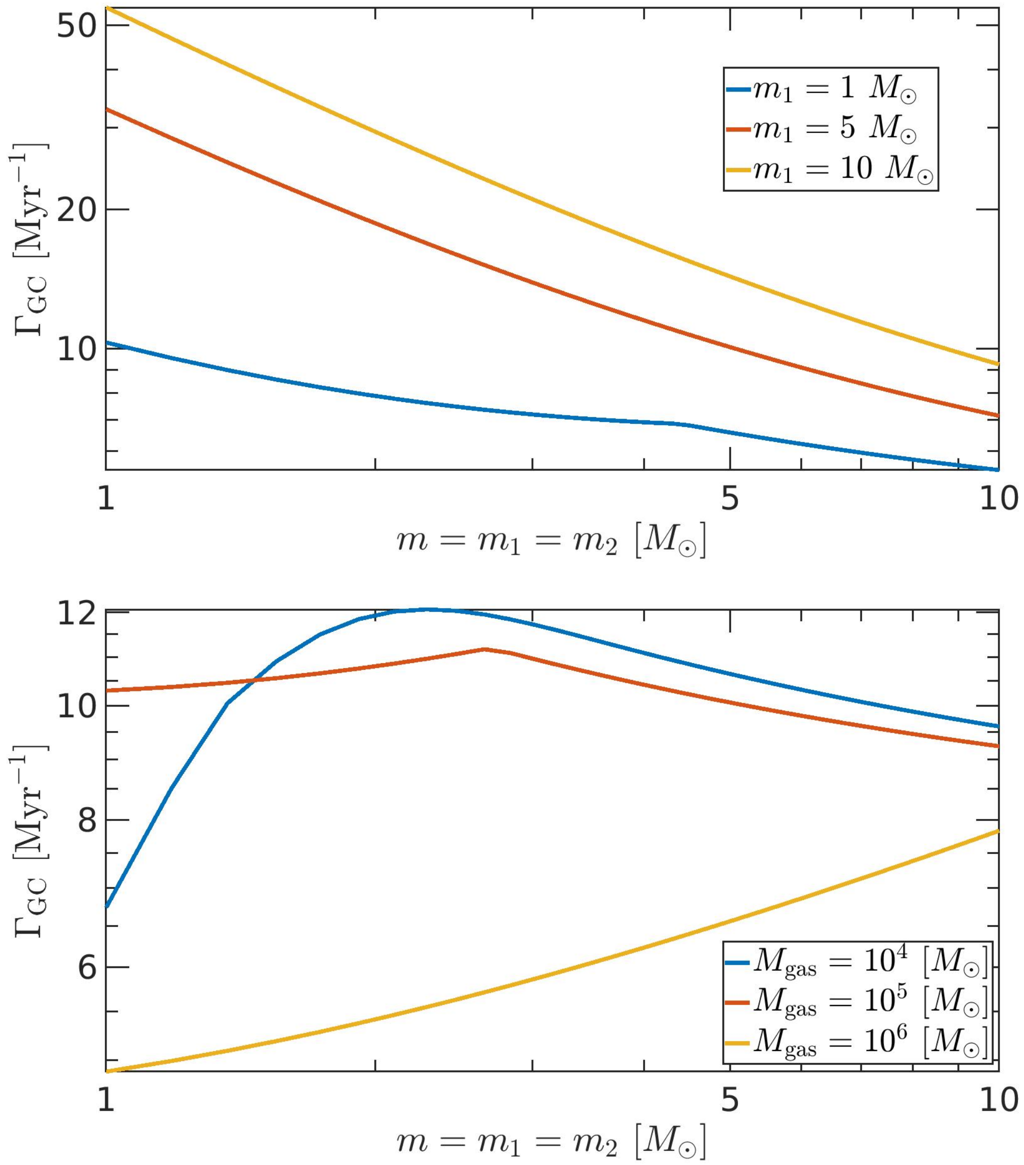

Figure 5.6: Left: The capture rate per object for different masses, in a second (or later) star generation environment in a GC, for our fiducial model specified above. Right: The capture rates for equal masses, given different gas masses.

This channel is efficient also in globular clusters (Fig 5.6). The dependence of gas-assisted capture rates do not scale monotonically with the gas mass, since for some masses, the Hill radius is determined mainly by the gas mass, while for others it is determined mainly by the captured objects.

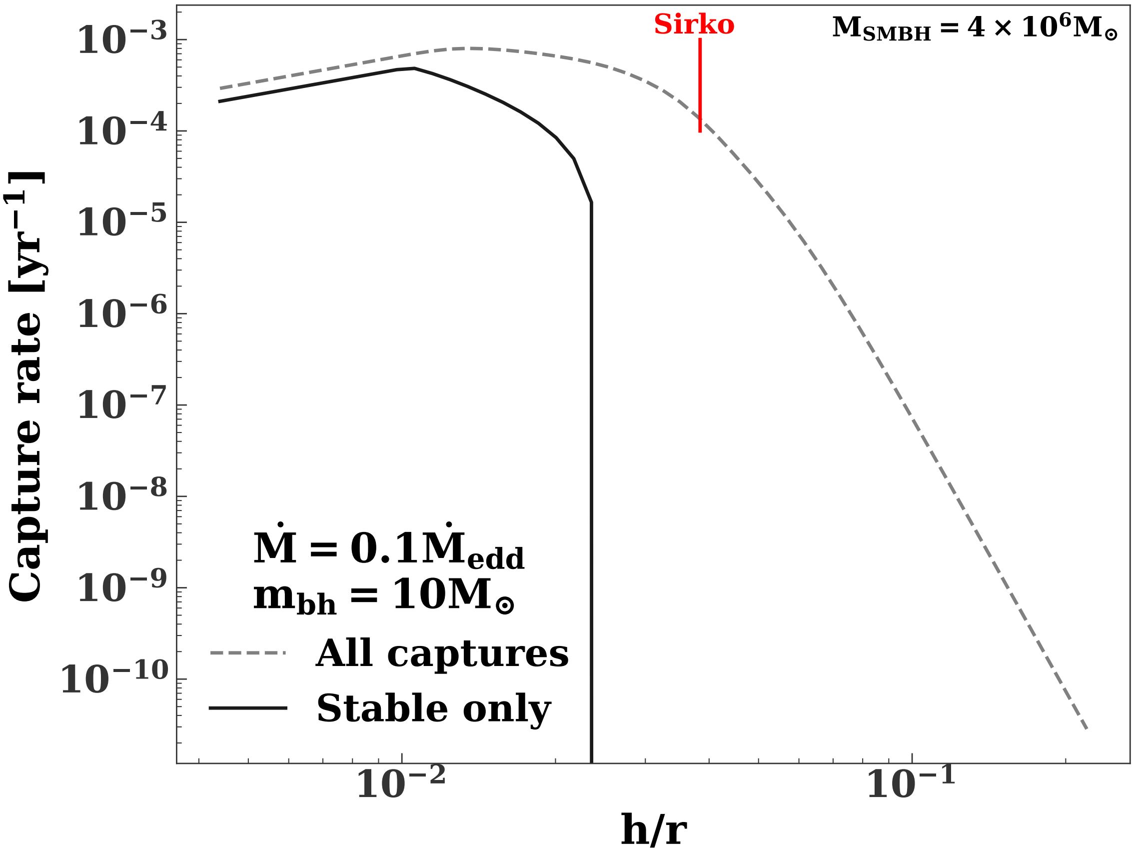

Figure 5.7: Total BH-BH capture rate for artificially inflated AGN disks with $\dot{M}/M_{\rm Edd} = 0.1$. The x-axis shows the aspect ratio at 3 pc after rescaling. If the disk aspect ratio is increased by an order of magnitude, the rate of stable captures goes to 0. The vertical red line shows the aspect ratio for the Sirko and Goodman (2003) model for fiducial parameters

In Fig. 5.7, we present the capture rates in AGN disks for different aspect ratios. Also here the gas-assisted binary formation mechanism is robust. It should be noted that a lot of captures might be unstable, i.e. only 'transient' captures. Similar properties are likely to occur also on the rest of the gaseous environments.

To summarize, here we studied gas-assisted binary formation in various gaseous environments. We formulated the analytical conditions for each capture and validated our results using N-body simulations. We also followed in detail the evolution circa the capture. It should be noted that here we studied the conditions for a capture rather than following the evolution long after the capture, and its properties.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **521,** 866–880 (2023)
Advance Access publication 2023 February 24

https://doi.org/10.1093/mnras/stad603

Binary formation through gas-assisted capture and the implications for stellar, planetary, and compact object evolution

Mor Rozner ,⋆ Aleksey Generozov and Hagai B. Perets
Technion - Israel Institute of Technology, Haifa 3200003, Israel

Accepted 2023 February 20. Received 2023 February 4; in original form 2022 December 1

ABSTRACT

Binary systems are ubiquitous and their formation requires two-body interaction and dissipation. In gaseous media, interactions between two initially unbound objects could result in gas-assisted binary formation, induced by a loss of kinetic energy to the ambient gas medium. Here, we derive analytically the criteria for gas-assisted binary capture through gas dynamical friction dissipation. We validate them with few-body simulations and explore this process in different gas-rich environments: gas-embedded star-forming regions (SFR), gas-enriched globular clusters, active galactic nucleus (AGN) discs, and protoplanetary discs. We find that gas-assisted binary capture is highly efficient in SFRs, potentially providing a main channel for the formation of binaries. It could also operate under certain conditions in gas-enriched globular clusters. Thin AGN discs could also provide a fertile ground for gas-assisted binary capture and in particular the formation of black hole/other compact object binaries, the production of gravitational-wave (GW) and other high-energy transients. Large-scale gaseous discs might be too thick to enable gas-assisted binary capture and previous estimates of the production of GW sources could be overestimated, and sensitive to specific conditions and the structure of the discs. In protoplanetary discs, while gas-assisted binary capture can produce binary Kuiper-belt objects, dynamical friction by small planetesimals is likely to be more efficient. Overall, we show that gas-assisted binary formation is robust and can contribute significantly to the binary formation rate in many environments. In fact, the gas-assisted binary capture rates are sufficiently high such that they will lead to multicaptures, and the formation of higher multiplicity systems.

Key words: black hole physics – galaxies: active – galaxies: star formation – *(Galaxy:)* globular clusters: general – *(stars:)* binaries *(including multiple)*: close.

1. INTRODUCTION

Binary systems are ubiquitous over a wide range of scales and in different astrophysical systems, from binary planetesimals in the Solar system, through stellar binaries and compact objects and up to the scales of binary massive black holes (MBHs). Indeed, the majority of stars reside in binaries, or even higher multiplicity systems (e.g. Raghavan et al. 2010; Sana et al. 2012; Duchêne & Kraus 2013; Moe & Di Stefano 2017), and a large fraction of Kuiper-belt objects (KBOs) reside in binaries (see a detailed review in Noll et al. 2008).

Binaries play a key role in the dynamics and evolution of stars and compact objects. In particular, close interactions between binary companions could lead to mass transfer or even mergers and collisions of the binary components. In turn, these interactions could give rise to the production of exotic stars and compact objects, which could otherwise not form from the evolution of single stars. Some compact binaries produce high energy emission (e.g. X-ray binaries) and mergers could result in explosive transient events such as supernovae (SNe), gamma-ray bursts, and/or the production of gravitational-wave (GW) sources.

Understanding the formation of binaries and their properties is therefore essential for decoding the evolution of stellar and planetary systems.

Several prominent binary formation channels were explored in the literature (see Lee et al. 2020 and the references therein). These could generally be divided between primordial formation of binaries, where the binary components form together as bound systems, and dynamical formation channels, where each of the binary component forms independently, and later dissipative processes bind them together to form a binary. The former involves the fragmentation of a bound blob of gas/dust in which two objects form and orbit each other (e.g. Nesvorný, Youdin & Richardson 2010). The latter involves dissipation mechanisms, where various channels were suggested to form binaries: (1) Tidal forces (Fabian, Pringle & Rees 1975; Press & Teukolsky 1977), GW emission, or even collisions, all of which become effective only through very close encounters between the progenitor unbound components; (2) three-body encounters, where the gravitational perturbation transfers kinetic energy between the components, until one is ejected with higher velocity, leaving behind a bound binary (Aarseth & Heggie 1976); and (3) dynamical friction (Goldreich, Lithwick & Sari 2002) and gas dynamical friction (GDF; Tagawa, Haiman & Kocsis 2020) where two objects embedded in a sea of far less-massive particles or in gas dissipate their excess kinetic energy to the ambient medium, leaving behind a bound binary.

⋆ E-mail: morozner@campus.technion.ac.il

Published by Oxford University Press on behalf of Royal Astronomical Society

Although gas-rich environments are quite common, the latter gas-assisted capture scenario was little studied, although recently these environments gained more focus in this context (e.g. Rowan et al. 2022; Boekholt, Rowan & Kocsis 2023; Li et al. 2023). Here, we explore this scenario analytically and using few-body simulations. We provide the specific conditions in which gas-assisted captures occur, and the dependence on the progenitor components, the ambient gas environment, and the limitations put by external potentials. We then use these calculations to explore the implications of gas-assisted capture to the formation of binaries over a wide range of environments, assess its importance, and the capture rates expected from this channel. In particular, we focus on active galactic nucleus (AGN) discs, where gas-assisted capture could play a key role in the production of GW sources; in star-forming regions (SFRs), where gas-assisted capture could serve as the main channel for binary formation; and in globular clusters (GCs) and protoplanetary disc, for which we find the gas-assisted capture is likely to be far less efficient.

In Section 2, we introduce the model of GDF. In Section 3, we discuss the conditions for gas-assisted binary formation and derive criteria for such a capture, using numerical and analytical methods. We then discuss the implications of our results and the probability for gas-assisted captures in several astrophysical environments (Section 5): SF environments (Section 5.1), second or later generations of GCs (Section 5.3), AGN discs (Section 5.2), and the Kuiper-belt (Section 5.4). In Section 6, we discuss the caveats of our model. In Section 7, we discuss the heating and cooling related to gas-assisted captures. Finally, in Section 8, we summarize our findings.

2. GAS DYNAMICAL FRICTION

There are several models to describe the dynamics of objects in gas, among them are evolution in gas-rich minidiscs (e.g. Stone, Metzger & Haiman 2017) and GDF. Here, unless stated otherwise, we will focus on GDF.

The GDF force on an object with mass m is (Ostriker 1999)

$$\mathbf{F}_{\rm GDF} = -\frac{4\pi G^2 m^2 \rho_{\rm g}}{v_{\rm rel}^3}\mathbf{v}_{\rm rel} I(v/c_{\rm s}) \tag{1}$$

where G is the gravitational constant, $\rho_{\rm g}$ is the gas density, $c_{\rm s}$ is the sound speed, and $\mathbf{v}_{\rm rel}$ is the relative velocity between the object and the gas. The function I is given by

$$I(\mathcal{M}) = \begin{cases} \frac{1}{2}\log(1-\mathcal{M}^{-2}) + \ln\Lambda, & \mathcal{M} > 1 \\ \frac{1}{2}\log\left(\frac{1+\mathcal{M}}{1-\mathcal{M}}\right) - \mathcal{M}, & \mathcal{M} < 1 \end{cases}. \tag{2}$$

Here, $\ln\Lambda$ is the Coulomb logarithm.[1] For $\mathcal{M} \gg 1$, I is nearly independent of the Mach number. Thus, for simplicity, we use the following modified function instead:

$$I(\mathcal{M}) = \begin{cases} \ln\Lambda & \mathcal{M} \geq 1 \\ \min\left\{\ln\Lambda, \frac{1}{2}\log\left(\frac{1+\mathcal{M}}{1-\mathcal{M}}\right) - \mathcal{M}\right\} & \mathcal{M} < 1 \end{cases}. \tag{3}$$

Following (Tagawa et al. 2020), we assume $\ln\Lambda = 3.1$. For numerical stability, we use the series expansion $I(\mathcal{M}) \approx \mathcal{M}^3/3 + \mathcal{M}^5/5$ for $\mathrm{M} < 0.02$ in our numerical calculations. The energy and angular momentum of the captured binary are (correspondingly)

$$E = -\frac{G m_1 m_2}{2a}, \quad L = \mu_{\rm bin}\sqrt{G M_{\rm bin} a(1-e^2)}, \tag{4}$$

respectively.

---

[1] For a finite time perturbation $\ln\Lambda$ is a function of time.

3. GAS-ASSISTED CAPTURE

Energy dissipation induced by GDF could lead, under conditions we describe later, to binary formation, similarly to the L2 mechanism (Goldreich et al. 2002). While L2 relies on dissipation induced by dynamical friction by other KBOs, we focus on GDF (see also Tagawa et al. 2020).

Generally, capture occurs if the energy dissipated during the passage of the objects is larger than the initial free unbound energy. Then the binary will be left bound at least momentarily. However, further evolution could unbind the binary or harden it. The discussion on further evolution is left for future work.

In this section, we derive both analytically and numerically the conditions for gas-assisted binary capture.

3.1 Threshold velocity for capture

The maximum initial velocity where capture occurs can be estimated by equating the work done by dynamical friction to the initial energy of the (unbound) orbit. We also require that capture occurs within the Hill sphere, where the gravity of the two-body system dominates tidal forces. If the separation of particles exceeds the Hill radius. Tidal forces from other objects [e.g. the central black hole (BH) or star], would dominate the gravity of the two bodies, and they would be torn apart.

Thus capture occurs if,

$$\frac{1}{2}\mu v_\infty^2 = \Delta E_{\rm GDF} \approx \mathbf{F}_{\rm GDF}(m_1, v_1, v_{\rm g}) \cdot \boldsymbol{\ell}_1 + \mathbf{F}_{\rm GDF}(m_2, v_2, v_{\rm g}) \cdot \boldsymbol{\ell}_2 \tag{5}$$

where ℓ_i is the typical length-scale in which mass, m_i, dissipates its energy; μ is the reduced mass of the two-body system; and $v_\infty = v_1 - v_2$ is the relative velocity at infinity.

For simplicity, we assume that the gas centre-of-mass is at rest with respect to the binary centre-of-mass, we discuss the effect of a headwind in Section 3.4. Thus, v_1 and v_2 are the initial velocities in the centre-of-mass frame and $m_1 v_1 = m_2 v_2$. It should be noted that by construction, the momentum could not be conserved, due to the action of the external force, but we consider only a local conservation as an approximation for short time-scales. In general, $\ell_1/\ell_2 = q^\alpha$. The power-law index, α, is 2 in the subsonic case and 5 in the supersonic case (see Appendix A). The maximum of ℓ_1 and ℓ_2 cannot exceed the Hill radius, to ensure a gravitational interaction between the objects.

In the supersonic regime, the time-scale for deceleration decreases rapidly with the stars' velocity. Thus, the time-scale for the binary elements to evolve after capture will be shorter than the initial capture time-scale. By geometry, the binary's initial period will be comparable to or greater than this capture time. Thus, the binary's orbital elements will evolve over a dynamical time-scale (i.e. they will change significantly over one orbital period).

We present a derivation of the threshold velocity for capture in Appendix B, and summarize the results below. In the supersonic regime, capture occurs if the initial relative velocity at infinity is less than

$$\begin{aligned} v_{c,1} &= \begin{cases} v_x q^{1/4}(1+q)^{3/4}, & v_x \gg v_{\rm esc} \\ \frac{v_x^2}{v_{\rm esc}}\frac{(1+q)^{1/2}}{q}, & v_x \ll v_{\rm esc} \end{cases} \\ v_x &= \left(8\pi G^2 \rho_{\rm g} m_{\rm bin} R_{\rm Hill} \ln\Lambda\right)^{1/4} \\ v_{\rm esc} &= \sqrt{\frac{2 G m_{\rm bin}}{R_{\rm Hill}}} \end{aligned} \tag{6}$$

Table 1. Maximum velocities for capture in different regimes of GDF, with no headwind.

	Supersonic	Subsonic
Unfocused	$v_x q^{1/4}(1+q)^{3/4}$	$v_s q$
Focused	$\frac{v_x^2}{v_{\rm esc}}\frac{(1+q)^{1/2}}{q}$	$\frac{\sqrt{8 q v_s v_{\rm esc}}}{1+q}$

Note. $v_x = (8\pi G^2 \rho_g m_{\rm bin} R_{\rm Hill} \ln\Lambda)^{1/4}$; $v_{\rm esc} = \sqrt{\frac{2 G m_{\rm bin}}{R_{\rm Hill}}}$; $v_s = \frac{8\pi G^2 \rho_g m_{\rm bin} R_{\rm Hill}}{3 c_s^3}$.

where $R_{\rm Hill}$ is the Hill radius, $m_{\rm bin}$ is the total mass, and q is mass ratio (between the secondary and primary masses). The first line corresponds to the threshold, neglecting the effects of gravitational focusing. This estimate is appropriate if the relative velocity at infinity is much greater than the escape speed at the Hill radius. Conversely, the second line corresponds to the threshold, assuming gravitational focusing is dominant (i.e. we approximate the particle trajectories as parabolic in estimating the work done as the particles cross the Hill sphere). This is appropriate if the relative velocity at infinity is much less than the escape speed at the Hill radius. In the supersonic regime, we use the unfocused estimate if $v_x > v_{\rm esc}(1+q)^{1/4} q^{5/4}$, where the two estimates are the same. Otherwise, we use the focused estimate. We refer to these two cases as the ‘unfocused regime’ and the ‘focused regime’.

In the subsonic regime capture occurs as long as the velocity at infinity is less than

$$v_{c,2} = \begin{cases} v_s q, & v_s \gg v_{\rm esc} \\ \sqrt{v_s v_{\rm esc} \frac{8q}{(1+q)^2}}, & v_s \ll v_{\rm esc} \end{cases}$$

$$v_s = \frac{8\pi G^2 \rho_g m_{\rm bin} R_{\rm Hill}}{3 c_s^3}. \tag{7}$$

Once again we assume the transition between the focused and unfocused regimes occurs where the two estimates for the threshold become equal (when $v_s = \frac{8 v_{\rm esc}}{q(1+q)^2}$).

The thresholds in equations (6) and (7) are approximate, since (in the unfocused regime) we use the initial velocity to estimate the energy dissipated. This is justified, because most energy will be dissipated at large velocities, due to the quadratic dependence of kinetic energy on velocity. For given gas properties and masses, one does not know a priori whether the threshold velocity for capture will be subsonic or supersonic (and whether to use the estimate in equation 6 or 7). Generally, only one of equation (6) and equation (7) will give a consistent result. If $v_{c,1} \lesssim v_{c,2}$ then $\frac{v_c}{c_s} \gtrsim 1$. Both estimates of the threshold are supersonic, and thus the supersonic estimate ($v_{c,1}$) should be used. Conversely, if $v_{c,1} \gtrsim v_{c,2}$ then $\frac{v_c}{c_s} \lesssim 1$. Both estimates are subsonic, and thus the subsonic estimate ($v_{c,2}$). In general, the threshold for capture can be estimated using

$$v_c = \min\{v_{c,1}, v_{c,2}\}. \tag{8}$$

We validate equation (8) with numerical simulations in Section 3.2.

Table 1 summarizes the threshold capture velocities in different regimes. Fig. 1 shows delineation between regimes as a function of the gas density and sound speed.

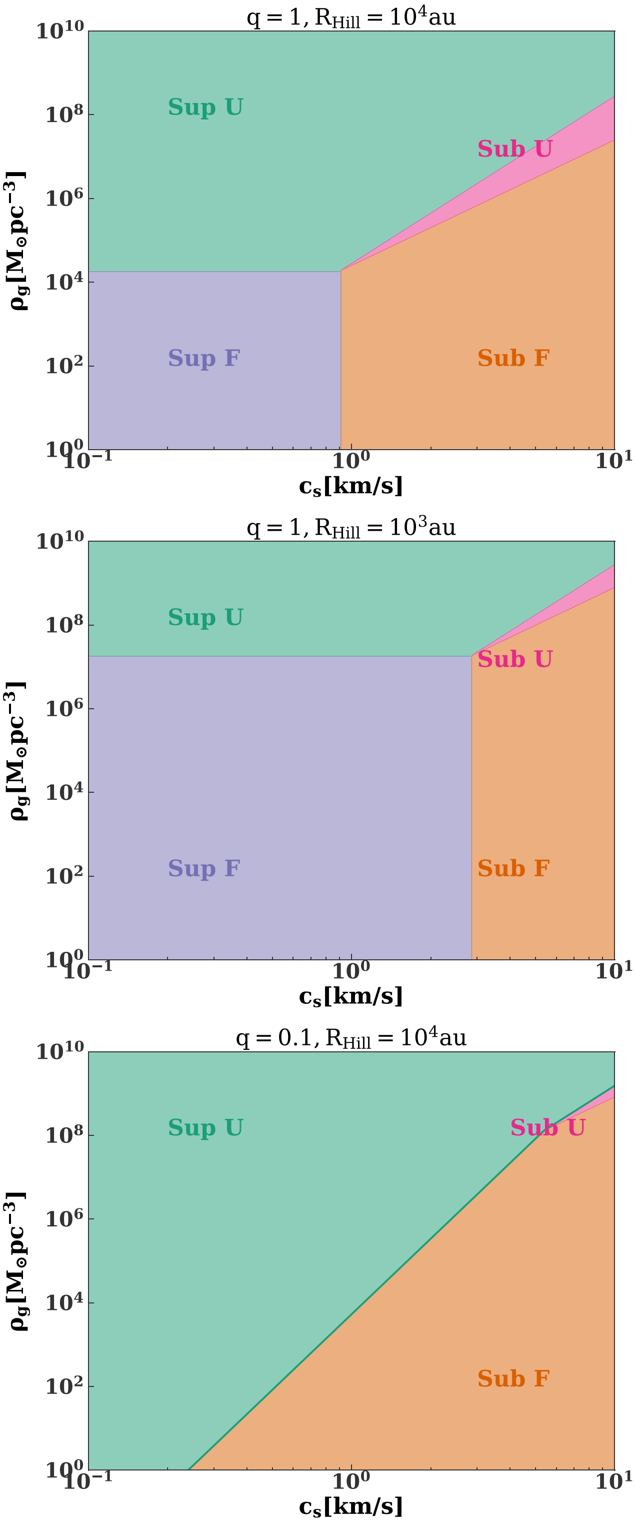

Figure 1. Separation between different capture regimes (see the text for details) as a function of gas density and sound speed for two 10-$M_\odot$ stars or compact objects. ‘Sup’ and ‘Sub’ correspond to supersonic and subsonic, respectively, while ‘F’ and ‘U’ correspond to focused and unfocused.

3.2 Numerical validation

In order to validate our analytic results, we make use of few-body numerical simulations with an added GDF force. We place two particles on an initially unbound orbit, and numerically integrate them forward in time under the influence of GDF. (See equation 1 and the discussion there.) We assume the initial separation corresponds to the Hill radius, although we did not explicitly include a tidal field in our simulations. For simplicity, the components of the separation parallel and perpendicular to the direction of motion are the same.

In other words, the initial relative velocity is misaligned by 45° with respect to initial separation.[2]

We evolve the two stars with the IAS15 integrator (Rein & Spiegel 2015) in `REBOUND` (Rein & Liu 2012). GDF is included via `REBOUNDx` (Tamayo et al. 2020). The gas medium has a constant density (in space and time) and is at rest with respect to the binary centre-of-mass initially.

Fig. 2 shows the maximum Mach number for which the two 10-$M_\odot$ stars are captured into a bound binary while crossing the Hill sphere, as a function of the sound speed for a handful of gas densities and Hill radii.[3] Fig. 3 shows the dependence of this capture threshold on the mass ratio of the two stars. For comparison, we also show the maximum Mach number from equation (8). This falls within a factor of ∼2 our numerical results.

Fig. 4 shows the stellar velocity, binary semimajor axis, and binary eccentricity as a function of time for a handful of gas parameters. We find reasonable agreement between our numerical and analytic solutions for the velocity (see Appendix A) at early times. At late times, gravitational acceleration (not included in Appendix A) causes the solutions to diverge. As expected, the binary orbital elements evolve over a dynamical time. Fig. 5 shows the trajectories of the stars during the capture in the second row of Fig. 4.

3.3 Stability of captured binaries

If a captured binary forms with a semimajor axis that is greater than the Hill radius, it will be short-lived. Requiring stability can significantly reduce the threshold capture velocity, as shown in Fig. 6.

Captured binaries will necessarily be unstable if the threshold capture velocity is in the gravitationally focused regime.

At the threshold velocity, the energy dissipated is precisely the kinetic energy at infinity. If the threshold velocity is in the focused regime this will be small compared to the potential energy at the Hill sphere. Furthermore, if the encounter is focused, the relative velocity of the object in the relevant environment (e.g. star/compact object or KBO; in cluster/AGN or Solar system environment, respectively) inside the Hill sphere (and hence the energy dissipated) is a weak function of the velocity at infinity. Thus, the energy dissipated will always be small compared to the potential energy at the Hill sphere in this regime and no stable binaries can form.

3.4 Effects of a headwind

So far we have neglected the binary's centre-of-mass motion through the gas. However, there might be a non-negligible headwind that could affect the capture and subsequent evolution. We perform additional few-body simulations with a headwind, and discuss the results here.

Fig. 7 shows the effect of this motion on the threshold capture velocity. In the unfocused, supersonic regime it reduces the threshold. For large centre-of-mass velocities, the reduction is significant. (It is up to a factor of ∼4 if the centre-of-mass velocity is twice the relative velocity, though this precise reduction depends on the orientation of the headwind). However, for two equal-mass objects

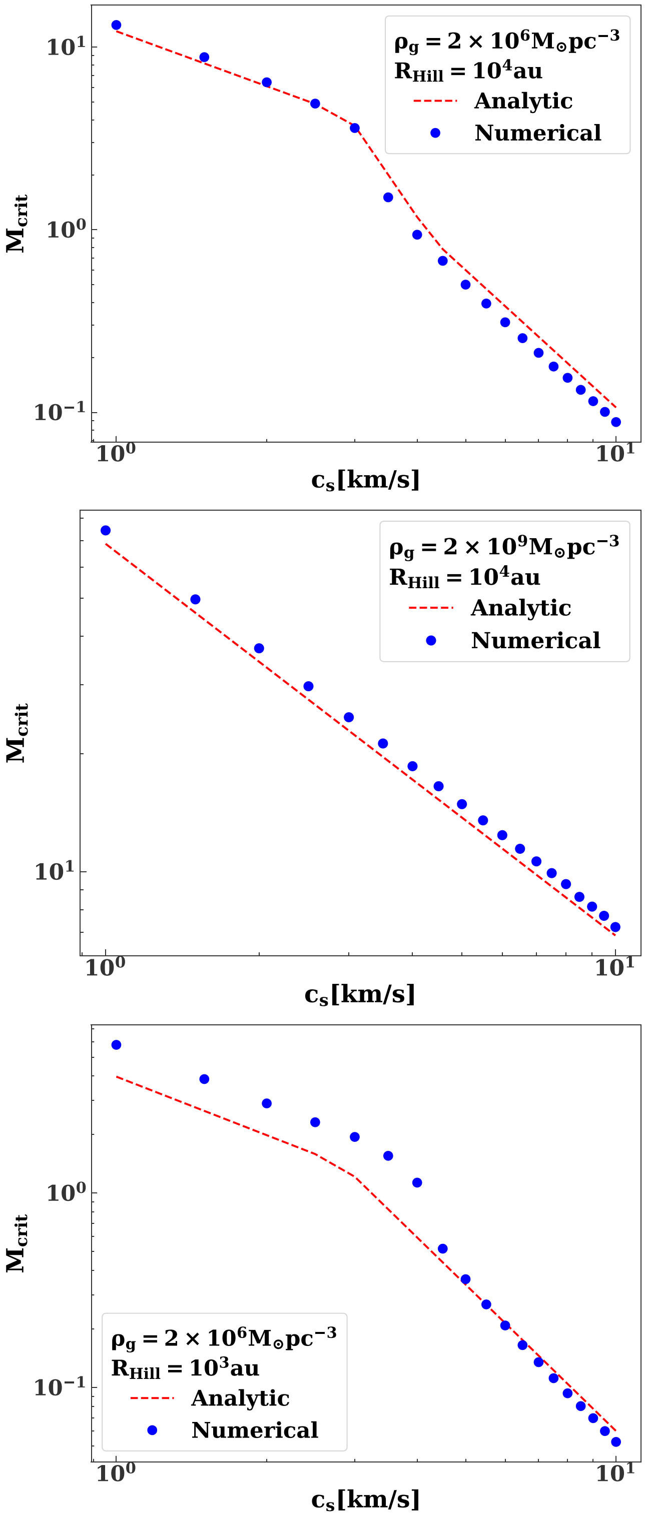

Figure 2. Maximum Mach number (at infinity) for which capture can occur as a function of sound speed for different gas densities and Hill radii. The stellar masses are 10 $M_\odot$. The blue points are from two-body simulations with GDF. The red, dashed lines show the analytic estimate for the maximum capture velocity (see equation 8 and Table 1).

---

[2] The threshold velocity has a weak dependence on the impact parameter, b, viz. $v_c \propto (1 - \frac{b^2}{R^2_{\rm Hill}})^\xi$, where ξ is 1/8 (1/2) in the supersonic (subsonic) regime. We neglect this correction.

[3] This is the Mach number at infinity. For each Mach number, the initial velocity corresponds to the velocity at the Hill sphere, accounting for gravitational focusing alone (neglecting the gas).

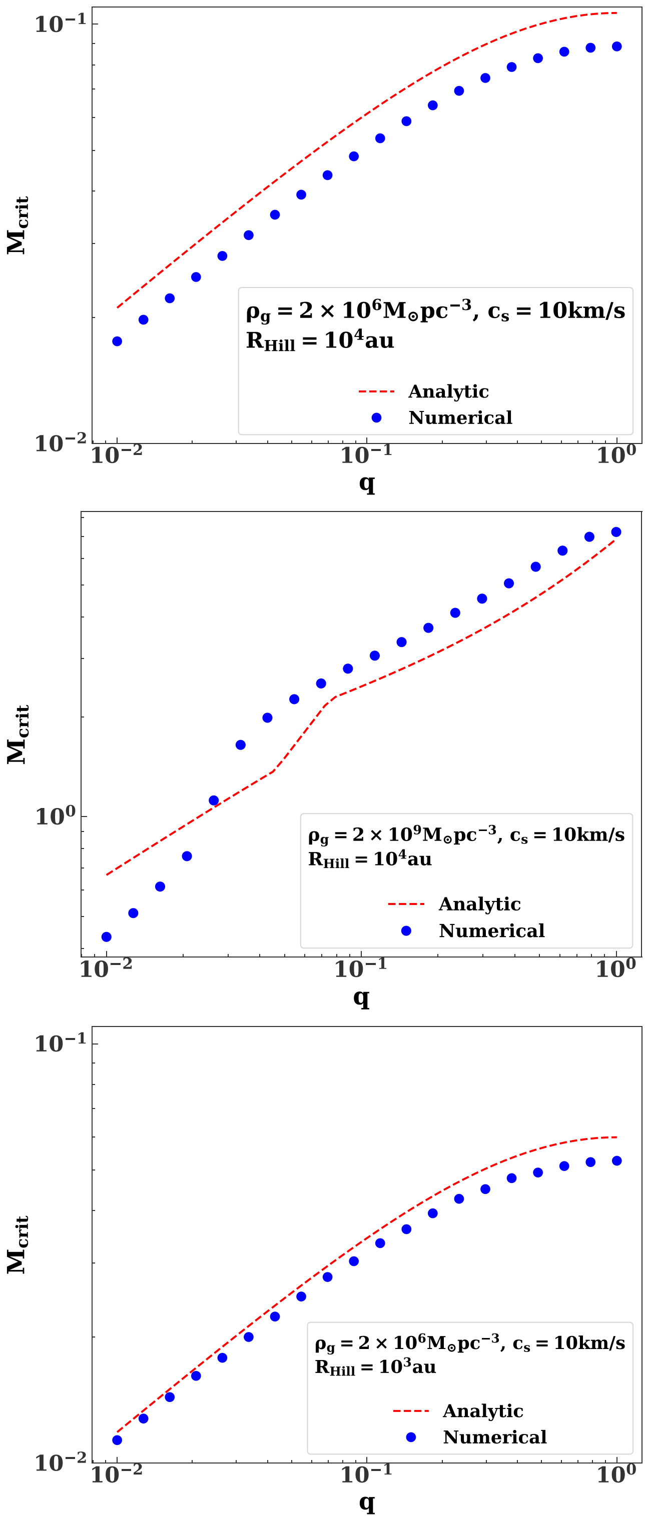

Figure 3. Maximum Mach number (at infinity) for which capture can occur as a function of mass ratio for different gas densities and Hill radii. The total mass of the binary is 20 $M_\odot$. The blue points are from two-body simulations with GDF. The red, dashed lines show the analytic estimate for the maximum capture velocity (see equation 8 and Table 1).

with isotropic Maxwellian velocities, the centre-of-mass velocity will be approximately half the relative velocity on average. In this case, the threshold is reduced by a factor of 0.68 on average. In the unfocused subsonic regime, the centre-of-mass motion increases the threshold velocity and aids capture.

Typically, the binaries' centre-of-mass motion will affect the threshold for capture by less than a factor of 2. Thus, we neglect this effect in our capture rate estimates.

4. CAPTURE RATES

For a given environment, the binary formation rate through gas-assisted capture could be written by

$$\Gamma(m_1, m_2) \approx \int_0^{v_{\rm crit}} n_\star(m_2|m_1)\mathcal{A}vp(v)\mathrm{d}v,$$
$$\mathcal{A} = R_{\rm Hill}z(1+\Theta^2)$$
$$z = \min\{R_{\rm Hill}, h_{\rm eff}\} \tag{9}$$

where m_1 is the mass of the capturer, m_2 is the captured mass, $n_\star(m_2|m_1)$ is the density of candidates for captured masses in the vicinity of m_1, $h_{\rm eff}$ is the effective scale height of the disc (if the environment has a disc-like configuration; e.g. an AGN disc, a gaseous disc in a cluster, or a protoplanetary disc), $p(v)$ the velocity distribution, and $\Theta = (v_{\rm esc}/v)^2$ is a correction for gravitational focusing. This correction is not valid in systems dominated by a massive central object like AGNs and protoplanetary discs, since it is derived assuming unperturbed two-body trajectories and neglects shearing motion. The critical velocity for capture, $v_{\rm crit}$, is calculated according to the regime (focused/unfocused), as specified in Table 1. We assume this is a Maxwellian distribution, such that $p(v) \propto v^2 e^{-v^2/2\sigma^2}$, where σ is the velocity dispersion. Thus, equation (9) simplifies to

$$\Gamma(m_1, m_2) = n_\star R_{\rm Hill} z\sigma\sqrt{\frac{2}{\pi}}\,[f_1 + f_2]$$
$$f_1 = \left(v_{\rm esc}^2/\sigma^2\right)\left(1 - e^{-v_{\rm crit}^2/(2\sigma^2)}\right)$$
$$f_2 = 2 - e^{-v_{\rm crit}^2/(2\sigma^2)}\left(2 + \left(\frac{v_{\rm crit}}{\sigma}\right)^2\right). \tag{10}$$

Note that this expression differs from the one derived for example in Tagawa et al. (2020) by the factor in the brackets (divided by $2\sqrt{2\pi}$). This term becomes significant if $v_{\rm crit} < \sigma$, where capture is dominated by the tail if the velocity distribution.

For small capture velocities, the capture for the unfocused case rate could be approximated by

$$\Gamma(m_1, m_2) \approx \sqrt{\frac{2}{\pi}}\frac{n_\star R_{\rm Hill}^2 v_{\rm crit}^4}{2\sigma^3}. \tag{11}$$

Below we derive the gas-assisted capture rate for different gaseous environments and set constraints on the available parameter space that enables such a capture. Throughout this paper, we use equation (9) to calculate the capture rate, dropping the focusing correction in shear-dominated environments (AGN and protoplanetary discs).

5. CAPTURE RATES IN DIFFERENT GAS-RICH ENVIRONMENTS

In this section, we study the conditions for binary formation in different environments, and summarize our results in Table 2. We consider following: (1) SF environments where binaries are formed from newly born stars or even pre-main-sequence stars or protostars, where gas-assisted capture may serve as an important

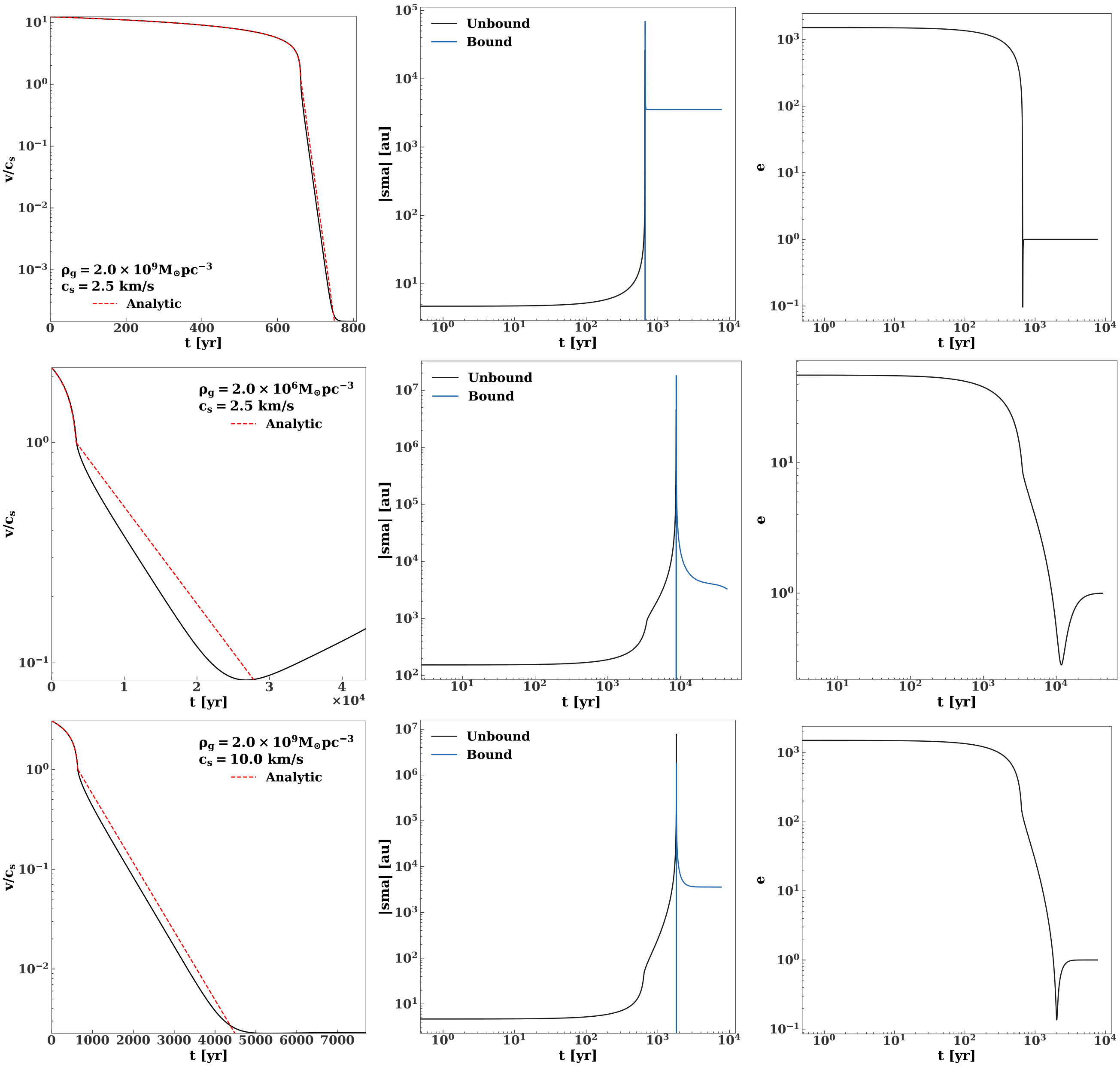

Figure 4. Evolution of the stellar velocities and binary orbital elements from numerical simulations with different gas density and sound speeds. In first column, we also plot the analytic solution of the velocity (from Section A) as a dashed, red line. In all cases, the stars are both 10 M$_\odot$.

channel for the fundamental formation of stellar binaries; (2) AGN discs around supermassive BHs, where gas-assisted captures could form stellar and compact object binaries, and may contribute to the formation of stellar binaries which could later inspiral and eventually give rise to merger products and explosive transients and GW sources from compact object mergers; (3) gas-enriched massive clusters, where the existence of multiple generation of stars suggest several epochs of gas-rich environments in which earlier generations of stars and compact objects could be embedded, and form stellar and compact object binaries, similar to the case of AGN discs; and (4) gaseous protoplanetary discs, where embedded planetesimals can form binaries through gas-assisted capture, and in particular, the early stages of planet formation in the Solar systems could give rise to the production of KBOs and asteroid binaries.

5.1 SF environments

Star formation takes place in cold gas-rich clumps embedded in molecular clouds. These clumps could constitute as a fertile ground for gas-assisted binary formation. The typical gas temperature in these regions is ∼10 K (Shu, Adams & Lizano 1987; Williams, Blitz & McKee 2000), which corresponds to a sound speed of 0.2 km s^{-1}. The typical mass of clumps is 10^3–10^4 M$_\odot$ and their radii are 2−5 pc (Shu et al. 1987).

The typical gas density in the clumps should exceed a threshold value to enable star formation, which is typically $n_{\rm th} \sim 10^4$ cm^{-3} (Bergin & Tafalla 2007 and references therein). Assuming the mass in stars is comparable to the mass in gas, the gas and stellar densities could be approximated by $\rho_{\rm stars} = \rho_{\rm gas} = n_{\rm g} m_H \approx 225$ M$_\odot$ pc^{-3}.

Unless stated otherwise, the radius of the clump is 2 pc, the clump mass is 10^3 M$_\odot$, and the gas density is $n_{\rm g} = 10^4$ cm^{-3}. Other

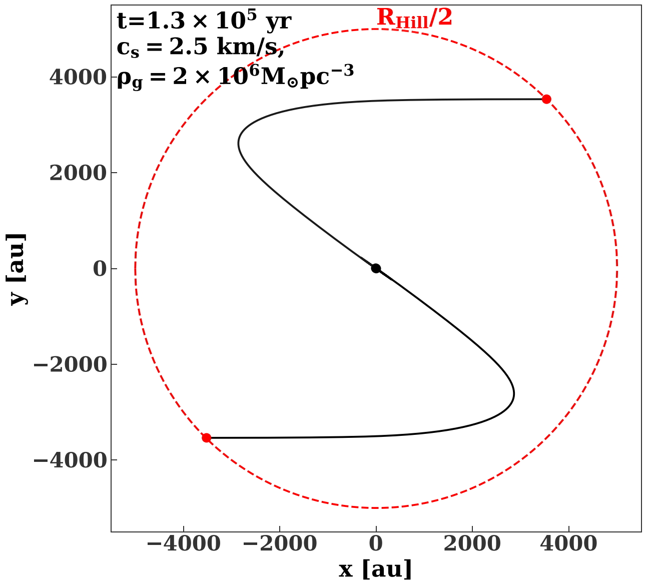

Figure 5. Particle paths for the gas-assisted capture in the second row of Fig. 4. The red dots show the initial position of the two 10-M$_\odot$ objects. The gas is assumed to be at rest with respect to the binary centre-of-mass initially. A movie of the capture is available at: https://www.youtube.com/watch?v = gh5KC_fjPp4.

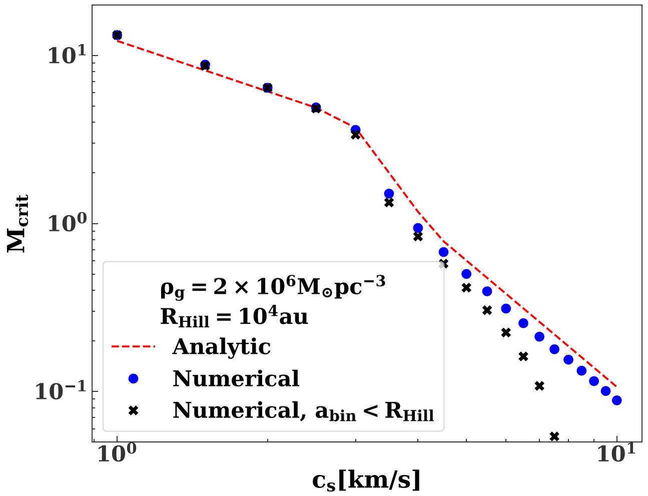

Figure 6. Maximum Mach number (at infinity) for which capture can occur as a function of sound speed for the density and Hill radius in the top panel of Fig. 2. The black crosses show the maximum Mach number for which capture into a stable binary (with semimajor axis less than the Hill radius) can occur. Capture into stable binaries is impossible for sound speeds $\gtrsim 8$ km s^{-1}.

parameters are derived from these choices. For the stars, we assume Kroupa mass function (Kroupa 2001). Using these parameters, we calculate the rate of gas-assisted binary captures using equation (9).

In Fig. 8, we present the capture rate per object in an SF region. The gas-assisted binary formation is robust and every object in this environment is likely to capture at least another object during the gas lifetime, and even more. As expected, the capture rate increases with the gas density. It should be noted that the capture rate presented corresponds to m_1 capturing m_2 – i.e. when calculating the rate (equation 9), the background number density changes with the captured species according to the background mass function. To calculate the total number of binaries with masses m_1 and m_2, one should sum up the contributions from m_1 capturing m_2 and vice versa.

Further evolution of these formed binaries is left out for future studies (in preparation), and could leave unique signatures on binaries distributions. It should be noted that past studies already discussed orbital decay of binaries due to gaseous background in similar context (e.g. Stahler 2010; Korntreff, Kaczmarek & Pfalzner 2012).

5.2 AGN discs

The evolution of binaries in AGN discs was studied extensively (e.g. McKernan et al. 2012; Stone et al. 2017; Tagawa et al. 2020, and references therein), especially as progenitors for GWs. Gas-assisted inspirals were modelled in different ways, either through a planetary migration modelling (McKernan et al. 2012; Stone et al. 2017), GDF (e.g. Bartos et al. 2017), or Bondi–Hoyle accretion (Antoni, MacLeod & Ramirez-Ruiz 2019), and was also explored explicitly through hydrodynamical simulations (e.g. Antoni et al. 2019; Li et al. 2023, and references therein), though the actual migration modelling is still debated. It was suggested by Tagawa et al. (2020), that the vast majority of merging binaries in AGN discs originate in gas-assisted binary formation. Hence, the conditions for binary formation effectively dictate the expected merger rates in such environments. There is a wide range of possible AGN configurations and masses of the central MBHs. We will consider a more specific case, but the same approach could be generalized to other AGN conditions.

We consider an MBH mass of 4 × 10^6 M$_\odot$ like Sgr A*. Unless otherwise specified, we adopt the Thompson, Quataert & Murray (2005) AGN disc model, with the modifications described in Tagawa et al. (2020). Fig. 9 shows radial profiles of gas density, temperature, and scale height for different mass accretion rates on to the MBH. Outside of the central ∼1 pc, the disc is Toomre unstable and forms stars. Heating from radiation pressure and SNe maintains the disc in a marginally stable state. The density profile is set by marginal Toomre stability and is independent of the mass accretion rate. The gas density on large scales is $\rho_g \approx 10^6 (r/1\ \mathrm{pc})^{-3}\ \mathrm{M}_\odot\ \mathrm{pc}^{-3}$. For the fiducial accretion rate ($0.1\dot{M}_{\mathrm{Edd}}$) in Tagawa et al. (2020), the gas temperature is 20 K at ∼1 pc. (Corresponding to a sound speed of ∼0.4 km s^{-1}. The disc is very thin with aspect ratio, $h/r \approx 10^{-3} - 10^{-2}$). We also consider lower accretion rates down to $10^{-4}\dot{M}_{\mathrm{Edd}}$, where there is no self-consistent solution extending to pc scales. In such cases, we assume an α-disc, whose outer radius is set by Toomre instability.

We now estimate the capture rates of different types of binaries within the disc: BH–BH, BH–star, and star–star. We assume a (number) density profile for the BHs of

$$n_{\mathrm{bh}}(r) = n_{\mathrm{o}} \left(\frac{r}{r_{\mathrm{o}}}\right)^{-2} \left(\frac{h}{r}\right)^{-1}$$

$$n_{\mathrm{o}} = \frac{N_{\mathrm{bh}}}{4\pi r_{\mathrm{o}}^3}, \qquad (12)$$

where N_{bh} is 1000 and r_{o} is the outer radius of the BH distribution (3pc). This is similar to the initial number of disc BHs in Tagawa et al. (2020), where the BH component is flattened and rotating. This was suggested to occur via vector resonant relaxation (Szölgyén & Kocsis 2018). However, the degree of flattening and hence the number of disc BHs will depend on the mass function, and for realistic conditions, it is not clear whether indeed such a flattened disc of BHs should indeed exist, nor why should it be aligned with the gaseous AGN disc. Nevertheless, in order to compare with the results of Tagawa et al. (2020), we consider similar conditions. Alternatively, multiple

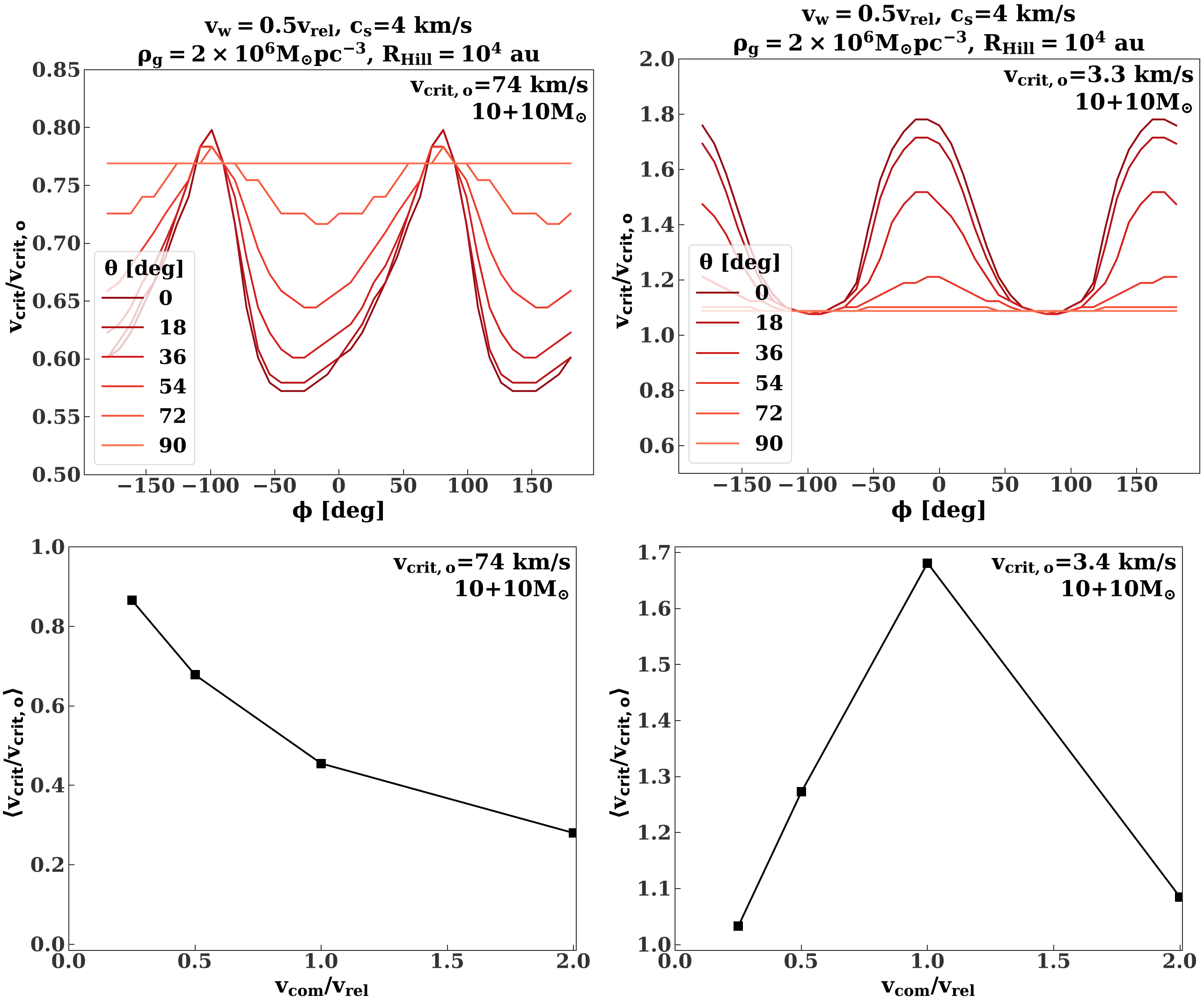

Figure 7. Two examples of the effect of binary motion on the threshold capture velocity. The left-hand (right-hand) panels correspond to the supersonic (subsonic), unfocused regime. Top panels show the change in the threshold capture velocity as a function of the direction of motion. Here, the bodies' centre-of-mass velocity (with polar angles θ and ϕ) is half the relative velocity. The bottom panels show the angle-averaged ratio between the thresholds with and without the centre-of-mass motion as a function of the centre-of-mass velocity (normalized to the relative velocity).

Table 2. Typical capture rates per object, for equal-mass binaries, in four different environments: SF regions, later generation formation in GCs, AGN discs, and protoplanetary discs. The columns refer correspondingly to the gas number density $n_{\rm g}$, binary mass $m_{\rm bin}$ for which the capture rates are presented, Hill radius $R_{\rm Hill}$, sound speed $c_{\rm s}$, background density of the captured objects $n_{\rm b}$, and finally the expected capture rate Γ, as calculated according to our model, taking into consideration the relevant regime of focusing. Note that the rate for AGN discs is averaged over a non-flat stellar density profile.

	$n_{\rm g}$ (cm^{-3})	$m_{\rm bin}$ (M$_\odot$)	$R_{\rm Hill}$ (au)	$c_{\rm s}$ (km s^{-1})	$n_{\rm b}$ (pc^{-3})	Γ (Myr^{-1})
SF	10^4	2	5×10^4	0.2	308	12.6
later SF – GCs	4×10^6	20	1.2×10^4	0.6	10^3	9.2
AGN*	4×10^7	20	2.4×10^3	0.4	9.5×10^4	1
PPD	4×10^{11}	4×10^{-12}	6×10^{-3}	0.15	3×10^{15}	$\ll$ (gas lifetime)$^{-1}$

star-formation epochs might give rise to new generation of BHs that form in the AGN disc (Stone et al. 2017), and provide a large number of BHs in the disc. The velocity distribution of the BHs is taken as a Maxwellian whose scale parameter is $(h/r)v_{\rm kep}$, where $v_{\rm kep}$ is Keplerian velocity. For simplicity, we assume all BHs are 10 M$_\odot$.

Fig. 10 shows the BH–BH capture rate from equation (9) integrated over the entire disc as a function of mass accretion rate (at the outer boundary), viz.

$$\Gamma_{\rm tot} = \int_{r_{\rm min}}^{3{\rm pc}} n_{\rm bh}(r)\Gamma(r)4\pi r h(r){\rm d}r. \tag{13}$$

The total rate is dominated by large scales and is a weak function of $r_{\rm min}$. The dashed line shows the total capture rate, while the solid line shows the rate of captures in the unfocused regime. Considering the

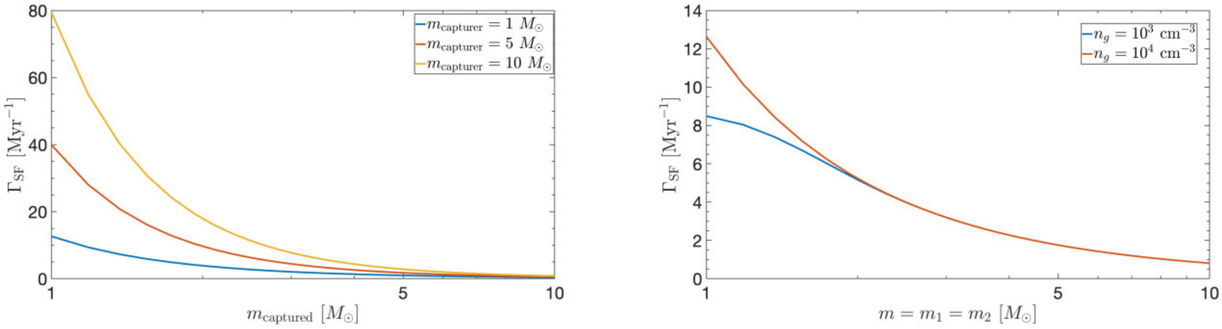

Figure 8. Left-hand panel: The capture rate per object for different masses, in an SF environment, for our fiducial model specified above. Right-hand panel: The capture rates for equal masses, given different gas number densities.

discussion in Section 3.3, only the latter can lead to long-lived, stable binaries. Thus, we expect a total binary formation rate of a few × 10^{-4} yr^{-1} in discs with Eddington ratios $\gtrsim 0.01$. This is comparable to the formation rate calculated by Tagawa et al. (2020) (cf their Fig. 7). The average capture rate per BH is 2×10^{-7} yr^{-1}.

However, this result is sensitive to the aspect ratio of the disc, as shown in Fig. 11. This figure shows the BH–BH capture rate after artificially rescaling the aspect ratio of the $\dot{M}/\dot{M}_{\rm edd} = 0.1$ disc in Fig. 9. For very thin discs, the capture rate increases linearly with the velocity dispersion and with the aspect ratio. However, for thicker discs the capture rate falls off steeply with aspect ratio, because captures only come from the tail of the velocity distribution. The capture rate of stable binaries is 0 for aspect ratios above a few × 10^{-2}.

The aspect ratio depends on the mechanism for angular momentum transport. In the above calculations, the radial gas velocity is 0.15 times the local sound speed in the outer disc, as in Tagawa et al. (2020). In Sirko & Goodman (2003), the radial velocity is $\approx\alpha(h/r)$ times the sound speed and can be much smaller. For $\alpha = 0.1$ and $\dot{M}/\dot{M}_{\rm Edd} = 0.1$, the aspect ratio at parsec scales is approximately an order of magnitude larger than in the Tagawa et al. (2020) model. Thus, the total capture rate is a factor of ∼2 smaller, and the rate of stable captures is 0.

So far we have focused on BH–BH captures. However, BH–star captures and star–star captures will also occur. We expect a few × 10^6 stars old, low-mass ($\lesssim$M$_\odot$) stars within the central ∼3 pc of the Galaxy. Geometrically, we expect $\sim 10^4$ stars within the disc. We estimate the rate of BH–star captures to be $\sim 2 \times 10^{-3}$ yr^{-1} and the rate of star–star captures to be $\sim 10^{-2}$ yr^{-1}.[4] This assumes the low-mass stars in the disc have an r^{-2} density profile like the BHs. However, in the studies of relaxation in spherical clusters, the density profile of low-mass species falls between $r^{-1.5}$ and $r^{-1.75}$ (Alexander & Hopman 2009). For an $r^{-1.5}$ stellar density profile, the BH–star and star–star capture rates are $\sim 1.4 \times 10^{-3}$ and 6×10^{-3} yr^{-1}, respectively. Thus, the overall capture rate per object is ∼1 Myr^{-1}. Note that the capture rate is dominated by stellar captures. For example, the BH–star capture rate is roughly one order of magnitude greater than the BH–BH capture rate.

5.3 Gas-enriched GCs/massive clusters

For decades, GCs were thought to contain a single-aged stellar population, i.e. originating from a single burst of star formation. However, over the last two decades it was found that the vast majority of GCs host at least two or even more population of stars (see detailed reviews in Renzini et al. 2015; Bastian & Lardo 2018; Gratton et al. 2019), which were suggested to form at different epochs. Although the exact origin of the multiple populations is still unknown, their existence is a smoking gun for gas-replenishment in GCs. As we pointed out in Rozner & Perets (2022), the dynamics and evolution of binaries in GCs should be revised, and the gas involved in the formation of the second or further generation should affect the evolution of previously formed stellar populations and binaries which become embedded in such gas-rich environment (Maccarone & Zurek 2012; Leigh et al. 2013, 2014; Roupas & Kazanas 2019; Rozner & Perets 2022). Similar to AGN discs, we therefore might expect that high gas abundance could also potentially give rise to gas-assisted binary formation in these environments.

The density of gas originating in the epoch of second (or further) star formation is highly uncertain (e.g. Bastian & Lardo 2018) and can be roughly estimated by $\rho_{\rm g} \sim M_{\rm 2P}/V_{\rm 2P} \sim 10^5$ M$_\odot$ pc^{-3}, where $M_{\rm 2P}$ is the mass of the second population and $V_{\rm 2P}$ is the volume in which it is enclosed, if we again consider specifically BHs, the typical BH number density is $n_\bullet \sim 10^3$ pc^{-3}. The temperature during the SF stage may differ from the current temperature. Hence, following Bekki (2010), we consider a gas temperature of 100 K, which corresponds to a sound speed of 0.6 km s^{-1}. Following Bekki (2010) and Mastrobuono-Battisti & Perets (2013, 2016), we consider the second population of stars as embedded in a disc, with an aspect ratio of $h/r \sim c_{\rm s}/v_{\rm K} \sim 3 \times 10^{-2}$, and a radius of 1 pc. Unless stated otherwise, these will be the fiducial parameters. We assume a stellar density of $n_\star = 10^4$ cm^{-3} for Solar star mass. The mass density of each stellar species is constant i.e. $\rho \equiv$ constant $= \rm m_i n_i$, where ρ is a constant.

The total mass of the cluster is 2×10^5 M$_\odot$.

In Fig. 12, we present the capture rates per object in multiple-population gas-enriched environment in GCs (or equivalent younger massive clusters objected in other galaxies), or nuclear clusters which do not host MBHs. As can be seen on the left-hand panel, the capture rate decreases for larger masses of captured objects, although the total capture rate is high for all the mass range. The right-hand panel shows the capture rate for different gas and stellar masses. As can be seen, the dependence on the gas mass is not trivial, as it also modifies the Hill radius, which depends on the enclosed stellar and gaseous mass. Sufficiently high densities lead to smaller probabilities for capture. The overall dependence on the central mass is stronger than the dependence on the gas density.

As we showed in Rozner & Perets (2022), the merger of pre-existing binaries in such environments could be catalyzed by the gaseous environment; the additional gas-catalyzed formation bina-

---

[4] For simplicity we assume all stars are 1 M$_\odot$.

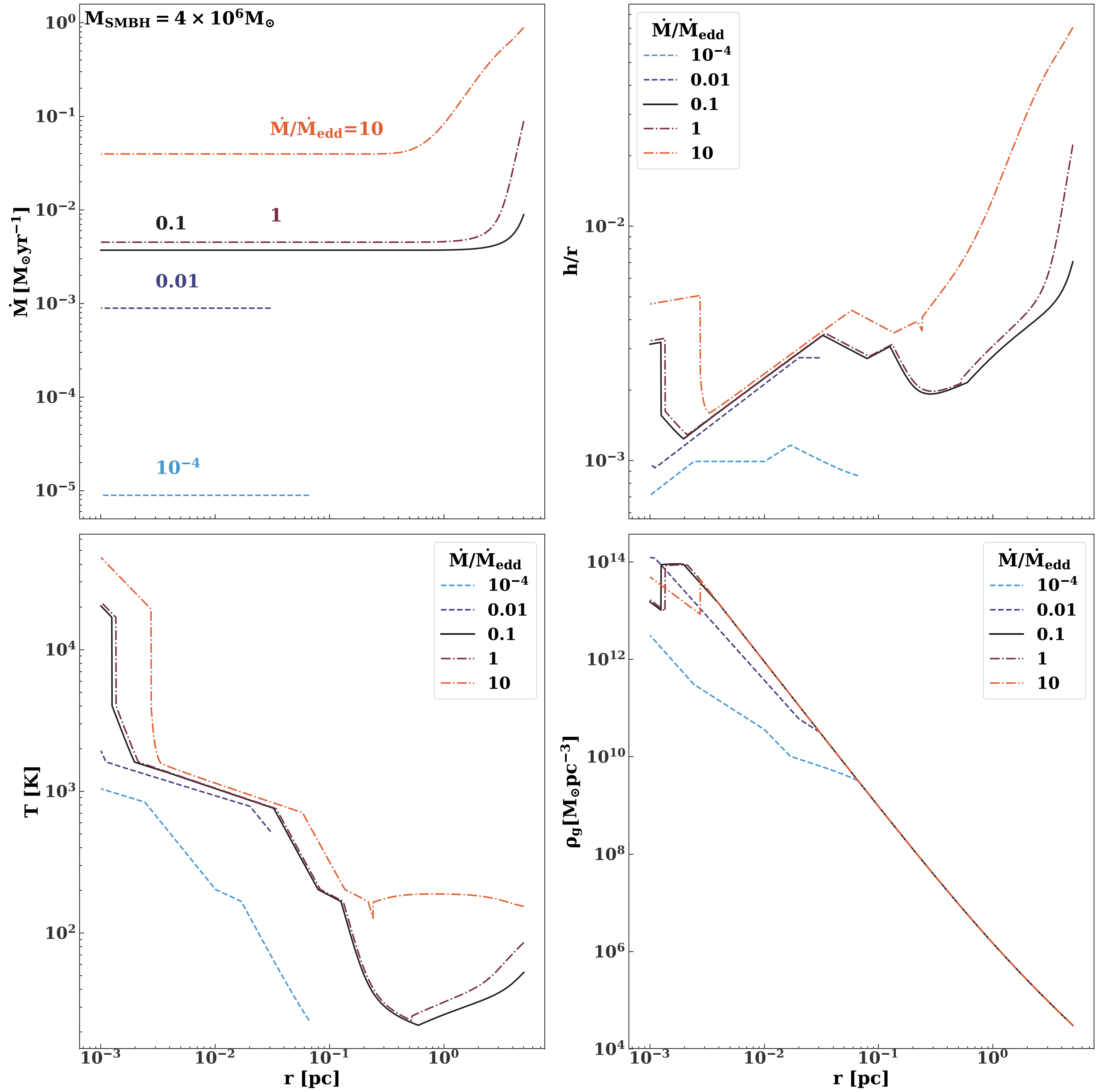

Figure 9. (Clockwise from top left) Radial profiles of mass accretion rate, aspect ratio, gas temperature, and gas density. Different colours correspond to different mass accretion rates at the outer boundary. The equation describing disc structure is in Thompson et al. (2005).

ries would therefore also further increase the binary mergers rate. We conclude that a significant fraction of the merged objects in this environments are the product of gas-assisted mergers, similarly to the conclusion in Tagawa et al. (2020) for AGN discs. That being said, this conclusion, like our results on AGN discs, strongly depend on the existence of a relatively flattened disc; thicker discs would not allow for significant capture rates.

5.4 Protoplanetary discs

The early stages of planet formation take place in protoplanetary discs that initially contain gas, with typical dust-to-gas ratio of ~1 per cent (Chiang & Goldreich 1997). Planetesimals vary in size, and the nature of their interaction with gas changes accordingly (Weidenschilling 1977). While small particles are well-coupled to the gas, large planetesimals interact differently and their interaction with the gas could be modelled using GDF (Grishin & Perets 2015, 2016). Goldreich et al. (2002) suggested that two initially unbound objects in the Kuiper-belt could form a wide binary with comparable masses, via dissipation induced by dynamical friction. Due to the high abundance of gas on the early stages of planet formation, GDF could potentially play a similar role and lead to gas-assisted binary formation in protoplanetary discs.

Following Armitage (2010) and Perets & Murray-Clay (2011), we consider a background gas density of $\rho_g = 3 \times 10^{-9}(a/\mathrm{au})^{-16/7}$ g cm^{-3} and sound speed of sound speed of $c_s =$

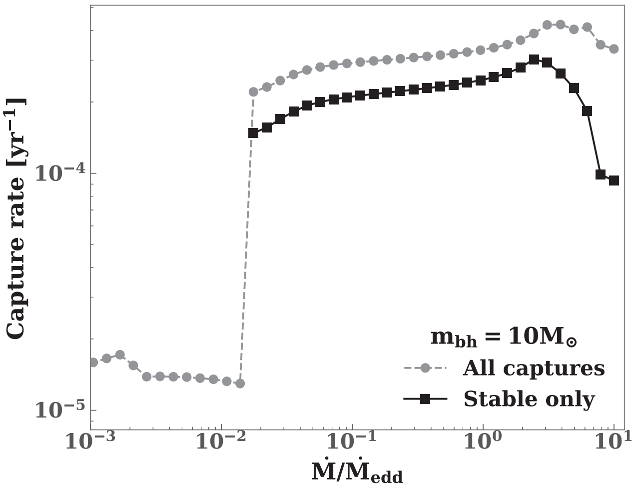

Figure 10. Total BH–BH capture rate due to GDF in model accretion discs (see text and Fig. 9 for details). Both BHs are 10 $M_\odot$. The dashed, grey line shows the total capture rate, while the solid, black line shows the rate of captures in the unfocused regime. Only the latter will form stable, long-lived binaries (see the discussion in Section 3.3).

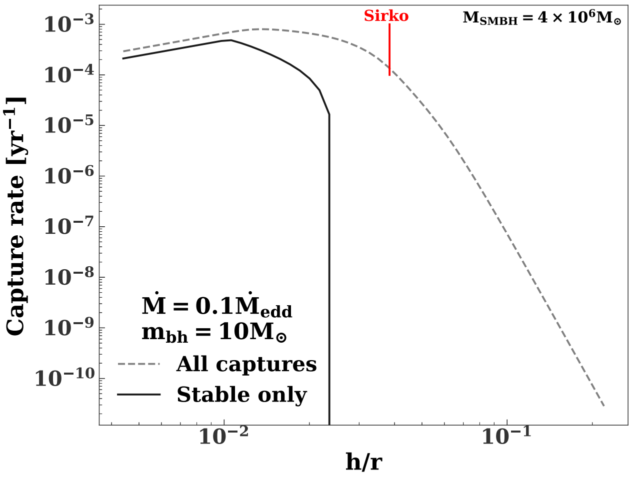

Figure 11. Total BH–BH capture rate for artificially inflated AGN discs with $\dot{M}/M_{\mathrm{Edd}} = 0.1$ (i.e. we use the Tagawa et al. 2020 model in Fig. 9, but aspect ratio is rescaled by a constant factor). The x-axis shows the aspect ratio at 3 pc after rescaling. If the disc aspect ratio is increased by an order of magnitude, the rate of stable captures goes to 0 (see the discussion in Section 3.3). The vertical red line shows the aspect ratio for the Sirko & Goodman (2003) model for fiducial parameters.

$0.6(a/au)^{-3/8}$ km s^{-1}. Following Goldreich et al. (2002), we choose a separation of $a = 40$ au from a Solar mass star and for a 100-km objects, the typical surface density is $\Sigma \sim 3 \times 10^{-4}$ g cm^{-2}. Then, the typical corresponding density is $\rho \sim \Sigma/r$ where r is the distance from the Sun, such that the number density of candidates for capture in this mass is $n_{\mathrm{b}} = \rho/m$. Given these assumptions, a capture of $\sim 10^{22}$-g object is not likely to occur via GDF within the gas lifetime. Since the typical velocity dispersion of large objects relative to the gas is supersonic, the energy dissipation induced by gas-dynamical friction is not efficient enough to enable such a capture, such that dynamical friction induced by smaller bodies will be more efficient under these assumptions.

6. POTENTIAL CAVEATS

Here we will briefly discuss potential caveats of our model:

(1) Interaction of objects in gaseous environments could be affected not only by GDF, but by migration in circumbinary discs (e.g. McKernan et al. 2012; Stone et al. 2017), or Bondi–Hoyle accretion (Antoni et al. 2019). Later stages of the evolution are more likely to be dominated by Bondi–Hoyle accretion rather than GDF. A more detailed description of the gaseous interaction is beyond of the scope of this paper and is left out for future studies.

(2) We showed that headwind will have a minor effect on capture (see Section 3.4). However, we considered only a linear constant headwind, and neglected shear.

(3) We use Ostriker (1999)'s prescription for GDF, which assumes straight-line trajectories. However, objects will have curved orbits following capture. This prescription also neglects interference between each object's wake. Nevertheless, while these could be critical for the evolution and inspiral of bound binaries, the conditions for the initial capture are generally consistent with our assumptions. Future hydrodynamical simulations may help resolve the potential importance of this issue.

(4) Gas accretion is not taken into account here, and could significantly change the mass distribution of objects in gaseous environments, as well as the heating rates. However, the dynamical time-scale for the capture is relatively short, and we do not expect mass-gain to be of significant importance for the capture. Nevertheless, accretion feedback could potentially change the GDF effect (see Gruzinov, Levin & Matzner 2020 and the references therein).

(5) Close to the Hill radius, the interaction of the two stars with the external potential could give rise to temporary captures, which potentially allow for longer close interaction and more significant energy dissipation, and hence higher capture rates (see Petit & Henon 1986; Boekholt et al. 2023 and references therein). It should be noted again that in this paper we focus on the capture process only and leave out further evolution, including stability for future studies (in preparation). Such evolution will be affected also by the external potential as well as other dissipation mechanisms such as GW radiation (e.g. Li, Lai & Rodet 2022; Boekholt et al. 2023).

(6) Feedback effects from the energy deposited into the gas due to the capture and subsequent migration could potentially change the conditions of the gaseous environment. We consider this issue in the following section.

7. FEEDBACK: HEATING AND COOLING

Up to this point, we have considered the effect of the gas on the capture formation of binaries. However, the capture, and more importantly the later inspiral of the binaries in gas could potentially give rise to feedback and heat the gaseous environments, potentially quenching further gas-assisted formation of binaries. In the following, we show that the heat generated due to the gas dissipation is mostly radiatively emitted and does not contribute significantly to the disc heating. Other processes such as gas accretion on to stars/compact objects might provide additional feedback (e.g. through jets); modelling the effect of these process depends on many different assumptions and uncertainties, and is beyond the scope of this study.

Capture and inspiral of binaries in the gas could heat the environment. The heating energy E_{heat} from the inspiral could be

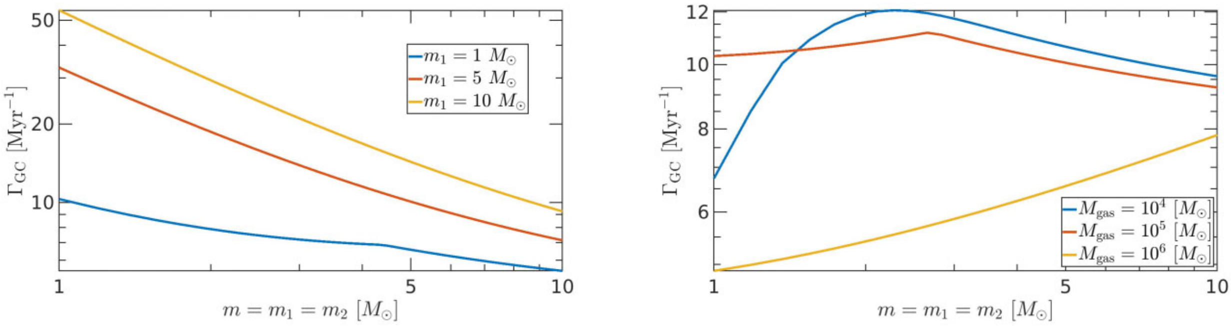

Figure 12. Left-hand panel: The capture rate per object for different masses, in a second (or later) star generation environment in a GC, for our fiducial model specified above. Right-hand panel: The capture rates for equal masses, given different gas masses.

approximated by

$$E_{\rm heat} \approx \frac{Gm_1m_2}{2a_{\rm final}}, \tag{14}$$

where $a_{\rm final}$ is the final semimajor axis of the binary prior to merger. Then, the average heating rate per area is

$$\ell_{\rm heat} = E_{\rm heat}\gamma, \tag{15}$$

where γ is the binary capture rate per area, i.e. $\gamma = \Gamma_{\rm cap}/A$ where $\Gamma_{\rm cap}$ is the capture rate as we calculated earlier, and A is a typical area. The typical cooling rate could be approximated by blackbody cooling:

$$\ell_{\rm cool} = \sigma_{\rm SB}T_{\rm eff}^4. \tag{16}$$

Here, we compare between the cooling and heating rates for the different environments we discussed, for our fiducial models specified in Table 2.

7.1 SF environments

The typical Hill radius is $R_{\rm Hill} = 2\times 10^4$ au, and the cloud lifetime is 5 Myr (Bergin & Tafalla 2007 and references therein). These parameters yield

$$\ell_{\rm cool} \approx 5\times 10^{36}\left(\frac{T_{\rm eff}}{10\ \rm K}\right)^4\ {\rm erg\ s^{-1}\ pc^{-2}}. \tag{17}$$

From integration of da/dt (as derived in Rozner & Perets 2022) for 5 Myr, and taking the relative velocity between the gas and the objects as half the Keplerian velocity, the separation after 5 Myr, $a_{\rm final}$ is $\approx$380 au, hence

$$\ell_{\rm heat} = 1.68\times 10^{34}\left(\frac{m}{1\ \rm M_\odot}\right)^2\left(\frac{380\ \rm au}{a_{\rm final}}\right) \times\left(\frac{\gamma}{1.4\ \rm Myr^{-1}\ pc^{-2}}\right)\ {\rm erg\ s^{-1}pc^{-2}}. \tag{18}$$

We then conclude that cooling is efficient in SF environments.

7.2 Second generation in GCs

For second-generation gas embedded in a disc, the energy radiated away within gas lifetime of 50 Myr is given by

$$\ell_{\rm cool} \approx 5.1\times 10^{40}\ {\rm erg\ s^{-1}\ pc^{-2}} \tag{19}$$

and the heating energy

$$\ell_{\rm heat} = 2.92\times 10^{36}\left(\frac{m}{10\ \rm M_\odot}\right)^2\left(\frac{0.07\ \rm au}{a_{\rm final}}\right) \times\left(\frac{\gamma}{\rm Myr^{-1}\ pc^{-2}}\right)\ {\rm erg\ s^{-1}\ pc^{-2}}. \tag{20}$$

Hence, cooling is efficient also here.

7.3 AGN discs

Over the 10^7-10^8 yr lifetime of an AGN disc, a 10-$\rm M_\odot$ binary black hole at 1 pc can inspiral to $\sim$1 au. At 1 pc

$$\ell_{\rm heat} = 4\times 10^{35}\left(\frac{m}{10\ \rm M_\odot}\right)^2\left(\frac{1\ \rm au}{a_{\rm final}}\right)\ {\rm erg\ s^{-1}\ pc^{-2}} \tag{21}$$

for the $\dot{M}/\dot{M}_{\rm edd} = 0.1$ model. On the other hand, the cooling luminosity per unit area at 1 pc is

$$\ell_{\rm cool} = \sigma_{\rm sb}T_{\rm eff}^4 \approx 5\times 10^{37}\ {\rm erg\ s^{-1}\ pc^{-2}}. \tag{22}$$

Thus, we do not expect BH–BH captures to significantly perturb the disc. The heating contribution from stellar captures is smaller.

Thus, we do not expect significant heating of the gas via binary inspiral, as cooling dominates heating in all environments we consider. Moreover, the heating rate may be overestimated. Binaries can merge by eccentricity excitation before significantly inspiraling in semimajor axis (see Fig. 4).

8. DISCUSSION AND SUMMARY

The evolution of binaries in gaseous environments was extensively studied over the last few years, in the context of various physical environments, and in particular AGN discs. Here, we focused on binary formation rather than the later stages of the evolution of pre-existing binaries. We made use of analytic arguments also validated with few-body simulations to derive the criteria for gas-assisted binary capture in different astrophysical environments, and discussed its expected rates and implications. We showed that interaction with gas could play a key role in binary formation, depending on the specific conditions of the gaseous and overall environment. We also pointed out several potential caveats, and potential processes that may affect these issues, but are not considered in depth in the current study. The conditions we derived are general and could be applied in principle to any type of gas-rich environment and used to characterize the formed gas-assisted binary population.

Here, we considered several typical gas-rich environments and conditions where gas-assisted binary formation could occur, including SFRs, AGN discs, gas-enriched clusters, and young protoplanetary discs. We find that all of these environments, besides protoplanetary discs, support high rates of gas-assisted binary formation, and thereby this formation channel is expected to significantly affect the binary population and its properties in these environments. In the following, we briefly discuss the implications of gas-assisted captures in specific environments.

8.1 Implications for different environments

(1) SFRs: Gas-assisted binary formation could then prove to be a major channel for the general formation of stellar binaries in SFRs, and hence in the universe at large.

(2) Gas-enriched clusters: In gas-enriched GCs, gas-assisted binary formation could alter the binary population during the early 100 Myr of evolution, if such clusters were gas enriched, as suggested by the existence of multiple stellar populations. Currently used models of GC stellar populations and their evolution do not consider such gas phase, nor its implications for the binary population and evolution. Fundamental aspects of such models should therefore be potentially reconsidered. In addition, the gaseous environment may also give rise to high productions rates of GW sources, even higher than those found by us in Rozner & Perets (2022), where we focused only on primordial binaries, where capture-formed binaries could further increase the rates making this a potential key channel for the origin of GW sources from stellar compact object binaries.

(3) AGN discs: For the case of AGN discs, we pointed out that the required conditions for efficient capture involve a very thin disc. Although such conditions might exist close to the MBH, it is not clear that such large-scale thin discs exist, and observations of large-scale maser discs suggest such discs are in fact very thick (e.g. Yamauchi et al. 2004; Mamyoda et al. 2009, and references therein). Studies suggesting high production rates of GW sources in AGN environments rely on a high supply rate of BH binaries into the disc. Since we find that such high supply rates can only likely be accommodated by the existence of large-scale (parsec scale) discs, and these might be short-lived or rare (if they exist at all), we suggest considering the AGN channel for GW sources with caution. We do note that the existence of a young stellar disc in the Galactic centre (Levin & Beloborodov 2003) suggest the past existence of at least a short-lived large-scale thin gaseous disc (it had to be thin to allow for star formation), but this case is quite different than that envisioned for AGNs.

(4) Protoplanetary discs: Dynamical friction-assisted binary formation was first suggested in this context by Goldreich et al. (2002), and it was shown to be highly efficient. Here, we find that the gas-phase and the generalization to GDF does not give rise to higher rates, as discussed above, and therefore play a lesser role in binary planetesimal formation in such environment.

Finally, we point out that sequential multicaptures may occur and give rise to fast growth of objects, and/or to the formation of high multiplicity systems which later become unstable. Such multicaptures are expected to take place whenever more than one capture occurs per gas lifetime. Detailed study of multicaptures is beyond the scope of this study but will be explored in depth in a dedicate study (in preparation).

ACKNOWLEDGEMENTS

We thank the referee, for his helpful and constructive comments. MR acknowledges the generous support of Azrieli fellowship. AG is supported at the Technion by a Zuckerman Fellowship. We thank Hiromichi Tagawa for insightful conversations.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

Aarseth S. J., Heggie D. C., 1976, A&A, 53, 259
Alexander T., Hopman C., 2009, ApJ, 697, 1861
Antoni A., MacLeod M., Ramirez-Ruiz E., 2019, ApJ, 884, 22
Armitage P. J., 2010, Astrophysics of Planet Formation. Cambridge Univ. Press, Cambridge
Bartos I., Kocsis B., Haiman Z., Márka S., 2017, ApJ, 835, 165
Bastian N., Lardo C., 2018, ARA&A, 56, 83
Bekki K., 2010, ApJ, 724, L99
Bergin E. A., Tafalla M., 2007, ARA&A, 45, 339
Boekholt T. C. N., Rowan C., Kocsis B., 2023, MNRAS, 518, 5653
Chiang E. I., Goldreich P., 1997, ApJ, 490, 368
Duchêne G., Kraus A., 2013, ARA&A, 51, 269
Fabian A. C., Pringle J. E., Rees M. J., 1975, MNRAS, 172, 15
Goldreich P., Lithwick Y., Sari R., 2002, Nature, 420, 643
Gratton R., Bragaglia A., Carretta E., D'Orazi V., Lucatello S., Sollima A., 2019, A&AR, 27, 8
Grishin E., Perets H. B., 2015, ApJ, 811, 54
Grishin E., Perets H. B., 2016, ApJ, 820, 106
Gruzinov A., Levin Y., Matzner C. D., 2020, MNRAS, 492, 2755
Korntreff C., Kaczmarek T., Pfalzner S., 2012, A&A, 543, A126
Kroupa P., 2001, MNRAS, 322, 231
Lee Y.-N., Offner S. S. R., Hennebelle P., André P., Zinnecker H., Ballesteros-Paredes J., Inutsuka S., Kruijssen J. M. D., 2020, Space Sci. Rev., 216, 70
Leigh N. W. C., Böker T., Maccarone T. J., Perets H. B., 2013, MNRAS, 429, 2997
Leigh N. W. C., Mastrobuono-Battisti A., Perets H. B., Böker T., 2014, MNRAS, 441, 919
Levin Y., Beloborodov A. M., 2003, ApJ, 590, L33
Li J., Lai D., Rodet L., 2022, ApJ, 934, 154
Li J., Dempsey A. M., Li H., Lai D., Li S., 2023, ApJ, 944, L42
Maccarone T. J., Zurek D. R., 2012, MNRAS, 423, 2
McKernan B., Ford K. E. S., Lyra W., Perets H. B., 2012, MNRAS, 425, 460
Mamyoda K., Nakai N., Yamauchi A., Diamond P., Huré J.-M., 2009, PASJ, 61, 1143
Mastrobuono-Battisti A., Perets H. B., 2013, ApJ, 779, 85
Mastrobuono-Battisti A., Perets H. B., 2016, ApJ, 823, 61
Moe M., Di Stefano R., 2017, ApJS, 230, 15
Nesvorný D., Youdin A. N., Richardson D. C., 2010, AJ, 140, 785
Noll K. S., Grundy W. M., Chiang E. I., Margot J. L., Kern S. D., 2008, in Barucci M. A., Boehnhardt H., Cruikshank D. P., Morbidelli A., Dotson R., eds, The Solar System Beyond Neptune. Univ. Arizona Press, Tucson, p. 345
Ostriker E. C., 1999, ApJ, 513, 252
Perets H. B., Murray-Clay R. A., 2011, ApJ, 733, 56
Petit J. M., Henon M., 1986, Icarus, 66, 536
Press W. H., Teukolsky S. A., 1977, ApJ, 213, 183
Raghavan D. et al., 2010, ApJS, 190, 1
Rein H., Liu S. F., 2012, A&A, 537, A128
Rein H., Spiegel D. S., 2015, MNRAS, 446, 1424
Renzini A. et al., 2015, MNRAS, 454, 4197
Roupas Z., Kazanas D., 2019, A&A, 621, L1
Rowan C., Boekholt T., Kocsis B., Haiman Z., 2022, preprint (arXiv:2212.06133)
Rozner M., Perets H. B., 2022, ApJ, 931, 149
Sana H. et al., 2012, Science, 337, 444
Shu F. H., Adams F. C., Lizano S., 1987, ARA&A, 25, 23

Sirko E., Goodman J., 2003, MNRAS, 341, 501

Stahler S. W., 2010, MNRAS, 402, 1758

Stone N. C., Metzger B. D., Haiman Z., 2017, MNRAS, 464, 946

Szölgyén Á., Kocsis B., 2018, Phys. Rev. Lett., 121, 101101

Tagawa H., Haiman Z., Kocsis B., 2020, ApJ, 898, 25

Tamayo D., Rein H., Shi P., Hernand ez D. M., 2020, MNRAS, 491, 2885

Thompson T. A., Quataert E., Murray N., 2005, ApJ, 630, 167

Weidenschilling S. J., 1977, MNRAS, 180, 57

Williams J. P., Blitz L., McKee C. F., 2000, in Mannings V., Boss A. P., Russell S. S., eds, Protostars and Planets IV. Univ. Arizona Press, Tucson, p. 97

Yamauchi A., Nakai N., Sato N., Diamond P., 2004, PASJ, 56, 605

APPENDIX A: ANALYTIC SOLUTION FOR VELOCITIES

Here, we present approximate closed-form solution for the velocity as a function of time. If $I(v/c_s) \approx \ln\Lambda$, as in the supersonic regime, then

$$v_{\rm sup}(t) = v_i\left(1-\frac{t}{t_d}\right)^{1/3},$$
$$t_d = \frac{v_i^3}{12\pi G^2 m\rho_{\rm g}\ln\Lambda}. \tag{A1}$$

The velocity decays to zero at t_d, but before this happens, the velocity will become subsonic. In this case $I \approx \frac{v^3}{3c_s^3}$, and

$$v_{\rm sub}(t) = v_o\exp\left[\frac{-(t-t_i)}{\tau}\right],\ \tau = \frac{3c_s^3}{4\pi G^2 m\rho_g}.$$

With the above solutions we can compare the length-scales over which the stars in the binary decelerate (ℓ_d). In the supersonic case,

$$\frac{\ell_{d,1}}{\ell_{d,2}} = \frac{v_{i,1}t_{d,1}}{v_{i,2}t_{d,2}} = \frac{v_{i,1}^4}{v_{i,2}^4}\frac{m_2}{m_1} = q^5. \tag{A2}$$

Aforementioned, subscripts 1 and 2 denote the primary and secondary star of the binary, respectively, and q is the ratio between the secondary and primary mass. In the subsonic case,

$$\frac{\ell_{d,1}}{\ell_{d,2}} = \frac{v_{i,1}\tau_1}{v_{i,2}\tau_2} = \frac{v_{i,1}}{v_{i,2}}\frac{m_2}{m_1} = q^2. \tag{A3}$$

APPENDIX B: DETAILED DERIVATION OF THE THRESHOLD VELOCITIES FOR CAPTURE

Here, we outline the derivations of the threshold velocities in Table 1 in more detail case-by-case.

B1 Supersonic, unfocused threshold

In the supersonic, unfocused limit the work done on star i is

$$W_i \approx \mathbf{F_{GDF,i}}\cdot\boldsymbol{\ell_i} \approx \frac{4\pi G^2\ln(\Lambda)\rho_{\rm g}m_i^2}{v_i^2}\ell_i, \tag{B1}$$

where m_i is the mass, v_i is the velocity, $\rho_{\rm g}$ is the gas density, $\ln\Lambda$ is the Coulomb logarithm, and ℓ_i is the path-length travelled. We assume the binary centre-of-mass is at rest with respect to the gas. Thus, the masses, velocities, and path-lengths can be rewritten in terms of the total mass ($m_{\rm bin}$), the mass ratio (q), and the Hill radius ($R_{\rm Hill}$), viz.

$$m_1 = \frac{m_{\rm bin}}{1+q},\ m_2 = \frac{m_{\rm bin}q}{1+q}$$
$$v_1 = \frac{v_\infty q}{1+q},\ v_2 = \frac{v_\infty}{1+q}$$
$$\ell_1 \approx q^5 R_{\rm Hill}\sqrt{1-\frac{b^2}{R_{\rm Hill}^2}},\ \ell_2 \approx R_{\rm Hill}\sqrt{1-\frac{b^2}{R_{\rm Hill}^2}}. \tag{B2}$$

Aforementioned, b is the impact parameter. We use the initial velocity to estimate the energy dissipated. This is justified because most energy will be dissipated at large velocities, due to the quadratic dependence of kinetic energy on velocity. The path-length of the secondary star is simply the straight-line distance through the Hill sphere.[5] The path-length of the primary is q^5 times this distance from the preceding Appendix. Then, the total work done is

$$W_{\rm tot} = \frac{4\pi G^2 m_{\rm bin}^2\ln(\Lambda)\rho_{\rm g}R_{\rm Hill}q^2(1+q)}{v_\infty^2}\sqrt{1-\left(\frac{b}{R_{\rm Hill}}\right)^2}. \tag{B3}$$

Finally, we equate $W_{\rm tot}$ with the energy of the unbound orbit ($\frac{1}{2}\frac{m_{\rm bin}q}{(1+q)^2}v_\infty^2$) to obtain the threshold capture velocity:

$$v_c = \underbrace{\left(8\pi G^2\rho_{\rm g}m_{\rm bin}R_{\rm Hill}\ln\Lambda\right)^{1/4}}_{v_x} q^{1/4}(1+q)^{3/4}\left(1-\frac{b^2}{R_{\rm hill}^2}\right)^{1/8}. \tag{B4}$$

For simplicity, we drop the last term.

B2 Subsonic, unfocused threshold

The derivation for the subsonic, unfocused case is similar, except the work done on star i is

$$W_i \approx \frac{4\pi G^2\ln(\Lambda)\rho_{\rm g}m_i^2}{3c_{\rm s}^3}v_i\ell_i, \tag{B5}$$

and the path-length of the primary star is q^2 times the path-length of the secondary (see Appendix A). Then the velocity threshold is

$$v_c = \underbrace{\frac{8\pi G^2\rho_{\rm g}m_{\rm bin}R_{\rm Hill}}{3c_{\rm s}^3}}_{v_{\rm s}} q\sqrt{1-\frac{b^2}{R_{\rm Hill}^2}}. \tag{B6}$$

B3 Supersonic, focused threshold

In the focused regime, we approximate the stellar trajectories as parabolic. The separation between the stars, r, and their relative velocity $v_{\rm rel}$, are

$$r = \frac{2xR_{\rm Hill}}{1+\cos(f)}$$
$$v_{\rm rel} = \sqrt{\frac{2Gm_{\rm bin}}{r}}, \tag{B7}$$

where f is the true anomaly of the orbit and x is the pericentre distance in units of the Hill radius. In the supersonic regime, the work done

---

[5] More precisely $\ell_2 = \frac{2R_{\rm Hill}}{1+q}\sqrt{1-\frac{b^2}{R_{\rm Hill}^2}}$

on star i is

$$W_i = 4\pi G^2 \rho_{\rm g} m_i^2 \ln(\Lambda) \int_{t_{\rm start}}^{t_{\rm end}} v_i^{-1} {\rm d}t$$
$$= 4\pi G^2 \rho_{\rm g} m_i^2 \ln(\Lambda) \int_{-f_{\rm end}}^{f_{\rm end}} v_i^{-1} \sqrt{\frac{(x R_{\rm Hill})^3}{2 G m_{\rm bin}}} \sec(f/2)^4 {\rm d}f. \quad \text{(B8)}$$

Above $\pm f_{\rm end} = \pm \arccos(2x - 1)$ are the true anomalies where the distance between the stars exceeds the Hill radius. Thus, the critical velocity capture

$$v_c = \frac{v_x^2}{v_{\rm esc}} \frac{\sqrt{1 + q + q^3 + q^4}}{q} h(x), \quad \text{(B9)}$$

where v_x is defined in equation (B4) and $v_{\rm esc}$ is the escape velocity at the Hill radius. $h(x)$ is a complicated function. However, $1 < h(x) < 2$, except very close to $x = 1$, and is dropped for simplicity. We also drop the last two terms under the square root. (Note $0 < q \leq 1$.)

B4 Subsonic, focused threshold

In the subsonic limit, the work done on star i is

$$W_i = \frac{4\pi G^2 \rho_{\rm g} m_i^2}{3 c_{\rm s}^3} \int_{t_{\rm start}}^{t_{\rm end}} v_i^2 {\rm d}t$$
$$= \frac{4\pi G^2 \rho_{\rm g} m_i^2}{3 c_{\rm s}^3} \int_{-f_{\rm end}}^{f_{\rm end}} v_i^2 \sqrt{\frac{(x R_{\rm Hill})^3}{2 G m_{\rm bin}}} \sec(f/2)^4 {\rm d}f. \quad \text{(B10)}$$

Thus, the velocity threshold is

$$v_c = \sqrt{8 v_{\rm esc} v_{\rm s}} \frac{\sqrt{q}}{1 + q} (1 - x)^{1/4}. \quad \text{(B11)}$$

For simplicity, we drop the last term on the right-hand side.

Aforementioned, we have ignored some edge cases. For example, the secondary may be supersonic, while the primary is subsonic. However, the approximate thresholds derived here are within a factor of ∼2 of the thresholds from our few-body simulations and are adequate for our purposes.

This paper has been typeset from a TEX/LATEX file prepared by the author.

5.3 Gas shielding of soft binaries

Based on Rozner & Perets 2024

As discussed in 1.4.2, Heggie-Hills law (Heggie, 1975; Hills, 1975) states that when encountering other stars, hard binaries (binaries with energy that exceed the mean energy of objects in the cluster, $|E| \gtrsim \langle m\sigma^2 \rangle$) tend statistically to get harder and soft binaries ($|E| \lesssim \langle m\sigma^2 \rangle$) tend statistically to get softer and finally disrupted. However, in gas-rich environments, this law could be revised.

The definition of soft-hard binary sets a critical semimajor axis, above it a binary will soften, and below it, a binary will harden. This critical separation could be calculated by equating the binary energy to the mean energy of the cluster, which yields

$$a_{\rm SH} = \frac{Gm_1m_2}{2\bar{m}\sigma^2} \tag{5.14}$$

However, in gas-rich environments, this critical separation is shifted, as all the binaries, including the soft ones, are considered hard relative to the gas, i.e. the separations of all the binaries will shrink due to the interaction with the gas. In these environments, the hardening/softening processes are governed by three components: (i) stellar hardening (for hard binaries) (ii) stellar softening (for soft binaries) and (iii) gas-hardening.

The softening rate could be calculated using the diffusion coefficients (Heggie, 1975; Binney and Tremaine, 2008),

$$\langle \dot{E}_{\rm soft} \rangle \approx \langle D[\Delta \tilde{E}] \rangle \approx \frac{8\sqrt{\pi}G^2\mu\bar{m}\rho_\star \ln \Lambda_{\rm bin}}{\sigma} \tag{5.15}$$

where the Coulomb factor is $\Lambda_{\rm bin} = a\sigma^2/(4G\bar{m})$.

The rate at which hard binaries become harder is (Heggie, 1975; Spitzer, 1987; **?**; Binney and Tremaine, 2008; Celoria et al., 2018)

$$\langle \dot{E}_{\rm hard} \rangle = 2\pi \frac{G^2 m_1 m_2 \rho_\star (M_{\rm bin} + \bar{m})}{M_{\rm bin}\sigma} \tag{5.16}$$

Up to a factor of order of unity, where $E_{\rm bin}$ is the energy of the binary and $\bar{m}$ is the (mean) mass of the perturber. The softening/hardening rates in terms of a are given correspondingly by

$$\left.\frac{da}{dt}\right|_{\rm soft} = \frac{16\sqrt{\pi} G\bar{m}\rho_\star \ln\Lambda_{\rm bin}}{M_{\rm bin}\sigma} a^2, \tag{5.17}$$

$$\left.\frac{da}{dt}\right|_{\rm hard} = -2\pi \frac{G\rho_\star (M_{\rm bin} + \bar{m})}{\sigma M_{\rm bin}} a^2 \tag{5.18}$$

where $\rho_\star$ is the background density of the stellar perturbers.

Gas-hardening could be modeled in different approaches, in each of them the rate of energy dissipation could change, but qualitatively, the effect would be the same – gas-hardening will add a term that could compete with the binary stellar softening. Here we will focus on gas dynamical friction (Ostriker, 1999). The GDF force on an object with mass m is (Ostriker, 1999),

$$\mathbf{F}_{\rm GDF} = -\frac{4\pi G^2 m^2 \rho_g}{v_{\rm rel}^3} \boldsymbol{v}_{\rm rel} I(v/c_s) \tag{5.19}$$

where G is the gravitational constant, ρ_g is the gas density, c_s is the sound speed, and $\mathbf{v}_{\rm rel}$ is the relative velocity between the object and the gas. The function I is given by

$$I(\mathcal{M}) = \begin{cases} \frac{1}{2}\log(1-\mathcal{M}^{-2}) + \ln\Lambda, & \mathcal{M} > 1 \\ \frac{1}{2}\log\left(\frac{1+\mathcal{M}}{1-\mathcal{M}}\right) - \mathcal{M}, & \mathcal{M} < 1 \end{cases} \tag{5.20}$$

where $\mathcal{M} = v/c_s$ is the Mach number.

The semimajor axis evolution is given by

$$\begin{aligned} \left.\frac{da}{dt}\right|_{\mathrm{GDF}} &= -\frac{8\pi G^{3/2}a^{3/2}}{\sqrt{m_1+m_2}}\rho_g(t)\frac{m_1}{v_{\mathrm{rel}}^2}I\left(\frac{v_{\mathrm{rel}}}{c_s}\right)\xi(q), \\ \xi(q) &= (1+q^{-1})^2 + q(1+q)^2 \end{aligned} \tag{5.21}$$

where $q = m_2/m_1$ is the mass ratio of the binary. Note that this equation differs from eq. 7 in Rozner and Perets (2022a) by a factor of unity, due to mass-ratio corrections. The relative velocity is taken as $\max\{\sigma, v_{\mathrm{Kep}}\}$. The energy evolution is given by

$$\left.\frac{dE}{dt}\right|_{\mathrm{GDF}} = -\frac{4\pi G^{5/2}m_1^2 m_2}{\sqrt{m_1+m_2}}\frac{\rho_g}{v_{\mathrm{rel}}^2}I\left(\frac{v_{\mathrm{rel}}}{c_s}\right)\xi(q)a^{-1/2} \tag{5.22}$$

The total separation evolution of the binary is then given by

$$\frac{da}{dt} = \begin{cases} \left.\frac{da}{dt}\right|_{\mathrm{GDF}} + \left.\frac{da}{dt}\right|_{\mathrm{hard}}, & \text{hard binary,} \\ \left.\frac{da}{dt}\right|_{\mathrm{GDF}} + \left.\frac{da}{dt}\right|_{\mathrm{soft}}, & \text{soft binary} \end{cases} \tag{5.23}$$

We define the *shielding radius* as the separation in which $\dot{a} = 0$, and essentially this is the new hard-soft boundary for binaries in gas-rich environments. In Fig. 5.8, we present the shielding radius for different background gas densities and masses. As can be seen, the shielding radius when the gas density is significant could be as large as the tidal radius of a cluster, indicating that all the binaries in this cluster during this epoch (assuming this density) could be treated as hard.

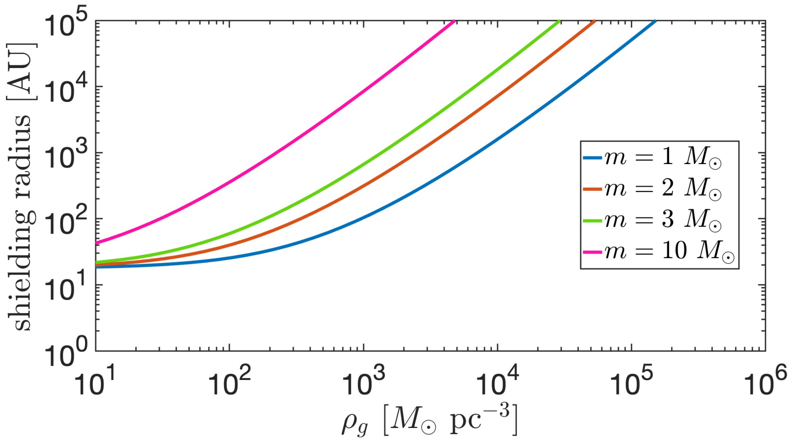

Figure 5.8: The shielding radius – the maximal radius for which a binary is considered hard when accounting for the effect of both gas hardening and stellar encounters, as a function of different gas densities, for different binaries of equal masses $m = m_1 = m_2$. Taken from Rozner and Perets (2024).

In Fig. 5.9, we compare the evolution of a binary when considering only the stellar effects, rather than the stellar interactions as well as the contribution from the gas shielding. As can be seen, without considering the contribution from gas hardening, this binary is considered as soft and is prone to disruption. With gas hardening, this binary will finally harden rather than soften.

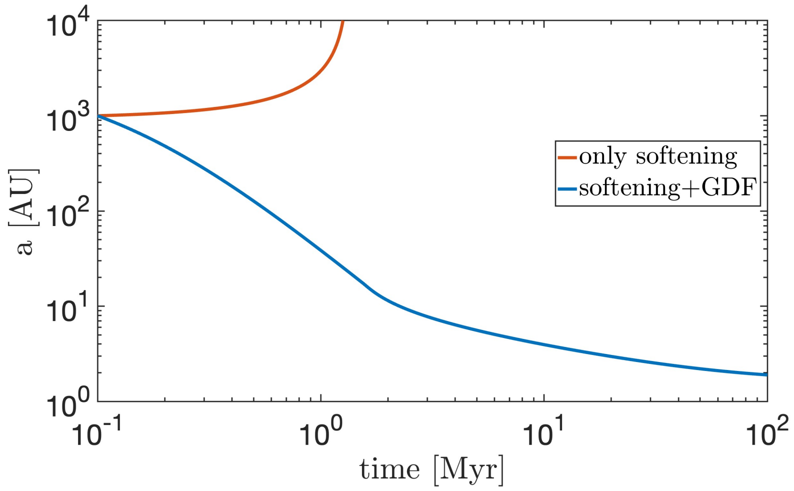

Figure 5.9: Comparison between the semimajor axes evolution with and without the contribution of gas-shielding, for a binary with $m = m_1 = m_2 = 1\ M_\odot$.

In Fig. 5.10, we present the gas density and semimajor axis evolution over time. The black dashed line in the lower panel is the soft-hard binary limit. In gas-dilute environments, binaries with separations lower than this separation will statistically tend to get harder while binaries with separations above the critical separation are expected to get statistically softer. However, in gas-rich environments, as can be seen, the binaries with separations larger than this critical separation could still get harder, as long as the separation is smaller than the shielding radius. The typical gas lifetime here is 50 Myr, and it could be seen that binaries that harden initially will soften after a while as the gas decay and gas shielding is not affected anymore.

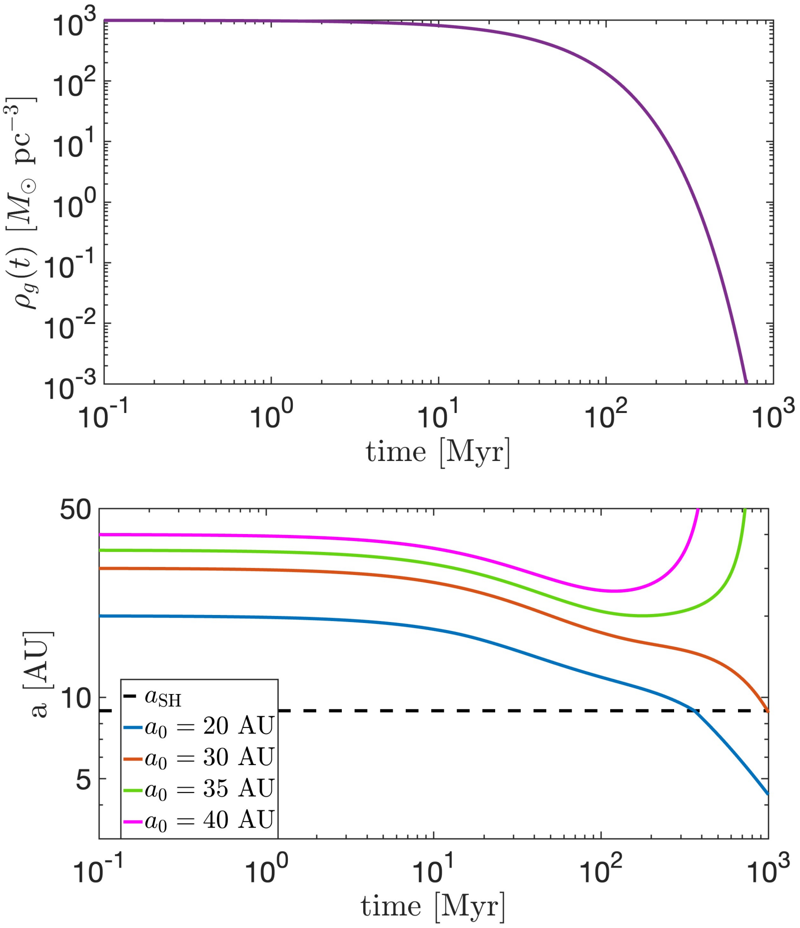

Figure 5.10: Upper panel: the gas density evolution over time, for initial gas density of $\rho_{g,0} = 10^3\ M_\odot\ \mathrm{pc}^{-3}$. Lower panel: The evolution of different initial semimajor axes for a background density with $\rho_{g,0} = 10^3\ M_\odot\ \mathrm{pc}^{-3}$, for a binary with masses $m = m_1 = m_2 = 1\ M_\odot$. The dashed line corresponds to the soft-hard limit boundary as derived based on stellar interactions only.

Gas shielding affects also the whole population of binaries and their properties. In Fig. 5.11 and 5.12 we present the separation distribution of binaries over time, for two gas densities, $rho_{g,0} = 10^3\ M_\odot\ \mathrm{pc}^{-3}$ (upper panels) and $rho_{g,0} = 10^4\ M_\odot\ \mathrm{pc}^{-3}$ (lower panels). As can be seen, binaries with separation smaller than the shielding radius will shrink their separations (note that the shielding radius changes with time as the gas density decays), and binaries with separations larger than this radius increase their separations. The overall fractions of hard binaries increase significantly.

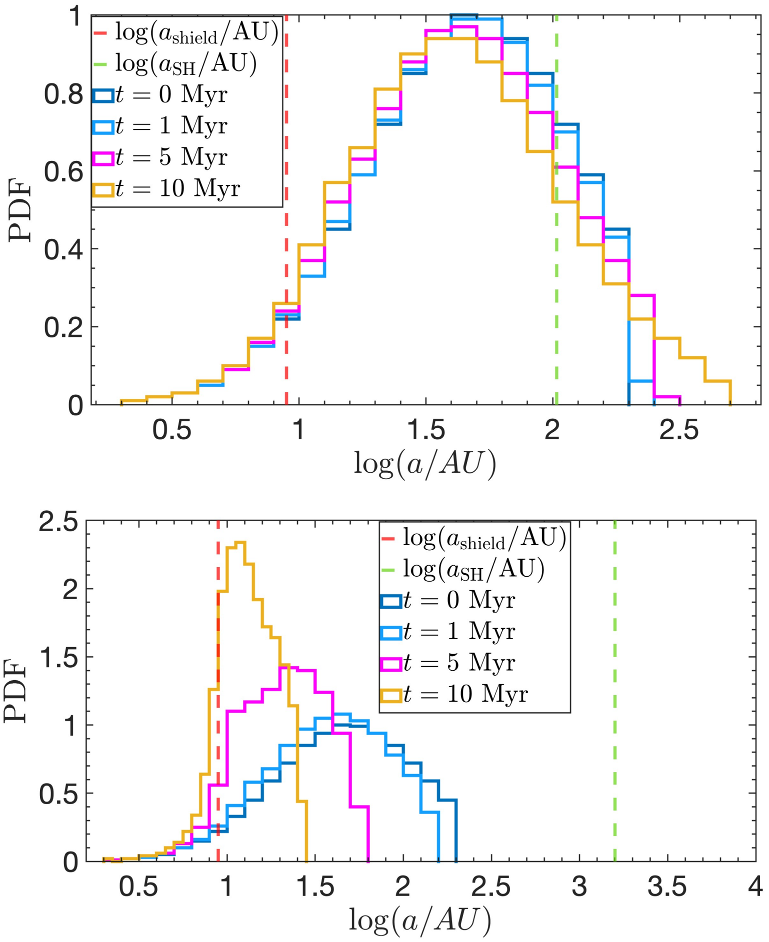

Figure 5.11: The results of a Monte Carlo simulation, with two equal masses $m = m_1 = m_2 = 1\ M_\odot$ averaged over 1000 iterations. The red dashed line corresponds to the critical semimajor axis between soft and hard binaries when considering stellar interactions only (without the effect of gas), and the green dashed line corresponds to the shielding radius as calculated relative to the *initial* gas density. The different solid lines correspond to different times. Upper panel: Initial background density of $\rho_{g,0} = 10^3\ M_\odot\ \mathrm{pc}^{-3}$. Lower panel: $\rho_{g,0} = 10^4\ M_\odot\ \mathrm{pc}^{-3}$. Note the different scales.

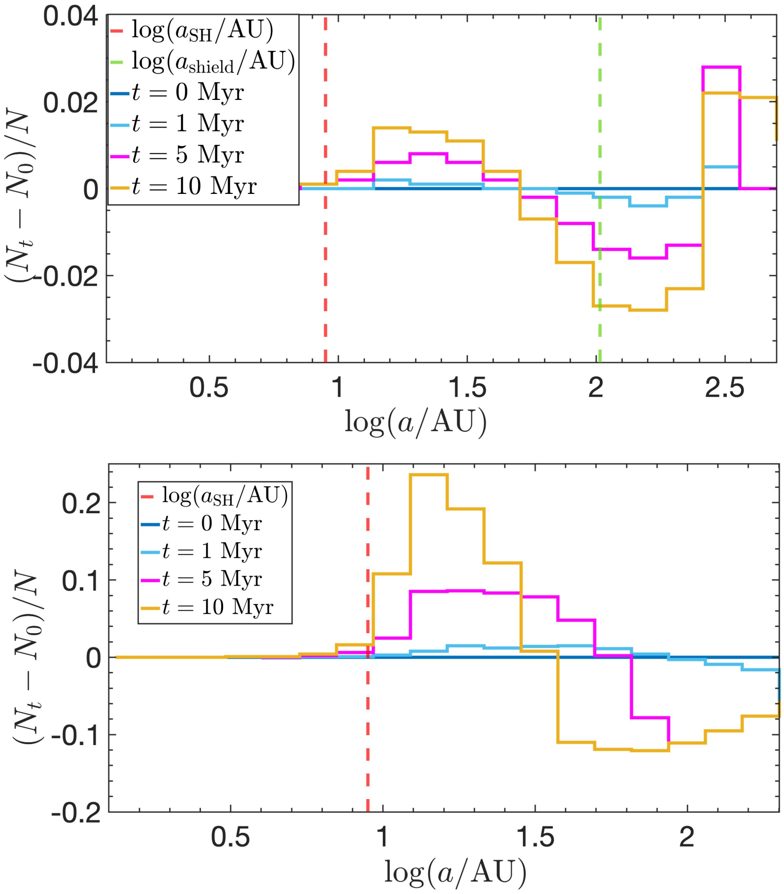

Figure 5.12: The normalized difference between the number of binaries in different separations, based on the results of our Monte Carlo simulation, with two equal masses $m = m_1 = m_2 = 1\ M_\odot$ averaged over 1000 iterations. The red dashed line corresponds to the critical semimajor axis between soft and hard binaries when considering stellar interactions only (without the effect of gas), and the green dashed line corresponds to the shielding radius as calculated relative to the *initial* gas density. The different solid lines correspond to different times. Upper panel: Initial background density of $\rho_{g,0} = 10^3\ M_\odot\ {\rm pc}^{-3}$. Lower panel: $\rho_{g,0} = 10^4\ M_\odot\ {\rm pc}^{-3}$. Note the different scales.

Gas shielding has various implications on binary evolution and the overall dynamics of clusters. As hard binaries give rise to many astrophysical phenomena such as supernovae and gravitational waves, higher fraction of hard binaries will lead to higher rates of these events. On the other direction, observed events could help us set constraints on the gas densities during different epochs.

Draft version April 3, 2024
Typeset using LaTeX **twocolumn** style in AASTeX631

Soft no more: gas shielding protects soft binaries from disruption in gas-rich environments

Mor Rozner[1] and Hagai B. Perets[1, 2]

[1] *Technion - Israel Institute of Technology, Haifa, 3200002, Israel*
[2] *Department of Natural Sciences, The Open University of Israel, 1 University Road, PO Box 808, Raanana 4353701, Israel*

ABSTRACT

Binaries in dense environments are traditionally classified as soft or hard based on their binding energy relative to the kinetic energy of surrounding stars. Heggie's law suggests that stellar encounters tend to soften soft binaries and harden hard binaries, altering their separations. However, interactions with gas in such environments can significantly modify this behavior. This study investigates the impact of gas on binary softening and its consequences. We find that gas interactions can actually harden binaries, extending the soft-hard boundary to larger separations. This introduces a "shielding radius" within which binaries are likely to harden due to gas interactions, surpassing the traditional soft-hard limit. Consequently, a notable portion of binaries initially classified as "soft" may become "hard" when both gas and stars are considered. We propose a two-stage formation process for hard binaries: initial soft binary formation, either dynamically or through gas-assisted capture, followed by gas-induced hardening before eventual disruption. In environments with low gas density but high gas content, the shielding radius could exceed the typical hard-soft limit by an order of magnitude, leading to a significant fraction of originally soft binaries effectively becoming hard. Conversely, in high gas-density environments, gas-induced hardening may dominate, potentially rendering the entire binary population hard. Gas hardening emerges as a crucial factor in shaping binary populations in gas-rich settings, such as clusters, star-forming regions, and possibly AGN disks. This highlights the complex interplay between gas dynamics and stellar interactions in binary evolution within dense environments.

1. INTRODUCTION

Binaries in dense environments could be categorized into two groups, based on their energy relative to the mean energy of their environment. Soft binaries are defined as binaries with low energies relative to the mean energy of binaries in the cluster, i.e. $|\tilde{E}| \lesssim \langle m\sigma^2 \rangle$, while hard binaries have energies that exceed the mean energy of the cluster, i.e. $|\tilde{E}| \gtrsim \langle m\sigma^2 \rangle$. The evolution of the two groups differs qualitatively from each other. Heggie's law states that soft binaries tend on average to get softer, while hard binaries tend to get harder (Heggie 1975; Hills 1975), albeit this could be somewhat modified when accounting for external potentials (Ginat & Perets 2021a).

The analysis of hard/soft dynamics has considered only dynamic interactions between stars in dense star clusters. However, under a wide range of conditions, dense stellar clusters can be highly enriched in ambient gas. This naturally occurs in the early star-forming epoch, at which stars are formed from the collapsing and fragmenting gaseous background and are embedded in gas until the gas dissipates and/or accreted. Furthermore, older clusters can also be enriched in gas, e.g. stars embedded in AGN (active galactic nucleus) disks, nuclear clusters enriched in gas from infalling gas clouds, and/or gas ejected by evolved stars through winds. Similarly, globular clusters, in which multiple populations of stars are observed suggest multiple star-formation epochs (e.g. Carretta et al. 2009; Renzini et al. 2015; Gratton et al. 2019; Bastian & Lardo 2018), in which earlier-formed stars from the previous generation(s) would be embedded in newly enriched gas needed for the formation of a later generation of gas.

In such gas-enriched environments, in addition to stellar interactions, the interaction with the gas leads to energy dissipation and further unique dynamical phenomena. Accretion onto stars/compact objects in gaseous environments could give rise to X-ray flares (Bahcall & Ostriker 1976) or explode via supernovae (Ostriker 1983; Artymowicz et al. 1993). Gas hardening was investigated thoroughly, in molecular clouds (Stahler 2010), protoplanetary disks (Perets & Murray-Clay 2011; Gr-

morozner@campus.technion.ac.il

ishin & Perets 2016), as a catalyzer for various astrophysical processes including gravitational waves (GWs) mergers (McKernan et al. 2012; Stone et al. 2017; Tagawa et al. 2020; Rozner & Perets 2022), gas-assisted binary formation (Tagawa et al. 2020; Rozner et al. 2023; Li et al. 2023; Rowan et al. 2023; Whitehead et al. 2023; Li & Lai 2024) and enhanced formation of massive black holes (Roupas & Kazanas 2019).

Binary formation in gas could be divided into three categories (e.g. Bonnell 2001; Kratter 2011 and references therein): in-situ formation in which the binaries are formed following the collapse of gas clumps in the separation in which they are observed today (which includes fission and fragmentation), collapse formation in a wider separation than their current, followed by a migration triggered by a dissipative force; and gas-assisted binary capture (Tagawa et al. 2020; Rozner & Perets 2022; Whitehead et al. 2023). Some of the suggested dissipation mechanisms following the formation take place either through an interaction with a circumstellar disk (Clarke & Pringle 1991; Hall et al. 1996), or through interaction with the ambient surrounding gas Other binary-formation channels include 3-body interaction (Mansbach 1970; Aarseth & Heggie 1976; Goodman & Hut 1993; Ginat & Perets 2024), tidal capture (Press & Teukolsky 1977), dynamical friction assisted capture, also known as L2 (Goldreich et al. 2002), and gas-drag formation of planetesimal binaries (Ormel & Klahr 2010; Perets & Murray-Clay 2011).

Although some of these mechanisms occur in naturally gas-rich environments and are also dense with stars, such as AGN disks, the hardness/softness of binaries, or the interconnection between interactions induced by the stellar and gaseous background, has been little explored in this context. Here we focus on these issues.

We examine the effect of gas-hardening on binaries in GCs, together with stellar hardening/softening. We derive analytically the new soft-hard boundary, accounting for the effect of the gas. We then discuss further implications on cluster dynamics and the population of hard/soft binaries.

In section 2, we briefly overview the components that different contributions to the dynamics of binaries in gas. In section 3 we introduce the concept of gas shielding. In section 4 we discuss our results, possible caveats and future implications. In section 5, we summarize our results and conclude.

2. DYNAMICS OF BINARIES IN GASEOUS ENVIRONMENTS

There are several approaches to model dynamics in gas. Some of them are mini-accretion disks (Artymowicz et al. 1991; McKernan et al. 2012; Stone et al. 2017; Tagawa et al. 2020), simulating the Bondi-Hoyle accretion supersonic flows and deriving the corresponding energy dissipation rate (Antoni et al. 2019) and gas dynamical friction (Ostriker 1999). In this paper, we will use the latter unless stated otherwise. While the details differ from one model to another, all of them model energy dissipation, and hence the concept of the process we present below would also hold for other descriptions of motion in gas, with the proper modifications. The motion of objects in gas is still not completely understood and is under active research, with some fundamental issues still debated, including the direction of the migration (e.g. Moody et al. 2019; Lai & Muñoz 2023; Grishin et al. 2023; Duffell et al. 2024).

2.1. *Hardening/softening through dynamical encounters*

A binary is called hard if its energy exceeds $\bar{m}\sigma^2$. This condition sets for every pair of masses a critical semimajor axis. Binaries with larger semimajor axes will be soft and binaries with smaller semimajor axes will be hard.

$$a_{\rm SH} = \frac{Gm_1m_2}{2\bar{m}\sigma^2} \tag{1}$$

Binary softening is dominated by a series of distant encounters, gradually increasing the internal energy of the binary, and could potentially lead to positive energy, i.e. disruption of the binary. The softening rate could be calculated using the diffusion coefficients (Heggie 1975; Binney & Tremaine 2008),

$$\langle \dot{E}_{\rm soft} \rangle \approx \langle D[\Delta \tilde{E}] \rangle \approx \frac{8\sqrt{\pi}G^2\mu\bar{m}\rho_\star \ln \Lambda_{\rm bin}}{\sigma} \tag{2}$$

where the Coulomb factor is $\Lambda_{\rm bin} = a\sigma^2/(4G\bar{m})$.

The rate at which hard binaries become harder is (Heggie 1975; Spitzer 1987; Heggie & Hut 1993; Binney & Tremaine 2008; Celoria et al. 2018)

$$\langle \dot{E}_{\rm hard} \rangle = 2\pi \frac{G^2 m_1 m_2 \rho_\star (M_{\rm bin} + \bar{m})}{M_{\rm bin}\sigma} \tag{3}$$

Up to a factor of order of unity, where $E_{\rm bin}$ is the energy of the binary and $\bar{m}$ is the (mean) mass of the perturber. For an analytic derivation of the hardening rate considering both energy and angular momentum in the equal masses case see Ginat & Perets 2021b.

The softening/hardening rates in terms of a are given correspondingly by

$$\left.\frac{da}{dt}\right|_{\rm soft} = \frac{16\sqrt{\pi}G\bar{m}\rho_\star \ln\Lambda_{\rm bin}}{M_{\rm bin}\sigma}a^2, \tag{4}$$

$$\left.\frac{da}{dt}\right|_{\rm hard} = -2\pi\frac{G\rho_\star(M_{\rm bin}+\bar{m})}{\sigma M_{\rm bin}}a^2 \tag{5}$$

where $\rho_\star$ is the background density of the stellar perturbers. It should be noted that while binary hardening/softening is usually considered relative to a background of $\bar{m}$, in some cases, less frequent interactions with more massive objects could lead overall to more significant effects. Black holes in the background of other black holes are considered hard for semimajor axes lower than $a_{\rm SH,\bullet} \approx 1.35$AU (see Quinlan 1996; Kritos & Cholis 2020; Rozner & Perets 2022), while for stellar background, they are considered hard for semimajor axes below $a_{\rm SH,\star} \approx 200.53$ AU. In between, these binaries are considered hard relative to other black holes but soft relative to stars. However, the contribution from stellar hardening dominates over the softening due to black holes, such that we can define all the stellar softening/hardening based on stellar background only.

$$\left|\frac{da/dt|_{\rm soft,\bullet}}{da/dt|_{\rm hard,\star}}\right| = \frac{8}{\sqrt{\pi}}\frac{M_\bullet}{M_{\rm bin}+m_\star}\frac{\rho_\bullet}{\rho_\star}\ln\Lambda_{\rm bin} \sim 10^{-2} \ll 1 \tag{6}$$

2.2. *Gas dynamical friction*

The GDF force on an object with mass m is (Ostriker 1999),

$$\mathbf{F}_{\rm GDF} = -\frac{4\pi G^2 m^2 \rho_g}{v_{\rm rel}^3}\boldsymbol{v}_{\rm rel} I(v/c_s) \tag{7}$$

where G is the gravitational constant, ρ_g is the gas density, c_s is the sound speed, and $\mathbf{v}_{\rm rel}$ is the relative velocity between the object and the gas. The function I is given by

$$I(\mathcal{M}) = \begin{cases} \frac{1}{2}\log(1-\mathcal{M}^{-2}) + \ln\Lambda, & \mathcal{M} > 1 \\ \frac{1}{2}\log\left(\frac{1+\mathcal{M}}{1-\mathcal{M}}\right) - \mathcal{M}, & \mathcal{M} < 1 \end{cases} \tag{8}$$

where $\mathcal{M} = v/c_s$ is the Mach number.

The semimajor axis evolution is given by

$$\left.\frac{da}{dt}\right|_{\rm GDF} = -\frac{8\pi G^{3/2}a^{3/2}}{\sqrt{m_1+m_2}}\rho_g(t)\frac{m_1}{v_{\rm rel}^2} I\left(\frac{v_{\rm rel}}{c_s}\right)\xi(q), \tag{9}$$
$$\xi(q) = (1+q^{-1})^2 + q(1+q)^2$$

where $q = m_2/m_1$ is the mass ratio of the binary. Note that this equation differs from eq. 7 in Rozner & Perets (2022) by a factor of unity, due to mass-ratio corrections. The relative velocity is taken as $\max\{\sigma, v_{\rm Kep}\}$. The energy evolution is given by

$$\left.\frac{dE}{dt}\right|_{\rm GDF} = -\frac{4\pi G^{5/2} m_1^2 m_2}{\sqrt{m_1+m_2}}\frac{\rho_g}{v_{\rm rel}^2} I\left(\frac{v_{\rm rel}}{c_s}\right)\xi(q)a^{-1/2} \tag{10}$$

3. GAS SHIELDING

The process of hardening/softening binaries could significantly change in the presence of gas. The interaction with gas leads to energy dissipation, which could potentially compete with the stellar softening of soft binaries. Binaries that were considered soft, accounting for stellar interactions only, could become hard when adding the contribution from gas-hardening. We term this process as *gas shielding*, as the gas shields soft binaries from disruption, by dissipating their orbits, making them hard binaries (with respect to stellar encounters) before encounters with other stars soften and eventually disrupt them. Hence, the definition of soft/hard binaries should be revised in the presence of gas, to include the contribution of gas hardening. Hereafter, we will demonstrate the effect of gas-shielding in GCs (and briefly consider also other gas-rich environments). This is a general process, that could occur in general in any gas-rich media, with the proper modifications.

For any set of parameters of the binary and its environments, one can define the *shielding radius*. The shielding radius is the critical semimajor axis between soft and hard binaries, i.e. the hardest soft binary/the softest hard binary, when we consider the combined effect of stellar hardening/softening and gas hardening.

For a given gas density, stellar number density, and typical velocity dispersion, the shielding semimajor axis could be calculated by setting $\dot{a}_{\rm tot} = \dot{a}_{GDF} + \dot{a}_{\rm soft} = 0$, which is given by

$$a_{\rm shield}^{1/2}\ln\left(\frac{a_{\rm shield}\sigma^2}{4G\bar{m}}\right) = \frac{\sqrt{\pi}}{2}\frac{\sqrt{GM_{\rm bin}}}{\sigma}\frac{\rho_g}{\rho_\star}\frac{m_1}{\bar{m}} I\left(\frac{\sigma}{c_s}\right)\xi(q) \tag{11}$$

which yields an analytical solution

$$a_{\rm shield} = \frac{B^2}{4W^2\left(-\frac{B\sigma}{4\sqrt{G\bar{m}}}\right)}, \tag{12}$$

$$B = \frac{\sqrt{\pi}}{2}\frac{\sqrt{GM_{\rm bin}}}{\sigma}\frac{\rho_g}{\rho_\star}\frac{m_1}{\bar{m}} I\left(\frac{\sigma}{c_s}\right)\xi(q) \tag{13}$$

where W is the Lambert W-function. We considered the relative velocity in eqs. 11,12 to be the velocity dispersion of stars.

name	notation	fiducial value
gas density	ρ_g	$10^5\ M_\odot\ \mathrm{pc}^{-3}$
stellar density	$\rho_\star$	$10^4\ M_\odot\ \mathrm{pc}^{-3}$
velocity dispersion	σ	10 km/sec
sound speed	c_s	10 km/sec
gas lifetime	τ_{gas}	50 Myr

Table 1. The fiducial values used along the paper unless stated otherwise.

The critical density for which a given semimajor axis a, which is considered soft when taking into account only stellar hardening, becomes hard due to the contribution from gas hardening, is given by

$$\rho_{\mathrm{g,crit}} = \frac{2}{\sqrt{\pi}}\rho_\star \frac{\ln\Lambda_{\mathrm{bin}}}{I(\sigma/c_s)\xi(q)}\frac{\bar{m}}{m_1}\frac{\sigma}{\sqrt{GM_{\mathrm{bin}}/a}} \qquad (14)$$

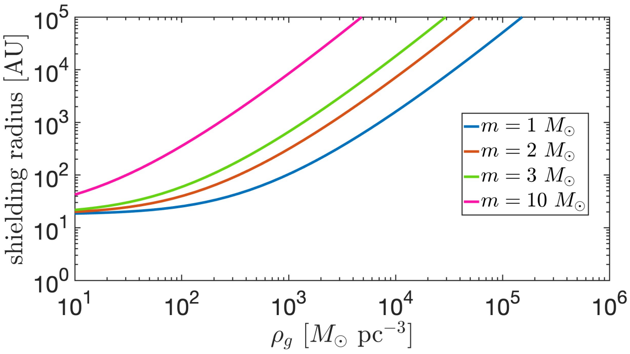

Figure 1. The shielding radius – the maximal radius for which a binary is considered hard when accounting for the effect of both gas hardening and stellar encounters, as a function of different gas densities, for different binaries of equal masses $m = m_1 = m_2$.

In Fig. 1, we present the shielding radius, i.e. the largest semimajor axis from which binaries will harden when taking into consideration both stellar hardening/softening and gas hardening, as a function of the gas density. As can be seen, higher gas densities dictate larger shielding radii, as expected, since then the effect of gas hardening is stronger. Higher masses give rise to larger shielding radii, as gas hardening in the regime of soft binaries scales as a positive power of the mass.

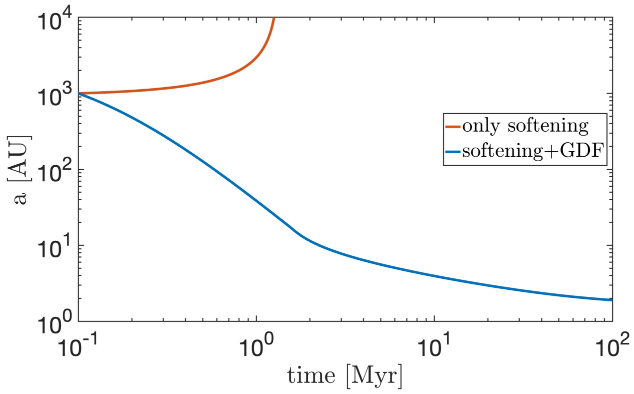

Figure 2. Comparison between the semimajor axes evolution with and without the contribution of gas-shielding, for a binary with $m = m_1 = m_2 = 1\ M_\odot$.

In Fig. 2, we examine the semimajor axis evolution of a binary with masses $m = m_1 = m_2 = 1\ M_\odot$ and initial semimajor axis of $a_0 = 10^3$ AU, with and without the contribution of gas hardening. As can be seen, such a binary in a gas-free environment will be considered a soft binary, i.e. its semimajor axis will grow, until finally disrupted. In a gas-rich environment, this binary is essentially hard. The semimajor axis decreases with time, until it reaches a final semimajor axis, as the gas decays and GDF is not efficient anymore.

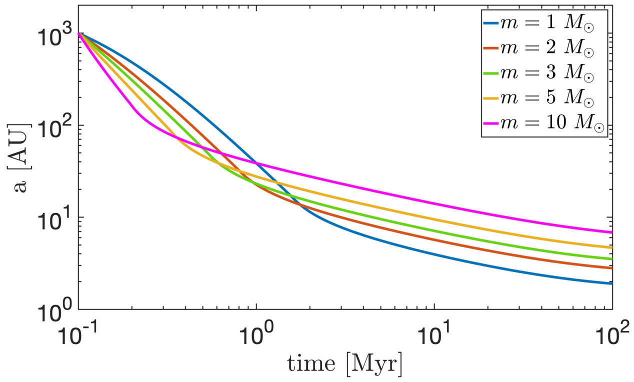

Figure 3. The evolution of binaries with different masses $m = m_1 = m_2$, under the effect of gas hardening and hardening/softening.

In Fig. 3, we present the semimajor axis evolution for different masses. The mass dependence is dictated by stellar softening, and scales as M_{bin}^{-1} , for hardening there is a weak dependence on the mass scales as $(m_1 + m_2 + \bar{m})/(m_1 + m_2)$ and for gas-hardening it scales as $m^{-1/2}$, assuming the relative velocity scales as the Keplerian one and as $\sqrt{m}$ when the relative velocity is taken as the velocity dispersion. For initially large separations, in the regime where gas hardening dominates, higher masses harden more efficiently. Then, at some point, the binary hardens to the hard-soft binary separation, in parallel to the gas decay, and from this point stellar-hardening dominates.

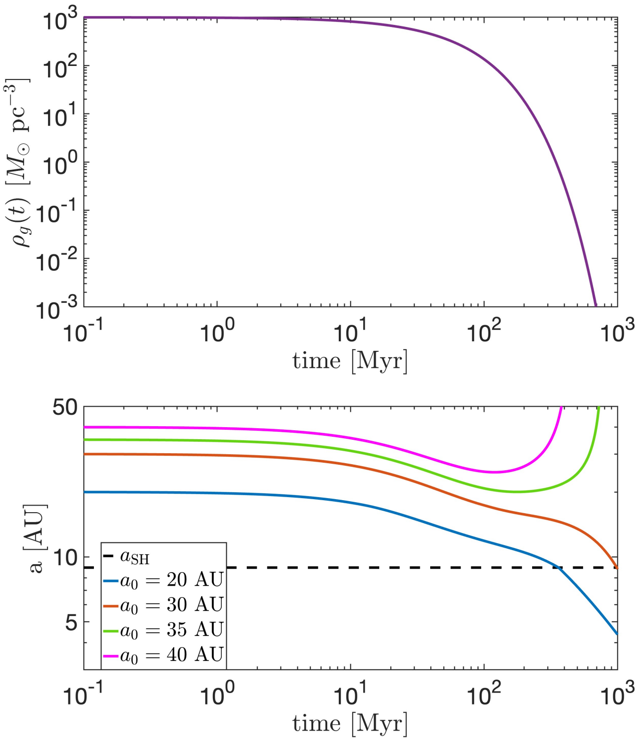

Figure 4. Upper panel: the gas density evolution over time, for initial gas density of $\rho_{g,0} = 10^3\ M_\odot\ \mathrm{pc}^{-3}$. Lower panel: The evolution of different initial semimajor axes for a background density with $\rho_{g,0} = 10^3\ M_\odot\ \mathrm{pc}^{-3}$, for a binary with masses $m = m_1 = m_2 = 1\ M_\odot$. The dashed line corresponds to the soft-hard limit boundary as derived based on stellar interactions only.

In Fig. 4, we present in solid lines the semimajor axis evolution of binaries with masses $m = m_1 = m_2 = 1\ M_\odot$ and initial background density of $\rho_{g,0} = 10^3\ M_\odot\ \mathrm{pc}^{-3}$. The dashed line corresponds to the soft-hard limit boundary as derived based on stellar interactions only. As can be seen, the qualitative behavior is different for different initial conditions. While for small semimajor axes, the semimajor axis is monotonically decreasing, for larger semimajor which are above the soft-hard limit, we see that after some decrement, the gas depletes to the point that softening by stellar encounters dominates, and the binary will eventually still be soft, even after the gas-hardening stage.

To assess the effect of gas hardening on the soft/hard binary population, we carried out a Monte-Carlo simulation, sampling binaries with initial semimajor axes from a log-normal distribution (Moe & Di Stefano 2017) between 1 AU to 200 AU, and letting them dynamically evolve. We then examined the semimajor axes distribution of the binaries at different times.

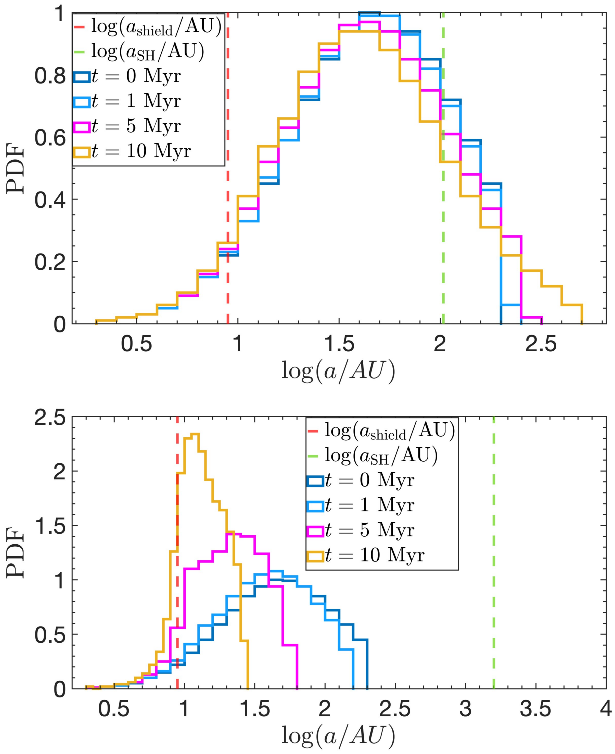

Figure 5. The results of a Monte Carlo simulation, with two equal masses $m = m_1 = m_2 = 1\ M_\odot$ averaged over 1000 iterations. The red dashed line corresponds to the critical semimajor axis between soft and hard binaries when considering stellar interactions only (without the effect of gas), and the green dashed line corresponds to the shielding radius as calculated relative to the *initial* gas density. The different solid lines correspond to different times. Upper panel: Initial background density of $\rho_{g,0} = 10^3\ M_\odot\ \mathrm{pc}^{-3}$. Lower panel: $\rho_{g,0} = 10^4\ M_\odot\ \mathrm{pc}^{-3}$. Note the different scales.

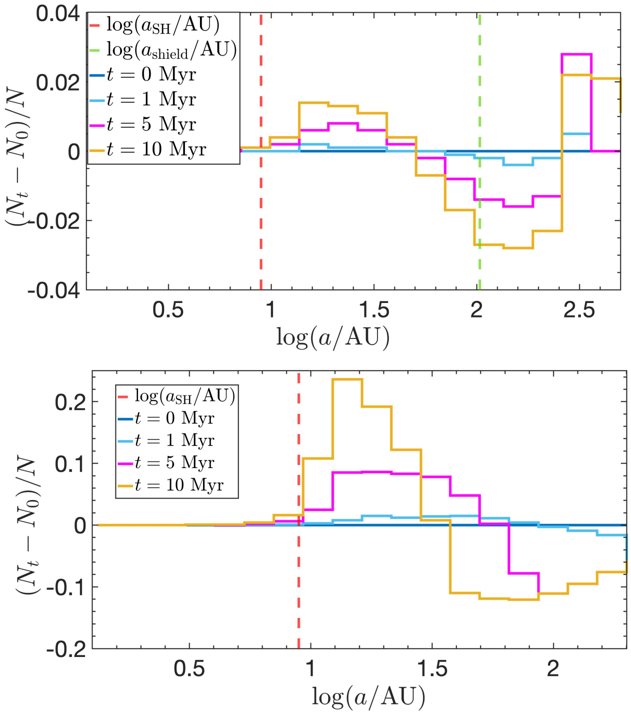

Figure 6. The normalized difference between the number of binaries in different separations, based on the results of our Monte Carlo simulation, with two equal masses $m = m_1 = m_2 = 1\ M_\odot$ averaged over 1000 iterations. The red dashed line corresponds to the critical semimajor axis between soft and hard binaries when considering stellar interactions only (without the effect of gas), and the green dashed line corresponds to the shielding radius as calculated relative to the *initial* gas density. The different solid lines correspond to different times. Upper panel: Initial background density of $\rho_{g,0} = 10^3\ M_\odot\ {\rm pc}^{-3}$. Lower panel: $\rho_{g,0} = 10^4\ M_\odot\ {\rm pc}^{-3}$. Note the different scales.

Gas shielding changes the distribution of orbital parameters of binaries, as well as their fractions. In gas-free environments, soft binaries are prone to frequent disruptions, that could be evaded in gas-rich environments, since gas hardening could dominate over the stellar softening, under appropriate conditions, leading to hardening instead of softening, as demonstrated earlier. This process could leave significant signatures on the population of binaries. To assess the statistical properties of soft/hard binaries in gas-rich environments, we carry out a Monte-Carlo simulation. For low-mass, FGK stars we follow Raghavan et al. (2010); Moe & Di Stefano (2017) findings and consider a log-normal distribution centered around 50 AU with a dispersion of $\sigma = 1 = \log(10AU/AU)$.

We draw binaries with separations between $1-200$ AU, equal mass binaries with $m = m_1 = m_2 = 1\ M_\odot$ and initial gas background density of $\rho_{g,0} = 10^3\ M_\odot\ {\rm pc}^{-3}$ and $\rho_{g,0} = 10^4\ M_\odot\ {\rm pc}^{-3}$ for the upper and lower subfigures correspondingly.

In Figs. 5,6, we present the semimajor axis distribution at different times. As expected, binaries with semimajor axes smaller than the shielding radius (which decreases with time, due to gas depletion) tend to decrease their semimajor axes, even if they are termed as soft binaries when considering stellar interactions only. Hence, the initial binary population is redistributed, such that binaries that are considered hard relative to the combined effect of gas and stars harden, and binaries considered soft, soften, and will eventually be disrupted. It should be noted that we didn't consider here replenishment of binaries and aborted the simulation when a separation of the tidal radius was reached. In a more realistic scenario, further binaries could be formed, either by gas-assisted binary formation, wide binary capture, or any other formation mechanism, adding further complications to our model. As expected, increasing the gas density enhances the effect and could lead in extreme cases to converting all the binaries to be hard. In these cases, gas-assisted binary formation which scales with the gas density, is expected to be highly efficient, potentially leading to the refilling of soft binaries population.

4. DISCUSSION

Gas shielding could lead to several important implications on the dynamics of binaries and their properties as a population. Here we will discuss our results and possible implications.

4.1. *Implications for other gas-rich environments*

Although we described here the effect of gas shielding in gas-rich clusters only, it is a general process that is expected to take place in any other gas-rich environment, e.g. affecting binary stars in AGN disks and star-forming regions, or binary-planetesimals in protoplanetary disks (with the necessary modifications). It should be noted that in gaseous disks, such as AGN or protoplanetary disks, the effect of shearing could modify the results significantly. In protoplanetary disks, dynamical friction could dominate over the effect of gas dynamical friction, depending on the evolutionary stage of the disk. In star-forming regions, the gaseous epoch is expected to be shorter, and gas shielding will apply mainly for stars rather than compact objects, which are likely not to have formed yet.

Our results could be used also to set constraints on the gas abundances and lifetimes in these environments, as we could derive expected hard/soft binary fractions for a given amount of gas.

4.1.1. *Star-forming regions*

In the following, we also demonstrate the process for star-forming regions. Star formation takes place in gas-rich environments, starting from molecular clouds that later collapse to prestellar cores and finally pre-main sequence stars (see a detailed review in Bergin & Tafalla 2007). The typical gas mass of the clouds is $\sim 10^6\ M_\odot$, within a radius of ~ 10 pc , which dictates a typical gas density of $10^3\ M_\odot\ \mathrm{pc}^{-3}$. The velocity dispersion (and the sound speed) in these regions is ~ 2 km/sec and the typical gas lifetime is ~ 2 Myr (Bergin & Tafalla 2007; Goodwin 2013). Assuming we have a similar density of gas and stars, we consider $\rho_\star = 10^3\ M_\odot\ \mathrm{pc}^{-3}$, and $n_\star = 10^3\ \mathrm{pc}^{-3}$.

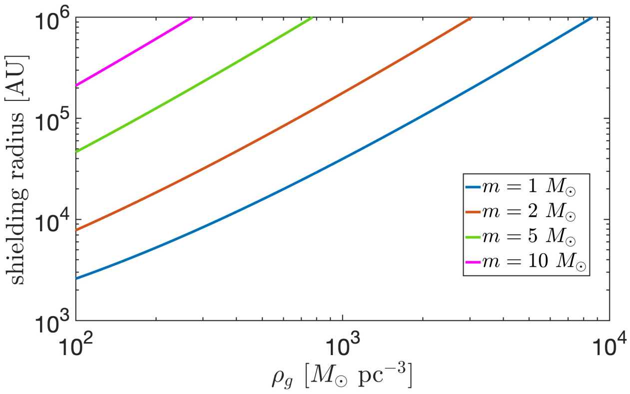

Figure 7. The shielding radius – the maximal radius for which a binary is considered hard when accounting also for the effect from gas hardening – as a function of different gas densities, for different binaries of equal masses $m = m_1 = m_2$, for the parameters of clouds of star-forming regions.

In Fig. 7, we present the shielding radius for the characteristic parameters of star-forming regions. It can be seen that the shielding radii are typically very large, allowing the formed stars to harden instead of soften and get disrupted. Therefore, in these environments, the vast majority if not all the formed binaries are likely to at least partially harden due to a significant contribution from gas hardening.

4.2. *Gravitational waves and other aspects of binary evolution*

Gas shielding naturally leads to an increment in the population of hard binaries, and the survivability of soft binaries. Hence, in gas-rich environments, the occurrence rate of events that require a hard binary could be elevated. In particular, the rates of binary-single and binary-binary encounters giving rise to close approaches, tidal interactions, GW inspirals and collisions will be increased, both due to the larger number of binaries, as well as the existence of more compact binaries; these should affect the rates of GW mergers, thermonuclear explosions from mergers of white dwarfs and other stellar merger, and mass-transfer interacting binaries products.

4.3. *Hard binary formation*

Most of the three-body encounters that dynamically form binaries produce wide soft binaries that quickly soften and disrupt (Aarseth & Heggie 1976; Atallah et al. 2024; Ginat & Perets 2024). Wide binaries could form also by captures (Kouwenhoven et al. 2010; Rozner & Perets 2023). In particular, modeling of soft binaries' steady-state formation and disruption suggests the existence of the order of a single soft binary in a cluster at any given time. However, this steady-state estimate does not account for the existence of gas. Once the gas is included soft binaries that form dynamically can be gas-shielded and become hard-binaries. In other words, the dynamical formation of binaries which is typically inefficient could become a major production form of binaries. The typical formation rate of hard binaries via gas shielding could be estimated by

$$\Gamma_{\mathrm{hard}}(a) \sim \frac{N_{\mathrm{soft(a)}}}{\tau_{\mathrm{shield}}} \approx \qquad (15)$$
$$\approx 0.12\mathrm{Myr}^{-1}\mathrm{N}_{\mathrm{soft}}(\mathrm{a}) \left(\frac{\mathrm{m}}{\mathrm{M}_\odot}\right)^{1/2} \left(\frac{\mathrm{a}}{10^3\mathrm{AU}}\right)^{1/2} \left(\frac{\rho_{\mathrm{g}}}{10^4\mathrm{M}_\odot\mathrm{pc}^{-3}}\right)$$

where N_{soft} is the number of formed soft binaries and τ_{shield} is the gas shielding timescale, given by $\tau_{\mathrm{shield}} = |a/\dot{a}_{\mathrm{GDF}}|$, where we considered two equal masses $m = m_1 = m_2 = 1\ M_\odot$, and since the binary is wide $v_{\mathrm{rel}} = \sigma = 10$ km/sec.

In addition, gas-assisted binary capture (Tagawa et al. 2020; Rozner et al. 2023; Rowan et al. 2023; Whitehead et al. 2023; DeLaurentiis et al. 2023) could take place in gas-rich environments, initially leading to the formation of soft binaries that will possibly refill the parameter space of hardened soft binaries.

4.4. *Binary distributions following the gas dispersal*

The gas-rich epoch of the second-generation star formation is truncated by the explosion of supernovae, that clear the cluster from gas. After that, the distribution achieved from the combined stellar-gas dynamics enters a freezout phase, from which another, different distribution of soft binaries could be obtained. In the absence of gas, and under the assumption of thermal equilibrium we derived analytically the separation distribution of soft binaries in clusters and showed that it obeys a powerlaw rule (Rozner & Perets 2023). However, after the gaseous phase, even if the distribution of binaries finally reaches equilibrium, the total number of hard binaries will grow, and some soft binaries could be shielded from disruption and remain soft by the dissipation of gas, while others will become hard. Moreover, there could be global effects in the cluster, such as enhanced mass segregation, as all the stars could be thought of as in a binary with the cluster potential, leading in turn to enhanced stellar scattering and further stellar interactions.

While we explored some of the effects of gas shielding on the distribution of soft binaries, the long-term effects of producing different binary populations than typically assumed should be further explored by a detailed population synthesis study, accounting for these effects.

4.5. *Caveats & future directions*

- Here we modeled the motion in gas using GDF. Qualitatively similar results should be obtained, in principle, using other models, substituting a different gas hardening law in our energy dissipation calculation.

- The gas density profile of a globular cluster during the second-generation star formation is currently unknown. We chose a typical density that corresponds to the total mass of second-generation stars enclosed by a typical volume of the core times an order of unity efficiency factor, but the actual density could vary.

- We considered only circular binaries. However, in a more realistic calculation, one should consider also eccentric binaries, for which gas hardening is even more efficient, assuming a flat gas density profile (e.g. Rozner & Perets 2022).

- We considered an initial log-normal distribution for the binary separation, which corresponds to the observed current distribution of FGK binaries (Moe & Di Stefano 2017). However, in the early stages of the cluster, there could be a significant contribution from binaries that were captured via gas-assisted binary capture, with a preference towards large separations comparable to the Hill radius (Rozner et al. 2023). These binaries will enrich the soft binary population. In addition, the distribution of more massive stars tends towards a log-uniform distribution, and would therefore give rise to initially harder binaries, on average. Stellar evolution, not considered here would also affect the distribution at late times; in particular compact-object binaries would not follow the field star binary distributions.
- We ignored gas accretion on the binaries. As the mass will increase due to accretion, the total effect of gas hardening should be strengthened when considering this effect as well.

5. SUMMARY

In this paper, we discussed the effect of gas hardening on soft/hard binaries and studied both particular examples and the total effect on the population. We defined the shielding radius of a binary as the largest separation in which a binary is hard relative to the joint contribution of stellar interaction and interaction with the gas. We showed this radius could be significantly larger than the 'standard' soft-hard boundary when considering stellar interactions only, and hence the effect of gas-hardening is significant and could revise the dynamics of binaries in gas-rich environments, as we demonstrated for globular clusters during their second-generation star formation.

While the interaction of binaries with gas has been extensively studied during the last few years, mainly in the context of AGN disks, there are still various unexplored directions. The dynamics of populations in gas are essentially different than the ones in gas-free regions, and the distributions of binary populations change accordingly.

Finally, we focused on binary shielding in gas-rich GCs, but the phenomenon applies in general for every gas-rich environment and hence should leave important signatures on the binary population in these environments, such as AGN disks and star-forming regions.

ACKNOWLEDGEMENTS

We thank Barry Ginat and Evgeni Grishin for fruitful discussions. MR acknowledges the generous support of Azrieli fellowship. MR and HBP acknowledge support from the European Union's Horizon 2020 research and innovation program under grant agreement No 865932-ERC-SNeX.

REFERENCES

Aarseth, S. J., & Heggie, D. C. 1976, A&A, 53, 259

Antoni, A., MacLeod, M., & Ramirez-Ruiz, E. 2019, ApJ, 884, 22, doi: 10.3847/1538-4357/ab3466

Artymowicz, P., Clarke, C. J., Lubow, S. H., & Pringle, J. E. 1991, ApJL, 370, L35, doi: 10.1086/185971

Artymowicz, P., Lin, D. N. C., & Wampler, E. J. 1993, ApJ, 409, 592, doi: 10.1086/172690

Atallah, D., Weatherford, N. C., Trani, A. A., & Rasio, F. 2024, arXiv e-prints, arXiv:2402.12429, doi: 10.48550/arXiv.2402.12429

Bahcall, J. N., & Ostriker, J. P. 1976, Nature, 262, 37, doi: 10.1038/262037a0

Bastian, N., & Lardo, C. 2018, ARA&A, 56, 83, doi: 10.1146/annurev-astro-081817-051839

Bergin, E. A., & Tafalla, M. 2007, ARA&A, 45, 339, doi: 10.1146/annurev.astro.45.071206.100404

Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition

Bonnell, I. A. 2001, in The Formation of Binary Stars, ed. H. Zinnecker & R. Mathieu, Vol. 200, 23

Carretta, E., Bragaglia, A., Gratton, R. G., et al. 2009, A&A, 505, 117, doi: 10.1051/0004-6361/200912096

Celoria, M., Oliveri, R., Sesana, A., & Mapelli, M. 2018, arXiv e-prints, arXiv:1807.11489, doi: 10.48550/arXiv.1807.11489

Clarke, C. J., & Pringle, J. E. 1991, MNRAS, 249, 584, doi: 10.1093/mnras/249.4.584

DeLaurentiis, S., Epstein-Martin, M., & Haiman, Z. 2023, in AAS/High Energy Astrophysics Division, Vol. 55, AAS/High Energy Astrophysics Division, 100.30

Duffell, P. C., Dittmann, A. J., D'Orazio, D. J., et al. 2024, arXiv e-prints, arXiv:2402.13039, doi: 10.48550/arXiv.2402.13039

Ginat, Y. B., & Perets, H. B. 2021a, MNRAS, 508, 190, doi: 10.1093/mnras/stab2565

—. 2021b, Physical Review X, 11, 031020, doi: 10.1103/PhysRevX.11.031020

—. 2024, forthcoming

Goldreich, P., Lithwick, Y., & Sari, R. 2002, Nature, 420, 643, doi: 10.1038/nature01227

Goodman, J., & Hut, P. 1993, ApJ, 403, 271, doi: 10.1086/172200

Goodwin, S. 2013, Star Formation, ed. T. D. Oswalt & M. A. Barstow (Dordrecht: Springer Netherlands), 243–277, doi: 10.1007/978-94-007-5615-1_5

Gratton, R., Bragaglia, A., Carretta, E., et al. 2019, A&A Rv, 27, 8, doi: 10.1007/s00159-019-0119-3

Grishin, E., Gilbaum, S., & Stone, N. C. 2023, arXiv e-prints, arXiv:2307.07546, doi: 10.48550/arXiv.2307.07546

Grishin, E., & Perets, H. B. 2016, ApJ, 820, 106, doi: 10.3847/0004-637X/820/2/106

Hall, S. M., Clarke, C. J., & Pringle, J. E. 1996, MNRAS, 278, 303, doi: 10.1093/mnras/278.2.303

Heggie, D. C. 1975, MNRAS, 173, 729, doi: 10.1093/mnras/173.3.729

Heggie, D. C., & Hut, P. 1993, ApJS, 85, 347, doi: 10.1086/191768

Hills, J. G. 1975, AJ, 80, 1075, doi: 10.1086/111842

Kouwenhoven, M. B. N., Goodwin, S. P., Parker, R. J., et al. 2010, MNRAS, 404, 1835, doi: 10.1111/j.1365-2966.2010.16399.x

Kratter, K. M. 2011, in Astronomical Society of the Pacific Conference Series, Vol. 447, Evolution of Compact Binaries, ed. L. Schmidtobreick, M. R. Schreiber, & C. Tappert, 47, doi: 10.48550/arXiv.1109.3740

Kritos, K., & Cholis, I. 2020, PhRvD, 102, 083016, doi: 10.1103/PhysRevD.102.083016

Lai, D., & Muñoz, D. J. 2023, ARA&A, 61, 517, doi: 10.1146/annurev-astro-052622-022933

Li, J., Dempsey, A. M., Li, H., Lai, D., & Li, S. 2023, ApJL, 944, L42, doi: 10.3847/2041-8213/acb934

Li, R., & Lai, D. 2024, MNRAS, 529, 348, doi: 10.1093/mnras/stae504

Mansbach, P. 1970, ApJ, 160, 135, doi: 10.1086/150412

McKernan, B., Ford, K. E. S., Lyra, W., & Perets, H. B. 2012, MNRAS, 425, 460, doi: 10.1111/j.1365-2966.2012.21486.x

Moe, M., & Di Stefano, R. 2017, ApJS, 230, 15, doi: 10.3847/1538-4365/aa6fb6

Moody, M. S. L., Shi, J.-M., & Stone, J. M. 2019, ApJ, 875, 66, doi: 10.3847/1538-4357/ab09ee

Ormel, C. W., & Klahr, H. H. 2010, A&A, 520, A43, doi: 10.1051/0004-6361/201014903

Ostriker, E. C. 1999, ApJ, 513, 252, doi: 10.1086/306858

Ostriker, J. P. 1983, ApJ, 273, 99, doi: 10.1086/161351

Perets, H. B., & Murray-Clay, R. A. 2011, ApJ, 733, 56, doi: 10.1088/0004-637X/733/1/56

Press, W. H., & Teukolsky, S. A. 1977, ApJ, 213, 183, doi: 10.1086/155143

Quinlan, G. D. 1996, NewA, 1, 35, doi: 10.1016/S1384-1076(96)00003-6

Raghavan, D., McAlister, H. A., Henry, T. J., et al. 2010, ApJS, 190, 1, doi: 10.1088/0067-0049/190/1/1

Renzini, A., D'Antona, F., Cassisi, S., et al. 2015, MNRAS, 454, 4197, doi: 10.1093/mnras/stv2268

Roupas, Z., & Kazanas, D. 2019, A&A, 632, L8, doi: 10.1051/0004-6361/201937002

Rowan, C., Boekholt, T., Kocsis, B., & Haiman, Z. 2023, MNRAS, 524, 2770, doi: 10.1093/mnras/stad1926

Rozner, M., Generozov, A., & Perets, H. B. 2023, MNRAS, 521, 866, doi: 10.1093/mnras/stad603

Rozner, M., & Perets, H. B. 2022, ApJ, 931, 149, doi: 10.3847/1538-4357/ac6d55

—. 2023, ApJ, 955, 134, doi: 10.3847/1538-4357/ace2c6

Spitzer, L. 1987, Dynamical evolution of globular clusters

Stahler, S. W. 2010, MNRAS, 402, 1758, doi: 10.1111/j.1365-2966.2009.15994.x

Stone, N. C., Metzger, B. D., & Haiman, Z. 2017, MNRAS, 464, 946, doi: 10.1093/mnras/stw2260

Tagawa, H., Haiman, Z., & Kocsis, B. 2020, ApJ, 898, 25, doi: 10.3847/1538-4357/ab9b8c

Whitehead, H., Rowan, C., Boekholt, T., & Kocsis, B. 2023, arXiv e-prints, arXiv:2309.11561, doi: 10.48550/arXiv.2309.11561

Chapter 6

The distribution of soft binaries in clusters and wide binaries in the field

Based on Rozner and Perets (2023)

Stellar systems could not achieve in general thermal equilibrium. Nevertheless, soft binaries are so weakly bound that they could approach equilibrium on faster timescales and hence could in principle be treated as being in equilibrium (Goodman and Hut, 1993; Binney and Tremaine, 2008). Here we will derive analytically the properties of soft binaries in clusters and wide binaries in the field. In our calculation, we define and characterize the available parameter space for wide binaries and investigate its properties.

To form a wide binary, two conditions must be satisfied: (I) the binary separation should be sufficiently small such that the binary should survive the tidal radius of the cluster (for soft binaries in clusters) or the Galactic tidal radius (for wide binaries in the field), and (II) the two components should become bound, i.e., their relative velocity should be smaller than the escape velocity at their instantaneous distance.

Consider a cluster with N stars with masses $\{m_i\}_{i=1}^{N}$, contained in a volume V. The temperature of the cluster is defined by $\beta^{-1} := k_B T = \bar{m}\sigma^2$, where $\bar{m}$ is the mean mass in the cluster and σ is its velocity dispersion. The distribution of a binary constitutes stars of masses m_1 and m_2, in thermal equilibrium, is given by Boltzmann distribution, i.e. $f_{1,2}(\mathbf{w_1}, \mathbf{w_2}) \propto \mathbf{e}^{-\beta\mathcal{H}}$ where $\mathbf{w_i}$ are the coordinates in the 6-dimensional phase space and $\mathcal{H}$ is the two-body corresponding Hamiltonian,

$$\mathcal{H} = \frac{1}{2}m_1 v_1^2 + \frac{1}{2}m_2 v_2^2 - \frac{Gm_1 m_2}{|x_1 - x_2|} \tag{6.1}$$

The number density of binaries with internal energy $\tilde{E}$ is then given by

$$n_{\rm eq}^{1,2}(\tilde{E}) = \frac{N_1 N_2}{V} \int f_{1,2}(w_1, w_2)\delta\left(\tilde{E}(w_1, w_2) - \tilde{E}\right) d^6 w_1 d^6 w_2 \tag{6.2}$$

where N_i is the number of stars from the i-th species. The total number of binaries is

$$n_{\rm eq}^{1,2}(\tilde{E}) = \frac{G^3 \rho_1 \rho_2 \pi^{3/2} (m_1 m_2)^2}{8 \mu^{3/2} \sigma^3 |\tilde{E}|^{5/2}} \left(\frac{\bar{m}}{M} \right)^{3/2} e^{|\tilde{E}|/\bar{m}\sigma^2} \tag{6.3}$$

where $\rho_i = n_i m_i$ are the densities of the different species $\mu_{12} = m_1 m_2/(m_1+m_2)$ is the reduced mass of a binary and $M = m_1 + m_2$ is the total mass of the binary. It should be noted that here we consider a uniform stellar density, as a simplifying assumption. However, in a more realistic model of clusters, other density profiles should be taken into consideration, which could modify the outcomes.

The number density for a given semi-major axis a is

$$n_{\rm eq}^{1,2}(a) = \frac{G^{3/2} n_1 n_2 \pi^{3/2} \bar{m}^{3/2}}{2^{3/2} \sigma^3} a^{1/2} e^{|\tilde{E}(a)|/\bar{m}\sigma^2} \tag{6.4}$$

The maximal energy for a soft binary is determined by the transition between hard and soft binaries and given by $\bar{m}\sigma^2$. The upper limit for the separation is the Hill radius defined by $R_{\rm Hill,c} = \left((m_1 + m_2)/M_{\rm cluster}\right)^{1/3} R_{\rm cluster}$ where $M_{\rm cluster}$ and $R_{\rm cluster}$ are the mass and radius of the cluster correspondingly. The galactic tidal field sets a larger cutoff separation, given by $R_{\rm Hill,g} \sim 1.7((m_1 + m_2)/2M_\odot)$ pc (Jiang and Tremaine, 2010).

We can integrate eq. 6.3 to obtain the total number density of soft binaries taking into account the contributions from the whole energy range. For simplicity, we approximate the exponential as unity.

We assume a uniform density of soft binaries inside the cluster, and approximate $N_{\rm soft}^{1,2} \approx n_{\rm eq}^{1,2} V$ and $\rho_i = m_i N_i / V$.

$$N_{\text{soft}}^{1,2} \approx R_{\text{cluster}}^{3} n_{\text{eq}}^{1,2} \approx s3d. \frac{N_1 N_2 \pi^{3/2}}{12 M_{\text{cl}}^{2}} \bar{m}^{3/2} M^{1/2} - \frac{N_1 N_2 \pi^{3/2}}{12} \left(\frac{m_1 m_2}{M_{\text{cl}}^{2}} \right)^{3/2} \tag{6.5}$$

For any given lower energy cutoff $|\tilde{E}_{\text{min}}|$, or equivalently a maximal separation R_{cut}, the number of binaries will be given by

$$N_{\text{bin}}^{1,2}(R_{\text{cut}}) \approx \frac{(2\pi)^{3/2} G^{3/2} n_1 n_2 \bar{m}^{3/2}}{12\sigma^3} R_{\text{cluster}}^{3} R_{\text{cut}}^{3/2} - \frac{N_1 N_2 \pi^{3/2}}{12} \left(\frac{m_1 m_2}{M_{\text{cl}}^{2}} \right)^{3/2} \tag{6.6}$$

Where for soft binaries, $R_{\text{cut}} = R_{\text{Hill}}$. For a general mass distribution $\xi(m)$,

$$N_{\text{bin}}^{\text{total}}(R_{\text{cut}}) = \int N_{\text{bin}}^{1,2}(R_{\text{cut}}) \xi(m_1) \xi(m_2) dm_1 dm_2 \tag{6.7}$$

For single mass clusters, it can be seen that the number of existing systems at any given point is of order unity, as expected.

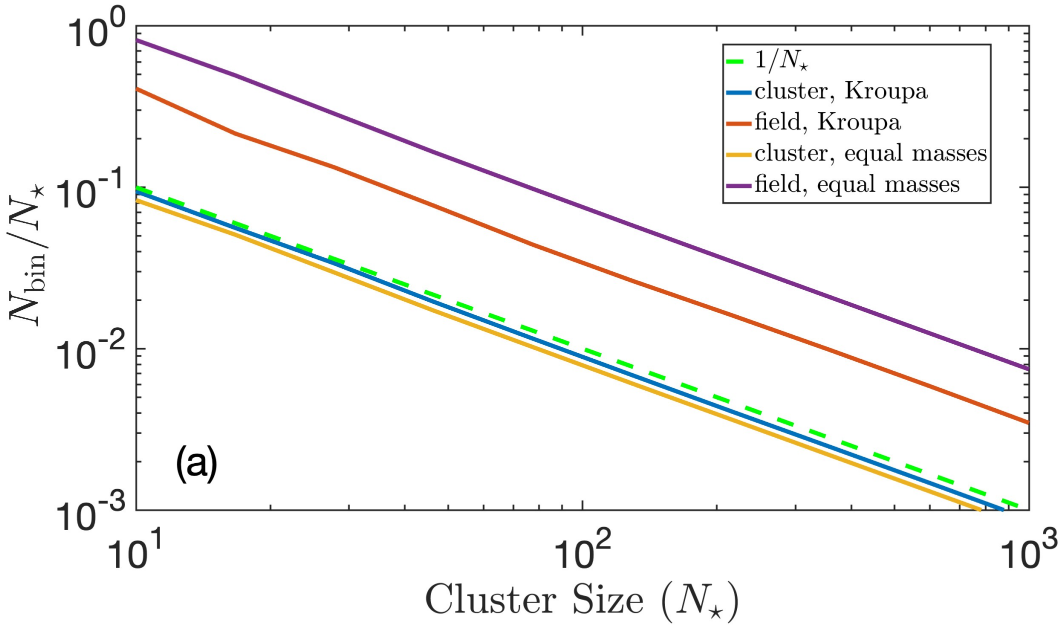

Figure 6.1: The fraction of soft binaries in clusters/wide binaries in the field for different choices of mass distribution, as derived from the Monte-Carlo simulation, is based on our analytical derivation. (averaged over 100 Monte-Carlo realizations).

In Fig. 6.1 we present the calculated fractions from our analytical model. The scaling of $1/N_\star$ is conserved for all the constellations we checked, up to a correction of multiplactory factor. The fraction of binaries in the field is higher than the fraction in clusters, as there is a wider range allowed for binary separations in the field. This dependence agrees with the freezeout distribution found in Moeckel and Clarke (2011), although Perets and Kouwenhoven (2012) found a shallower distribution, that might arise from different initial conditions, see a detailed discussion in Rozner and Perets (2023).

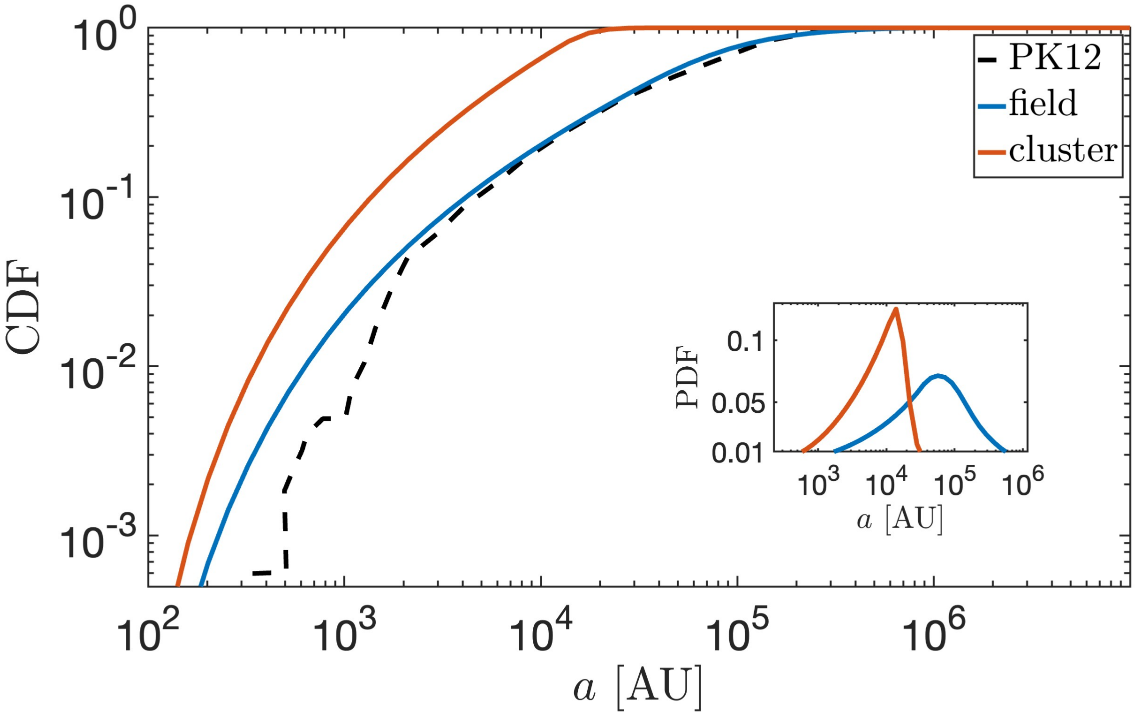

Figure 6.2: The semi-major axis distribution, as derived from the Monte-Carlo simulation, based on our analytical derivation, averaged over 1000 realizations, for a cluster containing both stars and planets ($N_\star = N_p$), in comparison to results from N-body simulations.

In Fig. 6.2, we introduce the semimajor axis distribution, as derived using our analytical model, vs. the numerical results from Perets and Kouwenhoven (2012). We found a good agreement between the two, apart from low semimajor axes, for which probably small statistical effects are significant. It could be seen that wide binaries in the field peak on higher semimajor axes, as expected since they are constrained by the Galactic tide rather than the Hill radius of the cluster.

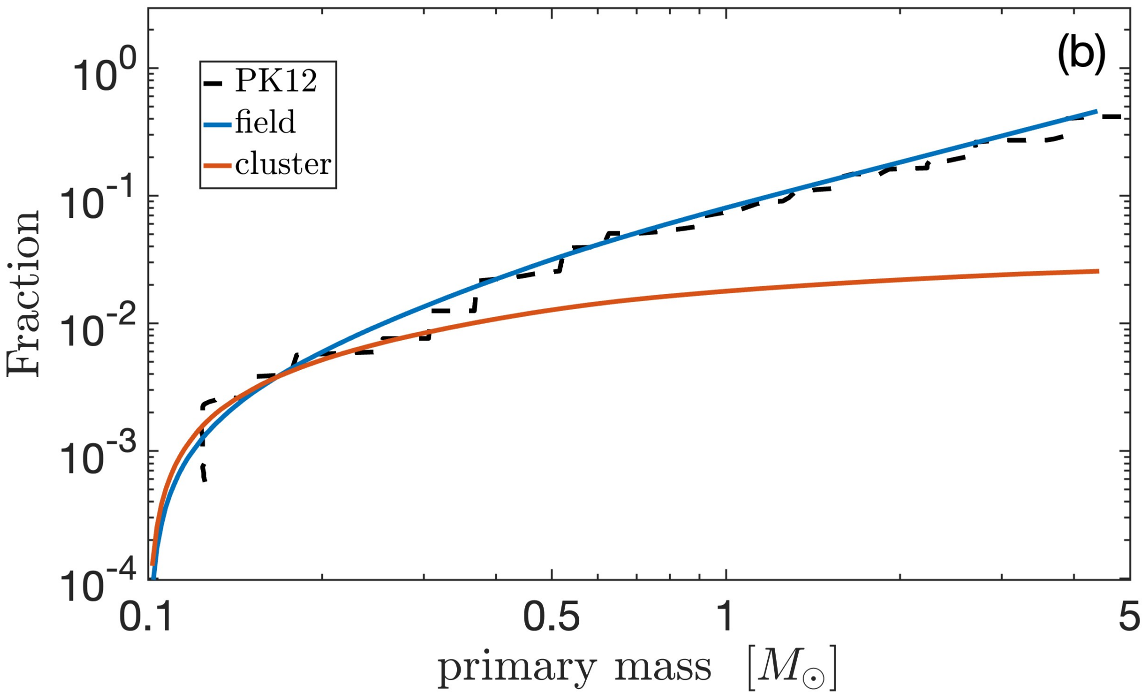

Figure 6.3: The binary fraction as a function of the primary mass, for different masses, as derived from an averaging over 1000 runs of the Monte-Carlo simulation, for a cluster containing 100 stars, drawn from a Kroupa mass function, and an equal number of planets, in comparison to N-body simulations.

In Fig. 6.3, we introduce the fraction of wide/soft binaries as a function of the primary mass. This fraction is a monotonically increasing function of the primary mass, with a steeper slope for field binaries. Also here, we have an excellent agreement with the N-body simulations in Perets and Kouwenhoven (2012). The vast majority of massive objects in the field have a wide companion, which stresses the importance of studying wide binaries.

To summarize, in this project we investigated the properties of wide binaries in the field and soft binaries in clusters in statistical mechanics tools. We found a good agreement between the N-body simulations and our analytical model, which will enable in the future to save expensive computational time.

Draft version June 29, 2023
Typeset using LaTeX **twocolumn** style in AASTeX631

Born to be wide: the distribution of wide binaries in the field and soft binaries in clusters

Mor Rozner[1] and Hagai B. Perets[1, 2]

[1] *Technion - Israel Institute of Technology, Haifa, 3200002, Israel*
[2] *Department of Natural Sciences, The Open University of Israel, 1 University Road, PO Box 808, Raanana 4353701, Israel*

ABSTRACT

Most stars, binaries, and higher multiplicity systems are thought to form in stellar clusters and associations, which later dissociate. Very wide binaries can be easily disrupted in clusters due to dynamical evaporation (soft binaries) and/or due to tidal disruption by the gravitational potential of the cluster. Nevertheless, wide binaries are quite frequent in the field, where they can sometimes play a key role in the formation of compact binaries, and serve as tools to study key physical processes. Here we use *analytic* tools to study the dynamical formation of soft binaries in clusters, and their survival as field binaries following cluster dispersion. We derive the expected properties of very wide binaries both in clusters and in the field. We analytically derive their detailed distributions, including wide-binary fraction as a function of mass in different cluster environments, binaries mass functions and mass ratios, and the distribution of their orbital properties. We show that our calculations agree well on most aspects with the results of N-body simulations, but show some different binary-fraction dependence on the cluster mass. We find that the overall fraction of wide binaries scales as $\propto N_\star^{-1}$ where $N_\star$ is the size of the cluster, even for non-equal mass stars. More massive stars are more likely to capture wide companions, with most stars above five solar mass likely to capture at least one stellar companion, and triples formation is found to be frequent.

1. INTRODUCTION

Binary and higher multiplicity stellar systems are quite frequent (e.g. Duquennoy & Mayor 1991; Raghavan et al. 2010; Sana et al. 2012; Duchêne & Kraus 2013; Moe & Di Stefano 2017), and play a key role in the dynamics and evolution of stellar systems. Most stars, even field stars, are thought to have formed in stellar clusters and associations (Lada & Lada 2003), and later released to the field once their host clusters dispersed. The early dynamics of binaries can therefore be significantly altered by their interactions with other stars, and/or due to the overall potential of their host clusters. In dense environments, binaries could be divided into two groups, based on the energy relative to the mean energy of their background: dynamically-hard binaries ($|E_{\rm bin}| \gtrsim \bar{m}\sigma^2$) and dynamically-soft binaries ($|E_{\rm bin}| \lesssim \bar{m}\sigma^2$). The evolution of soft and hard binaries differ qualitatively in such environments. While hard binaries become, on average, harder, due to interactions with other stars (more compact, shorter periods, or more general i.e. larger absolute binding energy, accounting for exchanges), soft binaries become softer (Heggie 1975; Hills 1975). It should be noted that hard binaries could go through exchanges when they encounter a third perturber. See a revised version, accounting for external cutoffs, of Heggie's law in Ginat & Perets (2021).

In the field, the low binding energy of wide binaries ($a \gtrsim 10^3$ AU) make them sensitive to even far flyby perturbations and other gravitational perturbations. This makes them an important tool to probe the galactic potential, MACHOs and primordial black-holes (e.g. Bahcall et al. 1985; Chanamé & Gould 2004; Quinn et al. 2009; Blas et al. 2017; Rozner et al. 2020). In addition, flyby and galactic tidal perturbations sometimes excite their eccentricities to extreme values, allowing the wide binary (and wide triple) components to closely interact through tidal, gravitational-wave of direct collisional interactions, giving rise to the formation of compact interacting binaries and/or merger products (Kaib & Raymond 2014; Michaely & Perets 2019, 2020; Grishin & Perets 2022; Michaely & Naoz 2022). Wide binaries could also constrain star formation (e.g. Larson 2001).

Given their important roles, it is important to understand the origins of wide binaries. Soft, wide binaries, can be dynamically formed in dense environments where perturbation by other stars can change

morozner@campus.technion.ac.il

their velocities, and give rise to random binding of two stars, with appropriate relative velocities and separations. It was suggested that the observed wide binaries in the field are surviving soft binaries that formed in the birth-cluster/association, after these clusters dispersed[1]. These are essentially the most recently formed soft binaries in the clusters which then survived as their host cluster dissolved/formed during the dissolution (Kouwenhoven et al. 2010; Moeckel & Bate 2010; Moeckel & Clarke 2011; Perets & Kouwenhoven 2012). Other formation channels are the formation of primordial wide binaries directly in the early star-formation phases in a gaseous environment, e.g. by fragmentation (e.g. Duchêne et al. 2004; Offner et al. 2010), gas-induced captures (Tagawa et al. 2020; Rozner et al. 2022; Xu et al. 2023) and dynamical unfolding of compact triples (Reipurth & Mikkola 2012). These originally formed wide binaries might not survive, due to perturbations, unless hardened by gas-induced inspiral, making them close binaries, or if they formed at late stages just before the cluster dispersal, at which point gas might not be available. Although the formation process of wide binaries is still unknown, chemical similarities were observed between binaries components, which indicate that the components of wide binaries were born together and were not formed due to random pairing (e.g. Andrews et al. 2018; Hawkins et al. 2020).

In the field, wide binaries may still experience infrequent flyby perturbations from field stars, and the widest ones could eventually be disrupted due to the tidal field of the Galaxy (Ambartsumian 1937; Chandrasekhar 1944; Yabushita 1966; Heggie 1975; King 1977; Heggie 1977; Retterer & King 1982; Bahcall et al. 1985; Jiang & Tremaine 2010), or potentially be excited to high eccentricities leading to strong peri-center interactions or collisions between the binary components due to Galactic tide induced secular evolution (Heisler & Tremaine 1986; Bonsor & Veras 2015; Hamilton & Rafikov 2019a,b).

Explaining the formation and survival of wide binaries is challenging, due to their wide separation that could in principle reach the size of the core of a young cluster, and their sensitivity to perturbations, which can destroy them.

Nevertheless, wide binaries are quite frequent e.g El-Badry & Rix 2018; El-Badry et al. 2019, 2021. Here we focus on the dynamical formation channel of soft binaries and their survival and provide the first analytic study of their detailed properties.

In this paper, we derive analytically the distributions of wide binaries in clusters and in the field, for general mass functions. We then compare our results with N-body simulations and observations.

In section 2, we derive analytically the distributions and the overall fractions of wide/soft binaries. In section 3 we present the results from our Monte Carlo simulations, based on the analytical derivations. In section 4 we discuss our results and future implications. In section 5, we summarize and conclude.

2. THE DISTRIBUTION OF SOFT/WIDE BINARIES

To enable the dynamical formation of a binary, two unbound stars need to be perturbed and change their relative velocities, to become bound. The conditions under which a bound binary is formed in this case are the following: (I) the binary separation should be sufficiently small such that the binary should survive the tidal radius of the cluster (for soft binaries in clusters) or the Galactic tidal radius (for wide binaries in the field). (II) The two components should become bound, i.e. their relative velocity should be smaller than the escape velocity at their instantaneous distance.

In our calculation, we define and characterize the available parameter space for wide binaries.

2.1. *Clusters in Virial equilibrium – soft binaries in clusters*

The distribution of energetically-soft binaries could be derived under the assumption of thermal equilibrium, using mechanical statistics methods, since their weak binding enables them to reach equilibrium faster than a whole stellar system, which in general has no maximum-entropy state (Goodman & Hut 1993; Binney & Tremaine 2008).

Consider a cluster with N stars with masses $\{m_i\}_{i=1}^N$, contained in a volume V. The temperature of the cluster is defined by $\beta^{-1} := k_B T = \bar{m}\sigma^2$, where $\bar{m}$ is the mean mass in the cluster and σ is its velocity dispersion. When planets are considered as well, the mass is still taken to be the mean mass of the stars, as the planets have a negligible effect on the cluster structure. Binaries with larger energies than the cluster temperature are hard binaries, while those with lower energies are soft.

The distribution of a binary constitutes stars of masses m_1 and m_2, in thermal equilibrium, is given by Boltzmann distribution, i.e. $f_{1,2}(\mathbf{w_1}, \mathbf{w_2}) \propto \mathbf{e}^{-\beta\mathcal{H}}$ where $\mathbf{w_i}$ are the coordinates in the 6-dimensional phase space and $\mathcal{H}$ is the two-body corresponding Hamiltonian,

---

[1] Other suggestions of dynamical formation in the field are unlikely (Goodman & Hut 1993).

$$\mathcal{H} = \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 - \frac{Gm_1m_2}{|x_1 - x_2|} \tag{1}$$

The number density of binaries with internal energy $\tilde{E}$ is then given by

$$n_{\rm eq}^{1,2}(\tilde{E}) = \frac{N_1N_2}{V}\int f_{1,2}(w_1,w_2)\delta\left(\tilde{E}(w_1,w_2) - \tilde{E}\right)d^6w_1d^6w_2 \tag{2}$$

where N_i is the number of stars from the i-th species. The total number of binaries is

$$n_{\rm eq}^{1,2}(\tilde{E}) = \frac{G^3\rho_1\rho_2\pi^{3/2}(m_1m_2)^2}{8\mu^{3/2}\sigma^3|\tilde{E}|^{5/2}}\left(\frac{\bar{m}}{M}\right)^{3/2}e^{|\tilde{E}|/\bar{m}\sigma^2} \tag{3}$$

where $\rho_i = n_im_i$ are the densities of the different species $\mu_{12} = m_1m_2/(m_1+m_2)$ is the reduced mass of a binary and $M = m_1 + m_2$ is the total mass of the binary. It should be noted that here we consider a uniform stellar density, as a simplifying assumption. However, in a more realistic model of clusters, other density profiles should be taken into consideration, which could modify the outcomes.

The number density for a given semi-major axis a is

$$n_{\rm eq}^{1,2}(a) = \frac{G^{3/2}n_1n_2\pi^{3/2}\bar{m}^{3/2}}{2^{3/2}\sigma^3}a^{1/2}e^{|\tilde{E}(a)|/\bar{m}\sigma^2} \tag{4}$$

For the simplified case of clusters composed of single-mass stars, there is an agreement with the analytical expressions derived in Goodman & Hut (1993); Binney & Tremaine (2008). The total number of soft/wide binaries could be derived via the integration of eq. 4, where the integral is dominated by the lower boundary, i.e. the minimal energy allowed for a soft/wide binary in the relevant context. While the separation of soft binaries in clusters could not exceed the Hill radius of the cluster, immediately after a cluster dispersion, leftover binaries from the cluster would become part of the field population and could survive with separations as large as the galactic tidal radius. Henceforth, we will split our analysis for soft binaries into two regimes: binaries in clusters and wide binaries in the field (i.e from dispersed clusters). Each regime dictates different regimes of integration and hence different distributions.

The maximal energy for a soft binary is determined by the transition between hard and soft binaries and given by $\bar{m}\sigma^2$. The upper limit for the separation is the Hill radius defined by $R_{\rm Hill,c} = \left((m_1+m_2)/M_{\rm cluster}\right)^{1/3}R_{\rm cluster}$ where $M_{\rm cluster}$ and $R_{\rm cluster}$ are the mass and radius of the cluster correspondingly. The galactic tidal field sets a larger cutoff separation, given by $R_{\rm Hill,g} \sim 1.7((m_1+m_2)/2M_\odot)$ pc (Jiang & Tremaine 2010).

We can integrate eq. 3 to obtain the total number density of soft binaries taking into account the contributions from the whole energy range. For simplicity, we approximate the exponential as unity.

We assume a uniform density of soft binaries inside the cluster, and approximate $N_{\rm soft}^{1,2} \approx n_{\rm eq}^{1,2}V$ and $\rho_i = m_iN_i/V$.

$$\begin{aligned} N_{\rm soft}^{1,2} &\approx R_{\rm cluster}^3 n_{\rm eq}^{1,2} \approx \\ &\approx \frac{N_1N_2\pi^{3/2}}{12M_{\rm cl}^2}\bar{m}^{3/2}M^{1/2} - \frac{N_1N_2\pi^{3/2}}{12}\left(\frac{m_1m_2}{M_{\rm cl}^2}\right)^{3/2}\end{aligned} \tag{5}$$

For any given lower energy cutoff $|\tilde{E}_{\rm min}|$, or equivalently a maximal separation $R_{\rm cut}$, the number of binaries will be given by

$$\begin{aligned} &N_{\rm bin}^{1,2}(R_{\rm cut}) \approx \\ &\approx \frac{(2\pi)^{3/2}G^{3/2}n_1n_2\bar{m}^{3/2}}{12\sigma^3}R_{\rm cluster}^3R_{\rm cut}^{3/2} - \frac{N_1N_2\pi^{3/2}}{12}\left(\frac{m_1m_2}{M_{\rm cl}^2}\right)^{3/2}\end{aligned} \tag{6}$$

Where for soft binaries, $R_{\rm cut} = R_{\rm Hill}$. For a general mass distribution $\xi(m)$,

$$N_{\rm bin}^{\rm total}(R_{\rm cut}) = \int N_{\rm bin}^{1,2}(R_{\rm cut})\xi(m_1)\xi(m_2)dm_1dm_2 \tag{7}$$

For single mass clusters, it can be seen that the number of existing systems at any given point is of order unity, as expected.

It should be noted that while our derivation was based on the assumption of thermal equilibrium, the distribution we obtained is in principle valid for a more general case, in which the only assumption is that the distribution of the background stars is Maxwellian, which is less demanding. Then, the final distribution is guaranteed based on the principle of detailed balance (Heggie 1975; Binney & Tremaine 2008).

2.2. *Expanding clusters – wide binaries in the field*

In later stages of cluster evolution, the cluster tends to dissolve, due to the expulsion of primordial gas and dynamical processes. Hence, a more realistic description of it should include expansion. We can extend our derivation to describe an expanding cluster. The Virial ratio is defined as the ratio between the kinetic and potential energies of the systems, i.e. $Q = -K/U$. Clusters with $Q = 1/2$ are in Virial equilibrium, $Q < 1/2$ corresponds to contracting clusters and $Q > 1/2$ to expanding clusters. The velocity dispersion of a cluster with a general Q could be written as $\sigma^2 = 2QGM_{\rm cl}/R_{\rm cl}$, and the distributions described in subsection 2.1 will change correspondingly, such that

$$n^{1,2}_{\rm exp}(\tilde{E}) \approx n^{1,2}_{\rm eq}(\tilde{E})/(2Q)^{3/2}, \; n^{1,2}_{\rm exp}(\tilde{a}) \approx n^{1,2}_{\rm eq}(\tilde{a})/(2Q)^{3/2} \tag{8}$$

It should be noted that not only the fraction for each type of binaries is changed, but also the integration boundaries, such that the overall number of binaries grows due to the expansion. The upper cutoff for an expanding cluster is now not the cluster Hill radius, but the Galactic tidal radius. Hence, not only the normalization changes but also the peak and the shape of the distributions. In principle, a more complicated dependence of the velocity dispersion on the radius could be introduced, which will also change the equilibrium stage of the cluster, i.e. the detailed balance between the creation and destruction of wide/soft binaries will no longer be sustained. However, for the cases we examined, and from the comparison to the N-body simulations, we conclude that the deviation from equilibrium is small.

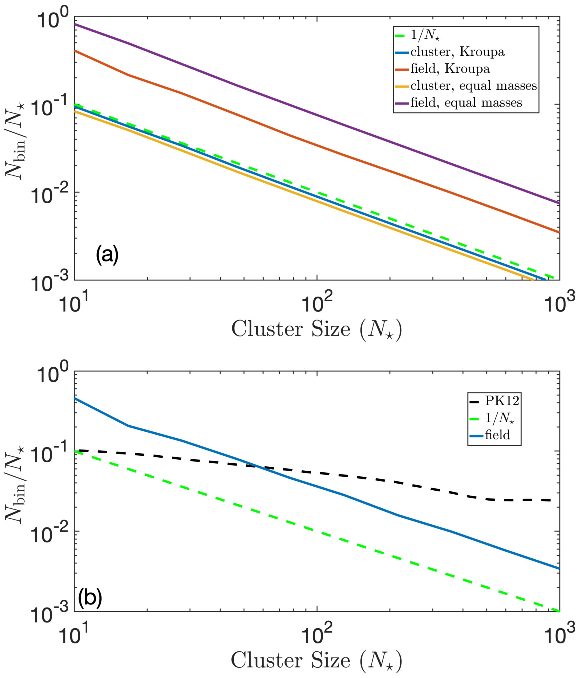

Figure 1. Fraction of soft/wide binary systems, as derived from the Monte-Carlo simulation, based on our analytical derivation. (a) The fraction of soft binaries in clusters/wide binaries in the field for different choices of mass distribution. (averaged over 100 Monte-Carlo realizations). (b) The fraction wide of binaries in the field where $N_\star = N_p$ is the number of stars and planets in the systems sampled, in comparison to N-body results (averaged over 50 Monte-Carlo realizations).

3. RESULTS

In the following, we present the distributions as derived from our model, as well as discuss the capture of triple and higher multiplicity systems and the capture of free-floating planets.

Unless stated otherwise, we consider a cluster with a Kroupa mass function (Kroupa 2001), in the range $0.1 \leq m \leq 7\ M_\odot$, following Perets & Kouwenhoven (2012). For the Monte Carlo simulation we perform, we draw the masses from a Kroupa mass function as implemented in `AMUSE` (Portegies Zwart et al. 2013) and its utility nstarman (2020). In several clusters we consider also planets, their mass is taken to be $1\ M_J$, and unless we stated otherwise, we consider in these clusters an equal number of stars and planets, and the size of the cluster by the number of stars. The radius of the cluster is set as $R_c = 0.1 N_\star^{1/3}$ pc (Perets & Kouwenhoven 2012), such that the number density of stars is kept constant even when the number of stars changes.

3.1. *Wide binary fractions*

In Fig. 1 we present the fraction of systems as a function of the cluster size. We present our results for expanding ($Q = 3/2$) and equilibrium/non-expanding ($Q = 1/2$) clusters, with the corresponding upper semi-major axis cutoffs: the Galactic tidal radius and the Hill radius of the cluster. As can be seen, the fraction of systems is a monotonically decreasing function of the cluster size, such that the overall scaling agrees well with the freeze-out distribution $1/N_\star$ (Moeckel & Clarke 2011), with different overall normalization factors for different choices of cutoffs and mass functions. It should be noted that the fraction derived in Perets & Kouwenhoven (2012) suggests a more flat dependence on the size of the cluster, which might arise from the specific type of realization of dispersing cluster used in those simulations.

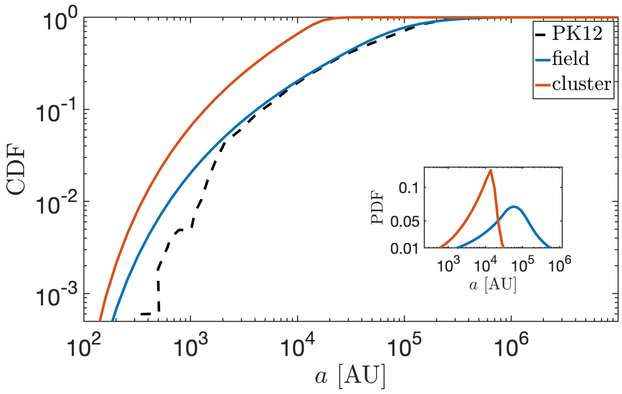

Figure 2. The semi-major axis distribution, as derived from the Monte-Carlo simulation, based on our analytical derivation, averaged over 1000 realizations, for a cluster containing both stars and planets ($N_\star = N_p$), in comparison to results from N-body simulations.

3.2. *The properties of wide binaries*

3.2.1. *Semi-major axis distribution*

In Fig. 2 we present the expected semi-major axis distribution of soft and wide binaries. The lower cutoff is determined by the transition between soft and hard binaries, i.e. at $a_{\rm SH} = Gm_1m_2/2\sigma^2$ and the upper cutoff is the galactic tidal radius ($R_{\rm tidal,g}$) or the Hill radius of the cluster ($R_{\rm Hill,c}$) correspondingly. The distribution peaks around $a \sim 10^5$AU for the galactic tidal cutoff in the field, and at $a \sim 10^4$AU for the Hill radius cutoff in clusters. nevertheless, the fraction of binaries shows a change by a factor of at most 2-3 over a wide range of semi-major axis (few times $10^3 - 10^4$ in clusters and few times $10^4 - 10^5$ AU in the field).

The expansion of the cluster changes the peak of the distribution, such that in expanding clusters the distribution peaks for higher separations, as expected given the larger allowed upper limit for the separation. Our results for the distribution are in good agreement with the N-body results presented in Perets & Kouwenhoven (2012), besides the small fraction of lower separations binaries (< 1000 AU), but at this regime, the small-statistics in the N-body results affect the apparent distribution.

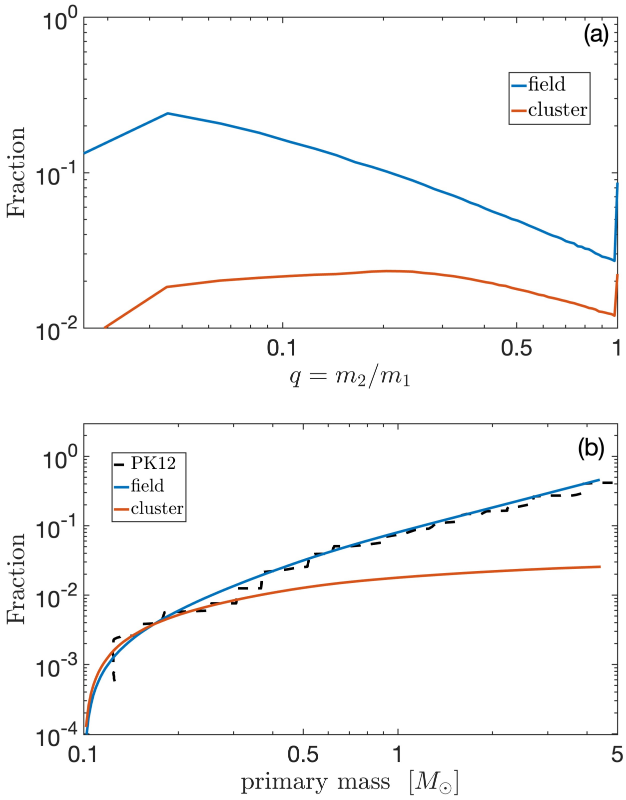

Figure 3. The binary fraction for different masses, as derived from an averaging over 1000 runs of the Monte-Carlo simulation, for a cluster containing 100 stars, drawn from a Kroupa mass function, and an equal number of planets, in comparison to N-body simulations. (a) The fraction of binaries as a function of the mass ratio, for stellar binaries only. (b) The fraction of binaries as a function of the primary mass.

3.2.2. *Masses and mass ratios*

In Fig. 3, we present the distribution of captured-formed binaries as a function of mass. We present both the mass ratio distribution (Subfig. a) and the primary mass distribution (Subfig. b). As expected from eq. 5, more massive stars are more likely to reside in binaries, and constitute primaries. Our results are in excellent agreement with the N-body results presented in Perets & Kouwenhoven (2012).

3.2.3. *Eccentricites & Inclinations*

In thermalized clusters, binaries are generally expected to show a thermal eccentricity distribution (e.g Ambartsumian 1937), which would then naturally be expected for the wide binaries, which is also consistent with N-body results (Perets & Kouwenhoven 2012). However, different cutoffs of the phase space might lead to deviations from thermal distribution (Rozner & Perets 2023, in prep.). Observationally, it was found that the eccentricity distribution could slightly deviate from thermal distribution Raghavan et al. (2010); Moe & Di Stefano (2017); Tokovinin (2020), with twin (same mass) wide binaries are particularly eccentric (Hwang et al. 2022a,b). However, the origin for the latter observation is still not understood, and the wide-binaries formation models explored here are not expected to produce such eccentric "twins". In populations of non-equal masses, the eccentricity distribution might show some mass dependence, however, here we only explore the distributions of semi-major axes and binary mass function and leave further discussion of eccentricities to later studies.

Since dynamical captures are generally random, the distribution of the orientations of the formed wide binaries is expected to be randomized. For binaries, this would suggest a random spin-orbit inclination for dynamically-formed wide binaries. It was found that binaries with small separations tend to be spin-orbit aligned, while for wide binaries the alignment appears to be random (Hale 1994). However, later studies showed that it might be more complicated, and some correlations may exist (e.g. Justesen & Albrecht 2020). After their formation, wide binaries undergo interaction with external perturbers, that can change their eccentricity and inclination distribution. For triples (see next section), this would suggest a random relative inclination between the inner and the outer orbits of wide triples (and quadruples), which has potential implications for secular processes to play an important role in the evolution, as we discuss below. It should be noted that the eccentricity, as well as inclination, could also be affected by the long-term secular evolution due to the Galactic tide (but it does not explain the observed superthermal distribution of eccentricities (Modak & Hamilton 2023).

3.3. *Wide triples and higher multiplicity systems*

Triple and higher multiplicity systems are also known to be abundant (Tokovinin 2008; Raghavan et al. 2010), and in particular, most of the massive O/B stars are observed to be part of triples or higher multiplicity systems (Sana et al. 2012; Moe & Di Stefano 2017).

Wide triple (and higher multiplicity) systems can also form through the dynamical capture mechanisms discussed above, through two possible channels: (I) Since primordial binaries are frequent, as discussed above, such systems can dynamically capture additional wide companions, similar to capture by single stars; a hard binary can capture a third distant companion (forming a triple), or another hard binary (forming a quadruple), and this could even result in even higher multiplicity systems if primordial triples etc. are considered. (II) A wide binary formed through dynamical capture, can then consequently capture additional wide companions.

The first channel describes the formation of a hierarchical triple in which $a_{\rm in} \ll a_{\rm out}$, while in the second channel, the capture of an additional companion to an already wide binary might give rise to more comparable inner and outer separations, and even $a_{\rm in}/a_{\rm out} \sim 1$. In the latter case, the newly formed systems might not be stable and would disrupt over a few dynamical timescales, ejecting one of the stellar components. Here we will briefly discuss each of these channels.

In the first capture scenario, the initial binary could be thought of as a single object with mass $M_{\rm in} = m_1 + m_2$ that captures another object with mass m_3, and then the calculation described in section 2 could be reiterated directly, with the appropriate modifications,

$$N_{\rm tri}^{12,3}(R_{\rm cut}) \approx \frac{(2\pi)^{3/2}G^{3/2}n_{12}n_3\bar{m}^{3/2}}{12\sigma^3}R^3_{\rm cluster}R^{3/2}_{\rm cut} - \\ - \frac{N_1N_2\pi^{3/2}}{12}\left(\frac{(m_1+m_2)m_3}{M^2_{\rm cl}}\right)^{3/2} \quad (9)$$

In this scenario, given the known high fractions of binary systems, the capture of a binary system by another binary system is highly likely. Therefore quadruple configurations of 2+2 (two close binary systems orbiting each other at a wide orbit) should be frequent, and secular effects may then later drive the inner binaries into compact configurations through Kozai-cycles and tidal friction. Observations suggest that such quadruple systems are indeed overabundant (Fezenko et al. 2022), and this scenario might provide a natural explanation.

The probability for the second channel, i.e. having two consequent captures is $f_{\rm tri,cap}(m_1, m_2, m_3) \approx f(m_1, m_2) \times f(m_2, m_3)$, assuming, without loss of generality, that m_2 captures m_3 and that two consequent captures are independent. In general, consequent captures could lead to the formation of high multiplicity systems, but with lower probability. In general, the probability of capturing the n-th object with mass m is given by $f_{\rm nm} = f_m^n$ where f_m is the probability of capturing m just once. These probabilities would be somewhat overestimated, in cases where the formed systems are not hierarchical, i.e. considering appropriate stability criteria, (e.g. Valtonen et al. 2008) would be destabilized and lost, lowering the fractions.

3.4. *Contributions to the field population*

The field population consists of contributions from many dispersed clusters. Here we will account for their weighted contributions. To compare our results with observations, we restrict the mass function from which we sample to the range $[0.1, 3]$ $M_\odot$, and the possible separations to be in the range of 100 AU $-$ 1 pc, based on the observational results from El-Badry et al. (2021) and references therein. We then sample clusters/associations with $N_\star = 10 - 10^5$ stars from a power-law distribution $dN/dN_\star \propto N_\star^{-2}$, following Lada & Lada (2003). We consider for these purposes only wide binaries in the field originated from dispersed clusters.

In Fig. 4, we present the separation distribution in different cluster sizes, and the integrated distribution for clusters with $10 - 10^5$ stars, sampled from $dN/dN_\star \propto N_\star^{-2}$. The separation distribution is affected by the size of the cluster twice: since $\sigma \propto N_\star^{1/3}$, the overall number density per separation scales as $n_{1,2}(a) \propto N_\star^{-1}$, and the lower separation cutoff scales as $a_{\rm SH} \propto N_\star^{-2/3}$. Hence, the contribution from larger clusters shifts the distribution towards lower separations. Since the powerlaw we chose for the number of stars is negative, the overall distribution is dominated by smaller clusters/associations and our choice of the lower cutoff for the number of stars in these.

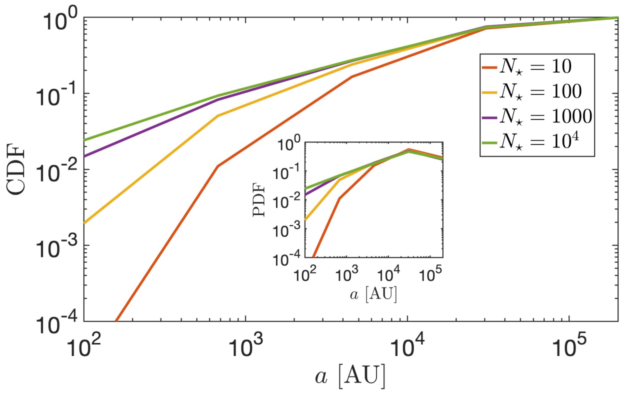

Figure 4. The semi-major axis distribution, as derived from the Monte-Carlo simulation, based on our analytical derivation, for clusters in different sizes.

4. DISCUSSION

In the following, we discuss our results for the distributions of wide/soft binaries, as well as discuss the formation of wide triples and higher multiplicity systems and their implications. Some of these were already discussed in Perets & Kouwenhoven (2012); we briefly summarize them, but also new issues and implications.

4.1. *Comparison of analytic and N-body results*

As discussed above, overall we can reproduce analytically the N-body results presented by Perets & Kouwenhoven (2012), with an excellent agreement, apart from the overall fraction of binaries, in which we find a steeper slope. However, our results for the fractions are in agreement with other N-body models which studied the freeze-out distribution of soft binaries in clusters (Moeckel & Clarke 2011).

4.2. *Comparison with observations*

Observational searches of wide binaries usually rely on probabilistic arguments to determine if the binary components are bound, given the uncertainties and/or unknowns in the inferred orbital elements. The era of GAIA data has dramatically expanded the sample of binaries in general and wide binary candidates in particular (Gaia Collaboration et al. 2016, 2018, 2021), and several studies explored the frequency and distribution of such binaries (e.g. Oh et al. 2017; El-Badry & Rix 2018; El-Badry et al. 2019; Hartman & Lépine 2020; Tian et al. 2020; El-Badry et al. 2021).

The observational wide binary distribution peaks around the separation of $10^3 - 10^4$ AU (Andrews et al. 2017; El-Badry et al. 2021 and references therein), while the distribution we find here peaks at 10^4 AU in clusters and a few times higher in the field. This could be explained by observational biases, the inclusion of cluster binaries in some surveys, the choice of the lower cutoff for the smallest clusters that contribute to the field population, and/or the contribution from other binary-formation channels at smaller separations. In addition, over time wide binaries can experience encounters with field stars. Given the typical velocity dispersion in the field, the widest binaries would evaporate and be disrupted, leaving behind wide binaries of smaller separations, as well as more realistic profiles for the densities of the clusters (e.g. Binney & Tremaine 2008; Jiang & Tremaine 2010; Michaely & Perets 2016).

In the equilibrium distribution explored here, The overall wide binary fraction is dominated by the contributions from small clusters/associations, as the cluster-size scales as $dN/dN_\star \propto N_\star^{-2}$, each cluster has $N_\star$ stars, such that the overall dependence scales as $1/N_\star \times$binary fraction. Hence, the lower size cutoff plays an important role in determining the final field properties. For the overall fraction, we consider stellar masses in the range $0.1 - 7$ $M_\odot$ and clusters in sizes $N_\star = 10 - 10^3$. The total wide binaries fraction for these masses as derived from our model is 0.2 for $N_{\star,\rm min} \approx 30$; and 0.006 for $N_{\star,\rm min} = 100$. The observational wide binary fraction ($a > 10^3$ AU) was found to be 0.115 (Raghavan et al. 2010; Moeckel & Clarke 2011). Hwang et al. (2021) found a fraction of 0.071 within 100 pc for binaries with separations larger than 1000 AU. Later studies considering a wide range of observations (Moe & Di Stefano 2017, and references therein) found a wide-binary fraction of 0.15 for periods $\log P(\mathrm{days}) = 6.1 - 7.4$ (corresponding to separations of $\sim 300 - 2000$ AU for solar mass stars).

Igoshev & Perets (2019) discussed the expected wide binary fractions for different mass groups. They studied the binarity of OB stars and found a wide multiplicity rate of 0.091 for sufficiently luminous secondary stars (G dwarfs or higher) with separations larger than 10^3 AU. Extrapolating for lower mass secondaries (assuming a Kroupa mass function for the secondaries), they infer an overall wide binary fraction of 0.27.

Setting $N_{\star,\min} = 30$, we find a fraction of 0.1 for G dwarfs (or higher masses) captured by OB-5 stars (cut for our case for $4.36 \leq m \leq 7\ M_{\odot}$) and a total fraction of 0.5 for OB captures. These fractions are roughly consistent with the fractions inferred by Igoshev & Perets (2019) mentioned above.

Overall, the fractions we derived are in the expected observational range but are sensitive to the choice for the lower cutoff of clusters'/stellar associations' contribution to field stars. Direct comparison, is difficult for wide binaries of low-mass stars which are long-lived and can be significantly affected by field perturbations and evaporation over time, in particular at larger separations. In that context, short-lived massive stars provide a better direct probe for the pure formation of wide binaries, not affected by long-term dynamical evolution.

As we discussed earlier, the overall binary fraction in clusters scales with $N_\star^{-1}$. Hwang et al. (2021) (and references therein) showed that clusters with lower metallicities, which might indicate higher masses, correspond to larger fractions of wide binaries, which might point on a general agreement with our results. Further comparison with these results is beyond the scope of this paper.

4.3. *Implications for the dynamical evolution of stellar binaries and higher multiplicity systems*

4.3.1. *Collisional field dynamics*

As briefly mentioned in the introduction, wide binaries/triples could serve as catalysts for the collisional dynamics of field stars where flyby perturbations of wide binaries/triples, excite their eccentricities to the point where the components strongly interact and may produce compact binaries, explosive transients and gravitational wave sources (Kaib & Raymond 2014; Michaely & Perets 2019, 2020; Grishin & Perets 2022; Michaely & Naoz 2022). Our results suggest that dynamical formation of wide binaries and triples could be quite frequent and can occur Myrs after the initial formation of the stellar components. In principle, stellar evolutionary scenarios could easily disrupt primordial wide binaries, e.g. due to even low natal kicks and/or Blaauw (prompt mass-loss) kicks (e.g. Blaauw 1961; Hansen & Phinney 1997; Igoshev & Perets 2019 and references inside), in particular for massive stars and their remnant black holes or neutron stars, known to experience prompt mass-loss/natal kicks following core-collapse supernova explosions. One would then think of excluding the possibility of wide BH/NS binaries, and their participation in collisional field dynamics (but see Raveh et al. 2022). However, our results provide a channel for the existence of such BH/NS wide binaries, since the dynamical capture, which can occur already after these remnants lost their companions, could allow them to acquire wide companions, which could then indeed play an important role in their evolution through collisional dynamics in the field. In particular, given their relatively large masses, NS, and, to a much higher extent BHs are very likely to capture companions, given our finding of the mass dependence.

4.3.2. *Secular evolution*

Triple secular and quasi-secular dynamics were suggested to catalyze a variety of phenomena, in particular when the inner and outer orbits are significantly misaligned. Such processes include the formation of short-period binaries (e.g. Mazeh & Shaham 1979; Kiseleva et al. 1998), mergers, gravitational waves merger sources (e.g. Blaes et al. 2002; Antognini et al. 2014; Antonini et al. 2017; Liu & Lai 2018; Michaely & Perets 2020), type Ia supernovae (e.g. Katz & Dong 2012; Thompson 2011), blue stragglers (Perets & Fabrycky 2009) and many others (see for a detailed review of many of these in Naoz 2016), and hence understanding their properties and distributions analytically could have significant implications.

The formation of likely misaligned triple and quadruple systems in the capture scenario could naturally provide the necessary conditions for significant secular evolution to take place, and therefore the capture scenario may play a key role in the initial production of secularly evolving systems, and their resulting strong interactions.

Furthermore, very wide systems are also affected by secular processes triggered by the Galactic tidal field (even for wide binaries, not only higher multiplicity systems). The dynamics in the galactic fields and other external perturbations were studied extensively (e.g. Heisler & Tremaine 1986; Jiang & Tremaine 2010; Hamilton & Rafikov 2019a,b, 2021; Grishin & Perets 2022 and references therein), and it is qualitatively similar to the evolution of quadruple hierarchical systems. Again, our finding on the formation of very wide binaries in the dynamical capture scenario suggests a large number of wide systems, sensitive to galactic tidal secular perturbations exist.

4.4. *Capture of planets*

Free-floating planets (FFPs) are planets that are not bound to any star or brown dwarf. The dynamical interactions between planets are thought to give rise to the frequent ejection of planets, and the production of unbound FFPs (Rasio & Ford 1996).

Regardless of the formation channel, since the majority of stars are thought to be born and evolve in clusters, so do planets. In principle, the derivation we used could describe the capture of FFPs by stars or other FFPs. Perets & Kouwenhoven (2012) already suggested that during the cluster dispersion planets could be captured, in a similar process. Here we should only mention the potential caveat that planets might not be expected to thermalize with the stars, due to the extreme mass ratio with respect to stars, which will change our distribution assumptions. Nevertheless, following their ejection the velocities of FFPs should initially follow the overall velocity distributions of their original host stars, and might not have time to significantly change before their host cluster disperses.

5. SUMMARY

In this paper, we analytically explored the formation of soft binaries in clusters and wide binaries in the field, and derived their properties. We show that our analytic results well reproduce detailed N-body simulations, and can be generally used to derive the properties of dynamically-capture wide binaries in any environment. We find that the capture formation can potentially explain the origin of most of the observed wide binaries, and can give rise to higher multiplicity systems, and possibly explain the observed overabundance of wide 2+2 quadruple systems. Soft and wide binaries and higher multiplicity systems are highly sensitive to collisional field dynamical processes, as well as secular dynamical processes, all of which may give rise to the formation of compact binaries and/or mergers and transient phenomena. Therefore, the understanding of the formation of wide systems and their properties is of significant importance for the evolution of stars and the production of explosive transients.

ACKNOWLEDGEMENTS

MR acknowledges the generous support of Azrieli fellowship.

REFERENCES

Ambartsumian, V. 1937, Astronomicheskii Zhurnal, 14

Andrews, J. J., Chanamé, J., & Agüeros, M. A. 2017, MNRAS, 472, 675, doi: 10.1093/mnras/stx2000

—. 2018, MNRAS, 473, 5393, doi: 10.1093/mnras/stx2685

Antognini, J. M., Shappee, B. J., Thompson, T. A., & Amaro-Seoane, P. 2014, MNRAS, 439, 1079, doi: 10.1093/mnras/stu039

Antonini, F., Toonen, S., & Hamers, A. S. 2017, ApJ, 841, 77, doi: 10.3847/1538-4357/aa6f5e

Bahcall, J. N., Hut, P., & Tremaine, S. 1985, ApJ, 290, 15, doi: 10.1086/162953

Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton University Press)

Blaauw, A. 1961, BAN, 15, 265

Blaes, O., Lee, M. H., & Socrates, A. 2002, ApJ, 578, 775, doi: 10.1086/342655

Blas, D., Nacir, D. L., & Sibiryakov, S. 2017, PhRvL, 118, 261102, doi: 10.1103/PhysRevLett.118.261102

Bonsor, A., & Veras, D. 2015, MNRAS, 454, 53, doi: 10.1093/mnras/stv1913

Chanamé, J., & Gould, A. 2004, ApJ, 601, 289, doi: 10.1086/380442

Chandrasekhar, S. 1944, ApJ, 99, 54, doi: 10.1086/144589

Duchêne, G., Bouvier, J., Bontemps, S., André, P., & Motte, F. 2004, A&A, 427, 651, doi: 10.1051/0004-6361:20041209

Duchêne, G., & Kraus, A. 2013, ARA&A, 51, 269, doi: 10.1146/annurev-astro-081710-102602

Duquennoy, A., & Mayor, M. 1991, A&A, 248, 485

El-Badry, K., & Rix, H.-W. 2018, MNRAS, 480, 4884, doi: 10.1093/mnras/sty2186

El-Badry, K., Rix, H.-W., & Heintz, T. M. 2021, MNRAS, 506, 2269, doi: 10.1093/mnras/stab323

El-Badry, K., Rix, H.-W., Tian, H., Duchêne, G., & Moe, M. 2019, MNRAS, 489, 5822, doi: 10.1093/mnras/stz2480

Fezenko, G. B., Hwang, H.-C., & Zakamska, N. L. 2022, MNRAS, 511, 3881, doi: 10.1093/mnras/stac309

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1, doi: 10.1051/0004-6361/201629272

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1, doi: 10.1051/0004-6361/201833051

—. 2021, A&A, 649, A1, doi: 10.1051/0004-6361/202039657

Ginat, Y. B., & Perets, H. B. 2021, MNRAS, 508, 190, doi: 10.1093/mnras/stab2565

Goodman, J., & Hut, P. 1993, ApJ, 403, 271, doi: 10.1086/172200

Grishin, E., & Perets, H. B. 2022, MNRAS, 512, 4993, doi: 10.1093/mnras/stac706

Hale, A. 1994, AJ, 107, 306, doi: 10.1086/116855

Hamilton, C., & Rafikov, R. R. 2019a, MNRAS, 488, 5489, doi: 10.1093/mnras/stz1730

—. 2019b, MNRAS, 488, 5512, doi: 10.1093/mnras/stz2026

—. 2021, MNRAS, 505, 4151, doi: 10.1093/mnras/stab1284

Hansen, B. M. S., & Phinney, E. S. 1997, MNRAS, 291, 569, doi: 10.1093/mnras/291.3.569

Hartman, Z. D., & Lépine, S. 2020, ApJS, 247, 66, doi: 10.3847/1538-4365/ab79a6

Hawkins, K., Lucey, M., Ting, Y.-S., et al. 2020, MNRAS, 492, 1164, doi: 10.1093/mnras/stz3132

Heggie, D. C. 1975, MNRAS, 173, 729, doi: 10.1093/mnras/173.3.729

—. 1977, RMxAA, 3

Heisler, J., & Tremaine, S. 1986, Icarus, 65, 13, doi: 10.1016/0019-1035(86)90060-6

Hills, J. G. 1975, AJ, 80, 809, doi: 10.1086/111815

Hwang, H.-C., El-Badry, K., Rix, H.-W., et al. 2022a, ApJL, 933, L32, doi: 10.3847/2041-8213/ac7c70

Hwang, H.-C., Ting, Y.-S., Schlaufman, K. C., Zakamska, N. L., & Wyse, R. F. G. 2021, MNRAS, 501, 4329, doi: 10.1093/mnras/staa3854

Hwang, H.-C., Ting, Y.-S., & Zakamska, N. L. 2022b, MNRAS, 512, 3383, doi: 10.1093/mnras/stac675

Igoshev, A. P., & Perets, H. B. 2019, MNRAS, 486, 4098, doi: 10.1093/mnras/stz1024

Jiang, Y.-F., & Tremaine, S. 2010, MNRAS, 401, 977, doi: 10.1111/j.1365-2966.2009.15744.x

Justesen, A. B., & Albrecht, S. 2020, A&A, 642, A212, doi: 10.1051/0004-6361/202039138

Kaib, N. A., & Raymond, S. N. 2014, 422.05

Katz, B., & Dong, S. 2012, arXiv e-prints, arXiv:1211.4584, doi: 10.48550/arXiv.1211.4584

King, I. R. 1977, Revista Mexicana de Astronomia y Astrofisica, vol. 3, 3, 167

Kiseleva, L. G., Eggleton, P. P., & Mikkola, S. 1998, MNRAS, 300, 292, doi: 10.1046/j.1365-8711.1998.01903.x

Kouwenhoven, M. B. N., Goodwin, S. P., Parker, R. J., et al. 2010, MNRAS, 404, 1835, doi: 10.1111/j.1365-2966.2010.16399.x

Kroupa, P. 2001, MNRAS, 322, 231, doi: 10.1046/j.1365-8711.2001.04022.x

Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57, doi: 10.1146/annurev.astro.41.011802.094844

Larson, R. B. 2001, in The Formation of Binary Stars, ed. H. Zinnecker & R. Mathieu, Vol. 200, 93, doi: 10.48550/arXiv.astro-ph/0006288

Liu, B., & Lai, D. 2018, ApJ, 863, 68, doi: 10.3847/1538-4357/aad09f

Mazeh, T., & Shaham, J. 1979, A&A, 77, 145

Michaely, E., & Naoz, S. 2022, ApJ, 936, 184, doi: 10.3847/1538-4357/ac8a92

Michaely, E., & Perets, H. B. 2016, MNRAS, 458, 4188, doi: 10.1093/mnras/stw368

—. 2019, ApJL, 887, L36, doi: 10.3847/2041-8213/ab5b9b

—. 2020, MNRAS, 498, 4924, doi: 10.1093/mnras/staa2720

Modak, S., & Hamilton, C. 2023, arXiv e-prints, arXiv:2303.15531, doi: 10.48550/arXiv.2303.15531

Moe, M., & Di Stefano, R. 2017, ApJS, 230, 15, doi: 10.3847/1538-4365/aa6fb6

Moeckel, N., & Bate, M. R. 2010, MNRAS, 404, 721, doi: 10.1111/j.1365-2966.2010.16347.x

Moeckel, N., & Clarke, C. J. 2011, MNRAS, 415, 1179, doi: 10.1111/j.1365-2966.2011.18731.x

Naoz, S. 2016, ARA&A, 54, 441, doi: 10.1146/annurev-astro-081915-023315

nstarman. 2020, amuse_util, Zenodo, doi: 10.5281/zenodo.3740579

Offner, S. S. R., Kratter, K. M., Matzner, C. D., Krumholz, M. R., & Klein, R. I. 2010, ApJ, 725, 1485, doi: 10.1088/0004-637X/725/2/1485

Oh, S., Price-Whelan, A. M., Hogg, D. W., Morton, T. D., & Spergel, D. N. 2017, AJ, 153, 257, doi: 10.3847/1538-3881/aa6ffd

Perets, H. B., & Fabrycky, D. C. 2009, ApJ, 697, 1048, doi: 10.1088/0004-637X/697/2/1048

Perets, H. B., & Kouwenhoven, M. B. N. 2012, ApJ, 750, 83, doi: 10.1088/0004-637X/750/1/83

Portegies Zwart, S., McMillan, S. L. W., van Elteren, E., Pelupessy, I., & de Vries, N. 2013, Computer Physics Communications, 184, 456, doi: 10.1016/j.cpc.2012.09.024

Quinn, D. P., Wilkinson, M. I., Irwin, M. J., et al. 2009, MNRAS, 396, L11, doi: 10.1111/j.1745-3933.2009.00652.x

Raghavan, D., McAlister, H. A., Henry, T. J., et al. 2010, ApJS, 190, 1, doi: 10.1088/0067-0049/190/1/1

Rasio, F. A., & Ford, E. B. 1996, Science, 274, 954, doi: 10.1126/science.274.5289.954

Raveh, Y., Michaely, E., & Perets, H. B. 2022, MNRAS, 514, 4246, doi: 10.1093/mnras/stac1605

Reipurth, B., & Mikkola, S. 2012, Nature, 492, 221, doi: 10.1038/nature11662

Retterer, J. M., & King, I. R. 1982, ApJ, 254, 214, doi: 10.1086/159725

Rozner, M., Generozov, A., & Perets, H. B. 2022, arXiv e-prints, arXiv:2212.00807, doi: 10.48550/arXiv.2212.00807

Rozner, M., Grishin, E., Ginat, Y. B., Igoshev, A. P., & Desjacques, V. 2020, JCAP, 2020, 061, doi: 10.1088/1475-7516/2020/03/061

Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Science, 337, 444, doi: 10.1126/science.1223344

Tagawa, H., Haiman, Z., & Kocsis, B. 2020, ApJ, 898, 25, doi: 10.3847/1538-4357/ab9b8c

Thompson, T. A. 2011, ApJ, 741, 82, doi: 10.1088/0004-637X/741/2/82

Tian, H.-J., El-Badry, K., Rix, H.-W., & Gould, A. 2020, ApJS, 246, 4, doi: 10.3847/1538-4365/ab54c4

Tokovinin, A. 2008, MNRAS, 389, 925, doi: 10.1111/j.1365-2966.2008.13613.x

—. 2020, MNRAS, 496, 987, doi: 10.1093/mnras/staa1639

Valtonen, M., Mylläri, A., Orlov, V., & Rubinov, A. 2008, in Dynamical Evolution of Dense Stellar Systems, ed. E. Vesperini, M. Giersz, & A. Sills, Vol. 246, 209–217, doi: 10.1017/S1743921308015627

Xu, S., Hwang, H.-C., Hamilton, C., & Lai, D. 2023, arXiv e-prints, arXiv:2303.16224, doi: 10.48550/arXiv.2303.16224

Yabushita, S. 1966, MNRAS, 133, 133, doi: 10.1093/mnras/133.1.133

Chapter 7

Discussion and Summary

In this dissertation, I studied dynamical phenomena in different scales.

In chapter 2, I described the process of aeolian erosion in protoplanetary disks, which is a mechanical destructive process that leads to a radius shrinkage of objects in the disk. I introduced an analytical model for the phenomenon, based on geological literature for aeolian erosion of sand dunes with the necessary modifications. I discussed several implications and interconnections between aeolian erosion and other processes that take place during the initial stages of planet formation. Then, I studied a similar process in white dwarf disks and discussed its implications. Aeolian erosion is most efficient for meter-sized objects and could affect significantly the object evolution and size distribution. It has symbiotic relations with other processes during the initial stages of planet formation and should be coupled to the planet formation models to obtain a fuller picture of these stages.

In chapter 3, I treated Pluto-Charon and the Sun as a hierarchical triple, and suggested that Pluto-Charon was formed from an initially highly inclined wide binary that was driven via secular/quasi-secular interaction towards its current state. I then investigate the available parameter space for such an event, using analytic tools and few-body simulations. The process described is general and could explain similarly other systems in the Kuiper belt.

In chapter 4, I discussed the formation and destruction of hot and warm Jupiters, considering inflated eccentric migration. By the end of core acrretion, gas giants are formed inflated, and the contraction time is such that during a significant fraction of their migration, they could still remain inflated. I coupled their thermal evolution them (black body cooling/heating due to an external source) to their dynamical evolution due to eccentric tidal migration. I employed analytical techniques and then compared them to numerical simulations and found an excellent agreement wherever the numerical simulations were valid. I then carried out a detailed population synthesis and investigated the properties of the migrating gas giants and concluded regarding the expected formed hot and warm Jupiters via this channel. While inflated eccentric migration elevates the fraction of formed warm Jupiters, it also increases the fraction of disrupted hot Jupiters.

In chapter 5, I discuss several projects, all of them centered around the dynamics of binaries in gas-rich environments. In section 5.1, I introduced a novel gravitational waves channel, of compact objects embedded in gas-rich clusters of the second (or later) generation star formation. While globular clusters are gas-dilute currently, during the formation of later generations, they are expected to be gas-rich, which provides further energy dissipation mechanism for the binaries, since gas-hardening could extract energy from the binary, leading it towards smaller separations in which gravitational waves are likely to occur within a Hubble time. In section 5.2 I discussed the conditions for gas-assisted binary formation. The dynamical formation of binaries requires an external dissipation mechanism. In gas-rich environments, gas hardening could be a natural solution to the problem. I derived analytically the conditions for gas-assisted capture and compared the results to N-body simulations with an external force. This formation mechanism is robust and could apply to various gas-rich environments such as gas-rich globular clusters or star-forming regions. In section

5.3, I introduced another dynamical process, termed "gas shielding". While soft binaries in dense environments tend to get disrupted shortly, in gas-rich environments gas hardening could compete with stellar softening and weaken the effect of softening, potentially protecting soft binaries from disruption and giving rise to a novel channel of hard binaries formation. This process affects the total binary distribution and its properties. Our studies suggest a new point of view of the dynamics in these environments. We studied not only new processes that could revise the properties of binaries in these environments, but also environments which are usually thought as gas-free, such as globular clusters.

In chapter 6, I studied the distribution of soft binaries in clusters and wide binaries in the field using statistical mechanics tools. I derived the semimajor axis and mass distributions, as well as the total fractions, and compared my results to N-body simulations. The results overall agree apart from slight differences. This formalism allows us to study the properties of these binaries without using expensive N-body simulations.

To summarize, I studied various problems in dynamics, mostly using analytical tools but also numerical ones. While they might seem unrelated at first sight, they share similar physical ideas and techniques, and I hope to keep and develop these ideas in the future.

Bibliography

S. J. Aarseth and D. C. Heggie. The probability of binary formation by three-body encounters. A&A, 53:259–265, December 1976.

B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. Phys. Rev. Lett., 116(6):061102, February 2016. doi: 10.1103/PhysRevLett.116.061102.

I. Adachi, C. Hayashi, and K. Nakazawa. The gas drag effect on the elliptical motion of a solid body in the primordial solar nebula. Progress of Theoretical Physics, 56: 1756–1771, December 1976. doi: 10.1143/PTP.56.1756.

M. E. Alexander. The Weak Friction Approximation and Tidal Evolution in Close Binary Systems. Ap&SS, 23(2):459–510, August 1973. doi: 10.1007/BF00645172.

Kassandra R. Anderson, Natalia I. Storch, and Dong Lai. Formation and stellar spin-orbit misalignment of hot Jupiters from Lidov-Kozai oscillations in stellar binaries. MNRAS, 456(4):3671–3701, March 2016. doi: 10.1093/mnras/stv2906.

J. M. O. Antognini. Timescales of Kozai-Lidov oscillations at quadrupole and octupole order in the test particle limit. MNRAS, 452(4):3610–3619, October 2015. doi: 10.1093/mnras/stv1552.

Andrea Antoni, Morgan MacLeod, and Enrico Ramirez-Ruiz. The Evolution of Binaries in a Gaseous Medium: Three-dimensional Simulations of Binary Bondi-

Hoyle-Lyttleton Accretion. ApJ, 884(1):22, October 2019. doi: 10.3847/1538-4357/ab3466.

Fabio Antonini and Hagai B. Perets. Secular Evolution of Compact Binaries near Massive Black Holes: Gravitational Wave Sources and Other Exotica. ApJ, 757 (1):27, September 2012. doi: 10.1088/0004-637X/757/1/27.

Fabio Antonini, Sourav Chatterjee, Carl L. Rodriguez, Meagan Morscher, Bharath Pattabiraman, Vicky Kalogera, and Frederic A. Rasio. Black Hole Mergers and Blue Stragglers from Hierarchical Triples Formed in Globular Clusters. ApJ, 816 (2):65, January 2016a. doi: 10.3847/0004-637X/816/2/65.

Fabio Antonini, Sourav Chatterjee, Carl L. Rodriguez, Meagan Morscher, Bharath Pattabiraman, Vicky Kalogera, and Frederic A. Rasio. Black Hole Mergers and Blue Stragglers from Hierarchical Triples Formed in Globular Clusters. ApJ, 816 (2):65, January 2016b. doi: 10.3847/0004-637X/816/2/65.

Fabio Antonini, Silvia Toonen, and Adrian S. Hamers. Binary Black Hole Mergers from Field Triples: Properties, Rates, and the Impact of Stellar Evolution. ApJ, 841(2):77, June 2017. doi: 10.3847/1538-4357/aa6f5e.

Philip J. Armitage. Astrophysics of Planet Formation. 2010.

Pawel Artymowicz, D. N. C. Lin, and E. J. Wampler. Star Trapping and Metallicity Enrichment in Quasars and Active Galactic Nuclei. ApJ, 409:592, June 1993. doi: 10.1086/172690.

Ralph A. Bagnold. The physics of blown sand and desert dunes. Methuen, London, 1941.

J. N. Bahcall, P. Hut, and S. Tremaine. Maximum mass of objects that constitute unseen disk material. ApJ, 290:15–20, March 1985. doi: 10.1086/162953.

M. Bailes, B. K. Berger, P. R. Brady, M. Branchesi, K. Danzmann, M. Evans, K. Holley-Bockelmann, B. R. Iyer, T. Kajita, S. Katsanevas, M. Kramer, A. Lazzarini, L. Lehner, G. Losurdo, H. Lück, D. E. McClelland, M. A. McLaughlin, M. Punturo, S. Ransom, S. Raychaudhury, D. H. Reitze, F. Ricci, S. Rowan, Y. Saito, G. H. Sanders, B. S. Sathyaprakash, B. F. Schutz, A. Sesana, H. Shinkai, X. Siemens, D. H. Shoemaker, J. Thorpe, J. F. J. van den Brand, and S. Vitale. Gravitational-wave physics and astronomy in the 2020s and 2030s. Nature Reviews Physics, 3(5):344–366, April 2021. doi: 10.1038/s42254-021-00303-8.

C. Baruteau, A. Crida, S. J. Paardekooper, F. Masset, J. Guilet, B. Bitsch, R. Nelson, W. Kley, and J. Papaloizou. Planet-Disk Interactions and Early Evolution of Planetary Systems. In Henrik Beuther, Ralf S. Klessen, Cornelis P. Dullemond, and Thomas Henning, editors, Protostars and Planets VI, page 667, January 2014. doi: 10.2458/azu_uapress_9780816531240-ch029.

N. Bastian, H. J. G. L. M. Lamers, S. E. de Mink, S. N. Longmore, S. P. Goodwin, and M. Gieles. Early disc accretion as the origin of abundance anomalies in globular clusters. MNRAS, 436(3):2398–2411, December 2013. doi: 10.1093/mnras/stt1745.

Nate Bastian and Carmela Lardo. Multiple Stellar Populations in Globular Clusters. ARA&A, 56:83–136, September 2018. doi: 10.1146/annurev-astro-081817-051839.

Matthew R. Bate, Ian A. Bonnell, and Nigel M. Price. Modelling accretion in protobinary systems. MNRAS, 277(2):362–376, November 1995. doi: 10.1093/mnras/277.2.362.

Matthew R. Bate, Ian A. Bonnell, and Volker Bromm. The formation of close binary systems by dynamical interactions and orbital decay. MNRAS, 336(3):705–713, November 2002. doi: 10.1046/j.1365-8711.2002.05775.x.

Kenji Bekki. Globular cluster formation with multiple stellar populations: selfenrichment in fractal massive molecular clouds. MNRAS, 469(3):2933–2951, August 2017. doi: 10.1093/mnras/stx982.

James Binney and Scott Tremaine. Galactic Dynamics: Second Edition. 2008.

Omer Blaes, Man Hoi Lee, and Aristotle Socrates. The Kozai Mechanism and the Evolution of Binary Supermassive Black Holes. ApJ, 578(2):775–786, October 2002. doi: 10.1086/342655.

Diego Blas, Diana López Nacir, and Sergey Sibiryakov. Ultralight Dark Matter Resonates with Binary Pulsars. Phys. Rev. Lett., 118(26):261102, June 2017. doi: 10.1103/PhysRevLett.118.261102.

J. Blum and G. Wurm. The growth mechanisms of macroscopic bodies in protoplanetary disks. ARA&A, 46:21–56, September 2008. doi: 10.1146/annurev.astro.46.060407.145152.

Jürgen Blum. Dust Evolution in Protoplanetary Discs and the Formation of Planetesimals. What Have We Learned from Laboratory Experiments? Space Sci. Rev., 214(2):52, March 2018. doi: 10.1007/s11214-018-0486-5.

Peter Bodenheimer, Olenka Hubickyj, and Jack J. Lissauer. Models of the in Situ Formation of Detected Extrasolar Giant Planets<FNOTEREF RID="FN1">1</FNOTEREF>,. Icarus, 143(1):2–14, January 2000. doi: 10.1006/icar.1999.6246.

Ian A. Bonnell. The Formation of Close Binary Stars. In H. Zinnecker and R. Mathieu, editors, The Formation of Binary Stars, volume 200, page 23, January 2001.

Richard A. Booth, Farzana Meru, Man Hoi Lee, and Cathie J. Clarke. Breakthrough revisited: investigating the requirements for growth of dust beyond the bouncing barrier. MNRAS, 475(1):167–180, March 2018. doi: 10.1093/mnras/stx3084.

J. A. Burns, P. L. Lamy, and S. Soter. Radiation forces on small particles in the solar system. Icarus, 40(1):1–48, October 1979. doi: 10.1016/0019-1035(79)90050-2.

Robin M. Canup. A Giant Impact Origin of Pluto-Charon. Science, 307(5709):546–550, January 2005. doi: 10.1126/science.1106818.

Robin M. Canup. On a Giant Impact Origin of Charon, Nix, and Hydra. AJ, 141(2):35, February 2011. doi: 10.1088/0004-6256/141/2/35.

E. Carretta, A. Bragaglia, R. G. Gratton, S. Lucatello, G. Catanzaro, F. Leone, M. Bellazzini, R. Claudi, V. D'Orazi, and Y. Momany. Na-O anticorrelation and HB. VII. The chemical composition of first and second-generation stars in 15 globular clusters from GIRAFFE spectra. A&A, 505(1):117–138, Oct 2009. doi: 10.1051/0004-6361/200912096.

Marco Celoria, Roberto Oliveri, Alberto Sesana, and Michela Mapelli. Lecture notes on black hole binary astrophysics. arXiv e-prints, art. arXiv:1807.11489, July 2018. doi: 10.48550/arXiv.1807.11489.

Julio Chanamé and Andrew Gould. Disk and Halo Wide Binaries from the Revised Luyten Catalog: Probes of Star Formation and MACHO Dark Matter. ApJ, 601(1):289–310, January 2004. doi: 10.1086/380442.

S. Chandrasekhar. On the Stability of Binary Systems. ApJ, 99:54, January 1944. doi: 10.1086/144589.

E. Chiang and A. N. Youdin. Forming Planetesimals in Solar and Extrasolar Nebulae. Annual Review of Earth and Planetary Sciences, 38:493–522, May 2010. doi: 10.1146/annurev-earth-040809-152513.

J. W. Christy and R. S. Harrington. The satellite of Pluto. AJ, 83:1005, August 1978. doi: 10.1086/112284.

C. J. Clarke. Pseudo-viscous modelling of self-gravitating discs and the formation of low mass ratio binaries. MNRAS, 396(2):1066–1074, June 2009. doi: 10.1111/j.1365-2966.2009.14774.x.

C. J. Clarke and J. E. Pringle. Star-disc interactions and binary star formation. MNRAS, 249:584–587, April 1991. doi: 10.1093/mnras/249.4.584.

G. H. Darwin. On the Bodily Tides of Viscous and Semi-Elastic Spheroids, and on the Ocean Tides upon a Yielding Nucleus. Philosophical Transactions of the Royal Society of London Series I, 170:1–35, January 1879.

Rebekah I. Dawson and John Asher Johnson. Origins of Hot Jupiters. ARA&A, 56: 175–221, September 2018. doi: 10.1146/annurev-astro-081817-051853.

S. E. de Mink and I. Mandel. The chemically homogeneous evolutionary channel for binary black hole mergers: rates and properties of gravitational-wave events detectable by advanced LIGO. MNRAS, 460(4):3545–3553, August 2016. doi: 10.1093/mnras/stw1219.

Tunahan Demirci and Gerhard Wurm. Accretion of eroding pebbles and planetesimals in planetary envelopes. A&A, 641:A99, September 2020. doi: 10.1051/0004-6361/202038690.

Tunahan Demirci, Niclas Schneider, Tobias Steinpilz, Tabea Bogdan, Jens Teiser, and Gerhard Wurm. Planetesimals in rarefied gas: wind erosion in slip flow. MNRAS, 493(4):5456–5463, April 2020. doi: 10.1093/mnras/staa607.

Erik Dennihy, Siyi Xu, Samuel Lai, Amy Bonsor, J. C. Clemens, Patrick Dufour, Boris T. Gansicke, Nicola Pietro Gentile Fusillo, Francois Hardy, R. J. Hegedus, J. J. Hermes, B. C. Kaiser, Markus Kissler-Patig, Beth Klein, Christopher J. Manser, and Joshua S. Reding. Five New Post-Main-Sequence Debris Disks with Gaseous Emission. arXiv e-prints, art. arXiv:2010.03693, October 2020.

Steven J. Desch. Density of Charon formed from a disk generated by the impact of partially differentiated bodies. Icarus, 246:37–47, January 2015. doi: 10.1016/j.icarus.2014.07.034.

Michal Dominik, Krzysztof Belczynski, Christopher Fryer, Daniel E. Holz, Emanuele Berti, Tomasz Bulik, Ilya Mandel, and Richard O'Shaughnessy. Double Compact Objects. I. The Significance of the Common Envelope on Merger Rates. ApJ, 759 (1):52, November 2012. doi: 10.1088/0004-637X/759/1/52.

Paul C. Duffell, Daniel D'Orazio, Andrea Derdzinski, Zoltan Haiman, Andrew MacFadyen, Anna L. Rosen, and Jonathan Zrake. Circumbinary Disks: Accretion and Torque as a Function of Mass Ratio and Disk Viscosity. ApJ, 901(1):25, September 2020. doi: 10.3847/1538-4357/abab95.

Albert Einstein. Näherungsweise Integration der Feldgleichungen der Gravitation. Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, pages 688–696, January 1916.

Kareem El-Badry, Hans-Walter Rix, Haijun Tian, Gaspard Duchêne, and Maxwell Moe. Discovery of an equal-mass 'twin' binary population reaching 1000 + au separations. MNRAS, 489(4):5822–5857, November 2019. doi: 10.1093/mnras/stz2480.

Kareem El-Badry, Hans-Walter Rix, and Tyler M. Heintz. A million binaries from Gaia eDR3: sample selection and validation of Gaia parallax uncertainties. MNRAS, 506(2):2269–2295, September 2021. doi: 10.1093/mnras/stab323.

Daniel Fabrycky and Scott Tremaine. Shrinking Binary and Planetary Orbits by Kozai Cycles with Tidal Friction. ApJ, 669(2):1298–1315, November 2007. doi: 10.1086/521702.

J. Farihi. Circumstellar debris and pollution at white dwarf stars. New A Rev., 71: 9–34, April 2016. doi: 10.1016/j.newar.2016.03.001.

Giacomo Fragione and Fabio Antonini. Massive binary star mergers in galactic nuclei: implications for blue stragglers, binary S-stars, and gravitational waves. MNRAS, 488(1):728–738, September 2019. doi: 10.1093/mnras/stz1723.

Pascale Garaud, Farzana Meru, Marina Galvagni, and Christoph Olczak. From Dust to Planetesimals: An Improved Model for Collisional Growth in Protoplanetary Disks. ApJ, 764(2):146, February 2013. doi: 10.1088/0004-637X/764/2/146.

Yonadav Barry Ginat, Hila Glanz, Hagai B. Perets, Evgeni Grishin, and Vincent Desjacques. Gravitational waves from in-spirals of compact objects in binary common-envelope evolution. MNRAS, 493(4):4861–4867, April 2020. doi: 10.1093/mnras/staa465.

Sivan Ginzburg and Eugene Chiang. The endgame of gas giant formation: accretion luminosity and contraction post-runaway. MNRAS, 490(3):4334–4343, December 2019. doi: 10.1093/mnras/stz2901.

Sivan Ginzburg and Re'em Sari. Hot-Jupiter Inflation due to Deep Energy Deposition. ApJ, 803(2):111, April 2015. doi: 10.1088/0004-637X/803/2/111.

Sivan Ginzburg and Re'em Sari. Hot-Jupiter core mass from Roche lobe overflow. MNRAS, 469(1):278–285, July 2017. doi: 10.1093/mnras/stx832.

J. Girven, C. S. Brinkworth, J. Farihi, B. T. Gänsicke, D. W. Hoard, T. R. Marsh, and D. Koester. Constraints on the Lifetimes of Disks Resulting from Tidally Destroyed Rocky Planetary Bodies. ApJ, 749(2):154, April 2012. doi: 10.1088/0004-637X/749/2/154.

Hila Glanz, Mor Rozner, Hagai B. Perets, and Evgeni Grishin. Inflated Eccentric Migration of Evolving Gas Giants II - Numerical Methodology and Basic Concepts. ApJ, 931(1):11, May 2022. doi: 10.3847/1538-4357/ac6807.

P. Goldreich and S. Tremaine. Disk-satellite interactions. ApJ, 241:425–441, October 1980. doi: 10.1086/158356.

Peter Goldreich and Steven Soter. Q in the Solar System. Icarus, 5(1):375–389, January 1966. doi: 10.1016/0019-1035(66)90051-0.

Peter Goldreich, Yoram Lithwick, and Re'em Sari. Formation of Kuiper-belt binaries by dynamical friction and three-body encounters. Nature, 420(6916):643–646, December 2002a. doi: 10.1038/nature01227.

Peter Goldreich, Yoram Lithwick, and Re'em Sari. Formation of Kuiper-belt binaries by dynamical friction and three-body encounters. Nature, 420(6916):643–646, December 2002b. doi: 10.1038/nature01227.

Jeremy Goodman and Piet Hut. Binary–Single-Star Scattering. V. Steady State Binary Distribution in a Homogeneous Static Background of Single Stars. ApJ, 403:271, January 1993. doi: 10.1086/172200.

Evgeni Grishin and Hagai B. Perets. Application of Gas Dynamical Friction for Planetesimals. I. Evolution of Single Planetesimals. ApJ, 811(1):54, September 2015. doi: 10.1088/0004-637X/811/1/54.

Evgeni Grishin and Hagai B. Perets. Application of Gas Dynamical Friction for Planetesimals. II. Evolution of Binary Planetesimals. ApJ, 820(2):106, April 2016. doi: 10.3847/0004-637X/820/2/106.

Evgeni Grishin and Hagai B. Perets. Chaotic dynamics of wide triples induced by galactic tides: a novel channel for producing compact binaries, mergers, and collisions. MNRAS, 512(4):4993–5009, June 2022. doi: 10.1093/mnras/stac706.

Evgeni Grishin, Hagai B. Perets, and Giacomo Fragione. Quasi-secular evolution of mildly hierarchical triple systems: analytics and applications for GW sources and hot Jupiters. MNRAS, 481(4):4907–4923, December 2018. doi: 10.1093/mnras/sty2477.

Evgeni Grishin, Hagai B. Perets, and Yael Avni. Planet seeding through gas-assisted capture of interstellar objects. MNRAS, 487(3):3324–3332, August 2019. doi: 10.1093/mnras/stz1505.

Evgeni Grishin, Uri Malamud, Hagai B. Perets, Oliver Wand el, and Christoph M. Schaefer. Origin of (2014) MU69-like Kuiper-belt contact binaries from wide binaries. arXiv e-prints, art. arXiv:2003.01720, March 2020a.

Evgeni Grishin, Mor Rozner, and Hagai B. Perets. Erosion-driven Size Redistribution of Protoplanetary Disk Solids and the Onset of Streaming Instability and Pebble Accretion. ApJ, 898(1):L13, July 2020b. doi: 10.3847/2041-8213/aba266.

T. Guillot, A. Burrows, W. B. Hubbard, J. I. Lunine, and D. Saumon. Giant Planets at Small Orbital Distances. ApJ, 459:L35, March 1996. doi: 10.1086/309935.

Jr. Haisch, Karl E., Elizabeth A. Lada, and Charles J. Lada. Disk Frequencies and Lifetimes in Young Clusters. ApJ, 553(2):L153–L156, June 2001. doi: 10.1086/320685.

Adrian S. Hamers. Secular dynamics of hierarchical multiple systems composed of nested binaries, with an arbitrary number of bodies and arbitrary hierarchical structure - II. External perturbations: flybys and supernovae. MNRAS, 476(3):4139–4161, May 2018. doi: 10.1093/mnras/sty428.

Adrian S. Hamers. Secular dynamics of hierarchical multiple systems composed of nested binaries, with an arbitrary number of bodies and arbitrary hierarchical structure - III. Suborbital effects: hybrid integration techniques and orbit-averaging corrections. MNRAS, 494(4):5492–5506, April 2020. doi: 10.1093/mnras/staa1084.

Adrian S. Hamers and Simon F. Portegies Zwart. Secular dynamics of hierarchical multiple systems composed of nested binaries, with an arbitrary number of bodies and arbitrary hierarchical structure. First applications to multiplanet and multistar systems. MNRAS, 459(3):2827–2874, July 2016. doi: 10.1093/mnras/stw784.

Adrian S. Hamers and Todd A. Thompson. The Impact of White Dwarf Natal Kicks and Stellar Flybys on the Rates of Type Ia Supernovae in Triple-star Systems. ApJ, 882(1):24, September 2019. doi: 10.3847/1538-4357/ab321f.

Adrian S. Hamers and Scott Tremaine. Hot Jupiters Driven by High-eccentricity Migration in Globular Clusters. AJ, 154(6):272, December 2017. doi: 10.3847/1538-3881/aa9926.

Adrian S. Hamers, Hagai B. Perets, Fabio Antonini, and Simon F. Portegies Zwart. Secular dynamics of hierarchical quadruple systems: the case of a triple system orbited by a fourth body. MNRAS, 449(4):4221–4245, June 2015. doi: 10.1093/mnras/stv452.

Adrian S. Hamers, Fabio Antonini, Yoram Lithwick, Hagai B. Perets, and Simon F. Portegies Zwart. Secular dynamics of multiplanet systems: implications for the formation of hot and warm Jupiters via high-eccentricity migration. MNRAS, 464(1):688–701, January 2017. doi: 10.1093/mnras/stw2370.

R. S. Harrington. Dynamical evolution of triple stars. AJ, 73:190–194, April 1968. doi: 10.1086/110614.

C. Hayashi. Structure of the Solar Nebula, Growth and Decay of Magnetic Fields and Effects of Magnetic and Turbulent Viscosities on the Nebula. Progress of Theoretical Physics Supplement, 70:35–53, January 1981. doi: 10.1143/PTPS.70.35.

D. C. Heggie. Binary evolution in stellar dynamics. MNRAS, 173:729–787, December 1975. doi: 10.1093/mnras/173.3.729.

J. G. Hills. Encounters between binary and single stars and their effect on the dynamical evolution of stellar systems. AJ, 80:809–825, October 1975. doi: 10.1086/111815.

P. Hut. Tidal evolution in close binary systems. A&A, 99:126–140, June 1981.

P. Hut. Binaries as a heat source in stellar dynamics - Release of binding energy. ApJ, 272:L29–L33, September 1983. doi: 10.1086/184111.

K. A. Innanen, J. Q. Zheng, S. Mikkola, and M. J. Valtonen. The Kozai Mechanism and the Stability of Planetary Orbits in Binary Star Systems. AJ, 113:1915, May 1997. doi: 10.1086/118405.

P. B. Ivanov and J. C. B. Papaloizou. On the tidal interaction of massive extrasolar planets on highly eccentric orbits. MNRAS, 347(2):437–453, January 2004. doi: 10.1111/j.1365-2966.2004.07238.x.

P. B. Ivanov and J. C. B. Papaloizou. Dynamic tides in rotating objects: orbital circularization of extrasolar planets for realistic planet models. MNRAS, 376(2): 682–704, April 2007. doi: 10.1111/j.1365-2966.2007.11463.x.

Yan-Fei Jiang and Scott Tremaine. The evolution of wide binary stars. MNRAS, 401: 977–994, January 2010. doi: 10.1111/j.1365-2966.2009.15744.x.

Anders Johansen, Jeffrey S. Oishi, Mordecai-Mark Mac Low, Hubert Klahr, Thomas Henning, and Andrew Youdin. Rapid planetesimal formation in turbulent circumstellar disks. Nature, 448(7157):1022–1025, August 2007. doi: 10.1038/nature06086.

M. Jura. Pollution of Single White Dwarfs by Accretion of Many Small Asteroids. AJ, 135(5):1785–1792, May 2008. doi: 10.1088/0004-6256/135/5/1785.

Nathan A. Kaib and S. N. Raymond. Stellar Collisions within Very Wide Binaries. art. 422.05, January 2014.

Scott J. Kenyon and Benjamin C. Bromley. Numerical Simulations of Collisional Cascades at the Roche Limits of White Dwarf Stars. ApJ, 844(2):116, August 2017a. doi: 10.3847/1538-4357/aa7b85.

Scott J. Kenyon and Benjamin C. Bromley. Numerical Simulations of Gaseous Disks Generated from Collisional Cascades at the Roche Limits of White Dwarf Stars. ApJ, 850(1):50, November 2017b. doi: 10.3847/1538-4357/aa9570.

Hiroshi Kinoshita and Hiroshi Nakai. Analytical Solution of the Kozai Resonance and its Application. Celestial Mechanics and Dynamical Astronomy, 75(2):125–147, October 1999. doi: 10.1023/A:1008321310187.

D. Koester, B. T. Gänsicke, and J. Farihi. The frequency of planetary debris around young white dwarfs. A&A, 566:A34, June 2014. doi: 10.1051/0004-6361/201423691.

Thaddeus D. Komacek and Andrew N. Youdin. Structure and Evolution of Internally Heated Hot Jupiters. ApJ, 844(2):94, August 2017. doi: 10.3847/1538-4357/aa7b75.

Thaddeus D. Komacek, Daniel P. Thorngren, Eric D. Lopez, and Sivan Ginzburg. Reinflation of Warm and Hot Jupiters. ApJ, 893(1):36, April 2020. doi: 10.3847/1538-4357/ab7eb4.

Yoshihide Kozai. Secular perturbations of asteroids with high inclination and eccentricity. AJ, 67:591–598, November 1962. doi: 10.1086/108790.

Leonardo Krapp, Pablo Benítez-Llambay, Oliver Gressel, and Martin E. Pessah. Streaming Instability for Particle-size Distributions. ApJ, 878(2):L30, June 2019. doi: 10.3847/2041-8213/ab2596.

K. M. Kratter. The Formation of Close Binaries. In L. Schmidtobreick, M. R. Schreiber, and C. Tappert, editors, Evolution of Compact Binaries, volume 447 of Astronomical Society of the Pacific Conference Series, page 47, September 2011. doi: 10.48550/arXiv.1109.3740.

Kaitlin Kratter and Giuseppe Lodato. Gravitational Instabilities in Circumstellar Disks. ARA&A, 54:271–311, September 2016. doi: 10.1146/annurev-astro-081915-023307.

Maximilian Kruss, Tunahan Demirci, Marc Koester, Thorben Kelling, and Gerhard Wurm. Failed Growth at the Bouncing Barrier in Planetesimal Formation. ApJ, 827(2):110, August 2016. doi: 10.3847/0004-637X/827/2/110.

Maximilian Kruss, Grzegorz Musiolik, Tunahan Demirci, Gerhard Wurm, and Jens Teiser. Wind erosion on Mars and other small terrestrial planets. arXiv e-prints, art. arXiv:1911.01692, Nov 2019.

Girish Kulkarni and Abraham Loeb. Formation of galactic nuclei with multiple supermassive black holes at high redshifts. MNRAS, 422(2):1306–1323, May 2012. doi: 10.1111/j.1365-2966.2012.20699.x.

Charles J. Lada and Elizabeth A. Lada. Embedded Clusters in Molecular Clouds. ARA&A, 41:57–115, January 2003. doi: 10.1146/annurev.astro.41.011802.094844.

Dong Lai. Dynamical Tides in Rotating Binary Stars. ApJ, 490(2):847–862, December 1997. doi: 10.1086/304899.

M. Lambrechts and A. Johansen. Rapid growth of gas-giant cores by pebble accretion. A&A, 544:A32, August 2012. doi: 10.1051/0004-6361/201219127.

Yueh-Ning Lee, Stella S. R. Offner, Patrick Hennebelle, Philippe André, Hans Zinnecker, Javier Ballesteros-Paredes, Shu-ichiro Inutsuka, and J. M. Diederik Kruijssen. The Origin of the Stellar Mass Distribution and Multiplicity. Space Sci. Rev., 216(4):70, June 2020. doi: 10.1007/s11214-020-00699-2.

Nathan W. C. Leigh, Alessandra Mastrobuono-Battisti, Hagai B. Perets, and Torsten Böker. Stellar dynamics in gas: the role of gas damping. MNRAS, 441(2):919–932, June 2014. doi: 10.1093/mnras/stu622.

Zoë M. Leinhardt and Sarah T. Stewart. Collisions between Gravity-dominated Bodies. I. Outcome Regimes and Scaling Laws. ApJ, 745(1):79, January 2012. doi: 10.1088/0004-637X/745/1/79.

Rixin Li, Andrew N. Youdin, and Jacob B. Simon. Demographics of Planetesimals Formed by the Streaming Instability. ApJ, 885(1):69, November 2019. doi: 10.3847/1538-4357/ab480d.

M. L. Lidov. The evolution of orbits of artificial satellites of planets under the action of gravitational perturbations of external bodies. Planet. Space Sci., 9(10):719–759, October 1962. doi: 10.1016/0032-0633(62)90129-0.

D. N. C. Lin and John Papaloizou. On the Tidal Interaction between Protoplanets and the Protoplanetary Disk. III. Orbital Migration of Protoplanets. ApJ, 309:846, October 1986. doi: 10.1086/164653.

Liantong Luo, Boaz Katz, and Subo Dong. Double-averaging can fail to characterize the long-term evolution of Lidov-Kozai Cycles and derivation of an analytical correction. MNRAS, 458(3):3060–3074, May 2016. doi: 10.1093/mnras/stw475.

Wladimir Lyra, Andrew N. Youdin, and Anders Johansen. Evolution of MU69 from a binary planetesimal into contact by Kozai-Lidov oscillations and nebular drag. arXiv e-prints, art. arXiv:2003.00670, March 2020.

Christopher J. Manser, Boris T. Gänsicke, Nicola Pietro Gentile Fusillo, Richard Ashley, Elmé Breedt, Mark Hollands, Paula Izquierdo, and Ingrid Pelisoli. The frequency of gaseous debris discs around white dwarfs. MNRAS, 493(2):2127–2139, April 2020. doi: 10.1093/mnras/staa359.

Rosemary A. Mardling. The Role of Chaos in the Circularization of Tidal Capture Binaries. I. The Chaos Boundary. ApJ, 450:722, September 1995a. doi: 10.1086/176178.

Rosemary A. Mardling. The Role of Chaos in the Circularization of Tidal Capture Binaries. II. Long-Time Evolution. ApJ, 450:732, September 1995b. doi: 10.1086/176179.

Michel Mayor and Didier Queloz. A Jupiter-mass companion to a solar-type star. Nature, 378(6555):355–359, November 1995. doi: 10.1038/378355a0.

B. McKernan, K. E. S. Ford, W. Lyra, and H. B. Perets. Intermediate mass black holes in AGN discs - I. Production and growth. MNRAS, 425(1):460–469, September 2012. doi: 10.1111/j.1365-2966.2012.21486.x.

W. B. McKinnon. On the origin of Triton and Pluto. Nature, 311(5984):355–358, September 1984. doi: 10.1038/311355a0.

W. B. McKinnon. On the Origin of the Pluto-Charon Binary. ApJ, 344:L41, September 1989. doi: 10.1086/185526.

Carl Melis, B. Klein, A. E. Doyle, A. J. Weinberger, B. Zuckerman, and P. Dufour. Serendipitous Discovery of Nine White Dwarfs With Gaseous Debris Disks. arXiv e-prints, art. arXiv:2010.03695, October 2020.

Brian D. Metzger, Roman R. Rafikov, and Konstantin V. Bochkarev. Global models of runaway accretion in white dwarf debris discs. MNRAS, 423(1):505–528, June 2012. doi: 10.1111/j.1365-2966.2012.20895.x.

Erez Michaely and Smadar Naoz. Ultrawide Black Hole-Neutron Star Binaries as a Possible Source for Gravitational Waves and Short Gamma-Ray Bursts. ApJ, 936 (2):184, September 2022. doi: 10.3847/1538-4357/ac8a92.

Erez Michaely and Hagai B. Perets. Tidal capture formation of low-mass X-ray binaries from wide binaries in the field. MNRAS, 458:4188–4197, June 2016. doi: 10.1093/mnras/stw368.

Erez Michaely and Hagai B. Perets. Gravitational-wave Sources from Mergers of Binary Black Holes Catalyzed by Flyby Interactions in the Field. ApJ, 887(2):L36, December 2019. doi: 10.3847/2041-8213/ab5b9b.

N. Miller, J. J. Fortney, and B. Jackson. Inflating and Deflating Hot Jupiters: Coupled Tidal and Thermal Evolution of Known Transiting Planets. ApJ, 702(2):1413–1427, September 2009. doi: 10.1088/0004-637X/702/2/1413.

Maxwell Moe and Kaitlin M. Kratter. Dynamical Formation of Close Binaries during the Pre-main-sequence Phase. ApJ, 854(1):44, February 2018. doi: 10.3847/1538-4357/aaa6d2.

N. Moeckel and C. J. Clarke. The formation of permanent soft binaries in dispersing clusters. MNRAS, 415:1179–1187, August 2011. doi: 10.1111/j.1365-2966.2011.18731.x.

Alessandro Morbidelli, William F. Bottke, David Nesvorný, and Harold F. Levison. Asteroids were born big. Icarus, 204(2):558–573, December 2009. doi: 10.1016/j.icarus.2009.07.011.

Walter Heinrich Munk and Gordon J. F. MacDonald. The rotation of the earth; a geophysical discussion. 1960.

Smadar Naoz. The Eccentric Kozai-Lidov Effect and Its Applications. ARA&A, 54:441–489, September 2016. doi: 10.1146/annurev-astro-081915-023315.

Smadar Naoz and Daniel C. Fabrycky. Mergers and Obliquities in Stellar Triples. ApJ, 793(2):137, October 2014. doi: 10.1088/0004-637X/793/2/137.

Smadar Naoz, Will M. Farr, Yoram Lithwick, Frederic A. Rasio, and Jean Teyssandier. Hot Jupiters from secular planet-planet interactions. Nature, 473(7346):187–189, May 2011a. doi: 10.1038/nature10076.

Smadar Naoz, Will M. Farr, Yoram Lithwick, Frederic A. Rasio, and Jean Teyssandier. Hot Jupiters from secular planet-planet interactions. Nature, 473(7346):187–189, May 2011b. doi: 10.1038/nature10076.

Smadar Naoz, Will M. Farr, and Frederic A. Rasio. On the Formation of Hot Jupiters in Stellar Binaries. ApJ, 754(2):L36, August 2012. doi: 10.1088/2041-8205/754/2/L36.

David Nesvorný, Rixin Li, Andrew N. Youdin, Jacob B. Simon, and William M. Grundy. Trans-Neptunian binaries as evidence for planetesimal formation by the streaming instability. Nature Astronomy, 3:808–812, June 2019. doi: 10.1038/s41550-019-0806-z.

Stella S. R. Offner, Kaitlin M. Kratter, Christopher D. Matzner, Mark R. Krumholz, and Richard I. Klein. The Formation of Low-mass Binary Star Systems Via Turbulent Fragmentation. ApJ, 725(2):1485–1494, December 2010. doi: 10.1088/0004-637X/725/2/1485.

Gordon I. Ogilvie. Tidal Dissipation in Stars and Giant Planets. ARA&A, 52:171–210, August 2014. doi: 10.1146/annurev-astro-081913-035941.

Ryan M. O'Leary, Frederic A. Rasio, John M. Fregeau, Natalia Ivanova, and Richard O'Shaughnessy. Binary Mergers and Growth of Black Holes in Dense Star Clusters. ApJ, 637(2):937–951, February 2006. doi: 10.1086/498446.

Ryan M. O'Leary, Bence Kocsis, and Abraham Loeb. Gravitational waves from scattering of stellar-mass black holes in galactic nuclei. MNRAS, 395(4):2127–2146, June 2009. doi: 10.1111/j.1365-2966.2009.14653.x.

C. W. Ormel and H. H. Klahr. The effect of gas drag on the growth of protoplanets. Analytical expressions for the accretion of small bodies in laminar disks. A&A, 520:A43, September 2010. doi: 10.1051/0004-6361/201014903.

Eve C. Ostriker. Dynamical Friction in a Gaseous Medium. ApJ, 513(1):252–258, March 1999. doi: 10.1086/306858.

Georgi B. Paraskov, Gerhard Wurm, and Oliver Krauss. Eolian Erosion of Dusty Bodies in Protoplanetary Disks. ApJ, 648(2):1219–1227, Sep 2006. doi: 10.1086/506138.

Bill Paxton, Lars Bildsten, Aaron Dotter, Falk Herwig, Pierre Lesaffre, and Frank Timmes. Modules for Experiments in Stellar Astrophysics (MESA). ApJS, 192(1): 3, January 2011. doi: 10.1088/0067-0049/192/1/3.

Hagai B. Perets. Binary Planetesimals and Their Role in Planet Formation. ApJ, 727(1):L3, January 2011. doi: 10.1088/2041-8205/727/1/L3.

Hagai B. Perets and Daniel C. Fabrycky. On the Triple Origin of Blue Stragglers. ApJ, 697(2):1048–1056, June 2009. doi: 10.1088/0004-637X/697/2/1048.

Hagai B. Perets and M. B. N. Kouwenhoven. On the Origin of Planets at Very Wide Orbits from the Recapture of Free Floating Planets. ApJ, 750:83, May 2012. doi: 10.1088/0004-637X/750/1/83.

Hagai B. Perets and Ruth A. Murray-Clay. Wind-shearing in Gaseous Protoplanetary Disks and the Evolution of Binary Planetesimals. ApJ, 733(1):56, May 2011. doi: 10.1088/0004-637X/733/1/56.

Hagai B. Perets and Smadar Naoz. Kozai Cycles, Tidal Friction, and the Dynamical Evolution of Binary Minor Planets. ApJ, 699(1):L17–L21, July 2009. doi: 10.1088/0004-637X/699/1/L17.

P. C. Peters. Gravitational Radiation and the Motion of Two Point Masses. Physical Review, 136(4B):1224–1232, November 1964. doi: 10.1103/PhysRev.136.B1224.

Cristobal Petrovich. Steady-state Planet Migration by the Kozai-Lidov Mechanism in Stellar Binaries. ApJ, 799(1):27, January 2015a. doi: 10.1088/0004-637X/799/1/27.

Cristobal Petrovich. Hot Jupiters from Coplanar High-eccentricity Migration. ApJ, 805(1):75, May 2015b. doi: 10.1088/0004-637X/805/1/75.

Simon Portegies Zwart, Steve McMillan, Stefan Harfst, Derek Groen, Michiko Fujii, Breanndán Ó. Nualláin, Evert Glebbeek, Douglas Heggie, James Lombardi, Piet Hut, Vangelis Angelou, Sambaran Banerjee, Houria Belkus, Tassos Fragos, John Fregeau, Evghenii Gaburov, Rob Izzard, Mario Jurić, Stephen Justham, Andrea Sottoriva, Peter Teuben, Joris van Bever, Ofer Yaron, and Marcel Zemp. A multiphysics and multiscale software environment for modeling astrophysical systems. New A, 14(4):369–378, May 2009. doi: 10.1016/j.newast.2008.10.006.

Simon F. Portegies Zwart and Stephen L. W. McMillan. Black Hole Mergers in the Universe. ApJ, 528(1):L17–L20, January 2000. doi: 10.1086/312422.

W. H. Press and S. A. Teukolsky. On formation of close binaries by two-body tidal capture. ApJ, 213:183–192, April 1977. doi: 10.1086/155143.

Roman R. Rafikov. Can Giant Planets Form by Direct Gravitational Instability? ApJ, 621(1):L69–L72, March 2005. doi: 10.1086/428899.

Roman R. Rafikov. Runaway accretion of metals from compact discs of debris on to white dwarfs. MNRAS, 416(1):L55–L59, September 2011a. doi: 10.1111/j.1745-3933.2011.01096.x.

Roman R. Rafikov. Metal Accretion onto White Dwarfs Caused by Poynting-Robertson Drag on their Debris Disks. ApJ, 732(1):L3, May 2011b. doi: 10.1088/2041-8205/732/1/L3.

Deepak Raghavan, Harold A. McAlister, Todd J. Henry, David W. Latham, Geoffrey W. Marcy, Brian D. Mason, Douglas R. Gies, Russel J. White, and Theo A. ten Brummelaar. A Survey of Stellar Families: Multiplicity of Solar-type Stars. ApJS, 190(1):1–42, Sep 2010. doi: 10.1088/0067-0049/190/1/1.

Frederic A. Rasio and Eric B. Ford. Dynamical instabilities and the formation of extrasolar planetary systems. Science, 274:954–956, November 1996. doi: 10.1126/science.274.5289.954.

Alexander Rasskazov and Bence Kocsis. The Rate of Stellar Mass Black Hole Scattering in Galactic Nuclei. ApJ, 881(1):20, August 2019. doi: 10.3847/1538-4357/ab2c74.

H. Rein and S. F. Liu. REBOUND: an open-source multi-purpose N-body code for collisional dynamics. A&A, 537:A128, January 2012a. doi: 10.1051/0004-6361/201118085.

H. Rein and S. F. Liu. REBOUND: an open-source multi-purpose N-body code for collisional dynamics. A&A, 537:A128, Jan 2012b. doi: 10.1051/0004-6361/201118085.

Hanno Rein and David S. Spiegel. IAS15: a fast, adaptive, high-order integrator for gravitational dynamics, accurate to machine precision over a billion orbits. MNRAS, 446(2):1424–1437, January 2015. doi: 10.1093/mnras/stu2164.

Mor Rozner and Hagai B. Perets. Binary Evolution, Gravitational-wave Mergers, and Explosive Transients in Multiple-population Gas-enriched Globular Clusters. ApJ, 931(2):149, June 2022a. doi: 10.3847/1538-4357/ac6d55.

Mor Rozner and Hagai B. Perets. Binary Evolution, Gravitational-wave Mergers, and Explosive Transients in Multiple-population Gas-enriched Globular Clusters. ApJ, 931(2):149, June 2022b. doi: 10.3847/1538-4357/ac6d55.

Mor Rozner and Hagai B. Perets. Binary Evolution, Gravitational-wave Mergers, and Explosive Transients in Multiple-population Gas-enriched Globular Clusters. ApJ, 931(2):149, June 2022c. doi: 10.3847/1538-4357/ac6d55.

Mor Rozner and Hagai B. Perets. Born to Be Wide: The Distribution of Wide Binaries in the Field and Soft Binaries in Clusters. ApJ, 955(2):134, October 2023. doi: 10.3847/1538-4357/ace2c6.

Mor Rozner and Hagai B. Perets. Soft no more: gas shielding protects soft binaries from disruption in gas-rich environments. arXiv e-prints, art. arXiv:2404.01384, April 2024. doi: 10.48550/arXiv.2404.01384.

Mor Rozner, Evgeni Grishin, Yonadav Barry Ginat, Andrei P. Igoshev, and Vincent Desjacques. Axion resonances in binary pulsar systems. J. Cosmology Astropart. Phys., 2020(3):061, March 2020a. doi: 10.1088/1475-7516/2020/03/061.

Mor Rozner, Evgeni Grishin, and Hagai B. Perets. The aeolian-erosion barrier for the growth of metre-size objects in protoplanetary discs. MNRAS, 496(4):4827–4835, June 2020b. doi: 10.1093/mnras/staa1864.

Mor Rozner, Evgeni Grishin, and Hagai B. Perets. The wide-binary origin of the Pluto-Charon system. MNRAS, 497(4):5264–5270, August 2020c. doi: 10.1093/mnras/staa2446.

Mor Rozner, Dimitri Veras, and Hagai B. Perets. Rapid destruction of planetary debris around white dwarfs through aeolian erosion. MNRAS, 502(4):5176–5184, April 2021. doi: 10.1093/mnras/stab329.

Mor Rozner, Hila Glanz, Hagai B. Perets, and Evgeni Grishin. Inflated Eccentric Migration of Evolving Gas Giants I - Accelerated Formation and Destruction of Hot and Warm Jupiters. ApJ, 931(1):10, May 2022. doi: 10.3847/1538-4357/ac6808.

Mor Rozner, Aleksey Generozov, and Hagai B. Perets. Binary formation through gas-assisted capture and the implications for stellar, planetary, and compact object evolution. MNRAS, 521(1):866–880, May 2023. doi: 10.1093/mnras/stad603.

V. S. Safronov. Evolution of the protoplanetary cloud and formation of the earth and planets. 1972.

Johan Samsing, Morgan MacLeod, and Enrico Ramirez-Ruiz. The Formation of Eccentric Compact Binary Inspirals and the Role of Gravitational Wave Emission in Binary-Single Stellar Encounters. ApJ, 784(1):71, March 2014. doi: 10.1088/0004-637X/784/1/71.

Johan Samsing, Daniel J. D'Orazio, Kyle Kremer, Carl L. Rodriguez, and Abbas Askar. Single-single gravitational-wave captures in globular clusters: Eccentric deci-Hertz sources observable by DECIGO and Tian-Qin. Phys. Rev. D, 101(12):123010, June 2020. doi: 10.1103/PhysRevD.101.123010.

Noemi Schaffer, Anders Johansen, Lukas Cedenblad, Bernhard Mehling, and Dhrubaditya Mitra. Erosion of planetesimals by gas flow. A&A, 639:A39, July 2020. doi: 10.1051/0004-6361/201935763.

Rainer Schräpler and Jürgen Blum. The Physics of Protoplanetesimal Dust Agglomerates. VI. Erosion of Large Aggregates as a Source of Micrometer-sized Particles. ApJ, 734(2):108, June 2011. doi: 10.1088/0004-637X/734/2/108.

Yasuhito Sekine, Hidenori Genda, Shunichi Kamata, and Taro Funatsu. The Charon-forming giant impact as a source of Pluto's dark equatorial regions. Nature Astronomy, 1:0031, January 2017. doi: 10.1038/s41550-016-0031.

Yaping. Shao. Physics and Modelling of Wind Erosion. Atmospheric and Oceanographic Sciences Library, 37. Springer Netherlands, Dordrecht, 2nd ed. edition, 2008. ISBN 1-4020-8895-7.

Yaping Shao and Hua Lu. A simple expression for wind erosion threshold friction velocity. J. Geophys. Res., 105(D17):22,437–22,443, Sep 2000. doi: 10.1029/2000JD900304.

Kedron Silsbee and Scott Tremaine. Lidov-Kozai Cycles with Gravitational Radiation: Merging Black Holes in Isolated Triple Systems. ApJ, 836(1):39, February 2017. doi: 10.3847/1538-4357/aa5729.

Edwin Sirko and Jeremy Goodman. Spectral energy distributions of marginally self-gravitating quasi-stellar object discs. MNRAS, 341(2):501–508, May 2003. doi: 10.1046/j.1365-8711.2003.06431.x.

Lyman Spitzer. Dynamical evolution of globular clusters. 1987.

Nicholas C. Stone, Brian D. Metzger, and Zoltán Haiman. Assisted inspirals of stellar mass black holes embedded in AGN discs: solving the 'final au problem'. MNRAS, 464(1):946–954, January 2017. doi: 10.1093/mnras/stw2260.

Hiromichi Tagawa, Zoltán Haiman, and Bence Kocsis. Formation and Evolution of Compact-object Binaries in AGN Disks. ApJ, 898(1):25, July 2020. doi: 10.3847/1538-4357/ab9b8c.

Daniel Tamayo, Hanno Rein, Pengshuai Shi, and David M. Hernand ez. REBOUNDx: a library for adding conservative and dissipative forces to otherwise symplectic N-body integrations. MNRAS, 491(2):2885–2901, Jan 2020. doi: 10.1093/mnras/stz2870.

Jean Teyssandier, Dong Lai, and Michelle Vick. Formation of hot Jupiters through secular chaos and dynamical tides. MNRAS, 486(2):2265–2280, June 2019. doi: 10.1093/mnras/stz1011.

Todd A. Thompson. Accelerating Compact Object Mergers in Triple Systems with the Kozai Resonance: A Mechanism for "Prompt" Type Ia Supernovae, Gamma-Ray Bursts, and Other Exotica. ApJ, 741(2):82, November 2011. doi: 10.1088/0004-637X/741/2/82.

Andrei Tokovinin. From Binaries to Multiples. I. Data on F and G Dwarfs within 67 pc of the Sun. AJ, 147(4):86, April 2014a. doi: 10.1088/0004-6256/147/4/86.

Andrei Tokovinin. From Binaries to Multiples. II. Hierarchical Multiplicity of F and G Dwarfs. AJ, 147(4):87, April 2014b. doi: 10.1088/0004-6256/147/4/87.

Clyde W. Tombaugh. The Search for the Ninth Planet, Pluto. Leaflet of the Astronomical Society of the Pacific, 5(209):73, January 1946.

Silvia Toonen, Adrian Hamers, and Simon Portegies Zwart. The evolution of hierarchical triple star-systems. Computational Astrophysics and Cosmology, 3(1):6, December 2016. doi: 10.1186/s40668-016-0019-0.

Dimitri Veras. Post-main-sequence planetary system evolution. Royal Society Open Science, 3:150571, February 2016. doi: 10.1098/rsos.150571.

Dimitri Veras and Kevin Heng. The lifetimes of planetary debris discs around white dwarfs. MNRAS, 496(2):2292–2308, June 2020. doi: 10.1093/mnras/staa1632.

Michelle Vick and Dong Lai. Dynamical tides in highly eccentric binaries: chaos, dissipation, and quasi-steady state. MNRAS, 476(1):482–495, May 2018. doi: 10.1093/mnras/sty225.

Michelle Vick, Dong Lai, and Kassandra R. Anderson. Chaotic tides in migrating gas giants: forming hot and transient warm Jupiters via Lidov-Kozai migration. MNRAS, 484(4):5645–5668, April 2019. doi: 10.1093/mnras/stz354.

H Von Zeipel. Arkiv för matematik. Astronomi och Fysik, 11(1), 1916.

S. J. Weidenschilling. Aerodynamics of solid bodies in the solar nebula. MNRAS, 180:57–70, July 1977. doi: 10.1093/mnras/180.2.57.

Stuart J. Weidenschilling and Francesco Marzari. Gravitational scattering as a possible origin for giant planets at small stellar distances. Nature, 384(6610):619–621, December 1996. doi: 10.1038/384619a0.

F. L. Whipple. On certain aerodynamic processes for asteroids and comets. In Aina Elvius, editor, From Plasma to Planet, page 211, January 1972.

F. L. Whipple. Radial Pressure in the Solar Nebula as Affecting the Motions of Planetesimals, volume 319, page 355. 1973.

F. Windmark, T. Birnstiel, C. Güttler, J. Blum, C. P. Dullemond, and Th. Henning. Planetesimal formation by sweep-up: how the bouncing barrier can be beneficial to growth. A&A, 540:A73, April 2012a. doi: 10.1051/0004-6361/201118475.

F. Windmark, T. Birnstiel, C. W. Ormel, and C. P. Dullemond. Breaking through: The effects of a velocity distribution on barriers to dust growth. A&A, 544:L16, August 2012b. doi: 10.1051/0004-6361/201220004.

Y. Wu and N. Murray. Planet Migration and Binary Companions: The Case of HD 80606b. ApJ, 589(1):605–614, May 2003. doi: 10.1086/374598.

Gerhard Wurm and Jürgen Blum. Experiments on Preplanetary Dust Aggregation. Icarus, 132(1):125–136, March 1998. doi: 10.1006/icar.1998.5891.

Andrew N. Youdin and Jeremy Goodman. Streaming Instabilities in Protoplanetary Disks. ApJ, 620(1):459–469, February 2005. doi: 10.1086/426895.

J. P. Zahn. Reprint of 1977A&A....57..383Z. Tidal friction in close binary stars. A&A, 500:121–132, May 1977.

A. Zsom, C. W. Ormel, C. Güttler, J. Blum, and C. P. Dullemond. The outcome of protoplanetary dust growth: pebbles, boulders, or planetesimals? II. Introducing the bouncing barrier. A&A, 513:A57, April 2010. doi: 10.1051/0004-6361/200912976.

B. Zuckerman, C. Melis, B. Klein, D. Koester, and M. Jura. Ancient Planetary Systems are Orbiting a Large Fraction of White Dwarf Stars. ApJ, 722(1):725–736, October 2010. doi: 10.1088/0004-637X/722/1/725.

זוגות קשים על ידי הפיכתם של זוגות רכים לקשים. מידול התהליכים המתוארים מעלה יכול לגרום לשינוי תפיסתי בהבנתנו את הדינמיקה של צבירים עתירי גז וסביבות עתירות גז באופן כללי, ולהוביל להבנה טובה יותר של מאפייני האוכלוסיות באיזורים אלו.

כמו כן, גזרתי את ההתפלגות של זוגות רכים בצבירים וזוגות מרוחקים בשדות כוכבים, בעזרת כלים סטטיסטיים. הודות להסכמה הטובה בין התוצאות האנליטיות שלנו ותוצרות מסימולציות נומריות רב-גופיות ברוב המקרים, נוכל להשתמש בעתיד בתוצאות כדי להחליף סימולציות נומריות הדורשות זמן חישוב רב וכן לקבל הבנה טובה יותר של מאפייני הזוגות הללו.

תקציר

במחקר המתואר בעבודת גמר זו התמקדתי בשלושה כיוונים מרכזיים, הכרוכים זה בזה ברעיונות משותפים, כלים ושיטות. אדון בהווצרות ודינמיקה של כוכבי-לכת, דינמיקה בסביבות רוויות גז והתפלגויות של זוגות רכים בצבירים וזוגות מרוחקים בשדות כוכביים. היווצרות כוכבי-לכת היא מסע ארוך ומלא אתגרים, וכולל כמה סדרי גודל, החל מגרגירי אבק קטנים ועד לגודל של כוכבי-לכת. באחד ממחקריי, חקרתי מחסום חדש בהיווצרות פלנטות - 'שחיקה רוחית' - אותו מידלתי בהשראת מודלים בגיאולוגיה. הראיתי כי תהליך זה יכול להוביל לכמה אפקטים משמעותיים שישפיעו גם על תהליכים אחרים בשלבים הראשונים של יצירת כוכבי-לכת, כמו גם לעיצוב מחדש של התפלגות הגדלים של עצמים בדסקה הקדם כוכבית. לאחר מכן, דנתי בשלבים מאוחרים יותר של התפתחות כוכבי-לכת ובהיווצרות של הצד פלוטו וירחו כארון מזוג מרוחק. זהו מכניזם כללי שיכול פוטנציאלית לתאר את היווצרותם של זוגות נוספים בחגורת קויפר. כמו כן, עסקתי גם בהתפתחות התרמית-דינמית של צדקים חמים וחמימים. בחנתי הגירה באקסנטריות גבוהה של ענקי-גז שהתחילו את דרכם מנופחים, ועקבתי אחרי התכווצותם בעזרת מודל תרמי, ועקבתי אחר השפעת צימוד זה על התכונות והכמויות של הצדקים החמים והחמימים הנוצרים. הראיתי כי האבולוציה התרמית היא מרכיב חשוב שמשפיע באופן משמעותי על ההתפתחות הדינמית ועל המאפיינים של צדקים לאחר היווצרם.

זוגות נפוצים בסקאלות אסטרופיסיקליות רבות ומגוונות, ובעוד שהם נחקרו באופן מעמיק עבור סביבות נטולות גז, ישנן תופעות שטרם נחקרו שייחודיות לסביבות רוויות גז. חקרתי את ההתפתחות של זוגות קומפקטיים בסביבות רוויות גז של הדור השני של היווצרות כוכבים ודנתי בערוץ חדשני להיווצרות של גלי-כבידה בסביבות אלו. גלי-כבידה כאלו יווצרו בשלבים המוקדמים של יצירת הצביר. גז יכול גם לסייע ביצירת זוגות. חקרתי את התנאים עבור יצירת זוגות בכמה סביבות גזיות כולל איזורי יצירת כוכבים, דסקות ספיחה -- גרעינים גלקטיים פעילים וכן סביבות גזיות של היווצרות הדור השני של כוכבים בצבירים כדוריים, ותיארתי תהליכים דינמיים אפשריים לאחר הלכידה. יצירת זוגות באופן דינמי דורשת מקור דיסיפציית אנרגיה חיצוני. בסביבות עתירות גז, הגז גורם לדעיכת אנרגיה שעשויה להיות משמעותית ולגרום ליצירת זוגות. במידה והזוגות האלו לא ייהרסו זמן קצר לאחר יצירתם, הם עשויים לתרום באופן משמעותי לאוכלוסיית הזוגות. לאחר מכן, דנתי בתהליך של הגנה על זוגות רכים בסביבות עתירות גז וערוץ חדשני להיווצרות זוגות קשים מזוגות רכים בסביבות אלו. בדרך כלל, זוגות רכים נידונים להריסה כתוצאה מאינטראקציות עם כוכבים אחרים. בסביבות עתירות גז המצב שונה, וכל אינטראקציה של זוגות עם הגז מחלצת אינטראקציה מהזוגות, בין אם רכים או קשים, ובכך מקשה גם זוגות רכים. אם צפיפות הגז גבוהה מספיק, בתום התקופה הגזית, אחוז ניכר מהזוגות הרכים הופכים רכים וניצלים מהריסה. התהליך הזה עשוי להשפיע על המספר הכולל של זוגות קשים ובכך לתרום לתהליכים שונים כגון סופרנובות או אינטראקציות גאותיות. מפני שרוב הזוגות נוצרים רכים בכמה ערוצי יצירה, התקשות דרך גז עשויה להיות ערוץ מרכזי ביצירת

המחקר נעשה בפקולטה לפיסיקה בהנחיית פרופ' חגי ב. פרץ.

מחברת חיבור זה מצהירה כי המחקר, כולל איסוף הנתונים, עיבודם והצגתם, התייחסות והשוואה למחקרים קודמים וכו', נעשה כולו בצורה ישרה, כמצופה ממחקר מדעי המבוצע לפי אמות המידה האתיות של העולם האקדמי. כמו כן, הדיווח על המחקר ותוצאותיו בחיבור זה נעשה בצורה ישרה ומלאה, לפי אותן אמות מידה.

חיבור דוקטורט זה מבוסס על תוצאות שפורסמו בכתבי-עת/התקבלו ויפורסמו בתקווה בעתיד הקרוב. אני הובלתי את המחקר בכל המאמרים המצויינים מטה, פרט לאלו המסומנים בכוכבית, בהם הובלתי במשותף עם הכותב/ת הראשונ/ה. להלן המאמרים שהופיעו בחיבור

Published/accepted

1. **Rozner M.** & Perets H. B., Soft no more: gas shielding protects soft binaries from disruption in gas-rich environments, accepted for publication in APJ

2. **Rozner M**. & Perets H. B. , *Born to be wide: the distribution of wide binaries in the field and soft binaries in clusters*, APJ, 955, 2 (2023)

3. **Rozner, M.**, Generozov A. & Perets H.B., *Binary formation through gas-assisted capture and the implications for stellar, planetary and compact-object evolution*, MNRAS, 521,1 (2023)

4. **Rozner, M.** & Perets H.B., *Binary evolution, gravitational-wave mergers and explosive transients in multiple-populations gas-enriched globular-clusters*, APJ, 931,2

5. **Rozner, M.**, Glanz H. , Perets, H. B. and Grishin, E., *Inflated Eccentric Migration of Evolving Gas-Giants I: Accelerated Formation and Destruction of Hot and Warm Jupiters*, APJ, 931, 1 (2022)

6. *Glanz H., **Rozner, M.** , Perets, H. B. and Grishin, E., *Inflated Eccentric Migration of evolving gas giants II: Numerical methodology and basic concepts*, APJ, 931, 1 (2022)

7. **Rozner, M.**, Veras, D. & Perets, H. B., Rapid destruction of planetary debris around WDs through wind erosion, MNRAS, 502, 4 (2021)

8. **Rozner, M.**, Grishin, E., Perets, H. B., *The Wide-Binary Origin of The Pluto-Charon System*, MNRAS, 497, 4 (2020)

9. *Grishin, E., **Rozner, M.** , Perets, H. B., *Erosion-driven Size Redistribution of Protoplanetary Disk Solids and the Onset of Streaming Instability and Pebble Accretion*, APJL, 898,1 (2020)

10. **Rozner, M.**, Grishin, E., Perets, H. B., *The aeolian-erosion barrier for the growth of metre-size objects in protoplanetary discs*, MNRAS, 496, 4 (2020)

דינמיקה של כוכבי-לכת, כוכבים ועצמים קומפקטיים

חיבור על מחקר

לשם מילוי חלקי של הדרישות לקבלת התואר
דוקטור לפילוסופיה

מור רוזנר

הוגש לסנט הטכניון — מכון טכנולוגי לישראל
ניסן התשפ״ד חיפה מאי 2024

דינמיקה של כוכבי-לכת, כוכבים ועצמים קומפקטיים

מור רוזנר