



Black hole binary systems

From dynamics to accretion

Von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades

Doktor der Naturwissenschaften
— Dr. rer. nat. —

genehmigte Dissertation von

M. Sc. Cristián Danilo Maureira Fredes

geboren am 19 Mai 1988 in San Antonio, Chile

2018

Referent

Prof. Dr. Bernard Schutz
Albert-Einstein-Institut
Leibniz Universität Hannover

Korreferent

Dr. Pau Amaro-Seoane
Albert-Einstein-Institut
Potsdam Universität

Tag der Promotion

27. Juni 2018

Nobody ever figures out what life is all about, and it doesn't matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough.

Richard Phillips Feynman

Abstract

The problem of the evolution of a large number of particles due to gravity is crucial to many astrophysical phenomena. An important problem is the dynamical evolution of a dense stellar system, such as a globular cluster (GC), a galactic nucleus (GN) or nuclear star cluster (NSC). Such loci are the breeding grounds of sources of tidal disruptions and gravitational waves. Right in the middle of these regions a massive black hole (MBH) might be lurking, which makes the problem even more interesting, because such massive objects can form a pair and later a binary, which could be powerful source of gravitational radiation for space-borne observatories.

The detailed tracking of the dynamical evolution of a set of N stars is a complex problem. Since we lack an analytical solution, it needs to be studied by approximations and numerical methods close to what we might expect from Nature. The close interactions between stars define the core mechanism that determines the global evolution of dense stellar systems. These interactions are responsible for defining the timescale in which catastrophic phenomena happen, such as the core collapse of the system; particularly relevant for the formation of a gravitational capture, that eventually will evolve mostly due to the emission of gravitational radiation.

Moreover, depending on the problem we are addressing we might need to add further layers of complexity. For instance, in the case of a GN the presence of gas can play an crucial role, so it needs to be considered, particularly in the massive black hole binary (MBHB) formation process.

The formation of the disc structure around the MBHB is in particular a very important problem which has received very little numerical investigation until the presentation of this work. It is usually assumed that the gas is supplied via the accumulated infall of gaseous clouds on to the binary, and hence this gas is distributed in a disc-like structure around it. Hence, it is relevant to address the formation of binaries taking into account such a gaseous disc around the system in different orbits, and the interaction of the gas with the black holes, not just dynamically, but also via the accretion on to them.

Motivated by the complexity and many open question of these fundamental problems, this thesis is (i) a detailed study of the non-linear dynamics that occur in dense stellar systems with state-of-the-art numerical techniques, (ii) a detailed study of the impact of gas on to the binary, in particular to address the role of circumbinary discs on the evolution of a MBHB, and (iii) how repeated infall events of gaseous clouds distribute and shape around such massive binaries, as well as the impact on the dynamical evolution of the binary itself.

All of these topics are intertwined and I have worked in them in a parallel way during my PhD. The most remarkable findings of my work are that (i) the use of a softening parameter is critical to analyse the long-term evolution of a dense stellar system, with an important impact on the timescale in which crucial events happen, including the formation of binaries, (ii) the way binaries of MBHs accrete gas in counter-rotating circumbinary discs, will determine the evolution of the massive binary, and (iii) the formation of disc-like structures around these binaries in a GN is, to say the least, challenging. Also, episodic circumbinary structures will modify the orbital evolution of MBHBs, altering their associated gravitational merger timescale.

Keywords: Black holes, Accretion, Stellar dynamics, Gravitational Waves.

Zusammenfassung

Die durch Schwerkraft verursachte Entwicklung einer großen Anzahl von Teilchen ist entscheidend für viele astrophysikalische Phänomene. Eine wichtige Fragestellung ist die dynamische Entwicklung eines dichten Sternsystems, wie etwa eines Kugelsternhaufens (KS) oder eines Galaxienkerns (GK). Solche Orte sind die Brutstätten von Quellen, die Gravitationswellen und Gezeiten-Sternzerissereignisse produzieren. Mitten in solchen Systemen könnte ein massives schwarzes Loch lauern, was das Problem noch interessanter macht, da sich solch massive Objekte zu einem Paar und danach zu einem Doppelsternsystem zusammenfinden können. Diese Gruppierungen sind mächtige Quellen von Gravitationsstrahlung für Raumfahrtobservatorien.

Die detaillierte Nachverfolgung der dynamischen Entwicklung von N Sternen ist ein komplexes Problem. Da es keine analytische Lösung gibt, müssen wir mit Näherungen und numerischen Methoden arbeiten, die möglichst gut die Natur abbilden. Die Wechselwirkung zwischen Sternen in nächster Nähe definieren den Kernmechanismus, der die globale Entwicklung des dichten Sternsystems bestimmt. Diese Wechselwirkungen sind für die Zeitskalen "katastrophaler" Ereignisse verantwortlich, wie etwa den Kernkollaps des Systems. Dieser Kollaps ist von essenzieller Bedeutung für die Bildung eines gravitativ bedingten Einfangs, dessen dynamische Entwicklung irgendwann vom Energieverlust durch Gravitationsstrahlung getrieben wird.

Darüber hinaus kann es bei manchen Fragestellungen notwendig sein, weitere Ebenen an Komplexität hinzuzufügen. Zum Beispiel kann bei einem GK das Vorhandensein von Gas eine entscheidende Rolle spielen. Deshalb muss Gas in diesem Fall berücksichtigt werden, insbesondere bei der Formation eines massiven binären schwarzen Loches (MBSL).

Die Bildung einer Scheibenstruktur um ein MBSL herum ist ein sehr wichtiges Problem, das bis zur Präsentation dieser Dissertation wenig Beachtung durch numerische Nachforschungen gefunden hat. Normalerweise nimmt man an, dass das Gas durch den Einfall von Gaswolken auf das Binärsystem ins Innere gelangt und sich deshalb eine scheibenartige Struktur bildet. Von daher ist es wichtig, die Formation von Binärsystemen unter der Anwesenheit einer solchen Gasscheibe auf unterschiedlichen Umlaufbahnen zu untersuchen, sowie die Wechselwirkung der MBSL mit dem Gas. Nicht nur die dynamische Wechselwirkung, sondern auch die Akkretion des Gases auf die MBSL ist von hoher Bedeutung.

Motiviert durch die Komplexität und vielen offenen Fragen dieser fundamentalen Themen ist diese Dissertation (i) eine detaillierte Studie der nicht-linearen Dynamik, die sich in dichten Sternsystemen abspielen durch modernste numerische Methoden, (ii) eine detaillierte Studie des Effekts von Gas auf Binärsysteme, insbesondere um die Rolle von Gasscheiben bei der Entwicklung von MBSL zu adressieren und (iii) wie sich wiederholte Einfälle von Gaswolken verteilen und formieren um solche MBSL, sowie die Wirkung auf die dynamische Entwicklung von Binärsystemen selbst.

Alle diese Themen sind stark verwandt und ich habe während meiner Dissertationsarbeit parallel an ihnen gearbeitet. Die bedeutendsten Erkenntnisse meiner Arbeit sind: (i) Der Einsatz eines "Softening Parameters" ist für die Analyse der langzeitlichen Entwicklung eines dichten Sternsystems entscheidend und hat einen wichtigen Einfluss auf die Zeitskalen, in denen wichtige Prozesse stattfinden, auch die Formation von Doppelsternsystemen (ii) Die Art und Weise, wie massive MBSL Gas akkretieren in entgegen-rotierenden Gasscheiben bestimmt die Entwicklung des massiven Binärsystems. (iii) Die Formation von scheibenartigen Strukturen in solchen Binärsystemen in GK ist, um es vorsichtig auszudrücken, sehr schwierig. Außerdem verändern sporadisch auftretende Strukturen um das Binärsystem herum die Entwicklung der Umlaufbahn von MBSL, was deren Zeitskala für eine durch Gravitation bedingte Verschmelzung beeinflusst.

Schlagworte: Schwarzes Loch, Akkretion, Stelardynamik, Gravitationswelle.

CONTENTS

I INTRODUCTION

1	GRAVITATIONAL WAVE ASTRONOMY	12
2	STELLAR DYNAMICS	13
2.1	Not that easy, actually	13
2.2	Getting into some more details	14
2.3	Different possible schemes	15
2.4	Modern hardware: GPUs	15
3	MASSIVE BLACK HOLES: FIRST PICTURE	18
3.1	Preamble	18
3.2	Massive black holes: from quasars to our Galactic Centre	18
3.3	Our Own Galaxy	19
4	BINARIES OF MASSIVE BLACK HOLES	21
4.1	Preamble	21
4.2	Formation and evolution: different phases	21
5	THE ORIGIN OF THE GAS: INTERACTING GALAXIES	24
5.1	Preamble	24
5.2	Major mergers	24
5.3	Unequal mergers	25
5.4	Minor mergers	26
6	GETTING THE GAS TO THE MASSIVE BLACK HOLES	27
6.1	Preamble	27
6.2	Nuclear-disc-driven	27
6.3	Circumbinary-disc-driven	28
7	GOALS OF THIS THESIS	29
7.1	From dynamics	29
7.2	To accretion	30
7.3	To study black hole binary systems	31

II PROJECTS

8	DYNAMICAL EVOLUTION OF DENSE STELLAR SYSTEMS	35
8.1	The current algorithm	35
8.2	The three flavours of GRAVIDY: tests	48
8.3	The role of softening on dynamics	55
8.4	Relativistic corrections	63
8.5	Conclusions	64
8.6	Appendix A: about the code	65
8.7	N -body visualisation tool	65
9	RETROGRADE MBHB EVOLUTION ON CIRCUMBINARY ACCRETION DISCS	67
9.1	Introduction and motivation	67
9.2	Numerical tool and description of the initial models	69
9.3	Accretion prescriptions	70
9.4	Results	72
9.5	A semi-analytical model for the evolution of a binary in an unperturbed retrograde disc	80
9.6	Discussion	85
9.7	Appendix	86

10 CIRCUMBINARY STRUCTURE FORMATION FROM MULTIPLE IN-	
FALLING CLOUDS	88
10.1 Introduction and motivation	88
10.2 Methods	90
10.3 Initial conditions and run description	96
10.4 Resolution and convergence tests	100
10.5 Results	104
10.6 Discussion and future work	117
10.7 Complementary material	119
III CONCLUSIONS	
11 MAIN CONCLUSIONS	125
11.1 From Stellar Dynamics...	125
11.2 ... To Retrograde Accretion Discs...	127
11.3 ... And The Formation Of Gaseous Structures...	128
11.4 ... To study black hole binary systems	129
Bibliography	148

LIST OF FIGURES

Figure 1	Messier 53	14
Figure 2	Nvidia GTX1070	16
Figure 3	Artist's impression of Hyperluminous quasar	19
Figure 4	Artist's impression of black holes	22
Figure 5	The Antennae galaxies	25
Figure 6	Artist's impression of infalling gas onto SMBHs	28
Figure 7	Accretion disc around BH gargantua	30
Figure 8	Class diagram of the code	40
Figure 9	Block time steps illustration	42
Figure 10	Hermite integration scheme illustration	45
Figure 11	Relation between updated particles	45
Figure 12	GPU Grid configuration	47
Figure 13	Parallelisation scheme	48
Figure 14	Cumulative energy error vs wall clock time	51
Figure 15	Wall clock time	52
Figure 16	Acceleration factor	54
Figure 17	Performance of the implementations	55
Figure 18	Cumulative energy error vs wall clock time	57
Figure 19	Comparison of the Lagrange radii	59
Figure 20	Energy conservation in a long time integration	60
Figure 21	Plummer sphere using Kroupa IMF	61
Figure 22	Plummer sphere using Kroupa IMF: core collapse	61
Figure 23	Time step distribution	62
Figure 24	SMBHB test using PN terms ($q=1$ and $q=2$)	63
Figure 25	SMBHB test using PN terms ($q=5$ and $q=100$)	64
Figure 26	GraviDyView snapshot	66
Figure 27	Surface density of the initial conditions	70
Figure 28	Relative MBH orbital separation ($e=0.0$)	72
Figure 29	Evolution of the secondary MBH ($e=0.0$)	73
Figure 30	Relative MBH orbital separation and eccentricity evolution ($e=0.0$)	74
Figure 31	Relative MBH orbital separation versus secondary mass for the $e=0.0$ case	75
Figure 32	Relative MBH orbital separation and eccentricity evolution ($e=0.6$)	76
Figure 33	Evolution of the secondary MBH ($e=0.6$)	77
Figure 34	Relative MBH orbital separation ($e=0.0$), including 3D	78
Figure 35	Relative MBH orbital separation ($e=0.6$), including 3D	79
Figure 36	Eccentricity evolution ($e=0.0$), including 3D	80
Figure 37	Eccentricity evolution ($e=0.6$), including 3D	81
Figure 38	Eccentricity vs secondary mass, semi-analytical model ($e=0.6$)	83
Figure 39	Eccentricity and semi major axis vs secondary mass, semi-analytical model ($e=0.01$)	84
Figure 40	Schematic representation of selected configurations of the simulations	92
Figure 41	Pericentre distance and injection of time of the clouds	97
Figure 42	Orientation of the angular momentum of each cloud	98

Figure 43	Time evolution of the relevant MBHB parameters for different values of r_{sink}	101
Figure 44	Accretion rate for all the resolutions	102
Figure 45	Angle-averaged density profile	102
Figure 46	Column density rendering	103
Figure 47	Angle-averaged mass density profile for all runs	104
Figure 48	Mass evolution of the circumbinary structures	105
Figure 49	Time evolution for the circumbinary structure	107
Figure 50	Snapshots of the simulations	108
Figure 51	Same as Fig. 49 but for the mini discs	111
Figure 52	RunA $F = 0.0$ last snapshot	112
Figure 53	Same as Fig. 49 but for the set of ‘forked’ simulations	115
Figure 54	Same as Fig. 51 but for the “forked” simulations.	115
Figure 55	Last snapshot visualisation of the “forked” simulations	116
Figure 56	Same as Fig. 49 for RunC.	120
Figure 57	Same as Fig. 51 for RunC.	122
Figure 58	Two SMBHs on their way to coalescence	126
Figure 59	Counter-rotating system snapshots from chapter 9	127
Figure 60	RunA $F = 0.0$ moment snapshot from chapter 9	129

LIST OF TABLES

Table 1	Specification of the different systems of the Albert Einstein Institute used for the tests.	50
Table 2	Cloud injection times	98
Table 3	Average mass of the mini-discs surrounding the M_{BHs} in all our runs.	113
Table 4	Initial 3-D position and velocity (x, y, z components) of the centre of mass of each cloud for RunA and RunB.	121

ACKNOWLEDGEMENTS

I thank the support from the Transregio 7 “Gravitational Wave Astronomy” and “Supermassive black holes, accretion discs, stellar dynamics and tidal disruptions” projects financed by the Deutsche Forschungsgemeinschaft (DFG) awarded to my adviser Pau Amaro-Seoane, which funded my doctorate.

I am honoured for being part of The International Max Planck Research Schools (IMPRS) on Gravitational Wave Astronomy, their people, training and education were crucial for my doctoral studies development.

To the Max-Planck-Institute für Gravitationsphysik (Albert-Einstein-Institut) for providing an environment for first-class scientist, to lead the research field of gravitational physics and gravitational astronomy.

To my adviser Pau Amaro-Seoane, for proving me the necessary “spiritual guidance” through my doctorate, opening multiple doors and opportunities in all these years, and guiding me not only on my research, but on my life in general, helping me to overcome every new challenge.

To all the scientist that accompanied and helped me through my doctorate, particularly to my closest collaborators and friends Alberto Sesana and Massimo Dotti, for their endless support; Jorge Cuadra for triggering this whole new chapter in my life; Xian Chen and Stanislav Babak for their scientific and personal support. To all the people that took the time of reading pieces of the thesis and gave me precious feedback that helped me to improve the document, specially to Pau, Alberto, Felipe, Massimo and Steve.

To Patrick Brem, for his friendship and crucial help at the beginning of my PhD, but most importantly for encouraging me to drink Club Mate and even more Beer. To Felipe G. Goicovic, for his patience and guidance since we started working together, and for his personal and professional support during all these time. To Luciano del Valle, for his amazing disposal on helping me through our projects and pen-and-paper skills.

To María José, you have been my constant support during the most difficult part of my PhD. You help me every day to be a better person, professionally and emotionally. Thanks for your kindness and for making me believe in myself again (And everything while you were finishing your PhD too!).

To my sister, for being there whenever I have needed her, and for her endless patience and guidance. I will always admire you, and encourage you to achieve all what you deserve. To my mother, for your infinite love and concern even when we are 12,000km apart. Thanks for being there, always, without asking for anything in return. To my father, for always encourage me to achieve more, and be a better person, but especially for that afternoon you helped me to prepare my Solar System exposition while sitting on the floor, drawing planets with plates and pot lids.

GLOSSARY

Acronym	Meaning
$1 M_{\odot}$	1 Solar Mass = 1.99×10^{30} kg
1 pc	1 parsec = 3.09×10^{16} m
1 Myr/Gyr	One million/billion years
AGN	Active Galactic Nucleus
BH	Black Hole
CDM	Cold Dark Matter
EMRI	Extreme Mass Ratio Inspiral
GC	Globular Cluster
GW	Gravitational Wave
MBH	Massive Black Hole ($M \approx 10^6 M_{\odot}$)
MBHB	Massive Black Hole Binary
NS	Neutron Star
PN	Post-Newtonian
SMBH	Super Massive Black Hole ($M > 10^6 M_{\odot}$)
SMBBH	Super Massive Binary Black Hole
SPH	Smoothed Particle Hydrodynamics

Part I

INTRODUCTION

GRAVITATIONAL WAVE ASTRONOMY

The final writing up and polishing of this thesis sadly coincides with the death of Donald Lynden-Bell last 5 February 2018. I envisage him as one of the founders of many different fields in modern astrophysics, in particular of the concept that supermassive black holes (SMBHs) are lurking at the core of active galaxies as a way to explain quasars. He also contributed to other fields, such as stellar and galactic dynamics, astrophysical jets and general relativity. My thesis addresses a few problems that arose within the new fields that he and others conceived.

When I enrolled for a thesis at the Max-Planck Institut for Gravitational Physics (Albert Einstein Institute), the existence of black holes (BHs) and gravitational waves (GWs) were questioned.

In these few years, the situation has drastically changed. In particular, the direct detection of gravitational waves by LIGO and Virgo has proved the existence of GWs (see the detection papers for the detections, Abbott et al. 2016a,b, 2017a,b,c,d). It must be noted that these detections are not “just” from BHs. The LIGO/Virgo team has also observed two inspiralling neutron stars, and the optical telescopes have impressively done what was considered by many unlikely: the joint detection of the system in the electromagnetic domain. The first event unleashed a flood of about 70 papers which I will not include here for legibility reasons.

Hence, in a matter of about two years, we have observed how Gravitational Wave Astronomy has followed a similar path to what Cosmology did in the 60’s, when it went from being a data-starved science (as many referred to it) to one of the most important fields of research in natural sciences. Since the announcement by LIGO about the first direct detection of GWs back in February 11th 2016, Gravitational Wave Astronomy has grown from being a purely theoretical discipline in astrophysics to one of its most vibrant fields, thanks to the ground-based GW detectors LIGO and Virgo providing us with (i) the direct detection of GWs, (ii) a very strong indication that objects consistent with general relativity BHs exist, (iii) the observation in GWs of the inspiral of neutron stars, (iv) the opening of multi-messenger astronomy, thanks to the joint detection of the electromagnetic radiation associated to that inspiral, and (v) a Nobel Prize.

The problem of formation and evolution of GW sources requires bringing together insights and techniques from a variety of field, which is what makes Gravitational Wave Astronomy such an interesting area. In particular, since plausibly the most straightforward way of making a changing quadrupole moment, i.e. a source of GWs, is a binary, my thesis deals with the formation of these in astrophysics.

STELLAR DYNAMICS

If allegedly the best understood way of creating a source of GWs is a binary, the easiest possible way of doing it is via stellar dynamics. Here we will consider dense stellar environments as the formation grounds for compact binaries. In these environments, compact binaries form through pure gravitational encounters between particles. To high approximation, we can model these particles as points, and ignore effects other than gravity in their evolution.

With this I mean the exchange of energy and angular momentum in a dense stellar environment such as a globular cluster (GCs) or a galactic nucleus (GN) until two stars are bound with each other. These astrophysical objects, GCs and GNs (or nuclear star clusters, NSC, which can be described as “observed GNs”) have densities that can reach some 10^6 or even 10^7 stars per cubic parsec. In such extreme environments, the exchange of angular momentum and energy lead the evolution of the whole system. Stellar binaries are, in particular, one of the main driving forces in this global evolution, and dictate whether the entire sample of millions of stars will collapse or not, to just mention one example. These binaries, if formed out of two compact objects, will become prominent sources of GWs.

This is the connection between stellar dynamics and sources of GWs.

2.1 NOT THAT EASY, ACTUALLY

The dynamical formation of compact object binaries and the emission of GWs is a captivating and very timely problem, but stellar dynamics per se represents also a fascinating and complex problem, in spite of the apparent simplicity of “just” solving this equation:

$$\ddot{\mathbf{r}}_i = -G \sum_{j=i, j \neq i}^{j=N} m_j \frac{(\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3}.$$

This simple equation has allowed us to study the solar system for the past 300 years, but also star clusters (i.e. GCs, GNs and NSCs), see Fig. 1 for an example of these accumulations of stars. We can also study whole galaxies, systems of about 10^{11} stars, and even clusters of galaxies. However, the kind of dynamics I have dealt with during the work of my thesis is collisional dynamics, meaning that we care about the long-term effects of close (as well as not-so-close) stellar encounters. As heat conduction in the air in a room, the evolution of these dense stellar systems is governed by the slow diffusion of “heat” through the system from the inside towards the edge.

In Newtonian gravity, we can only solve analytically the two-body problem (and sometimes, with some particular exceptions, the restricted three-body problem). If we take into account relativistic effects, which are imperative in the addressing of the formation and evolution of GW sources, we must rely on approximations provided that the two stars are far from each other, and on computer simulations when they get closer and closer.



Figure 1: Messier 53 (also known as M53 or NGC 5024) is a globular cluster in the Coma Berenices constellation. It has a core radius of about 2pc and a half-light radius of 6pc, and a total mass of about $10^6 M_{\odot}$.

But we are not just interested in the relativistic two-body problem. Since the binaries that will eventually turn into sources of GWs are not perfectly isolated in these systems, but surrounded by up to ten million of other stars, we must resort to numerical simulations to investigate how dense stellar systems evolve and how binaries of compact objects form.

This “clean”, dynamics-driven formation of binaries of stars is what motivated the first phase of my doctorate. I started developing from scratch a new direct-summation N -body code in spite of the fact that there are a few of them publicly available. The reason for that is twofold. From the one handside, I found difficult to modify existing codes, some of which have been under development for over 50 years by many different people with coding styles, and last, my background in computer science.

2.2 GETTING INTO SOME MORE DETAILS

The dynamical evolution of a dense stellar system such as a globular cluster or a galactic nucleus, has been addressed extensively by a number of authors. For Newtonian systems consisting of more than two stars we must rely on numerical approaches which provide us with solutions that are more or less accurate. In this sense, one could make the following coarse categorisation of integration schemes for pure stellar dynamics: those which are particle-based and those which are not. In the latter, the system is treated as a continuum, so that while we know the general properties of the stellar system such as the mean stellar density, of the average velocity dispersion, we do not have information about specific orbits of stars. To this group, belongs direct integration of the Fokker-Planck equation [Inagaki and Wiyanto, 1984, Kim et al., 1998] or moments of it [Amaro-Seoane et al., 2004, Schneider et al., 2011], including Monte Carlo approaches to the numerical integration of this equation [Spitzer and Hart, 1971]. A particle-based algorithm, however, assumes that a particle is tracing a star, or a group of them. In this group, the techniques go back to the early 40’s and involved light bulbs [Holmberg, 1941]. The first computer simulations were performed at the Astronomisches Rechen

Institut, in Heidelberg, Germany, by [von Hoerner, 1960, 1963], using 16 and 25 particles. These first steps led to the modern N -body algorithms.

2.3 DIFFERENT POSSIBLE SCHEMES

We can distinguish two types of N -body algorithms: the so-called collisionless, where a star just sees the background potential of the rest of the stellar system [e.g. the Barnes-Hut treecode or the fast multipole method Barnes and Hut, 1986, Greendard, 1987, which scale as $O(N \log N)$ and $O(N)$, with N the particle number, respectively], and the more expensive collisional one, or “direct-summation”, in which one integrates all gravitational forces for all stars to take into account the graininess of the potential and individual time steps, to avoid large numerical errors. This is important in situations in which close encounters between stars play a crucial role, such as in galactic nuclei and globular clusters, because of the exchange of energy and angular momentum. The price to pay however is that they typically scale as $O(N^2)$ or even $O(N^3)$.

A very well known example is the family of direct-summation NBODY integrators of Aarseth [see e.g. Aarseth, 1999, Spurzem, 1999, Aarseth, 2003]¹ or also KIRA [see Portegies Zwart et al., 2001]². The progress in both software and hardware has reached a position in which we start to get closer and closer to simulate realistic systems.

However, the scaling $O(N^2)$ requires supercomputers, such as traditional Beowulf clusters, which requires a parallelisation of the code, such as the version of NBODY6 developed by Spurzem and collaborators, NBODY6++³ [Spurzem, 1999], or special-purpose hardware, like the GRAPE (short for GRAVity PipE⁴) system. The principle behind GRAPE systems is to run on a special-purpose chip the most time consuming part of an N -body simulation: the calculation of the accelerations between the particles. The remainder is calculated on a normal computer which serves as host to the accelerator board(s) containing the special purpose chips. Such a system achieves similar or even higher speeds than implementations of the N -body problem on supercomputers [see e.g. Taiji et al., 1996, Makino and Taiji, 1998, Makino, 1998, Fukushige et al., 2005].

2.4 MODERN HARDWARE: GPUS

On the other hand, modern graphics processing units (GPUs) offer a very interesting alternative. They have been mostly used in gaming devices, embedded systems and mobile phones. In Fig. 2 we can see an example of a modern GPU. They were originally used to perform calculations related to 3D computer graphics. Nevertheless, due to their highly parallel structure and computational speed, they can very efficiently be used for complex algorithms. This involves dealing with the parallel computing architecture developed by NVIDIA⁵, the Compute Unified Device Architecture (CUDA). This is the main engine in NVIDIA GPUs, and it has been made accessible to developers via standard programming languages, such as C with NVIDIA extensions compiled thanks to a PathScale Open64 C compiler. This is what allows us to cre-

1 All versions of the code are publicly available at the URL <http://www.ast.cam.ac.uk/~sverre/web/pages/nbody.htm>
 2 <http://www.sns.ias.edu/~starlab/>
 3 Available at this URL <http://silkroad.bao.ac.cn/nb6mpi>
 4 <http://grape.c.u-tokyo.ac.jp/grape>
 5 <http://www.nvidia.com>



Figure 2: The Nvidia GTX 1070 GPU (Pascal chip), with 1920 CUDA cores, reaching a theoretical peak of 6.5 TFLOPs (Single precision).

ate binary modules to be run on the GPUs. Another option is Open Computing Language (OpenCL)⁶, which offers a framework to write parallel programmes for heterogeneous systems, including also computational nodes with field-programmable gate arrays (FPGAs), digital signal processors (DSPs), among others. CUDA, and also OpenCL are “the doors” to the native instruction set and memory of the parallel elements in the GPUs. This means that these can be handled as open architectures like GPUs with the enormous advantage of having a parallel-cores configuration. More remarkably, each core can run *thousands* of processes at the same time. We selected CUDA over OpenCL, because our systems are equipped with NVIDIA GPUs, even though we note that OpenCL has shown similar performance to CUDA in N -body simulations [Capuzzo-Dolcetta and Spera, 2013].

There has been recently an effort at porting existing codes to this architecture, like e.g. the work of Portegies Zwart et al. [2007], Hamada and Iitaka [2007], Belleman et al. [2008] on single nodes or using large GPU clusters [Berczik et al., 2011, Nitadori and Aarseth, 2012, Capuzzo-Dolcetta et al., 2013] and recently, the work by [Berczik et al., 2013] using up to 700 thousand GPU cores for a few million bodies simulation with the ϕ -GPU⁷ code, which reached in their work about the half of the peak of the new Nvidia Kepler K20 cards.

Large-scale (meaning number of particles) simulations have recently seen an important improvement with the work of Wang et al. [2015, 2016]. In his more recent work of 2016, Wang and collaborators integrated systems of one million bodies in a globular cluster simulation, using from 2,000 to 8,600 hours of computing time.⁸

In chapter 8 I present the development from scratch of a direct-summation N -body code with a relativistic treatment of compact binaries, aimed to GPUs, but also for traditional CPU clusters. I present a suit of different well-known tests in

⁶ <https://www.khronos.org/opencl/>

⁷ <ftp://ftp.mao.kiev.ua/pub/berczik/phi-GPU/>

⁸ This impressive achievement was rewarded with a bottle of Scotch whisky (not whiskey), kindly and generously offered to him by Douglas Heggie during the excellent MODEST 15-S in Kobe.

stellar dynamics that the code successfully passed, as well as a study of the impact of adopting a softening in the integration of a dense stellar system on its global evolution.

MASSIVE BLACK HOLES: FIRST PICTURE

3.1 PREAMBLE

To follow the line of thought of Donald Lynden-Bell, in section 4.5 of Lynden-Bell and Wood [1968], the authors state that

“Only a fool tries the harder problem when he does not understand the simplest special case”.

Whilst the dynamical formation of binaries in a dense stellar system is a well-defined and clean problem, future space-borne GW observatories will be aiming at much more massive binaries. In particular, the Laser Interferometer Space Antenna (LISA) mission [see e.g. Amaro-Seoane et al., 2017, 2013a, 2012a,b] will observe binaries of massive black holes (MBHs), ranging between 10^4 and $10^7 M_{\odot}$. These binaries form and spend the last phase of their existence in galactic nuclei, where gas piles up and distributes around the MBHs.

Before we get into details, and in particular how you form a binary of them, it is necessary to understand the motivation for the existence of MBHs.

3.2 MASSIVE BLACK HOLES: FROM QUASARS TO OUR GALACTIC CENTRE

Schmidt [1963] identified a star-like object with the radio source 3C 273. This source is very powerful, and has the highest redshift for an object of those characteristics, of $z = 0.158$. This object turned out to be the second most distant known object to astronomers at that time, with an associated luminosity of ten times that of a galaxy. This object, which was a quasar, a “quasi-stellar radio source” (QSO), was the first discovery, but many quickly followed. QSOs turned out to be the most luminous, powerful, and energetic objects known in the whole universe. They were found in the centres of active galaxies, with an energy output up to thousand of times the entire luminosity of our Galaxy, the Milky Way (to distinguish between our and others galaxies, I will use a capital “G” to refer to the Milky Way galaxy).

The key question here is how to produce a luminosity of some 10^{40} W. The only plausible way for physicists was accretion of matter on to an object [Hoyle and Fowler, 1963, Salpeter, 1964, Zel’dovich, 1964, Lynden-Bell, 1969, 1978, Lynden-Bell and Rees, 1971], more specifically on to MBHs. Observers quested from that moment for MBHs, and the first solid observations started to appear by the 80’s and 90’s [e.g. Tonry, 1984, 1987], which was an accomplishment considering the technology: optical instruments, using absorption-line spectroscopy techniques, see in particular M31 [Dressler and Richstone, 1988, Kormendy, 1988], NGC3115 [Kormendy and Richstone, 1992], and NGC3377 [Kormendy and Richstone, 1995, Kormendy et al., 1998].

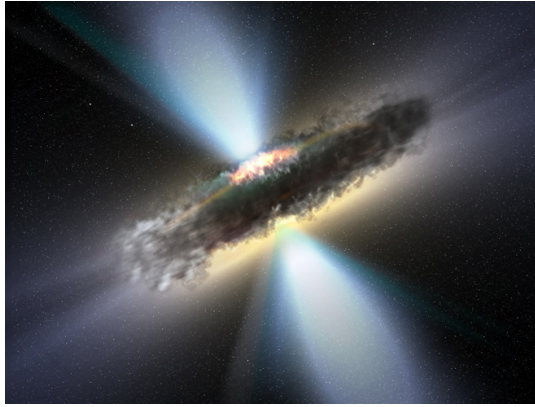


Figure 3: Artist’s impression of a hyperluminous quasar similar to those harbouring supermassive black holes such as e.g. TON 618 or S5 0014+81 (with masses of 66 billion and 40 billion solar masses, respectively). The gas is distributed in this illustration as an accretion disc.

Later things changed. We came to the idea of putting a whole telescope out of the problems of Earth’s atmosphere, which provided us with a spatial distribution 5 times larger than anything before. With the advent of the Hubble Space Telescope in 1990, we captured extremely high resolution images with a lower background light compared to ground-based telescopes. We could “view deeply into space and time”.

3.3 OUR OWN GALAXY

An obvious place to look for MBHs, if they indeed lurk in galactic nuclei, is our own Galactic Centre (GC, and I follow the same convention I mentioned previously of using capital letters when talking about the Milky Way). This is because it is the galactic nucleus which is closest to us, with an estimated distance of about $D = 8.28 \pm 0.33 \text{ kpc}$ (see the excellent review of Genzel et al. 2010).

The GC is however a “messy” place. The stellar density, which is a strong source of light, has an increasing density towards the centre [Gallego-Cano et al., 2018, Schödel et al., 2018] which very closely follows the theoretical prescriptions of cuspy stellar distributions around MBHs [Peebles, 1972, Bahcall and Wolf, 1976, Preto and Amaro-Seoane, 2010, Baumgardt et al., 2017]. This strong source of light impede observations. Moreover, the distribution of dust and gas contents makes it also difficult to obtain clear observations [Becklin and Neugebauer, 1968, Rieke and Rieke, 1988, Genzel and Townes, 1987, Genzel et al., 1994, Mezger et al., 1996, Baganoff et al., 2001, 2003, Muno et al., 2004]. In spite of all of these challenges, observers have discovered one of the most important results in Astronomy in the last decades: by following the motion of a cluster of a few hundreds of young stars, the “source” stars (S-stars), we have arrived to the conclusion that a very massive, dark object resides at the innermost centre of the GC. A mass of about $4 \times 10^6 M_{\odot}$ enclosed within a volume of $\sim 1/3$ the distance between the Earth and the Sun [see e.g. Genzel et al., 2010, and references therein] seems to suggest for a SMBH and is indeed our very best indication for their existence. Most of the alternatives, more exotic models like agglomerations of compact objects [Maoz, 1995, 1998, Genzel et al., 1997, 2000, Ghez et al., 1998, 2005], Fermion

balls [Ghez et al., 2005, Genzel et al., 2010] have been ruled out with maybe the exception of boson balls or stars [Torres et al., 2000, Schunck and Mielke, 2003, Amaro-Seoane et al., 2010, Liebling and Palenzuela, 2012]. Our MBH has a mass which actually falls in the low-end of the distribution. We know that these objects can have masses of up to a few $10^{10} M_{\odot}$, as illustrated in Fig. 3.

Massive black holes seem to inhabit the centres of most massive galaxies. Our own Galactic Centre harbours the best evidence for one of them. By following the orbits of young stars around a radio source, observers have derived with a very small error the mass and volume of the dark massive object that are revolving around. Although we do not yet have proof that this object is a general relativity black hole, the evidence is very compelling. We will assume for the rest of this thesis that such dark, massive objects are super/massive black holes.

4

BINARIES OF MASSIVE BLACK HOLES

4.1 PREAMBLE

If MBHs exist, it stands to reason, as in the case of stellar-mass black holes, that two of them could form a binary due to galaxy mergers. Such a binary would be a very powerful source of GWs but would emit in a different frequency window for ground-based detectors. As explained before, in order to observe these systems we would need to go to space so as to investigate the milli-Hz window of Gravitational Wave Astronomy.

These sources would be much longer-lived than LIGO/Virgo sources. Instead of speaking of timescales well below one minute (for stellar-mass black holes, although neutron star binaries will be approximately limited to that amount of time), with detectors such as LISA we could accumulate months, if not years of data (depending on different parameters). The kind of science that one can do in that situation is very appealing and is described in, e.g. Amaro-Seoane et al. [2017, 2013a, 2012a,b].

4.2 FORMATION AND EVOLUTION: DIFFERENT PHASES

Physicists realised about the potential of these sources a long time ago, and Begelman et al. [1980] presented the seminal paper about the formation of binaries of SMBHs. A long list of other authors have devoted years of their research to further look into this problem.

Although there are many open questions, there is an emerging consensus regarding the formation and evolution of these binaries, as well as concerning the different phases through which the binary has to go before it successfully merges within a Hubble time. These phases are the following:

- I. *Pairing*. If modern Cosmology is right, galaxies interact many times with other galaxies, and they eventually merge with each other. After the merger, and with the proviso that they both harbour a MBH in their centres, dynamical friction in the stellar bulge is the main mechanism for bringing the two MBHs together to successfully form a Keplerian binary. Their progressive sinking towards the centre [Chandrasekhar, 1943, Begelman et al., 1980, Colpi et al., 2009, Yu, 2002], is due to the dynamical friction on each BH, both because of the gas and the stars. Regarding the stars, and assuming an isothermal sphere formed by N stars, the density profile can be approximately described by $\rho_* = \sigma_*^2 / (2\pi Gr^2)$. One can derive that the timescale associated to the dynamical friction is:

$$\tau_{df} = 2 \times 10^8 \ln^{-1} N \left(\frac{10^6 M_\odot}{M_\bullet} \right) \left(\frac{r}{100 \text{pc}} \right)^2 \left(\frac{\sigma_*}{100 \text{kms}^{-1}} \right) \text{yr}. \quad (1)$$

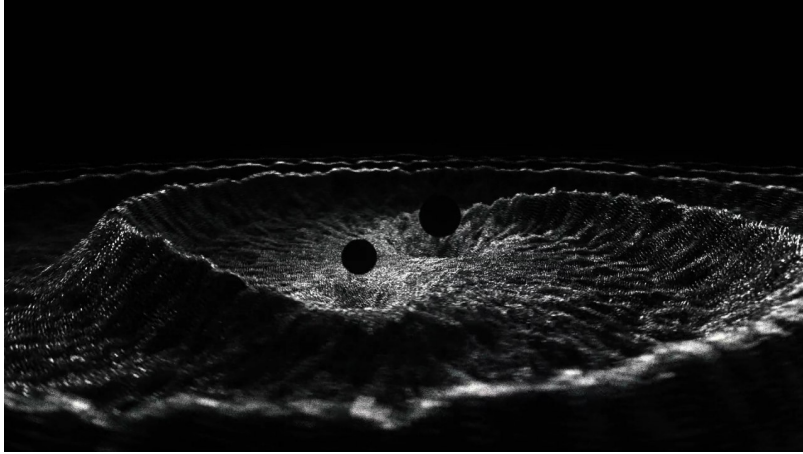


Figure 4: Artist’s impression of two black holes about to merge with each other. The masses are different, as one can see from the imprint in space-time. Credit: Marshmallow Laser Feast (visual), and Samaya Nissanke (science advisor). See video at <https://vimeo.com/254467760> (music: Arthur Jeffes).

We have introduced the black hole mass M_{\bullet} , the velocity dispersion σ_{\star} , and the separation r . It is important to note that there is no significant amplification of the eccentricity during this process [Colpi et al., 1999].

- II. *Hardening*. The two MBHs form now a binary, they are gravitationally bound. From distances of about some tens of parsecs (for MBHs in the range of masses that LISA can detect), dynamical friction is not efficient anymore. The binary will now shrink due to a different physical process. The *individual* dynamical interaction with stars originating from the surrounding stellar system helps the semi-major axis of the binary to further shrink. Three-body interactions with stars happen within a timescale given by

$$\tau_{\text{hard}} \sim \frac{\sigma_{\star}}{\pi G \rho_{\star} a} \sim 70 \left(\frac{\sigma_{\star}}{100 \text{ km s}^{-1}} \right) \left(\frac{10^4 M_{\odot} \text{ pc}^{-3}}{\rho_{\star}} \right) \left(\frac{10^{-3} \text{ pc}}{a} \right) \text{ Myr}. \quad (2)$$

Where we have introduced the semimajor axis a . Contrary to τ_{df} , τ_{hard} increases when the semimajor axis decreases. There is however a problem because for spherical nuclei, and for BH masses in the LISA range, the binary needs to interact (actually eject) the equivalent of $\sim 2 - 3 \times (M_{\bullet,1} + M_{\bullet,2})$ stars, while the reservoir of stars is, being generous, of about one if not two orders of magnitude less than that amount. The two BHs are hence doomed to dance around each other for eternity, and they will never enter the GW-dominated regime. Since this happens at distances of about one parsec, it has been coined as “the last parsec problem” [Milosavljević and Merritt, 2001, Yu, 2002, Merritt and Milosavljević, 2005]. However, it seems that this is an artefact of the oversimplification of current numerical models, in particular probably the sphericity [see, e.g. Preto et al., 2011, Khan et al., 2011] although there has been a lot of work in this direction in the last years.

- III. *GW inspiral*. If the binary manages to shrink its semimajor axis and surmount the last parsec problem, the main driving mechanism in the evolu-

tion is the loss of energy due to the emission of GWs. Assuming Keplerian orbits and neglecting the effect of periaapsis shift, one can obtain an analytical expression for the timescale within which the binary will merge [Peters, 1964]:

$$\tau_{\text{GW}} \sim 5.4 \times 10^8 f(e)^{-1} \frac{(1+q)^2}{q} \frac{a^4}{M_{\bullet,*}^3} \left(\frac{1}{10^{-3} \text{pc}} \right)^4 \left(\frac{10^6 M_{\odot}}{M_{\bullet,*}} \right)^3, \quad (3)$$

with $M_{\bullet,*} = M_{\bullet,1} + M_{\bullet,2}$, and $f(e) = (1 + (73/24)e^2 + (37/96)e^4)(1 - e^2)^{-7/2}$.

- IV. *Gravitational recoiling.* If the two BHs have a spin different of zero and they do not have exactly the same mass, general relativity predicts that the centre of mass of the resulting merger product will receive a recoiling velocity in a preferred direction. This is due to the fact that the emission of GWs in that situation is not isotropic. This prediction has been corroborated thanks to numerical relativity [Campanelli et al., 2006, Baker et al., 2006, González et al., 2007]. Depending of the orientation and magnitude of the spins and orbital angular momentum, this speed can achieve values of a few thousand km/s and hence leave the galactic bulge and the entire galaxy [Gualandris and Merritt, 2008, Merritt et al., 2009, Devecchi et al., 2009].

Binaries of massive black holes form and evolve in the context of pure stellar dynamics and perfect sphericity of the host stellar system. Now that we have relatively well understood this approximate first narrative, we must proceed and add the next most important layer of complexity: gas.

THE ORIGIN OF THE GAS: INTERACTING GALAXIES

5.1 PREAMBLE

According to modern Cosmology, structures in the Universe form from smaller to bigger. This means that baryonic matter, typically proto galaxies in the young Universe, accompanied by their dark matter haloes, merge many times over the Hubble time with each other. Hence, these proto galaxies can be envisaged as building blocks which lead to the formation of galaxies in the local Universe. This is important and relevant for my study, because this has a crucial implication: since galaxies merge with each other, the gas contents they harbour will not follow the usual Keplerian orbit around the centre. Gaseous clouds will interact among each other. Some gas contents might be set on such an orbit that it might arrive to the centre of the galaxy, where the two SMBHs are sinking towards the new centre of the two merged galaxies. Gas will distribute around the binary, maybe in the form of a disc, as it has been put forward in the related literature many times. This has been the working assumption for many theoretical articles written in this field. Motivated by this, and the evolution of a binary of two SMBHs, I decided to address this problem in two different ways, as I will explain later. But, first, as more detailed introduction to this subject, we will succinctly review the most relevant ideas.

In the jargon of this field, depending on the mass ratio of the galaxies interacting, we distinguish between “major” or “minor” mergers (i.e. with similar or very different masses, respectively, see e.g. Colpi 2014).

5.2 MAJOR MERGERS

Major mergers have been addressed in a number of articles, of which probably the one I would highlight is Mayer et al. [2007]. In this article, Mayer and collaborators simulated a gas-rich galaxy merger using a Smooth Particle Hydrodynamics (SPH) code with a relatively large number of particles for each galaxy. This line of work encouraged others to investigate the fate of the gas and, more interestingly for the goals I want to achieve in this PhD, the impact on the evolution of the SMBH binary. The work of Mayer et al. [2007], Colpi et al. [2009], Colpi and Dotti [2011], Chapon et al. [2013], Mayer [2013] addressed the first shrinkage of the SMBH pair, from separations $\gtrsim 10\text{kpc}$ down to the formation of the (Keplerian) binary, which typically happens at a distance of about $\lesssim 10\text{pc}$. Other authors analysed the hardening phase of the binary with dedicated simulations [Escala et al., 2005, Dotti et al., 2006, 2007, 2009, Fiacconi et al., 2013]. The situation is somewhat more complicated for disc galaxy mergers, because the multi-component nature of the physics requires the inclusion of both the dark matter and stellar dynamics (although in a collision-less way). Moreover, it is crucial to include the evolution of the



Figure 5: The Antennae galaxies are a good example of interacting galaxies. These two galaxies, which are located about 19 Mpc from Earth, are shown in this composite image using results from the Chandra X-ray Observatory (in blue), from the Hubble Space Telescope (in gold and brown), and from the Spitzer Space Telescope (in red).

gas as well, which makes the whole simulation setup not only very expensive, but also very challenging.

Van Wassenhove et al. [2012, 2014] later improved the picture by looking into the further $\lesssim 10\text{pc}$ dynamics of the SMBH binary with N-body/SPH simulations, starting from initial separations of 100kpc. Mayer et al. [2007], Chapon et al. [2013] used Adaptive Mesh Refinement (AMR) simulations and reached the impressive separation of $\sim 1\text{pc}$ for the binary. More recently, Roškar et al. [2014] has added more layers of complexity in the physics behind these N-body/SPH simulations, by including radiative cooling and heating, feedback and star formation processes in the process of the merger of two Milky-Way-like galaxies. More and more levels of complexity are being added to the general picture, as in the work of Fiacconi et al. [2013], which add the formation, dynamics, and implication of gaseous clumps in the pairing phase of two BHs.

5.3 UNEQUAL MERGERS

When two merging galaxies differ in their masses, besides dynamical friction other mechanisms play a crucial role in the evolution of the process, as shown, for instance, by Governato et al. [1994], Taffoni et al. [2003]. More particularly, (i) *tidal stripping*, i.e. the effect of tidal fields on to the primary halo, can significantly reduce the mass of the secondary, and (ii) *tidal heating*, the effect of short impulses by bound particles on to the secondary galaxy, heats the system and can lead to a partial dissolution of the primary galaxy. Due to this, we could have configurations in which the less massive black hole is set on a peripheral orbit within the primary galaxy.

Addressing the configuration of unequal mergers have laid bare features that were hiding in the equal mass ratio case. Simulations of this configuration with different techniques have led to the conclusion that there is a relationship between the central concentration of the interacting galaxies and

the specific geometry of the encounter. This is important to understand how the secondary SMBH can sink towards the primary [Kazantzidis et al., 2005, Callegari et al., 2009, 2011, Van Wassenhove et al., 2012, 2014].

Depending on the amount of gas on the disc in the secondary galaxy, we have different situations. If the starburst fraction close to the secondary SMBH is large, $\gtrsim 10\%$, there will be important transformations of the structure and dynamics, and the merger is said to be *wet*. On the other hand, if the secondary remains in a peripheral orbit far away, of some $\sim 1\text{kpc}$, the potential merger between the SMBHs will merely depend on dynamical friction, due to the lack of starburst, and the merger is called *dry* [Callegari et al., 2009, 2011, Khan et al., 2012].

5.4 MINOR MERGERS

An extreme case of unequal mergers are minor mergers, with mass ratios of about 1:10. In this case, the probability that the two MBHs do not pair and form a binary is high. For minor mergers not only the gas contents and distribution matters, but also the initial orbital configuration of the MBHs. Moreover, their masses might significantly change during the process, which renders the whole problem even more complicated [see e.g. Callegari et al., 2011].

If the merger of the two galaxies is co-planar and prograde, the outcome has a higher mass ratio MBHs pair, with a small separation. However, if there is some inclination the binary is very likely doomed to never form, because tidal shocks prevent accretion on to the secondary MBH. Nonetheless, if the initial setup for the orbits of the galaxies is such that the pericentre distance is small, and considering a gas-rich environment, the pairing will be even more difficult due to the relative higher velocities of the small periapsis distance. In this case, ram pressure strips the gas and notably reduces starburst and accretion, hence reducing the possibility of pairing. Where the threshold lies turns out to be a complicated problem which depends on the geometry, amount of gas, internal structure for the galaxies, and the dynamics of the MBHs at the shortest separations [Khan et al., 2012].

6

GETTING THE GAS TO THE MASSIVE BLACK HOLES

6.1 PREAMBLE

If minor mergers are efficient, the secondary MBH can sink towards the centre of the merged galaxies to meet and form a pair with the primary MBH and, eventually, a binary, as showed by Escala et al. [2005], Dotti et al. [2006, 2007, 2009], Fiacconi et al. [2013]. These simulations, and others, typically reach distances of $\lesssim 100\text{pc}$ down to a few parsecs, which is still very far away from the GW-led phase, which starts at about $\lesssim 10^{-3}\text{pc}$. The ulterior evolution from these numerical models is not just a problem of resolution, but of “micro-physics”, such as star formation episodes, gas thermodynamics, self-gravity, gas dissipation (radiative cooling), viscosity, turbulence and rotation, to mention the most relevant physical phenomena.

We mainly distinguish two well-separated regimes in this problem. Firstly, we have the nuclear-disc-driven one, typically from 100pc to $\sim 0.1\text{pc}$. The MBHs evolve due to perturbations in the density field [Escala et al., 2005, Dotti et al., 2006, 2007, 2009]. Later, we have the disc-driven migration, the torques exerted from the binary on to the gas causes the opening of a gas cavity. The binary is hence located in a region in which gas density is orders of magnitude below what we can expect to have in the main disc [Gould and Rix, 2000, MacFadyen and Milosavljević, 2008, Hayasaki, 2009, Cuadra et al., 2009, Roedig et al., 2011, Noble et al., 2012, Kocsis et al., 2012, Roedig et al., 2012, D’Orazio et al., 2013a, Farris et al., 2014]. There is an interplay between the disc and the binary itself, because of angular momentum transfer, and the gravitational and viscous torques.

6.2 NUCLEAR-DISC-DRIVEN

The circumnuclear disc (CND) is usually modelled as a Mestel distribution, which is self-gravitating and vertically pressure-supported, with an aspect ratio of between 0.1 and 0.05. So as to prevent gravitational instabilities, one adapts a Toomre parameter $Q \gtrsim 3$, and the effect of the stellar bulge is represented with a Plummer sphere more massive than the disc itself [Escala et al., 2005, Dotti et al., 2006, 2007, 2009].

In this setup the secondary is dragged into a co-planar co-rotating orbit around the primary and remains in a circular orbit [Dotti et al., 2006, 2007, 2009], which is faster if the disc is denser (colder). However, retrograde discs lead to a *flip* in the angular momentum [Dotti et al., 2009]. This renders the problem particularly complex and interesting, which is why we decided to study it in detail, and formed part of my research for my PhD.

As usual, and going back to the thought of Donald Lynden-Bell, once we have understood the ideal consideration in which the disc is treated as perfectly smooth, the question about the effect of fragmentation arises. We indeed expect this to happen in nature, because of gravitational instabilities



Figure 6: Gas distributed around a binary of two SMBHs from the multiple infall of gaseous clouds. The black holes are depicted as black dots. This is the same image as in the cover page, and is the result of one of the many different simulations I did to address the problem of how gas distributes around binaries of SMBHs.

triggering stellar formation. Hence clumpyness needs to be addressed. Indeed, I did this also during my PhD with a colleague of mine in Chile, and we published the results in del Valle et al. [2015] but I do not include this work here because I did not lead it.

6.3 CIRCUMBINARY-DISC-DRIVEN

Once the binary forms, the gas should be distributed as a circumbinary disc, dominated by the binary’s gravity and quadrupolar field. However, as I mentioned before, this relies on a number of theoretical assumptions that needed to be checked. In particular, the migration phase is dominated by the torques excited by resonances, which are in charge of bringing the secondary very close to the centre. This is essential to understand whether the binary will ever merge or stall, with the two MBHs dancing around each other for ever.

How do we form a gap? How is it formed? del Valle and Escala [2012, 2014] find that the binary needs to be surrounded by a geometrically thick disc for a gap to be open. There is a strong coupling between the binary and the disc due to viscous torques in the disc that leads to a slow orbital decay [Artymowicz and Lubow, 1994, 1996, Gould and Rix, 2000, Armitage and Natarajan, 2002, Armitage et al., 2013]. If this gap does indeed open, the associated timescale of the hardening process is slower than that of the nuclear-disc-driven phase.

Moreover, there seems to be small persistent discs around each BH forming within the inner cavity, because of the porous nature of the inner edge of the main circumbinary disc [Farris et al., 2014, Roedig et al., 2011, 2012]. These “mini discs” are essential to understand the ulterior evolution of the binary, because the torques might be significantly enhanced. Complementary, as observed by Roedig et al. [2011], the binary seems to have a tendency to saturate its eccentricity around the value of 0.6 [Roedig et al., 2011].

All of these observations have far-reaching implications for the evolution of both the binary and, to some extent, also of the CND.

7

GOALS OF THIS THESIS

The main question I wanted to address when I first arrived to the Albert Einstein Institute was the following: if black holes exist and form binaries, where do these binaries form and what characteristics do they have? One of my strongest motivations was, obviously, the generation of sources of gravitational waves.

Sooner than I thought, I had already spent all of my available time with this question, which had turned out to be even more interesting than I previously thought. From “pure” stellar dynamics systems (conceptually thinking) such as galactic nuclei or globular clusters, to the cores of interacting galaxies with plenty of gas, the problem showed to be anything but straightforward. Moreover, the implications of addressing this problem increased with time, just as the complications did.

7.1 FROM DYNAMICS

I started writing a new direct-summation N -body code to study the dynamical formation of binaries of black holes. For this, I adopted a softening into the integrator, which means I added a constant value for the square of the separation between two stars or black holes when they come too close to each other. This is a common trick in the field of stellar dynamics, because a proper integration of very hard binaries requires very delicate and convoluted numerical techniques that would have exceeded my time to deliver this PhD. Nevertheless, in spite of this trick, the integrator can give a robust idea of what is happening in a dense stellar system such as a globular cluster, down to the moment in which it collapses. This is the most interesting moment, because the concentration of compact objects at the centre is higher, and one expects binaries to form binaries that eventually will emit GWs. However, I was intrigued by the role of softening and investigated this in detail. The results regarding the softening that one adopts are very interesting, because different values for this constant will have a global impact on the whole dynamics of the sample of thousands of stars. I will keep working on this integrator in the future, which I have made publicly available to anyone. The integrator and several tools can be found in

<http://gravity.xyz/>

As Sverre Aarseth told us when I gave him the draft of the paper which is now published, now “it is just a matter of another ten years until you have included a proper treatment of binaries and regularisation into your integration scheme”.



Figure 7: Artist impression of the accretion disc around the black hole gargantua, warped because of gravitational lensing. Credit: Warner Bros, Entertainment Inc. and Paramount Pictures Corporation. Author: double negative, <http://www.dneg.com>

7.2 TO ACCRETION

As I mentioned many times, my main goal was not solely the running of numerical experiments, but the understanding of the formation of binaries of black holes, either of stellar masses or supermassive. This is what led me to address different problems. The first problem was the impact of a realistic treatment of gas accretion on to a secondary MBH on its way to merge with a primary MBH embedded in a gaseous disc. The more difficult question to investigate were retrograde orbits, and the effort led us to the publication of a paper in which we gave a detailed description of this problem with different interpretations of what accretion should be in reality.

This was my first contact with binaries of SMBHs and gas, and I found the problem motivating, but one fundamental question emerged when I started reading more and more literature on this subject. Although it has been assumed by many authors, to my surprise I did not find a detailed analysis of how gaseous structures form around binaries of SMBHs. This is a pivotal, as explained in this introduction a few times, because how gas is distributed around the binary, and how the binary interacts with it will settle (i) whether the binary will merge within a Hubble time or not, (ii) the orbital characteristics of the binary, if any, when it reaches the GW led regime and (iii) any potential electromagnetic counterpart that might be triggered in the process.

I hence decided to lead the effort of a rather gargantuan (in the sense of titanic, not of the black hole of Fig. 7) effort to analyse the effect of repeated gas clouds infalls on to a binary of two SMBHs in a galactic centre. I was interested by the formation of structures and their architecture around the binary, and also by the fate that this binary would have depending on the infall properties. The most interesting result I found was that the formation of disc-like structures in the gas is very challenging, as one can see for instance in Fig. 6. The scientific results are available to anyone, because the papers are published, but the tools, scripts, and code as well, because science should be reproducible and open to anyone. Everything can be found in the following web site:

<http://multipleclouds.xyz/>

7.3 TO STUDY BLACK HOLE BINARY SYSTEMS

In view of all I explained and introduced in this first part of the PhD, the main aim and goal of my study was the addressing of a a very simple question.

Where, when and how do black holes form binaries? I have studied the dynamical formation of binaries in dense stellar systems by developing a completely new direct-summation N -body code for the GPUs and multiple CPUs which I have released publicly. In a completely different range of masses, differing by some 5 orders of magnitude, the formation of binaries of supermassive black holes in galactic nuclei has the additional complication of gas dynamics. As I studied massive binaries in retrograde discs, I realised that, although it has always been assumed to be so, the gas does not necessarily need to be distributed around the binary in a disc-like structure. The implications of my findings are important for a plethora of different fields in astrophysics, not just in gravitational-wave astronomy.

Part II

PROJECTS

DYNAMICAL EVOLUTION OF DENSE STELLAR SYSTEMS

The project presented on this chapter was published under:

GRAVIDY, A GPU MODULAR, PARALLEL DIRECT-SUMMATION
 N -BODY INTEGRATOR: DYNAMICS WITH SOFTENING.

Cristián Maureira-Fredes & Pau Amaro-Seoane
 Monthly Notices of the Royal Astronomical Society,
 Volume 473, Issue 3, 21 January 2018, Pages 3113–3127.

In this chapter we present the initial version of GRAVIDY (Gravitational dynamics), a highly-modular, direct-summation N -body code written from scratch using GPU technology ready to integrate a pure dynamical gravitational system. In §8.1 we present in detail the structure of the code, the most relevant and innovative parts of the algorithm, and their implementation of the scheme in the idiom of GPU computing. In §8.2 we check our code with a series of well-known tests of stellar dynamics for a dense stellar system and evaluate global dynamical quantities and we also evaluate the performance of the GPU version against the CPU one. In §8.3 we study the implications of using a softening parameter in these kind of simulations. In §8.4 we present the implementation of the relativistic corrections, and a set of tests. In §8.5 we summarise our work and give a short description of the immediate goals that will be described in upcoming publications.

We have decided to focus on single-node clusters (meaning one or more GPU cards embedded in a host PC) and traditional multi-CPU clusters (e.g. Beowulf clusters), since this setup is more common to most users who aim to run middle-scale simulations. In the appendices we give a succinct description on how to download the code, how to compile it, and the structure of the data. We also include a set of python tools to analyse the results. Moreover, we also introduce a simple visualisation tool based on OpenGL, which can provide us with information sometimes difficult to obtain with two-dimensional graphics. In particular, we have made a significant effort in documentation and modularity, since it is our wish that the code is used, shaped and modified at will.

8.1 THE CURRENT ALGORITHM

8.1.1 *The integration scheme*

In this section we give a very brief introduction to the numerical N -body problem. We refer the reader to e.g. [Aarseth, 2003, Heggie and Hut, 2003] or the excellent on-line course “The art of computational science”¹. The evolution of an N -body system is described by the second order ordinary differential equation

¹ <http://www.artcompsci.org/>

$$\ddot{\mathbf{r}}_i = -G \sum_{\substack{j=1 \\ j \neq i}}^N m_j \frac{(\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3}, \quad (4)$$

where G is the gravitational constant, m_j is the mass of the j th particle and \mathbf{r}_j the position. We denote vectors with bold fonts. The basis of the problem is purely dynamical, because the orbital evolution is determined exclusive by the gravitational interaction.

The total energy of the system is a useful quantity to keep track of every time step in the integration. It is given by the expression

$$E = \frac{1}{2} \sum_{i=1}^N m_i \mathbf{v}_i^2 - \sum_{i=1}^N \sum_{j>i}^N \frac{Gm_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (5)$$

where \mathbf{v}_i is the velocity of the particle i .

To numerically integrate the system of equations we adopt the 4th-order Hermite integrator (H4 from now onwards) presented in [Makino, 1991, Makino and Aarseth, 1992] [and see also Aarseth, 1999, 2003]. H4 is a scheme based on a predictor-corrector scenario, which means that we use an extrapolation of the equations of motion to get a predicted position and velocity at some specific time. We then use this information to get the new accelerations of the particles, later we correct for the predicted values using interpolation based on finite differences terms. One can use polynomial adjustment in the gravitational forces evolution among the time because the force acting over each particle changes smoothly (which is the reason why adding a very massive particle representing e.g. a supermassive black hole will give you sometimes a headache). To advance the system to the following integration time we approximate the equations of motion with an explicit polynomial. This prediction is less accurate, but it is improved in the corrector phase, which consist of an implicit polynomial that will require good initial values to scale to a good convergence.

This is a fourth-order algorithm in the sense that the predictor includes the contributions of the third-order polynomial, and after deriving the accelerations, adds a fourth-order corrector term. In the remaining of this chapter we focus on the implementation of the scheme into our GPU (and CPU) code and how to maximise all of the computational resources available. For a detailed description of the idea behind H4, we refer the reader to the article in which it was presented for the first time, [Makino and Aarseth, 1992].

An advantage of the choice for H4 is that we can use the family of Aarseth's codes (among others) as a test-bed for our implementation. These codes – some of which adopt H4, but not all of them– have been in development for more than 50 years. The codes are public and have been widely used and validated, improved and checked a number of times by different people, they have been compared to other codes and even observational data. In this regard, to test our implementation and parallelisation of H4, the access to the sources of the codes is an asset.

8.1.2 Numerical strategy

A main goal in the development of GRAVIDY is its *legibility*. We have focused in making it easy to read and modify by other users or potential future developers without compromising the computational performance of the algorithm.

Listing 8.1: `gprof` output of a serial GRAVIDY simulation, using 1024 particles up to 1 N -body time unit, showing the amount of time of the particle-particle force interaction. Some output columns are omitted.

```

Flat profile:

Each sample counts as 0.01 seconds.
%   self
time seconds   calls name
92.5   12.05 125186180 Hermite4::force_calculation(...)
 1.1    0.15   5161 Hermite4::predicted_pos_vel(...)
 0.3    0.04   5161 Hermite4::correction_pos_vel(...)
 0.2    0.03   5161 Hermite4::next_integration_time(...)
 0.2    0.03   5161 Hermite4::find_particles_to_move(...)
 0.0    0.00   5161 Hermite4::update_acc_jrk(...)
...

```

This means that we have made a significant effort in keeping a clear structure in the source code so that, in principle, it can be well understood by somebody who has not previously worked with it with relatively little effort. The modularity of the code should allow new users to easily implement new physics or features into it or adapt it to the purposes they seek. It is unfortunately easy—at least to a certain extent—to miss either clarity in coding or performance, when trying to have both in a code. For instance, if we want to obtain the best performance possible, one has to use low-level instructions that for an outside user might result into something difficult to understand when reading or trying to modify the source code. On the other hand, name conventions for files, functions and variables might become a burden to certain applications.

While most existing N -body codes have achieved certain balance between the two to some degree, it is difficult to adapt them to new architectures and technology to boost their performance. For the development of GRAVIDY, we have followed the next steps:

SERIAL IMPLEMENTATION: The first logical step in developing a code using GPU technology is—in our opinion—to have a full, one-thread CPU version. This allows us to check for the robustness of the code, as it is usually easier to find and fix any potential bugs when introducing new features or improving the performance.

PROFILING AND ASSESSMENT: After that, we localise the main bottlenecks of the code. One obvious one is the computation of the gravitational forces, but it is in general a good idea to check for other potential hindrances. For this, we typically employ `gprof`².

In the case of our CPU code, for instance, we see that more than 93% of the wall clock time goes into the updating of a particular function, the gravitational interaction, as displayed in listing 8.1. There are some tools to perform the profiling of CUDA application, but they are not discussed in this work. The idea of keeping an updated CPU version of the code, helps to find potential errors or bottlenecks of the parallel version.

GRANULARITY AND HOT-SPOTS: After having identified the bottlenecks of the code, the objective is to use parallel computing to attack them to speed up the calculations. Ideally, this would be a one-time task, but

² GNU `Gprof` Documentation: <http://www.cs.utah.edu/dept/old/texinfo/as/gprof.html>

the fact is that the implementation, which will be described in detail later, depends on the hardware at our disposal, which keeps quickly evolving in the case of GPU technology, the release of new libraries and new ideas to speed up the force calculation. This means that the code is on a cycle process of constant improvement, and has to be revisited regularly.

Before we get into the details of the parallelisation, it is important to introduce the concept of *granularity*. This is a classification of the algorithm at play based on the communication and performance. We usually distinguish two types of granularity: *fine-grained* algorithms are computationally lightweight, because the task performed (in this case by threads) is small compared with communication, which is a frequent process. On the other hand, *coarse-grained* algorithms do the opposite. They require less communication but are computationally more demanding for each task.

A good example of the difference in granularity is the comparison of Message Passing Interface (MPI) and GPU computing: usually, a code relying on MPI splits a bigger task in many sub-tasks, which are sent to different computational nodes. After finishing the computation, the data are gathered in a so-called “master” node. Programming in GPU is quite different; there is a frequent communication between the CPU and the GPU and the tasks assigned to the many threads available in the GPUs are quite small. This is why in GPU computing, the finer the granularity, the more efficiently we can follow a parallelisation scheme.

OPTIMISATION: In the process of parallelisation of a code in CUDA one has to bear in mind the following characteristics, inherent to this programming language and GPU architecture: (i) while the use of local memories improves the performance of a CUDA code, the GLOBAL memory (the GPU off-chip memory) is accessed in “chunks”, multiples of 32 bytes, the so-called *warps*. Failing to use them completely leads to a so-called lack of *occupancy*, (ii) Indeed, the usage of SHARED memory (the CUDA-block shared memory) can lead to a number of issues if not employed properly: since we are accessing chunks of memories in warps, we *have* to retrieve information from the GLOBAL memory also in chunks. Moreover, if threads within a warp access different banks of memory, we will confront a *bank conflict*. E.g. two threads access the same value from the local memory. (iii) *Divergence* is another issue to take into account. When a control flow statement (e.g if-else) is inside a kernel execution and it splits the threads to perform different pieces of code, the problem is that one of those tasks could be more expensive computationally. Hence, at the moment of synchronisation, the threads executing the less expensive code will need to wait for the others to achieve a synchronisation. While it is not forbidden to use control flow, we need to at least leave one warp do some task. This means that control flow must be at warp level; different warps can do different tasks without running into divergence. (iv) The hardware that we have at our disposal plays a crucial role: GPU chips are formed by several Streaming Multiprocessor (SM), which contain the GPU cores (Streaming Processor, SP). All SM can execute only one warp per core at a time, so that we must try to use them constantly, which is what we refer to as the *occupancy*. This will be determined by the usage of every thread in a block respect to the registers and the SHARED memory, which is GPU-dependent.

8.1.3 *Particular choices*

OBJECT ORIENTED PROGRAMMING: Object oriented programming (OOP) is a powerful paradigm that allows us to program an algorithm as objects interactions. In GRAVIDY, we use OOP. The reason beneath it is related to our parallelisation scheme, which is described below, more concretely with the data structure we have chosen.

We have mainly two possible choices for data structures: classes with arrays, or Arrays of Objects, which follows the basic idea of Struct of Arrays (SoA) and Array of Structs (AoS). For GRAVIDY we have chosen *classes* with arrays for the main units of the program structure. It is a good strategy to minimise the data transfer between Host and Device, so as to avoid having large communication times.

Calling the GPU kernels³, we only transfer a piece of information relative to all the bodies, e.g. only the masses of the bodies or only the position of a certain amount of bodies, but not all of them. Because of this, we deem it more practical to handle the data transfer of our scheme by using arrays containing information on different particles.

In practice, this means that we use the GPU to only compute some specific integration steps like the force calculation, and the energy. Hence, we have to transfer to the GPU at every iteration the predicted positions and velocities, to then transfer from the GPU the new calculated forces (i.e. the acceleration and its first derivative). It is also possible to have all the integration steps on the GPU, but we do not see a significant improvement for our integrator scheme. Higher integration schemes show a better performance using the whole process on the GPU [Capuzzo-Dolcetta et al., 2013]

It is not required to update the forces of all the particles, so that we encapsulate the information of the active particles, and then we transfer the AoS to the GPU. All the remaining attributes of the bodies (i.e. those not transferred to the GPU) are just class-members (arrays), and need to be in the host CPU. An example of this could be large linear arrays, such as the time steps of the particle.

CLASS DISTRIBUTION: Since our code is using OOP, we describe a brief interaction between the classes in Fig. 8. The main header, `common.hpp`, contains the definition of the constants, structures, macros, etc. The idea behind this model is to easily be able to add more features in upcoming versions of our code, from new utilities functions to new integration schemes.

Every class is in charge of a different mechanism, from getting the integration options from command-line, to the different integration methods using parallelism or not⁴.

DOUBLE-PRECISION (DP) OVER SINGLE-PRECISION (SP): Using DP or SP in N -body codes has been already addressed by different authors in the related literature [see e.g. Hamada and Iitaka, 2007, Nitadori, 2009, Gaburov et al., 2009]. Using DP is not the best scenario for GPU computing, because there is a decrease factor in the maximum performance that a code can reach. We can reach only half of the theoretical maximum performance peak, which depends on each individual card: for

³ We depict as “kernel” the set of functions which can be executed on the GPU

⁴ For more information, please refer to the code documentation

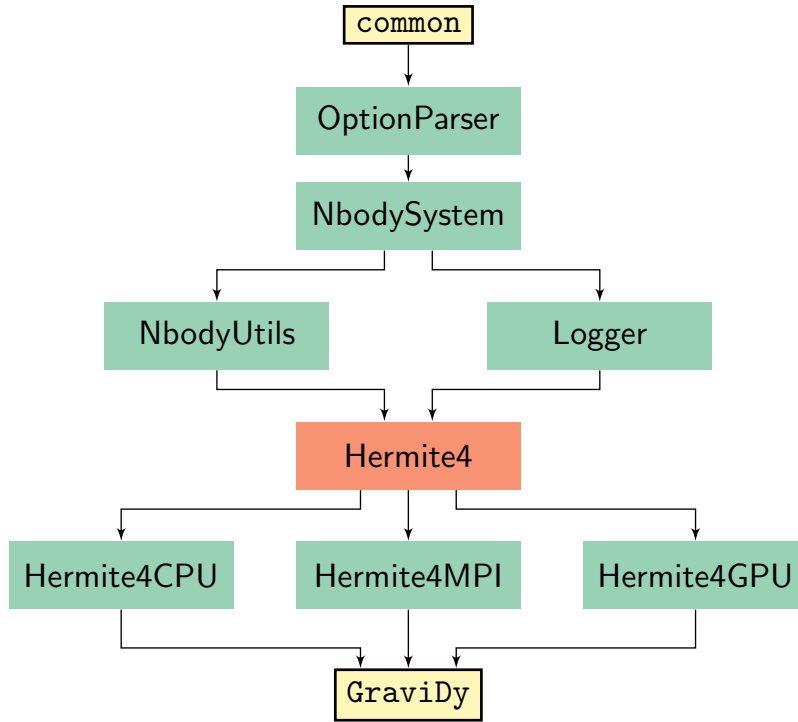


Figure 8: Class diagram of the code that shows the hierarchy of the application structure (GRAVIDY).

example, the NVIDIA Tesla C2050/M2050 has a peak of the processing power in GFLOPs 1030.46 with SP, but only 515.2 with DP.

We choose DP for a more accurate numerical representation, because it provides us a simple way of getting better energy conservation, at the expenses of performance. There are different approaches, like the mixed-precision, [Aarseth, 1985], and pseudo DP (Nitadori 2009, currently used in the code ϕ -GPU, Berczik et al. 2011). These offer a relatively more accurate representation (compared to SP) without a big impact in performance.

8.1.4 The implementation scheme

These are the steps that GRAVIDY follows when running a simulation:

1. Memory allocation of the CPU and GPU arrays.
2. Initialisation of the variables related to the integration.
3. Copy the initial values of positions, velocities and masses of the particles to the GPU to calculate the initial system energy, and calculate the initial acceleration and its first time derivative, the so-called “jerk”. The cost of this force calculation is $O(N^2)$.
4. Copy the initial forces from the GPU to CPU.
5. Find the particles to move in the current integration time, N_{act} , with a cost $O(N)$.
6. Save the current values of the forces, to use them in the correction step, with a cost $O(N)$.

7. Integration step:

- a) Predict the particle's positions and velocity up to the current integration time, with cost $O(N)$.
- b) Copy of the predicted positions and velocities of all the particles from the CPU to the GPU.
- c) Update the N_{act} particles on the GPU, which is explained in detail in §8.1.5.
 - i. Copy the N_{act} particles to a temporary array on the GPU.
 - ii. Calculate the forces between the particles on the GPU, with a cost $O(N_{\text{act}} \cdot N)$.
 - iii. Reduce forces on the GPU.
 - iv. Copy the new forces from the GPU to the CPU.
- d) Correct the position and velocity of the N_{act} updated particles on the CPU, $O(N_{\text{act}})$.
- e) Copy the positions and velocities of the corrected N_{act} particles from the CPU to the GPU.

GRAVIDY adheres to the usual good practises of the beginning of the development of every direct-summation N -body code:

- *Direct-summation*, also known as particle-particle strategy, This approach is the simplest way to address the task of calculating the exerted force by all the $N - 1$ bodies on a single body that we need to update at certain time step. This brute-force procedure has an order $O(N^2)$, which represents the bottleneck of the algorithm.
- *Softened point-mass potential*, as an alternative in this version of the code to a proper close encounter regularisation. All particles are represented by a dimensionless point mass. We introduce a softening parameter (ϵ) in the distance calculation between two bodies while we get the new forces,

$$\ddot{\mathbf{r}}_i = -G \sum_{\substack{j=1 \\ j \neq i}}^N \frac{m_j}{(r_{ij}^2 + \epsilon^2)^{3/2}} \mathbf{r}_{ij}, \quad (6)$$

so as to handle the situation in which two bodies get closer.

- *Block time steps*, It is not straightforward to have an N -body code using individual time steps in parallel computing, because the idea behind massive parallelism is to perform the same task on different data chunks. We use the block time steps algorithm [Press, 1986], to update group particles simultaneously. This scheme has been adopted by a number of authors [Portegies Zwart et al., 2001, Hut, 2003, Aarseth, 1999, 2003, Harfst et al., 2008, Nitadori and Aarseth, 2012].

The main idea is to have several blocks (groups) of particles sharing the same time steps. This decreases the amount of operations of the integration process, which allows us to obtain a similar accuracy to the individual time steps scheme.

In this scenario the parallelisation is ideal, because compared with the individual scheme, we have several chunks of threads working on different blocks of particles.

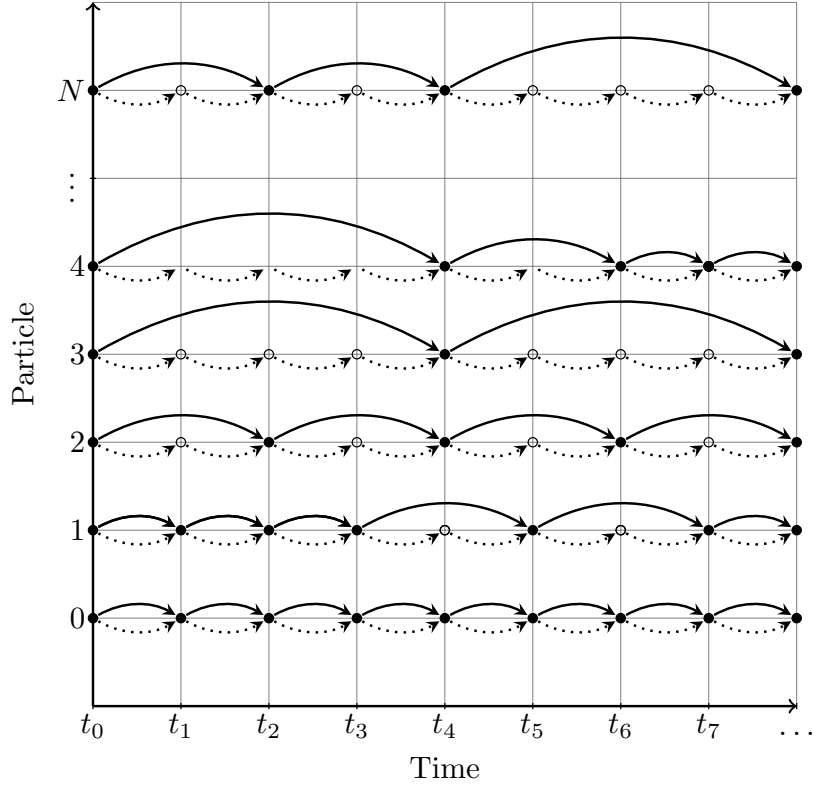


Figure 9: Block time steps illustration. The different blocks are represented by the length of the jump. Each particle is predicted (but not moved) at every time t (dotted arrows), even if it is not the turn of their block time step (empty circles). The particles will be updated (and moved) only in their corresponding block time step (filled circles). Whenever a particle is updated, its block time step can be change. In this illustration, particles 1, 4 and N change their block (length of the jump).

In this approach, particle i is part of the lower n th block time step between two quantities,

$$2^n \Delta t_s \leq \Delta t_i < 2^{n+1} \Delta t_s. \quad (7)$$

Δt_i is determined by equation 24, and Δt_s is a constant.

The particles distribution among the different blocks is determined by the following condition,

$$\Delta t_{i,new} = 2^{\lceil \log_2 \Delta t_i \rceil - 1}, \quad (8)$$

and is described in Fig. 9.

For the boundaries of the time steps, we use the Aarseth criterion [Aarseth, 2003] for the lower limit of the time steps allowed in the system:

$$\Delta t_{min} = 0.04 \left(\frac{\eta_I}{0.2} \right)^{1/2} \left(\frac{R_{cl}^3}{\bar{m}} \right)^{1/2}, \quad (9)$$

where η_I is the initial parameter for accuracy, typically 0.01; R_{cl}^3 is the close encounter distance and \bar{m} the mean mass.

Usually $\Delta t_{\text{min}} = 2^{-23}$. On the other hand, we set a maximum time step $\Delta t_{\text{max}} = 2^{-3}$. When updating a particle's time step, if it is out the this boundaries, we modify the value to Δt_{min} if $\Delta t < \Delta t_{\text{min}}$, and to Δt_{max} if $\Delta t > \Delta t_{\text{max}}$.

The algorithm uses a Hermite of 4th order to integrate the evolution, which follows these steps:

1. Select the particles with the minimum $t_{i,0} + \Delta t_{i,0}$ and assign this minimum time to the global integration time t .
2. Calculate the predicted position ($\mathbf{r}_{i,\text{pred}}$) and velocity ($\mathbf{v}_{i,\text{pred}}$) at time t , for all i particles, using these expression:

$$\mathbf{r}_{i,\text{pred}} = \mathbf{r}_{i,0} + \mathbf{v}_{i,0}\Delta t_i + \frac{1}{2!}\mathbf{a}_{i,0}\Delta t_i^2 + \frac{1}{3!}\dot{\mathbf{a}}_{i,0}\Delta t_i^3 \quad (10)$$

$$\mathbf{v}_{i,\text{pred}} = \mathbf{v}_{i,0} + \mathbf{a}_{i,0}\Delta t_i + \frac{1}{2!}\dot{\mathbf{a}}_{i,0}\Delta t_i^2 \quad (11)$$

in this case, the $\Delta t_{i,0}$ is calculated as follows:

$$\Delta t_i = t - t_{i,0} \quad (12)$$

3. Calculate the acceleration ($\mathbf{a}_{i,1}$) and the jerk ($\dot{\mathbf{a}}_{i,1}$) using the predicted position and velocity only for the i -particles in which the global time t , is equal to $t_{i,0} + \Delta t_{i,0}$ ⁵.

$$\mathbf{a}_{i,1} = \sum_{\substack{j=0 \\ j \neq i}}^N Gm_j \frac{\mathbf{r}_{ij}}{(r_{ij}^2 + \epsilon^2)^{\frac{3}{2}}}, \quad (13)$$

$$\dot{\mathbf{a}}_{i,1} = \sum_{\substack{j=0 \\ j \neq i}}^N Gm_j \left[\frac{\mathbf{v}_{ij}}{(r_{ij}^2 + \epsilon^2)^{\frac{3}{2}}} - \frac{3(\mathbf{v}_{ij} \cdot \mathbf{r}_{ij})\mathbf{r}_{ij}}{(r_{ij}^2 + \epsilon^2)^{\frac{5}{2}}} \right], \quad (14)$$

where

$$\mathbf{r}_{ij} = \mathbf{r}_{j,\text{pred}} - \mathbf{r}_{i,\text{pred}}, \quad (15)$$

$$\mathbf{v}_{ij} = \mathbf{v}_{j,\text{pred}} - \mathbf{v}_{i,\text{pred}}, \quad (16)$$

$$r_{ij} = |\mathbf{r}_{ij}| \quad (17)$$

It is important to note, that $(\mathbf{v}_{ij} \cdot \mathbf{r}_{ij})$ correspond to the *dot product*, and not a simple multiplication.

4. Calculate the 2nd and the 3rd derivative of the acceleration ($\mathbf{a}_{i,1}^{(2)}, \mathbf{a}_{i,1}^{(3)}$) using the third-order Hermite interpolation polynomial constructed using $\mathbf{a}_{i,0}$ and $\dot{\mathbf{a}}_{i,0}$:

⁵ Eq.(2) of Makino and Aarseth [1992] has a typo in the sign of the second term in the sum of $\dot{\mathbf{a}}_{i,1}$.

$$\mathbf{a}_{i,1}(t) = \mathbf{a}_{i,0} + \dot{\mathbf{a}}_{i,0}\Delta t_{i,0} + \frac{1}{2}\Delta t_{i,0}^2\mathbf{a}_{i,0}^{(2)} + \frac{1}{6}\Delta t_{i,0}^3\mathbf{a}_{i,0}^{(3)} \quad (18)$$

where $\mathbf{a}_{i,0}$ and $\dot{\mathbf{a}}_{i,0}$ are the acceleration and jerk calculated at the previous time t , the second and third acceleration derivatives $\mathbf{a}_i^{(2)}$ and $\mathbf{a}_i^{(3)}$ are given by:

$$\mathbf{a}_{i,0}^{(2)} = \frac{-6(\mathbf{a}_{i,0} - \mathbf{a}_{i,1}) - \Delta t_{i,0}(4\dot{\mathbf{a}}_{i,0} + 2\dot{\mathbf{a}}_{i,1})}{\Delta t_{i,0}^2} \quad (19)$$

$$\mathbf{a}_{i,0}^{(3)} = \frac{-12(\mathbf{a}_{i,0} - \mathbf{a}_{i,1}) - 6\Delta t_{i,0}(\dot{\mathbf{a}}_{i,0} + \dot{\mathbf{a}}_{i,1})}{\Delta t_{i,0}^3} \quad (20)$$

where $\mathbf{a}_{i,1}$ and $\dot{\mathbf{a}}_{i,1}$ are the acceleration and the jerk at the time $t_i + \Delta t_i$.

5. After the previous calculation, it is necessary to add the corrections to the position and the velocity for the particle i at the time $t_i + \Delta t_i$

$$\mathbf{r}_{i,1} = \mathbf{r}_{i,pred} + \frac{1}{24}\Delta t_i^4\mathbf{a}_{i,0}^{(2)} + \frac{1}{120}\Delta t_i^5\mathbf{a}_{i,0}^{(3)} \quad (21)$$

$$\mathbf{v}_{i,1} = \mathbf{v}_{i,pred} + \frac{1}{4}\Delta t_i^3\mathbf{a}_{i,0}^{(2)} + \frac{1}{24}\Delta t_i^4\mathbf{a}_{i,0}^{(3)} \quad (22)$$

6. We then need to calculate the next time step for the i particle ($\Delta t_{i,1}$) and time t using the following formulae:

$$t_{i,1} = t_{i,0} + \Delta t_{i,0} \quad (23)$$

$$\Delta t_{i,1} = \sqrt{\eta \frac{|\mathbf{a}_{i,1}| |\mathbf{a}_{i,1}^{(2)}| + |\mathbf{j}_{i,1}|^2}{|\mathbf{j}_{i,1}| |\mathbf{a}_{i,1}^{(3)}| + |\mathbf{a}_{i,1}^{(2)}|^2}}. \quad (24)$$

Here η is the accuracy control parameter, $\mathbf{a}_{i,1}$ and $\dot{\mathbf{a}}_{i,1}$ are already known, $\mathbf{a}_{i,1}^{(3)}$ has the same value as $\mathbf{a}_{i,0}^{(3)}$ due the third-order interpolation, and $\mathbf{a}_{i,1}^{(2)}$ is given by:

$$\mathbf{a}_{i,1}^{(2)} = \mathbf{a}_{i,0}^{(2)} + \Delta t_i \mathbf{a}_{i,0}^{(3)} \quad (25)$$

We depict the process of the integration in Fig. 10.

8.1.5 The parallelisation scheme

As we have already mentioned, the bottleneck of any N -body code is the force calculation Fig. 10. In this respect, GRAVIDY is not different and a quick performance test to get the profile of our serial code yields almost 100% of the execution time in this calculation 8.1. We hence introduce a parallelisation scheme, which we discuss in detail now.

GRAVIDY is based on a direct-summation Hermite 4th order integrator and uses block time steps, so that in the force update process we have a nested loop for every i -active particle (which we will refer to from now with the subscript “act”). This means that for every particle which needs to be updated we

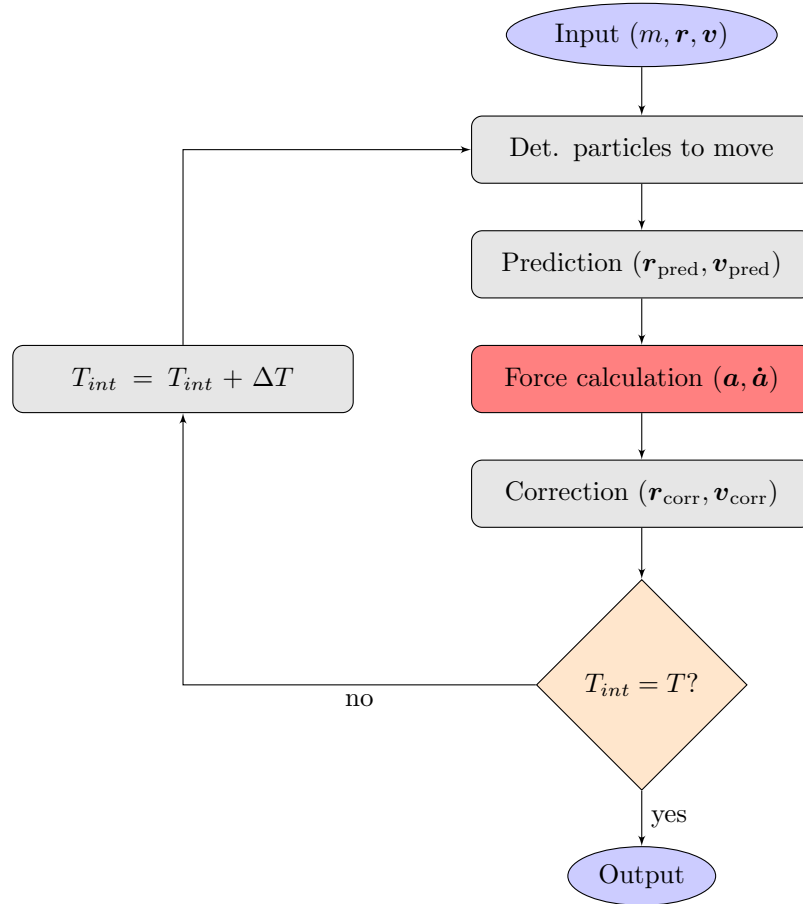


Figure 10: Hermite integration scheme illustration. This flow diagram shows the steps of every iteration. The force calculation is marked with red, because it is the bottleneck of the computational time.

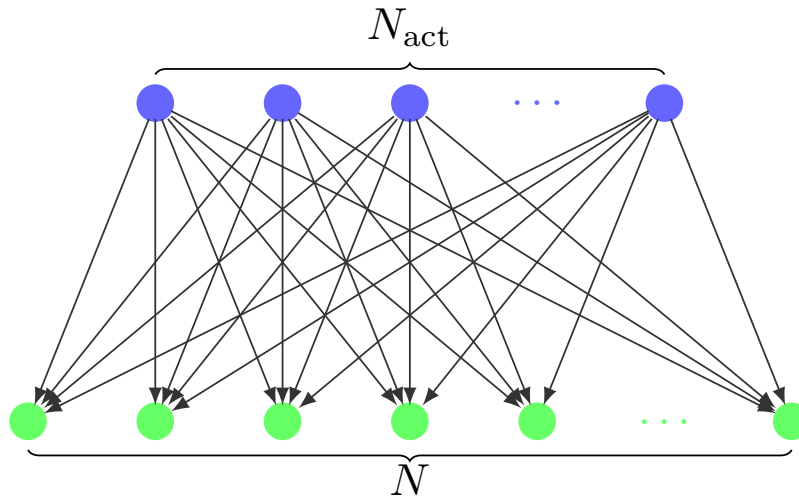


Figure 11: Relation between the particles which will be updated in a certain integration time (N_{act}) and the whole set of particles (N). The relation between the active particles and the others is $N_{\text{act}} \ll N$ in non-synchronisation times.

have a loop run on the whole set of particles of our system to check whether particle j is interacting with i .

The whole process scales with the amount of i -particles, as we can see in Fig. 11.

We then need to parallelise the loop corresponding to each of the i -particles. For each of them we circulate through all of the j -particles, and this is the process which needs to be parallelised. Although this is in principle a straightforward scheme, since we focus on GPUs, we run into the following issues:

1. A GPU can launch a large number of threads, easily up to thousands of them. In our scenario, however, the number of active particles N_{act} is very small compared to the total amount of particles (N). This has an impact on the performance: we do not use all available threads, we are integrating a grid of $N_{\text{act}} \times N$ forces. When the number of active particles is very low our *occupancy* will be bad.
2. Contrary, in the case in which we have to move all particles, we will have an $O(N^2)$ parallelism, which maximises the GPU power. In this case, however, the *memory bandwidth* is the limitation factor, since every particle requires all information about all other $N - 1$ particles.

The $N \times N$ forces grid allows us to use *tiles*, as explained in Nguyen [2007], and introduced in a direct-summation N -body for the first time by Nitadori [2009]. Using tiles means that we work with a square of the GPU grid of p rows and p columns, as depicted in Fig. 12. Instead of having the usual p^2 power that arises from all particles interacting with all particles, we need to only take into account the sum of p blocks in the SHARED memory of the GPU. Synchronisation only takes place at some particular moments, displayed as bold vertical lines in the figure, and we only have a load from GLOBAL to SHARED memory of p particles per tile, which implies a reusing of existing information instead of loading new one. The circles in the figure represent that, given a certain moment, the calculation can be in any column, because the threads are synchronised during the computation. We evaluate the interactions row by row and have parallelism in every column.

Whilst tiles are optimal when we need to evaluate force interaction for all particles in the system, GRAVIDY rarely deals with this situation due to the nature of the base algorithms and main goals behind it. The amount of particles that we need to update at every step is $N_{\text{act}} \ll N$, with N the total number of them. This means that we will be wasting computational resources, since some threads will not do the entire work. This is why we have to look for an alternative scheme which handles better the fact that we only need to move a subgroup of particles instead of the whole system.

It is better to have all particles handled by the GPU, and not only the active ones, because even though this subgroup is smaller, or even much smaller, it is more efficient from the point of view of the GPU, since the *occupancy* is improved. The parallelisation happens at j -level (i.e. when calculating the forces between active particles with the rest of the system). This idea was first implemented by Nitadori [2009], and has proven to yield very good performance.

The main ideas behind the j -parallelisation is how force calculation is done and the summation of the forces (“reduction”):

- *Force calculation*: The interaction between the i -particle and the rest of the system is distributed among the GPU threads, which means that we launch N threads, and each of them calculates its contribution with the

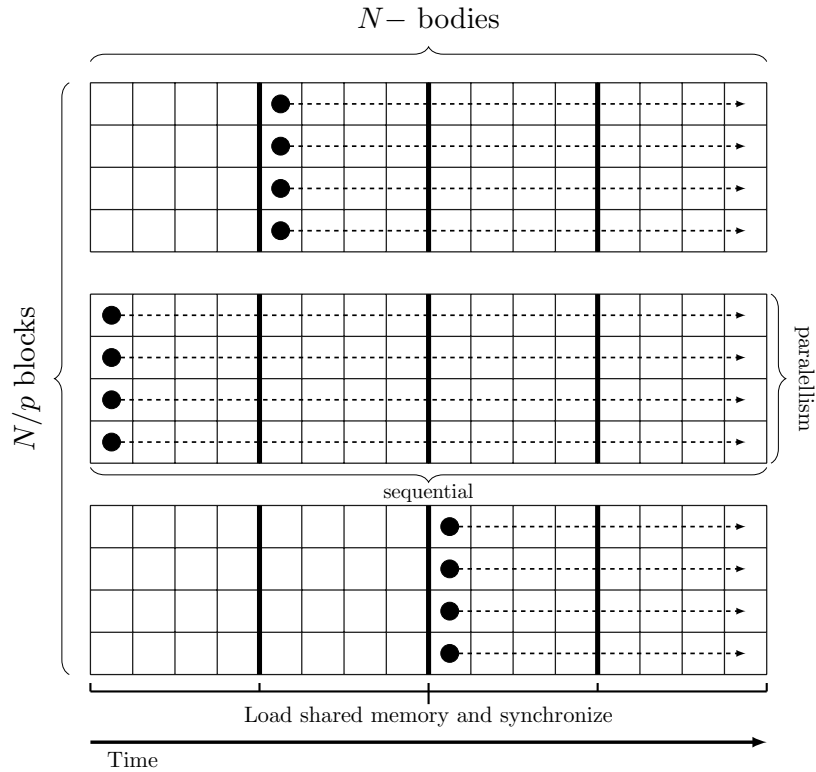


Figure 12: Grid configuration using the *tiles* approach (This figure is based on GPU-Gems3 book [Nguyen, 2007])

i -particle. After this calculation, we have an array where each element contains the contributions all the particles. j This corresponds to the upper part of Fig. 13, which illustrates a set-up of two GPUs. After the force calculation we end up with an array containing the information about the forces for all particles.

- *Force reduction*: In the lower part of the same Fig. we depict the summation of all of these forces, which is also performed in parallel, so that we use the blocks distribution of the GPU for this task.

So as to illustrate these steps, we give a particular example of a system with 1024 particles: let us assume that we need to move 300 of them ($N_{\text{act}} = 300$) at the moment of launching the kernel. Before this, we need to set up the kernel configuration. We define the grid size (number of thread blocks), and the block size (number of threads per block). This sizes are 3-component vectors, to help an easy identification of the *thread-ID* and *block-ID* at every kernel launch.

Since we have a two-phase force calculation process, we use a different configuration for each kernel call (block and grid size): one for the preliminary force calculation, which calculates the total force on an active particle, in 16 values (the JPBLOCK size), and another one to reduce those sets into the final total force on an active particle. We base our configuration on the idea presented in Nitadori and Aarseth [2012], with some variations of the data scheme. We use two-dimensional blocks to handle the indexes of the threads. In the case of the force calculation, we have:

```

1 dim3 blockSize(BSIZE, 1, 1); // (64, 1, 1)
2 dim3 gridSize(1 + Nact/BSIZE, JPBLOCKS, 1); // (5, 16, 1)

```

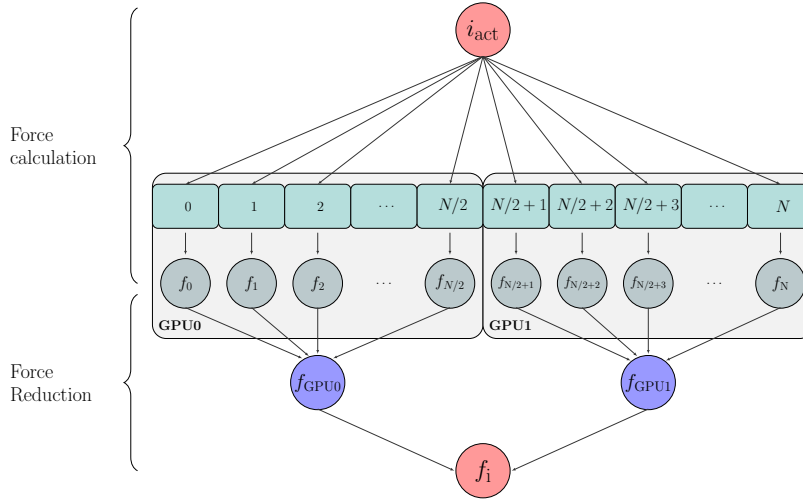



Figure 13: parallelisation scheme to split the j -loop instead of the i -loop. Two GPUs are depicted to represent how the code works with multiple devices. In this case, we have two sections, the first is to calculate the force interactions of the i -particle with the whole system but by different threads (upper part). Then a reduction, per device, is necessary to get the new value for the i -particle force (f_i).

Firstly, we define the block size to 64 (number of threads). Next, we configure the grid: the first dimensions of the blocks (`blockIdx.x = 5`) are thought to handle the $N_{\text{act}} = 300$ particles (5 blocks using 64 threads each = 300 total threads), and the second dimension of the blocks (`blockIdx.y`) JPBLOCKS, to handle the N particles in groups of $N/\text{JPBLOCK} = 64$. Lastly, we need 64 more threads to use the JPBLOCKS, which is the reason for having the 1+ at the beginning of the first dimension of the grid. After this process, the temporary forces are allocated in an $N \times \text{JPBLOCKS}$ array, so that for every particle we continuously save their force interaction with the other particles of the system. We then perform reductions (summations) to obtain the final forces for every N_{act} particle, which are distributed in JPBLOCKS (16) blocks. Hence, the kernel configuration must be different, and we use a simple one. In this case, we use a simple parallel CUDA reduction of N_{act} arrays of JPBLOCK (16) elements.

```

1 dim3 blockSize(JPBLOCKS, 1, 1); // (16, 1, 1)
2 dim3 gridSize(Nact, 1, 1); // (300, 1, 1)

```

8.2 THE THREE FLAVOURS OF GRAVIDY: TESTS

Thanks to the fact that there is a number of codes implementing similar approaches to ours, we are in the position of running exhaustive tests on GRAVIDY. Indeed, the global dynamics of a dense stellar system (typically an open cluster, because of the limitation in the number of particles we can integrate) has been addressed numerically by a large number of authors in the field of stellar dynamics. Therefore, we have decided to focus on the dynamical evolution of a globular cluster with a single stellar population. We present in this section a number of tests to measure the performance and the accuracy

of the three versions of GRAVIDY which we present using different amount of particles. Our goal is to be able to offer an OpenSource code that fits different needs and requirements. This is why this first release of GRAVIDY offers three different choices, which are general enough for different users with different hardware configurations. These are:

- (I) **THE CPU VERSION** consists in the more basic implementation in this work, a CPU version. I.e. This version uses OpenMP and is intended for a system without graphic processing units, but with many cores. This flavour can be used for debugging purposes by disabling the OpenMP directives (`#pragma omp`). This is the basis for our further development of the code.
- (II) **THE MPI VERSION** is virtually the same serial implementation, but with OpenMPI directives added to improve the performance of the hot-spots of the algorithm, in particular the force and energy calculation. In this case we use the MPI library, and hence it can be run on a single machine using a certain amount of cores as “slave” processes or on a large cluster with separated machines as slaves.
- (III) **THE GPU VERSION** discards all CPU usage and only relies on the GPU to integrate all gravitational interactions. As we mention later, we tried to use CPU combined with GPU, but we did not see any benefit in it, and the approach was hence neglected. We use CUDA to be able to interact with NVIDIA graphics processing units. The code is designed to detect the amount of present GPUs and use all of them, unless otherwise required by the user. This means that this version can use in a parallel way as many GPU cards as the host computer can harbour in a very simple and efficient way. The communication between the different GPU cards in the host computer is internal and run through Peripheral Component Interconnect Express (PCIe), a high-speed serial computer expansion bus standard, so that the data flows rapidly because of the low overhead.

The specifications of the hardware (CPU, GPU and available RAM) and operating systems we used are summarised in table 1.

8.2.1 *Initial conditions and N -body units*

For all of our tests we choose an equal-mass Plummer sphere [Plummer, 1911] for the sake of comparison with other codes. We choose standard N -body units (NBU, hereon) for the calculations and in the resulting output [Hénon, 1971, Heggie and Mathieu, 1986]. This means that

- The total mass of the system is 1: $\sum_{i=0}^N m_i = 1$.
- The gravitational constant (G) is set to 1: $G = 1$.
- The total energy of the system is equal to -0.25 : $E_{\text{tot}} = K + U = -0.25$, with K and U the total kinetic and potential energy of the system, respectively.
- The virial radius is set to ≈ 1 .

The Plummer spheres have a fixed half-mass radius of 0.8 and a Plummer radius of 0.6.

We used the code by Küpper et al. [2011] (McLuster) to generate all the initial conditions for the test we performed on the current work.

System A		datura (165 nodes)
CPU	Intel(R) Xeon(R) CPU X5650 @ 2.67GHz (24 cores)	
GPU		none
RAM		24 GB
OS		Scientific Linux 6.0
System B		gpu-01 (1 node)
CPU	Intel(R) Xeon(R) CPU E5504 @ 2.00GHz (4 cores)	
GPU	4 x Tesla M2050 @ 575 Mhz (448 cores)	
RAM		24 GB
OS		Scientific Linux 6.0
System C		krakatoa (1 node)
CPU	AMD Opteron 6386SE @ 2.8 GHz (32 cores)	
GPU	2 x Tesla K20c @ 706 MHz (2496 cores)	
RAM		256 GB
OS		Debian GNU/Linux 8
System D		sthelens (1 node)
CPU	Intel(R) Xeon(R) CPU E5-2697v2 @ 2.7GHz (24 cores)	
GPU	2 x Tesla C2050 / C2070 @ 1.15 Ghz (448 cores)	
RAM		256 GB
OS		Debian GNU/Linux 8

Table 1: Specification of the different systems of the Albert Einstein Institute used for the tests.

8.2.2 Accuracy, performance and speed

For GRAVIDY, as we have seen, we have chosen a Hermite 4th-order integrator. The numerical error introduced scales hence as $O(\Delta t^4)$ assuming a shared time step, which means that the previous is true *only* if all particles are updated at every integration step. Since we use a block time step scheme, certain groups of particles share a time step value, but not all of them. Thanks to this approach, the numerical error which we reach in our integrations is slightly less than the value mentioned previously.

Nitadori and Makino [2008] use a 6th- and 8th-order Hermite integrator scheme, which leads to higher-order (2nd and 3rd) acceleration derivatives and hence to better results in energy conservation, in particular the 8th-order case, but also in parallelisation efficiency. The reason for this is that by going to higher order in the truncation, we allow particles to populate ranges of time steps forbidden in the lower-order approximation. This leads to a larger number of block steps, and hence more parallelism efficiency. This efficiency comes from the amount of particles updated in every block step, because using this higher order integrator, we have bigger groups of particles in every block. One of our goals for the future version of our code is to try different integration schemes, like the Hermite 6th-order scheme described before. However, for the current purpose of GRAVIDY with softening, as presented in this work, we deem it sufficient to employ a 4th-order, which yields a good conservation of the energy, as we will see below. Although GRAVIDY aims at users with access to a handful of GPU cards on one node, we note that the work of Capuzzo-Dolcetta et al. [2013], which is based on the same integrator, shows a good performance in large GPU clusters, which means that theoretically one could run the current version of GRAVIDY in this kind of facility.

We have introduced in Eq. (24) a free parameter, η , responsible for determining the calculation of every time step of the system, from the initial

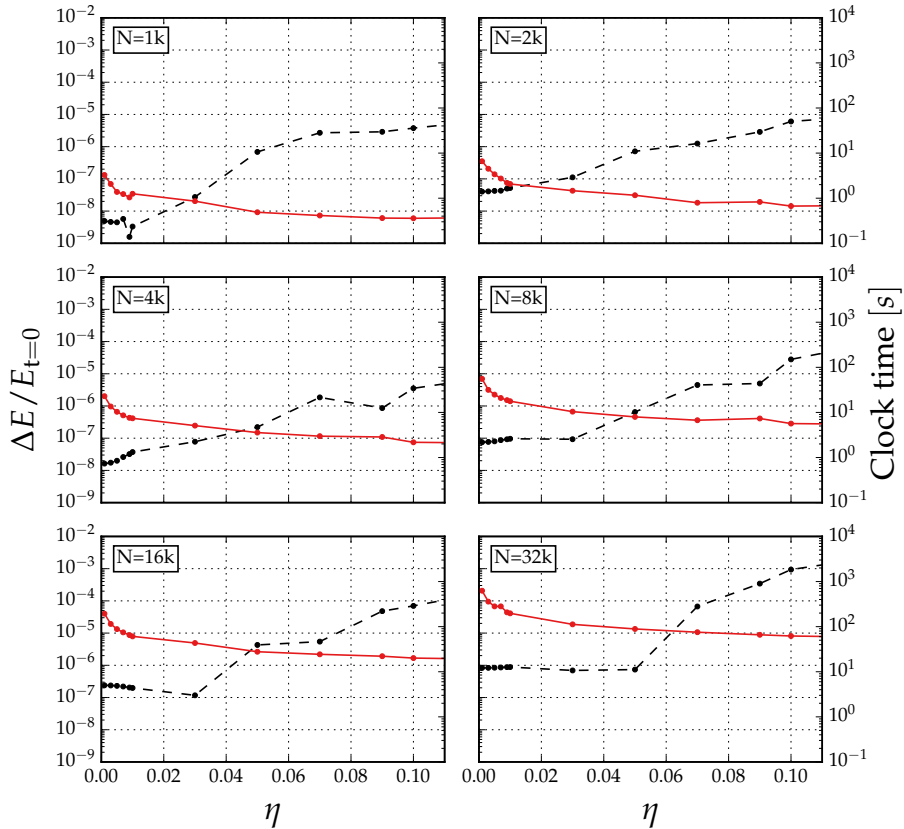


Figure 14: Cumulative energy error (dashed black line) and wall clock time (solid red line) in function of η for six different systems consisting of a Plummer sphere with $N = 1, 2, 4, 8, 16, 32k$ particles, with $k := 1000$, from the top to the bottom, left to right. The integration corresponds to one time unit, namely from $t = 1$ to $t = 2$ in the wall clock time analysis, and for $t = 2$ in the energy error calculation. The reason for choosing the elapse between 1 and 2 is to get rid of any initial numerical error at the simulation startup, from 0 to 1. All tests have been performed on System B of Tab. (1).

calculation to the update after every iteration. Hence, so as to assess an optimal value for it, we perform different tests to find a balance between a good energy conservation and a minimum wall clock time. We explore values between 0.001 and 0.1 integrating a series of systems with N ranging between 1024 to 32768, for convenience⁶, and up to 2 NBU. We show the results in Fig. 14 performed on System B of Tab. (1). For small values of η , the cumulative energy error approximately stays constant, because the error is small enough to leave accuracy in hands of the integrator scheme and the hardware. Increasing η leads to larger errors. This is particularly evident when we use systems with a larger number of particles. The system with $N = 32768$ particles, and a $\epsilon = 10^{-4}$, achieves $\Delta E / E_0 \approx 10^{-3}$ for $\eta = 0.1$, while it is as low as $\Delta E / E_0 \approx 10^{-6}$ for the same value and 1024 particles.

In the same figure we describe the performance in function of η by using the wall clock time in seconds for the code to reach one NBU for the same values of the parameter. We can see that the value of η is inversely proportional to the time, since increasing its value results in decreasing the execution time.

⁶ Any number of particles can be also handle properly, not necessarily powers of 2.

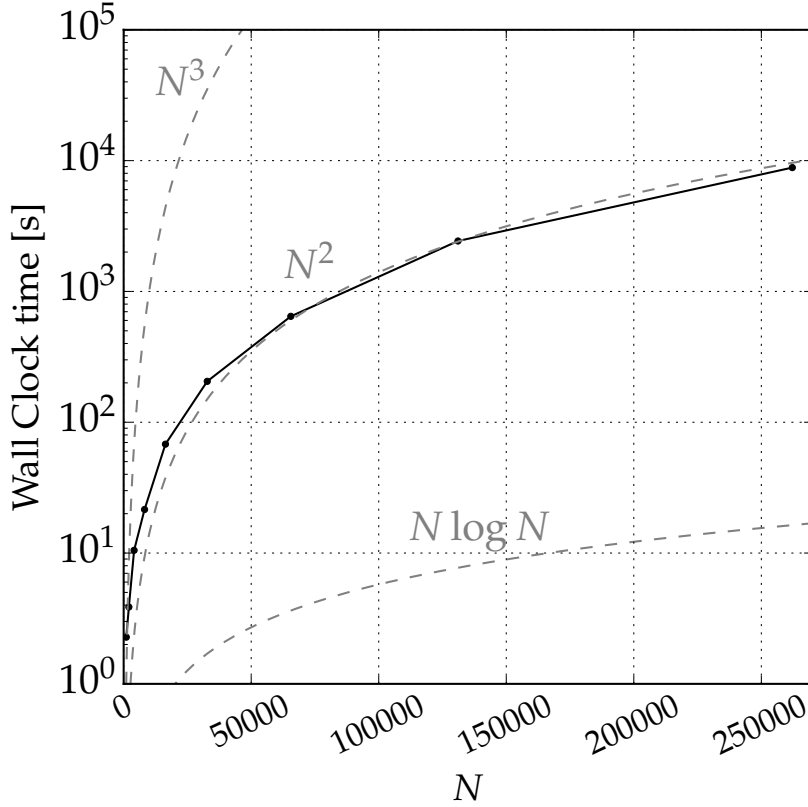


Figure 15: Wall clock time of integration from $t = 1$ NBU up to $t = 2$ NBU, using $\eta = 0.01$ and $\epsilon = 10^{-4}$ using different amount of particles on System C of Tab.(1).

When we increase η we implicitly increase the time step of every particle, so that one unit of time is reached sooner. We find that a value of about 0.01 is the best compromise for most of our purposes, yielding an accuracy of about $\Delta E/E_0 = 10^{-7}$ in most of the cases.

To measure the execution speed of our code we perform a set of tests by integrating the evolution for one NBU of a Plummer sphere with different particle numbers, ranging from $N = 1024$ to $N = 262144$. For the analysis, we choose the time starting at $t = 2$ and finishing at $t = 3$, since the first time unit is not representative because the system can have some spurious numerical behaviour resulting from the fact that it is not *slightly* relaxed. When testing the parameters η and ϵ , we picked the time starting at $t = 1$ and finishing at $t = 2$ because we wanted to understand their impact not right at the beginning of the simulation. Now we allow the system to further relax so as to obtain a more realistic system. In particular, the distribution time steps drifts away from the initial setup. For the simulations we choose $\eta = 0.01$ and $\epsilon = 10^{-4}$ which, as we discussed previously, are a good compromise in terms of accuracy and performance.

We display the wall clock time of each integration in Fig. 15. We also display reference curves for the powers of N^3 , N^2 and $N \log N$, multiplied by different factors to adapt them to the figure. We see that GRAVIDY scales very closely as a power of 2. The deviations arise from the fact that not all particles are being updated at every time step.

We employ different combinations of techniques during the parallelisation, both at GPU and CPU level. One of the most important points in the develop-

ment of the algorithm is the low amount of work required by the GPU when $N_{\text{act}} \ll N$, as we have seen. To address this we have explored other algorithms, such as using only the CPU when $N_{\text{act}} \propto N$, with α a free parameter to be determined by the experiment of interest, or when $N_{\text{act}} < N_{\beta}$, with N_{β} a (a priori) fixed amount of particles. In both cases, none turned out to be better than the j -parallelisation, even with very small amount of N_{act} , because of the prevailing GPU usage.

In Fig. 16 we show the acceleration factor for all parallel scenarios as compared to the single-thread CPU case, which we use as a reference point. Due to the design of the code, the maximum performance is achieved with the larger particle number. The most favourable scenario for GRAVIDY is, as we can see in the figure, System B. The 4 GPUs available boost the performance up to a factor of 193 as compared with the single-thread CPU case. A similar speed up is achieved on System D, which reaches a factor of 92 for the 2 GPUs. The CPU-parallel version lies behind this performance: only reaching a factor of 58 for System A, using up to 240 cores.

8.2.3 *Scaling of the three different flavours of the code*

An obvious question to any user of a numerical tool is that of scaling. In this subsection we present our results for the three different versions of GRAVIDY of how wall clock time scales as a function of threads or cores, or what is the acceleration of the multiple-GPU version of the code in function of the particle number as compared with a single GPU run, which we use as reference point.

In Fig. 17 we depict this information for the CPU, MPI and GPU versions. We can see in the CPU version that for small amounts of particles, in particular for 2k and 1k, we have an increase in the execution time with more threads, contrary to what we would expect. This is so because the amount of parallelism is not enough and the code spends more time splitting data and synchronising threads than performing the task itself, a usual situation in tasks with a low level of computation.

The MPI version uses the same j -parallelisation idea from the GPU one. In this case the code splits the whole system to the amount of available slaves (be it cores or nodes), performs the force calculation and finally sums up (“reduces”) the final forces for all active particles. This procedure was performed developing our own forces datatype operations and reduction, based on structures. This means that we define our own operations to be able to “sum” two forces (which are two three-dimensional arrays per particle). The simulations with small amount of particles (1k, 2k, 4k, 8k and 16k) are a clear example of a parallelisation “overkill”: using more resources than what is actually needed. Additionally, the communication process plays a role in scaling, which can be seen in the curves corresponding to these simulations for a number larger than 200 cores - the execution time increases instead of decreasing. On the other hand, large amount of particles (cases with 32k, 64k, 128k and 256k) show the expected behaviour, a better execution time with more nodes or cores. Surely this is not a solution for all simulations, since at some point the curves flatten.

The GPU version is a different scenario, since every device has its own capability, limitations and features that makes it difficult to compare their performances. For this reason we have decided to present the acceleration factor of every case normalised to a single-GPU run in the same system. This flavour of GRAVIDY should always have a better performance when increasing the particle number. Although having a good occupancy is in principle the ideal

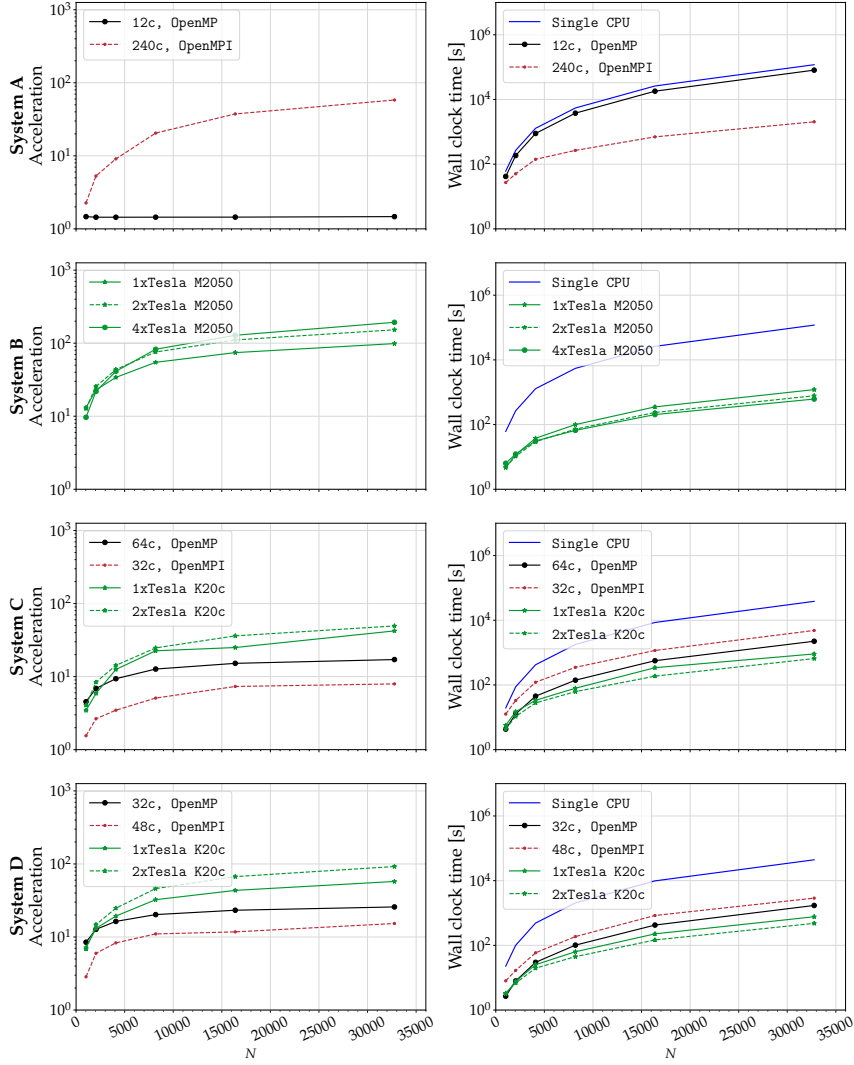


Figure 16: Acceleration factor and wall clock time for the different parallel versions of the integrator (see Tab. 1). The acceleration factor is normalised to the single CPU version (1 thread), up to $T = 3$ NBU. For the CPU parallel version of the code, we give information about the number of cores with the letter “c”. The GPU-parallel cases display the information on the number of cards with multiplying numbers.

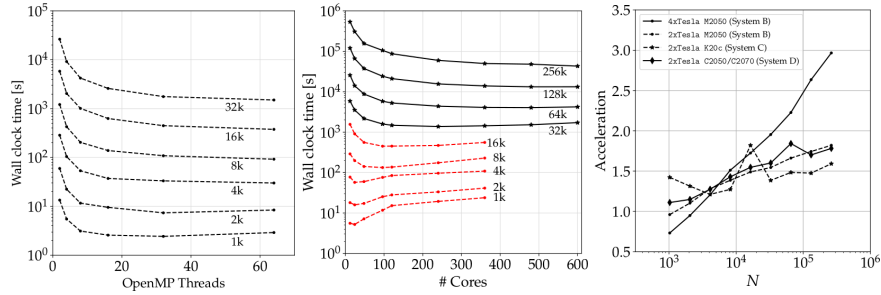


Figure 17: Performance of the three different flavours of GRAVIDY, as a function of the number of OpenMP threads, and number of cores and GPUs, for the CPU, MPI and GPU versions, respectively and from left to right. The integration corresponds to up to $t = 2$ NBU. *Left panel*: The CPU version runs on a single node with different numbers of threads. The experiments were performed on system C of Tab.(1). *Mid panel*: The MPI version running on different numbers of cores, using up to 600 of them and particles, up to 262144. In this case we use system A of the same table. *Right panel*: The GPU flavour using different amount of devices in parallel and particles. We show the acceleration factor as compared to a single GPU run for three different setups with different GPU specifications, corresponding to systems B, C and D of the same table.

scenario in this case, it is not necessarily the best reference point to assess the efficiency of the CUDA kernels, because it is related to register uses, but also to the amount of redundant calculations and the arithmetic intensity. We show the acceleration factor of two and four Tesla M2050 devices as compared to a single-GPU run which have hardware and performance differences⁷ but they nonetheless reach a similar acceleration factor. We have access to two Tesla K20c, which have more than the double peak performance in double precision floating point compared to the other mentioned models. The scaling between using one and two devices has a factor of 1.6.

Every GPU is a different device, so that in order to obtain a proper optimisation we need to first do a study in terms of kernel calls configuration. The current work present a fixed configuration of 32 threads per block, using a number of blocks corresponding to $N/32$. A deeper study on each GPU-device configuration is planned for future publication, where speeding up the first GPU implementation will be one of the main concerns.

8.3 THE ROLE OF SOFTENING ON DYNAMICS

For the current version of GRAVIDY, and quoting Sverre Aarseth on a comment he got some years ago during a talk, “we have denied ourselves the pleasure of regularisation”[Kustaanheimo and Stiefel, 1965, Aarseth and Zare, 1974, Aarseth, 1999, 2003]. This means that the code resorts to softening, via the parameter ϵ , introduced in Eq. 14. This quantity can be envisaged as a critical distance within which gravity is, for all matters, nonexistent. This obviously solves the problem of running into large numerical errors when the distance between two particles in the simulation become smaller and smaller, because

⁷ The primary difference is that model M is designed for Original Equipment Manufacturer (OEM) for an integrated system, without active cooling, while model C includes the active cooling and can be installed on any standard computer.

since they are 0-dimensional, this induces an error which grows larger and larger as they approach. This comes at a price, however. The relaxation time of the system is, approximately [see e.g. section on Two-body relaxation by Amaro-Seoane, 2012],

$$t_{\text{rlx}} \sim N_* \frac{t_{\text{dyn}}}{\ln(d_{\text{max}}/d_{\text{min}})}. \quad (26)$$

In this equation d_{min} and d_{max} are the minimum and maximum impact parameters. In an unsoftened N -body problem they are of the order of $d_{\text{min}} \approx Gm/\sigma^2$, and the size of the cluster, respectively. In other words, $d_{\text{min}} \approx R_{\text{cl}}/N$, with R_{cl} the radius of the self-gravitating cluster, if the system is virialised, and d_{max} is of the half-mass radius order. Now suppose the code uses a softening parameter ϵ . If the value of ϵ is smaller than d_{min} , then softening should play only a minor role in two-body relaxation, and the global dynamical evolution of the cluster must be similar to that of another cluster using regularisation. In the contrary case in which $\epsilon > d_{\text{min}}$, the relaxation time is artificially modified, as we can read from the last equation. The larger the quantity $\ln(d_{\text{max}}/d_{\text{min}})$, the more efficient is relaxation, and hence the shorter the relaxation time.

8.3.1 “Best” value for the softening?

We perform a series of simulations to assess the relevance of ϵ in the global dynamical evolution of an autogravitating stellar system. In Fig. 18 we depict the energy error and wall clock time for six different particle numbers as a function of the softening. The lower its value, the faster the simulation. However, by using larger values of the softening, we must understand that we are evolving a system in which two-body deflections are not being taking into account. This is the most important aspect of two-body relaxation, and therefore a critical factor in the general evolution. Thus, the fundamental feature which drives the global evolution of the system is non-existing below larger and larger distances. In particular, the larger values correspond to about 10% of the virial radius of the system. From these panels it seems that a value of $\epsilon \approx 10^{-4}$ is a good compromise *for this particular test that we are running in this example*. A good practice would be that the user tests different softening values for the case which is being addressed before making a decision for the softening. This choice is left for the user of the code, because we deem it difficult, if not impossible, to implement a self-regulating scheme in which the best value for the softening is calculated a priori.

8.3.2 Core collapse

Single-mass calculations

A good reference point to assess the global dynamical evolution of a dense stellar system is the core collapse of the system [see e.g. Spitzer, 1987, Aarseth et al., 1974, Giersz and Spurzem, 1994]. We present here the evolution of the so-called “Lagrange radii” (the radii of spheres containing a certain mass fraction of the system) in Fig. 19, for three representative values of the softening, the three upper panels, as calculated with GRAVIDY, and depict also the results of one calculation performed with NBODY6GPU [Nitadori and Aarseth, 2012], the lower panel, which uses KS regularisation [Kustaanheimo and Stiefel, 1965, Aarseth, 2003]. This can be envisaged as the “best answer”, which provides the reference point with which the other calculations should

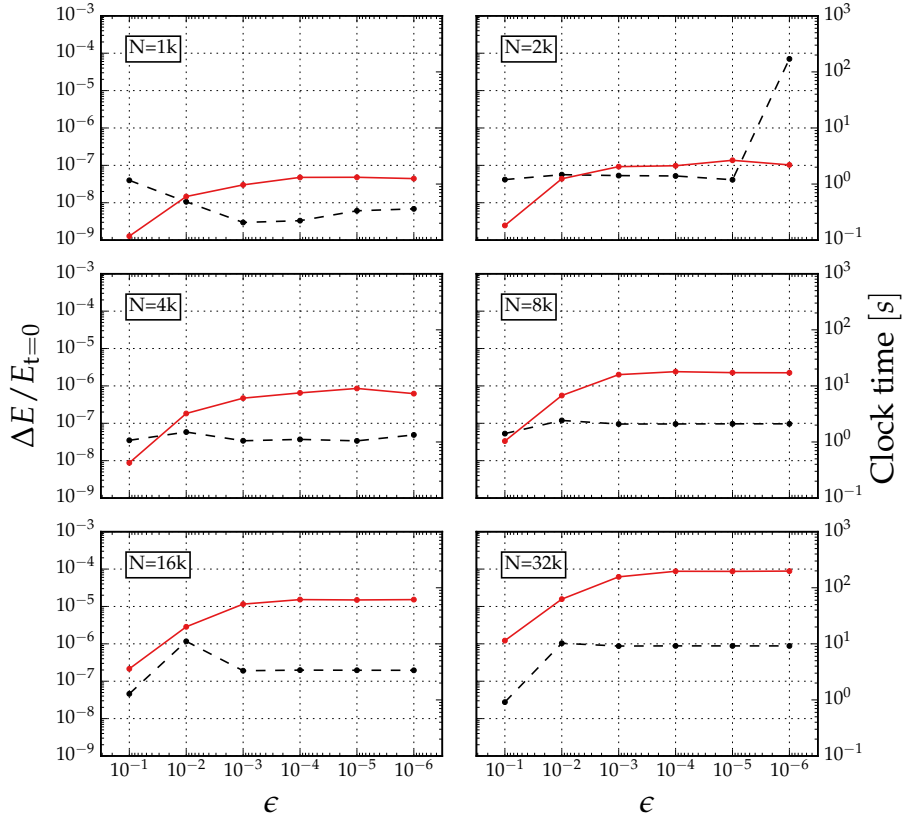


Figure 18: Cumulative energy error (dashed black line) and wall clock time (solid red line) using different values of the softening (ϵ). We integrate different amounts of particles N up to $t = 2$ NBU. The wall clock time corresponds to the execution time between $t = 1$ and $t = 2$ NBU while the energy error is the one at $t = 2$ NBU.

be compared. In the figures we use the half-mass relaxation time, which we introduce as

$$t_{\text{rh}} = 0.138 \left(\frac{Nr_h^3}{Gm} \right)^{\frac{1}{2}} \frac{1}{\ln(\Lambda)}, \quad (27)$$

where N is the number of particles of the system, m the average mass of a star, r_h the half-mass radius, and $\Lambda := \gamma N$, with $\gamma = 0.1$ the argument of the Coulomb logarithm.

From the panels we can easily see the impact of the softening parameter in the calculations: the collapse of the core is retarded for larger values. Our default choice for the softening, 10^{-4} is just $2 T_{\text{rh}}$ earlier than a NBODY6GPU calculation that we performed to compare with our code.

Another way of looking at the core collapse is in terms of energy. In Fig. 20 we display the evolution of the energy for the same systems of Fig. 19. As the collapse develops, the average distance between particles becomes smaller and smaller. There is an obvious correlation between the conservation of energy and the value of the softening. The transition between a fairly good energy conservation and a bad one happens more smoothly for larger and larger values of the softening, since the error has been distributed since the beginning of the integration. This means that, the smaller the value of the softening, the more abrupt the transition between the good and bad energy conservation, which leads to a big jump for the lowest value, 10^{-5} . We stop the simulations at this point because of the impossibility of GRAVIDY to form binaries, the main way to stop the core collapse.

As discussed previously, and as we can see in Fig. 20 and 19, the introduction of softening in the calculations has an impact on the global dynamical behaviour of the system. We find factors of 1.001, 1.08 and 1.55 of delay to reach the core collapse for the softening values $\epsilon = 10^{-5}$, $\epsilon = 10^{-4}$ and $\epsilon = 10^{-3}$, respectively.

The NBODY6GPU simulation was run on a different system, using a GeForce GTX 750 (Tesla M10) GPU, which is why we compared with the overall system evolution instead of the wall clock time.

Calculations with a spectrum of masses

Additionally to the single-mass calculations, we have also addressed multi-mass systems. The fact of having an Initial Mass Function (IMF) accelerates the core collapse of the system, as shown by many different authors [Inagaki and Wiyanto, 1984, Spitzer, 1987, Kim and Lee, 1997, Kim et al., 1998]. In our calculations, we use a Plummer sphere with a Kroupa IMF [Kroupa, 2001] and 8192 particles. In Fig. 21 we present the evolution of the Lagrange radii and the energy conservation of the system. We can see that the core collapse happens around $2 T_{\text{rh}}$, which is the point from which the energy conservation becomes worse and worse, to achieve a value of about 6 orders of magnitude worse than in phases before the collapse. Another way of depicting the collapse is by identifying the heaviest 10% of the stellar population and how it distributes in the core radius as calculated at $T = 0$. We can see this in Fig. 22.

The equilibrium of the system can be evaluated by analysing the distribution of the time steps. As we have mentioned previously, in Section (8.1.4), the initial distribution of time steps in the system has a log-normal distribution, which in a balanced system must remain similar, or close. In Fig. 23 we show the step distribution after the core collapse for the single-mass system with $\epsilon = 10^{-4}$

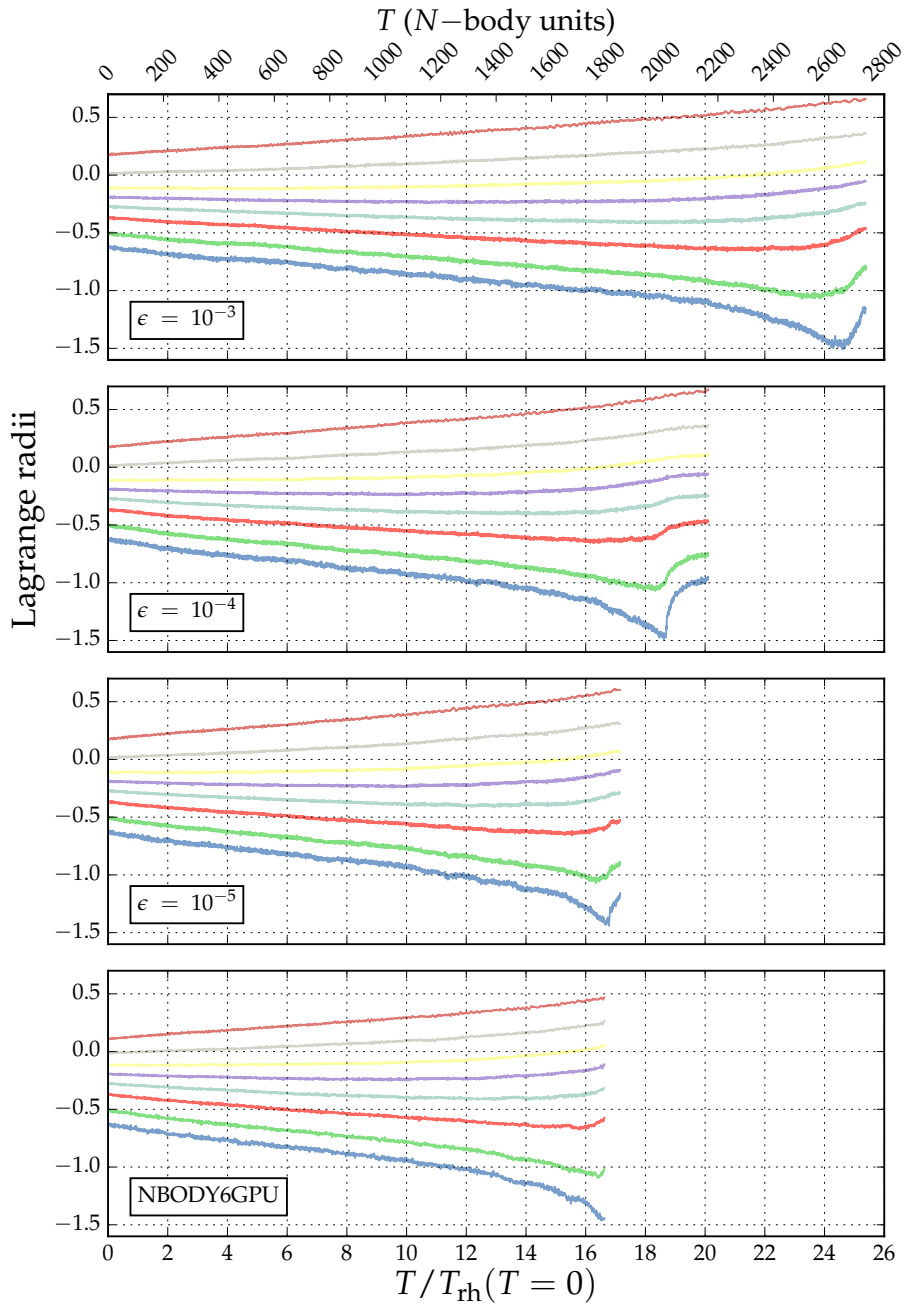


Figure 19: Comparison of the Lagrange radii of a Plummer Sphere with $N = 8192$ particles, using different values of ϵ (softening) for GRAVIDY and the NBODY6GPU code, from upper to bottom. The mass percentages are 0.5, 1, 2, 3, 4, 5, 6, 75 and 90% of the total mass, from the bottom to the upper part of each plot. The core collapse is reached at $\approx 24, 18$ and $16 T_{\text{rh}}$ for $\epsilon = 10^{-3}, 10^{-4}$ and 10^{-5} respectively. The half-mass relaxation time for this system is $T_{\text{rh}} = 112.186$ NBU. The NBODY6GPU code does not include a softening parameter, and treat binary evolution with a KS-regularisation.

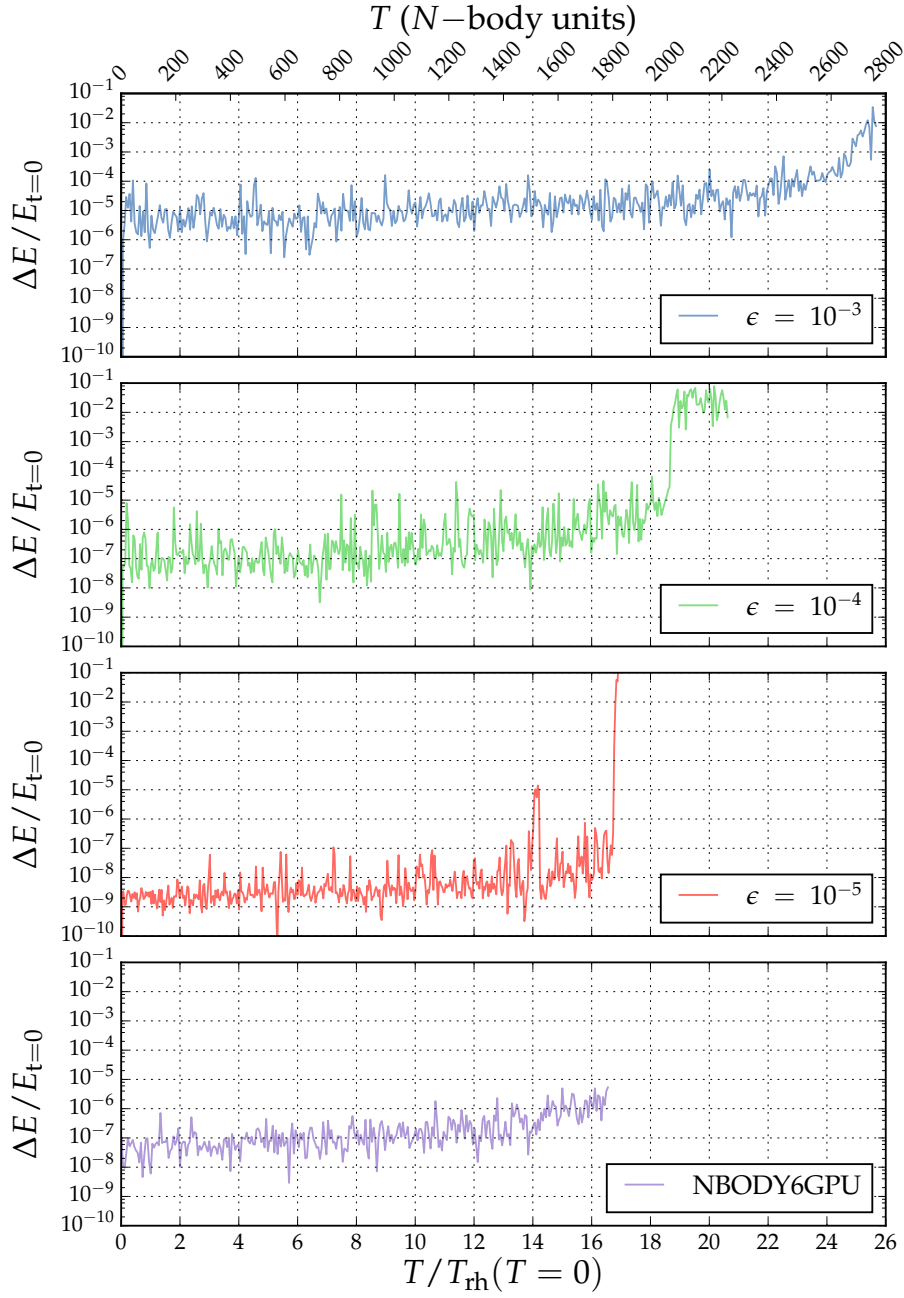


Figure 20: Energy conservation in a long time integration of a system with $N = 8102$ Comparison of the Energy conservation of a Plummer Sphere with $N = 8192$ particles, using different values of ϵ (softening) for GRAVIDY and the NBODY6GPU code, from upper to bottom. The core collapse is reached at $\approx 24, 18$ and $16 T_{\text{rh}}$ for $\epsilon = 10^{-3}, 10^{-4}$ and 10^{-5} respectively. The half-mass relaxation time for this system is $T_{\text{rh}} = 112.186$ NBU The NBODY6GPU code does not include a softening parameter, and treat binary evolution with a KS-regularisation. All the runs were stopped after the core collapse.

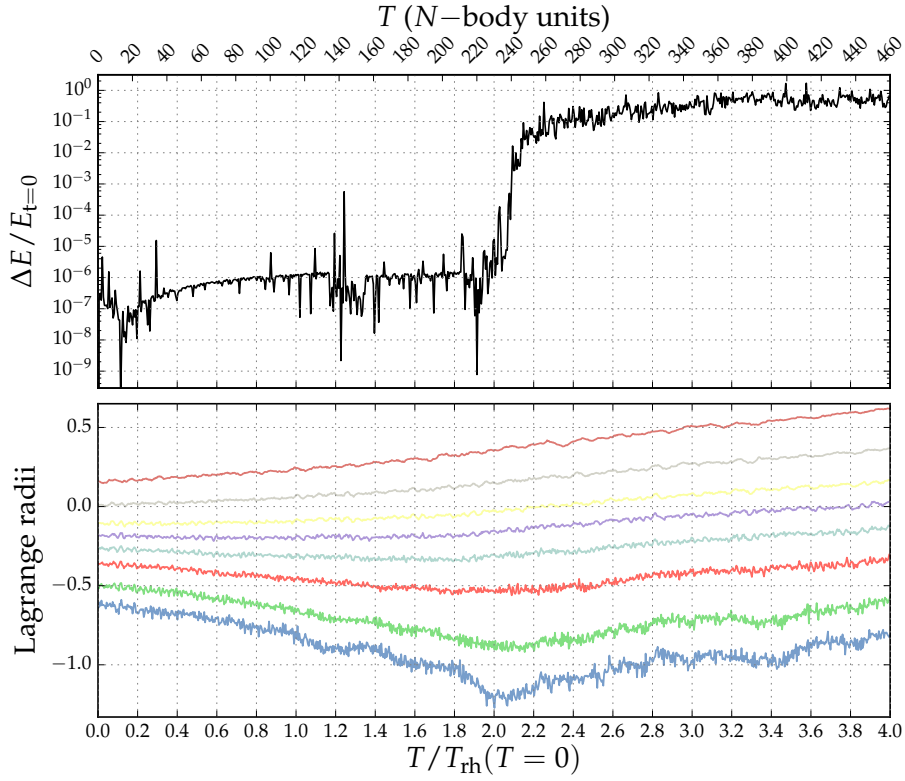


Figure 21: Plummer sphere using 8192 particles and following a Kroupa IMF. *Top Panel:* Cumulative energy of the system. *Bottom Panel:* Lagrange radii distribution for 0.5, 1, 2, 3, 4, 5, 6, 75 and 90% of the total mass.

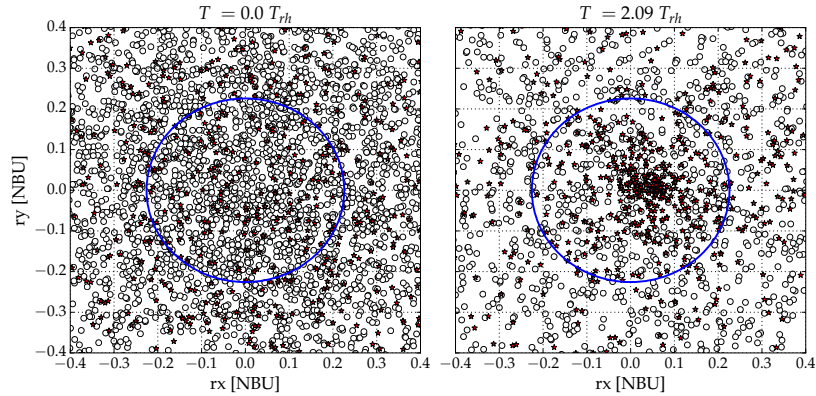


Figure 22: Inner section of a Plummer sphere with 8192 particles Plummer sphere following a Kroupa IMF before (left) and after the core collapse (right) at $T = 2.09 T_{rh}$. The blue circle depicts the core radius at $T = 0$. The top 10% of the heaviest particles in the system are marked as red stars, while all other particles as empty circles.

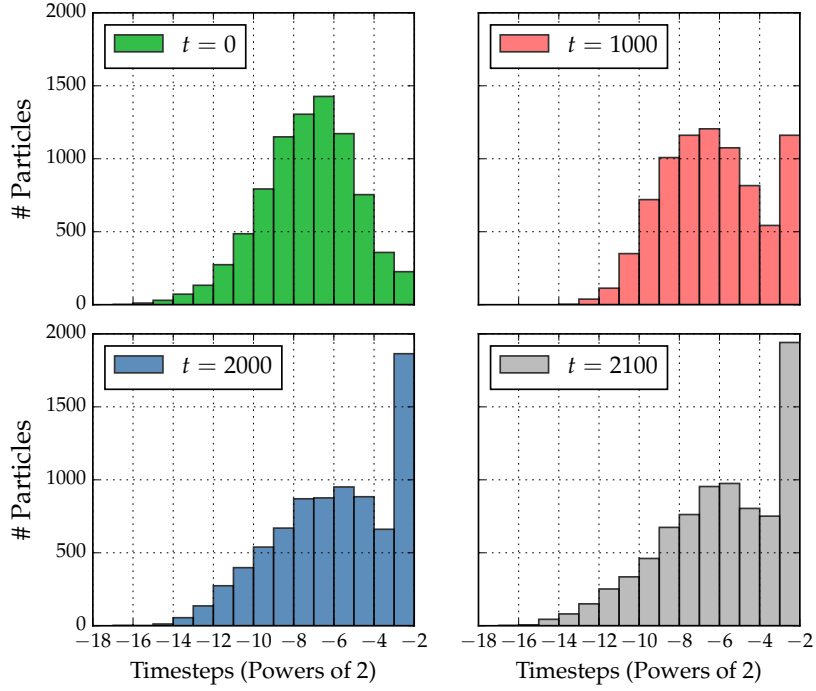


Figure 23: Time step distribution of a Plummer sphere with $N = 8192$ particles. Four different times are shown, (1) $t = 0$ NBU, for an initial distribution (upper left panel) (2) $t = 1000$ NBU, a few half-mass relaxation times $\sim 9 T_{\text{rh}}$ (upper right panel), (3) $t = 2000$ NBU, a pre core-collapse stage with many particles leaving the core (lower left panel), (4) $t = 2100$ NBU, a post core-collapse stage with a few particles (mostly binaries) reaching smaller time steps (lower right panel).

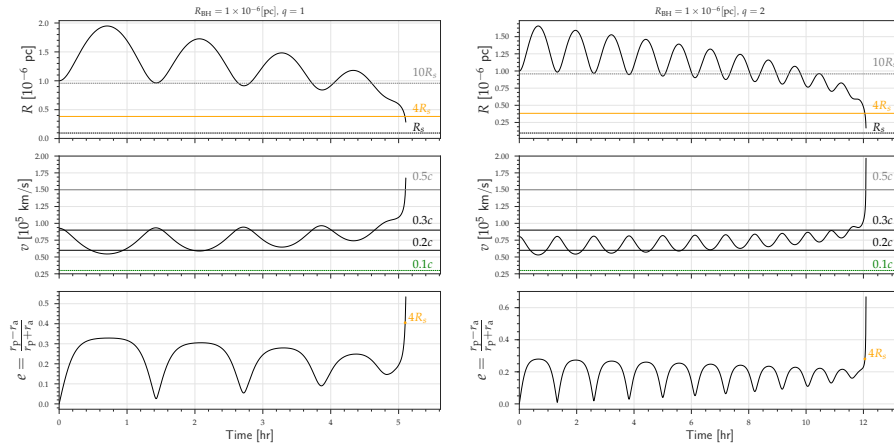


Figure 24: *Upper panel:* Evolution of the initial distance R_{BH} between the two supermassive black holes for a mass ratio $q = 1$ (and $q = 2$ in the set of three right panels). The three horizontal lines correspond, from the top to the bottom, to a distance of $10 R_S$, with R_S the Schwarzschild radius, $4 R_S$, and $1 R_S$. *Mid panel:* Evolution of the relative velocity between the two supermassive black holes. The four horizontal lines correspond, from the top to the bottom, to a fraction of the speed of light c of 50%, 30%, 20% and 10%. *Bottom panel:* Evolution of the eccentricity of the binary as a function of time in hours. We mark the point in the evolution at which the separation is $4 R_S$ with an orange dot.

8.4 RELATIVISTIC CORRECTIONS

GRAVIDY includes a treatment of relativistic orbits. This has been implemented for the code to be able to study sources of gravitational waves. The approach we have used is the post-Newtonian one, presented for the first time in an N -body code in the work of Kupi et al. [2006] (and see also Amaro-Seoane and Chen 2016) and later expanded to higher orders in Brem et al. [2013]. The idea is to modify the accelerations in the code to include relativistic corrections at 1PN, 2PN (periapsis shifts) and 2.5PN (energy loss in the form of gravitational wave emission). Contrary to the scheme of Kupi et al. [2006], which implements the modification in the regularised binaries, in the case of GRAVIDY, the corrections are active for a pair of two particles for which we set the softening to zero. The expressions for the accelerations, as well as their time derivatives can be found in the updated review of 2017 Amaro-Seoane [2012].

We run a series of different tests for binaries with different mass ratios and initial semi-major axis. In Fig. 24 we display the evolution of a binary of two supermassive black holes of total mass $1.33^6 M_\odot$ and mass ratios of 1 and 2. In Fig. 25 we show mass ratios of 5 and 100, and the latter starts with a smaller initial semi-major axis. For each of these cases we plot the geometric distance, the relative velocity and the eccentricity. Higher mass ratios lead to a more complex structure in the evolution. We can see how the relative velocity increases up to a significant fraction of the speed of light c as the separation grows smaller. We however note that the post-Newtonian approach should not be trusted for velocities larger than about 20% c .

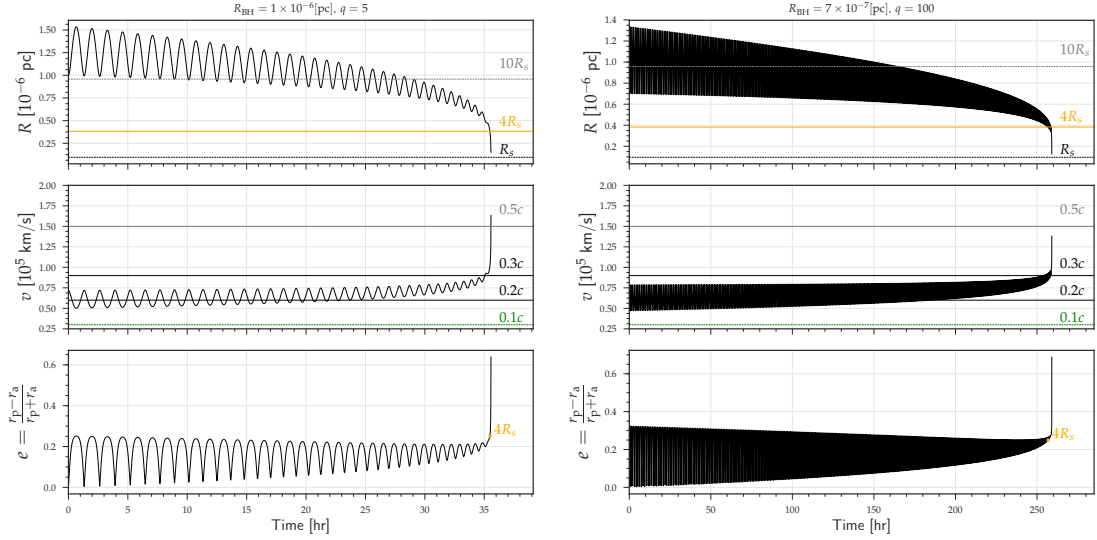


Figure 25: Same as Fig. 24 but for mass ratios $q = 5$ (left panels) and $q = 100$ (right panels), and a different initial separation, of 7×10^{-7} pc (also right panels).

8.5 CONCLUSIONS

In this chapter we have presented the first version of our new N -body code, written purely in C/C++, using OpenMPI and CUDA, which we call GRAVIDY. The current version of our code provides an environment to evolve a self-gravitating stellar system, and uses a Hermite 4th-order integration scheme, using block time steps and softening, and features relativistic corrections (periapsis shift and energy loss) for sources of gravitational radiation. This first release of GRAVIDY has been mainly focused on users who can have access to a machine hosting few GPUs, or usual parallel CPU systems.

We summarise here the main features of GRAVIDY:

1. The code is written using an iterative and incremental development, which is methodology similar to the Assess, Parallelise, Optimise, Deploy (APOD) development cycle presented by NVIDIA.
2. The code organisation is designed to be highly modular. Every critical process of the integrator is represented by a separate function or chain of functions. Our goal is to produce a code which can be read without difficulties, which makes easier future modifications or forks.
3. Since maintainability is one of our main goals, the documentation is also a critical factor. We document every function in the inner procedure of the integrator.
4. We use a Hermite 4th order integrator scheme.
5. The code uses block time steps to improve the performance of the integrator. We evolve particles in groups of block time steps, which allows for an update of several particles at the same time.
6. We use GPU computing techniques, OpenMP and OpenMPI to parallelise the calculation of the gravitational interactions of our system after having localised the hot-spots of our algorithm. The main objective here

Listing 8.2: Example run of the integrator. Columns, decimals, and information were modified to fit the output on this document.

```

$ ./gravidy-gpu -i ../input/04-nbody-p1024_m1.in -p -t 1
[2017-01-28 01:60:56] [INFO] GPUs: 1
[2017-01-28 01:60:56] [INFO] Sp1. 1024 particles in 1 GPUs
[2017-01-28 01:60:56] [INFO] GPU 0 particles: 1024
Time  Iter  Nsteps  Energy  RelE  CumE  ETime
0.000   0      0      -2.56e-01  0.00e+00  0.00e+00  3.08e-02
0.125  698    30093  -2.56e-01  2.41e-07  2.41e-07  3.84e-01
0.250  1262   61319  -2.56e-01  1.10e-07  1.30e-07  6.60e-01
0.375  1897   91571  -2.56e-01  4.19e-08  8.84e-08  9.49e-01
0.500  2530  121963  -2.56e-01  8.51e-08  3.30e-09  1.23e+00
0.625  3132  150924  -2.56e-01  2.89e-08  3.23e-08  1.52e+00
0.750  3725  180446  -2.56e-01  1.39e-08  1.83e-08  1.76e+00
0.875  4354  212425  -2.56e-01  5.23e-07  5.41e-07  2.02e+00
1.000  5160  244165  -2.56e-01  2.32e-07  3.09e-07  2.32e+00
[2017-01-28 01:60:59] [SUCCESS] Finishing...

```

was to be able to update a relatively small amount of particles which share a common time step in a given moment, a situation which is against the design of GPU cards, developed to reach a high parallelism.

8.6 APPENDIX A: ABOUT THE CODE

GRAVIDY is a C/C++ and CUDA application, that uses the CUDA, OpenMPI and boost libraries.

As an overview, the compilation can be done with: `make <flavour>`, for the `cpu`, `mpi` and `gpu` versions. A simple run of the code is displayed in the Listing 8.2.

The URL hosting the project is:

- <http://gravidy.xyz>

where you can find the prerequisites, how to get, compile and use the code more detailed. Additionally, documentation regarding the code, input and output files is included.

Inside the repository, there is a `scripts` directory with a set of classes to be able to handle all the output files of the code.

The code was compiled using `gcc (4.9.2)`, `openmpi(1.6.5)`, `CUDA(6.0)` and `boost (1.55)`.

The following compilation FLAGS were used `-O3 -Wall -fopenmp -pipe -fstack-protector -Wl,-z,relro -Wl,-z,now -Wformat-security -Wpointer-arith -Wformat-nonliteral -Wl,-O1 -Wl,-discard-all -Wl,-no-undefined -rdynamic`.

8.7 *n*-BODY VISUALISATION TOOL

A graphical representation of *N*-body simulations is always an attractive idea to display how the simulation was performed (Fig. 26). Due to this reason, we decided to write a small application to have a simple 3D visualisation of GRAVIDY snapshots, based in OpenGL.

GRAVIDYVIEW is a lightweight and simple OpenGL *N*-body visualisation tool, written in C/C++. It can be downloaded from:

- <https://gitlab.com/cmaureir/gravidy-view>.

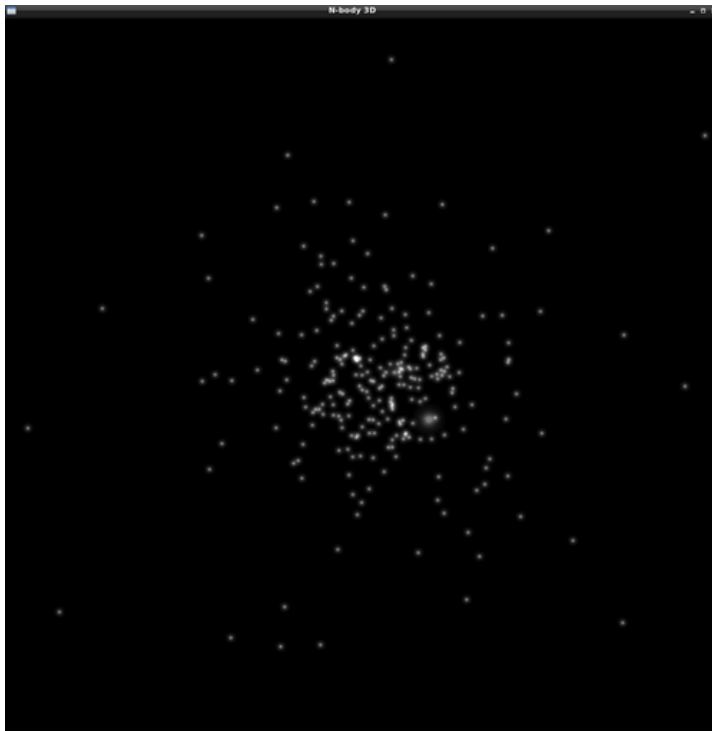


Figure 26: Snapshot pre-visualisation with GraviDyView using a $N = 1024$ system.

 RETROGRADE MBHB EVOLUTION ON CIRCUMBINARY
 ACCRETION DISCS

The project presented on this chapter was published under:

RETROGRADE BINARIES OF MASSIVE BLACK HOLES
 IN CIRCUM-BINARY ACCRETION DISCS

P. Amaro-Seoane, **C. Maureira-Fredes**, M. Dotti & M. Colpi
 Astronomy & Astrophysics, Volume 591, July 2016, id.A114, 12 pp.

In this chapter we present a suite of 2D hydro-dynamic simulations to study in detail the evolution of a MBHB embedded in a counter-rotating disc with focus on how different prescriptions on the accretion onto the secondary MBH can influence the evolution of the orbital elements. In §9.2 and §9.3 we describe the numerical tools, the initial configurations and the accretion prescriptions. In §9.4, we evaluate the effect of the different accretion recipes on the dynamics of the binary. Furthermore, we present results obtained from a set of 3D Smoothed-particle hydrodynamics simulations of MBHBs in a retrograde non-self gravitating disc to highlight commonalities and differences with the results from 2D simulations. In §9.5 we present a semi-analytical model in order to explore the long term evolution of the binary. We discuss the implications of our findings in §9.6.

9.1 INTRODUCTION AND MOTIVATION

We have seen in the introduction that massive black hole pairs are thought to be the natural outcome of galaxy mergers along the cosmic history [Begelman et al., 1980]. When two galaxies collide, the gravitational interaction of their galactic cores with the underlying dark matter, stellar, and gaseous background guides the sinking of the two massive black holes (MBH) at the centre of the galaxy remnant, which leads to the formation of a Keplerian binary. This occurs when the mass in gas and stars enclosed within the MBH orbit is smaller than the masses of the two MBHs, typically of \sim parsec for MBH masses of million suns [Colpi, 2014]. If the binary further hardens to attain separations as small as $\lesssim 0.001 \text{ pc}^1$, the emission of gravitational waves forces the two MBHs to coalesce in less than a Hubble time. Their final pairing and coalescence is measurable with a LISA-like observatory, e.g., LISA, if the total mass is $\lesssim 10^7 M_\odot$ [Amaro-Seoane et al., 2012b].

The evolution of the massive MBH binary (MBHB) on sub-pc scales depends on the properties of the cores of their hosts. In gas-poor remnants, MBHBs lose energy and angular momentum via scattering individual stars. The final fate of the binary depends on the effective reservoir of stars the MBHBs can

¹ The exact separation at which an MBH binary coalesces in less than a Hubble time depends on the binary eccentricity, mass, and mass ratio [see e.g., Peters, 1964, for an approximation based on Keplerian ellipses].

interact with. It has been shown that the presence of some degree of triaxiality [naturally present in galaxy merger remnants, Preto et al., 2011, Khan et al., 2011, 2012] can excite centrophilic orbits in a high number of the stars of the galaxy remnant, bringing the two MBHs to the scales of gravitational wave-driven inspiral and coalescence [see also Berczik et al., 2006, Berentzen et al., 2009, Gualandris and Merritt, 2012, Khan et al., 2013].

The presence of dense gas structures could accelerate the evolution of the binary, resulting in a faster coalescence [Begelman et al., 1980]. A fundamental difference between gas-poor and gas-rich environments is that, in the presence of a consistent amount of gas, accretion onto the MBHs is expected to be significant. This would enable us to detect dual AGN as spatially resolved nuclear sources and MBHBs at shorter separations from peculiar signatures in their optical/X-ray spectra [e.g., Begelman et al., 1980, Shen and Loeb, 2010, Tsalmantza et al., 2011, Eracleous et al., 2012, Montuori et al., 2011, 2012, Sesana et al., 2012, Tanaka et al., 2012, D’Orazio et al., 2013b, Bogdanović, 2015, D’Orazio et al., 2015], enabling us to map the orbital decay of MBHBs in the electromagnetic spectrum. In gas-poor environments, the most robust way of identifying MBHBs is through the detection of gravitational radiation during the inspiral phase. For large MBH $\gtrsim 10^8 M_\odot$, the Pulsar Timing Array experiment, operating at nano-Hz frequencies, might reveal their signal [Hobbs et al., 2010]. Furthermore, MBHBs can be detected during the inspiral, merger and ring-down in experiments such as LISA, at shorter wavelengths (around 0.1 mHz-1 Hz), and for lighter MBHB coalescence ($\sim 10^7 M_\odot$) [Amaro-Seoane et al., 2012b].

The presence of massive gas structures close to MBHBs is not unexpected. If the two merging galaxies initially had a significant amount of gas, their reciprocal perturbation drives gas inflows toward the centre of the two structures. These inflows result in massive gas discs in the centre of the galaxy remnant [e.g., Escala et al., 2005, Mayer et al., 2007, Dotti et al., 2007, Hopkins and Quataert, 2010, Dotti et al., 2012]. However, the details of the interaction between the binary and a gas disc are still debated [Fiacconi et al., 2013, Roškar et al., 2015, Lupi et al., 2015b, del Valle et al., 2015]. The gas disc can either be corotating or counter-rotating with respect to the MBHB [Nixon et al., 2011b, Roedig and Sesana, 2014]. The corotating case seems to be the more natural outcome of a gas rich galaxy merger [see, e.g., Mayer et al., 2007], since the MBHs bind in a binary during the natal process that forms the nuclear gas disc. For this reason, MBHBs embedded in co-rotating circumbinary discs have been extensively studied [Goldreich and Tremaine, 1980, Lin and Papaloizou, 1986, Artymowicz and Lubow, 1994, Ivanov et al., 1999, Escala et al., 2005, Hayasaki et al., 2007, Dotti et al., 2007, 2009, MacFadyen and Milosavljević, 2008, Cuadra et al., 2009, Lodato et al., 2009, Farris et al., 2011, Roedig et al., 2011, 2012, Noble et al., 2012, D’Orazio et al., 2013b]. However, the degree of misalignment between the gas in the remnant nucleus and the binary could depend on the parameters of the merger [Blecha et al., 2011, Hopkins et al., 2012] so that counter-rotating accretion is not ruled out. Furthermore, if the binary does not coalesce on a short timescale (comparable with the timescale over which star formation depletes the central corotating gas), subsequent inflows of gas could be uncorrelated to the angular momentum of the binary, possibly resulting in counter-rotating circumbinary discs [Goicovic et al., 2016].

Regarding the evolution of an MBH binary, retrograde and prograde discs differ in a few important aspects: (i) in the prograde scenario, the disc-binary interaction leads to the opening of a gap, i.e., a hollow region surrounding the

MBHs of a size comparable to twice the binary separation, and binary hardening is somewhat reminiscent of a planet type II migration. In the retrograde scenario resonances are either absent (for circular binaries) or weak [Nixon and Lubow, 2015]. In many cases the binary does not manage to excavate a gap [Bankert et al., 2015], and the MBHB-disc interaction actually enhances the inflow of matter toward the centre [Nixon et al., 2011a]; (ii) retrograde gas interacting with the MBHs can remove more angular momentum per unit of mass than in the prograde case, since its initial angular momentum has sign opposite to that of the binary and this leads to an increase in the eccentricity [Nixon et al., 2011a, Roedig and Sesana, 2014, Schnittman and Krolik, 2015]; (iii) in the retrograde case, the relative velocities between the MBHs (in particular between the secondary MBH, in an unequal mass binary) and the disc are significantly larger, so that the interaction between the gas and the MBHs is confined to smaller regions.

9.2 NUMERICAL TOOL AND DESCRIPTION OF THE INITIAL MODELS

We consider the case of MBH binaries orbiting in the orbital plane of an accretion disc. This enable us to limit the dimensions in our simulations to two. Thus, we can use `FARGO`², a two-dimensional hydrodynamical grid program that integrates the isothermal Navier-Stokes equations using a staggered polar grid [Masset, 2000].

The code is particularly suited for quasi-Keplerian scenarios, since it separates the azimuthal averaged motions from azimuthal and radial perturbations, resulting in longer time steps [the `FARGO` algorithm was originally presented in Masset, 2000]. This algorithm speeds up significantly the calculations and hence enable us to study the parameter space more efficiently than with other schemes. `FARGO` is parallelised by splitting radially the grid in rings, which are calculated in different CPUs.

We run a set of 14 simulations of unequal mass MBHBs in gaseous discs. In every run the primary MBH (M_1) is treated as an external potential, and is not evolved during the simulations; it is considered a point-like source. In internal units, the mass of M_1 is 1, and is placed at rest at the centre of the disc. The disc is confined within an outer radius defined by the radial limits of the grid. In internal units, the outer radius is 25, and the inner radius is 0.1. The disc follows a Mestel profile, as we depict in Fig. 27, with a surface density $\Sigma_{gas}(\hat{R}) \propto \hat{R}^{-1}$, and a total mass of 0.078, i.e., $\approx 1/12$ of the mass of the primary.

The disc follows an isothermal equation of state, with a thermal profile resulting in an aspect ratio $H/\hat{R} = 0.04$ constant throughout the disc. The dynamics of the disc is initially quasi-Keplerian, since, globally, the potential is dominated by M_1 . The initial angular speed is however not strictly Keplerian, since the code accounts for the pressure support to the rotational equilibrium. The computational domain is divided in a grid of 128 radial and 384 azimuthal sectors.

In every simulation a second MBH, M_2 , of initial mass $M_{2,0} = 0.1 M_1$ is placed in the disc, at a distance $d = 10$ from M_1 . The secondary is initially moving on a bound orbit, and counter-rotates with respect to the disc. M_2 can either be on a circular or on an eccentric orbit. In this last case we choose the initial eccentricity to be $e_0 = 0.6$, corresponding to an initial apocentre $\hat{R}_{apo} = 16$. At the beginning of each simulation, the secondary MBH is implanted in the disc as a particle with mass increasing from 0 to $M_{2,0}$ over one orbital

² <http://fargo.in2p3.fr>

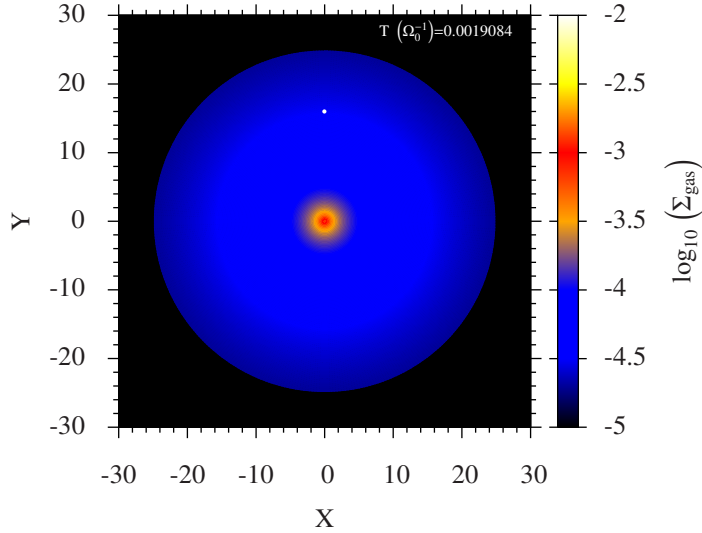


Figure 27: Face-on, colour coded map of the disc’s surface density at the onset of each simulation. The position of the secondary is marked with a solid green circle. For all quantities, we use the internal units of the code, except for the time, which is in units of Ω_0^{-1} , i.e., the inverse of the initial binary’s rotational frequency.

timescale. This choice prevents the growth of unphysical disc perturbations close to the secondary.

We pay particular attention on how the gas is added and removed from the disc. The amount of gas present can change only if it crosses the radial edges of the computational domain, or if it is accreted by the secondary. At the outer and inner edge of the computational domain we set out-flowing boundary conditions, i.e., matter crossing the boundaries disappears from the computational domain, but no gas can inflow into the computational domain.

So as to check the stability of the initial conditions, we first corroborate that the disc is stable by setting the mass of the secondary to a very small value (a 100th of the primary). We confirm the stability for what indeed is the range of mass ratios characteristic to the code -written to study the migration of a planet of mass much smaller than that of the central star- for some tens of initial periods of the binary.

Every set of initial conditions has been run five times, using different prescriptions for the accretion onto M_2 . The secondary is modelled either as a sink particle that can accrete gas from the disc, or as a point mass whose mass is fixed in time. Since the accretion of mass has strong consequences on the dynamical evolution of M_2 we use different prescriptions for the mass accretion, as it is discussed in the next section.

9.3 ACCRETION PRESCRIPTIONS

Accretion onto the secondary MBH is a key process affecting the dynamical evolution of the MBHB as (i) the accreting gas changes the mass and velocity of M_2 , according to conservation of the total momentum of the system, and (ii) the process of accretion itself decreases the gas density close to the secondary. In addition, the perturbation induced by the motion of M_2 further changes the underlying disc density pattern, back reacting on the dynamics of the MBHs in the binary. It is therefore important to implement an accre-

tion prescription that does not bias the evolution of the binary. The safest approach would be to follow the hydrodynamics of the gas down to the innermost stable circular orbit around each MBH, or at least around the secondary MBH. Our simulations however do not model all the physics needed to evolve gas down to such extremely small radii. Further out from this physical limit, the numerical nature of our investigation prevents us to set a too small value of the sink radius, that in a number of cases is a free parameter of the simulation. When gas reaches the sink radius, the gas itself is removed from the computational domain and its mass and momentum is added to the MBH. As a consequence, we decided to use different prescriptions for the gas accretion onto the secondary, to test if any accretion prescription results in an artificial orbital evolution of the secondary.

The first prescription we use is the standard FARGO implementation that we will refer as RL model, hereon: gas is accreted onto the secondary if its distance from M_2 , denoted as R_2^{gas} , is less than a given fraction (0.75) of the Roche Lobe (RL) radius R_{RL} around M_2 . The mass accretion is modelled by reducing the gas density within the RL by a factor $(1 - f_{red})$, at every timestep Δt . To prevent spurious high density and pressure jumps, $f_{red} = 1/3$ if $0.45R_{RL} < R_2^{gas} < 0.75R_{RL}$, while $f_{red} = 2/3$ if $R_2 < 0.45R_{RL}$ [see for more details Kley, 1999].

However, we note that for a binary counter-rotating with respect to the accretion disc in which it is embedded, the Roche Lobe of the secondary MBH is far from being comparable to its gravitational sphere of influence radius R_{bound} . The counter-rotating gas crossing the Roche Lobe in the region $R_{bound} < R_2^{gas} < R_{RL}$ is moving too fast with respect to the secondary MBH to bind to it. For $M_2 \ll M_1$, the Roche Lobe radius is

$$R_{RL} \sim \frac{1}{2}d \left(\frac{M_2}{M_1} \right)^{1/3}, \quad (28)$$

where d is the separation between the two MBHs, while

$$R_{bound} \sim \frac{G M_2}{V_{rel}^2} \sim \frac{1}{4}d \frac{M_2}{M_1}, \quad (29)$$

where G is the gravitational constant, and V_{rel} is the modulus of the relative velocity between the gas and the secondary. The ratio between the two radii is then

$$\frac{R_{bound}}{R_{RL}} \sim \frac{1}{2} \left(\frac{M_2}{M_1} \right)^{2/3}. \quad (30)$$

As a consequence, the Roche Lobe based standard implementation of accretion could result in an overestimated accretion rate, hence in an unphysical dynamical evolution of the secondary.

A second prescription we implement is based on the choice of a fixed sink radius R_{fix} . To prevent spurious pressure jumps we use the same two zones implementation discussed above, accreting 1/3 of the material present in the $0.5 R_{fix} < R_2^{gas} < R_{fix}$ shell and 2/3 of the material with $R_2^{gas} < 0.5 R_{fix}$ at each timestep. The size of the fixed sink radius around M_2 could affect the dynamics of M_2 in an unphysical manner if $R_{fix} > R_{bound}$. To check for this spurious effect we run the same set-up with different values of $R_{fix} = 0.5, 0.25$, and 0.05 , in code units. For a circular binary at the onset of the simulation R_{fix}/R_{bound} is approximately 2, 1, and 0.2 for $R_{fix} = 0.5, 0.25$, and 0.05 , respectively; similarly $R_{fix} = 0.5$ corresponds to $R_{fix}/R_{RL} \approx 0.2$.

A third prescription we test requires gas to be gravitationally bound to M_2 in order to get accreted. This automatically solves the problem of unbound

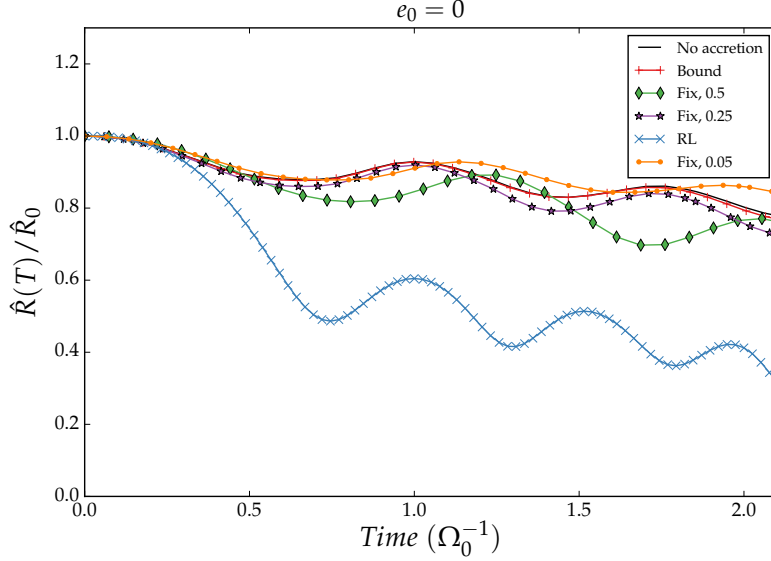


Figure 28: Relative MBH orbital separation (in units of the initial separation) as a function of time in units of Ω_0^{-1} for a binary with initial eccentricity $e_0 = 0$. Only the first ~ 2 orbits are shown to highlight the differences in the orbital evolution caused by the different accretion prescriptions. Black solid line refers to the run without any implementation of accretion onto the secondary (the no-accretion model). The red plus line refers to the Bound model and the blue crosses line to the RL model (lowest line). Green diamonds, purple stars and orange circles refer to the Fix models with sink radii $R_{fix} = 0.5, 0.25,$ and 0.05 , respectively. Colour version is available in the on-line version.

gas spuriously binding (and accreting) to the secondary. In this case, the gas density reduction factor is either $f_{red} = 1/3$ if the total gas energy per unit mass (with respect to the secondary) is $(3/4)W < E < (1/2)W$, or $f_{red} = 2/3$ if $E < (3/4)W$, where $W = -GM_2/R_2$.

Finally, we test the dynamical evolution of an accreting secondary against a non-accreting one. In this case, the secondary MBH is allowed to bind gas according to the above prescription, but the bound gas is not removed from the simulation, and can either remain bound to M_2 , co-moving with it on retrograde orbits, or can be stripped by either the tidal field of the primary or by the ram pressure of the gas disc. In the following, for the standard FARGO implementation we will refer to as the “RL model”, as “Fix” 0.5, 0.25, and 0.05 for the models with fixed sink radii, and as “Bound” and “No-accretion” for the remaining last two.

9.4 RESULTS

9.4.1 Effect of the accretion prescription on the dynamics

Figure 28 shows the evolution of the binary separation as a function of time, for different accretion prescriptions over a short timescale (i.e., an interval of 2 orbital times). Initially the MBHs move on a circular orbit.

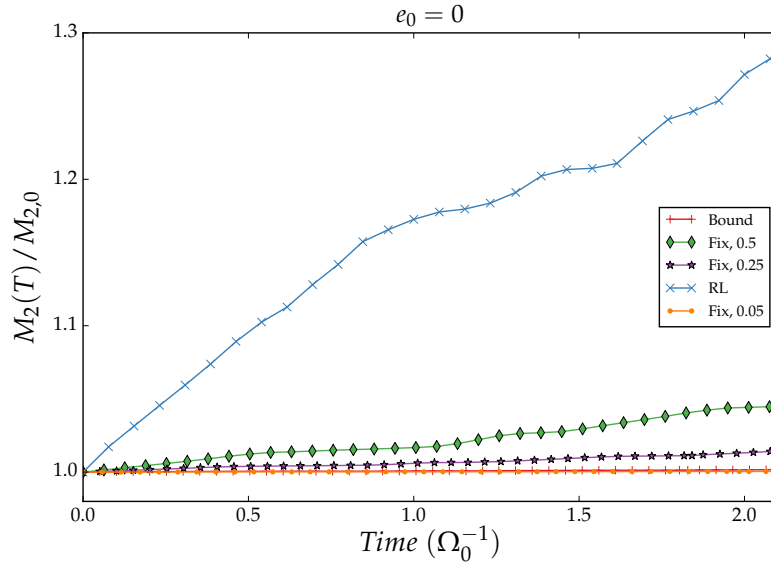


Figure 29: Mass of the secondary MBH, in units of the initial mass $M_{2,0}$, as a function of time in units of Ω_0^{-1} . Line colour and style codes are as in Figure 28.

The timescale of MBH evolution strongly depends on the accretion prescription assumed. Over a timescale of a few orbital times, the binary hardening in the RL model is faster than in the Bound model, as illustrated in Figure 28. The reason for this is simple. Gas at a distance from M_2 in between the influence radius of the secondary and its Roche Lobe ($R_{bound} < R_2^{gas} < R_{RL}$) is, by definition, not bound. When the sink radius is forced to be equal to R_{RL} , the full momentum of the gas (i.e., also the momentum of unbound gas) is added to the MBH, resulting in an artificially fast dynamical evolution of the binary, because of the unrealistically high accretion rate.

We can test the above interpretation exploring additional cases varying R_{fix} . Whenever a fixed sink radius is assumed, the dynamics depends on the size of the sink radius itself. As expected, larger sink radii correspond to faster migration. In Figure 28 we show how the evolution of the MBH separation, which is sensitive to R_{fix} , converges with continuity to the Bound model when decreasing the values of the sink radius. We further find that the No-accretion model in which the MBH does not accrete results in a binary decay very similar to the Bound model. The reason is straightforward: in both cases gas is allowed to bind to the secondary MBH. In the No-accretion run, the gas transfers its (negative) linear momentum to the secondary when it starts co-moving with it, i.e., during the binding process. The same happens in the Bound model in which the gas either binds to the MBH (transferring its linear momentum) or gets immediately accreted during the binding process. In this last case the linear momentum conservation is forced by the accretion prescription. Furthermore, material that would not bind to the secondary is not accreted by default. We further notice that in the Fix models, when the sink radius of M_2 is larger than its gravitational influence radius, the MBH absorbs the linear momentum of gas that, in reality, would not be fated to interact with it. Under these conditions the MBHB separation again has a fast artificial decay.

The amount of mass that is accreted by the secondary MBH can be used as a tracer of the linear (and angular) momentum that the disc transfers to

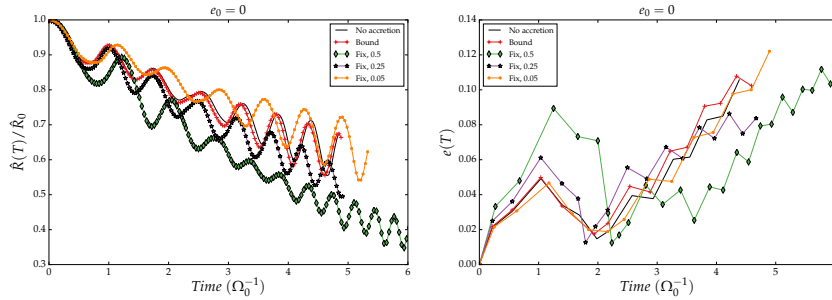


Figure 30: MBHB orbital separation (left) and eccentricity (right) as a function of time, for $e_0 = 0$. We follow the same notation as in Figure 28. The RL model is not considered, here.

the MBH, and in turn to the MBHB. Figure 29 shows the evolution of the secondary mass as a function of time, for the different accretion prescriptions. The MBH in the RL run displays the fastest mass growth and higher amount of accreted mass. In particular in the RL model, the MBH accretes 30% of its mass in order to reduce the separation by 25%. By contrast a reduction of 25% in the relative MBH distance is attained with a fractional mass increase $\sim 0.3\%$ in the Bound model. This indicates that in the process of binary hardening, the fractional decay per unit accreted mass is less effective in the RL model. In this case, the significant mass accreted by the MBH creates an empty region resulting in a much weaker frictional drag on the MBH from the torque of surrounding gas particles that contribute to the deceleration without being accreted. An interpretation of the results will also be discussed in §9.5.

9.4.2 Evolution of circular binaries

The evolution of a circular binary is studied further in this section. Fig. 30 illustrates the run of the BH separation versus time over $\sim 6 \Omega_0^{-1}$, for the different accretion prescriptions. The RL case is not included in the figure because of the too large, unrealistic accretion prescription. On such a long timespan, all the other runs but for the Fix 0.5 show similar evolution. The discrepancy between Fix 0.5 and the other runs can be ascribed to the large sink radius of Fix 0.5 resulting in an overestimate of the accretion rate and of the orbital brake. As shown in Fig. 30 the binary eccentricity, increasing slightly at start, fluctuates around a mean of 0.08 and never exceeds 0.16 during the whole evolution.

Figure 31 shows the MBHB separation as a function of M_2 , the mass the secondary MBH, to illustrate again that the RL model represents the less efficient mechanism of binary hardening in terms of normalised accreted mass. In our case, the high efficiency is caused by the drastically reduced amount of matter accreted onto M_2 (see Figure 29). The not-accreted gas still exerts a non negligible gravitational torque onto the secondary, breaking its orbit and driving its pairing. Such a torque is incorrectly computed (and overestimated) in the simulations with unphysically high accretion rates, where the gas responsible for most of the torque is removed from the simulations and its mass artificially added to the secondary.

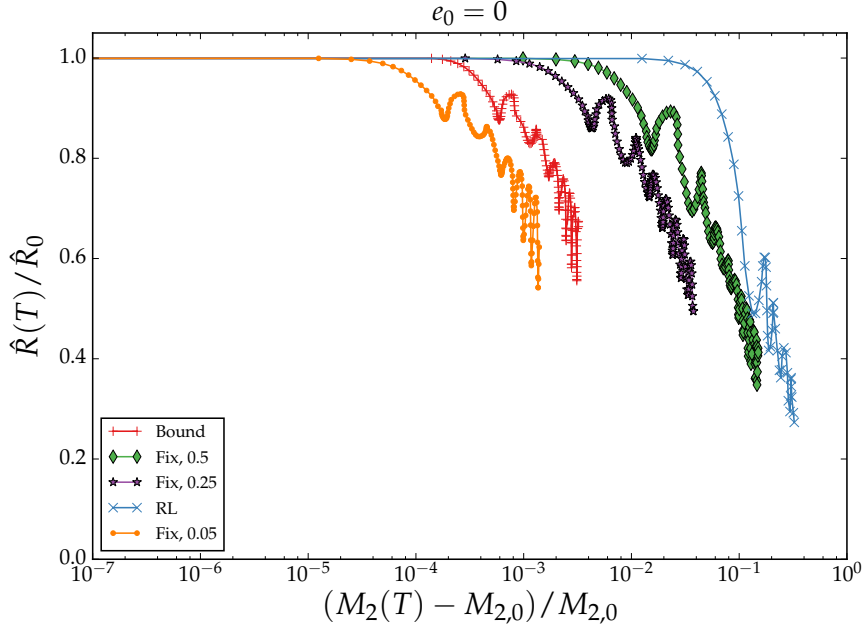


Figure 31: Binary separation as a function of the secondary mass, for different accretion prescriptions. The binary separation and M_2 are normalised to their initial values. The evolution is followed for a time equal to $6\Omega_0^{-1}$. Line colour and style codes are the same as in Figure 28.

9.4.3 Evolution of eccentric binaries

In this section we study the hardening of an initially eccentric binary, in the retrograde disc. Because of the results about the accuracy of the dynamical evolution of the binary presented for the circular case, we confine the analysis of the eccentric cases to the Bound, No-accretion and $R_{fix} = 0.05$ runs. The evolution of MBHB separation (upper panel) and eccentricity (lower panel) for a binary with initial eccentricity $e_0 = 0.6$ is shown in Fig. 32.

The initial ($t \lesssim 2\Omega_0^{-1}$) evolution of the MBHB is quite similar in the three cases. The binary hardens and contemporary the eccentricity grows considerably, up to $e \gtrsim 0.7$, in qualitative agreement with the analytical predictions of Nixon et al. [2011a]. The slight differences between the different cases are due to the different amount of mass accreted. As shown in Fig. 33 the secondary accretes more gas when assuming the bound prescription with respect to, e.g., the $R_{fix} = 0.05$ run.

Such differences depend on the accretion prescriptions used. The Bound and No-accretion prescriptions avoid implicitly the inconsistency of maintaining a sink radius R_{fix} constant throughout the orbital phase, over the time span explored. In the eccentric case the secondary MBH experiences, along a single orbit, different regions of the disc. The disc density is the highest at pericentre where the relative velocity between the gas and the MBH is also the highest. The notion of R_{bound} becomes thus time (phase) dependent. The outcome of the Bound, No-accretion and Fix models can thus be different. The run with R_{fix} can give a consistent description of the MBH dynamics only if, along the orbital phase, R_{fix} remains smaller than the gravitational influence radius. On the other hand on an eccentric orbit, the bound prescription

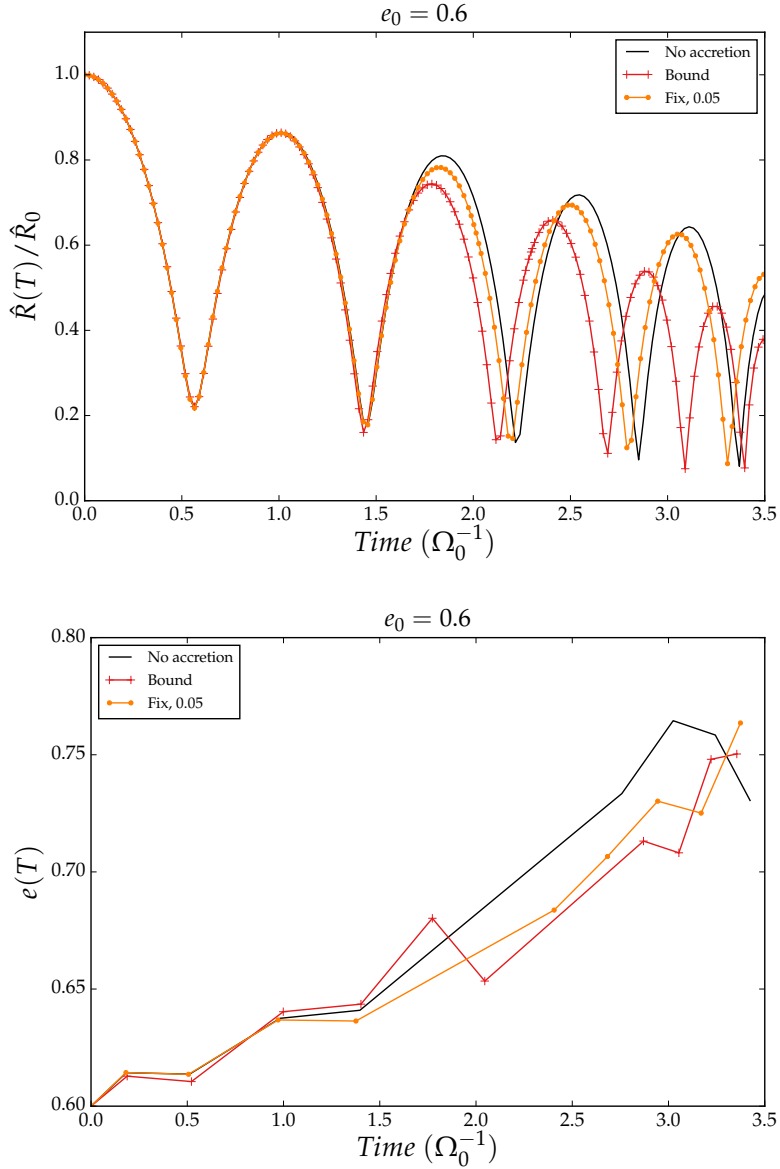


Figure 32: Upper panel: binary separation versus time, for an eccentric binary with $e_0 = 0.6$. Units and colour codes are as in the previous figures. Lower panel: evolution of the binary eccentricity versus time for an eccentric binary with $e_0 = 0.6$.

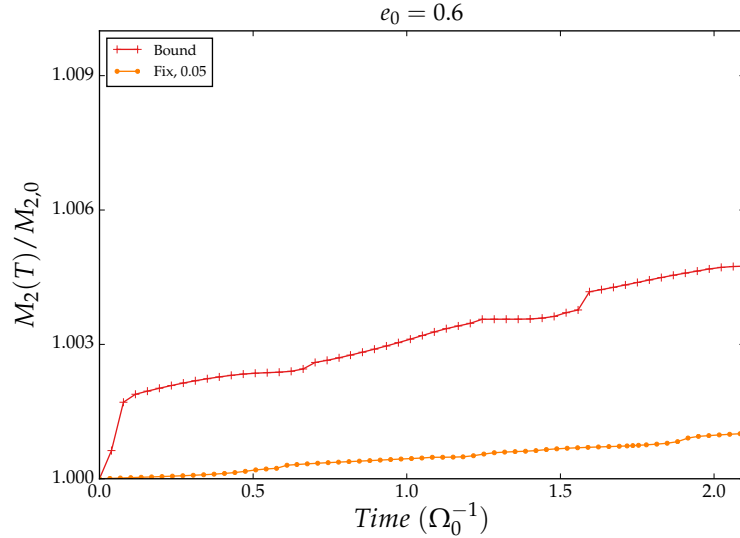


Figure 33: Evolution of the mass of the secondary ($M_2/M_{2,0}$) as a function of time (in units of Ω_0^{-1}) for the eccentric case, with $e_0 = 0.6$.

over-predicts the amount of matter accreted. Gas that is bound at apocentre is accreted promptly when using the bound prescription. However, gas bound at apocentre would unbind at pericentre due to the closer action of tidal forces from the primary MBH. The instantaneous capture prescription in the Bound model thus over-predicts the mass accreted, affecting the underlying dynamics of the disc. In other terms, a small R_{fix} , smaller than the bound radius R_{bound} (at any time over orbital evolution), or the No-accretion model that allows gas to remain dynamically active inside the grid, can account for dynamical processes that can not be captured by the Bound prescription.

9.4.4 Three-dimensional Smoothed-particle hydrodynamics experiments

A clear advantage of using a 2D modelling is that we are able to run many different cases with a relatively low computational cost. However, the approach upon which the 2D models rely must be tested³. In particular, it is important to assess the role of disc thickness in affecting the results.

We hence run a few representative cases using full 3D Smoothed-particle hydrodynamics simulations with GADGET-2⁴ (Springel 2005). The gaseous disc is modelled with 2×10^5 Smoothed-particle hydrodynamics particles, following an isothermal equation of state but for the possible heating term associated with the Smoothed-particle hydrodynamics artificial viscosity. Such viscosity is needed in order to properly recover the occurrence of shocks in the gas, as it acts when converging flows are in place. The viscosity term follows a modified Monaghan-Balsara prescription [Monaghan and Gingold, 1983, Balsara, 1995], with the viscosity parameter $\alpha=0.5$ and $\beta = 2 \times \alpha$. Gravity is computed on a oct-tree, with close encounters among particles being softened through a [Monaghan and Lattanzio, 1985] spline kernel with the same softening parameter of 0.1 for the BHs and gas particles (in internal units).

³ We note that a recent release of the code includes the possibility of 3D models, as described in <http://fargo.in2p3.fr/>

⁴ <http://www.mpa-garching.mpg.de/~volker/gadget/index.html>

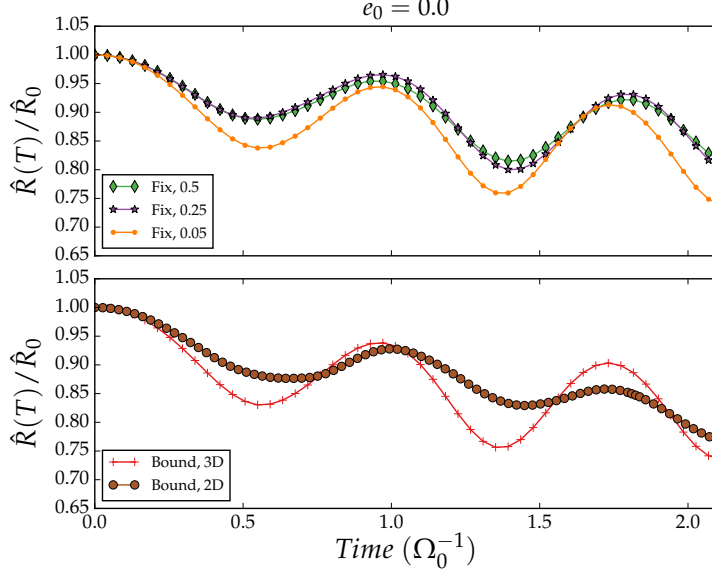


Figure 34: Same as Figure 28, MBHB Orbital separation as a function of time (in units of Ω_0^{-1}) for the circular case, with $e_0 = 0$. Upper Panel: Evolution of the 3D runs assuming the R_{fix} and R_{bound} accretion prescriptions. Lower Panel: Comparison between the R_{bound} prescription between the 2D and 3D experiments.

We implemented a fixed sink radius R_{fix} and a bound radius R_{bound} to mimic the *second* and *third* prescriptions presented in §9.3. The same R_{fix} used in the 2D simulations is introduced here, i.e., 0.05, 0.25 and 0.5. For the "third" prescription, we select gas particles within a fix radius of 1 from each MBH, and among these we check which particles are gravitationally bound to the black hole. In order to do so, we include a parameter $\alpha = 0.3$, similarly to the implementation discussed in [Dotti et al., 2009]. This α parameter enable us to require that a gas particle is more bound to the secondary MBH using the Bondi-Hoyle-Lyttleton radius,

$$R_{bound} = \frac{2 G M_{BH}}{v_{bh}^2 + C_s^2}, \quad (31)$$

where v_{bh} is the relative velocity between the secondary MBH and the gas particle, and C_s is the sound speed of the gas. We accrete particles whenever their separations from the MBHs are less or equal to α times R_{bound} . We initialise the disc using GD_BASIC⁵ [Lupi et al., 2015a] following the same prescription for the 2D case, presented in §9.2, preserving the same aspect ratio $H/\hat{R} = 0.04$. We depict the orbital separation between the two MBHs for the three-dimensional simulations in Fig. 34 and 35 for the circular and eccentric case, respectively. So as to facilitate the comparison with the 2D case, we have added a panel in which we display the bound cases for both the 2D and 3D simulations. After 5 orbital periods, which require about one week of computation on 12 CPUs, the 3D experiment reaches a value of ≈ 0.7 instead of the ≈ 0.8 value of the 2D simulation. Two important results can be inferred from the comparison of the different runs:

- All the discussion about the importance of the accretion prescription already presented for the 2D runs is equally valid in their 3D counterparts. The very

⁵ <http://www.dfm.uninsubria.it/alupi/software.html>

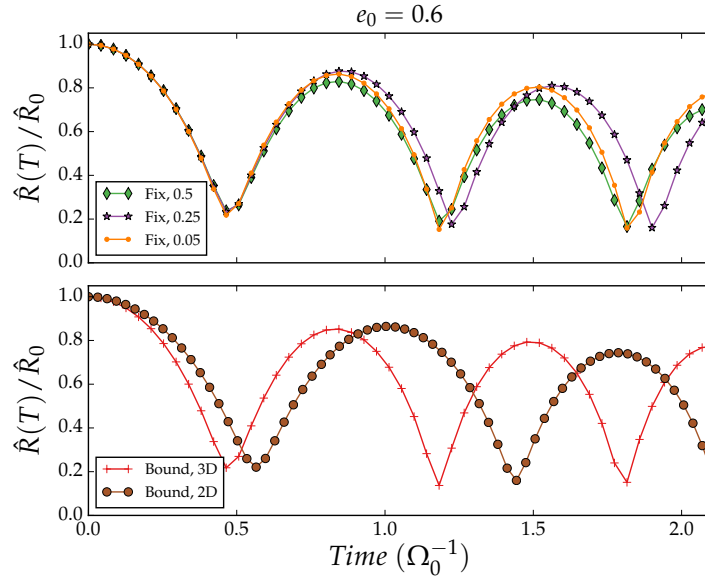


Figure 35: Same as Fig. 32: orbital separation of the binary as a function of time (in units of Ω_0^{-1}) for the eccentric case, with $e_0 = 0.6$. Upper Panel: Evolution of the 3D runs assuming the R_{fix} and R_{bound} accretion prescriptions. Lower Panel: Comparison between the R_{bound} prescription between the 2D and 3D experiments.

similar evolution of the MBH separation as a function of time for the $R_{fix} = 0.05$ and R_{bound} case strengthen the claim that resolving the MBH sphere of influence is a necessary requirement for a proper dynamical evolution.

- We can readily see that the agreement between the 2D and 3D runs in Fig. 34 is satisfactory: in spite of the very different numerical approaches and the simplifications introduced in the 2D case, the evolution of the distance of the binary is very similar, although not identical. The small differences present in the MBH evolution are due to the fact that, in the 2D case the gas and the MBHs are confined to the plane of the disc and, hence, the MBHs are forced to interact with more gas than in the 3D case. In the 3D case, however, the non-negligible disc thickness allows some of the gas not to interact with the MBH.

The effect of the disc vertical profile on the secondary dynamics is way more significant in the eccentric case. In Fig. 35 we see a much more pronounced difference between the 2D and 3D bound cases (lower panel).

As a note of caution we remark that neither the 2D nor the 3D simulations are realistic and accurate representations of circumbinary discs. The idealised thermodynamics of the gas does not catch all the cooling processes taking place in such high density regions. If the disc were allowed to radiate a significant amount of the energy acquired from the interaction with the MBH, it would settle in a geometrically thinner configuration, more similar to the 2D case. In this study we aimed at isolating the effect of the accretion prescriptions on the MBH dynamics, and a detailed study of the effect of the gas thermodynamics is beyond the scope of this thesis.

The comparison of the evolution of the secondary eccentricity in the different runs gives similar results. Fig. 36 shows the eccentricity as a function of time for the circular cases. The trends are similar, although not identical in the 2D and 3D bound cases, reaching the same maximum value (≈ 0.12).

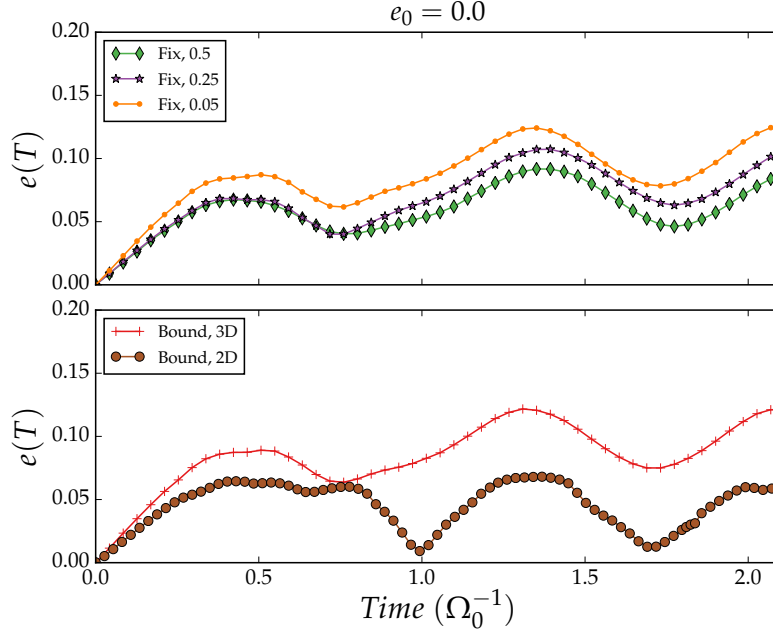


Figure 36: Same as Fig. 30, Evolution of the MBHB eccentricity as a function of time (in units of Ω_0^{-1}) for the circular case, with $e_0 = 0$. Upper Panel: Evolution of the R_{fix} accretion prescriptions. Lower Panel: Comparison between the R_{bound} prescription between the 2D and 3D experiments.

The differences between these two cases are due to the reduced amount of gas mass the secondary interacts with in the 3D case, in which the disc thickens as the time goes by. The comparison among the 3D runs demonstrates again the need of resolving the secondary sphere of influence, with the bound and $R_{fix} = 0.05$ resulting in the same dynamical evolution of the binary. The same comments apply to the eccentric cases shown in Fig. 37.

9.5 A SEMI-ANALYTICAL MODEL FOR THE EVOLUTION OF A BINARY IN AN UNPERTURBED RETROGRADE DISC

In a retrograde disc, the secondary MBH experiences a drag force resulting from gas-dynamical friction on scales larger than R_{bound} and from accretion on scales smaller than R_{bound} . Here we propose an analytical scheme that helps interpreting the run of the eccentricity and semi-major axis of the MBHB versus time (or equivalently the accreted mass), varying the slope of the underlying gas density profile.

In the simplifying assumption that the gaseous background is stationary and that the MBH motion is supersonic, the deceleration force can be approximated as

$$F_{drag} = -4\pi\lambda G^2 M_2^2 \rho_{gas} \frac{V_{rel}}{V_{rel}^3}, \quad (32)$$

where ρ_{gas} is the density of the gas at distances near R_{bound} , and $V_{rel} = V_{rel} \tilde{V}_{rel}$ is the velocity of the accreting MBH relative to the gas velocity. The factor λ identifies with the eigenvalue of the Bondi-Hoyle-Lyttleton model for accretion (equal to 1.12 for a isothermal gas) but here λ , suitably rescaled, can also account for the gas dynamical drag according to Ostriker [1999]. In equation

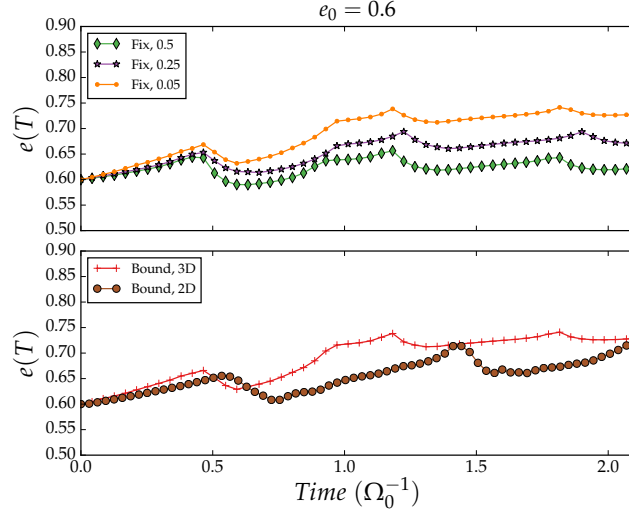


Figure 37: Same as Fig. 32, Evolution of the MBHB eccentricity as a function of time (in units of Ω_0^{-1}) for the eccentric case, with $e_0 = 0.6$. Upper Panel: Evolution of the R_{fix} accretion prescriptions. Lower Panel: Comparison between the R_{bound} prescription

32, $\mathbf{V}_{rel} = \mathbf{V}_2 - \mathbf{V}_{gas}$, where \mathbf{V}_2 is the velocity of M_2 relative to the centre of mass of the binary (we here consider the limit $M_2 \ll M_1$ for simplicity, so that the total mass $M = M_1 + M_2$ of the binary is approximated to M_1 , and the reduced mass μ as M_2) and \mathbf{V}_{gas} the Keplerian velocity of the gas, relative to M_1 , at the current position \hat{R} of M_2 , i.e., $\mathbf{V}_{gas} = (GM_1/\hat{R})^{1/2}\tilde{\mathbf{V}}_{gas}$, with $\tilde{\mathbf{V}}_{gas}$ a unit vector in the direction of \mathbf{V}_{gas} .

To explore the MBHB dynamics and describe the sinking of the secondary MBH in the retrograde disc, we consider the drag force as a perturbation on the Keplerian motion of M_2 in the gravitational potential of the primary MBH, M_1 Vecchio et al. [1994]. We then trace the dynamics of M_2 computing the change of the orbital elements, i.e., the energy and angular momentum, or equivalently the semi-major axis a and eccentricity e under the action of the mean drag force,

$$\langle \mathbf{F}_{drag} \rangle_T = \frac{(1-e^2)^{3/2}}{2\pi} \int_0^{2\pi} \mathbf{F}_{drag}(\psi) \frac{d\psi}{(1+e\cos\psi)^2}, \quad (33)$$

where ψ the orbital phase, and the mean is over the orbital period. In this way we separate the instantaneous motion of M_2 from the motion averaged over an orbital period: here on we will refer to as secular motion, and secular evolution.

The drag force can be cast in the following form in order to separate the modulus from the direction, both time (phase) dependent:

$$\mathbf{F}_{drag} = -\zeta(\hat{R}, V_{rel}) (\tilde{\mathbf{V}}_2 - \tilde{\mathbf{V}}_{gas}) \quad (34)$$

with

$$\zeta(\hat{R}, V_{rel}) = 4\pi\lambda(GM_2)^2\rho_{gas,0} \left(\frac{\hat{R}_0}{\hat{R}}\right)^n \frac{1}{V_{rel}^2}. \quad (35)$$

In equation 34 and 35, the distance \hat{R} and the velocity vectors are function of phase ψ , along the Keplerian motion, while the constants \hat{R}_0 and $\rho_{gas,0}$ denote a reference radius and density in the retrograde, in-homogeneous disc.

In equation 35, the power-law exponent n describes the distribution of the gas density as a function of the distance \hat{R} : $\rho_{gas}(\hat{R}) = \rho_{gas,0}(\hat{R}_0/\hat{R})^n$. The value $n = 1$ corresponds to the disc's density profile at the onset of the hydrodynamical simulations.

After some calculation (sketched in Appendix) the equations for the secular evolution of the MBH mass m_2 , in units of $M_{2,0}$, of the semi-major axis $\tilde{a} = a/a_0$, in units of the initial semi-major axis a_0 , and of the eccentricity e can be cast in a simple form:

$$\frac{\dot{m}_2}{m_2} = \Gamma_0 (1 - e^2)^{(3-n)} \tilde{a}^{(3/2-n)} \langle \mathcal{A}(e) \rangle_\psi, \quad (36)$$

where dot denotes the ‘‘secular’’ time derivative, $m_2 = M_2/M_{2,0}$, and $\langle \mathcal{A}(e) \rangle_\psi$ a dimensionless function of the running value of e (eq. in Appendix). Γ_0 is a constant equal to

$$\Gamma_0 = 4\pi\lambda G \rho_{gas,0} \frac{1}{\Omega_0} \frac{M_{2,0}}{M_1} \left(\frac{\hat{R}_0}{a_0} \right)^n. \quad (37)$$

In this last expression Ω_0 is the Keplerian frequency of the MBHB at a_0 : $\Omega_0 = (GM_1/a_0^3)^{1/2}$. The equation for the dimensionless semi-major axis reads

$$\frac{\dot{\tilde{a}}}{\tilde{a}} = -2 \Gamma_0 m_2 \tilde{a}^{(3/2-n)} [1 - e^2]^{(2-n)} \langle \mathcal{B}(e) \rangle_\psi, \quad (38)$$

where $\langle \mathcal{B}(e) \rangle_\psi$ is given in Appendix.

Similarly, one can calculate the rate of change of the angular momentum and in turn of the eccentricity e . As the orbit of M_2 is coplanar, the direction of the binary orbital angular momentum does not vary with time as the drag is causing only a decrease in the modulus of \mathbf{J} . According to equation 34, the eccentricity evolves as

$$\frac{\dot{e}}{e} = \Gamma_0 m_2 \tilde{a}^{(3/2-n)} \frac{(1 - e^2)^{(4-n)}}{e^2} \langle \mathcal{C}(e) \rangle_\psi, \quad (39)$$

where $\langle \mathcal{C}(e) \rangle_\psi$ is a dimensionless function of the running value of the eccentricity e (eq. in Appendix).

Equations 36, 38 and 39 are coupled and can be solved numerically for $m_{2,0} = 1$, $\tilde{a}_0 = 1$ and initial eccentricity e_0 at time $t = 0$. The results can then be rescaled for any arbitrary value of $M_{2,0}$ and a_0 . In 36, 38 and 39, the timescale that enters the equations is $\tau_0 = \Gamma_0^{-1}$ that can be displayed in the form

$$\tau_0 \sim \frac{1}{8\lambda} \frac{V_{K,0}}{\pi G \Sigma_{gas,0}} \frac{H}{\hat{R}_0} \frac{M_1}{M_{2,0}} \left(\frac{a_0}{\hat{R}_0} \right)^n, \quad (40)$$

where $V_{K,0}$ is the Keplerian velocity at \hat{R}_0 , and $\Sigma_{gas,0} \sim 2H\rho_{gas,0}$. Recalling that $H/\hat{R}_0 \sim c_{s,0}/V_{K,0}$ in a thin isothermal disc (with isothermal sound speed $c_{s,0}$), τ_0 scales as

$$\tau_0 \Omega_0 \sim \frac{1}{8\lambda} Q_0 \frac{M_1}{M_{2,0}}, \quad (41)$$

where we have introduced the Toomre parameter $Q_0 = c_{s,0}\Omega_0/(\pi G \Sigma_{gas,0})$ for disc stability. In the following we will describe the solutions of this simple model in the e versus $M_2/M_{2,0}$ and \tilde{a} versus $M_2/M_{2,0}$ planes.

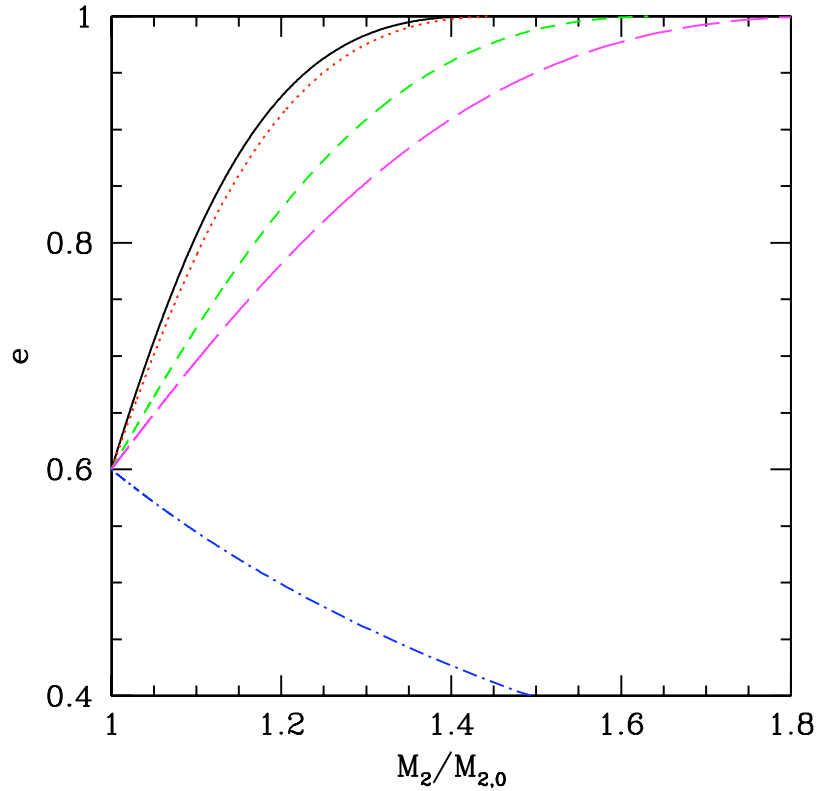


Figure 38: Orbital eccentricity e versus $M_2/M_{2,0}$ for $e_0 = 0.6$, as computed within the semi-analytical model. Solid (black) line refers to a background gas density scaling with distance \hat{R} as a power law with $n = 1$, corresponding to the profile of the hydrodynamical simulation. Red (dotted), green (dashed), magenta (long dashed) and blue (dot-dashed) refer to retrograde discs with $n = 1.3, 2, 2.2$ and 3 , respectively.

9.5.1 Binary evolution trends

In a fixed and non steep gaseous background ($n < 3$), eccentric binaries evolves into more eccentric binaries⁶.

Fig. 38 shows the run of the eccentricity e versus $M_2/M_{2,0}$ for $e_0 = 0.6$, and $n = 1, 1.3, 2, 2.2$ and 3 . The eccentricity grows monotonically up to unity, and this occurs before \tilde{a} has decayed significantly (i.e., by more than two orders of magnitude). The limit $e \rightarrow 1$ is reached after the secondary has accreted a mass comparable to 30% – 60% of the initial mass $M_{2,0}$.

In the case of a steep density background, i.e., $n = 3$, the eccentricity shows an opposite trend and decreases with time, since the gas disc density at pericentre is so high that the secondary MBH experiences the largest drag there. A higher braking force at pericentre reduces the eccentricity and the MBH spirals inwards along orbits progressively less eccentric. The semi-major axis drops dramatically by more than three orders of magnitude, on a time τ_0 .

Nearly circular orbits (with $e_0 = 0.01$) show interesting behaviours in their long-term evolution. We first notice that $n = 3/2$ is a critical value, separating

⁶ We notice that a rapid increase in the eccentricity is also found in an analytical study by Schnittman and Krolik [2015] who expressed the negative torque of a retrograde disc on a MBHB as function of the mass accretion rate, motivated by magneto-hydrodynamical simulations by Bankert et al. (2014).

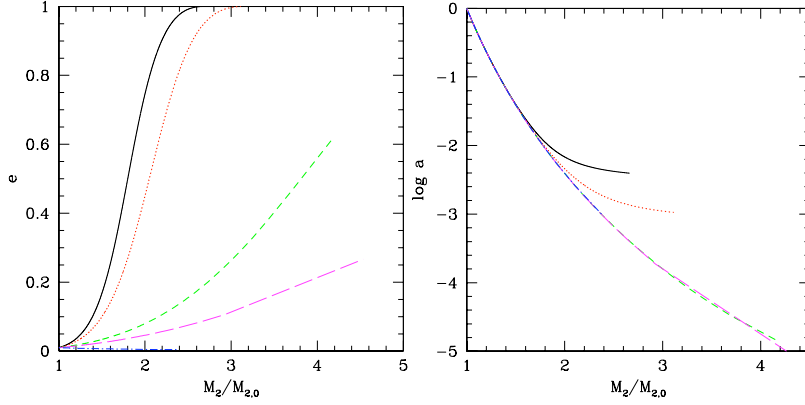


Figure 39: Eccentricity e (left panel) and semi-major axis \tilde{a} (right panel) versus $M_2/M_{2,0}$ for a nearly circular orbit, with $e_0 = 0.01$, as computed within the semi-analytical model. Line colours and style codes are the same as in Fig. 38. For a retrograde disc with background density $\rho_{gas} \propto \hat{R}^{-n}$ scaling with $n < 3/2$ (solid-black, red-dotted lines) the evolution stops at the time the eccentricity $e \rightarrow 1$. By contrast, when $n > 3/2$ (as in the remaining cases, i.e., $n = 2, 2.2$ and 3) the semi-major axis decays very rapidly before the eccentricity has time to rise up to unity.

two trends. Solutions with $n < 3/2$ evolve toward $e \rightarrow 1$ before the semi-major axis has decreased significantly. By contrast, when $n > 3/2$, the decay of the semi-major axis is faster than the increase of e . The value of $n = 3/2$ is critical since the logarithmic derivatives of our variables all scale as $\tilde{a}^{(3/2-n)}$, as indicated in equations 36, 38, and 39. For $n > 3/2$ the decay of the semi-major axis accelerates with decreasing \tilde{a} .

Fig. 39 shows the run of e and \tilde{a} versus $M_2/M_{2,0}$. As a rule of thumb and for circular binaries, the secondary needs to accrete a mass of the order of (2-4) $M_{2,0}$ to slide on the path leading to coalescence by gravitational waves. For $n = 3$, the eccentricity remains small for the entire evolution, and never rises.

In the early phase of the binary evolution (corresponding to $M_2 < 0.2M_{2,0}$), nearly circular binaries remains almost circular, as their eccentricity growth occurs after a sizeable increase of the MBH mass (see Fig. 39). This explains why the initially circular binary models studied with FARGO with different accretion prescription do not display a sizeable increase of e , on the time span of the simulation. This is in line with the findings by Roedig and Sesana [2014].

Our semi-analytical model has applicability in the limit of $M_2/M_1 \ll 1$. Only if $M_2/M_1 \ll 1$ the secondary, lighter MBH causes minor perturbations in the underlying disc so that the background density can be treated as a constant over the binary evolution timescale. For high MBH ratios (as the one considered in this chapter) changes in the density background occur in response to the mutual perturbation excited by the secondary MBH in the retrograde disc, and the evolution needs to be traced only via direct numerical simulations.

9.6 DISCUSSION

In this chapter we explored the evolution of a MBHB embedded in a counter-rotating circumbinary disc. As mentioned in the introduction, the retrograde case differs from the prograde one in three main points: (i) the gravitational torque responsible for the binary shrinking does not halt inflows of gas around the MBHs embedded in the disc; (ii) retrograde gas interacting with the MBHs can remove more angular momentum per unit of mass, since its initial angular momentum has sign opposite to that of the binary; (iii) in the retrograde case, the relative velocities between the MBHs and the disc are significantly larger, so that the interaction between the gas and the MBHs is limited to smaller regions. Since points (i) and (ii) facilitate the binary shrinking while point (iii) limits its strength, it is not obvious a priori if the interaction between a MBHB and a counter-rotating circumbinary disc results in rapid hardening of the binary.

In particular, we pointed out that, as a consequence of the high relative speed between the secondary and the gas, simulations of counter-rotating MBHB-disc systems are strongly affected by the prescription assumed for the accretion of gas onto the MBHs. We argued that in order to model the disc-MBH hydro-dynamics in an appropriate way, a necessary condition is that gas bound accretes onto the lighter, secondary MBH. We confirmed our claim running a suite of 2D hydrodynamical simulations, showing that assuming a too large accretion (sink) radius, such as e.g., the Roche Lobe radius of the secondary (Nixon et al. [2011a]), results in a too fast spurious evolution of the binary (see Figures 28, 29, 31). We further noticed that, while for secondaries moving on quasi-circular orbits the "Bound" prescription results in the correct dynamical evolution of the secondary, for MBHs moving on very eccentric orbits this prescription is not sufficient. This is due to the fact that gas bound to the secondary at the apocentre can unbind during the close pericentre passages, avoiding to be accreted. The non-accreted gas changes the effective mass of the secondary (i.e., the mass of the secondary plus the mass of the gas co-moving to it), and can rejoin the background gas distribution, possibly further interacting with the secondary at later times. For this reason, a sink radius smaller than the influence radius of the MBH, varying over the orbital phase, has to be set properly in order to predict the binary dynamical evolution (see Fig. 32, 33).

The 3D Smoothed-particle hydrodynamics simulations run in parallel for selected cases have shown a close match with the results of the 2D simulations validating the findings in 2D, despite the intrinsic physical difference related to the different dimensionality and the difference in the numerical method.

We focused on the evolution of the binary orbit giving particular emphasis to the effects of the accretion prescriptions. The present analysis still lacks of some possibly important physical effects. We limited the mass of the disc to its initial value, and we do not study the effect of a continuous (or episodic) re-fuelling of the disc from the outer regions of the nucleus. Such a fuelling is necessary for the coalescence of a binary in a prograde disc [as discussed in Dotti et al., 2012], and can boost the brake of the binary in the retrograde case as well. Finally, we did not model any disc fragmentation and star-formation in the disc, that (together with the fuelling of the binary from large scales) is the main unknown in the evolution of binaries in prograde discs as well [see e.g., Lodato et al., 2009, Amaro-Seoane et al., 2013b]. The effects of these processes, together with a detailed discussion of the peculiar observational properties of counter-rotating systems, is postponed to future investigations.

In this Appendix we shortly outline the key steps for deriving the evolution equations 36, 38 and 39 of Section 5. We follow closely the derivation by [Vecchio et al., 1994].

First we specify the instantaneous motion of the secondary MBH, described by the velocity vector \mathbf{V}_2 and the separation vector $\hat{\mathbf{R}}_2$, relative to the centre of mass of the binary. Since the accretion drag is a weak perturbing force, the motion is determined by the driving force by the MBH primary. The instantaneous values of $\hat{\mathbf{R}}_2$ and \mathbf{V}_2 are expressed in term of the orbital phase ψ , along the Keplerian motion. They are defined uniquely by the instantaneous values of the energy and angular momentum per unit mass that we here denote as E and J , or alternatively a and e . For the distance modulus we have

$$\hat{R}_2 = \frac{a(1 - e^2)}{1 + e \cos \psi}. \quad (42)$$

The velocity is decomposed along the radial and tangential directions

$$V_{2,\hat{R}_2} = \left[\frac{GM}{a(1 - e^2)} \right]^{1/2} e \sin \psi, \quad (43)$$

$$V_{2,t} = \left[\frac{GM}{a(1 - e^2)} \right]^{1/2} (1 + e \cos \psi), \quad (44)$$

where $M = M_1 + M_2 \sim M_1$ in the limit of a massive primary. Furthermore,

$$V_2 = \frac{GM_1}{a(1 - e^2)} (1 + 2e \cos \psi + e^2). \quad (45)$$

The instantaneous deceleration on the secondary MBH due to accretion can be written as

$$\dot{V}_{drag} = -\frac{1}{M_2} \zeta(\hat{\mathbf{R}}_2, V_{rel}) (\tilde{\mathbf{V}}_2 - \tilde{\mathbf{V}}_{gas}) \quad (46)$$

with

$$\zeta(\hat{\mathbf{R}}_2, V_{rel}) = 4\pi\lambda(GM_2)^2 \rho_{gas,0} \left(\frac{\hat{R}_0}{\hat{R}_2} \right)^n \frac{1}{V_{rel}^2}. \quad (47)$$

The Keplerian velocity of the fluid elements at distance \hat{R} is $\mathbf{V}_{gas} = (GM_1/\hat{R})^{1/2} \tilde{\mathbf{V}}_{t,gas}$ which is vector tangential to $\hat{\mathbf{R}}/\hat{R}$, as the gas in the retrograde disc is moving on circular orbits. According to equations 42, 43, 44 and 46,

$$\mathbf{V}_2 \cdot \mathbf{V}_{gas} = -\frac{GM_1}{a(1 - e^2)} (1 + e \cos \psi)^{3/2}. \quad (48)$$

The instantaneous rate of change of the energy per unit mass and angular momentum per unit mass (in the direction of \mathbf{J} , as the orbit is coplanar with the disc's plane) read:

$$\dot{E} = \dot{V}_{drag} \cdot \mathbf{V}_2 \quad (49)$$

$$\dot{j} = (\hat{\mathbf{R}} \times \dot{V}_{drag}) \cdot \frac{\mathbf{J}}{J} \quad (50)$$

The rate of change of the energy per unit mass correlates with the rate of change of the semi-major axis through the relation $\dot{E}/E = -\dot{a}/a$, where

$E = -GM_1/2a$. Similarly, the rate of change of the eccentricity e scales as $e\dot{e} = -(1-e^2)[\dot{J}/J + (1/2)\dot{E}/E]$, where $J^2 = GM_1a(1-e^2)$ [Vecchio et al., 1994]. As the secondary MBH accretes gas, the instantaneous rate of change of the mass M_2 is set equal to the Bondi-Holye-Littleton accretion rate:

$$\dot{M}_2 = \dot{M}_{\text{BHL}}. \quad (51)$$

The drag force due to accretion induces secular changes in the orbital elements of the binary. We thus calculate the secular rate of change of a, e and M_2 averaging all physical quantities over the Keplerian period P . As along a Keplerian orbit, the phase ψ is related to the time coordinate by

$$\frac{1}{P}dt = \frac{1}{2\pi}(1-e^2)^{3/2} \frac{1}{(1+e\cos\psi)^2} d\psi, \quad (52)$$

we convert the means in the time domain to the phase domain, defining $\langle \cdot \rangle_T = (1-e^2)^{3/2} \langle \cdot (1+e\cos\psi)^{-2} \rangle_\psi$, hereon.

Using the above relations, we derive equations 36, 38 and 39. If we define $f = 1 + e^2 + 2e\cos\psi$ and $g = 1 + e\cos\psi$, the expressions for the means in the phase domain introduced in equation 36, 38 and 39 read:

$$\langle \mathcal{A}(e) \rangle_\psi = \frac{1}{2\pi} \int_0^{2\pi} d\psi \frac{g^{(n-2)}}{(f+g+2g^{3/2})^{3/2}} \quad (53)$$

$$\langle \mathcal{B}(e) \rangle_\psi = \frac{1}{2\pi} \int_0^{2\pi} d\psi \frac{g^{(n-2)}(f^{1/2}+g)}{(f+g+2g^{3/2})} \quad (54)$$

$$\langle \mathcal{C}(e) \rangle_\psi = \frac{1}{2\pi} \int_0^{2\pi} d\psi \frac{g^{(n-2)}}{(f+g+2g^{3/2})} \mathcal{I}(\psi, e) \quad (55)$$

where

$$\mathcal{I}(\psi, e) = \left(f^{-1/2} + g^{-1} - f^{1/2}/(1-e^2) - g/(1-e^2) \right). \quad (56)$$

Notice that in the limit of $e \rightarrow 0$, \dot{e} is order $\mathcal{O}(e)$. The expansion analysis around $e \rightarrow 0$ shows that \dot{e}/e is proportional to $(11/4 - n)$. Thus, for small initial eccentricities and $n > 11/4$, the orbit circularises while in the opposite case ($n < 11/4$), e grows in time. In the same limit \dot{a} is finite, and the MBH continues to spiral in.

CIRCUMBINARY STRUCTURE FORMATION FROM
MULTIPLE INFALLING CLOUDS

The project presented on this chapter was published under:

ACCRETION OF CLUMPY COLD GAS ONTO MASSIVE BLACK HOLES BINARIES:
THE CHALLENGING FORMATION OF EXTENDED CIRCUMBINARY STRUCTURES

C. Maureira-Fredes, F. G. Goicovic, P. Amaro-Seoane & A. Sesana
Monthly Notices of the Royal Astronomical Society,
Volume 478, Issue 2, 1 August 2018, Pages 1726–1748.

and

ACCRETION OF CLUMPY COLD GAS ONTO MASSIVE BLACK HOLE BINARIES:
A POSSIBLE FAST ROUTE TO BINARY COALESCENCE

F. G. Goicovic, **C. Maureira-Fredes**, A. Sesana, P. Amaro-Seoane & J. Cuadra
Monthly Notices of the Royal Astronomical Society,
Volume 479, Issue 3, 21 September 2018, Pages 3438–3455.

10.1 INTRODUCTION AND MOTIVATION

We have mentioned in the introduction that thanks to advances in high angular resolution instrumentation, our understanding of the central regions in galaxies has gone through a major breakthrough. Space-borne telescopes such as the Hubble Space Telescope and adaptive optics from the ground allow us to investigate the kinematics and distribution of gas and stars at sub-parsec scales for external galaxies [see e.g. Kormendy, 2003, Kormendy and Ho, 2013] and to milli-pc for the Milky Way (see e.g. Schödel et al. 2003, 2014, Gallego-Cano et al. 2018, Schödel et al. 2018 and, in particular, the review of Genzel et al. 2010). A capital conclusion is that massive dark compact objects, very likely massive black holes (M_{BH}), with a mass typically ranging between 10^6 – $10^9 M_{\odot}$ are lurking at the innermost centre of most large galaxies in the observable range. Moreover, the formation and evolution of these objects and their host galaxies seem to be linked [Magorrian et al., 1998, Ferrarese and Merritt, 2000, Gebhardt et al., 2000, Häring and Rix, 2004].

On the other hand, from a theoretical standpoint hierarchical models best explain the formation of structures in the Universe, down to the size of a galaxy. This implies that galaxies, during their lifetime, have suffered at least one merger, if not several [see e.g. Volonteri et al., 2003, for hierarchical merger studies in Λ CDM Cosmology]. If these galaxies harbour a M_{BH} in their centre, during the collision they will sink to the centre of the merged galaxy, form a binary and shrink its semi-major axis and become “harder” thanks to the interaction with stars coming from the surrounding stellar system in which they are embedded [see e.g. Begelman et al., 1980, Quinlan, 1996, Sesana et al., 2007, Colpi and Dotti, 2011]. A star will strongly interact with the binary of M_{BH} s by removing energy and angular momentum out of it, so that it will be re-ejected into the stellar system with a higher kinetic energy, and

the semi-major axis of the binary will shrink a bit more. Nevertheless, the loss-cone [see e.g. Frank and Rees, 1976, Amaro-Seoane et al., 2004], which provides the M_{BH} binary (MBHB) with stars to interact with, is soon empty, halting the evolution of the binary. The pace of subsequent shrink of the MBHB is dictated by the rate at which new stars are scattered into the loss cone. For spherical stellar distributions, the relevant scale is set by the two-body scattering timescale, that can exceed the Hubble time [Milosavljević and Merritt, 2001]. When this happens, the binary stalls: the MBHBs are still bound but they will not coalesce in a Hubble time. They are emitting gravitational waves, but the energy loss is too weak.

This situation is known as “the last parsec problem”, because the separation between the M_{BH} s is typically of the order of ~ 1 pc [Begelman et al., 1980, Milosavljević and Merritt, 2003, Merritt and Milosavljević, 2005]. Whether or not these binaries will merge in a Hubble time is a question that depends on various issues and is currently debated. However, the emerging consensus is that such binaries will in most of the cases overcome this hang-up, due to efficient loss cone re-population in the triaxial, dynamically un-relaxed stellar distribution produced by galaxy mergers [Berczik et al., 2006, Preto et al., 2011, Khan et al., 2011, Vasiliev et al., 2015, Sesana and Khan, 2015]. MBHBs will therefore get to the phase in their evolution in which they efficiently emit gravitational waves to be detected with a GW observatory such as the Laser Interferometer Space Antenna [LISA, see Amaro-Seoane et al., 2017, 2013a, 2012b].

In gas rich galaxies, which become dominant at low galaxy masses and/or high redshifts, a key factor to surmount this last stretch is the role of the gas, which may be crucial in the evolution of the binary [Escala et al., 2004, 2005]. For instance, the work of Cuadra et al. [2009] found that gaseous discs should commonly help in the merger of M_{BH} s with masses of interest for LISA, whilst this mechanism fails for masses larger than $\sim 10^7 M_{\odot}$.

The evolution of MBHBs interacting with a circumbinary gaseous structure has been investigated by a number of authors. In most studies, gas is assumed to be distributed in a steady, extended circumbinary disc that is either co-rotating [see e.g. Armitage and Natarajan, 2005, Cuadra et al., 2009, Haiman et al., 2009, Kocsis et al., 2012, Amaro-Seoane et al., 2013b, D’Orazio et al., 2013a, Tang et al., 2017] or a counter-rotating [see e.g. Roedig and Sesana, 2014, Nixon and Lubow, 2015, Amaro-Seoane et al., 2016] with respect to the rotation of the MBHB. The binary-disc system is assumed to evolve in isolation, and almost no attempt has been made to connect the system with the larger scale galactic environment that is providing the mass supply to the gaseous structure. The violent interaction of two merging, gas rich galaxies is expected to produce a turbulent environment in which cold clumps and filaments of gas continuously interact exchanging angular momentum and eventually infalling toward the centre of the merger remnant. The initial orbit of the galaxy merger provides a large source of angular momentum, the geometry and angular momentum distribution of the infalling matter is therefore set by the competing effect of ordered dynamics due to the large scale rotation of the merger remnant and chaotic motions driven by turbulence and clumps and filament scattering [Sesana et al., 2014].

It has been proposed that accretion of gas on to the binary can happen either (i) in a coherent way, so that the angular momentum direction is basically kept constant through all the episodes of accretion [Dotti et al., 2010], (ii) stochastically, meaning that the gas accretes on to the black holes in randomly oriented planes [King et al., 2005, King and Pringle, 2006], for

which we have observational evidences, such as in the nucleus of the Abell 2597 galaxy cluster [Tremblay et al., 2016] or in the AGN PKS B1718-649 [Maccagni et al., 2018]. Cold chaotic accretion has been linked to several physical mechanisms acting on the interstellar medium, such as turbulence, rotation, AGN/Supernovae feedback, among others [Hobbs et al., 2011, Gaspari et al., 2013, 2015, 2017]. Also, it has been pointed that (iii) actually accretion might be a mixture of both, coherent and stochastic accretion, with different degrees of anisotropy, as investigated by Dotti et al. [2013] to address possible anisotropies present in the gas in the nuclear regions of active galaxies.

Accretion on to a MBHB of single clouds has been investigated numerically by Dunhill et al. [2014], Goicovic et al. [2016] and Goicovic et al. [2017], taking into account different orbital configurations and cooling rates. We have found that the interaction prompts a transient phase of high accretion onto the MBHB, while part of the leftover gas settles into a circumbinary disc of various masses and sizes depending on the initial orbit of the cloud. Whether such gaseous structure can be kept stable and grow in time under the influence of a series of episodic accretion events is unclear, as the successive infall of various clouds on to a MBHB has not been addressed yet. In this chapter, we present for the first time to our knowledge dedicated, detailed numerical simulations of repeated gaseous infall episodes and accretion on to a MBHB in a post-merger galactic nucleus to assess the architecture of the gas forming around the binary. We consider stochastic feeding of the binary, assuming different degrees of anisotropy in the distribution of the infalling clouds, as well as different feeding rates.

This chapter is organised as follows. In §10.2 we describe the main technical features of the simulation, focusing on physical ingredients and numerical details that have been developed specifically for this suite of simulations. In §10.3 we define the initial conditions of each individual run, providing all the cloud parameter details necessary to reproduce our runs. We test the stability and convergence of our numerical scheme against critical parameters such as the assumed sink radius and the number of particles used to simulate each cloud in §10.4. The results of all simulations, including an extensive description of the transient and long term circumbinary structures, and individual mini-discs are presented in §10.5 and the significance of our main finding and future outlook are discussed in §10.6. The impact on the evolution of the binary itself is presented in a complementary paper (Goicovic et al. [2018]) to this chapter.

Additional data and multimedia material can be found in the website of the project ¹.

10.2 METHODS

We study the evolution of the M_{BH} -clouds system using GADGET-3, a Smooth Particle Hydrodynamics (SPH) and N -body code, which is a modified version of GADGET-2 ². Every particle is represented as a point-mass, characterised by its 3-D position and velocity, and the code solves for the hydro-dynamical and gravitational interaction of the elements, which are organised in a tree structure [Barnes and Hut, 1986].

Specific to our implementation is the modelling of the MBHB as two equal-mass sink particles, initialised in a Keplerian circular orbit with centre of mass

¹ <http://multipleclouds.xyz/>

² <http://wwwmpa.mpa-garching.mpg.de/galform/gadget/>

at rest in the origin of our Cartesian reference frame. We established the code units such as the initial semi-major axis, orbital period and mass of the black hole binary are equal to one, i.e. $a_0 = P_0 = M_0 = 1^3$, which can be re-scaled to an astrophysical system as detailed in Goicovic et al. [2016]. Along the chapter, we will present results for a physical MBHB with initial semi-major axis of $a = 0.2\text{pc}$, a total initial mass of $M_{\text{bin}} = 10^6 M_\odot$, and an orbital period of $P = 8400\text{ yr}$. We will refer to this astrophysical rescaling as our *fiducial system*.

Each infalling molecular cloud is modelled as a spherical distribution of equal-mass SPH particles. Each cloud has a mass of $M_c = 0.01 M_0$ and, to verify the convergence of our simulations, is modelled with increasing resolution, using four different number of particles

- $N_c = 5 \times 10^4$ (50k, default simulation value),
- $N_c = 2 \times 10^5$ (200k),
- $N_c = 5 \times 10^5$ (500k),
- $N_c = 1 \times 10^6$ (1m).

The main configuration of our simulation set-up is described in Goicovic et al. [2016] and visualised in Fig. 40. It is an adaptation of the scheme presented in Bonnell and Rice [2008], in which we:

- replace the central black hole by a MBHB;
- set the initial separation between the cloud and the MBHB to 15 (in code units, corresponding to 3 pc rescaled to our fiducial system);
- adopt an initial velocity modulus of the cloud of $v_c = 0.5 v_0 = 0.5 \sqrt{GM_0/a_0}$;
- gauge the encounter impact parameter by varying the θ_{v_c} angle formed by the cloud velocity vector with the direction connecting the centre of masses of the cloud and of the MBHB.

In the top panels of Fig. 40, we sketch the set up of a co-rotating and a counter-rotating encounter. In the first case, the angular momentum vector of the cloud orbit, \vec{L}_c , is aligned with the orbital angular momentum vector of the MBHB, \vec{L}_{bin} ; in the latter, the two angular momenta have opposite direction. In general, the two angular momenta can have arbitrary relative orientation. In our set-up, we fix the reference frame so that the MBHB is initially at rest in the origin and \vec{L}_{bin} is along the positive z axis. We then generally refer to a cloud as “co-rotating” (“counter-rotating”) if the z component of its \vec{L}_c is aligned (counter-aligned) to \vec{L}_{bin} .

When added to the simulation, the cloud is composed by spherically distributed particles with uniform density, sustained against gravitational collapse by a turbulent internal velocity field. We set up this turbulence by drawing from a Gaussian random distribution with a power spectrum $P_v(k) \propto k^{-4}$, where k is the wave-number of the velocity perturbation in Fourier space. This distribution is chosen to match the observed velocity of molecular clouds [Larson, 1981].

³ Throughout the chapter the subscript ₀ will refer to initial system parameters.

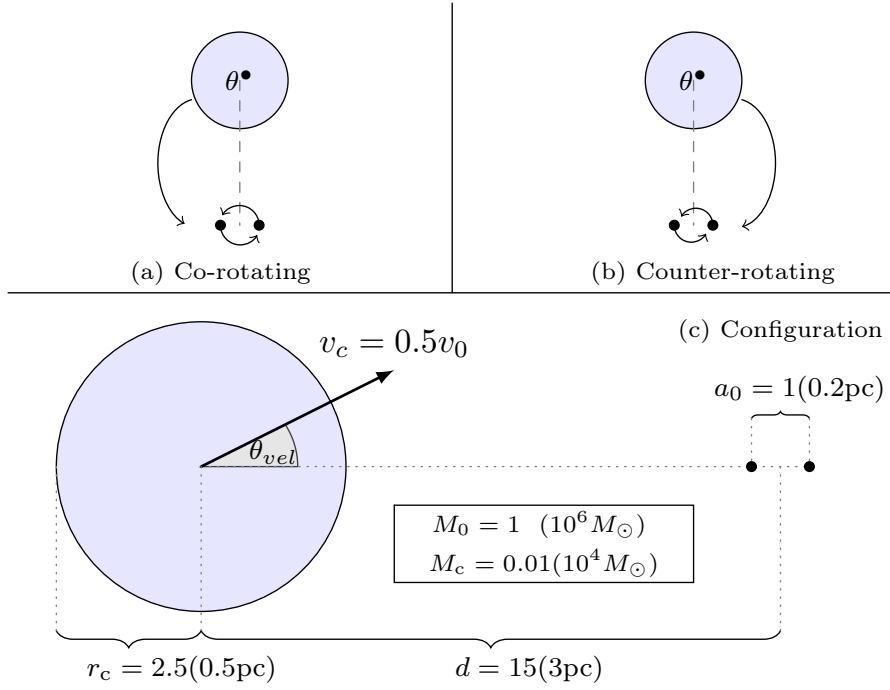


Figure 40: Schematic representation of selected configurations of the simulations: (a) Co-rotating case, the gas cloud infalls following the MBHB orbit; (b) Counter-rotating case, the gas cloud infalls oppositely to the MBHB orbit. Panel (c) sketches how the cloud impact parameter is tuned by changing the angle θ_{vel} (see description in main text). The parameters of the system (initial separation between the cloud and MBHB, initial cloud velocity, masses and sizes of both the cloud and the MBHB) are listed in the figure.

10.2.1 *Additional physics**Accretion*

Each M_{BH} in the simulation is represented by a sink particle, which means that it will accrete any particle entering a pre-defined sink radius, but otherwise it interacts only gravitationally with other SPH particles. We do not use the built-in stochastic accretion model that is included in GADGET, which is not appropriate for the problem in hand, where a reliable accretion procedure for individual gas particles, beyond a stochastic selection scheme, is required. Entering the sink-radius must be a necessary but not sufficient condition for accretion. In fact physically unbound particles can simply fly-by across the sink radius. Therefore, to avoid spurious accretion, we verify that each candidate particle entering the sink radius is bound to the associated M_{BH} . In practice, a gas particle will be accreted if the following conditions are satisfied:

$$r_{\text{gas}} \leq r_{\text{sink}} \quad (57)$$

$$E_{\text{bind}} < 0 \quad (58)$$

where r_{gas} is the relative distance between the M_{BH} and gas particle, and we take a fiducial $r_{\text{sink}} = 0.1$ which is a 10% of the initial separation of the MBHB. A convergence study of the results against the chosen value of r_{sink} is presented in §10.4.1. In order to be able to verify angular momentum conservation, it is important to keep track of the accreted gas particles. The positions and velocity of those particles at the moment of accretion are locally stored, to be used in the computation of the angular momentum of the system when required.

Thermodynamics

The thermodynamics of the gas is modelled using a barotropic equation of state [Bonnell, 1994, Escala et al., 2005, Dotti et al., 2006],

$$P = K\rho^\gamma, \quad (59)$$

where the constant γ depends on the gas density. The value of K represents the entropic function of each gas particle in the code [Springel and Hernquist, 2002], and because it depends explicitly of γ , it must be calculated so that the pressure behaves as a continuous function of density. We choose an effective equation of state that represents the behaviour of a collapsing protostellar cloud, whereby the low density gas evolve isothermally up to some critical density (ρ_c), at which point it becomes adiabatic [Bate et al., 1995]. This can be represented as:

$$\gamma = 1.0 \quad \text{for} \quad \rho \leq \rho_c; \quad (60)$$

$$\gamma = 1.4 \quad \text{for} \quad \rho > \rho_c \quad (61)$$

As explained in Goicovic et al. [2016], the addition of these two regimes breaks the scale-free nature of our simulations, although the results can be scaled within a certain range of critical densities, as explained in Goicovic et al. [2016].

For the simulations presented in this chapter we choose $\rho_c = 10^{-2} M_0 a_0^{-3}$ in code units, corresponding to $10^{-16} \text{g cm}^{-3} \approx 1.5 \times 10^6 M_\odot \text{pc}^{-3}$ when scaled to our fiducial system. Note that this is two orders of magnitude smaller than the value adopted in Goicovic et al. [2016]. This modification was introduced to save computational cost, as some of the configurations produce a large number of gas clumps. Evolving these high density regions requires very small time-steps. Hence, by effectively stopping the collapse at much lower densities, this modification prevents simulations from stalling due to excessively small time-steps.

With this approach we are mimicking the evolution of a multi-phase gas without implementing sophisticated cooling mechanisms or radiative transfer schemes. The isothermal treatment of the gas is qualitatively representative of optically thin media, where cooling is very efficient. Since our goal is to study the dynamics of the gas and not the detailed thermodynamic evolution of the dense regions, this treatment suits our purposes. Because the gas dynamics is mostly dominated by the gravitational potential generated by the binary, we do not expect the thermodynamics to have a major impact in our results. Nevertheless, future studies should include a proper thermodynamic treatment of the gas.

External potential

MBHBs live in the dense environment of galactic nuclei, sitting at the bottom of the galactic potential well. Although this potential does not significantly affect the cloud-MBHB individual interaction, it is important to take it into account when performing an extensive experiment including the interaction of multiple clouds. In fact, the external potential is relevant in two ways:

- it changes the dynamics of the gas flung away by the binary, keeping it bound to the system and allowing it to come back for further interactions;
- it acts as a restoring force, keeping the binary close to the origin of the reference frame (i.e. to the bottom of the potential well).

In practice the inclusion of the potential does not greatly affect the dynamics of the close MBHB-cloud encounters – which remains dominated by the gravity exerted by the MBHB – and it ensures angular momentum and energy conservation due to its spherically symmetric nature. Most importantly, it avoids drifting of the system away from the coordinate origin, which is problematic when clouds are added to the system at different times. Instead, the MBHB experiences a gentle wandering with no secular effects Fig. 43 and, together with its surrounded gas structures, is kept close to the reference frame origin.

To determine the potential, we assume that matter is distributed around the origin following an Hernquist density profile [Hernquist, 1990]:

$$\rho(r) = \frac{M_* a_*}{2\pi r} \frac{1}{(r + a_*)^3}, \quad (62)$$

which implies a cumulative mass distribution given by

$$m_*(r) = M_* \frac{r^2}{(r + a_*)^2}, \quad (63)$$

where M_* and a_* are the total mass and scale radius of the distribution. Based on our fiducial system where $M_{\text{bin}} = 10^6 M_\odot$, we derived M_* by assuming the hole to bulge mass relation of Magorrian et al. [1998] and we computed the scale radius using the radius to stellar mass relation of Dabringhausen et al. [2008]. In code units, this gives

- $M_* = 4.78 \times 10^2 M_0$,
- $a_* = 3.24 \times 10^2 a_0$,

which implies $m_*(< a_0) \approx 5 \times 10^{-3} M_0$, thus a negligible effect on the Keplerian nature of the MBHB.

10.2.2 Other technical adjustments and numerical calibration

Dynamics of the sink particles

The simple inclusion of dynamical 'sinks' several order of magnitude more massive than the other SPH particles, introduced a number of issues with the SPH scheme. This was already noticed in Cuadra et al. [2009], who proposed to extract the sink particles from the tree for a better integration of their trajectories. We adopted the same strategy here, integrating the MBHB orbit with a fixed time step $\Delta T = 0.02 P_0$, thus allowing the binary positions and velocities to be updated more often than typical SPH particles. We verified that this made the evolution more reliable, ensuring that no crucial interactions between particles and the M_{BHs} were missed along the integration.

Still, close inspection of the MBHB evolution showed unphysical jumps in the angular momentum of the system. We verified that this was related to the frequency of update of the SPH tree. In SPH simulations one can choose how often the particle tree (that defines how particles are grouped in computing mutual forces) is generated and updated, which can be controlled by an adjustable parameter in GADGET. So long as the system does not experience dramatic changes, simulations run smoothly with sparse tree updates. However, we are dealing here with multiple clouds infalling onto a MBHB from different directions, triggering violent episodes of accretion, which is clearly not the standard system handled by SPH codes of this type. We found that this required reconstructing the tree 100 times more often than in the default GADGET configuration.

Injection of clouds

At the beginning of each simulation, only one cloud is present in the system besides the MBHB. All the following clouds interact with the binary at later times and therefore need to be included into the system 'on the fly'. Once the injection time for the new cloud is reached, the simulation is stopped and the new cloud is added. Each cloud is characterised by a specific set of ids, so that particles can always be tracked back to their original cloud. This is useful to track the relative importance of each cloud in the accretion process, or in the formation of specific circumbinary structures. Once the cloud is added, the simulation is resumed and the integrator can adapt to the new particles, forming the tree again, and handling this new scenario. The procedure is repeated for each cloud. Note that due to the inclusion of new clouds, the total angular momentum is not conserved. The angular momentum of each injected cloud is, however, known, and it is therefore easy to track angular momentum conservation along the integration of the system.

With the exception of the ones accreted by the MBHB, we do not remove any particle from the simulation. In fact, because of the way the tree is constructed, particles that are flung far away from the binary are grouped in large structures and integrated rarely, representing a negligible contribution to the computational burden. Moreover, the addition of the external potential keeps the structure compact, minimising the number of particles escaping at distances larger than $100a_0$.

Softening and sink radius

Finally it is important the selection of appropriate softening parameter. We choose for the sink particles a value of $\epsilon_{\text{BH}} = 0.01a_0 = 0.002\text{pc}$, and for SPH particles $\epsilon_{\text{gas}} = 0.001a_0 = 0.0002\text{pc}$. These values are small enough to ensure we are not bypassing gravitational interactions, and are an order or magnitude smaller than the maximum value recommended by Bonnell and Rice [2008]. As mentioned above, we fix the sink radius at $r_{\text{sink}} = 0.1a_0$. We performed (see §10.2.2) a series of tests ensuring that neither the dynamics of the MBHB nor the accretion of SPH particles is sensitive to the specific choice of r_{sink} .

10.3 INITIAL CONDITIONS AND RUN DESCRIPTION

In the previous Section, we defined the main physical ingredients and technical features of our simulations, we now proceed in detailing the initial conditions of our set of runs. Our goal is to simulate a series of clouds interacting with a central MBHB sitting at the bottom of a fixed potential well. The MBHB is initially in the coordinate frame origin and has a separation $a_0 = 1$ in code units. Each cloud is injected at a distance $d = 15$ and needs the specification of a time of injection, impact parameter and direction of the orbit.

We construct two series of 30 events, drawing the time between each event from a cumulative Gamma distribution with $k = 2$ and $\theta = 2.5P_0$. We made several draws from the Gamma distribution and picked two markedly distinct sets. In the first set, hereinafter RunA, we perform an “aggressive” feed to the MBHB, with an average $\Delta t \approx 3P_0$ between events. On the other hand, in the second set, hereinafter RunB, clouds are fed to the MBHB with an average $\Delta t \approx 6P_0$, allowing the system more time to relax in between each infalling cloud. Note that when scaled to our fiducial system, the above infall rates correspond to $0.4M_{\odot}\text{yr}^{-1}$ and $0.2M_{\odot}\text{yr}^{-1}$ entering the inner parsec respectively, comparable to what is typically found in high resolution simulations of gas-rich high redshift galaxies [see, e.g. Prieto et al., 2017], and is in broad agreement with observations of post-merger galaxies [see e.g. Sanders and Mirabel, 1996, Naab and Burkert, 2001].

Cloud impact parameters are drawn so that the periapsis passage is uniformly distributed in the range $r_p \in [0, 2a_0]$, if the MBHB was replaced by a single M_{BH} sitting at the origin of the coordinate system and the potential well was ignored. A uniform periapsis distribution corresponds to a standard impact parameter distribution $p(b) \propto b$ at infinity, when the trajectory is dominated by gravitational focusing of the central object, as it is the case in our simulations. The injection time and r_p value of each cloud for both RunA and RunB are shown in Fig. 41.

After specifying the time of injection and impact parameter, we define the orbit of the incoming cloud by assigning a direction to its orbital angular momentum, \vec{L}_c , with respect to \vec{L}_{bin} . We explore three different sets of initial

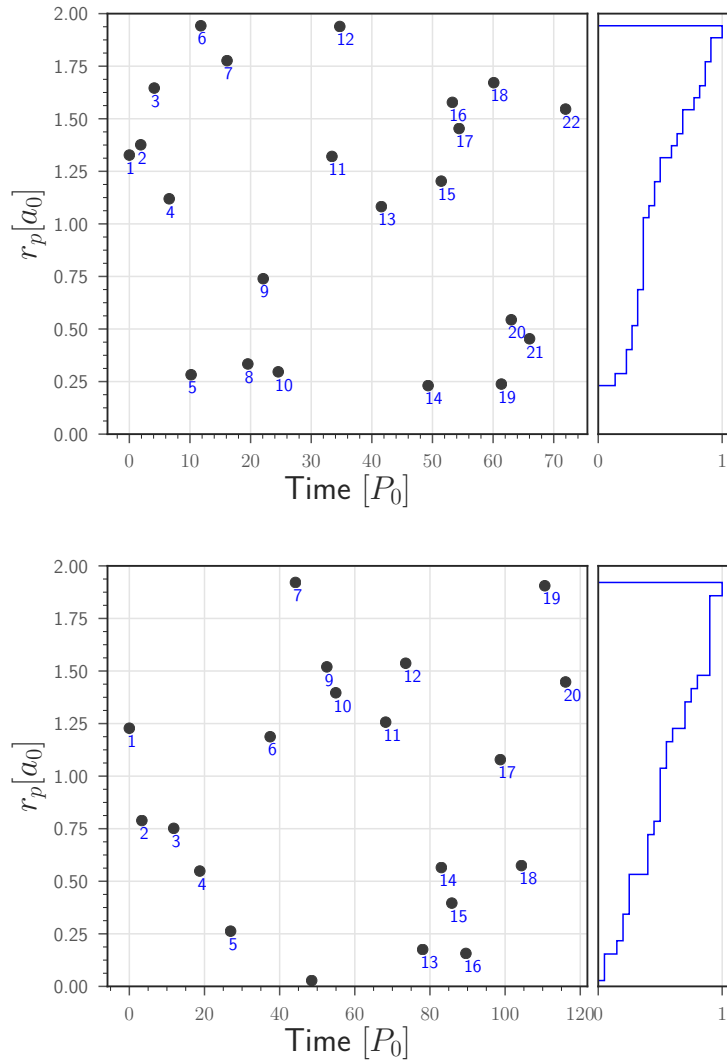


Figure 41: Pericentre distance r_p and injection time of the clouds in RunA (top panel) and RunB (bottom panel). Each individual cloud is represented as a black dot (numbered in ascending order). The small panels to the right show the r_p cumulative distribution of the injected clouds.

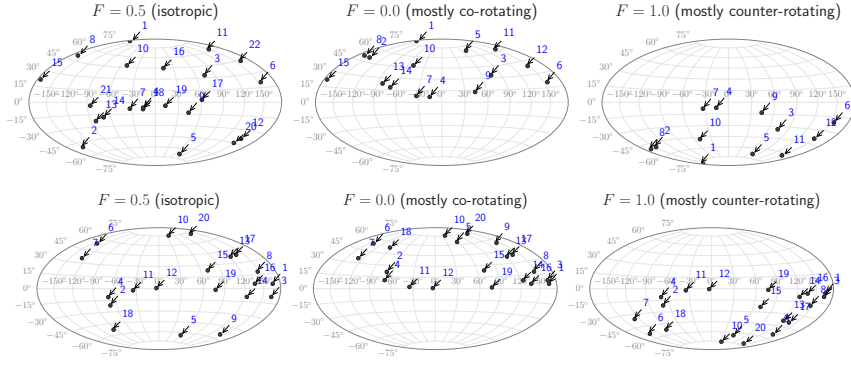


Figure 42: Orientation of the angular momentum of each cloud on the sphere for the three F distributions. In this representation \vec{L}_{bin} points to the north pole. Each panel reports only the clouds that were integrated in that specific run, numbered by injection order. The top row is for RunA and the bottom row for RunB.

Cloud	RunA		RunB	
	$t_i[P_0]$	$\Delta t_i[P_0]$	$t_i[P_0]$	$\Delta t_i[P_0]$
1	0.0	1.87	0.00	3.33
2	1.88	2.22	3.33	8.48
3	4.10	2.46	11.81	6.93
4	6.57	3.61	18.74	8.20
5	10.18	1.60	26.94	10.53
6	11.78	4.32	37.47	6.74
7	16.11	3.42	44.21	4.31
8	19.54	2.53	48.52	4.03
9	22.08	2.47	52.55	2.38
10	24.55	8.85	54.94	13.28
11	33.41	1.29	68.21	5.33
12	34.71	6.86	73.54	4.52
13	41.57	7.70	78.06	4.98
14	49.27	2.16	83.04	2.78
15	51.44	1.84	85.82	3.75
16	53.28	1.10	89.57	9.16
17	54.39	5.71	98.73	5.63
18	60.10	1.25	104.36	6.20
19	61.36	1.63	110.55	5.53
20	62.99	3.01	116.09	3.12
21	66.01	5.94	—	—
22	71.95	3.49	—	—

Table 2: Cloud injection times in RunA and RunB. For each cloud we report the time of injection (second and fourth columns) and the time to the next cloud injection (third and fifth columns). The two horizontal lines in column 2 and 3 identify the last cloud injection for RunA $F = 1.0$ (after cloud 12) and RunA $F = 0.0$ (after cloud 15), while RunA $F = 0.5$ reached cloud 22. Conversely all RunB configurations reached cloud 20.

conditions, defined by the fraction F of counter-rotating clouds interacting with the MBHB [Dotti et al., 2013]:

1. $F = 0.5$: \vec{L}_c are randomly distributed on the sphere. In this case, on average, 50% of the clouds will be co-rotating and 50% will be counter-rotating with respect to the MBHB. This is known as ‘chaotic’ accretion scenario and is visualised in the left panels of Fig. 42,
2. $F = 0.0$: all clouds are co-rotating with the MBHB, i.e. they all have $L_{c,z}$ aligned to \vec{L}_{bin} , as shown in the central panels of Fig. 42,
3. $F = 1.0$: all clouds are counter-rotating with the MBHB, i.e. they all have $L_{c,z}$ counter-aligned to \vec{L}_{bin} , as shown in the right panels of Fig. 42.

We first generate 30 random clouds ($F = 0.5$ case) and obtain the $F = 0.0$ and $F = 1.0$ cases by simply ‘mirroring’ \vec{L}_c with respect to the equatorial plane, as shown in Fig. 42. The mirroring procedure is crucial to single out the effect of co- and counter-rotation both on the formation of gaseous structures and on the evolution of the binary, because it allows us to consider systems that, besides the flipping of \vec{L}_c , are otherwise identical. Note that once \vec{L}_c and r_p are specified, one still has the freedom to rotate the orbit of the cloud within its orbital plane. To define the orientation of the cloud orbit, we consider the intersection of its orbital plane with the x, y plane defined by our coordinate system, and we place r_p at an angle Θ randomly drawn in the range $[0, 2\pi]$. All the mathematical details of the generation of the initial conditions are given in Appendix 10.7.1.

Although we generate 30 clouds for each set of initial conditions, we only show initial conditions for 22 clouds for RunA, and 20 clouds for RunB in Fig. 41 and 42. Due to time constraints and necessary maintenance of the computer clusters employed for the calculation, we were in fact only able to integrate RunA $F = 0.0$ up to cloud 15, RunA $F = 0.5$ up to cloud 22 and RunA $F = 1.0$ up to cloud 12. All RunB were integrated up to cloud 20. The full information about the initial condition of each cloud, including initial positions and velocities, are given in Appendix 10.7.1.

In summary, we generated two sets of runs, RunA and RunB defined by different cloud impact parameters and injection times. For each of the runs we considered three angular momenta distributions, $F = 0.0$, $F = 0.5$ and $F = 1.0$, for a total of six different sets of initial conditions. Each of the sets is integrated at four single cloud resolutions: 50k, 200k, 500k and 1m. In the following, we will concentrate on the results of the 50k simulations, which reached the larger number of clouds in the system. Runs at higher resolutions are obviously slower; for example, only 4-5 clouds are generally injected in the 1m case. Higher resolution runs are used as benchmark for comparison and to assess convergence of the simulations.

To study the relaxation of the system after the infall of several clouds, we also ‘forked’ each of the 50k runs after the injection of 5 and 10 clouds. In practice, we ran in parallel two additional sets of simulations in which the system was allowed to evolve unperturbed after 5 and 10 clouds interacted with the binary, to study the long term properties of the relaxed system.

Commonly for this type of numerical investigations, a large computational infrastructure was needed to handle the required set of runs and tests.

The simulations presented in this work feature complex dynamics of multiple clouds interacting with a MBHB, it is therefore important to test the incidence of our main numerical assumptions on the evolution of the system, to keep the impact of spurious numerics under control. The tree reconstruction frequency has been tuned to optimise angular momentum conservation, as described in §10.2.2, and softening has been chosen to guarantee a proper resolution of the gravitational interaction between particles. The other numerical ‘degrees of freedom’ are the choice of r_{sink} and of the number of particles used to simulate each cloud, N_c (i.e. the ‘resolution of the simulation’). In this section we check the robustness of our set-up against our choice of these parameters.

10.4.1 r_{sink} value and accretion convergence

Ideally, a particle will be accreted when it approaches the M_{BH} at about $r = 6GM/c^2 = 3R_S$, where R_S is the Schwarzschild radius. If we consider our fiducial system, this distance is $\approx 5 \times 10^{11}$ cm, equivalent to 2×10^{-6} in code units. It is clear that a realistic condition for particle accretion is beyond any feasible resolution in our numerical scheme. Thus a fictitious sink radius is introduced by hand, as explained in §10.2.2. The numerical value of r_{sink} is set *ad-hoc* for numerical convenience. To test its impact on the dynamics of the system we ran four otherwise identical simulations with $r_{\text{sink}} = 0.1$ (standard model), 0.05, 0.2 and 0.5 in code units thus spanning an order of magnitude. For these test simulations we considered RunA $F = 0.0$, with 200k resolution.

Fig. 43 shows the evolution of the key parameters describing the evolution of the MBHB in the four runs, evolved for about 25 initial binary orbital periods, sufficient to follow the strong dynamical interaction with eight subsequent clouds. Results match so well across the runs that we had to offset the lines, otherwise they would overlap almost perfectly. The value of r_{sink} does not appreciably impact any of the MBHB parameters, not even the eccentricity evolution, which depends on a fine balance between energy and angular momentum exchanges, and is therefore sensitive to minor fluctuations in the dynamics. Note that the pool of interacting clouds span a large dynamical range, including clouds with $r_p < a_0$ (clouds 5 and 8, see Fig. 42) whose dynamics might in principle be severely affected by an improper treatment of r_{sink} .

Critically, the evolution of the two M_{BH} masses is independent on r_{sink} , which indicates that gas accretion is not affected by its unphysically large value. This is because of the conditions spelled in §10.2.2, whereby particles are required to be *bound* to the M_{BH} for being accreted. In practice only particles that settle into orbits enclosed in the M_{BH} Roche Lobe can be accreted. These particles form eccentric (either transient or persistent) mini-disks that are continuously perturbed by infalling material and are swiftly drained into the sink. Therefore, setting a smaller sink radius only causes a small delay in the time at which a particle is recorded as accreted. Note that this does not mean that all particles crossing r_{sink} will be accreted in reality, as we will discuss in the next section.

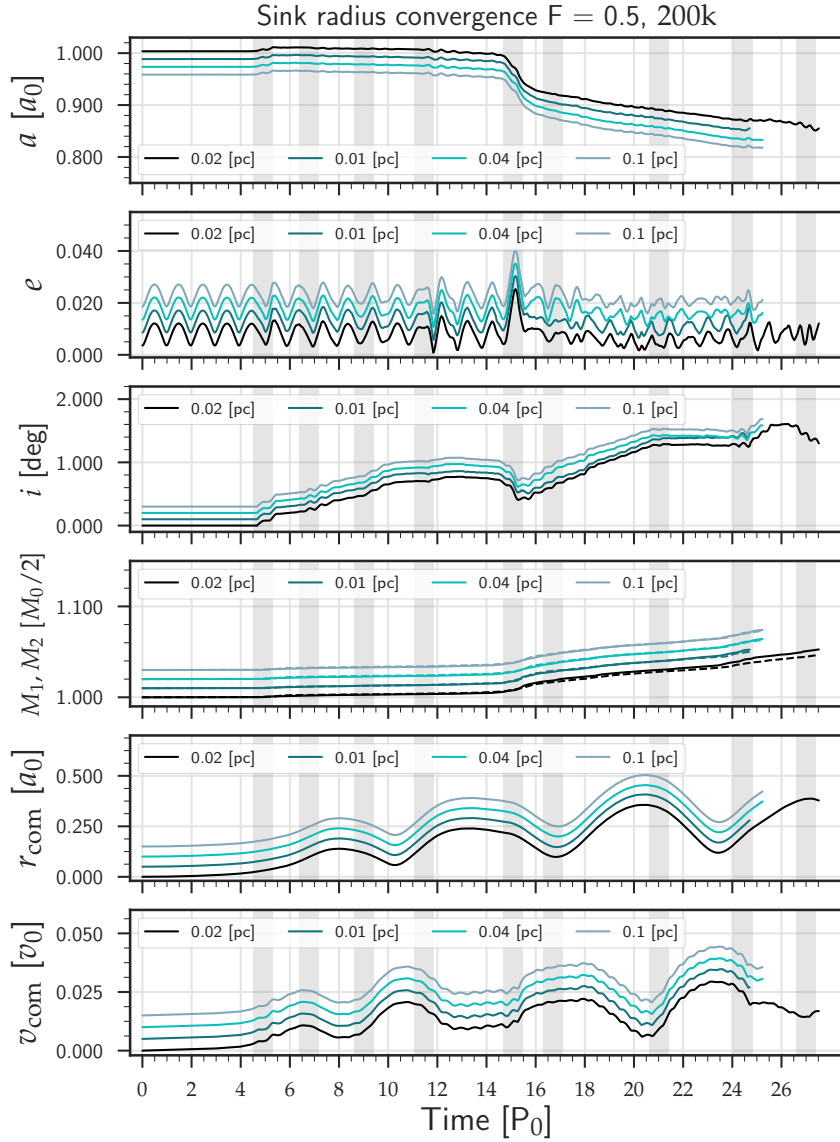


Figure 43: Time evolution of the relevant MBHB parameters in RunA $F = 0.0$, with 200k, for different values of r_{sink} , displayed in simulation units. In each plot, the black line is for $r_{\text{sink}} = 0.1a_0$ (the adopted default values), while different shades of blue are for $r_{\text{sink}} = 0.05a_0, 0.2a_0, 0.5a_0$. Note that lines have been progressively shifted upwards for clarity, since they would otherwise almost perfectly overlap. Grey vertical stripes indicate the ‘arrival time’ of each new cloud, i.e. the time of first periastron passage in its orbit around the MBHB. In the mass panel (fourth from the top), solid and dashed lines represent M_1 and M_2 respectively, normalised to their respective initial values.

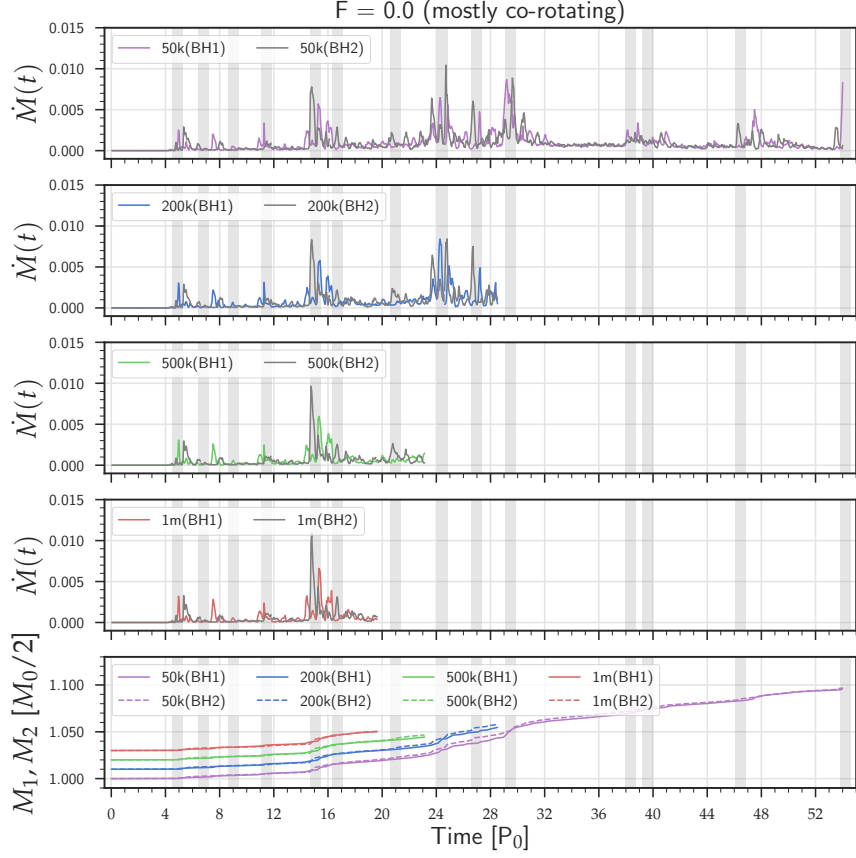


Figure 44: Accretion rate for RunA $F = 0.0$ for all the investigated resolutions: 50k, 200k, 500k and 1m, from top to bottom. The two lines on each panel represent accretion onto each M_{BH} . Orange vertical lines mark the moment each cloud is injected into the system, and grey vertical stripes the time of the first close interaction with the MBHB. Rates are displayed in simulation units $[M_0/P_0]$. The bottom panel shows the mass evolution of each MBH at all resolutions. Lines have been progressively shifted upwards for clarity.

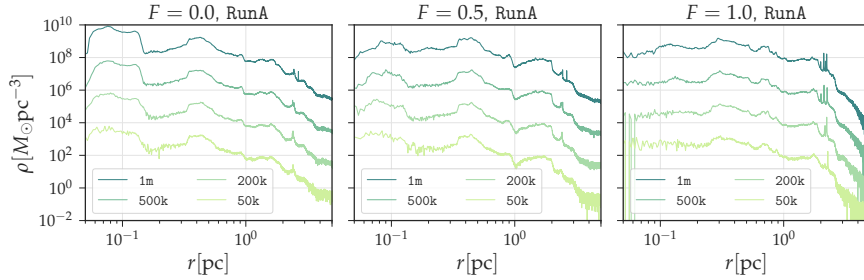


Figure 45: Angle-averaged density profile at $T = 10.2[P_0]$. In each panel, from bottom to top, lines corresponds to resolution 50k, 200k, 500k and 1m, and have been shifted upwards to ease comparison. The three different F -distributions are represented from left to right, as indicated in figure. The y -axis normalisation of the 50k resolution is fixed to the scale of our fiducial system, $M_{\text{bin}} = 10^6 M_\odot$.

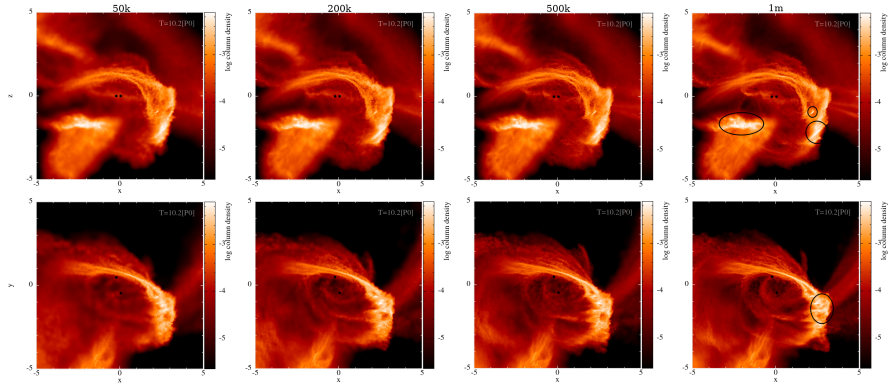


Figure 46: Column density rendering of RunA, $F = 1.0$ at $T = 10.2[P_0]$. Each column represents a cloud resolution, 50k, 200k, 500k and 1m from left to right. In each column, the top and the bottom panels represent views in the $x - z$ and $x - y$ plane respectively.

10.4.2 Robustness against N_c

It is obvious that the level of dynamical detail that traceable in a numerical simulation critically depends on its resolution, and the SPH technique is no exception to this rule. Simulations shall be performed with a number of particles sufficient to resolve the physical features of interest in an N -independent fashion. Of particular interest for this work are accretion onto the MBHB and the distribution of non accreted gas around the binary. To test that our simulations are ‘well behaved’ we ran each of them at the four particle resolutions 50k, 200k, 500k and 1m particles, thus spanning a range of 20.

In Fig. 44, we show the accretion rate on each individual M_{BH} in the case RunA $F = 0.0$, for all values of N_c . Accretion is highly variable, showing prominent intermittent spikes in correspondence to the arrival times of clouds with small impact parameters. Note that, at peak, $dM/dt > 0.005P_0^{-1}$ on each individual M_{BH} . Converted to our fiducial system this is about $0.6M_{\odot}\text{yr}^{-1} \approx 60\dot{M}_{\text{Edd}}$. The actual fate of the gas during these high accretion episodes is unclear. Photon trapping might allow gas to be accreted at super-Eddington rates [e.g. Abramowicz et al., 1988, Ohsuga et al., 2005], or alternatively, radiation pressure might cause the expulsion of the majority of the gas in powerful winds, as observed in [e.g. Tombesi et al., 2010, 2015]. As discussed in Goicovic et al. [2017], the fate of the gas in itself has only a minor impact on the dynamical evolution of the system, which is driven more by gas *capture* from the MBHB, rather than gas *accretion*. Winds can, however, strongly interact with the surrounding infalling clouds, affecting their dynamics. We caution that this effect is not captured in our simulations. Nevertheless, Fig. 44 shows that, despite minor differences in the definition of the accretion peaks, neither the accretion rate nor the total mass growth of the M_{BH} s have an appreciable dependence on the amount of particles used in the simulation.

To test the dependence of the gas distribution on N_c , we show in Fig. 45 the angle-averaged gas density as a function of radius for all RunA at $T = 10.2P_0$, after the disruption of the third cloud. The density is displayed in physical units, by scaling the results to our fiducial system ($M_{\text{bin}} = 10^6 M_{\odot}$). Density profiles are equivalent at all resolutions, even though they are noisier in the 50k and 200k runs, due to the smaller N_c . Increasing the number of particles

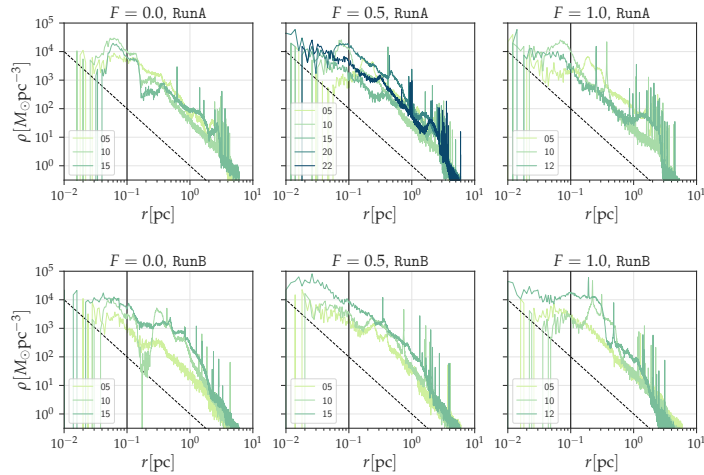


Figure 47: Angle-averaged mass density profile for all 50k runs (as indicated in each panel) at different snapshots, in physical units, scaled to our fiducial system. Curves are labelled by the number of clouds present in the system, and represent the density profile at the last recorded snapshot before the subsequent cloud is injected in the system (e.g.: 05 represents the status of the system at the last snapshot before adding the 6th cloud). In each panel, the dashed line refers to a $\rho \propto r^{-2}$ profile, for comparison. The solid vertical black line shows the initial position of each of the MBHB components.

allows to better capture fine details in the gas distribution. For example in the $F = 0.0$ case (left panel), the sharp density drop around 0.15pc is better resolved in the 1m run. Likewise, in the $F = 1.0$ case (right panel) the larger amount of particles in run 500k and 1m allows the resolution of two density peaks around 2.2pc, which are blended into a single peak in runs 50k and 200k. Nevertheless, differences are minor, and the overall structure of the gas distribution is preserved across resolutions.

A visualisation of the 3-D particle distribution is shown in Fig. 46, where snapshots of the $F = 1.0$ case at $T = 10.2P_0$ are shown in the $x - y$ and $x - z$ plane. The overall gas distribution is exquisitely consistent at all resolutions, even though structures appear slightly blurrier going from the right to the left. The aforementioned difference in the density peaks at 2.2pc, is due to the regions highlighted with black circles on the rightmost column. Looking at this critical areas, and moving left in the panel sequence (going down in resolution), the dense areas become less defined eventually blending the fine structures in larger clumps. These differences, however, do not affect the overall dynamical evolution of the systems and are relevant only in the determination of the statistics of dense clumps prone to star formation. Since this specific investigation is beyond the scope of the current work, we deem the 50k runs sufficiently accurate for our purposes.

10.5 RESULTS

In this section we present the main results of the simulations concerning the formation and evolution of gaseous structures around the MBHB ⁴. In a complementary paper (Goicovic et al. [2018]) to this chapter, we focus on the

⁴ Animations of all the simulations can be seen in <http://multipleclouds.xyz/movies/>

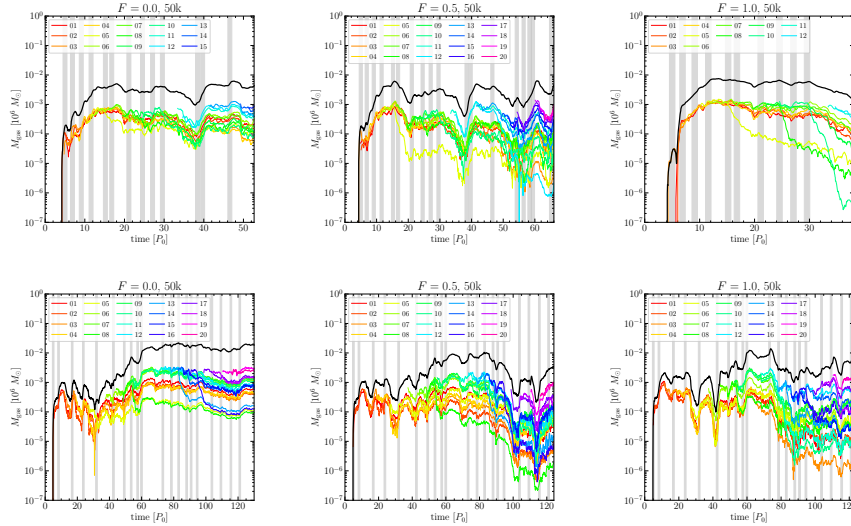


Figure 48: Mass evolution of the circumbinary structure for the models $F = 0.0$ (left), $F = 0.5$ (centre) and $F = 1.0$ (right). The first row represents RunA while the bottom row represents RunB. The different coloured lines are the mass contribution from each infalling cloud while the black line is the sum of the contribution from all clouds

evolution of the MBHB. We will describe the general outcome of both RunA and RunB, in order to make comparisons between the two. When a specific aspect is investigated in more depth, we consider RunA as default case. We also ran an extra suite of simulations, that we call RunC, mixing the cloud configuration of RunA and the time distribution of RunB. Some relevant results for this extra suite of runs is presented in Appendix 10.7.2. In the figures of this Section, relevant quantities are displayed in physical units, scaled to our fiducial system (i.e. a MBHB with initial mass $M_{\text{bin}} = 10^6 M_{\odot}$ and initial separation $a = 0.2\text{pc}$), whereas run snapshots are shown in simulation units ($M_{\text{bin}} = a = 1$).

10.5.1 Overall evolution of the gas distribution

We first look at the evolution of the (angle averaged) gas density profile as a function of distance to the MBHB centre of mass. This is shown Fig. 47 as runs advance and more clouds are added to the system. As a general trend, we see that in all cases the gas density outside the binary orbit tends to follow a $\rho \propto r^{-2}$ distribution, even though the profile is necessarily time dependent and there is a large scatter. In the long run, this has to be expected; by throwing gas at the binary from all directions, the resulting envelope will be almost at rest with respect to the binary centre of mass. One can then estimate the Bondi radius as $R_B = 2GM/c_s^2$. For our fiducial system, the gas temperature is $T = 100\text{K}$, resulting in $c_s \approx 2\text{km s}^{-1}$. This means that the Bondi radius is several tens of parsecs i.e. much bigger than the domain of the simulation. We are hence injecting within the Bondi radius of the binary gas at a roughly constant rate, which is therefore expected to settle into an r^{-2} density profile.

The ‘‘humps’’ – observed for example in the middle top-panel on the 22th cloud line around 5pc, or bottom-right panel at 0.5pc and 1pc for the 10th

and 15th cloud lines, respectively – correspond to new clouds infalling into the system, and are not properties generated by bound material that is falling back to the binary after cloud disruption. Each new infalling cloud will generate a “hump” that will move from right to left until is disrupted by the MBHB. Conversely, the sharp lines accumulating at $R > 1\text{pc}$ are due to high density clumps formed during the phase of cloud compression in the interaction with the MBHB and then ejected at large distances.

The figure also highlights few interesting features specific to each of the F -distributions. On the $F = 0.0$ panels (left column) we can observe a clear dip in the density profile around 0.2-0.3pc, comparable to the binary orbital separation, building up over time as more clouds get into the system. This is a clear indication that the action of the MBHB is carving a cavity in the gas distribution. In fact in the $F = 0.0$ case, clouds are mostly co-rotating and we expect a lot of gas will re-arrange in a co-rotating circumbinary structure. Lindblad resonances are then expected to carve a hole in the gas distribution of a size of $\approx 2a$ [Artymowicz and Lubow, 1994]. The large overdensities at the M_{BH} location are instead indicative of prominent mini-discs, that are also expected to form in the co-rotating case [Goicovic et al., 2016]. Note, moreover, how the density just outside the cavity in RunB builds up with time much more prominently than in RunA, which is a sign that a more massive circumbinary disc is being built in the former case (as we will see below, cf Fig. 48 and Fig. 50). Conversely, clouds are mostly counter-rotating in the $F = 1.0$ case (right panels), and Lindblad resonances do not operate. No steady dip is in fact observed in this case, however prominent transient overdensities can be seen forming around 0.2pc, which are indicative of the formation of compact dense rings that get disrupted in the interaction with new incoming clouds (again, see example in Fig. 50). The isotropic nature of the $F = 0.5$ runs can be also appreciated (central panels), which shows a relatively smooth and steady shape. In the following, we will examine in detail the evolution of these gas structures, paying particular attention to the formation of mini-discs, circumbinary discs and rings.

10.5.2 Circumbinary structures

Following the infall of each cloud, some of the non-accreted material will remain bound to the binary, forming structures around it, as well as around each M_{BH} . Previous work by Goicovic et al. [2016] studied the impact of the orbital configuration of the infalling cloud on the formation of such structures, showing a rich phenomenology depending on the initial orbital inclination relative to the MBHB plane and impact parameter. In particular they found that the binary is generally unable to change significantly the orientation of the gas, which produces discs that follow the initial cloud’s inclination.

For the simulations presented in this chapter, the incoherence of the accretion events produces a variety of gaseous streams that continue interacting with the binary and between each other, making a clear identification of circumbinary disc structures much harder. To study any type of stable structure around the MBHB is essential to devise a set of conditions defining whether an SPH particle belongs or not to that structure. In practice we define circumbinary structures by considering all particles that:

- are bound to the MBHB (to avoid including unbound streams of gas flung away in the interaction with the MBHB);

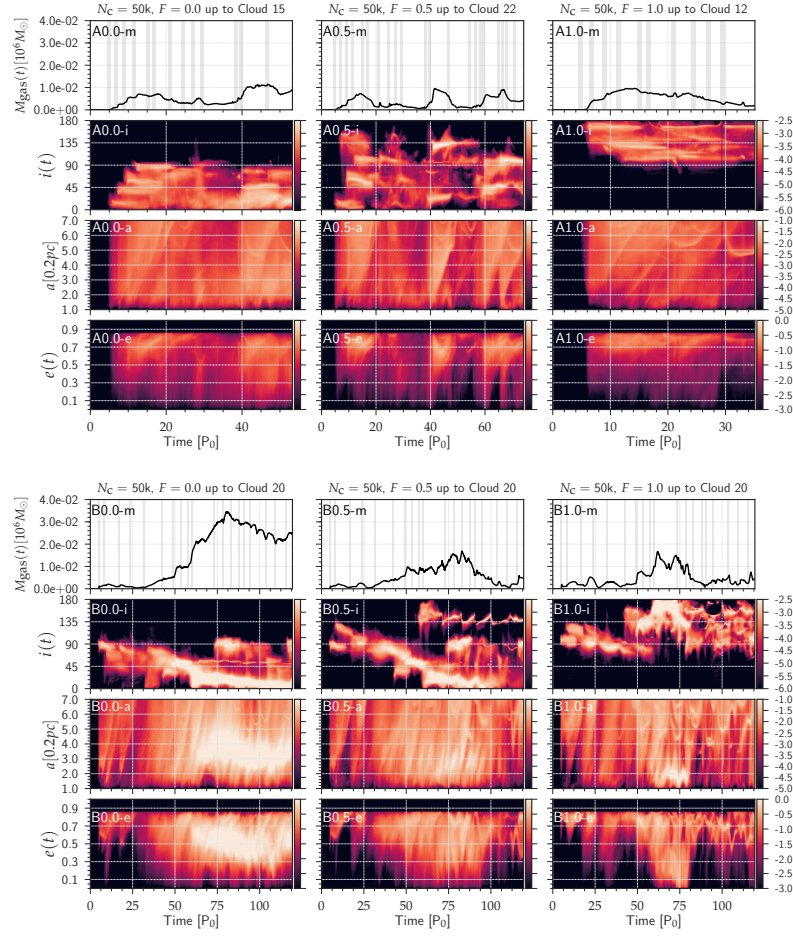


Figure 49: Time evolution of the circumbinary structure in each of the simulations. The three columns, left to right, are for the $F = 0.0$, $F = 0.5$ and $F = 1.0$ scenarios; the top 3×4 grid of plots is for RunA and the bottom for RunB. In each grid, the top row shows the time evolution of the total mass of the structure, M_{gas} . All the other rows show in log-scale the time evolution of selected particle distributions, normalised so that the integral of any vertical slice of any panel gives the total mass M_{gas} at that specific time. With such normalisation, the second row represents dM_{gas}/di , the third row dM_{gas}/da (a is displayed in units of initial binary separation) and the fourth row dM_{gas}/de . The colour gradient is displayed in log-scale, as indicated by the bars at the far right of each row. Grey lines on the M_{gas} panels, indicate the time of arrival of infalling clouds.

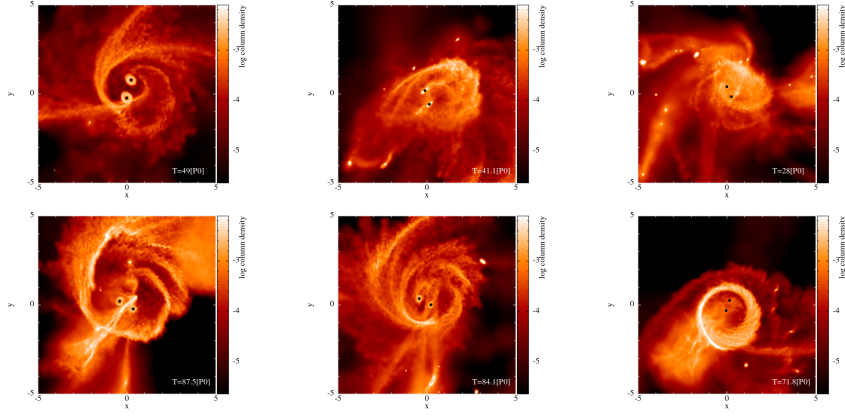


Figure 50: Snapshots of RunA (top row) and RunB (bottom row) at selected time (as indicated in each panel), highlighting features described in Fig. 49. In all panels, the binary is viewed face-on in the $x - y$ plane, and the gas column density (integrated along the z -axis) is shown in log-scale as indicated on the right of each panel.

- are inside a critical radius $r < 7a_0$. The specific choice is arbitrary, but is motivated by the typical size of circumbinary discs found in Goicovic et al. [2016].

For each SPH particles identified by this first cut, we can take 3-D position and velocity, replace the MBHB with a point mass centred in the centre of mass of the system, and compute a Keplerian orbit around this point. This calculation is an approximation of the true orbit of the particle, since it ignores the binary nature of the central massive object, as well as the external potential. The advantage of this approach is that it allows to define orbital elements such as semimajor axis a , eccentricity e and inclination i for each particle. We therefore impose the further condition that the particle orbit

- has $a(1 - e) > 1$, to make sure that it does not intersect the MBHB;
- has $e < 0.9$. Although this is also a somewhat arbitrary choice, it allows to exclude particles belonging to new infalling clouds (which are on almost radial orbits).

Fig. 48 shows amount of gas that meets the aforementioned conditions as a function of time, together with the breakdown of the contribution of each individual cloud. Despite stochasticity and differences between the runs, we identify some general trends. As new clouds get into the system, their contribution to the circumbinary structure initially rumps up to $M \approx 10^{-3} M_{\text{bin}}$ (i.e. 10% of the initial mass cloud), since material settles on orbits with $a < 7a_0$ and circularises to $e < 0.9$. This process is evident for the initial clouds in RunA which have fairly large periapsis passage and tend to not interact with each other (cf the top panel of Fig, 41). Up to $T \approx 15P_0$, the circumbinary structure mass tends to steadily grow in time to about 1% of the MBHB mass. Eventually, incoming new clouds on intersecting orbits prompts accretion of pre-existing circumbinary gas, preventing the structure to significantly further grow in mass. This is why the total mass in these structures never grows to much more than few% of the MBHB mass. This process is more evident right from the start in RunB. In this case, the first clouds have rather small periapsis and they strongly interact with each other upon arrival onto the MBHB.

The net effect is that the circumbinary structure does not grow much beyond $M \approx 10^{-3} M_{\text{bin}}$ until a number of clouds with larger periapsis contribute a substantial mass budget from $T \approx 40P_0$ onwards (cf the bottom panel of Fig. 41).

It is clear that the contribution of each individual cloud to the circumbinary structure tend to decrease in time (although with large fluctuation). With “older” clouds contributing less to the mass budget. This causes the large rainbow-like spread towards the end of the runs, whereby the contribution of each individual cloud ranges between 10% to about 0.01% of their initial mass. It is interesting to note that the spread is much larger in the $F = 0.5$ and $F = 1.0$ cases, in which clouds experience more violent interactions among themselves and with the binary. In the $F = 0.0$, conversely, most of the circumbinary material is stored into a disc held-up by Lindblad resonances. The income of new clouds in mostly co-rotating orbit tends to add new material to the disc without a substantial disruption of the pre-existing conditions. For example, the first incoming cloud (red curve) still contributes about 5% of its mass to the circumbinary disc by the end of the simulations in both RunA and RunB.

Although Fig. 48 quantifies the amount of material that forms a circumbinary structure, it does not provide much information about the nature of that structure. For example, the gas might be configured in a disc or in a cloud or in multiple rings and it would not make a difference. For each SPH particle belonging to the structure, we reconstructed the orbit and we computed a , e and i . We can now construct the distributions of these quantities, and follow their evolution in time.

This is shown in Fig. 49 for all our runs, together with a replication of the total circumbinary gas mass shown in Fig. 48 (top rows). Each panel is built as follows. We take a uniform grid in the desired quantity and, at each simulation snapshot, we construct an histogram by adding particles to the bins and normalising so that the integral over the bins gives the total mass in the structure. Histograms at subsequent steps are then concatenated along the x -axis and smoothed to produce the 2-D density maps displayed in the figure (in logarithmic scale). Fig. 49 clearly show the connection between the F -distribution and the geometry of the circumbinary structures for both RunA and RunB. We now describe the diverse phenomenology of each F -distribution, referring to representative examples of each individual case shown in Fig. 50.

For the co-rotating case ($F = 0.0$, left column), after an initial transient phase, most of the particles have less than 50° respect to the binary, and tend to distribute in a co-rotating, extended circumbinary disc, displaying a variety of eccentricities. In RunA, the disc does not build up as a coherent structure, as hinted by the low total mass (panel A0.0-m) and large range of inclinations (panel A0.0-i). The presence of large amount of gas at high inclinations, prevents resonances from being efficient, and a well defined cavity cannot be seen in panel A0.0-a, where we see gas been distributed in the whole range $1 < a < 7$, although a concentration of gas at $a < 5$ appears from $T \approx 40P_0$ onwards. Conversely, in RunB, a prominent thin disc builds up coherently and progressively aligns with the MBHB. This is demonstrated by the mass build-up up to about $M_{\text{gas}} = 0.03M$ (panel B0.0-m) and by the narrow range of inclinations decreasing with time (panel B0.0-i). The density contrast in panel B0.0-a highlights that the bulk of the disc lies in the range $2 < a < 4$, with the decline at $a < 2$ indicative of the resonance-sustained cavity. Gas

within the disc remains in fairly eccentric orbits spread around $e = 0.5$ (panel B0.0-e).

The left column of Fig. 50 displays representative snapshots of these two simulations. The top panel is taken when the 13th cloud of RunA $F = 0.0$ interacts with the system, at $T \approx 49P_0$. An extended circumbinary structure, almost in the binary orbital plane ($i \approx 20^\circ$) is clearly visible, but it is about to be partially disrupted by the 13th cloud streaming-in from the left (which will cause the subsequent drop in M_{gas} observed in panel A0.0-m). The bottom panel shows the status of RunB $F = 0.0$ at the peak of M_{gas} at $T \approx 95P_0$. Comparisons between the two highlights the prominence of the circumbinary disc forming in RunB, featuring a massive spiral that extends to $a \approx 5$,

Not surprisingly, in the counter-rotating case ($F = 1.0$, right column), the gas distribution around the binary has high inclinations. Also in this case, the two runs display quite different behaviours. In RunA, we see the formation of a persistent structure. The mass gets to $M_{\text{gas}} \approx 10^{-2}M$ at the peak, declining to $M_{\text{gas}} \approx 10^{-3}M$ by the end of the simulation (panel A1.0-m), is spread across the whole a range (panel A1.0-a), and displays a wide range of inclinations (panel A1.0-i). The gas essentially configures into a tenuous and low mass counter-rotating cloud. This is clearly shown in the top right panel of Fig. 50 taken at $T = 28P_0$; no clear disc-like structure is present, and the binary is surrounded by a relatively compact cloudy envelope with streams extending to $a > 5$.

The situation is strikingly different in RunB. Here we see that the violent interaction with the binary causes most of the gas to be either accreted or expelled and no steady circumbinary structure is formed. A prominent eccentric transient ring forms around $T = 60P_0$, being disrupted about 20 periods later (note that in our fiducial system the ring would last for about 0.2 Myr), as clear from panel B1.0-m. The structure is almost perfectly counter-aligned to the binary (panel B1.0-i), is confined within $a < 2.5$ (panel B1.0-a) and has an average eccentricity $e \approx 0.3$ (panel B1.0-e). The ring is clearly visualised in the bottom right panel of Fig. 50.

The $F = 0.5$ simulations, shown in the central column of Fig. 49, highlight the impact of the specific cloud distribution on the formation of circumbinary structures. RunA features a number of transient, incoherent structures extending at all radii (panel A0.5-a) and inclinations (panel A0.5-i). The top central panel of Fig. 50 shows the system at $T = 41P_0$. Cloud 12 and 13 are bringing fresh gas into the system, causing a temporary enhancement of the mass in the circumbinary structure, that looks like an incoherent, extended envelope. Conversely, RunB shows a behaviour similar to the $F = 0.0$ case; at $T \approx 60P_0$, a co-rotating circumbinary structure forms, progressively aligning with the MBHB as it gains mass in subsequent interactions (panels B0.5-m and B0.5-i). The structure persists for about 30 binary orbits (≈ 0.3 Myr, when scaled to our fiducial system), before being completely disrupted by the arrival of new clouds on small pericentre orbits (clouds 13-to-16, cf Fig. 41). A snapshot of the system at $T = 89P_0$, shown in the bottom central panel of Fig. 50, highlights the similarity with the $F = 0.0$ run; in this case, the circumbinary spiral is less massive and slightly more compact.

Interestingly, in both RunA and RunB, there is the tendency to form more co-rotating than counter-rotating structures. This can be clearly seen in panels A0.5-i and B0.5-i of Fig. 49, where the particles tend to cluster below 90 degrees in inclination even though half of the events come from the southern hemisphere. This occurs simply because resonances with the MBHB orbital motion efficiently transfer angular momentum to co-rotating gas, which can

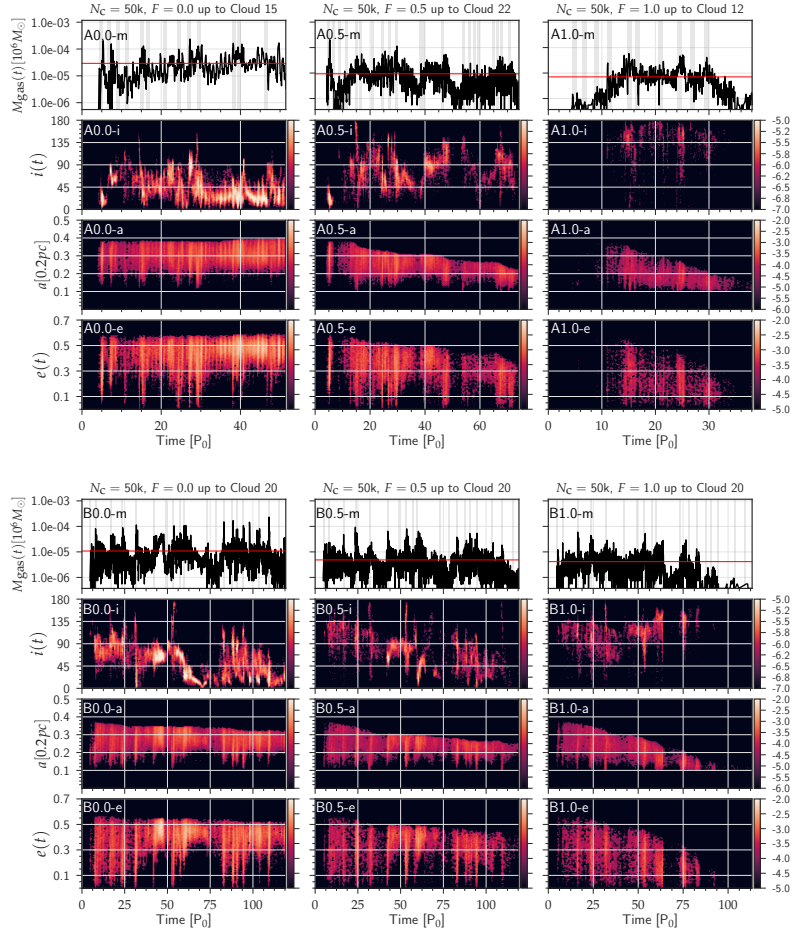


Figure 51: Same as Fig. 49 but for the mini-disc of $M_{\text{BH}1}$ in the $F = 0.0$, $F = 0.5$ and $F = 1.0$ cases (left to right), for RunA (top 3×4 grid) and RunB (bottom grid). Each panel is labelled as in Fig. 49, note the log-scale in the top rows showing the total mass in the mini-discs, M_{gas} .

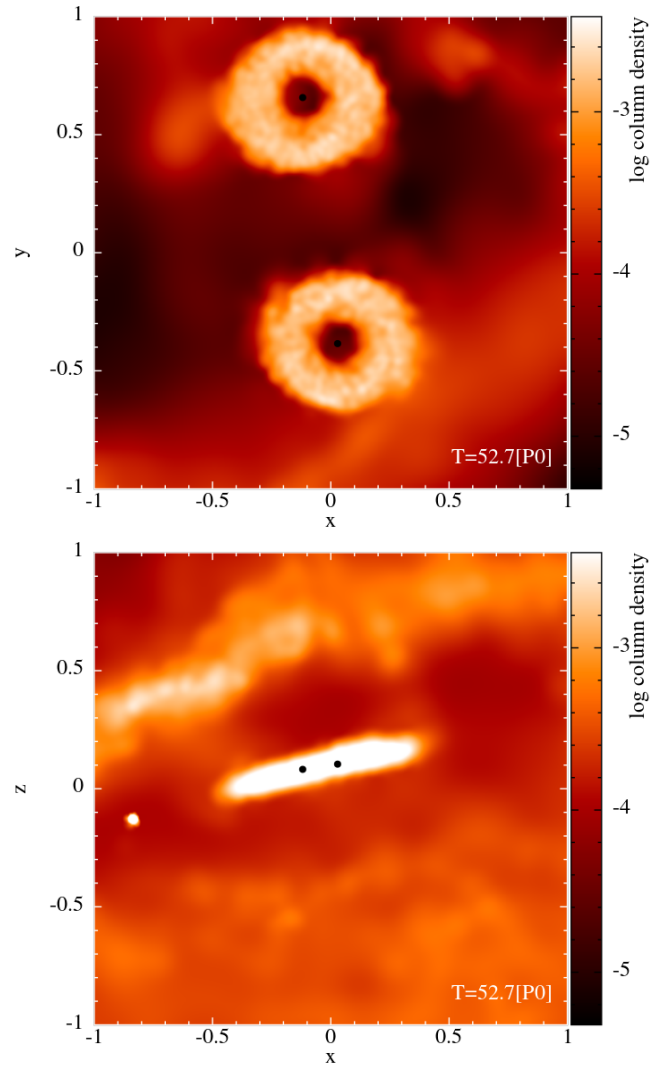


Figure 52: Snapshot of the last stage of RunA $F = 0.0$ (shown in Fig. 51, first column of the top grid). The top and bottom panels show the $x - y$ and $x - z$ views respectively.

Simulation	Distribution	$M_{MD1} [10^6 M_\odot]$	$M_{MD2} [10^6 M_\odot]$
RunA	$F = 0.0$	3.58×10^{-5}	3.39×10^{-5}
	$F = 0.5$	6.64×10^{-6}	6.82×10^{-6}
	$F = 1.0$	1.17×10^{-6}	9.93×10^{-7}
RunB	$F = 0.0$	1.08×10^{-5}	1.07×10^{-5}
	$F = 0.5$	4.84×10^{-6}	5.15×10^{-6}
	$F = 1.0$	4.15×10^{-6}	4.32×10^{-6}

Table 3: Average mass of the mini-discs surrounding the M_{BH} s in all our runs.

therefore settle into more extended and stable structures. An effect that is absent for the counter-rotating material.

10.5.3 Mini-discs

It is also interesting to study the dynamics of gaseous structures *inside* the MBHB corotation radius, which are generally directly responsible for the binary feeding and the associated high energy electromagnetic radiation. In the standard picture of a steady circumbinary disc, the forcing imposed by the binary quadrupolar potential induces gas streams that feed prominent mini-discs around each M_{BH} [D’Orazio et al., 2013a]. Conversely, in the counter-rotating case, the absence of Lindblad resonances allows gas at the edge of a putative circumbinary disc to impact directly onto the binary. Since the M_{BH} and the gas have opposite directions, $|\vec{v}_{\text{BH}} - \vec{v}_{\text{g}}| \approx 2v_{\text{BH}}$; this means that the capture cross section of the gas is very small, and only gas *inside* the sink radius becomes bound to the M_{BH} being promptly accreted. In practice, mini-discs extending beyond r_{sink} are unlikely to form in this case.

Similarly to §10.5.2, we now select all particles inside either of the M_{BH} Roche Lobes, defined as circles around each MBH with radius

$$R_{\text{RL}} = \frac{0.49 q^{2/3}}{0.6 q^{2/3} + \log(1 + q^{1/3})} a \quad (64)$$

[Eggleton, 1983], where q is the binary mass ratio, which we assume to be unity throughout all of our simulations. We then use 3-D positions and velocities to compute their Keplerian orbit around the closest M_{BH} . We define the mini-disc around each M_{BH} as the collection of gas particles on orbits that:

- are bound to that M_{BH} ;
- have a pericentre larger than the sink radius, i.e.: $a(1 - e) > r_{\text{sink}} = 0.1$ (equivalent to 0.02pc for our fiducial system);
- have an apocentre smaller than the Roche Lobe size, i.e.: $a(1 + e) < R_{\text{RL}}$.

The evolution of the mini-disc around $M_{\text{BH}1}$ in each run is summarised in Fig. 51. There we show the time evolution of the mini-disc total mass (top row), and the distribution of inclination (second row), semimajor axis (third row) and eccentricity (bottom row) of the particles belonging to it, constructed as detailed in §10.5.2. The first thing we notice is that, due to the complex dynamics triggered by the infall of incoherent gas clouds, mini-discs are far from being well defined, stable structures in our simulations. However, some general features of relaxed ‘steady state simulations’ are preserved.

In particular, we see that the prominence and persistence of mini-discs is a strong function of F . This is also supported by numbers reported in Table 3, demonstrating that the average mass in the mini-discs is much higher in the $F = 0.0$ runs. Note that this is particularly true for RunA: here the frequent supply of fresh clouds has the effect of feeding the mini-discs in the $F = 0.0$ case, whereas increases the chance of disrupting them quickly in the less coherent $F = 0.5$ or in the counter-rotating $F = 1.0$ cases. In RunB, clouds are supplied at a lower rate, leaving more time for the gas in the mini-discs to be accreted. This mostly affects the $F = 0.0$ run, in which we see a smaller amount of mass accumulating in the mini-discs on average.

In general Fig. 51 show that mini-discs are more massive and persistent in the $F = 0.0$ runs (left column). Note that both in RunA and RunB there is a significant scatter in the particle inclination distribution. This is because particles partially preserve memory of the inclination of their parent cloud and tend to form mini-discs aligned with their incoming orbital angular momentum. In the long run, however, the MBHB potential torques the disks causing a partial alignment with its orbital angular momentum (effect visible both in panels A0.0-i and B0.0-i). An example of such persistent mini-discs forming in the co-rotating case is shown in Fig. 52, displaying the last snapshot of RunA $F = 0.0$.

As expected, mini-discs becomes much more intermittent as F increases. This is essentially because mini-disc formation is a more natural outcome in the interaction with co-rotating clouds. In fact, we can see that in the $F = 0.5$ case (central column) the formation of structures with $i < 90^\circ$ is strongly preferred (panels A0.5-i and B0.5-i). In the extreme $F = 1.0$ case, mini-discs are very intermittent and, contrary to the circumbinary structures, often show a significant fraction of co-rotating material (i.e. with $i < 90^\circ$, see panel B1.0-i). The shrinking of the binary as the simulations advance, is clearly noticeable in the a and e panels of all the $F = 0.5$ and $F = 1.0$ simulations. In fact, as the two M_{BH} s get closer to each other, their Roche Lobes contract and the size of the mini-discs that they can accommodate shrink accordingly.

Finally, we note that the study of mini-discs is a delicate matter in this kind of simulations, due to the relatively large sink radius and to the inherent small size of these structures which makes their resolution difficult. In fact, even the most prominent mini-discs in the $F = 0.0$ simulations have an average mass of about 10^{-5} , which means that in the 50k simulations they are resolved with about 50 particles. We therefore checked convergence of our results by comparing mini-discs at 50k and 500k resolutions finding essentially no difference neither in the intermittent behaviour nor in the average masses, indicating that our results are robust. Likewise, we tested that shrinking the sink radius by a factor of two does not appreciably affect the mini-discs evolution.

10.5.4 ‘Forked’ simulations: stopping the supply of clouds

The variability of a system being constantly affected by multiple chaotic accretion events has clearly emerged in the previous sections, where we showed the continuous formation and disruption of structures as new clouds were added to the system. An interesting question is what happens when the supply of new clouds ceases and the system is allowed to relax. To answer this question, we took two snapshots from all the RunA simulations at the moment when the 5th and the 10th clouds were added to the system and ‘forked’ them, thus starting two parallel series of runs in which no further clouds were added, allowing the system to evolve unperturbed. We show in this section results

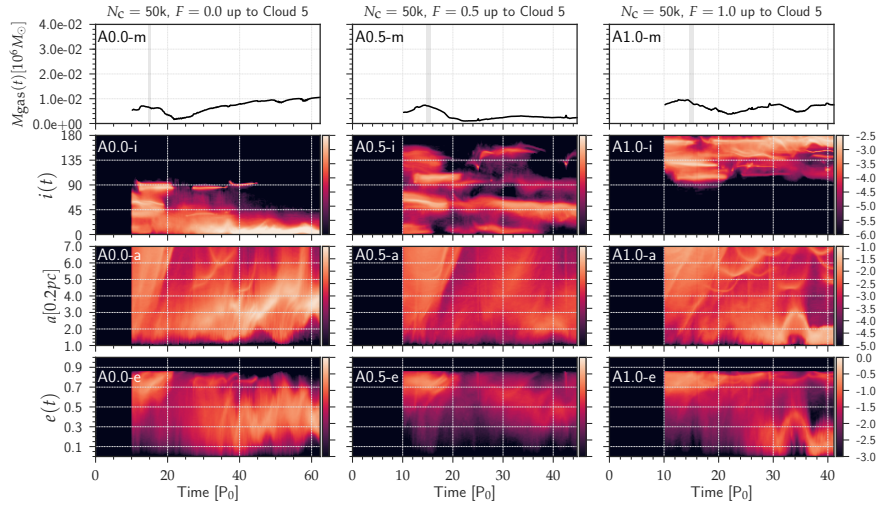


Figure 53: Same as Fig. 49 but for the set of ‘forked’ simulations. The $F = 0.0$, $F = 0.5$ and $F = 1.0$ distributions are shown from left to right. The vertical grey stripes in the first row mark the time of arrival of the 5th cloud. No further cloud is included into the system.

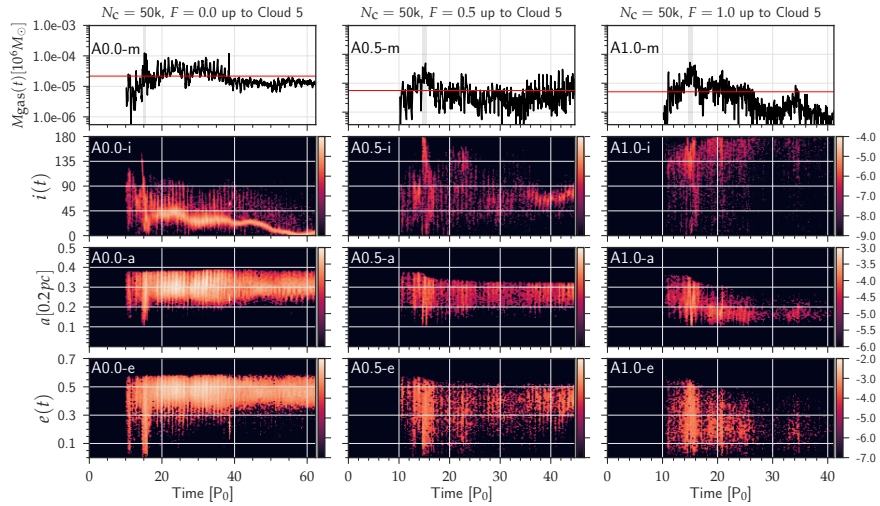


Figure 54: Same as Fig. 51 but for the “forked” simulations. Note the log-scale in the top row.

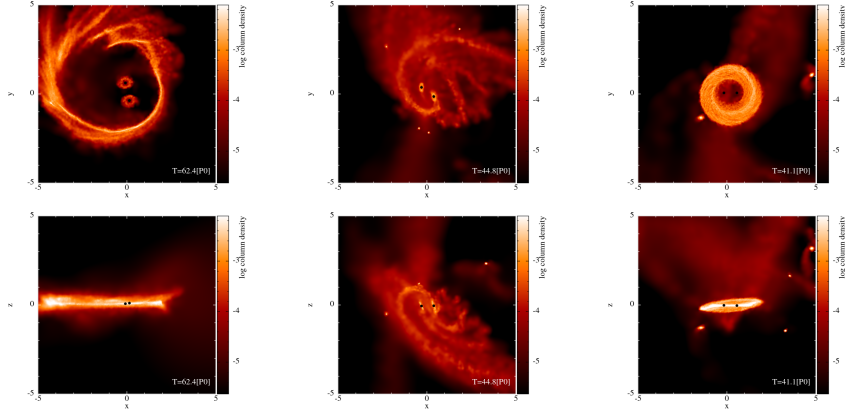


Figure 55: Visualisation of the last snapshot of the “forked” simulations. $F = 0.0$, $F = 0.5$ and $F = 1.0$ runs are shown left to right. The top and bottom row display $x - y$ and $x - z$ views of the system.

from simulations forked after the 5th clouds, the outcome of the ones forked after the 10th cloud are qualitatively very similar.

To illustrate the evolution of the system, we perform the same analysis described in §10.5.2 and build 2-D density plots to study the main properties of the particles forming both the circumbinary disc, and the mini-discs around each M_{BH} . Those are shown in Fig. 53 and 54 respectively. We now discuss the main long term features of the $F = 0.0$, $F = 1.0$ and $F = 0.5$ run separately.

In the left column of Fig. 53 we can see that in the $F = 0.0$ run, after the infall of the 5th cloud, the remaining material from the other disrupted clouds falls back forming a prominent circumbinary disc around the MBHB (panel 0.0-m). The disc is relatively thin, with an inclination of about 10° with respect to the binary orbital plane (panel 0.0-i), and it most of the mass is concentrated at a distance $2 \leq a \leq 5$ from the binary centre of mass, with a broad range of eccentricities $0.1 < e < 0.7$ (panels 0.0-a and 0.0-e). The left column of Fig. 54 allow us to confirm the long term stability of the mini-discs, which is almost aligned with the MBHB orbital plane.

The circumbinary disc in the $F = 1.0$ simulation (right column) displays a very similar behaviour. Disrupted cloud material falling back on orbits with comparable inclinations, interact with each other leading to the formation of a relatively massive circumbinary disc (panel 1.0-m). From panel 1.0-i, we note that, due to the high initial inclination and to the fact that $L_d \ll L_B$, the disc tends to counter-align with the MBHB [King et al., 2005]. By the end of the simulation, an almost perfectly counter-rotating circumbinary disc has formed, with $i \approx 170^\circ$. The circumbinary disc is much better defined than in the $F = 0.0$ case; particles are mostly confined within $a \leq 2.5$ (panel 1.0-a) and have lower eccentricities $e < 0.3$ (panel 1.0-e). The right column of Fig. 54 shows that mini-discs are much lighter, intermittent, and tend to be consumed with time. Note the increase in the mini-disc mass around $T = 35P_0$ (panel 1.0-m), in correspondence of the decrease in the circumbinary disc mass. This is due to a stream of gas partially disrupting the circumbinary structure and feeding gas to the central MBHB.

The $F = 0.5$ simulation does not seem to converge towards a specific relaxed state. The very diverse orbits of infalling clouds prevent the stream to efficiently interact, dissipate angular momentum and circularise into any

specific circumbinary structure. In fact, the mass in a putative ‘circumbinary disc’ does not grow in time (panel 0.5-m) and streams at different inclinations are clearly recognisable until the end of the simulation (panel 0.5-i). Fig. 54 shows that mini-discs are not very prominent, and also tend to be consumed with time.

A summary of the overall structures formed by the end of these “forked” simulations is displayed on Fig. 55. The $x - y$ and $x - z$ views highlight the thin circumbinary discs forming in both the $F = 0.0$ and $F = 1.0$ runs, along with the well defined mini-discs in the former case. Conversely, in the $F = 0.5$ case, no coherent structure is recognisable, and at least four distinct streams originating from different clouds are clearly visible. The aftermath of an epoch of incoherent accretion events is therefore strongly dependent on the properties of the infalling clouds. In general, ordered circumbinary structures persist only if the sum of all infalling material has a substantial net angular momentum (either co- or counter-aligned with the MBHB).

10.6 DISCUSSION AND FUTURE WORK

In this work we have used detailed SPH simulations performed with a modified version of the code GADGET-3 to study the interaction with a circular, equal mass MBHB with a series of infalling clouds in very eccentric orbits. We performed six main runs, considering two distributions of cloud pericentre distances and arrival times (defining RunA and RunB), and three distribution of angular momenta with different degree of anisotropy that we labelled $F = 0.0$ (co-rotating clouds), $F = 0.5$ (isotropically distributed clouds) and $F = 1.0$ (counter-rotating clouds). The goal is to study the dynamics of the MBHB-gas interaction when the binary is supplied with gas in incoherent discrete ‘pockets’, which might be a typical situation in the turbulent environment of high redshift, gas-rich galaxies [e.g. Prieto et al., 2017] and merger remnants [e.g. Perret et al., 2014], relevant to the early build-up of M_{BHs} [see Volonteri, 2010, for a review] and to future low frequency gravitational wave observations with LISA [Amaro-Seoane et al., 2017]. The main focus of this chapter was on the formation and evolution of bound gaseous structures, in a complementary paper (Goicovic et al. [2018]) to this chapter, we will turn our attention on the evolution of the MBHB.

Our main findings are summarised in the following points.

Post interaction gas distribution. In general, the density profile distribution of the gas post-interaction with the MBHB follows a $\rho_{\text{gas}} \propto r^{-2}$ distribution. This is expected since the speed of sound of the gas is just few km s^{-1} , implying a large Bondi radius of the MBHB, extending beyond the bulk of the extended gas. Compression of the clouds at pericentre also causes the formation of several, extremely dense self-gravitating clumps that can be favourable sites for *in-situ* star formation. In this work, we stopped gravitational collapse of the clumps by imposing an adiabatic behaviour of gas region above a critical gas density.

Circumbinary structures. Many theoretical and numerical studies of MBHB evolution in gaseous environments rely on the presence of a relatively stable, extended circumbinary disc that can efficiently extract energy and angular momentum from the binary [e.g. MacFadyen and Milosavljević, 2008, Cuadra et al., 2009, Roedig et al., 2011, Kocsis et al., 2012, Tang et al., 2017]. However, the route to the formation of such stable, extended disc has hardly been investigated. Our simulations show that, when the gas is fed to the binary in incoherent pockets, it is hard to form a massive circumbinary structure.

Adding more clouds to the system is a two-edged blade: on the one hand, new clouds supply fresh material that can add-up to the mass budget of the disc; on the other hand, clouds intersecting at high inclinations with a pre-existent disc can cause its partial disruption. Even in the co-rotating simulations ($F = 0.0$), after the addition of 20 clouds, injecting in the system a gas mass equal to 20% of the MBHB mass, the circumbinary disc mass does not exceed a mere 1% of the binary mass ($\approx 10^4$ solar masses for our fiducial system). Extrapolating from Goicovic et al. [2016], the naive expectation would be that about 50% of the injected mass ends up in a circumbinary structure ($\approx 10^5$ solar masses for our fiducial system). More isotropic and counter-rotating cloud distributions ($F = 0.5$ and $F = 1.0$) result in even lighter and more transient circumbinary discs.

Counter-rotating rings. The interaction with counter-rotating gas clouds result in the formation of compact counter-rotating rings. Due to the absence of Lindblad resonances for co-rotating orbits, the inner edge of the ring has about the size of the MBHB orbit and is not sustained by its forcing quadrupolar potential, resulting in an extremely unstable structure. The net consequence is that interaction with clouds on small impact parameters easily disrupt those rings, triggering copious accretion on the MBHB. Therefore, building-up extended counter-rotating structures is virtually impossible in our scenario, since it would require a much more gentle supply of material with larger angular momentum, within a relatively small range of inclinations.

Mini-discs formation. As expected from previous studies [Goicovic et al., 2016], a co-rotating cloud distribution ($F = 0.0$) results in the formation of prominent mini-discs. Although those mini-discs can also be significantly affected by the infall of new clouds, they are usually able to maintain certain stability, and they re-build quite efficiently following disruption from particularly aggressive cloud interactions. There is a notable tendency of the mini-discs to align after a few orbital periods (see, Fig. 51). mini-discs are much less prominent in the $F = 0.5$ and $F = 1.0$ runs. Because of the higher relative velocity between the gas particles and the M_{BH} s, the cross section of gas capture within the M_{BH} Roche Lobe is often smaller than the sink radius. It is possible that mini-discs form also in those cases, but on much smaller scales, that cannot be resolved by our current set-up.

Post interaction relaxation. When the supply of gas clouds ceases, the system tends to relax into a configuration that depends on the overall angular momentum distribution of the gas. This was investigated in our “forked” simulations, in which we limited the supply of gas to the 5th cloud. The stopping of cloud infall avoid further disruption of the circumbinary structures, that grow their mass in time, approaching a stable configuration. After about $20P_0$ from the infall of the last cloud, a prominent co-rotating circumbinary disc forms in the $F = 0.0$ case, whereas a well defined compact thin ring forms in the $F = 1.0$ case. The mass in these structures is about 10% of the total mass content of the supplied clouds. Post-infall relaxation does not lead to any well defined structure in the isotropic case $F = 0.5$, and several incoherent streams are still present at the end of the simulation. We speculate that in the long term, interaction within the stream will result in a relatively homogeneous, tenuous gas envelope.

Accretion onto the MBHB. A consequence of the violent interaction between different clouds and the continuous disruption of circumbinary structure is the triggering of the infall of enormous amount of gas onto the M_{BH} s. This has been seen in larger scale simulations of M_{BH} fuelling [Hobbs et al., 2011, Carmona-Loaiza et al., 2015], and in simulations of accretion onto a MBHB

due to ‘disk tearing’ [Doğan et al., 2015]; the interaction of gas streams coming from different direction cause efficient cancellation of their respective angular momenta, resulting in efficient infall onto the binary. In the test runs presented here, we showed that the associated accretion rate can be as high as $\approx 100\dot{M}_{\text{Edd}}$, when scaled to our fiducial system. This has a strong effect on the dynamics of the binary, which we investigate in detail in a complementary paper (Goicovic et al. [2018]) to this chapter.

Most importantly, our results show that the effect of multiple cloud interactions with a MBHB, does not sum-up to the effect of the single clouds taken individually. Cloud-cloud interactions have a fundamental role in shaping the gaseous structures forming around the binary, promoting continuous formation and disruption of circumbinary discs or ring and triggering episodes of enhanced accretion onto the MBHB. In general, we found that it is difficult to grow extended, massive circumbinary structures. This means that the evolution of the MBHB in this incoherent-feeding scenario is driven mostly by direct gas capture and accretion rather than resonant torques exerted by a circumbinary disc, as we explore in Goicovic et al. [2018]. Our simulations are the first to explore in detail this incoherent MBHB feeding scenario. We focused on some specific aspects of the system evolution, but there is a number of different properties of the systems that can be further investigated. For example, the distribution of dense clumps can be used to simulate star cluster formation in the vicinity of the binary and to study their further interaction with the binary. Likewise, the final state of the system can be evolved for longer time to better assess the stability and fate of the gaseous structures on longer timescales. Due to computational constraints, we could follow “forked” simulations only for $20\text{-}25P_0$, corresponding to only 0.2 Myr for our fiducial system. Finally, we remark that we implemented an extremely simplified hydrodynamic scheme, featuring an effective isothermal/adiabatic equation of state and ignoring any feedback from accretion onto the MBHB. Eventually, our result should be tested against enhanced simulations including realistic cooling prescription capturing disc fragmentation, together with a scheme tracing accretion feedback on the surrounding gas.

10.7 COMPLEMENTARY MATERIAL

10.7.1 *Generating the clouds’ initial conditions from the angular momentum vector*

The initial angular momentum vector of each cloud orbit is determined by sampling the orientation and the pericentre distance. However, the orbit of each cloud is not fully determined, as a Keplerian trajectory is defined by a total of 6 parameters. Having $(\theta_\ell, \phi_\ell, r_p)$, together with the initial distance and speed of the cloud, leaves one degree of freedom, which basically means that we can choose the initial position to be oriented in *any* direction as long as it lies on the plane defined by the unitary vector

$$\hat{\ell} = (\cos \phi_\ell \sin \theta_\ell, \sin \phi_\ell \sin \theta_\ell, \cos \theta_\ell), \quad (65)$$

where θ_ℓ and ϕ_ℓ are the polar and azimuthal angles of the angular momentum direction, respectively.

The initial position and velocity vectors are generated by first obtaining an arbitrary unitary vector \hat{e} lying on the aforementioned plane by taking the cross product with the x -axis

$$\hat{e} = \hat{\ell} \times \hat{x}. \quad (66)$$

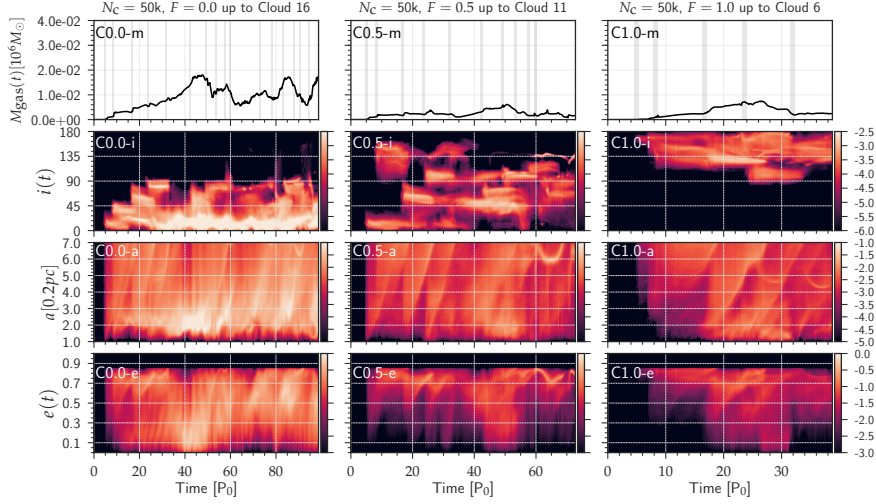


Figure 56: Same as Fig. 49 for RunC.

Then we rotate this vector by a random angle β , drawn uniformly between 0 and 2π , using the expression

$$\mathbf{r}' = \hat{\mathbf{e}} \cos \beta + (\hat{\mathbf{l}} \times \hat{\mathbf{e}}) \sin \beta. \quad (67)$$

This ensures that there is no preferential direction of incoming gas.

On the other hand, the velocity vector also lies on the plane defined by the angular momentum ($\hat{\mathbf{l}}$), but is rotated by an angle θ_{vel} with respect to the position vector, and therefore it can be obtained using a similar expression

$$\mathbf{v}' = \hat{\mathbf{e}} \cos(\beta - \theta_{\text{vel}}) + (\hat{\mathbf{l}} \times \hat{\mathbf{e}}) \sin(\beta - \theta_{\text{vel}}), \quad (68)$$

where β has the same value as used in eq. (67). The angle θ_{vel} is directly related to the pericentre distance as follows

$$\theta_{\text{vel}} = \arcsin\left(\frac{v_p}{v_{\text{ini}}} \frac{r_p}{d_{\text{ini}}}\right), \quad (69)$$

where $d_{\text{ini}} = 15a_0$ is the initial distance to the binary centre of mass, $v_{\text{ini}} = 0.25\sqrt{GM_0/a_0}$ is the initial velocity and v_p is the velocity at periastris.

Finally, we normalise these vectors to the initial distance and velocity,

$$\mathbf{r} = \frac{d_{\text{ini}}}{\|\mathbf{r}'\|} \mathbf{r}', \quad (70)$$

$$\mathbf{v} = -\frac{v_{\text{ini}}}{\|\mathbf{v}'\|} \mathbf{v}', \quad (71)$$

which yields the orbit defined by the sampled angular momentum vector. The initial position and velocity vectors of each cloud are displayed in Table 4, listed here so that the results can be reproduced by future studies.

10.7.2 Supplementary suite of simulations (RunC)

We show here some results for the extra suite of simulations (RunC) combining the cloud angular momentum distribution and pericentre distances of RunA (see Fig. 42 and 41) with the cloud time of arrivals of RunB (see Table 2). Due to limiting computing power, only 16 10 and 5 clouds arrived onto the binary

Distribution	Cloud	RunA						RunB					
		$\vec{r} [a_0]$			$\vec{v} [v_0]$			$\vec{r} [a_0]$			$\vec{v} [v_0]$		
F = 0.0	01	14.63	1.24	3.04	-0.22	0.09	-0.04	1.68	10.45	-10.63	-0.05	-0.23	0.09
	02	-10.06	-5.58	-9.62	0.21	-0.01	0.15	-11.00	-7.22	-7.21	0.22	0.10	0.05
	03	-12.28	6.96	-5.08	0.12	-0.14	0.17	-0.38	6.18	-13.66	-0.02	-0.17	0.18
	04	-4.13	-10.12	10.28	0.06	0.09	-0.22	-10.65	-5.85	-8.80	0.21	0.10	0.09
	05	12.21	7.61	-4.25	-0.22	-0.08	0.07	-13.12	7.25	0.43	0.19	-0.16	0.01
	06	1.54	-7.94	12.63	0.05	0.21	-0.12	-12.88	-0.82	-7.65	0.22	-0.08	0.09
	07	6.10	6.20	-12.22	-0.04	0.00	0.25	-4.80	11.49	8.36	-0.03	-0.15	-0.20
	08	11.08	3.20	9.60	-0.19	0.00	-0.15	6.41	12.84	-4.38	-0.11	-0.22	0.06
	09	10.63	-10.37	-2.10	-0.16	0.19	-0.04	-13.07	-0.03	-7.35	0.18	-0.11	0.14
	10	13.52	4.21	-4.93	-0.22	-0.02	0.12	12.04	8.12	-3.75	-0.24	-0.03	0.05
	11	10.23	10.79	1.97	-0.21	-0.09	-0.07	-5.79	-9.17	-10.36	0.13	0.20	0.08
	12	-1.35	11.93	-8.99	-0.07	-0.23	0.07	-0.85	-14.98	-0.06	0.01	0.22	-0.11
	13	-2.67	-6.03	-13.47	0.14	0.10	0.19	8.24	-5.29	11.36	-0.11	0.11	-0.20
	14	8.18	-1.46	-12.49	-0.09	0.05	0.23	-10.31	-8.45	6.88	0.20	0.15	-0.05
	15	4.68	5.60	13.11	-0.09	0.01	-0.22	6.82	3.50	-12.90	-0.16	-0.03	0.19
	16	—	—	—	—	—	—	-9.59	-10.59	4.56	0.17	0.18	-0.04
	17	—	—	—	—	—	—	-4.74	-12.65	6.52	0.15	0.20	-0.04
	18	—	—	—	—	—	—	-13.91	4.85	2.84	0.20	-0.13	-0.09
	19	—	—	—	—	—	—	14.41	-3.42	2.38	-0.19	0.05	-0.16
	20	—	—	—	—	—	—	-12.34	-8.50	0.73	0.13	0.21	0.02
	F = 0.5	01	14.63	1.24	3.04	-0.22	0.09	-0.04	1.68	10.45	-10.63	-0.05	-0.23
02		-10.06	10.81	2.62	0.21	-0.09	-0.10	-11.00	0.12	-10.20	0.22	0.04	0.11
03		-12.28	6.96	-5.08	0.12	-0.14	0.17	-0.38	-8.35	-12.46	-0.02	0.07	0.24
04		-4.13	-14.40	-0.76	0.06	0.23	-0.08	-10.65	-1.93	-10.39	0.21	0.06	0.12
05		12.21	-8.52	-1.82	-0.22	0.11	0.02	-13.12	-4.87	-5.39	0.19	0.12	0.10
06		1.54	-7.94	12.63	0.05	0.21	-0.12	-12.88	-0.82	-7.65	0.22	-0.08	0.09
07		6.10	11.37	-7.64	-0.04	-0.12	0.22	-4.80	11.49	8.36	-0.03	-0.15	-0.20
08		11.08	3.20	9.60	-0.19	0.00	-0.15	6.41	12.84	-4.38	-0.11	-0.22	0.06
09		10.63	-9.54	4.59	-0.16	0.12	-0.14	-13.07	-5.28	5.12	0.18	0.18	-0.02
10		13.52	4.21	-4.93	-0.22	-0.02	0.12	12.04	8.12	-3.75	-0.24	-0.03	0.05
11		10.23	10.79	1.97	-0.21	-0.09	-0.07	-5.79	-7.62	-11.55	0.13	0.18	0.11
12		-1.35	-13.21	-6.97	-0.07	0.16	0.19	-0.85	-14.98	-0.06	0.01	0.22	-0.11
13		-2.67	5.47	-13.71	0.14	-0.06	0.21	8.24	-5.29	11.36	-0.11	0.11	-0.20
14		8.18	6.74	-10.61	-0.09	-0.10	0.21	-10.31	-5.12	9.61	0.20	0.11	-0.10
15		4.68	5.60	13.11	-0.09	0.01	-0.22	6.82	3.50	-12.90	-0.16	-0.03	0.19
16		-1.93	-13.95	5.17	0.11	0.17	-0.15	-9.59	-10.59	4.56	0.17	0.18	-0.04
17		12.22	-6.54	-5.73	-0.22	0.12	-0.02	-4.74	-12.65	6.52	0.15	0.20	-0.04
18		0.21	7.90	-12.75	0.03	-0.02	0.25	-13.91	-4.08	3.86	0.20	0.12	-0.10
19		2.26	-5.10	13.92	-0.03	0.04	-0.24	14.41	-3.28	2.56	-0.19	0.04	-0.16
20		6.53	-8.72	-10.31	-0.15	0.09	0.18	-12.34	7.32	-4.38	0.13	-0.20	0.07
21		13.38	0.58	6.75	-0.24	-0.01	-0.05	—	—	—	—	—	—
22		-10.89	-3.93	-9.53	0.17	-0.04	0.19	—	—	—	—	—	—
F = 1.0	01	14.63	-1.46	-2.94	-0.22	-0.08	0.06	1.68	1.55	-14.83	-0.05	-0.12	0.21
	02	-10.06	10.81	2.62	0.21	-0.09	-0.10	-11.00	0.12	-10.20	0.22	0.04	0.11
	03	-12.28	-2.25	-8.31	0.12	0.10	0.20	-0.38	-8.35	-12.46	-0.02	0.07	0.24
	04	-4.13	-14.40	-0.76	0.06	0.23	-0.08	-10.65	-1.93	-10.39	0.21	0.06	0.12
	05	12.21	-8.52	-1.82	-0.22	0.11	0.02	-13.12	-4.87	-5.39	0.19	0.12	0.10
	06	1.54	11.66	9.31	0.05	-0.10	-0.21	-12.88	6.29	4.43	0.22	-0.01	-0.12
	07	6.10	11.37	-7.64	-0.04	-0.12	0.22	-4.80	-7.31	12.19	-0.03	0.18	-0.17
	08	11.08	-6.62	-7.64	-0.19	0.06	0.15	6.41	1.77	-13.45	-0.11	-0.04	0.22
	09	10.63	-9.54	4.59	-0.16	0.12	-0.14	-13.07	-5.28	5.12	0.18	0.18	-0.02
	10	13.52	3.18	5.66	-0.22	-0.09	-0.06	12.04	-8.91	0.81	-0.24	0.05	-0.04
	11	10.23	-8.59	-6.82	-0.21	0.05	0.12	-5.79	-7.62	-11.55	0.13	0.18	0.11
	12	-1.35	-13.21	-6.97	-0.07	0.16	0.19	-0.85	-13.70	-6.04	0.01	0.25	-0.01
	13	—	—	—	—	—	—	8.24	10.96	6.07	-0.11	-0.19	-0.13
	14	—	—	—	—	—	—	-10.31	-5.12	9.61	0.20	0.11	-0.10
	15	—	—	—	—	—	—	6.82	-7.47	-11.08	-0.16	0.12	0.15
	16	—	—	—	—	—	—	-9.59	-8.97	7.25	0.17	0.16	-0.09
	17	—	—	—	—	—	—	-4.74	8.48	11.43	0.15	-0.07	-0.19
	18	—	—	—	—	—	—	-13.91	-4.08	3.86	0.20	0.12	-0.10
	19	—	—	—	—	—	—	14.41	-3.28	2.56	-0.19	0.04	-0.16
	20	—	—	—	—	—	—	-12.34	-8.50	0.73	0.13	0.21	0.02

 Table 4: Initial 3-D position and velocity (x, y, z components) of the centre of mass of each cloud for RunA and RunB.

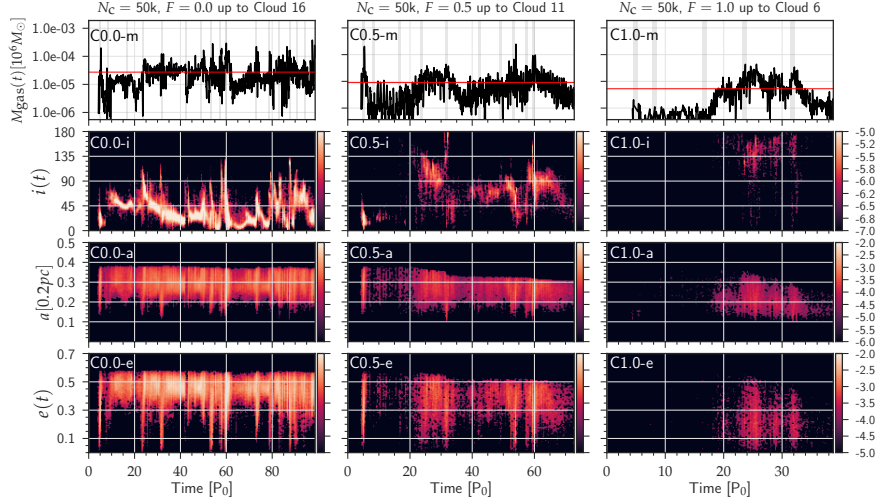


Figure 57: Same as Fig. 51 for RunC.

in run $F = 0.0$, $F = 0.5$ and $F = 1.0$ respectively. And main properties of the gas structures forming along the simulations are shown in Fig. 56 and 57.

The overall behaviour of simulations is quite similar to what shown in Fig. 49 and 51 for RunA. Indicating, not surprisingly, that the distribution of the cloud parameters, rather than their times of arrival, leaves a strong imprint on the forming gas structures. The $F = 0.0$ case (left column in both figures) initially builds up a prominent, co-rotating circumbinary structure confined to $i < 30^\circ$ and $a < 3$ (Fig. 56), which is partially destroyed by clouds incoming at $T \approx 50P_0$. A similar behaviour is seen in the mini-discs evolution (Fig. 57). Note, however that the mini-disc inclination shows much larger fluctuations often reaching values larger than 45° . Mass accumulation in circumbinary structures is much smaller in the $F = 0.5$ and $F = 1.0$ cases (central and right columns, respectively), and mini-discs are much lighter and intermittent (especially in the $F = 1.0$ case), in line with the general finding of RunA and RunB. Overall, this new suite of runs confirm the robustness of the general features highlighted in the main body of the chapter.

Part III

CONCLUSIONS

 MAIN CONCLUSIONS

Most of the work of my PhD has been devoted to answering the question about how do black hole binary systems form and evolve. I have addressed (i) the dynamics of dense stellar systems with a direct-summation N -body code which I have written from scratch. The developing and testing of the code took a considerable amount of time of my PhD. It has proved to be a robust integrator, yielding results very similar to other codes, in particular to Sverre Aarseth's well known NBODY, which he has developed for over 50 years. I have been able to investigate the impact of the softening on to the global evolution of the whole cluster, and I have also included a treatment for relativistic encounters, vital to simulate sources of GWs. Since the first version of GRAVIDY was ready I have been exploring how to manage the formation and evolution of binaries of stellar-mass black holes formed dynamically together with the evolution of globular clusters, this is part of current, on-going work, which I will hopefully present in the near future as a follow-up of my research.

On a very different regime of masses, I have studied (ii) binaries of supermassive black holes, such as the one depicted in the illustration of Fig. 58. Because of the characteristics of the problem, in this case we need to take into account another layer of complexity, the interaction with gas. It may seem a trivial addition to the problem at first sight, but the truth is that by adding gas dynamics to the vacuum problem, a plethora of -not quite but almost-completely free parameters seems to be similar to opening Pandora's box.

When I got involved in the research project of retrograde accretion discs around binaries of SMBHs, I got more and more interested by the problem and decided to investigate a maybe more fundamental issue: how do accretion discs form around binaries of SMBHs, and do they form "clean" discs, as often used in the literature, or may we expect other episodic architectures, and what is the interplay between the binary itself and the disc, if any? Responding this question turned out to be such a complex problem that I spent the remaining of my available time with it, but the findings were very relevant.

In this final chapter I give an overview of the most interesting findings of my work. This list is not exhaustive, because each chapter has its own conclusions, more detailed and technical. In this chapter I relax the technical terms on purpose to help gather a general impression rather than some particular aspects.

11.1 FROM STELLAR DYNAMICS...

As I have just explained, in chapter 8 I introduced a new direct-summation N -body code, GRAVIDY. My code focuses on more familiar computational environments, like a personal computer or a single cluster node. However, I have significantly improved the serial performance using both Central Processing Units (CPU) and Graphical Processing Units (GPU).

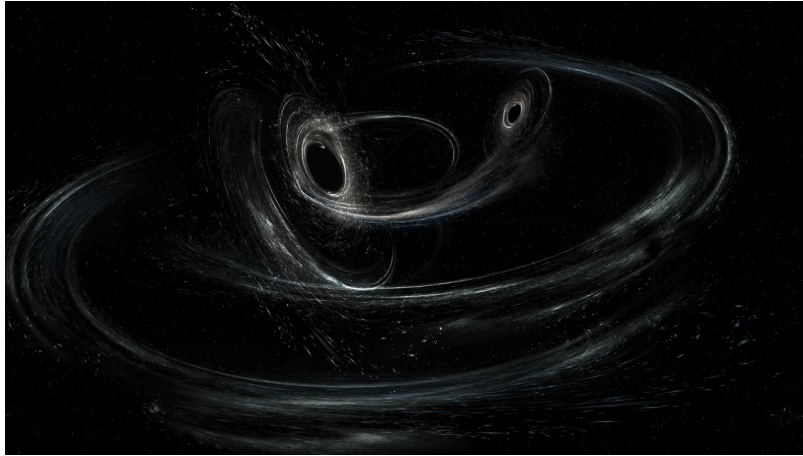


Figure 58: Two supermassive black holes on their way to the final coalescence can be tracked for up to years with a space-borne mission like LISA. Credit: LIGO/Caltech/MIT/Sonoma State (Aurore Simonnet)

In this first release of GRAVIDY, I have presented a series of classical tests of the code, as well as a study of the performance of its different “flavours”: the single CPU, the MPI and the GPU versions. One important question I have addressed is the role of the “softening” in the global evolution of a system, as integrated with our code. This softening is a fixed value added to the denominator in the calculation of the forces to make sure the code does not diverge numerically when two stars in the system come too close to each other. This is a very common practice in the field, but I wanted to understand what does it mean to effectively “switch off” gravity below a determined threshold. The dynamics is non-linear, and neither the impact of this assumption.

I find that the value of the softening is crucial in determining the global dynamics, and should not be taken lightly, in particular if one is interested in studying physical phenomena for which relaxation is important. Using a softening translates into a maximum increase of the forces and the a smoothly declination to zero, which is approximate. To study a dynamical process, such as e.g. the collision of two clusters, focusing on the short-term (i.e. for times well below a relaxation time) dynamical behaviour of the system, using a softening should be fine, but the role of the parameter should be assessed carefully by exploring different values.

The development of my code continues, and it is been adopted by other researchers who have approached me or simply downloaded the code from the public web page. The on-going development of GRAVIDY includes a close encounter solver, with a time-symmetric integration scheme to treat binaries, such as the one presented in the work of Konstantinidis and Kokkotas [2010]. The next step of this project is to run detailed, dedicated simulations of globular clusters so as to investigate the features that compact binaries have, as a way to understand both, sources of GWs and the formation and evolution of dense stellar systems.

Another immediate goal is to include a central massive particle and the required corrections to the gravitational forces so as to ensure a good conservation of the energy in the system. This massive particle could be envisaged

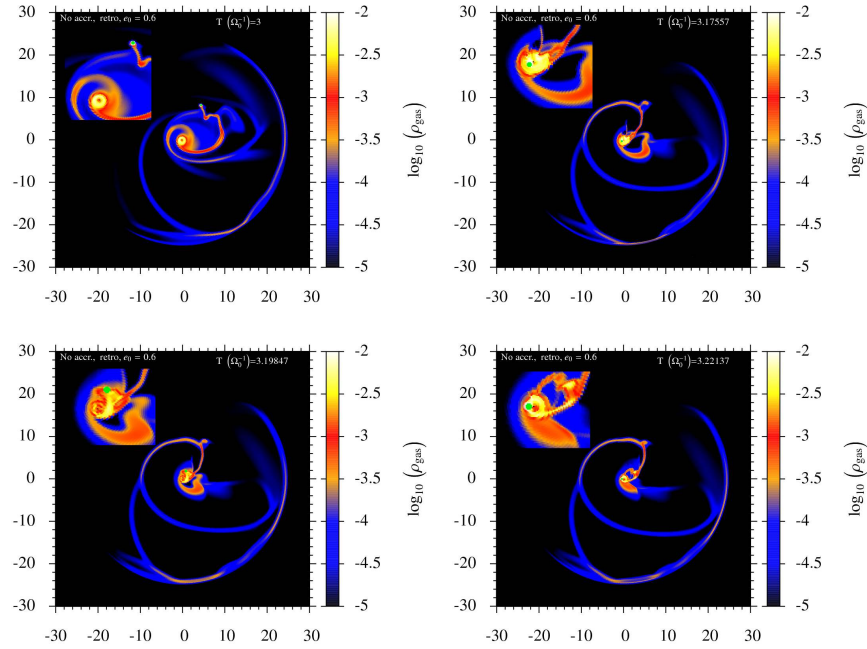


Figure 59: Four snapshots in the evolution of a counter-rotating system as presented in chapter 9. In this particular case, the secondary suffers a sudden braking which leads to a shrinkage of the semi-major axis of the binary. For each snapshot, I also show a zoom in of the area of interest.

as a massive black hole in a galactic centre (or a star in a protoplanetary system). I also plan on bridging the gap between spherical nucleus models that focus on collisional effects and simulations of larger structure that are able to account for complex, more realistic non-spherical geometry. Finally, a future goal is to include stellar evolution routines, from which the modularity of our code will provide an easy scenario. One of the candidate modules for this could be SEVN [Spera et al., 2015].

11.2 ... TO RETROGRADE ACCRETION DISCS...

When I was developing and testing the dynamics code I got interested by the general problem of how binaries of massive black holes assemble in galaxies and, in particular, by the role of gas. I got involved in a research project about the effect of circumbinary discs on the black hole binary evolution, from two circumstances regarding the rotation, a co-rotating and a counter-rotating scenario. The results are presented on chapter 9. For this project, my collaborators and I used two very different hydrodynamical codes. One of them is a two-dimensional code (Fargo), and a three-dimensional N -body/SPH code (Gadget). We modified the codes to include different accretion prescriptions so as to study a possible dependence with the overall evolution of the system. For the reasons mentioned in the introduction, we focus on an unequal mass binary to analyse the effects of the secondary black hole revolving in opposite direction to the gas. The secondary MBH plays a crucial role on the braking torque that is exerted on the binary. In Fig. 59 I depict an example of the work we did with Fargo.

The 2D models are fast and easy to integrate, but needed to be compared to 3D cases. We observed a very good comparison between the two, which allowed us to validate the 2D code.

Our study confirms a strong dependence on the accretion prescription for the evolution of the binary of SMBHs. Overestimating the accretion radius in the counter-rotating case leads to a fast, spurious evolution of the binary. In the quasi-circular case it is fine to assume that accretion only happens for bound gas, but this approach breaks when we have eccentric systems, and leads to wrong results. My description of accretion based on bound gas inside the Roche Lobe, should be used for these cases to ensure correct results.

11.3 ... AND THE FORMATION OF GASEOUS STRUCTURES...

As I have explained in the introduction of this thesis, and in chapter 10, one of the most interesting questions I tackled during the work that I presented in chapter 9 was about the formation of these gaseous structures. It is often assumed that, in theory, gas should be distributed in the shape of a disc around the binary of the two SMBHs, but I wondered whether other architectures were also possible. Also, the episodic nature and life duration of these structures was very important for me to address. The disc is crucial in the evolution of the binary because of the interplay between the two, as presented in chapter 10. If we want to study the evolution of the disc, we need to study the evolution of the binary, and its impact on to the disc, and vice-versa.

If the disc is so important for the binary, and the binary for the disc, is it correct to assume that we have both? The existence of binaries of SMBHs is still a subject of debate, as explained in chapter 9 and 10, and my approach in this regard was to assume that they do exist and inhabit the nuclei of galaxies. But what about the discs? I felt that it is relevant and important, as well as timely, to address this question, and I led an effort that got us important results.

I organised and led the development of an important amount of smooth particle hydrodynamical simulations to get a general idea of how these structures form. We investigated the possibility that gaseous discs should form from many discrete, incoherent accretion events. The main focus is twofold, because besides studying the formation, evolution and endurance of gaseous structures around binaries of SMBHs, I also wanted to understand what the impact on the evolution of the binary itself would be. In particular I was interested in knowing whether the semi-major axis would shrink in a short enough time so as to ensure its observation via the detection of the emission of the GWs.

I focused on different configurations, co-rotating, isotropic, and counter-rotating infalls of the gaseous clouds in terms of the angular momentum direction. Each case was complemented by a different distribution of pericentre passage of each cloud, and a different arrival time. This numerical effort led to interesting findings.

Circumbinary structures form in most of the cases. The formation of a disc structure, however, is challenging and strongly depends on the characteristics of the host galaxy and time interval for cloud infall. Besides the extended disc, I find that smaller, detached discs form around the individual SMBHs in

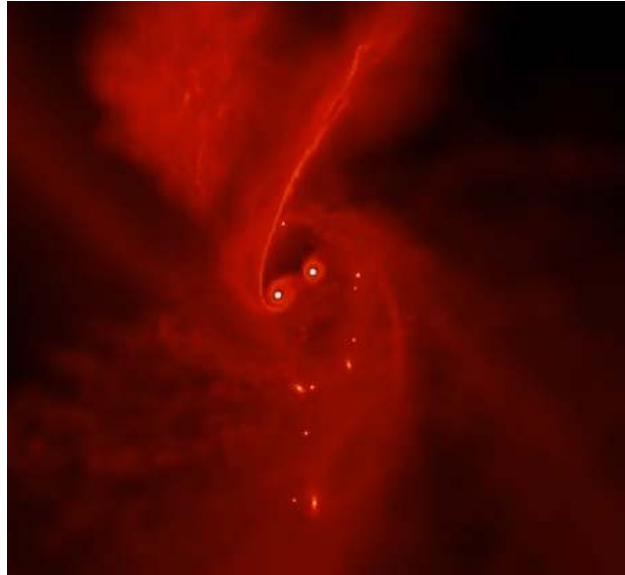


Figure 60: A moment in the evolution of a binary of two SMBHs with a gaseous structure forming around them upon the infall of various clouds. We can see clumps of gas forming because of the squeezing of gas around the binary. These gaseous lumps can be tidally disrupted by the binary. This snapshot corresponds to RunA, $F = 0$, as described in chapter 9.

the co-rotating scenarios only, with a tendency of discs alignment after a few orbital periods. In the case of counter-rotating infalling clouds, we observe gaseous clumps forming as in Fig. 60, which strongly interact with the binary and are tidally disrupted. The interaction between the material of each cloud is a key element on the global evolution of our systems, and can not be described as the sum of single cloud infall events.

11.4 ... TO STUDY BLACK HOLE BINARY SYSTEMS

In this PhD I have addressed the formation of black hole binary systems, for two particular regimes: dense stellar systems, where we expect binaries to form among stellar-mass black holes, and the nuclei of galaxies, where super-massive black holes are very likely lurking.

The first part of this thesis is devoted to the presentation of a new direct-summation N -body code which I have written totally from scratch, along with a suit of very-well known tests in the community of stellar dynamics, which the code successfully passed. The code is a Swiss army knife that can be used to investigate many different problems, although the prime goal was the formation of binaries of black holes. This is the reason why I included a relativistic treatment for the evolution of orbits of black holes when they reach relativistic speeds. My immediate future goal is to start doing realistic simulations of clusters to follow the formation and evolution of such binaries, as a way to understand what the orbital parameters of black holes are when they are on their way to coalesce. The code is public and it has been downloaded by many people which have contacted me to get involved in their research.

The second part deals with massive black holes and accretion discs. In particular, I was interested in understanding the interplay between such mas-

sive binaries and the gas which might be surrounding it. The first important aspect I investigated was the assumption made by many different groups as regards the amount of gas accreted by counter-rotating binaries of massive black holes, and I have shown that the way gas is usually assumed to be accreted leads to wrong results. In my study I derive a prescription that should be adopted for a correct evolution of the system.

As I am interested in understanding the very fundamental questions in which we base our progress for the problem of binaries, I questioned the formation of disc-like structures around these binaries of massive black holes, and the evolution of the binary itself in a more realistic scenario, such as galactic centres. What I found is that the formation of disc-like structures around the binary seems to be a particular, episodic event in a much more complex situation. Various structures form, and they have different impacts on to the evolution of the binary itself.

I envisage the research I present in this PhD as not complete and just a first step towards the understanding of how binaries of stellar-mass- or supermassive black holes form. Luckily, we live a very vibrant time in this field, which will lead to the answers of many questions and, undoubtedly, to many more new questions.

BIBLIOGRAPHY

- S. J. Aarseth. Direct methods for N-body simulations. In J. U. Brackbill and B. I. Cohen, editors, *Multiple time scales*, p. 377 - 418, pages 377–418, 1985.
- S. J. Aarseth. From NBODY1 to NBODY6: The Growth of an Industry. *The Publications of the Astronomical Society of the Pacific*, 111:1333–1346, November 1999.
- S. J. Aarseth. *Gravitational N-Body Simulations*. ISBN 0521432723. Cambridge, UK: Cambridge University Press, November 2003., November 2003.
- S. J. Aarseth and K. Zare. A regularization of the three-body problem. *Celestial Mechanics*, 10:185–205, October 1974. doi: 10.1007/BF01227619.
- S. J. Aarseth, M. Henon, and R. Wielen. A comparison of numerical methods for the study of star cluster dynamics. *A&A*, 37:183–187, December 1974.
- B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters*, 116(6):061102, February 2016a. doi: 10.1103/PhysRevLett.116.061102.
- B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence. *Physical Review Letters*, 116(24):241103, June 2016b. doi: 10.1103/PhysRevLett.116.241103.
- B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Physical Review Letters*, 118(22):221101, June 2017a. doi: 10.1103/PhysRevLett.118.221101.
- B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al. GW170608: Observation of a 19 Solar-mass Binary Black Hole Coalescence. *ApJ Lett.*, 851:L35, December 2017b. doi: 10.3847/2041-8213/aa9f0c.
- B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al. GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence. *Physical Review Letters*, 119(14):141101, October 2017c. doi: 10.1103/PhysRevLett.119.141101.
- B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Physical Review Letters*, 119(16):161101, October 2017d. doi: 10.1103/PhysRevLett.119.161101.
- M. A. Abramowicz, B. Czerny, J. P. Lasota, and E. Szuszkiewicz. Slim accretion disks. *ApJ*, 332:646–658, September 1988. doi: 10.1086/166683.

- P. Amaro-Seoane. Stellar dynamics and extreme-mass ratio inspirals. *ArXiv e-prints*, May 2012.
- P. Amaro-Seoane and X. Chen. Relativistic mergers of black hole binaries have large, similar masses, low spins and are circular. *MNRAS*, 458:3075–3082, May 2016. doi: 10.1093/mnras/stw503.
- P. Amaro-Seoane, M. Freitag, and R. Spurzem. Accretion of stars on to a massive black hole: a realistic diffusion model and numerical studies. *MNRAS*, 352:655–672, August 2004. doi: 10.1111/j.1365-2966.2004.07956.x.
- P. Amaro-Seoane, J. Barranco, A. Bernal, and L. Rezzolla. Constraining scalar fields with stellar kinematics and collisional dark matter. *Journal of Cosmology and Astroparticle Physics*, 11:2–+, November 2010. doi: 10.1088/1475-7516/2010/11/002.
- P. Amaro-Seoane, S. Aoudia, S. Babak, P. Binétruy, E. Berti, A. Bohé, C. Caprini, M. Colpi, N. J. Cornish, K. Danzmann, J.-F. Dufaux, J. Gair, O. Jennrich, P. Jetzer, A. Klein, R. N. Lang, A. Lobo, T. Littenberg, S. T. McWilliams, G. Nelemans, A. Petiteau, E. K. Porter, B. F. Schutz, A. Sesana, R. Stebbins, T. Sumner, M. Vallisneri, S. Vitale, M. Volonteri, and H. Ward. eLISA: Astrophysics and cosmology in the millihertz regime. *ArXiv e-prints*, January 2012a.
- P. Amaro-Seoane, S. Aoudia, S. Babak, P. Binétruy, E. Berti, A. Bohé, C. Caprini, M. Colpi, N. J. Cornish, K. Danzmann, J.-F. Dufaux, J. Gair, O. Jennrich, P. Jetzer, A. Klein, R. N. Lang, A. Lobo, T. Littenberg, S. T. McWilliams, G. Nelemans, A. Petiteau, E. K. Porter, B. F. Schutz, A. Sesana, R. Stebbins, T. Sumner, M. Vallisneri, S. Vitale, M. Volonteri, and H. Ward. Low-frequency gravitational-wave science with eLISA/NGO. *Classical and Quantum Gravity*, 29(12):124016, June 2012b. doi: 10.1088/0264-9381/29/12/124016.
- P. Amaro-Seoane, S. Aoudia, S. Babak, P. Binétruy, E. Berti, A. Bohé, C. Caprini, M. Colpi, N. J. Cornish, K. Danzmann, J.-F. Dufaux, J. Gair, I. Hinder, O. Jennrich, P. Jetzer, A. Klein, R. N. Lang, A. Lobo, T. Littenberg, S. T. McWilliams, G. Nelemans, A. Petiteau, E. K. Porter, B. F. Schutz, A. Sesana, R. Stebbins, T. Sumner, M. Vallisneri, S. Vitale, M. Volonteri, H. Ward, and B. Wardell. eLISA: Astrophysics and cosmology in the millihertz regime. *GW Notes, Vol. 6, p. 4-110*, 6:4–110, May 2013a.
- P. Amaro-Seoane, P. Brem, and J. Cuadra. Tidal Disruptions in Circumbinary Disks. I. Star Formation, Dynamics, and Binary Evolution. *ApJ*, 764:14, February 2013b. doi: 10.1088/0004-637X/764/1/14.
- P. Amaro-Seoane, C. Maureira-Fredes, M. Dotti, and M. Colpi. Retrograde binaries of massive black holes in circumbinary accretion discs. *A&A*, 591:A114, June 2016. doi: 10.1051/0004-6361/201526172.
- P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender, E. Berti, P. Binétruy, M. Born, D. Bortoluzzi, J. Camp, C. Caprini, V. Cardoso, M. Colpi, J. Conklin, N. Cornish, C. Cutler, K. Danzmann, R. Dolesi, L. Ferraioli, V. Ferroni, E. Fitzsimons, J. Gair, L. Gesa Bote, D. Giardino, F. Gilbert, C. Grimani, H. Halloin, G. Heinzl, T. Hertog, M. Hewitson, K. Holley-Bockelmann, D. Hollington, M. Hueller, H. Inchauspe, P. Jetzer, N. Karnesis, C. Killow, A. Klein, B. Klipstein, N. Korsakova, S. L. Larson, J. Livas, I. Lloro, N. Man, D. Mance, J. Martino, I. Mateos, K. McKenzie, S. T. McWilliams,

- C. Miller, G. Mueller, G. Nardini, G. Nelemans, M. Nofrarias, A. Petiteau, P. Pivato, E. Plagnol, E. Porter, J. Reiche, D. Robertson, N. Robertson, E. Rossi, G. Russano, B. Schutz, A. Sesana, D. Shoemaker, J. Slutsky, C. F. Sopuerta, T. Sumner, N. Tamanini, I. Thorpe, M. Troebs, M. Vallisneri, A. Vecchio, D. Vetrugno, S. Vitale, M. Volonteri, G. Wanner, H. Ward, P. Wass, W. Weber, J. Ziemer, and P. Zweifel. Laser Interferometer Space Antenna. *ArXiv e-prints*, February 2017.
- P. J. Armitage and P. Natarajan. Accretion during the Merger of Supermassive Black Holes. *ApJ Lett.*, 567:L9–L12, March 2002. doi: 10.1086/339770.
- P. J. Armitage and P. Natarajan. Eccentricity of Supermassive Black Hole Binaries Coalescing from Gas-rich Mergers. *ApJ*, 634:921–927, December 2005. doi: 10.1086/497108.
- P. J. Armitage, J. B. Simon, and R. G. Martin. Two Timescale Dispersal of Magnetized Protoplanetary Disks. *ApJ Lett.*, 778:L14, November 2013. doi: 10.1088/2041-8205/778/1/L14.
- P. Artymowicz and S. H. Lubow. Dynamics of binary-disk interaction. 1: Resonances and disk gap sizes. *ApJ*, 421:651–667, February 1994. doi: 10.1086/173679.
- P. Artymowicz and S. H. Lubow. Mass Flow through Gaps in Circumbinary Disks. *ApJ Lett.*, 467:L77, August 1996. doi: 10.1086/310200.
- F. K. Baganoff, M. W. Bautz, W. N. Brandt, G. Chartas, E. D. Feigelson, G. P. Garmire, Y. Maeda, M. Morris, G. R. Ricker, L. K. Townsley, and F. Walter. Rapid X-ray flaring from the direction of the supermassive black hole at the Galactic Centre. *Nat*, 413:45–48, September 2001. doi: 10.1038/35092510.
- F. K. Baganoff, Y. Maeda, M. Morris, M. W. Bautz, W. N. Brandt, W. Cui, J. P. Doty, E. D. Feigelson, G. P. Garmire, S. H. Pravdo, G. R. Ricker, and L. K. Townsley. Chandra X-Ray Spectroscopic Imaging of Sagittarius A* and the Central Parsec of the Galaxy. *ApJ*, 591:891–915, July 2003. doi: 10.1086/375145.
- J. N. Bahcall and R. A. Wolf. Star distribution around a massive black hole in a globular cluster. *ApJ*, 209:214–232, October 1976.
- J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, J. R. van Meter, and M. C. Miller. Getting a Kick Out of Numerical Relativity. *ApJ Lett.*, 653:L93–L96, December 2006. doi: 10.1086/510448.
- D. S. Balsara. von Neumann stability analysis of smooth particle hydrodynamics—suggestions for optimal algorithms. *Journal of Computational Physics*, 121:357–372, 1995. doi: 10.1016/S0021-9991(95)90221-X.
- J. Bankert, J. H. Krolik, and J. Shi. Structure of Retrograde Circumbinary Accretion Disks. *apj*, 801:114, March 2015. doi: 10.1088/0004-637X/801/2/114.
- J. Barnes and P. Hut. A hierarchical $O(N \log N)$ force-calculation algorithm. *Nat*, 324:446–449, December 1986. doi: 10.1038/324446a0.
- M. R. Bate, I. A. Bonnell, and N. M. Price. Modelling accretion in protobinary systems. *MNRAS*, 277:362–376, November 1995.

- H. Baumgardt, P. Amaro-Seoane, and R. Schödel. The distribution of stars around the Milky Way's black hole III: Comparison with simulations. *ArXiv e-prints*, January 2017.
- E. E. Becklin and G. Neugebauer. Infrared Observations of the Galactic Center. *ApJ*, 151:145, January 1968. doi: 10.1086/149425.
- M. C. Begelman, R. D. Blandford, and M. J. Rees. Massive black hole binaries in active galactic nuclei. *Nat*, 287:307–309, September 1980. doi: 10.1038/287307a0.
- R. G. Belleman, J. Bédorf, and S. F. Portegies Zwart. High performance direct gravitational N-body simulations on graphics processing units II: An implementation in CUDA. *New Astronomy*, 13:103–112, February 2008. doi: 10.1016/j.newast.2007.07.004.
- P. Berczik, D. Merritt, R. Spurzem, and H.-P. Bischof. Efficient Merger of Binary Supermassive Black Holes in Nonaxisymmetric Galaxies. *ApJ Lett.*, 642:L21–L24, May 2006. doi: 10.1086/504426.
- P. Berczik, K. Nitadori, S. Zhong, R. Spurzem, T. Hamada, X. Wang, I. Berentzen, A. Veles, and W. Ge. High performance massively parallel direct N-body simulations on large GPU clusters. In *International conference on High Performance Computing, Kyiv, Ukraine, October 8-10, 2011.*, p. 8-18, pages 8–18, October 2011.
- P. Berczik, R. Spurzem, L. Wang, S. Zhong, and S. Huang. Up to 700k GPU cores, Kepler, and the Exascale future for simulations of star clusters around black holes. In *Third International Conference "High Performance Computing", HPC-UA 2013*, p. 52-59, pages 52–59, October 2013.
- I. Berentzen, M. Preto, P. Berczik, D. Merritt, and R. Spurzem. Binary Black Hole Merger in Galactic Nuclei: Post-Newtonian Simulations. *ApJ*, 695: 455–468, April 2009. doi: 10.1088/0004-637X/695/1/455.
- L. Blecha, T. J. Cox, A. Loeb, and L. Hernquist. Recoiling black holes in merging galaxies: relationship to active galactic nucleus lifetimes, starbursts and the $M_{BH}-\sigma_*$ relation. *MNRAS*, 412:2154–2182, April 2011. doi: 10.1111/j.1365-2966.2010.18042.x.
- T. Bogdanović. Supermassive Black Hole Binaries: The Search Continues. *Astrophysics and Space Science Proceedings*, 40:103, 2015. doi: 10.1007/978-3-319-10488-1_9.
- I. A. Bonnell. A New Binary Formation Mechanism. *MNRAS*, 269, August 1994. doi: 10.1093/mnras/269.3.837.
- I. A. Bonnell and W. K. M. Rice. Star Formation Around Supermassive Black Holes. *Science*, 321:1060, August 2008. doi: 10.1126/science.1160653.
- P. Brem, P. Amaro-Seoane, and R. Spurzem. Relativistic mergers of compact binaries in clusters: the fingerprint of the spin. *MNRAS*, 434:2999–3007, October 2013. doi: 10.1093/mnras/stt1220.
- S. Callegari, L. Mayer, S. Kazantzidis, M. Colpi, F. Governato, T. Quinn, and J. Wadsley. Pairing of Supermassive Black Holes in Unequal-Mass Galaxy Mergers. *ApJ Lett.*, 696:L89–L92, May 2009. doi: 10.1088/0004-637X/696/1/L89.

- S. Callegari, S. Kazantzidis, L. Mayer, M. Colpi, J. M. Bellovary, T. Quinn, and J. Wadsley. Growing Massive Black Hole Pairs in Minor Mergers of Disk Galaxies. *ApJ*, 729:85, March 2011. doi: 10.1088/0004-637X/729/2/85.
- M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower. Accurate Evolutions of Orbiting Black-Hole Binaries without Excision. *Physical Review Letters*, 96(11):111101, March 2006. doi: 10.1103/PhysRevLett.96.111101.
- R. Capuzzo-Dolcetta and M. Spera. A performance comparison of different graphics processing units running direct N-body simulations. *Computer Physics Communications*, 184:2528–2539, November 2013. doi: 10.1016/j.cpc.2013.07.005.
- R. Capuzzo-Dolcetta, M. Spera, and D. Punzo. A fully parallel, high precision, N-body code running on hybrid computing platforms. *Journal of Computational Physics*, 236:580–593, March 2013. doi: 10.1016/j.jcp.2012.11.013.
- J. M. Carmona-Loaiza, M. Colpi, M. Dotti, and R. Valdarnini. Overlapping inflows as catalysts of AGN activity - II. Relative importance of turbulence and inflow-disc interaction. *MNRAS*, 453:1608–1618, October 2015. doi: 10.1093/mnras/stv1749.
- S. Chandrasekhar. Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction. *ApJ*, 97:255, March 1943. doi: 10.1086/144517.
- D. Chapon, L. Mayer, and R. Teyssier. Hydrodynamics of galaxy mergers with supermassive black holes: is there a last parsec problem? *MNRAS*, 429:3114–3122, March 2013. doi: 10.1093/mnras/sts568.
- M. Colpi. Massive Binary Black Holes in Galactic Nuclei and Their Path to Coalescence. *ssr*, 183:189–221, September 2014. doi: 10.1007/s11214-014-0067-1.
- M. Colpi and M. Dotti. Massive Binary Black Holes in the Cosmic Landscape. *Advanced Science Letters*, 4:181–203, February 2011. doi: 10.1166/asl.2011.1205.
- M. Colpi, L. Mayer, and F. Governato. Dynamical Friction and the Evolution of Satellites in Virialized Halos: The Theory of Linear Response. *ApJ*, 525:720–733, November 1999. doi: 10.1086/307952.
- M. Colpi, S. Callegari, M. Dotti, and L. Mayer. Massive black hole binary evolution in gas-rich mergers. *Classical and Quantum Gravity*, 26(9):094029, May 2009. doi: 10.1088/0264-9381/26/9/094029.
- J. Cuadra, P. J. Armitage, R. D. Alexander, and M. C. Begelman. Massive black hole binary mergers within subparsec scale gas discs. *MNRAS*, 393:1423–1432, March 2009. doi: 10.1111/j.1365-2966.2008.14147.x.
- J. Dabringhausen, M. Hilker, and P. Kroupa. From star clusters to dwarf galaxies: the properties of dynamically hot stellar systems. *MNRAS*, 386:864–886, May 2008. doi: 10.1111/j.1365-2966.2008.13065.x.
- L. del Valle and A. Escala. Binary-Disk Interaction: Gap-opening Criteria. *ApJ*, 761:31, December 2012. doi: 10.1088/0004-637X/761/1/31.
- L. del Valle and A. Escala. Binary-Disk Interaction. II. Gap-opening Criteria for Unequal-mass Binaries. *ApJ*, 780:84, January 2014. doi: 10.1088/0004-637X/780/1/84.

- L. del Valle, A. Escala, Maureira-Fredes C., J. Molina, J. Cuadra, and P. Amaro-Seoane. Supermassive Black Holes in a Star-forming Gaseous Circumnuclear Disk. *ApJ*, 811:59, September 2015. doi: 10.1088/0004-637X/811/1/59.
- L. del Valle, A. Escala, C. Maureira-Fredes, J. Molina, J. Cuadra, and P. Amaro-Seoane. Supermassive black holes in a star-forming gaseous circumnuclear disk. *ApJ*, 811(1):59, 2015. URL <http://stacks.iop.org/0004-637X/811/i=1/a=59>.
- B. Devecchi, E. Rasia, M. Dotti, M. Volonteri, and M. Colpi. Imprints of recoiling massive black holes on the hot gas of early-type galaxies. *MNRAS*, 394: 633–640, April 2009. doi: 10.1111/j.1365-2966.2008.14329.x.
- D. J. D’Orazio, Z. Haiman, and A. MacFadyen. Accretion into the central cavity of a circumbinary disc. *MNRAS*, 436:2997–3020, December 2013a. doi: 10.1093/mnras/stt1787.
- D. J. D’Orazio, Z. Haiman, and A. MacFadyen. Accretion into the central cavity of a circumbinary disc. *MNRAS*, 436:2997–3020, December 2013b. doi: 10.1093/mnras/stt1787.
- D. J. D’Orazio, Z. Haiman, and D. Schiminovich. Relativistic boost as the cause of periodicity in a massive black-hole binary candidate. *Nat*, 525:351–353, September 2015. doi: 10.1038/nature15262.
- M. Dotti, M. Colpi, and F. Haardt. Laser Interferometer Space Antenna double black holes: dynamics in gaseous nuclear discs. *MNRAS*, 367:103–112, March 2006. doi: 10.1111/j.1365-2966.2005.09956.x.
- M. Dotti, M. Colpi, F. Haardt, and L. Mayer. Supermassive black hole binaries in gaseous and stellar circumnuclear discs: orbital dynamics and gas accretion. *MNRAS*, 379:956–962, August 2007. doi: 10.1111/j.1365-2966.2007.12010.x.
- M. Dotti, M. Ruzzkowski, L. Paredi, M. Colpi, M. Volonteri, and F. Haardt. Dual black holes in merger remnants - I. Linking accretion to dynamics. *MNRAS*, 396:1640–1646, July 2009. doi: 10.1111/j.1365-2966.2009.14840.x.
- M. Dotti, M. Volonteri, A. Perego, M. Colpi, M. Ruzzkowski, and F. Haardt. Dual black holes in merger remnants - II. Spin evolution and gravitational recoil. *MNRAS*, 402:682–690, February 2010. doi: 10.1111/j.1365-2966.2009.15922.x.
- M. Dotti, A. Sesana, and R. Decarli. Massive Black Hole Binaries: Dynamical Evolution and Observational Signatures. *Advances in Astronomy*, 2012: 940568, 2012. doi: 10.1155/2012/940568.
- M. Dotti, M. Colpi, S. Pallini, A. Perego, and M. Volonteri. On the Orientation and Magnitude of the Black Hole Spin in Galactic Nuclei. *ApJ*, 762:68, January 2013. doi: 10.1088/0004-637X/762/2/68.
- S. Doğan, C. Nixon, A. King, and D. J. Price. Tearing up a misaligned accretion disc with a binary companion. *MNRAS*, 449:1251–1258, May 2015. doi: 10.1093/mnras/stv347.
- A. Dressler and D. O. Richstone. Stellar dynamics in the nuclei of M31 and M32 - Evidence for massive black holes? *ApJ*, 324:701–713, January 1988. doi: 10.1086/165930.

- A. C. Dunhill, R. D. Alexander, C. J. Nixon, and A. R. King. Misaligned accretion on to supermassive black hole binaries. *MNRAS*, 445:2285–2296, December 2014. doi: 10.1093/mnras/stu1914.
- P. P. Eggleton. Approximations to the radii of Roche lobes. *ApJ*, 268:368, May 1983. doi: 10.1086/160960.
- M. Eracleous, T. A. Boroson, J. P. Halpern, and J. Liu. A Large Systematic Search for Close Supermassive Binary and Rapidly Recoiling Black Holes. 201:23, August 2012. doi: 10.1088/0067-0049/201/2/23.
- A. Escala, R. B. Larson, P. S. Coppi, and D. Mardones. The Role of Gas in the Merging of Massive Black Holes in Galactic Nuclei. I. Black Hole Merging in a Spherical Gas Cloud. *apj*, 607:765–777, June 2004. doi: 10.1086/386278.
- A. Escala, R. B. Larson, P. S. Coppi, and D. Mardones. The Role of Gas in the Merging of Massive Black Holes in Galactic Nuclei. II. Black Hole Merging in a Nuclear Gas Disk. *ApJ*, 630:152–166, September 2005. doi: 10.1086/431747.
- B. D. Farris, Y. T. Liu, and S. L. Shapiro. Binary black hole mergers in gaseous disks: Simulations in general relativity. *Ph.Rv. D*, 84(2):024024, July 2011. doi: 10.1103/PhysRevD.84.024024.
- B. D. Farris, P. Duffell, A. I. MacFadyen, and Z. Haiman. Binary Black Hole Accretion from a Circumbinary Disk: Gas Dynamics inside the Central Cavity. *ApJ*, 783:134, March 2014. doi: 10.1088/0004-637X/783/2/134.
- L. Ferrarese and D. Merritt. A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies. *ApJ Lett.*, 539:L9–L12, August 2000. doi: 10.1086/312838.
- D. Fiacconi, L. Mayer, R. Roškar, and M. Colpi. Massive Black Hole Pairs in Clumpy, Self-gravitating Circumnuclear Disks: Stochastic Orbital Decay. *ApJ Lett.*, 777:L14, November 2013. doi: 10.1088/2041-8205/777/1/L14.
- J. Frank and M. J. Rees. Effects of massive central black holes on dense stellar systems. *mnras*, 176:633–647, September 1976. doi: 10.1093/mnras/176.3.633.
- T. Fukushige, J. Makino, and A. Kawai. GRAPE-6A: A Single-Card GRAPE-6 for Parallel PC-GRAPE Cluster Systems. *PASJ*, 57:1009–1021, December 2005.
- E. Gaburov, S. Harfst, and S. P. Zwart. SAPPORO: A way to turn your graphics cards into a GRAPE-6. *New Astronomy*, 14:630–637, October 2009. doi: 10.1016/j.newast.2009.03.002.
- E. Gallego-Cano, R. Schödel, H. Dong, F. Nogueras-Lara, A. T. Gallego-Calvente, P. Amaro-Seoane, and H. Baumgardt. The distribution of stars around the Milky Way’s central black hole. I. Deep star counts. *A&A*, 609:A26, January 2018. doi: 10.1051/0004-6361/201730451.
- M. Gaspari, M. Ruszkowski, and S. P. Oh. Chaotic cold accretion on to black holes. *MNRAS*, 432:3401–3422, July 2013. doi: 10.1093/mnras/stt692.
- M. Gaspari, F. Brighenti, and P. Temi. Chaotic cold accretion on to black holes in rotating atmospheres. *A&A*, 579:A62, July 2015. doi: 10.1051/0004-6361/201526151.

- M. Gaspari, P. Temi, and F. Brighenti. Raining on black holes and massive galaxies: the top-down multiphase condensation model. *MNRAS*, 466:677–704, April 2017. doi: 10.1093/mnras/stw3108.
- K. Gebhardt, D. Richstone, J. Kormendy, T. R. Lauer, E. A. Ajhar, R. Bender, A. Dressler, S. M. Faber, C. Grillmair, J. Magorrian, and S. Tremaine. Axisymmetric, Three-Integral Models of Galaxies: A Massive Black Hole in NGC 3379. *The Astronomical Journal*, 119:1157–1171, March 2000. doi: 10.1086/301240.
- R. Genzel and C. H. Townes. Physical conditions, dynamics, and mass distribution in the center of the Galaxy. *ARA&A*, 25:377–423, 1987. doi: 10.1146/annurev.aa.25.090187.002113.
- R. Genzel, D. Hollenbach, and C. H. Townes. The nucleus of our Galaxy. *Reports on Progress in Physics*, 57:417–479, May 1994. doi: 10.1088/0034-4885/57/5/001.
- R. Genzel, A. Eckart, T. Ott, and F. Eisenhauer. On the nature of the dark mass in the centre of the Milky Way. *MNRAS*, 291:219–234, October 1997. doi: 10.1093/mnras/291.1.219.
- R. Genzel, C. Pichon, A. Eckart, O. E. Gerhard, and T. Ott. Stellar dynamics in the Galactic Centre: proper motions and anisotropy. *MNRAS*, 317:348–374, September 2000. doi: 10.1046/j.1365-8711.2000.03582.x.
- R. Genzel, F. Eisenhauer, and S. Gillessen. The Galactic Center massive black hole and nuclear star cluster. *Reviews of Modern Physics*, 82:3121–3195, October 2010. doi: 10.1103/RevModPhys.82.3121.
- A. M. Ghez, B. L. Klein, M. Morris, and E. E. Becklin. High Proper-Motion Stars in the Vicinity of Sagittarius A*: Evidence for a Supermassive Black Hole at the Center of Our Galaxy. *ApJ*, 509:678–686, December 1998. doi: 10.1086/306528.
- A. M. Ghez, S. Salim, S. D. Hornstein, A. Tanner, J. R. Lu, M. Morris, E. E. Becklin, and G. Duchêne. Stellar Orbits around the Galactic Center Black Hole. *ApJ*, 620:744–757, February 2005. doi: 10.1086/427175.
- M. Giersz and R. Spurzem. Comparing direct N-body integration with anisotropic gaseous models of star clusters. *MNRAS*, 269:241, July 1994.
- F. G. Goicovic, J. Cuadra, A. Sesana, F. Stasyszyn, P. Amaro-Seoane, and T. L. Tanaka. Infalling clouds on to supermassive black hole binaries - I. Formation of discs, accretion and gas dynamics. *MNRAS*, 455:1989–2003, January 2016. doi: 10.1093/mnras/stv2470.
- F. G. Goicovic, A. Sesana, J. Cuadra, and F. Stasyszyn. Infalling clouds on to supermassive black hole binaries - II. Binary evolution and the final parsec problem. *MNRAS*, 472:514–531, November 2017. doi: 10.1093/mnras/stx1996.
- F. G. Goicovic, C. Maureira-Fredes, A. Sesana, P. Amaro-Seoane, and J. Cuadra. Accretion of clumpy cold gas onto massive black hole binaries: a possible fast route to binary coalescence. *ArXiv e-prints*, January 2018.
- P. Goldreich and S. Tremaine. Disk-satellite interactions. *ApJ*, 241:425–441, October 1980. doi: 10.1086/158356.

- J. A. González, U. Sperhake, B. Brügmann, M. Hannam, and S. Husa. Maximum Kick from Nonspinning Black-Hole Binary Inspiral. *Physical Review Letters*, 98(9):091101, March 2007. doi: 10.1103/PhysRevLett.98.091101.
- A. Gould and H.-W. Rix. Binary Black Hole Mergers from Planet-like Migrations. *ApJ Lett.*, 532:L29–L32, March 2000. doi: 10.1086/312562.
- F. Governato, M. Colpi, and L. Maraschi. The fate of central black holes in merging galaxies. *MNRAS*, 271, November 1994. doi: 10.1093/mnras/271.2.317.
- Leslie Greendard. *The rapid evaluation of potential fields in particles systems*. PhD thesis, Yale University, New Haven, CT, 1987.
- A. Gualandris and D. Merritt. Ejection of Supermassive Black Holes from Galaxy Cores. *ApJ*, 678:780–797, May 2008. doi: 10.1086/586877.
- A. Gualandris and D. Merritt. Long-term Evolution of Massive Black Hole Binaries. IV. Mergers of Galaxies with Collisionally Relaxed Nuclei. *ApJ*, 744:74, January 2012. doi: 10.1088/0004-637X/744/1/74.
- Z. Haiman, B. Kocsis, and K. Menou. The Population of Viscosity- and Gravitational Wave-driven Supermassive Black Hole Binaries Among Luminous Active Galactic Nuclei. *ApJ*, 700:1952–1969, August 2009. doi: 10.1088/0004-637X/700/2/1952.
- T. Hamada and T. Iitaka. The Chamomile Scheme: An Optimized Algorithm for N-body simulations on Programmable Graphics Processing Units. *New Astronomy*, March 2007.
- S. Harfst, A. Gualandris, D. Merritt, and S. Mikkola. A hybrid N-body code incorporating algorithmic regularization and post-Newtonian forces. *MNRAS*, 389:2–12, September 2008. doi: 10.1111/j.1365-2966.2008.13557.x.
- N. Häring and H.-W. Rix. On the Black Hole Mass-Bulge Mass Relation. *ApJ Lett.*, 604:L89–L92, April 2004. doi: 10.1086/383567.
- K. Hayasaki. A New Mechanism for Massive Binary Black-Hole Evolution. *PASJ*, 61:65–74, February 2009. doi: 10.1093/pasj/61.1.65.
- K. Hayasaki, S. Mineshige, and H. Sudou. Binary Black Hole Accretion Flows in Merged Galactic Nuclei. *PASJ*, 59:427–441, April 2007.
- D. Heggie and P. Hut. *The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics*. Cambridge University Press (1886), February 2003.
- D. C. Heggie and R. D. Mathieu. Standardised Units and Time Scales. In P. Hut and S. L. W. McMillan, editors, *The Use of Supercomputers in Stellar Dynamics*, volume 267 of *Lecture Notes in Physics*, Berlin Springer Verlag, page 233, 1986. doi: 10.1007/BFb0116419.
- M. H. Hénon. The Monte Carlo Method (Papers appear in the Proceedings of IAU Colloquium No. 10 Gravitational N-Body Problem (ed. by Myron Lecar), R. Reidel Publ. Co., Dordrecht-Holland.). *A&AS*, 14:151–167, November 1971. doi: 10.1007/BF00649201.
- L. Hernquist. An analytical model for spherical galaxies and bulges. *ApJ*, 356: 359–364, June 1990. doi: 10.1086/168845.

- A. Hobbs, S. Nayakshin, C. Power, and A. King. Feeding supermassive black holes through supersonic turbulence and ballistic accretion. *MNRAS*, 413: 2633–2650, June 2011. doi: 10.1111/j.1365-2966.2011.18333.x.
- G. Hobbs, A. Archibald, Z. Arzoumanian, D. Backer, M. Bailes, N. D. R. Bhat, M. Burgay, S. Burke-Spolaor, D. Champion, I. Cognard, W. Coles, J. Cordes, P. Demorest, G. Desvignes, R. D. Ferdman, L. Finn, P. Freire, M. Gonzalez, J. Hessels, A. Hotan, G. Janssen, F. Jenet, A. Jessner, C. Jordan, V. Kaspi, M. Kramer, V. Kondratiev, J. Lazio, K. Lazaridis, K. J. Lee, Y. Levin, A. Lommen, D. Lorimer, R. Lynch, A. Lyne, R. Manchester, M. McLaughlin, D. Nice, S. Osłowski, M. Pilia, A. Possenti, M. Purver, S. Ransom, J. Reynolds, S. Sanidas, J. Sarkissian, A. Sesana, R. Shannon, X. Siemens, I. Stairs, B. Stappers, D. Stinebring, G. Theureau, R. van Haasteren, W. van Straten, J. P. W. Verbiest, D. R. B. Yardley, and X. P. You. The International Pulsar Timing Array project: using pulsars as a gravitational wave detector. *Classical and Quantum Gravity*, 27(8):084013, April 2010. doi: 10.1088/0264-9381/27/8/084013.
- E. Holmberg. On the Clustering Tendencies among the Nebulae. II. a Study of Encounters Between Laboratory Models of Stellar Systems by a New Integration Procedure. *ApJ*, 94:385, November 1941. doi: 10.1086/144344.
- P. F. Hopkins and E. Quataert. How do massive black holes get their gas? *MNRAS*, 407:1529–1564, September 2010. doi: 10.1111/j.1365-2966.2010.17064.x.
- P. F. Hopkins, L. Hernquist, C. C. Hayward, and D. Narayanan. Why are active galactic nuclei and host galaxies misaligned? *MNRAS*, 425:1121–1128, September 2012. doi: 10.1111/j.1365-2966.2012.21449.x.
- F. Hoyle and W. A. Fowler. On the nature of strong radio sources. *MNRAS*, 125:169–+, 1963.
- P. Hut. The Starlab Environment for Dense Stellar Systems. In J. Makino and P. Hut, editors, *Astrophysical Supercomputing using Particle Simulations*, volume 208 of *IAU Symposium*, page 331, 2003.
- S. Inagaki and P. Wiyanto. On equipartition of kinetic energies in two-component star clusters. *PASJ*, 36:391–402, 1984.
- P. B. Ivanov, J. C. B. Papaloizou, and A. G. Polnarev. The evolution of a supermassive binary caused by an accretion disc. *MNRAS*, 307:79–90, July 1999. doi: 10.1046/j.1365-8711.1999.02623.x.
- S. Kazantzidis, L. Mayer, M. Colpi, P. Madau, V. P. Debattista, J. Wadsley, J. Stadel, T. Quinn, and B. Moore. The Fate of Supermassive Black Holes and the Evolution of the $M_{BH}-\sigma$ Relation in Merging Galaxies: The Effect of Gaseous Dissipation. *ApJ Lett.*, 623:L67–L70, April 2005. doi: 10.1086/430139.
- F. M. Khan, A. Just, and D. Merritt. Efficient Merger of Binary Supermassive Black Holes in Merging Galaxies. *ApJ*, 732:89, May 2011. doi: 10.1088/0004-637X/732/2/89.
- F. M. Khan, I. Berentzen, P. Berczik, A. Just, L. Mayer, K. Nitadori, and S. Callegari. Formation and Hardening of Supermassive Black Hole Binaries in Minor Mergers of Disk Galaxies. *ApJ*, 756:30, September 2012. doi: 10.1088/0004-637X/756/1/30.

- F. M. Khan, K. Holley-Bockelmann, P. Berczik, and A. Just. Supermassive Black Hole Binary Evolution in Axisymmetric Galaxies: The Final Parsec Problem is Not a Problem. *ApJ*, 773:100, August 2013. doi: 10.1088/0004-637X/773/2/100.
- S. S. Kim and H. M. Lee. Core-collapse times of two-component star clusters. *Journal of Korean Astronomical Society*, 30:115–122, October 1997.
- S. S. Kim, H. M. Lee, and J. Goodman. Two-Component Fokker-Planck Models for the Evolution of Isolated Globular Clusters. *ApJ*, 495:786–+, March 1998.
- A. R. King and J. E. Pringle. Growing supermassive black holes by chaotic accretion. *MNRAS*, 373:L90–L92, November 2006. doi: 10.1111/j.1745-3933.2006.00249.x.
- A. R. King, S. H. Lubow, G. I. Ogilvie, and J. E. Pringle. Aligning spinning black holes and accretion discs. *MNRAS*, 363:49–56, October 2005. doi: 10.1111/j.1365-2966.2005.09378.x.
- W. Kley. Mass flow and accretion through gaps in accretion discs. *MNRAS*, 303:696–710, March 1999. doi: 10.1046/j.1365-8711.1999.02198.x.
- B. Kocsis, Z. Haiman, and A. Loeb. Gas pile-up, gap overflow and Type 1.5 migration in circumbinary discs: application to supermassive black hole binaries. *MNRAS*, 427:2680–2700, December 2012. doi: 10.1111/j.1365-2966.2012.22118.x.
- S. Konstantinidis and K. D. Kokkotas. MYRIAD: a new N-body code for simulations of star clusters. *A&A*, 522:A70, November 2010. doi: 10.1051/0004-6361/200913890.
- J. Kormendy. Evidence for a supermassive black hole in the nucleus of M31. *ApJ*, 325:128–141, February 1988. doi: 10.1086/165988.
- J. Kormendy. The Stellar-Dynamical Search for Supermassive Black Holes in Galactic Nuclei. In L. Ho, editor, “*Coevolution of Black Holes and Galaxies*”, *Carnegie Observatories, Pasadena*, 2003. astro-ph/0306353.
- J. Kormendy and L. C. Ho. Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. *araa*, 51:511–653, August 2013. doi: 10.1146/annurev-astro-082708-101811.
- J. Kormendy and D. Richstone. Evidence for a supermassive black hole in NGC 3115. *ApJ*, 393:559–578, July 1992. doi: 10.1086/171528.
- J. Kormendy and D. Richstone. Inward Bound—The Search For Supermassive Black Holes In Galactic Nuclei. *ARA&A*, 33:581–+, 1995.
- J. Kormendy, R. Bender, A. S. Evans, and D. Richstone. The Mass Distribution in the Elliptical Galaxy NGC 3377: Evidence for a $2 \times 10^8 M_{\odot}$ Black Hole. *AJ*, 115:1823–1839, May 1998. doi: 10.1086/300313.
- P. Kroupa. On the variation of the initial mass function. *MNRAS*, 322:231–246, April 2001.
- G. Kupa, P. Amaro-Seoane, and R. Spurzem. Dynamics of compact object clusters: a post-Newtonian study. *MNRAS*, 371:L45–L49, September 2006. doi: 10.1111/j.1745-3933.2006.00205.x.

- A. H. W. Küpper, T. Maschberger, P. Kroupa, and H. Baumgardt. Mass segregation and fractal substructure in young massive clusters - I. The McLuster code and method calibration. *MNRAS*, 417:2300–2317, November 2011. doi: 10.1111/j.1365-2966.2011.19412.x.
- P. E. Kustaanheimo and E. L. Stiefel. Perturbation theory of Kepler motion based on spinor regularization. *J. Reine Angew. Math.*, 218:204–219, 1965.
- R. B. Larson. Turbulence and star formation in molecular clouds. *MNRAS*, 194:809–826, March 1981. doi: 10.1093/mnras/194.4.809.
- S. L. Liebling and C. Palenzuela. Dynamical Boson Stars. *Living Reviews in Relativity*, 15:6, May 2012. doi: 10.12942/lrr-2012-6.
- D. N. C. Lin and J. Papaloizou. On the tidal interaction between protoplanets and the primordial solar nebula. II - Self-consistent nonlinear interaction. *ApJ*, 307:395–409, August 1986. doi: 10.1086/164426.
- G. Lodato, S. Nayakshin, A. R. King, and J. E. Pringle. Black hole mergers: can gas discs solve the ‘final parsec’ problem? *MNRAS*, 398:1392–1402, September 2009. doi: 10.1111/j.1365-2966.2009.15179.x.
- A. Lupi, F. Haardt, and M. Dotti. Massive black hole and gas dynamics in galaxy nuclei mergers - I. Numerical implementation. *MNRAS*, 446:1765–1774, January 2015a. doi: 10.1093/mnras/stu2223.
- A. Lupi, F. Haardt, M. Dotti, and M. Colpi. Massive black hole and gas dynamics in mergers of galaxy nuclei - II. Black hole sinking in star-forming nuclear discs. *MNRAS*, 453:3437–3446, November 2015b. doi: 10.1093/mnras/stv1920.
- D. Lynden-Bell. Galactic Nuclei as Collapsed Old Quasars. *Nat*, 223:690–+, 1969.
- D. Lynden-Bell. Gravity power. *Physica Scripta*, 17:185–191, March 1978. doi: 10.1088/0031-8949/17/3/009.
- D. Lynden-Bell and M. J. Rees. On quasars, dust and the galactic centre. *MNRAS*, 152:461–+, 1971.
- D. Lynden-Bell and R. Wood. The gravo-thermal catastrophe in isothermal spheres and the onset of red-giant structure for stellar systems. *MNRAS*, 138:495, 1968. doi: 10.1093/mnras/138.4.495.
- F. M. Maccagni, R. Morganti, T. A. Oosterloo, J. B. R. Oonk, and B. H. C. Emonts. ALMA observations of AGN fuelling: the case of PKS B1718-649. *ArXiv e-prints*, January 2018.
- A. I. MacFadyen and M. Milosavljević. An Eccentric Circumbinary Accretion Disk and the Detection of Binary Massive Black Holes. *ApJ*, 672:83-93, January 2008. doi: 10.1086/523869.
- J. Magorrian, S. Tremaine, D. Richstone, R. Bender, G. Bower, A. Dressler, S. M. Faber, K. Gebhardt, R. Green, C. Grillmair, J. Kormendy, and T. Lauer. The Demography of Massive Dark Objects in Galaxy Centers. *The Astronomical Journal*, 115:2285–2305, June 1998. doi: 10.1086/300353.
- J. Makino. Optimal order and time-step criterion for Aarseth-type N-body integrators. *ApJ*, 369:200–212, March 1991. doi: 10.1086/169751.

- J. Makino. Grape-6. *Highlights in Astronomy*, 11:597, 1998.
- J. Makino and S. J. Aarseth. On a hermite integrator with ahmad-cohen scheme for gravitational many-body problems. *PASJ*, 44:141–151, apr 1992.
- J. Makino and M. Taiji. *Scientific simulations with special-purpose computers : The GRAPE systems*. Scientific simulations with special-purpose computers : The GRAPE systems /by Junichiro Makino & Makoto Taiji. Chichester ; Toronto : John Wiley & Sons, c1998., 1998.
- E. Maoz. A Stringent Constraint on Alternatives to a Massive Black Hole at the Center of NGC 4258. *ApJ Lett.*, 447:L91, July 1995. doi: 10.1086/309574.
- E. Maoz. Dynamical Constraints on Alternatives to Supermassive Black Holes in Galactic Nuclei. *ApJ Lett.*, 494:L181–L184, February 1998. doi: 10.1086/311194.
- F. Masset. FARGO: A fast eulerian transport algorithm for differentially rotating disks. *A&AS*, 141:165–173, January 2000. doi: 10.1051/aas:2000116.
- L. Mayer. Massive black hole binaries in gas-rich galaxy mergers; multiple regimes of orbital decay and interplay with gas inflows. *Classical and Quantum Gravity*, 30(24):244008, December 2013. doi: 10.1088/0264-9381/30/24/244008.
- L. Mayer, S. Kazantzidis, P. Madau, M. Colpi, T. Quinn, and J. Wadsley. Rapid Formation of Supermassive Black Hole Binaries in Galaxy Mergers with Gas. *Science*, 316:1874, June 2007. doi: 10.1126/science.1141858.
- D. Merritt and M. Milosavljević. Massive Black Hole Binary Evolution. *Living Reviews in Relativity*, 8, November 2005. doi: 10.12942/lrr-2005-8.
- D. Merritt, J. D. Schnittman, and S. Komossa. Hypercompact Stellar Systems Around Recoiling Supermassive Black Holes. *ApJ*, 699:1690–1710, July 2009. doi: 10.1088/0004-637X/699/2/1690.
- P. G. Mezger, W. J. Duschl, and R. Zylka. The Galactic Center: a laboratory for AGN? *The Astronomy and Astrophysics Review*, 7:289–388, 1996. doi: 10.1007/s001590050007.
- M. Milosavljević and D. Merritt. Long-Term Evolution of Massive Black Hole Binaries. *ApJ*, 596:860–878, October 2003.
- M. Milosavljević and D. Merritt. Formation of Galactic Nuclei. *ApJ*, 563:34–62, December 2001. doi: 10.1086/323830.
- J. J. Monaghan and R. A. Gingold. Shock Simulation by the Particle Method SPH. *Journal of Computational Physics*, 52:374–389, November 1983. doi: 10.1016/0021-9991(83)90036-0.
- J. J. Monaghan and J. C. Lattanzio. A refined particle method for astrophysical problems. *A&A*, 149:135–143, August 1985.
- C. Montuori, M. Dotti, M. Colpi, R. Decarli, and F. Haardt. Search for sub-parsec massive binary black holes through line diagnosis. *MNRAS*, 412: 26–32, March 2011. doi: 10.1111/j.1365-2966.2010.17888.x.
- C. Montuori, M. Dotti, F. Haardt, M. Colpi, and R. Decarli. Search for sub-parsec massive binary black holes through line diagnosis - II. *MNRAS*, 425: 1633–1639, September 2012. doi: 10.1111/j.1365-2966.2012.21530.x.

- M. P. Muno, F. K. Baganoff, M. W. Bautz, E. D. Feigelson, G. P. Garmire, M. R. Morris, S. Park, G. R. Ricker, and L. K. Townsley. Diffuse X-Ray Emission in a Deep Chandra Image of the Galactic Center. *ApJ*, 613:326–342, September 2004. doi: 10.1086/422865.
- T. Naab and A. Burkert. Gas Dynamics and Inflow in Gas-Rich Galaxy Mergers. In J. H. Knapen, J. E. Beckman, I. Shlosman, and T. J. Mahoney, editors, *The Central Kiloparsec of Starbursts and AGN: The La Palma Connection*, volume 249 of *Astronomical Society of the Pacific Conference Series*, page 735, 2001.
- Hubert Nguyen. *Gpu gems 3*. Addison-Wesley Professional, first edition, 2007. ISBN 9780321545428.
- K. Nitadori and S. J. Aarseth. Accelerating NBODY6 with graphics processing units. *MNRAS*, 424:545–552, July 2012. doi: 10.1111/j.1365-2966.2012.21227.x.
- K. Nitadori and J. Makino. Sixth- and eighth-order Hermite integrator for N-body simulations. *New Astronomy*, 13:498–507, October 2008. doi: 10.1016/j.newast.2008.01.010.
- Keigo Nitadori. *New approaches to high-performance N-body simulations with high-order integrator, new parallel algorithm, and efficient use of SIMD hardware*. PhD thesis, University of Tokyo, 2009.
- C. Nixon and S. H. Lubow. Resonances in retrograde circumbinary discs. *MNRAS*, 448:3472–3483, April 2015. doi: 10.1093/mnras/stv166.
- C. J. Nixon, P. J. Cossins, A. R. King, and J. E. Pringle. Retrograde accretion and merging supermassive black holes. *MNRAS*, 412:1591–1598, April 2011a. doi: 10.1111/j.1365-2966.2010.17952.x.
- C. J. Nixon, A. R. King, and J. E. Pringle. The final parsec problem: aligning a binary with an external accretion disc. *MNRAS*, 417:L66–L69, October 2011b. doi: 10.1111/j.1745-3933.2011.01121.x.
- S. C. Noble, B. C. Mundim, H. Nakano, J. H. Krolik, M. Campanelli, Y. Zlochower, and N. Yunes. Circumbinary Magnetohydrodynamic Accretion into Inspiring Binary Black Holes. *ApJ*, 755:51, August 2012. doi: 10.1088/0004-637X/755/1/51.
- K. Ohsuga, M. Mori, T. Nakamoto, and S. Mineshige. Supercritical Accretion Flows around Black Holes: Two-dimensional, Radiation Pressure-dominated Disks with Photon Trapping. *ApJ*, 628:368–381, July 2005. doi: 10.1086/430728.
- E. C. Ostriker. Dynamical Friction in a Gaseous Medium. *ApJ*, 513:252–258, March 1999. doi: 10.1086/306858.
- P. J. E. Peebles. Star Distribution Near a Collapsed Object. *ApJ*, 178:371–376, December 1972. doi: 10.1086/151797.
- V. Perret, F. Renaud, B. Epinat, P. Amram, F. Bournaud, T. Contini, R. Teyssier, and J.-C. Lambert. Evolution of the mass, size, and star formation rate in high redshift merging galaxies. MIRAGE - A new sample of simulations with detailed stellar feedback. *A&A*, 562:A1, February 2014. doi: 10.1051/0004-6361/201322395.
- P. C. Peters. Gravitational Radiation and the Motion of Two Point Masses. *Physical Review*, 136:1224–1232, November 1964.

- H. C. Plummer. On the problem of distribution in globular star clusters. *MNRAS*, 71:460–470, March 1911.
- S. F. Portegies Zwart, S. L. W. McMillan, P. Hut, and J. Makino. Star cluster ecology - IV. Dissection of an open star cluster: photometry. *MNRAS*, 321: 199–226, February 2001.
- S. F. Portegies Zwart, R. G. Belleman, and P. M. Geldof. High-performance direct gravitational N-body simulations on graphics processing units. *New Astronomy*, 12:641–650, November 2007. doi: 10.1016/j.newast.2007.05.004.
- W. H. Press. Techniques and tricks for n -body computation. In P. Hut and S. L. W. McMillan, editors, *The Use of Supercomputers in Stellar Dynamics*, page 184. Springer-Verlag, 1986.
- M. Preto and P. Amaro-Seoane. On Strong Mass Segregation Around a Massive Black Hole: Implications for Lower-Frequency Gravitational-Wave Astrophysics. *ApJ Lett.*, 708:L42–L46, January 2010. doi: 10.1088/2041-8205/708/1/L42.
- M. Preto, I. Berentzen, P. Berczik, and R. Spurzem. Fast Coalescence of Massive Black Hole Binaries from Mergers of Galactic Nuclei: Implications for Low-frequency Gravitational-wave Astrophysics. *ApJ Lett.*, 732:L26, May 2011. doi: 10.1088/2041-8205/732/2/L26.
- J. Prieto, A. Escala, M. Volonteri, and Y. Dubois. How AGN and SN Feedback Affect Mass Transport and Black Hole Growth in High-redshift Galaxies. *ApJ*, 836:216, February 2017. doi: 10.3847/1538-4357/aa5be5.
- G. D. Quinlan. The dynamical evolution of massive black hole binaries I. Hardening in a fixed stellar background. *New Astronomy*, 1:35–56, July 1996. doi: 10.1016/S1384-1076(96)00003-6.
- G. H. Rieke and M. J. Rieke. Stellar velocities and the mass distribution in the Galactic center. *ApJ Lett.*, 330:L33–L37, July 1988. doi: 10.1086/185199.
- C. Roedig and A. Sesana. Migration of massive black hole binaries in self-gravitating discs: retrograde versus prograde. *MNRAS*, 439:3476–3489, April 2014. doi: 10.1093/mnras/stu194.
- C. Roedig, M. Dotti, A. Sesana, J. Cuadra, and M. Colpi. Limiting eccentricity of subparsec massive black hole binaries surrounded by self-gravitating gas discs. *MNRAS*, 415:3033–3041, August 2011. doi: 10.1111/j.1365-2966.2011.18927.x.
- C. Roedig, A. Sesana, M. Dotti, J. Cuadra, P. Amaro-Seoane, and F. Haardt. Evolution of binary black holes in self-gravitating discs. Dissecting the torques. *A&A*, 545:A127, September 2012. doi: 10.1051/0004-6361/201219986.
- R. Roškar, R. Teyssier, O. Agertz, M. Wetzstein, and B. Moore. A systematic look at the effects of radiative feedback on disc galaxy formation. *MNRAS*, 444:2837–2853, November 2014. doi: 10.1093/mnras/stu1548.
- R. Roškar, D. Fiacconi, L. Mayer, S. Kazantzidis, T. R. Quinn, and J. Wadsley. Orbital decay of supermassive black hole binaries in clumpy multiphase merger remnants. *mn*, 449:494–505, May 2015. doi: 10.1093/mnras/stv312.

- E. E. Salpeter. Accretion of Interstellar Matter by Massive Objects. *ApJ*, 140: 796–800, August 1964.
- D. B. Sanders and I. F. Mirabel. Luminous Infrared Galaxies. *ARA&A*, 34:749, 1996. doi: 10.1146/annurev.astro.34.1.749.
- M. Schmidt. 3C 273 : A Star-Like Object with Large Red-Shift. *Nat*, 197:1040, March 1963. doi: 10.1038/1971040a0.
- J. Schneider, P. Amaro-Seoane, and R. Spurzem. Higher-order moment models of dense stellar systems: applications to the modelling of the stellar velocity distribution function. *MNRAS*, 410:432–454, January 2011. doi: 10.1111/j.1365-2966.2010.17454.x.
- J. D. Schnittman and J. H. Krolik. Evolution of a Binary Black Hole with a Retrograde Circumbinary Accretion Disk. *apj*, 806:88, June 2015. doi: 10.1088/0004-637X/806/1/88.
- R. Schödel, T. Ott, R. Genzel, A. Eckart, N. Mouawad, and T. Alexander. Stellar Dynamics in the Central Arcsecond of Our Galaxy. *ApJ*, 596:1015–1034, October 2003. doi: 10.1086/378122.
- R. Schödel, A. Feldmeier, D. Kunneriath, S. Stolovy, N. Neumayer, P. Amaro-Seoane, and S. Nishiyama. Surface brightness profile of the Milky Way’s nuclear star cluster. *A&A*, 566:A47, June 2014. doi: 10.1051/0004-6361/201423481.
- R. Schödel, E. Gallego-Cano, H. Dong, F. Nogueras-Lara, A. T. Gallego-Calvente, P. Amaro-Seoane, and H. Baumgardt. The distribution of stars around the Milky Way’s central black hole. II. Diffuse light from sub-giants and dwarfs. *A&A*, 609:A27, January 2018. doi: 10.1051/0004-6361/201730452.
- F. E. Schunck and E. W. Mielke. TOPICAL REVIEW: General relativistic boson stars. *Classical and Quantum Gravity*, 20:R301–R356, October 2003. doi: 10.1088/0264-9381/20/20/201.
- A. Sesana and F. M. Khan. Scattering experiments meet N-body - I. A practical recipe for the evolution of massive black hole binaries in stellar environments. *MNRAS*, 454:L66–L70, November 2015. doi: 10.1093/mnras/slv131.
- A. Sesana, F. Haardt, and P. Madau. Interaction of Massive Black Hole Binaries with Their Stellar Environment. II. Loss Cone Depletion and Binary Orbital Decay. *ApJ*, 660:546–555, May 2007. doi: 10.1086/513016.
- A. Sesana, C. Roedig, M. T. Reynolds, and M. Dotti. Multimessenger astronomy with pulsar timing and X-ray observations of massive black hole binaries. *MNRAS*, 420:860–877, February 2012. doi: 10.1111/j.1365-2966.2011.20097.x.
- A. Sesana, E. Barausse, M. Dotti, and E. M. Rossi. Linking the Spin Evolution of Massive Black Holes to Galaxy Kinematics. *ApJ*, 794:104, October 2014. doi: 10.1088/0004-637X/794/2/104.
- Y. Shen and A. Loeb. Identifying Supermassive Black Hole Binaries with Broad Emission Line Diagnosis. *ApJ*, 725:249–260, December 2010. doi: 10.1088/0004-637X/725/1/249.

- M. Spera, M. Mapelli, and A. Bressan. The mass spectrum of compact remnants from the PARSEC stellar evolution tracks. *MNRAS*, 451:4086–4103, August 2015. doi: 10.1093/mnras/stv1161.
- L. Spitzer. *Dynamical evolution of globular clusters*. Princeton, NJ, Princeton University Press, 1987, 191 p., 1987.
- L. J. Spitzer and M. H. Hart. Random gravitational encounters and the evolution of spherical systems. II. models. *ApJ*, 166:483, June 1971.
- V. Springel and L. Hernquist. Cosmological smoothed particle hydrodynamics simulations: the entropy equation. *MNRAS*, 333:649–664, July 2002. doi: 10.1046/j.1365-8711.2002.05445.x.
- R. Spurzem. Direct N-body Simulations. *Journal of Computational and Applied Mathematics*, 109:407–432, September 1999.
- G. Taffoni, L. Mayer, M. Colpi, and F. Governato. On the life and death of satellite haloes. *MNRAS*, 341:434–448, May 2003. doi: 10.1046/j.1365-8711.2003.06395.x.
- M. Taiji, J. Makino, T. Fukushige, T. Ebisuzaki, and D. Sugimoto. Grape-4: A teraflops machine for n-body simulations. In Piet Hut and Junichiro Makino, editors, *IAU Symp. 174: Dynamical Evolution of Star Clusters: Confrontation of Theory and Observations*, page 141, 1996.
- T. Tanaka, K. Menou, and Z. Haiman. Electromagnetic counterparts of supermassive black hole binaries resolved by pulsar timing arrays. *MNRAS*, 420:705–719, February 2012. doi: 10.1111/j.1365-2966.2011.20083.x.
- Y. Tang, A. MacFadyen, and Z. Haiman. On the orbital evolution of supermassive black hole binaries with circumbinary accretion discs. *MNRAS*, 469:4258–4267, August 2017. doi: 10.1093/mnras/stx1130.
- F. Tombesi, R. M. Sambruna, J. N. Reeves, V. Braitto, L. Ballo, J. Gofford, M. Cappi, and R. F. Mushotzky. Discovery of Ultra-fast Outflows in a Sample of Broad-line Radio Galaxies Observed with Suzaku. *ApJ*, 719:700–715, August 2010. doi: 10.1088/0004-637X/719/1/700.
- F. Tombesi, M. Meléndez, S. Veilleux, J. N. Reeves, E. González-Alfonso, and C. S. Reynolds. Wind from the black-hole accretion disk driving a molecular outflow in an active galaxy. *Nat*, 519:436–438, March 2015. doi: 10.1038/nature14261.
- J. L. Tonry. Evidence for a central mass concentration in M32. *ApJ Lett.*, 283:27–30, August 1984. doi: 10.1086/184326.
- J. L. Tonry. A central black hole in M32. *ApJ*, 322:632–642, November 1987. doi: 10.1086/165758.
- D. F. Torres, S. Capozziello, and G. Lambiase. Supermassive boson star at the galactic center? *Phys. Rev. D*, 62(10):104012, November 2000. doi: 10.1103/PhysRevD.62.104012.
- G. R. Tremblay, J. B. R. Oonk, F. Combes, P. Salomé, C. P. O’Dea, S. A. Baum, G. M. Voit, M. Donahue, B. R. McNamara, T. A. Davis, M. A. McDonald, A. C. Edge, T. E. Clarke, R. Galván-Madrid, M. N. Bremer, L. O. V. Edwards, A. C. Fabian, S. Hamer, Y. Li, A. Maury, H. R. Russell, A. C. Quillen, C. M. Urry, J. S. Sanders, and M. W. Wise. Cold, clumpy accretion onto an active

- supermassive black hole. *Nat*, 534:218–221, June 2016. doi: 10.1038/nature17969.
- P. Tsalmantza, R. Decarli, M. Dotti, and D. W. Hogg. A Systematic Search for Massive Black Hole Binaries in the Sloan Digital Sky Survey Spectroscopic Sample. *ApJ*, 738:20, September 2011. doi: 10.1088/0004-637X/738/1/20.
- S. Van Wassenhove, M. Volonteri, L. Mayer, M. Dotti, J. Bellovary, and S. Callegari. Observability of Dual Active Galactic Nuclei in Merging Galaxies. *ApJ Lett.*, 748:L7, March 2012. doi: 10.1088/2041-8205/748/1/L7.
- S. Van Wassenhove, P. R. Capelo, M. Volonteri, M. Dotti, J. M. Bellovary, L. Mayer, and F. Governato. Nuclear coups: dynamics of black holes in galaxy mergers. *MNRAS*, 439:474–487, March 2014. doi: 10.1093/mnras/stu024.
- E. Vasiliev, F. Antonini, and D. Merritt. The Final-parsec Problem in the Collisionless Limit. *ApJ*, 810:49, September 2015. doi: 10.1088/0004-637X/810/1/49.
- A. Vecchio, M. Colpi, and A. G. Polnarev. Orbital evolution of a massive black hole pair by dynamical friction. *ApJ*, 433:733–745, October 1994. doi: 10.1086/174683.
- M. Volonteri. Formation of supermassive black holes. *The Astronomy and Astrophysics Review*, 18:279–315, July 2010. doi: 10.1007/s00159-010-0029-x.
- M. Volonteri, F. Haardt, and P. Madau. The Assembly and Merging History of Supermassive Black Holes in Hierarchical Models of Galaxy Formation. *ApJ*, 582:559–573, January 2003. doi: 10.1086/344675.
- S. von Hoerner. Die numerische Integration des n-Körper-Problemes für Sternhaufen. I. *Z. Astrophys.*, 50:184–214, 1960.
- S. von Hoerner. Die numerische Integration des n-Körper-Problems für Sternhaufen, II. *Z. Astrophys.*, 57:47–82, 1963.
- L. Wang, R. Spurzem, S. Aarseth, K. Nitadori, P. Berczik, M. B. N. Kouwenhoven, and T. Naab. NBODY6++GPU: ready for the gravitational million-body problem. *MNRAS*, 450:4070–4080, July 2015. doi: 10.1093/mnras/stv817.
- L. Wang, R. Spurzem, S. Aarseth, M. Giersz, A. Askar, P. Berczik, T. Naab, R. Schadow, and M. B. N. Kouwenhoven. The DRAGON simulations: globular cluster evolution with a million stars. *MNRAS*, 458:1450–1465, May 2016. doi: 10.1093/mnras/stw274.
- Q. Yu. Evolution of massive binary black holes. *MNRAS*, 331:935, April 2002.
- Y. B. Zel’dovich. The Fate of a Star and the Evolution of Gravitational Energy Upon Accretion. *Soviet Physics Doklady*, 9:195, September 1964.