Cover page: *Artist's impression of the accretion disc around a supermassive black hole at the centre of an active galaxy. The black hole is rotating and emits a relativistic jet.*
Credit: NASA/Dana Berry, SkyWorks Digital

Alessandro Lupi

Black holes in galactic nuclei - seed formation from stellar mass black holes and massive black hole pairing in galaxy mergers.

SUPERVISOR:
Francesco Haardt

CO-SUPERVISORS:
Massimo Dotti and Monica Colpi

University of Insubria | Department of Science and Hi-Tech
Como, Italy

“Science is a way of life. Science is a perspective. Science is the process that takes us from confusion to understanding in a manner that’s precise, predictive and reliable - a transformation, for those lucky enough to experience it, that is empowering and emotional.”

– Brian Greene

Contents

Abstract

Black holes (BHs) are a very important class of astrophysical objects. They are the most compact objects in the Universe, hence they represent the most extreme sources of gravity. BHs come in two flavours: the stellar mass BHs (SBHs) relic of young massive stars ($1-20\,\mathrm{M}_\odot$) and the massive BHs (MBHs), with masses of $10^6-10^9\,\mathrm{M}_\odot$, dwelling in the nuclei of the most massive galaxies. While the formation mechanisms of SBHs are well understood, no clear consensus exists about MBH formation. According to the Soltan arguments (Soltan, 1982), MBHs gain the largest fraction of their mass via radiative efficient accretion of gas. As a consequence, we expect that MBH formed early in the Universe as smaller mass *seeds*. Recently, observations of high redshift quasars (e.g.; Mortlock et al., 2011; Fan et al., 2006) showed that MBHs with masses above $10^9\,\mathrm{M}_\odot$ were already in place when the Universe was less than 1 Gyr old and posed tight constraints on the models for the formation and growth of MBHs. Two main scenarios have been developed for MBH seed formation: the *light seed* scenario, where seeds formed as relic of the first generation of stars with masses of up to few hundred solar masses (Madau & Rees, 2001), and the *heavy seed* scenarios, where seeds formed from the direct collapse of massive gas clouds in primordial haloes with masses of up to few $10^5\,\mathrm{M}_\odot$ (Haehnelt & Rees, 1993). Despite the large number of studies about MBH formation models, each model still has its own caveats, which make the study of MBH formation worth of further investigations.

According to the Λ-CDM cosmology, galaxies form when gas cools down within dark matter haloes, which assembly in a hierarchical fashion from small density perturbations. Galaxies grow via accretion and mergers, and the central MBHs evolve in the same way. So, when a galaxy merger occur, the MBHs hosted in the nucleus of the galaxy progenitors can sink towards the centre of

the merger remnant, forming a MBH binary (MBHB). Despite galaxy mergers are usually observed, no clear detections of MBHBs exist to date. The formation and evolution of MBHBs is a complex process, since it occurs in a rapidly varying environment where gas, star formation and SNa feedback play a pivotal role. Several studies have been performed to date, but a clear understanding of the whole process is still far from being reached.

In this thesis I cover both aspects of MBH formation and evolution. In the first study I consider an alternative route for seed BH formation. Using two different codes, the AMR code RAMSES (Teyssier, 2002) and the mesh-free code GIZMO (Hopkins, 2015), I studied the evolution of a single massive circum-nuclear gaseous disc embedding a population of SBHs. The disc was subject to radiative cooling, star formation and supernova feedback and becomes unstable to fragmentation, which led to the formation of clumps as massive as $10^4 - 10^5\,\mathrm{M}_\odot$. My simulations showed that during the disc evolution, some SBHs can be gravitationally captured by a clump. Within the clumps, such BHs can experience episodes of super-critical accretion, which make them grow up to $10^3 - 10^4\,\mathrm{M}_\odot$ in few Myr. Thanks to the very low radiative efficiency associated to the slim accretion disc (Abramowicz et al., 1988), the energy released to the surrounding gas is too small to halt the accretion flow, hence BHs can accrete almost unimpeded until one of these events occur: the clump is totally accreted by the BH, the clump is consumed by star formation or the clump is destroyed by supernova explosions.

In the second study, instead, I consider the intermediate stages of a galaxy merger, when the MBHs originally dwelling in the centre of their own progenitor galaxies reach few hundred separations in the nucleus of the merger remnant. I assumed that each MBH was embedded in a self-gravitating circum-nuclear gaseous disc. With the code RAMSES I studied the evolution of the MBHs and their surrounding discs, including physical processes like radiative cooling, star formation and supernova feedback, which are implemented in the code as sub-grid recipes. First, I implemented a new refinement prescription aimed at improving the orbital evolution of massive particles, an already known major issue in AMR codes, like observed by Gabor & Bournaud (2013); Dubois et al. (2014). Secondly, I evolved the discs assuming different sub-grid recipes to study how the MBH and gas dynamics could be affected by the different choices. I found that the MBH dynamics is almost independent of the physical modelling, if one assumes that no previous star formation

occurred in the discs, while the gas evolution and its final distribution can be significantly affected. On the other side, if one assumes that star formation was already ongoing, even the BH dynamics can be modified, if supernovae are powerful enough to disrupt gas clumps forming in the discs.

A general introduction to the work is reported in Chapter 1. In Chapter 3 I discuss the first study about an alternative model for seed BH formation. In Chapter 4, instead, I describe the second study concerning the evolution of the MBH pair in the intermediate stages of a galaxy merger. The reader interested in the main results of the work can directly move to Chapters 3 and 4. Finally, Chapter 5 reports my conclusions.

Chapter 1

Introduction

Black holes (BHs) are a very important class of astrophysical objects and they are simple objects which can be described with only two parameters: mass and spin. They are the most compact objects in the Universe, hence they represent the most extreme sources of gravity. BHs come in two flavours: the stellar mass BHs (SBHs) relic of young massive stars ($1 - 20\,\mathrm{M}_\odot$) and the massive BHs (MBHs), with masses of $10^6 - 10^9\,\mathrm{M}_\odot$, dwelling in the nuclei of the most massive galaxies. The observed BH mass function can be represented as two disjoint intervals, as shown in Fig. 1.1 for the local population of BHs.

Since the discovery of **Cygnus X-1** (Bowyer et al., 1965), the first confirmed SBH, and **Sagittarius A*** , (Balick & Brown, 1974) , the $4 \times 10^6\,\mathrm{M}_\odot$ BH inhabiting our Galaxy, we got compelling evidences of the presence of this class of objects in our Universe. In particular, we know that all nearby massive galaxies (corresponding to $M_{\mathrm{galaxy}} \gtrsim 10^{11}\ \mathrm{M}_\odot$) house a supermassive BH (SMBH) in their core, while nothing can be said yet about less massive galaxies. SBHs are observed in X-ray binaries as they accrete gas from a companion star, while SMBHs shine as bright quasi-stellar objects (QSOs) or less luminous active galactic nuclei (AGNs). SMBHs are also observed during quiescent phases as massive dark objects in the centre of nearby galaxy spheroids (the Milky Way case; Kormendy & Ho, 2013).

The missing interval between 10^2 and $10^5\,\mathrm{M}_\odot$, namely that of Intermediate Mass BHs (IMBHs) is supposed to exist, but to date evidences of this BH population are still poor.

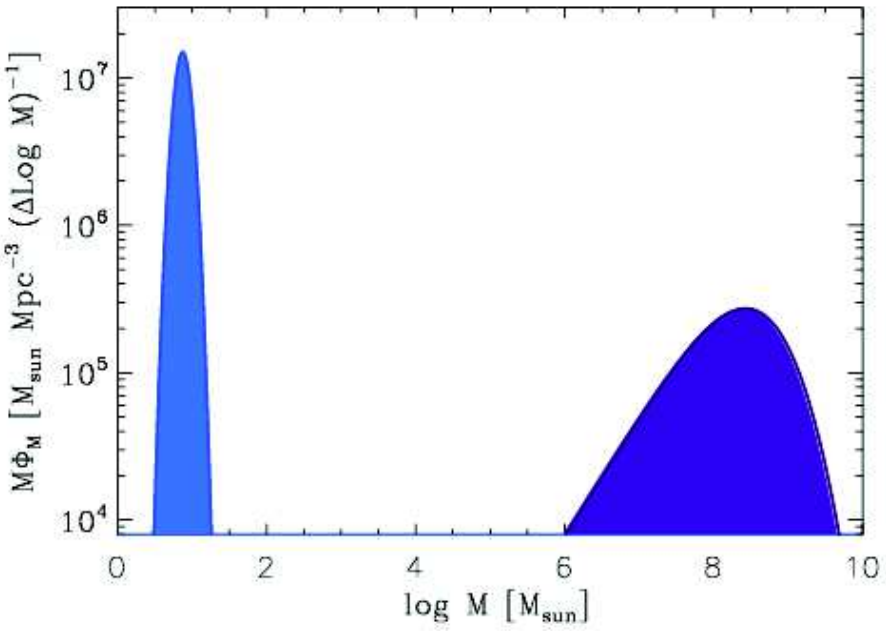

Figure 1.1: *Mass function of the local population of BHs. The SBH peak has been drawn assuming a log-normal distribution with mean mass equal to* $5\,\mathrm{M}_\odot$*, width of* 0.1 *dex and a normalisation yielding a density of about* $1.1 \times 10^7\,\mathrm{M}_\odot \cdot \mathrm{Mpc}^{-3}$ *(Fukugita & Peebles, 2004). The SMBH peak, instead, contributes to an overall density of about* $4.3 \times 10^5\,\mathrm{M}_\odot \cdot \mathrm{Mpc}^{-3}$*. Figure taken from Merloni (2008).*

The discovery of tight correlations between the BH mass and the stellar velocity dispersion[1] of the hot stellar component (see Fig. 1.2, Ferrarese & Merritt, 2000; Gültekin et al., 2009; McConnell et al., 2011) and between the BH mass and the host galaxy spheroidal component, i.e. $M_{\rm BH} \sim 10^{-3} M_{\rm spheroid}$ (Magorrian et al., 1998; Häring & Rix, 2004), suggests that over time galaxies and MBHs co-exist and undergo a symbiotic evolution.

Galaxies form following the baryonic infall of gas into collapsing dark matter haloes and build up their mass via accretion and mergers. During this hierarchical assembly of structure, MBHs co-evolve with their galaxy hosts through accretion of inflowing gas and coalescences driven by galaxy mergers. As a consequence, we expect that MBHs formed early in the Universe as smaller mass *seeds*.

In order to understand the origin of these correlations we should look at the early Universe. Recent observations demonstrate that MBHs must form very early and grow rapidly up to $10^9\,{\rm M}_\odot$ within 1 Gyr from the Big Bang, like in the case of ULAS J1120+0641 (Mortlock et al., 2011) and a handful of bright quasars observed by the *Sloan Digital Sky Survey* at $z \gtrsim 6$ (Fan et al., 2006), and set tight constraints on any model for the formation and growth of MBH at early epochs.

However, these massive quasars are only the high mass tail of the MBH mass function, and we totally lack a consensus about the low luminosity population, which would determine whether these correlations can be extended to lower mass galaxies (Treister et al., 2011; Fiore et al., 2012). According to these correlations, lower mass galaxies are expected to host lighter MBHs and would be the best candidates for the search of IMBHs and even of MBH seeds.

Thanks to recents efforts we have some hints about the evolution of the quasar and galaxy luminosity function with time, which strengthen the idea of a MBH-galaxy symbiotic evolution. Moreover, the direct formation of SMBHs from the collapse of a $\sim 10^9$ ${\rm M}_\odot$ gas cloud is also unlikely from theoretical arguments and observational constraints which indicate that gas dynamics in galaxies tends to halt the global collapse and to fragment into clouds that further fragment to form stars.

[1]The velocity dispersion is the root mean square of the stellar velocities and gives an estimate of the stellar motion due to the galaxy gravitational potential.

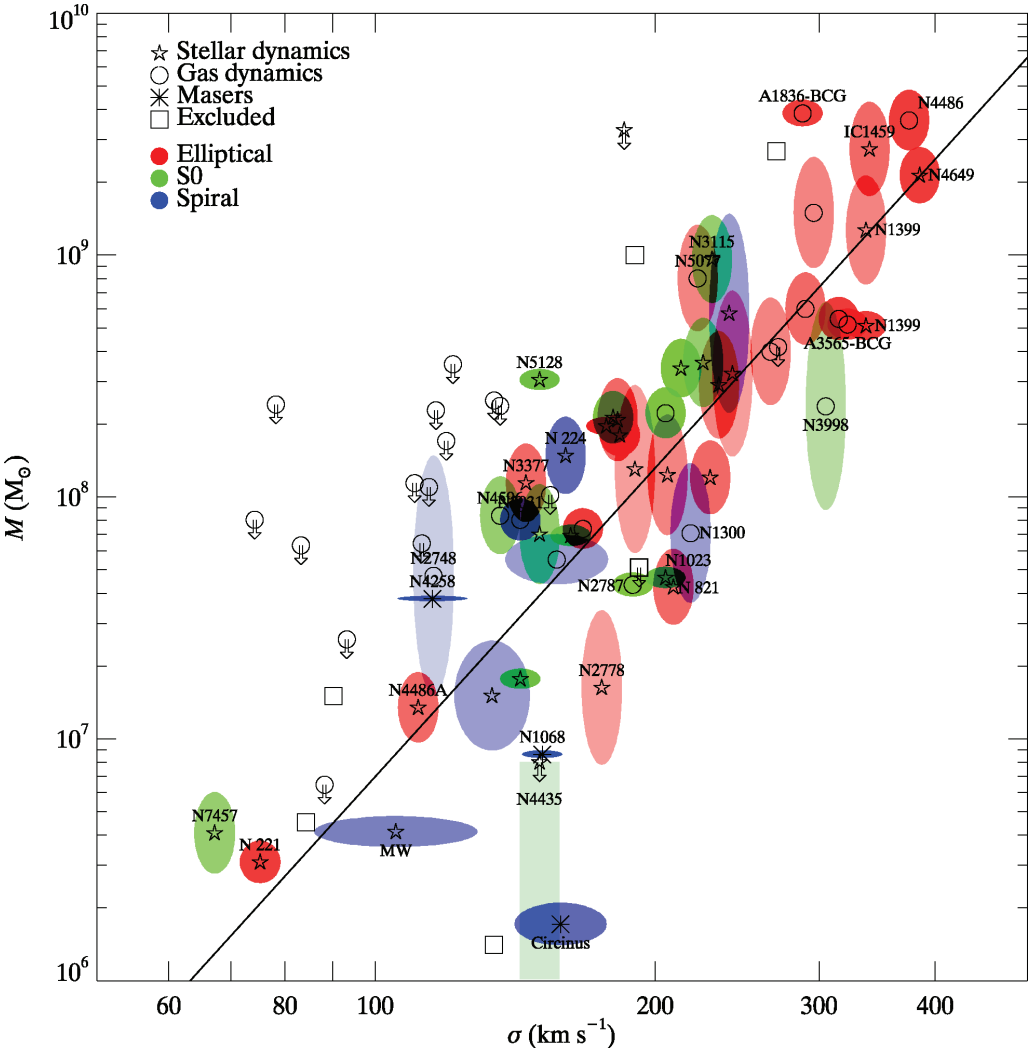

Figure 1.2: *$M - \sigma$ relation for galaxies with dynamical measurements from (Gültekin et al., 2009). The symbol indicates the method of BH mass measurement: stellar dynamical (pentagrams), gas dynamical (circles), masers (asterisks). Arrows indicate upper limits to BH mass. The colour of the error ellipse indicates the Hubble type of the host galaxy: elliptical (red), S0 (green), and spiral (blue). The saturation of the colours in the error ellipses or boxes is inversely proportional to the area of the ellipse or box. Squares are galaxies not included in the fit. The line is the best fit relation to the full sample:* $M_{\rm BH} = 10^{8.12}\,{\rm M}_\odot(\sigma/200\,{\rm km\,s^{-1}})^{4.24}$.

From all these evidences some crucial questions arise, concerning when and how MBH seed form, how efficiently a MBH grows in mass with time and what is the frequency of MBH within galaxies. While a good agreement on the formation mechanisms of the stellar population has been reached, with SBHs forming at the end of the life of stars with masses $\gtrsim 20\,\mathrm{M}_\odot$ or when an already formed neutron star living in a binary systems exceeds the critical mass for stability[2], the formation of MBH seeds is still poorly understood.

Several mechanisms have been proposed, from the collapse of the first generations of stars (PopIII stars) to dynamical processes in nuclear stellar clusters, from the runaway merger of SBHs to the direct collapse of a massive gas cloud. All these models can be summarised in the "flow chart", initially proposed by Rees (1978) and updated in Fig. 1.3 to comprehend more recent models.

1.1 MBH formation and growth

The observation of high redshift quasars up to redshift ~ 7 tells us that MBHs were already in place when the Universe was less than 1 Gyr old and implies masses larger than $10^9\,\mathrm{M}_\odot$ to explain the observed luminosities of $\sim 10^{48}\,\mathrm{erg/s}$ (Barth et al., 2003; Willott et al., 2005). This huge luminosity is produced through accretion of gas on to the BH and can be written as

$$L_{\mathrm{BH}} = \eta \dot{M}_{\mathrm{in}} c^2, \tag{1.1}$$

where $\dot{M}_{\mathrm{in}}$ is the gas accretion rate, c is the speed of light and η is an efficiency parameter corresponding to the fraction of the rest-mass energy of the inflowing gas which is radiated away. Typically it is assumed that accretion can occur only when the emitted luminosity is below the so-called 'Eddington limit', defined as the maximum luminosity for which radiation pressure does not overcome gravity. The Eddington luminosity can be written as

$$L_{\mathrm{Edd}} = \frac{4\pi c m_{\mathrm{p}} M_{\mathrm{BH}}}{\sigma_{\mathrm{T}}}, \tag{1.2}$$

[2]The critical mass for neutron stars, also known as the Tolman-Oppenheimer-Volkoff limit (**TOV** limit) is the maximum mass a neutron star can have to balance gravity with the neutron degeneration pressure. It is the analogous of the Chandrasekaar mass for white dwarves and it assumes values in the range $1.5 - 3\,\mathrm{M}_\odot$, where the uncertainties are related to the poorly known equation of state for very dense matter.

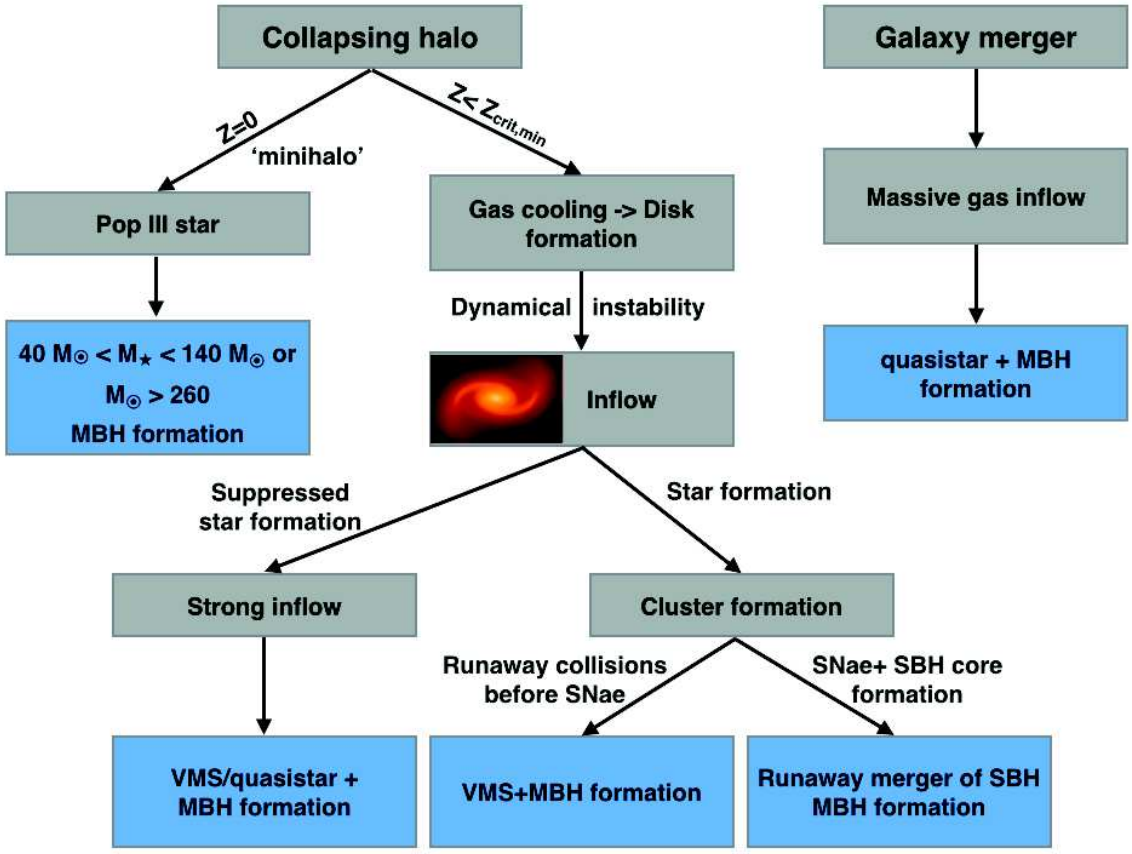

Figure 1.3: *Possible channels for MBH formation in isolated haloes (tree on the left) or through galaxy mergers (tree on the right). Updated version of the chart in Volonteri (2010).*

where $m_{\rm p}$ is the proton mass and $\sigma_{\rm T}$ is the Thomson scattering cross-section. Above this limit, gas is expelled instead of being accreted and the BH growth is halted. If one considers that only a fraction $1-\eta$ of the inflowing mass $M_{\rm in}$ is actually accreted, the mass growth rate can be written as $\frac{{\rm d}M_{\rm BH}}{{\rm d}t} = (1-\eta)\dot{M}_{\rm in}$. According to General Relativity, the expected radiative efficiency in the case of a test particle which is spiralling towards the BH on nearly circular orbits depends on the BH spin a, and varies between 0.057 for a Schwarzchild BH ($a = 0$) and 0.42 for a maximally rotating Kerr BH ($a = 1$). The typical assumed value is $\eta = 0.1$ (according to the Soltan argument; see section §1.1.5). Therefore, the mass growth for a BH accreting at the Eddington limit is an exponential function with an e-folding timescale $t_{\rm Salpeter} = \frac{\eta}{1-\eta}\frac{\sigma_{\rm T} c}{4\pi G m_{\rm p}} \sim 45$ Myr. Considering a BH with an initial mass M_0, the growth time is then defined as

$$t_{\rm growth} = t_{\rm Salpeter} \ln \frac{M_{\rm BH}}{M_0}. \tag{1.3}$$

A BH with an initial mass in the range $10^2 - 10^5\,{\rm M}_\odot$ would need ~ 0.5 Gyr to reach $10^9\,{\rm M}_\odot$, hence we need already massive seeds to explain the observed MBH at high redshift.

According to the current paradigm of the Λ-CDM cosmology, structures form via gravitational amplification of small perturbations in a cold dark matter, dark energy dominated Universe where 'baryons' account for 10% only of the total matter budget. During the cosmic history, the small dark matter perturbations grow to larger and larger scales in a hierarchical fashion, until they reach the critical point for collapse and virialization and form a self-gravitating halo. Baryons within these virialized haloes cool down and condense, eventually becoming dense enough to fragment into clumps which subsequently form stars.

This is the framework where MBH seeds form, and the mechanisms responsible for their formation necessarily depend on the halo conditions. For example, in a halo where star formation is suppressed, the gas could flow towards the centre forming an already massive object, while, in a halo with efficient star formation, one needs a dynamical process able to build a massive object from the coalescence of stellar objects (stars of SBHs) before a large fraction of the system is ejected in the so-called 'evaporation' process.

1.1.1 Pop III scenario

One of the most popular scenarios for MBH formation asserts that seeds form as remnants of Population III stars (PopIII), i.e. stars formed from zero metallicity gas[3] (Madau & Rees, 2001). These stars are expected to form at $z \sim 20 - 25$ in 'minihaloes', haloes with masses $\approx 10^6\,\mathrm{M}_\odot$ and a virial temperature of $\sim 10^3$ K (Tegmark et al., 1997; Madau & Rees, 2001; Haiman & Loeb, 2001; Volonteri, Haardt & Madau, 2003; Madau et al., 2004).

All models of PopIII star formation predict initial stellar masses larger than today's stars (up to few hundreds solar masses; Couchman & Rees, 1986; Omukai & Nishi, 1998; Bromm & Larson, 2004). The fate of these massive stars, if they retain all their mass until death, depends on their exact mass after a very short lifetime (i.e. few Myr). Fig. 1.4 shows the stellar remnant for different initial stellar masses and metallicities, as reported in Heger et al. (2003).

Low-metallicity stars between 25 and $140\,\mathrm{M}_\odot$ are expected to collapse into a BH via fallback of gas ejected by a faint SNa explosion. The corresponding mass of the BH would be between 10 and $40\,\mathrm{M}_\odot$. However, such small BHs would be dynamically unstable in the galaxy potential well and then would wander within its host, interacting with other stars, without settling in the galaxy centre.

More massive stars, between 140 and $260\,\mathrm{M}_\odot$ lie in the pair-instability regime, in which the production of electron-positron pairs would lead to the complete disruption of the star.

If stars with masses above $260\,\mathrm{M}_\odot$ exist, they would directly collapse into a BH without triggering a SNa explosion.

Though the first studies predicted the formation of these very massive stars, recent simulations including more detailed physics and higher resolution seemed to change this picture. Indeed, the inclusion of turbulence, radiative feedback and also new numerical techniques showed that fragmentation in the primordial gas is stronger than previously thought and would lead to the formation of binaries and/or clusters of smaller mass PopIII stars ($\lesssim 100\,\mathrm{M}_\odot$; Clark et al., 2011; Greif et al., 2011; Stacy, Greif & Bromm, 2012). Additional constraints come from the works of McKee & Tan (2008); Trenti, Stiavelli & Michael Shull (2009), where different conditions like the feedback produced

[3] All elements above Helium are globally considered as *metals*.

by the accretion of the envelope on to a pre-formed small mass core or the presence of a UV background field could limit the effective mass of the newly formed PopIII star.

If one consider Eddington-limited gas accretion, these lighter seeds could grow to the supermassive variety by $z \sim 6$ only if (i) gas accretion continued unimpeded at the Eddington rate for $\gtrsim 0.6$ Gyr, and if (ii) the mass-to-light conversion efficiency of the accretion process was not high, $\epsilon \lesssim 0.1$ (Tanaka & Haiman, 2009). The first condition seems hard to satisfy in the shallow potential of low-mass dark matter haloes, as radiative feedback from the progenitor and from BH accretion itself dramatically affects the gas inflow and its supply to the hole, resulting in sub-Eddington rates, therefore negligible mass growth (e.g., Wise, Turk & Abel, 2008; Milosavljević et al., 2009; Alvarez, Wise & Abel, 2009). The second condition is problematic too, as it requires a radiative efficiency well below that proper of accretion onto rapidly rotating black holes. Indeed, there are mounting evidences that the most massive holes at high redshifts power radio-loud AGNs (see, e.g., Ghisellini et al., 2014). These are thought to be associated with Kerr holes - though observational evidences of the widely accepted jet-spin connection are, at best, scarce, even in the well studied Galactic stellar black hole candidates (see, e.g., Russell, Gallo & Fender, 2013).

It has also been proposed that dark matter (as weakly-interacting massive particles, WIMPs) could affect the formation of the first stars in the centre of galaxies (Ripamonti, Mapelli & Ferrara, 2007; Iocco, 2008). The condition for this to occur is that the halo profile is sufficiently steep to provide an additional heating source through WIMP annihilation when dark matter is compressed by the inflowing gas and this would result in the formation of a ‘dark’ star (Freese, Spolyar & Aguirre, 2008), another MBH formation mechanism.

1.1.2 Direct collapse

Another important model to date is the direct collapse of a massive gas cloud into a single MBH (Haehnelt & Rees, 1993; Bromm & Loeb, 2003; Begelman, Volonteri & Rees, 2006; Lodato & Natarajan, 2006). The high density conditions required by this model can be only fulfilled in the centre of galaxies, and only in primordial haloes where fragmentation (highly enhanced by the efficient cooling of metals) is suppressed and the collection of large amount of

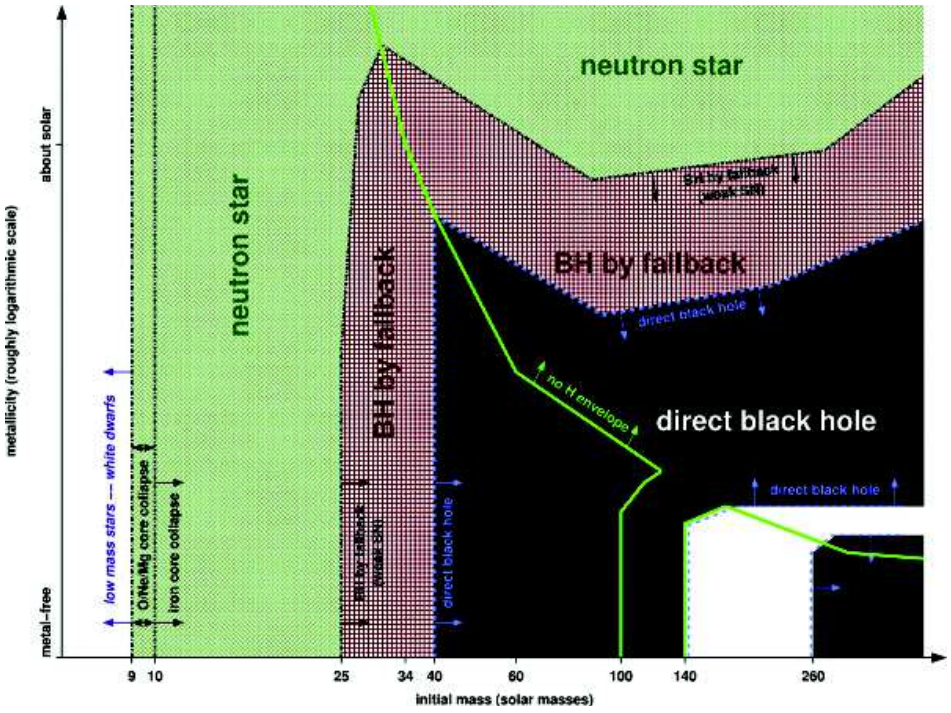

Figure 1.4: *Remnants of massive single stars as a function of initial metallicity ($y-axis$; qualitatively) and initial mass ($x-axis$). The thick green line separates the regimes where the stars keep their hydrogen envelope (left and lower right) from those where the hydrogen envelope is lost (upper right and small strip at the bottom between* 100 *and* $140\,\mathrm{M}_\odot$*). The dashed blue line indicates the border of the regime of direct black hole formation (black). This domain is interrupted by a strip of pair-instability supernovae that leave no remnant (white). Outside the direct black hole regime, at lower mass and higher metallicity, follows the regime of BH formation by fallback (red cross-hatching and bordered by a black dot-dashed line). Outside of this, green cross-hatching indicates the formation of neutron stars. The lowest mass neutron stars may be made by O/Ne/Mg core collapse instead of iron core collapse (vertical dot-dashed lines at the left). At even lower mass, the cores do not collapse and only white dwarfs are made (white strip at the very left). Figure taken from Heger et al. (2003)*

gas in galaxy nuclei is favoured. These haloes have typical virial temperature of $\sim 10^4$ K and a mass $M_{\rm halo} \sim 10^8\,{\rm M}_\odot$. However, in order to inhibit fragmentation and funnel this large amount of gas towards the centre, one needs a dissociating background able to dissociate molecular hydrogen (the other coolant responsible for fragmentation; Bromm & Loeb, 2003) and an efficient transfer of angular momentum.

Since the Lyman-Werner background is very high compared to the average value at the considered redshift, suppression can only occur if the halo lives very close to another halo where PopIII star formation produces the necessary flux (Dijkstra et al., 2008) or if the radiated Lyman α photons can be trapped within the high density collapsing cloud (Spaans & Silk, 2006).

Recent studies also suggested that in a highly turbulent medium fragmentation is automatically reduced, even in metal enriched systems (Begelman & Shlosman, 2009), hence loosing this constraint.

If the conditions to avoid dissociation are fulfilled, cooling via atomic hydrogen can continue until the temperature reaches ~ 4000 K and then contraction proceeds nearly adiabatically.

The second issue is the angular momentum barrier. Dark matter haloes, and their gas component, both posses angular momentum, which tends to counteract gravity, then halting the collapse and leading to the formation of a rotationally supported disc (Mo, Mao & White, 1998; Oh & Haiman, 2002). Some authors studied the direct collapse in haloes with very low angular momentum (Eisenstein & Loeb, 1995) or by considering only the low angular momentum material in efficiently cooling haloes (Koushiappas, Bullock & Dekel, 2004). However, even in these models, substantial angular momentum transport is required.

Another possible path has been proposed by Lodato & Natarajan (2006), where authors discuss how a marginally stable self-gravitating disc develops spiral structures which redistribute angular momentum, causing mass inflows without triggering extended fragmentation in the disc. The maximum inflow rate sustainable by the disc is $\dot{M}_{\rm in} = 2\alpha_{\rm c}\frac{c_{\rm s}^3}{G}$, where $\alpha_{\rm c} \sim 0.12$ is the viscosity.

The most appealing route to efficiently transfer angular momentum outwards is the so-called ‘bars within bars’ instability (Shlosman, Frank & Begelman, 1989; Begelman, Volonteri & Rees, 2006). Bar formation in galaxies is particularly efficient in transferring angular momentum outwards, through gravitational and hydrodynamical torques. Begelman, Volonteri & Rees (2006)

suggests that when gas exceeds a certain threshold a bar instability arises, which funnels the gas inwards. If the inflowing gas is able to cool, another phase of bar instability on smaller scales is triggered, and the process cascades. The most attractive aspect of this mechanism is the very short timescale required, i.e. the galaxy dynamical time. However, the inflow rate required by this cascade to emerge is very high, with values of roughly $1\,\mathrm{M}_\odot/\mathrm{yr}$ (e.g, Ferrara, Haardt & Salvaterra, 2013; Latif et al., 2015).

In this extreme conditions, the gas can trap its own radiation, forming a supermassive quasi-static object supported by radiation pressure, which burns hydrogen for about 1 Myr and grows up to $10^6\,\mathrm{M}_\odot$ (Begelman, Volonteri & Rees, 2006; Begelman, 2010) before collapsing into a BH. After the collapse, the BH, still embedded in the massive envelope accreted, starts swallowing the available gas reservoir at rates corresponding to the Eddington limit for the envelope (which are hundreds of times the Eddington limit for the BH itself). This object, named 'quasistar' appears as a red giant star with a luminosity comparable to that of an AGN. When accretion stops, the residual gas disperses, leaving behind a naked MBH with a typical mass between 10^4 and $10^5\,\mathrm{M}_\odot$.

If the gas inflow rate is lower, the evolution is different. Gas contracts forming a supermassive stars (SMS, with masses $\gtrsim 5 \cdot 10^4\,\mathrm{M}_\odot$). Uniformly rotating supermassive stars can evolve in isolation, cooling and contracting until the collapse conditions are reached (Baumgarte & Shapiro, 1999; Zeldovich & Novikov, 1971). Saijo et al. (2002) also included post-Newtonian correction to investigate whether the fast rotation can result in disc formation or if it the SMS can collapse into a BH. Numerical experiments of a maximally rotating SMS have been performed by Shibata & Shapiro (2002), who found a Kerr-like BH with 90% of the initial SMS mass.

Another possible route for direct collapse has been proposed by (Mayer et al., 2010, 2015), where a massive ($10^8\,\mathrm{M}_\odot$) unstable gas cloud forms at the end of a major merger between two gas-rich galaxies. Such a massive clouds could evolve as a VMS and then collapse in a MBH.

1.1.3 Stellar dynamics

Many galaxies host in the centre both a MBH and a nuclear stellar cluster (NSC). Though the link between them is not yet understood, one can con-

sider the formation of MBH seeds from the evolution of a pre-existent NSC at high redshift. As before described, one needs to prevent efficient star formation to form seeds via direct collapse. However, if one consider small minihaloes where PopIII stars form, at the time more massive haloes virialize the environment will have been enriched with metals released by SNe. In these conditions, efficient formation of PopII stars starts, possibly leading to the formation of a NSC of 'normal' low mass stars (Omukai, Schneider & Haiman, 2008). If the NSC is particularly compact, stellar collisions can lead to the formation of a very massive star (VMS), which could then evolve into a MBH seed with a mass in the range $10^2 - 10^4\,\mathrm{M}_\odot$ (Devecchi & Volonteri, 2009). Seed formation via stellar dynamical processes has been initially proposed by Begelman & Rees (1978) and then studied by several authors (e.g., Quinlan & Shapiro, 1987; Portegies Zwart & McMillan, 2002; Gürkan, Fregeau & Rasio, 2006).

The bottom line of all these models is the dynamical relaxation of the NSC. Two-body interactions between stars attempt to produce equipartition of the kinetic energy in the system. However, as a result, energy conservation leads to the contraction of the cluster's core and the evaporation of the less bound stars. The contraction proceeds until the core decouples from the system (Spitzer, 1987). If this process occurs on very short timescales, the decoupled core becomes so dense that stellar collision lead to the formation of a VMS (Portegies Zwart et al., 1999; Devecchi et al., 2010, 2012).

Another model not relying on low metallicity NSCs has been proposed by Davies, Miller & Bellovary (2011). In this model, a NSC invested by a massive gas inflow will experience rapid deepening of its potential well. When the escape velocity from the NSC exceeds $\sim 1000\,\mathrm{km/s}$, the heating accomplished by binaries in the core becomes ineffective and the core collapse can proceeds until SBH-SBH mergers occur in a runaway fashion, leading to the formation of a single MBH seed. Because of the high escape velocity, the BH merger remnant cannot recoil out of the cluster, even in the most extreme conditions ($v_{\mathrm{kick}} \sim 4000\,\mathrm{km/s}$). This model has been investigated in a cosmological context by Lupi et al. (2014), where authors found a maximum seed mass of $10^3\,\mathrm{M}_\odot$.

1.1.4 Primordial black holes

Primordial BHs have been postulated by several authors (Zel'dovich & Novikov, 1967; Hawking, 1971; Khlopov, Rubin & Sakharov, 2005) with masses ranging from the Planck mass (formed at the Planck epoch) to $\sim$ $\mathrm{M}_{\odot}$ (formed at the QCD transition) and up to $10^5\,\mathrm{M}_{\odot}$. However, constraints from Hawking radiation predictions and observations of the gamma-ray background and micro-lensing techniques put into questions the abundance of this class of BHs.

1.1.5 MBH growth across cosmic epochs

All the observations of MBHs in the Universe come from the electromagnetic emission associated with an 'active' phase of the galaxy nucleus. Observations of quiescent BHs at high redshift are unfeasible with the current facilities, and we must wait the upcoming gravitational waves detectors like *the Einstein Telescope* and the *eLISA* observatory to find them. Moreover, these upcoming facilities will give us the ability to investigate in detail the population of BHs at high redshift, and hopefully will allow us to recognise the correct model for seed formation. Though each model predicts different mass functions, formation redshifts and formation efficiency and we could in principle discriminate between them by means of constraints on these parameters, this information is typically smeared out during the cosmic evolution because of the BH growth (via accretion and mergers).

A viable route to find unevolved BH seeds is to look at low mass galaxies. Though the low redshift MBH-host correlations are well established, their evolution at high redshift is still under debate. While the high mass tail is already present at early epochs and it is well traced by AGNs (where the central MBH has already grown significantly), the low mass tail is the most sensitive to the original seeding mechanism. If the first seeds were light, they would start below the correlations and migrate toward them via sustained accretion unaffected by self-regulation. On the other side, massive seeds would lie above the correlations and then experience limited growth to match the correlations at later epochs.

Considering the MBH-host relations in dwarf galaxies, we expect that IMBH would inhabit their nuclei, and that their mass today is close to the original seed mass. However, due to the small region of influence of these BHs

and the low accretion rates, they would be unlikely to be observed (Volonteri et al., 2011). Moreover, the stellar density in dwarf galaxy is small, so that measures based on stellar dynamics are challenging (van Wassenhove et al., 2010), the galaxy centre is often not constrained and the IMBH could be wandering around within the galaxy core, making its identification extremely difficult. Anyway, observations of the low mass end of the galaxy mass function would be very attracting to find imprints of the seeding mechanisms for MBHs.

An alternative route to detect these low mass MBHs would be via radiation produced in tidal disruption events, but we have to wait the new generation telescopes like *ELT* and *JWST* to increase our sensitivity up to the level at which these events will be observable.

Thanks to the Soltan argument (Soltan, 1982; Novak, 2013; Comastri et al., 2015) we know that most of the BH mass has been accreted via radiatively efficient accretion (rather than via mergers or radiatively inefficient accretion), while only a small fraction of the BH mass ($\sim 10\%$) can be accreted via radiatively inefficient accretion at early times. This evidence is very important, since a radiatively inefficient accretion at early epochs would limit the detectability of this kind of objects.

The rapid growth of MBH needed to explain high redshift quasars at $z \simeq 7$ is still an open question (Haiman, 2004; Volonteri & Rees, 2006; Tanaka & Haiman, 2009) and several studies have been conducted to investigate the maximum accretion rates on to MBHs taking into account internal and external effects. The external effect is due to the net amount of gas which could flow downwards to the MBH, while the internal effect is determined by the AGN feedback, i.e. the energy released upon gas accretion by the MBH (Silk & Rees, 1998).

There are mounting evidences that MBHs in high redshift quasars experienced very fast accretion up to billion solar masses. Studies by Ghisellini et al. (2009, 2015), for example, find that all blazars[4] at high redshift have masses above billion solar masses. Being the jet emission collimated, we expect that for each observed blazar $\sim 2\Gamma^2$ unobserved jetted quasars should exist. If one takes into account this population, the estimated MBH density at $z \gtrsim 6$ corresponds to the total MBH density predicted by the cosmological model. The presence of these 'monsters' implies that the 'standard' AGN feedback mech-

[4] A blazar is a quasar powered by a rotating BH emitting a powerful jet pointing to us.

anism cannot play a role in their growth and very high accretion rates might be possible. In these conditions radiation would be trapped and advected inwards, making radiative efficiency to drop (Abramowicz & Lasota, 1980). Radiation trapping would decrease the disc luminosity, and the presence of the jet would dissipate angular momentum, favouring sustained accretion at super-critical rates.

Apart from these peculiar cases, AGN feedback has been advocated as a responsible for the evolution of star forming galaxies to elliptical 'red and dead', which would explain the underabundance of star forming galaxies with respect to many predictions of theoretical models of galaxy formation (Schawinski et al., 2007). In massive galaxies AGN feedback could blow gas out of the halo, decreasing the star forming potential of galaxies.

However, the picture is more complex, since star formation and SNe can limit the MBH growth at early times (Davies et al., 2007; Schawinski et al., 2009) but stellar mass loss can also help fuelling the MBH itself (Ciotti et al., 1991; Ciotti & Ostriker, 1997, 2007; Wild, Heckman & Charlot, 2010). The MBH self-regulation due to alternating active and quiescent phases and the AGN feedback effect on to the galactic medium are considered as the main driver of the establishment of the MBH-host correlations observed.

To study the interplay between galaxies and MBHs, with the goal of understanding how and when the correlations arise, astronomers rely on secondary indicators like the occupation fraction the masses in dwarf galaxies. Theorists, on the other side, use numerical simulations and semi-analytical models, in which they try to include different physical processes which enable to trace the evolution across cosmic time of the baryonic component within galaxies. However, both techniques barely resolve the scales on which the MBH formation and evolution processes operate (Di Matteo, Springel & Hernquist, 2005; Dotti et al., 2007; Johnson & Bromm, 2007; Booth & Schaye, 2009; Debuhr et al., 2010; Power, Nayakshin & King, 2011; Kim et al., 2011; Dubois et al., 2014). Moreover, since galaxies and MBH, as well, evolve via a 'merger driven' scenario (Volonteri & Natarajan, 2009), the memory of the initial conditions when seeds formed is rapidly erased.

Therefore, accurate models for star formation, MBH accretion, stellar and AGN feedback are fundamental to compare numerical results with observations and then discriminate the right path for MBH formation and subsequent evolution. From an observational point of view, upcoming facilities like *JWST,*

Euclid and *ALMA* will be essential to zoom in on the nuclear region of quasars and to extend our limit to the edge of Dark Ages, when the first galaxies, and their embedded BH seeds, started to form.

As already discussed, a promising alternative route to electromagnetic signatures of MBH formation and evolution is that of gravitational waves, which will enable us to look for infant BHs during galaxy mergers at $z \gtrsim 10$, just before the reionisation of the intergalactic medium takes place and immediately after the end of the Dark Ages. Studies bySesana, Volonteri & Haardt (2007) and Arun et al. (2009) show that the mass ratio distribution for coalescing MBHBs depends on the seeding mechanism and the binary evolution considered. In a direct collapse scenario, where seeds are massive, the detected events would involve equal mass binaries. In a low mass scenario, instead, all the events at high redshift would be undetectable with the upcoming facilities and we have to wait for mergers at lower redshift, when BHs have already experienced large accretion and the mass ratio has become more uniformly distributed. Fig. 1.5 shows the rate of binary mergers detectable by *eLISA* for both 'light seed' and 'heavy seed' scenarios, as reported in Volonteri (2010).

A more detailed discussion of MBH formation and evolution at early epochs can be found in Volonteri (2010); Volonteri & Bellovary (2012).

1.2 Binary black holes

According to our 'merger driven' scenario galaxies assembly hierarchically through mergers of smaller units. BH seeds growing in these pristine merging haloes pair in a MBH binary (MBHB) and definitively undergo coalescence (Volonteri, Haardt & Madau, 2003). The coalescence is driven by gravitational wave emission, which offers a unique environment to measure the MBH masses and spins with exquisite precision. Therefore, understanding the MBHB formation path in galaxy mergers is a key step towards understanding the mode of assembly of MBHs across cosmic history (Di Matteo, Springel & Hernquist, 2005; Sesana et al., 2014; Dubois, Volonteri & Silk, 2014).

An example of the characteristics tracks of BHs along cosmic history computed from semi-analytical models of galaxy formation is reported in Fig. 1.6 (Volonteri & Natarajan, 2009).

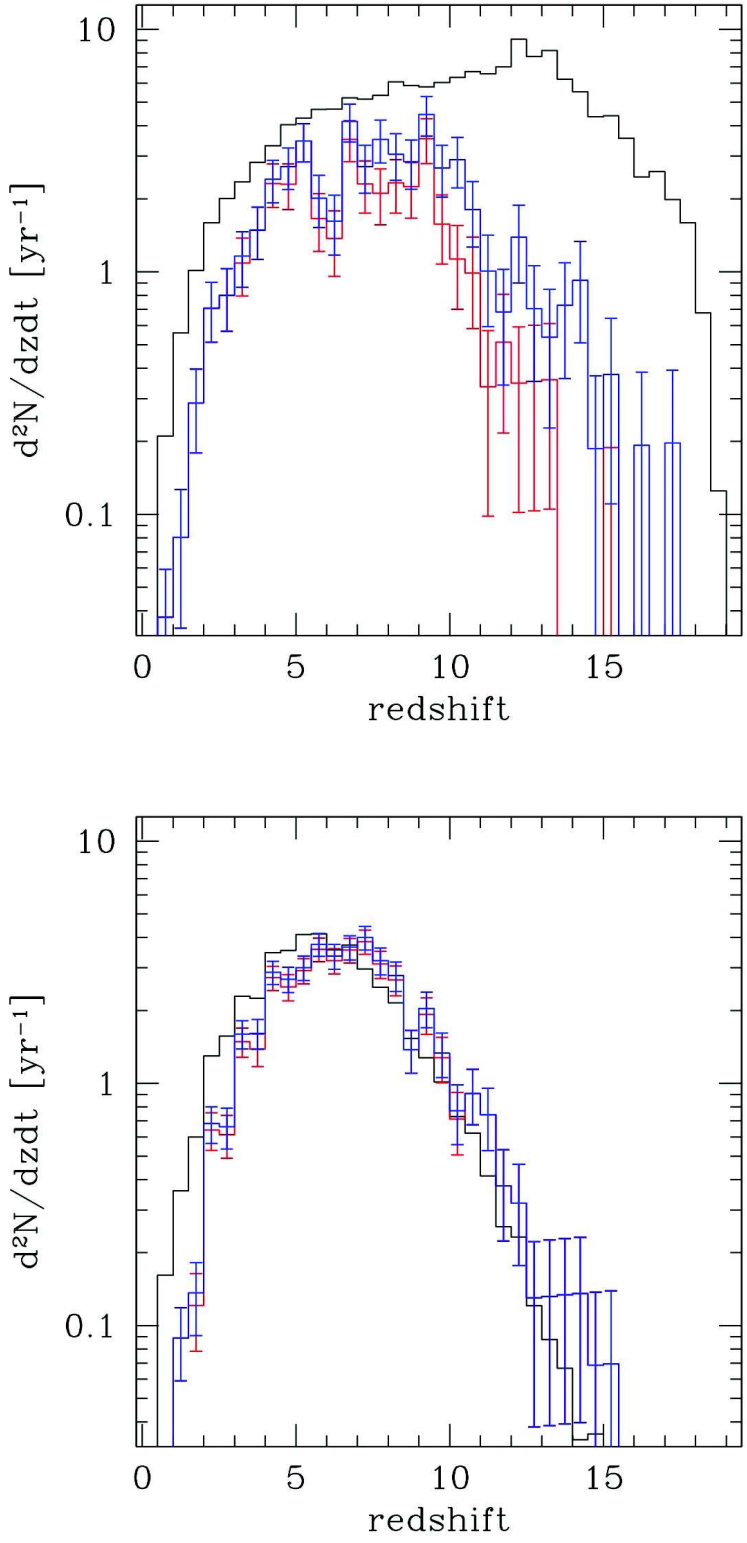

Figure 1.5: *Merger rate of MBHs for two different seed models. Top: 'light seeds' from PopIII remnants. Bottom: 'heavy seeds' from gas-dynamical collapse. Black: all mergers. Red: mergers detectable with* $S/N > 10$ *in eLISA's baseline configuration. Blue: mergers detectable with* $S/N > 10$ *in the old 6-link configuration (three independent channels). Figures taken from Volonteri (2010).*

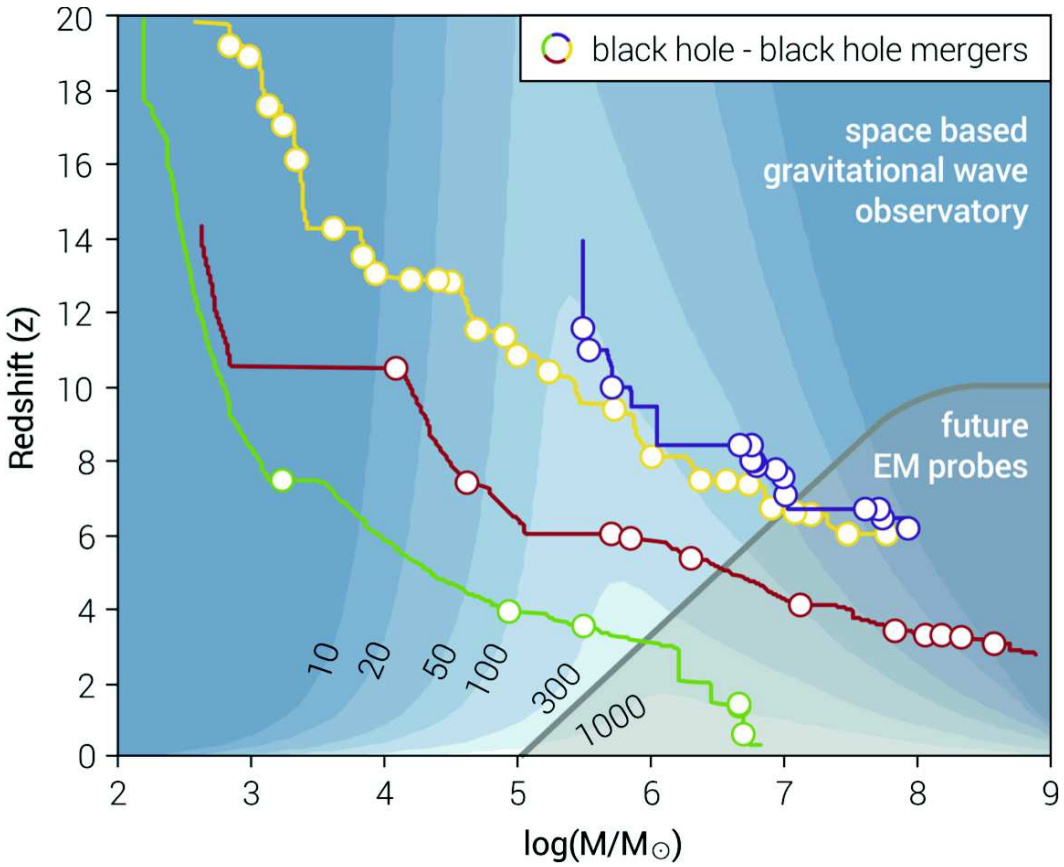

Figure 1.6: *Paths of black holes forming at high redshift from light ($10^{2-3}\,\mathrm{M}_\odot$) and heavy ($10^{5-6}\,\mathrm{M}_\odot$) seeds. The black holes evolve along tracks, in the mass versus redshift diagram, as they experience accretion episodes and coalescences with other black holes. Circles mark the loci of black hole coalescences. Four paths are selected: two ending with a black hole powering a $z \sim 6$ QSO (starting from a massive seed, blue curve, and from a seed resulting from the collapse of a massive metal-free star, yellow curve); a third ending with a typical $10^9\,\mathrm{M}_\odot$ black hole in a giant elliptical galaxy (red curve); and finally the fourth ending with the formation of a Milky Way-like black hole (green curve). The tracks are obtained using state-of-the-art semi-analytical merger tree models. The grey transparent area in the bottom right corner roughly identifies the parameter space accessible by future electromagnetic probes which will observe black holes powered by accretion. Over-lied are contour levels of constant sky and polarisation angle-averaged Signal-to-Noise-Ratios (SNRs) for eLISA, for equal mass non-spinning binaries as a function of their total rest frame mass (Amaro-Seoane et al., 2013). It is remarkable that black hole mergers can be detected by eLISA with a very high SNR across all cosmic ages. Figure taken from Amaro-Seoane et al. (2013).*

1.2.1 Binary black holes: formation and evolution

In a seminal work Begelman, Blandford & Rees (1980) discussed how the bending and apparent precession of radio jets observed in a number of galactic nuclei could be related to the presence of a MBH binary and explored the dynamics of its formation.

They depicted the existence of three main phases along the way to coalescence:

- (I) an early phase of *pairing* under the dynamical friction by stars of the merger remnant, resulting in the formation of a *Keplerian binary*. As dynamical friction is proportional to the background density of stars and to the square of the MBH mass (it acts on each MBH individually), the more massive the BHs and denser the environments are, the more rapidly they sink. Moreover, being the dynamical friction timescale a function of the distance from the galaxy nucleus r ($\tau_{\rm DF} \propto M_{\rm BH}^{-1} r^2 \sigma_*$, where σ_* is the 1D stellar velocity dispersion), the process becomes ever more rapid with orbital decay. A binary forms approximatively when the mass in stars enclosed in their orbits becomes smaller than two times the binary total mass. During the inspiral the eccentricity does not vary significantly. As the acquired velocity of the BHs becomes large, the effect of dynamical friction weakens, leading to the end of the first phase;

- (II) a *migration/hardening* phase during which the binary shrinks due to energy loss by close encounters with single stars and gas torques (if gas is present) (Amaro-Seoane, Brem & Cuadra, 2013). Each encounter with a star can extract a fraction $\sim \varepsilon m_*/M_{\rm MBHB}$ of the MBHB energy (where $\varepsilon \sim 0.2 - 1$ is obtained from averaging over many interactions, m_* is the stellar mass and $M_{\rm MBHB}$ is the binary mass), hence a large number of scattering is required to significantly reduce the binary separation. Unlikely the first phase, the timescale for the second phase increases with decreasing separation, and indicate that hardening proceeds more slowly with orbital decay.

  The transition to the third phase occurs when the timescale for gravita-

tional wave emission drops below the hardening timescale

$$a^*_{\rm II\to III} = \left(\frac{256G^2}{5\pi c^5}\right)^{1/5}\left(\frac{\sigma_*}{\rho_*}\right)^{1/5} f^{1/5}(e)\left[\frac{q}{(1+q)^2}\right]^{1/5} M^{3/5}_{\rm MBHB}, \tag{1.4}$$

where ρ_* is the stellar density around the binary, q is the mass ratio between the two BHs and $f(e) = [1+(73/24)e^2+(37/96)e^4](1-e^2)^{-7/2}$. In order to observe the binary coalescence the transition must occur in less than a Hubble time, and this typically occurs for separations of $1.4\cdot 10^4 (M_{\rm MBHB}/10^6)^{-1/4} r_{\rm g}$ for an equal mass circular binary (where $r_{\rm g}$ is the Schwarzchild radius associated to the binary $2GM_{\rm MBHB}/c^2$).

Though in principle the stellar environment conditions are compatible with rapid coalescence of the binary, nature operates in a different way. Indeed, close encounters only occur with stars orbiting within the so-called 'loss cone', i.e. the domain, in phase-space, with sufficiently low angular momentum to interact with the binary. Since in typical spheroids the loss cone is not enough populated and stars interacting with the binary are ejected from it, the binary finally stalls at $\sim 0.1-1$ pc. At this separation the coalescence cannot occur in a Hubble time, and this is referred to as 'last parsec problem'. The ejection of stars from the loss cone results in the formation of a stellar core around the binary, which has been observed in some core, missing-light elliptical galaxies (Milosavljević & Merritt, 2001; Kormendy, 2013; Merritt, 2013).

- (III) a *gravitational wave driven inspiral* phase ending with the coalescence of the two black holes due to the emission of gravitational waves. At the coalescence, the BH grows in mass, acquire a new spin according to mass/energy conservation and can receive a recoil up $\lesssim 5000\,{\rm km/s}$, which depends on the orientation and magnitude of the BH spins and orbital angular momentum.

After Begelman, Blandford & Rees (1980), the last parsec problem has been investigated in detail, representing a bottleneck to the path to coalescence, at least for MBHs with masses larger than $\sim 2\times 10^6\,{\rm M}_\odot$ (see e.g., Merritt & Milosavljević, 2005; Merritt, Mikkola & Szell, 2007). However, in the case of not spherically symmetric systems, like those obtained in recent galaxy merger's simulations (Berczik et al., 2006; Khan et al., 2013; Wang

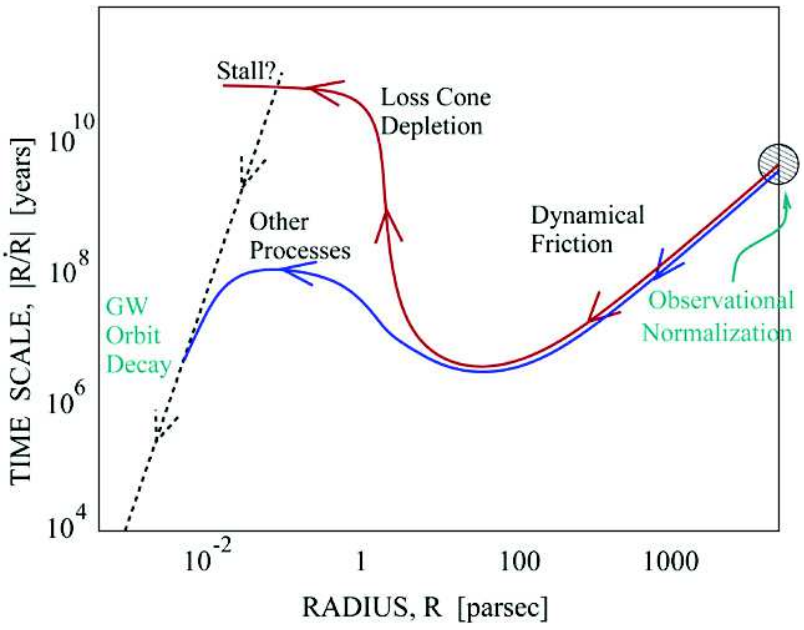

Figure 1.7: *Evolution of MBH separation during a galaxy merger, as depicted by Begelman, Blandford & Rees (1980). In phase (i) dynamical friction brings the MBHs together in a Keplerian binary. In phase (ii) the MBHB slowly shrinks via three-body scattering with stars in the loss-cone. Finally, when the two MBHs are sufficiently close, the loss of orbital energy and angular momentum via gravitational radiation leads to rapid coalescence. Figure taken from Backer, Jaffe & Lommen (2004).*

et al., 2014), this problem can be alleviated (Yu, 2002; Merritt & Poon, 2004; Vasiliev, Antonini & Merritt, 2015; Sesana & Khan, 2015). As a byproduct of these less simplified assumptions, also the eccentricity has been seen to increase to $\sim$ 1 (Preto et al., 2011; Khan, Just & Merritt, 2011), indicating rapid transfer of angular momentum to stars via scattering. Finally, additional mechanisms like recycling of stars ejected by the binary on returning eccentric orbits or perturbers as massive clouds can accelerate the hardening phase. A schematic picture of the three phases is depicted in Fig. 1.7.

The study of these three phases is particularly interesting for IMBHs, since these objects are the primary sources for the *eLISA* observatory and, as already discussed, the best candidates to find an imprint of seed formation mechanisms. However, to reach the separations for IMBH binary coalescence at high redshift high stellar densities and low velocity dispersions are essential, and it is unclear whether these conditions can be fulfilled in the high redshift NSCs where IMBHs are supposed to live.

1.2.2 Black hole dynamics in mergers

While the study of black hole binary formation and evolution in purely stellar systems has been explored in detail with direct N-body codes (e.g., Khan et al., 2013), black hole dynamics in mergers between disc galaxies (especially gas rich galaxies) has been studied from cosmologically motivated orbits down to scales of $\lesssim 10$ pc when a Keplerian binary forms (Mayer et al., 2007; Colpi et al., 2009; Chapon, Mayer & Teyssier, 2013; Mayer, 2013). The subsequent hardening phase has been later investigated with dedicated simulations (e.g., Escala et al., 2005; Dotti et al., 2007; Fiacconi et al., 2013).

Disc galaxies are multi-component systems comprising a dark matter halo, a stellar disc coexisting with a multi-phase gas disc, a central stellar bulge (if any) and also a MBH, if present. The study of the dynamical evolution of MBHs in the time-varying environment of a galaxy merger is complex as it occurs on the same time-scales on which stars and gas, turning into new stars stars, evolve (see Colpi, 2014, for a review). Despite the large number of simulations of galaxy mergers, only few of them follow the BHs down to pc scales, while most of the simulations assume prompt BH coalescence.

A leap in understanding the role of gas during the pairing phase has been taken when studying minor mergers, i.e. mergers with nominal 1:4 mass ratios and less. Early works on collisionless mergers of unequal-mass spherical dark matter haloes indicated that besides dynamical friction other mechanisms are at play: (i) tidal stripping of the secondary, which delays the sinking by dynamical friction; (ii) tidal heating, which (partially) dissolves the system (Taffoni et al., 2003). Depending on the parameters of the orbit and the halo properties, the encounter can lead to a rapid orbital decay, disruption or survival of the secondary halo. These results show that under certain conditions the secondary BH can orbit in the galaxy periphery for more than a Hubble time, without forming a binary with the primary BH.

It has been recently highlighted that gas can play a pivotal role in unequal-mass galaxy mergers. In wet mergers (where the gas fraction is above 10%) tidal torques during the last peri-centre passage before merging triggers inflows towards the nucleus of the secondary, giving rise to a starburst which deepens the potential well in the BH neighbourhood, thus reducing the action of tides by the primary (Kazantzidis et al., 2005). This prevents the wandering to the lighter MBH in the periphery of the primary when sufficient gas is

present in the host galaxy, raising the question as to whether minor mergers lead in general to binary formation (Callegari et al., 2009, 2011; Khan et al., 2012). Fig. 1.8 shows two examples of unequal-mass wet mergers (one with 1:4 mass ratio and one with 1:10).

An additional effect which must be considered to have a clearer picture of the fate of BHs in minor mergers is accretion. Indeed, in wet mergers both BHs can accrete from the surrounding gas and increase their mass. Since higher masses correspond to shorter dynamical friction timescales, the secondary BH can sink more rapidly towards the primary BH. Moreover, the interaction with the gas of the main galaxy, can trigger episodes of accretion, whose result is to decrease the mass ratio between the two BHs, sometimes giving rise to equal mass binaries.

Follow-up studies have indicated that the dividing line from success and failure in forming an MBHB is around 1:10 mass ratios (Bellovary et al., 2010; Van Wassenhove et al., 2012), but still depends on details such as the encounter geometry and gas content. In summary, the fate of BHs in minor mergers depends on several parameters and the boundary between failure and success is poorly constrained.

On the other side, major mergers among gas-rich galaxies represent a natural path for MBH pairing and binary formation. The orbital braking is in these cases driven by gas-dynamical friction which is faster than dynamical friction from stars (Escala et al., 2005; Dotti, Colpi & Haardt, 2006; Mayer et al., 2007; Chapon, Mayer & Teyssier, 2013). During the merger, the two progenitor galaxies first experience a close fly-by, which triggers the formation of tails and plumes due to tidal forces. The discs sink via dynamical friction against the dark matter halo, dragging the central MBHs with them, collide and develop strong shocks which redistribute/cancel angular momentum. This process triggers a large inflow towards the centre, leading to the formation of a turbulent, rotationally supported, massive circum-nuclear disc of $10^9\,\mathrm{M}_\odot$ with a radius of $\lesssim 100$ pc (Chapon, Mayer & Teyssier, 2013). The gravo-turbulence in the disc is large enough to make the disc Toomre stable[5], thus preventing gas fragmentation and star formation, and helps the MBH to

[5]The Toomre criterion is a stability criterion for discs, the analogous of Jeans criterion for spherical systems. The parameter used to describe the disc stability is called 'Toomre parameter' and it is defined as

$$Q_{\mathrm{Toomre}} = \frac{c_{\mathrm{s}} k}{\pi G \Sigma}, \tag{1.5}$$

sink and form a Keplerian binary on typical timescales of few Myr after the completion of the merger.

The limit and drawback of these high-resolution simulations is that gas is treated as a single phase medium described by a polytropic equation of state (with index 7/5) which mimics the thermodynamics of a generic star-forming region. Only recently state-of-the-art simulations of major mergers have achieved enough resolution to detail the MBH dynamics down parsec scales, in presence of a multiphase gas (Capelo et al., 2015; Roškar et al., 2015). Capelo et al. (2015) studied the large-scale dynamics of MBH in a variety of mergers with mass ratio 1:1 down to 1:10 to explore the black hole accretion history and their dynamics during the pairing phase, down to separations of several parsecs. Their focus was mainly in exploring the possibility of triggering 'dual' AGN activity along the course of the merger. Thanks to improved recipes for cooling, SF and SNa feedback and the increased the resolution through excised zoom-in techniques on the nuclear region of the merger remnant, Roškar et al. (2015) accurately describe the pairing phase of MBHs during the major merger between two galaxies with moderate gas fractions, finding that gas is mostly evacuated by the strong starburst triggered by the merger and the galactic disc is rebuild in ~ 10 Myr. Even when the disc reforms, it exhibit non axisymmetric features like spiral arms and a population of stochastically distributed clumps. MBHs moving in this environment experience gas torquing and scattering (due to massive clumps) whose net effect is to delay the binary formation. However, despite the delay, the typical timescale for the inspiral is ~ 100 Myr, which is shorter than what found in collisionless simulations. This result is in agreement with what found by Fiacconi et al. (2013), who describes BH dynamics in clumpy nuclear discs.

1.2.3 Black hole dynamics on nuclear scales

Besides the *ab initio* simulations of entire galaxy mergers, a good benchmark for studying the binary formation and the path to coalescence is through sim-

where $c_{\rm s}$ is the sound speed, $k = \frac{2\Omega}{R}\frac{\mathrm{d}}{\mathrm{d}R}(R^2\Omega)$ is epicyclic frequency, Ω is the orbital frequency, R is the cylindrical radius and Σ is the disc surface density. It has been named after Toomre (1964), who discussed the stability of stellar discs (the stability criterion for gaseous flattened systems has been studied by Safronov, 1960) and it asserts that gravitational instabilities can develop in a disc when gravity exceeds the stabilising support of both pressure and rotation.

ulations of massive circum-nuclear gaseous discs, like in Escala et al. (2005); Dotti et al. (2009); Fiacconi et al. (2013); del Valle et al. (2015).

Gaseous discs can cool down, develop turbulence and inhomogeneities like massive clumps which will then form stars. Gas can also dissipate the kinetic energy of the moving black holes via radiative cooling in a nearly Keplerian circum-binary disc (Amaro-Seoane, Brem & Cuadra, 2013). A compelling question is whether angular momentum transfer via gas-BH interactions is faster than that from star-BH scattering.

In a number of studies (Escala et al., 2005; Dotti et al., 2007, 2009; Fiacconi et al., 2013; del Valle et al., 2015), the massive disc has been modelled as a self-gravitating axisymmetric Mestel disc (Mestel, 1963). The disc is pressure supported vertically and the Toomre parameter is everywhere greater than 3, preventing the development of gravitational instabilities. In order to model the innermost region of the galactic bulge the disc is embedded in a stellar spherical background, modelled as a Plummer sphere (Plummer, 1911). The primary BH resides in the centre of the disc, and a secondary BH is assumed to wander in the outer region of the disc.

These studies have highlighted key differences in the BH dynamics with respect to collisionless studies, especially the dragging of the secondary BH in a co-rotating co-planar circular orbit. The simulations showed that any orbit with large initial eccentricity is forced into circular rotation in the disc and that circularisation is faster the cooler is the disc. Moreover, if the initial orbit of the secondary is counter-rotating with respect to the disc, dynamical friction acts against the secondary, turning counter-rotating orbits into co-rotating ones.

The limit of these studies is that smooth, stable discs are ideal. Gas can cool and fragment, forming stars, and the feedback associated to SNa explosions and stellar winds can feed back energy into the gas, inducing the formation of a multi-phase medium. Due to the complex interplay between cooling, SF and feedback from SNe and stellar winds, a first attempt has been to insert a phenomenological prescription for radiative cooling to allow the controlled formation of clumps (Fiacconi et al., 2013). Clumps forming in the disc have masses between 10^5 and $10^7\,\mathrm{M}_\odot$ and radii of few pc and evolve via mass segregation, mergers with other clumps and collisions with clumps or the secondary BH. They act as massive perturbers and disturb the smooth orbital decay of the BH due to not coherent torques. Because of these torques, the BH

deviates from its original orbit, moving inwards or outwards, or also out of the disc plane. Depending on the interactions experienced the orbital decay can be either accelerated or delayed. In a number of interactions, a temporary BH-clump pair form, which rapidly sinks towards the centre. This stochasticity of the BH dynamical evolution in a clumpy medium suggests that a detailed description of the multi-phase interstellar medium on nuclear scale is essential to achieve a clearer understanding of the BH pairing process.

When the mass enclosed within the binary orbit becomes smaller than the binary total mass, a Keplerian binary forms, surrounded by a less massive disc, named *circum-binary disc*. At this stage, the tidal force exerted by the binary excites non axisymmetric density perturbations in the disc, resulting in the formation of a cavity, called *gap* (e.g., Pringle, 1991; Cuadra et al., 2009; Roedig et al., 2012; Farris et al., 2014, 2015), with a radius close to $\sim 2a(t)$, where a is the binary semi-major axis. Under these conditions, the orbital decay proceeds at a slow rate (Gould & Rix, 2000; Armitage, 2013), set by the viscous time of the gas flowing within the cavity to accrete on the BHs. However, recent 2D and 3D simulations suggested that as many as three discs exist in the binary+disc system, i.e. the circum-binary disc plus two mini-discs around each BH of the binary, and that these discs persist being fed by gas flowing through the gap. Fig. 1.9 shows an example of a circum-binary disc with the two associated mini-discs, taken from Roedig et al. (2012).

The eccentricity e tends to increase during the binary shrinking (Armitage & Natarajan, 2005) due to vicinity of the secondary BH to the cavity's edge, where a trailing density wave induced by the BH reduces the BH tangential velocity, leading to an increase of e. This process proceeds until the BH velocity at the apocentre becomes smaller than the gas velocity. In these conditions, the density trail moves ahead the BH and increases its velocity. The process reaches saturation at $e \sim 0.6 - 0.8$. During this phase, also the binary mass tends to increase, due to repeated accretion episodes from the mini-discs around the BHs, modulated by the orbital period of the binary. This process is particularly interesting, since it can produce electromagnetic signatures of the binary shrinking before the transition to the gravitational wave driven phase.

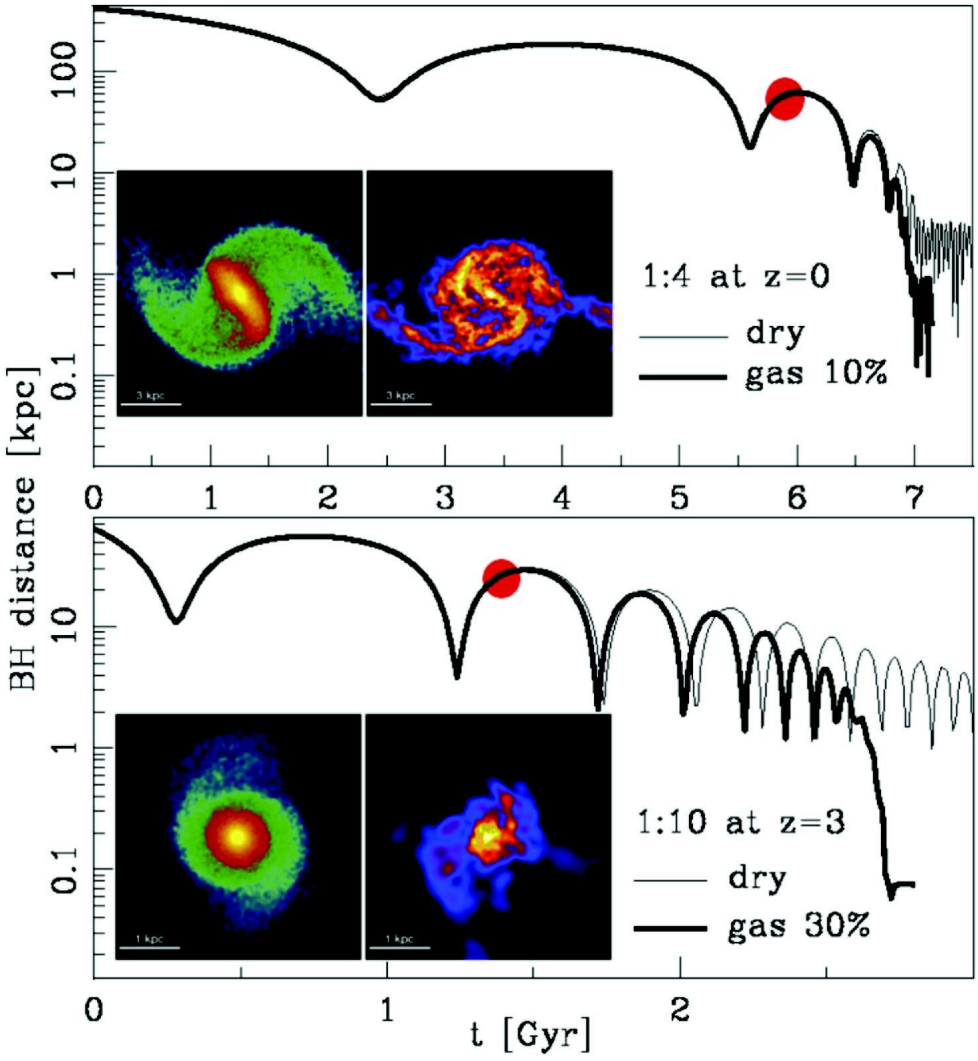

Figure 1.8: *Upper panel: black hole separation as a function of time for a 1:4 merger. The thin and thick lines refer to the dry (gas free) and wet (with gas fraction of 10%) cases, respectively. The inset shows the colour-coded density of stars (left) and gas (right) for the wet case at* $t = 5.75$ *Gyr (marked with a red dot on the curve); each image is 12 kpc on a side, and colours code the range* $10^{-2} - 1$ $\mathrm{M}_{\odot}$ *pc*$^{-3}$ *for stars, and* $10^{-3} - 10^{-1}$ $\mathrm{M}_{\odot}$ *pc*$^{-3}$ *for the gas. Lower panel: black hole separation as a function of time for a 1:10 merger (upper panel). The thin and thick line refer to the dry and wet (with gas fraction of 30%) cases, respectively. The inset shows density maps at* $t = 1.35$ *Gyr for the wet merger: images are 4 kpc on a side (colour coding as in upper panel). Figure taken from Callegari et al. (2009).*

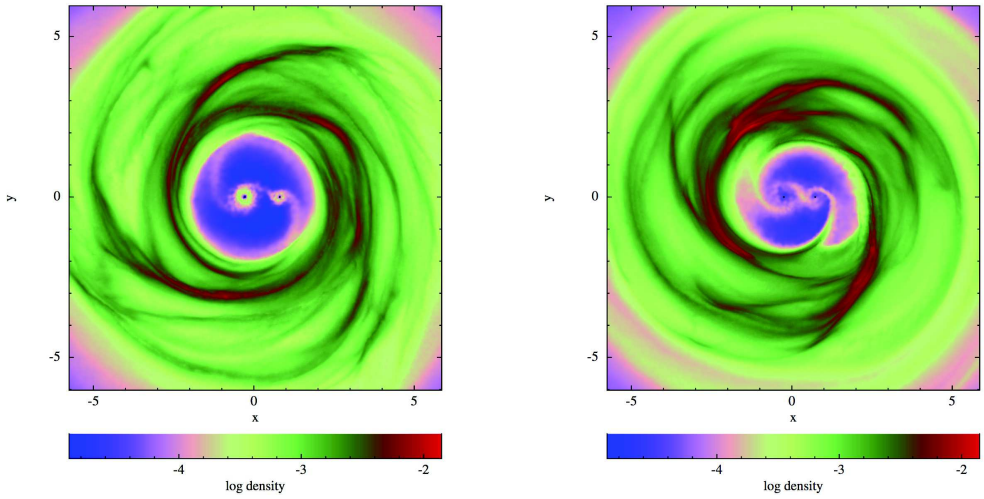

Figure 1.9: *Colour-coded gas surface density of two Newtonian, self-gravitating circum-binary discs, showing the presence of a binary region with the two black holes and their mini-discs, a porous cavity filled with streams, the inner rim or edge working as a dam, and the body disc. Left (right) panel refers to a run with gas in the cavity treated with an isothermal (adiabatic) equation of state. Figures taken from Roedig et al. (2012).*

Chapter 2

Numerical techniques

Numerical methods are becoming ever more important in astrophysics. The interplay among the fundamental physical processes describing our Universe cannot be reduced to simple analytical estimates, unless one limits to highly approximated treatments. The key feature of the computer is its capability to solve complex systems of equations, which allows us to have a powerful complementary tool to the analytic reasoning. In this chapter I will describe the two codes I used for the simulations performed in this work, based on two different approaches: the Eulerian technique, used in mesh-based codes and the Lagrangian technique, used in particle-based technique.

In the mesh-based approach the fluid properties are discretised on a regular mesh fixed in time. Collisionless components like dark matter, stars and BHs are mapped on to the grid as well, via suitably build deposition schemes, in order to build a density field. Gravity is then computed on the mesh by solving the Poisson equation and then the accelerations are calculated at the original particle positions. Fluid, instead, is followed by solving the Euler equations for hydrodynamics. Being the Euler equations a system of hyperbolic partial differential equations, a solution is found solving a Riemann problem at the cell interfaces through the so-called Godunov methods (Godunov, 1959). The Godunov schemes are conservative finite volume methods able to track the fluid properties even when shocks occur in the gas flow.

One of the main limits of mesh-based codes is that the grid is fixed in time, thus they suffer from numerical diffusion and poor angular momentum conservation.

In the particle-based approach, instead, fluid is sampled with a discrete set of tracers. In the smoothed particle hydrodynamics (SPH) scheme the tracers represent particles of the fluid, while in the recently developed unstructured mesh schemes they represent the centre of a set of cells which can move with the fluid. Gravity is typically computed for both fluid tracers and collisionless particles via a hierarchical multipole method (often called 'tree method' because of the geometrical structure used to represent the particle distribution), where the contribution of a distant group to the acceleration is described via a multipole expansion of the group's gravity instead of the using each particle like in direct N-body codes.

In SPH schemes, fluid evolution is followed by solving the Euler equations in Lagrangian form. Because of the intrinsically local validity of the Lagrangian form, shocks are poorly resolved, and one needs to add an artificial viscosity term to take them into account.

In the recently developed moving mesh or mesh-free methods, instead, particles are only tracers of a set of moving cells, hence the Euler equations are solved as in mesh-based codes, avoiding the use of artificial viscosity. These new methods have been developed with the aim at capturing the advantages of both SPH and mesh-based codes, i.e. the intrinsic adaptivity, no preferred directions (hence perfect conservation of angular momentum) and optimal gravity coupling from SPH codes and the very accurate description of shocks, fluid instabilities and shear flows from mesh codes.

The reader interested in a more detailed discussion of the numerical methods used in both mesh-based and particle-based codes to describe gravity and hydrodynamics can read appendix A.

In the next sections I will describe the codes used in this work, RAMSES and GIZMO, focussing on the details of the sub-grid physics implementations and on the differences from the standard schemes just presented.

2.1 A Eulerian case: the AMR code RAMSES

RAMSES (Teyssier, 2002) is a mesh-based code with an additional adaptive refinement scheme. The idea behind the refinement technique is to adaptively increase the resolution of the coarse grid when and where required by specified properties of the fluid (Berger & Oliger, 1984; Berger & Colella, 1989). When a cell's refinement criterion is satisfied, a new grid level is created with a resolution 2^ν (with ν the number of dimensions) times better than the underlying one and the cell's quantities are interpolated to map the coarse values to higher resolution. On the other side, when a refinement criterion is not fulfilled anymore, cells are de-refined, reducing resolution if not needed. Each level is evolved in time in an iterative fashion, interpolating and projecting quantities between adjacent levels to preserve consistency. Such a scheme results in a reduced computational cost with respect to a fixed grid simulation, and allows for a better description of phenomena occurring on very different scales (like, for example, the clustering of galaxies on scales of tens of Mpc and the formation of stars within the molecular clouds with radii of few pc). Despite the great advantages of the AMR technique, some spurious effects can arise, especially for massive particle orbits. I will discuss them in Chapter 4.

Gravity is computed in the code through a spectral method at the coarse level and a multi-grid approach for refined levels. The gravitational contribution of dark matter and stellar particles is consistently included in the code by means of a PM scheme (see section §A.1.1).

Hydrodynamics is computed with a Godunov-type scheme at each level, and a sub-stepping algorithm is also implemented for refined levels to improve the accuracy of hydrodynamical processes.

The code has been released to the scientific community and also implements some simple recipes aimed at modelling physical processes which cannot be resolved on the simulated scales (the so-called sub-grid physics). The processes implemented in the code are:

- Optically thin radiative cooling, with the cooling function from Sutherland & Dopita (1993) for primordial composition gas and also taking into account metal line cooling;

- Star formation (SF), enabled for low-temperature ($T < 2 \times 10^4$ K) and high-density gas ($\rho > \rho_{\rm thr}$, where $\rho_{\rm thr}$ is a user-defined value).

Since simulations on galactic scales cannot resolve the molecular clouds where star forms, SF is implemented converting a fraction of the gas within a cell into a discrete number of stellar particles, through a Poisson sampling procedure aimed at recovering the Kennicutt-Schmidt relation (Kennicutt, 1998). This procedure is described in details in Rasera & Teyssier (2006);

- Supernova (SNa) feedback, driven by stellar particles that are allowed to explode as type II SNae releasing mass and energy into the surrounding medium. In the standard implementation a fixed mass fraction of the stellar particle explodes after a typical life-time of 10 Myr and releases into the cell hosting the SNa progenitor $10^{50}\,\mathrm{erg}/\,\mathrm{M}_{\odot}$ as thermal energy only. The SNa yield can be arbitrarily chosen to model a stellar population described by a given initial mass function (IMF). Such a thermal prescription, however, cannot properly describe the momentum-driven blast wave typically associated with SNae. In order to model the non-thermal processes energising the blast wave an alternative model has been also implemented, termed "blast wave-like feedback", in which the released energy is decoupled from the gas thermal budget for a typical time-scale $\tau_{\mathrm{delay}} = 20$ Myr, so that it cannot be immediately lost via radiative cooling, inhibiting the SNa effect (Teyssier et al., 2013).

- accretion on to sink particles, implemented with three different prescriptions (Bleuler & Teyssier, 2014): (i) a fraction of the total gas mass within the accretion radius is accreted when the gas density exceeds a user-defined threshold ("threshold accretion"), (ii) gas is continuously accreted with a rate estimated through the Bondi-Hoyle formula ("Bondi-Hoyle accretion") or (iii) the accretion rate is computed as the net mass flux flowing within the accretion radius of the sink particle ("flux accretion"). For all of these prescriptions the accretion radius is set to 4 cells and the accretion rate is capped at the Eddington limit.

- BH thermal feedback (Dubois et al., 2014), in which the radiation produced by accretion is stored until the total budget is large enough to heat up the surrounding gas to at least 10^7 K. This prescription has been initially proposed by Booth & Schaye (2009) to prevent the gas to immediately loose the small amount of additional thermal energy

gained after each time-step, which would result in an ineffective feedback. The RAMSES prescription assumes a fixed accretion radiative efficiency $\epsilon = 0.1$ (the standard value used for the optically thin Shakura & Sunyaev accretion disc; Shakura & Sunyaev, 1973) and a fixed fraction $= 15\%$ of the accretion energy to be released to gas.

2.2 A Lagrangian case: the mesh-free code GIZMO

With respect to standard Lagrangian codes based on the SPH formalism, GIZMO implements the new mesh-free Godunov-type method described in section §A.2.2.

Gravity is computed using a tree-based algorithm derived from the gravity algorithm used in GADGET2.

The code has been publicly released to the scientific community in a basic version, where only hydrodynamics and gravity are available. Since our studies also need additional sub-grid physical processes, I implemented in the code the additional recipes necessary to model gas cooling, SF, type II SNa feedback, gas accretion onto BHs and BH feedback, in a fashion similar to RAMSES :

- I included radiative cooling by means of the GRACKLE[1] chemistry and cooling library, which provides both equilibrium and non-equilibrium chemistry (The Enzo Collaboration et al., 2014; Kim et al., 2014). In my runs I employed the equilibrium cooling curve for primordial species (atomic H and HE), and tabulated metal cooling and heating from CLOUDY (Ferland et al., 2013), in order to be consistent with RAMSES prescriptions.

- Gas particles are eligible to star formation when they match the same criteria for density and temperature adopted in RAMSES and belong to a converging flow (i.e., $\nabla \cdot \mathbf{v} < 0$). Resulting star particles are generated locally according to the Kennicutt-Schmidt law (Kennicutt, 1998), and using a stochastic prescription as described in Stinson et al. (2006).

- My implementation of SNa feedback slightly differs from RAMSES one. While in RAMSES energy is wholly released within the cell hosting

[1] `http://grackle.readthedocs.org`

the progenitor, in GIZMO I distributed it among the gas particles lying within the SNa maximum extension radius $R_{\rm E}$[2] (Chevalier, 1974), according to their distance from the SNa, weighted through the kernel function used in the code. For the SNa blast wave-like model I implemented a "delayed cooling" prescription, in which I temporarily inhibit radiative cooling for gas particles within $R_{\rm E}$. Because of these differences one must be careful in choosing the cooling shut-off time, in order to get consistent results.

- I enabled gas accretion onto sink particles from gas particles lying within a distance encompassing an effective number of neighbours $N = 32$, weighted through the kernel function used in the code, so that the accretion zone implicitly adapts following the sink particle motion with the flow, unlike in RAMSES . Though this choice allows for better resolution in very high-density regions (where the kernel size becomes very small), spurious accretion events could in principle occur when the particle moves in a low-density environment, i.e. when the kernel size is very large. In order to prevent this undesired effect I enabled accretion only for sink particles with a kernel radius smaller than 10 times the softening length associated with the sink particles themselves. The accretion rate is then computed following RAMSES recipes.

- My recipe for BH thermal feedback strictly follows RAMSES implementation.

[2]The SNa maximum extension radius is defined as $R_{\rm E} = 10^{1.74} E_{51}^{0.32} n_0^{-0.16} \tilde{P}_{04}^{-0.20}$ pc, where $E_{\rm SN} = E_{51} 10^{51}$ erg, n_0 is the ambient hydrogen density and $\tilde{P}_{04} = 10^{-4} P_0 k_{\rm B}^{-1}$ with P_0 the ambient pressure

Chapter 3

Massive black hole seed formation: stellar mass black hole growth via super-critical accretion

As already described in Chapter 1, MBHs are typically observed as AGNs (or high redshift quasars) powered by the gas accretion process on to the hole. Mass estimates show that along cosmic history MBHs coevolve with the galaxy host, as highlighted by the $M_{\rm BH} - \sigma$ and the $M_{\rm BH} - M_{\rm spheroid}$ relations (Gültekin et al., 2009; Häring & Rix, 2004).

According to the Λ-CDM cosmology paradigm, galaxies form via accretion and mergers with smaller sub-units, and MBHs grow in symbiosis with them. Therefore, we expect that the first BHs were lighter, and then they have grown up to billion solar masses. Some of them must have grown very rapidly, in order to explain the high redshift quasars observed (up to $z \sim 7$), powered by $\sim 10^9\,{\rm M}_\odot$ MBHs (Mortlock et al., 2011). Different models for seed formation have been proposed to date, from the collapse of PopIII stars (Madau & Rees, 2001) to dynamical processes in stellar clusters (Devecchi & Volonteri, 2009; Davies, Miller & Bellovary, 2011), to the direct collapse of a massive gas cloud (Haehnelt & Rees, 1993). Each model carries its own pros and cons, and a clear consensus about the correct one is still missing, in particular because of the technical limits of our observations (both the electromagnetic and

gravitational ones).

Recently Madau, Haardt & Dotti (2014) proposed an alternative scenario, in which super-critical (i.e., super-Eddington) accretion episodes onto stellar mass seeds could help to bypass the difficulties associated with all the other scenarios. They considered the radiatively-inefficient "slim-disc" solution (Abramowicz et al., 1988) - advective, optically thick flows that generalise the standard Shakura & Sunyaev solution (Shakura & Sunyaev, 1973) - and showed how mildly super-Eddington accretion significantly eases the problem of assembling MBHs in less than a billion year. Because of the (accretion-rate dependent) low radiative efficiencies of slim discs around non-rotating *as well as* rapidly rotating holes, the accretion time-scale in this regime is almost independent of the spin parameter. It is this unique feature of slim discs that makes such models so appealing. In the paper, they briefly discussed (see also Volonteri, Silk & Dubus, 2015) how conditions for super-critical accretion could be physically plausible in the dense environment of high redshift massive protogalaxies.

Here, I elaborate upon this concept by means of high resolution simulations of a cluster of stellar mass black holes orbiting the central ~ 200 pc of a gas-rich galaxy. I will focus on the effect of a radiatively inefficient BH feedback on the growth of SBHs embedded in a circum-nuclear gaseous disc, showing how the interplay between gas dynamics and the black holes can easily lead to the formation of a massive MBH seed in the centre of the system within few million years. Though these simulations are highly idealised and should be thought as a proof-of-concept of the proposed scenario, they highlight the basic point, i.e. that super-Eddington accretion in well-formed, evolved galaxies is an attractive route to the formation of MBH seeds. In fact, a population of SBHs is expected to reside in the inner ~ 200 pc, the circum-nuclear disc can provide enough gas to be accreted, and negative feedback is negligible in the high-density clumps developed in the disc.

3.1 Initial conditions

I considered the nuclear region of a high-redshift massive spiral galaxy, assuming that it hosts a gaseous circum-nuclear disc (CND hereafter). I modelled the CND following an exponential surface density profile (see Fig. 3.1, top panel), with total mass $10^8\,\mathrm{M}_\odot$ and scale radius 50 pc. The disc has been em-

bedded in a stellar spherical background following an Hernquist profile, with scale radius 100 pc and total mass $2 \times 10^8\,\mathrm{M}_\odot$. Note that the adopted gas mass from the CND is at the lower end of that of CNDs observed in low z merger remnants (e.g., Medling et al., 2014), and should thus be regarded as a conservative choice for the purpose of the investigation. At high redshift galaxies are indeed expected to be more gas rich, which is supported by both theoretical arguments and observational evidence.

The disc, modelled as an ideal gas with polytropic index $\gamma = 5/3$, has been set in hydrostatic equilibrium in the global potential with an initial temperature $T = 10^4$ K. In Fig. 3.1 I report the surface density (top panel) and the velocity profile (bottom panel) of the gas component in the disc.

I assumed that previous star-formation episodes left a population of stellar mass black holes in the galaxy nucleus. The mass of such "black hole seeds" $M_{\rm BH}$ has been alternatively set 20 or $100\mathrm{M}_\odot$. I initially distributed the BHs uniformly within the inner 150 pc of the CND. The BHs laid in the disc plane and had an initial velocity equal to the local circular velocity. I added a randomly oriented velocity component sampled from a normal distribution with standard deviation $\sigma \sim 20\%$ of the maximum circular velocity.

In order to create the initial conditions for this study I developed a code, named GD_BASIC[1]. The code computes the density and velocity profiles for both gas and stellar components to guarantee global equilibrium. The code produces two kinds of output, suitable for both grid and particle based codes. The output for particle based codes is stored in the standard GADGET2 binary format, while in the case of grid codes, the gas must be initialised from the computed density and velocity profiles, with the addition of the stellar particles extracted from the standard binary output. The code has also been described in Lupi, Haardt & Dotti (2015).

[1]The name is an acronym for 'Gas disc, BH and stars initial conditions'.The code is publicly available on my personal webpage `http://www.dfm.uninsubria.it/alupi/software.html`

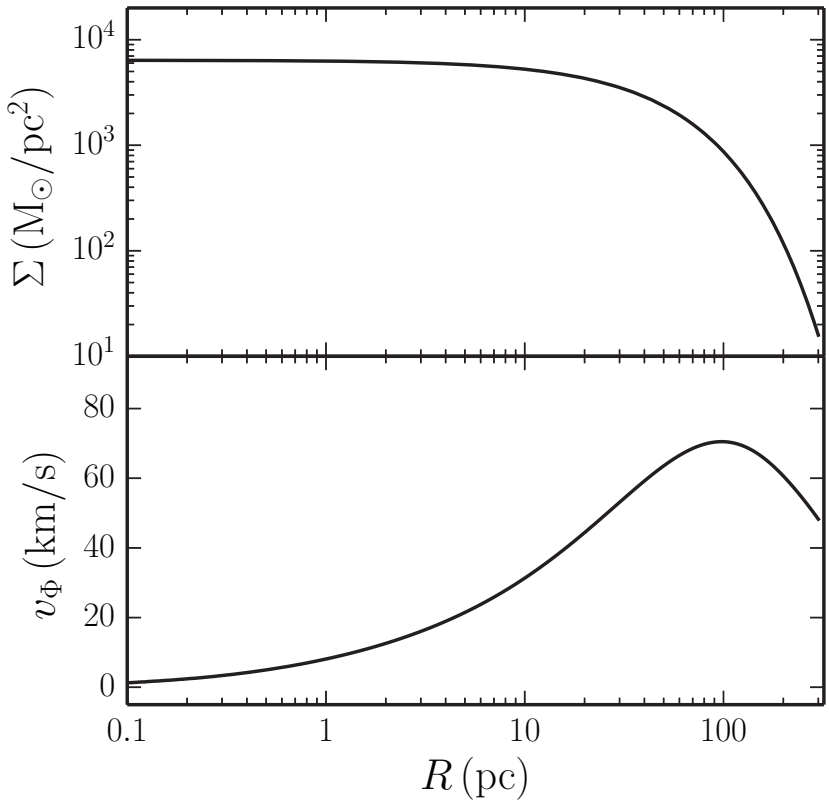

Figure 3.1: *Initial surface density (top panel) and circular velocity (bottom panel) profiles of the circum-nuclear disc.*

3.2 GD_BASIC: A Montecarlo code to model circumnuclear discs in galactic nuclei

I assume that a galactic nucleus is constituted by three different components:

- a stellar spherical structure (termed "nucleus" hereinafter) described by an Hernquist profile (Hernquist, 1990), defined in spherical coordinates as

$$\rho_b(r) = \frac{M_b}{2\pi} \frac{a}{r\,(r+a)^3}, \tag{3.1}$$

  where $\rho_b(r)$ is the density as a function of radius r, $M_b = 2 \times 10^8\,\mathrm{M}_\odot$ the total nucleus mass, and $a = 100$ pc the nucleus scale radius.

- an exponential gaseous disc with surface density profile defined in cylindrical coordinates as

$$\Sigma_d(R) = \frac{M_d}{2\pi R_d^2} \exp(-R/R_d), \tag{3.2}$$

  where R is the disc radius, $R_d = 50$ pc the disc scale radius, and $M_d = 10^8\,\mathrm{M}_\odot$ the total disc mass.

- a MBH with mass $M_{\mathrm{BH}} = 10^7\,\mathrm{M}_\odot$, at rest in the centre of the disc. This last component will be included in the study described in Chapter 4 only.

GD_BASIC solves the disc hydrostatic equilibrium equations for a user-defined surface density profile subject to the additional potentials of the Hernquist nucleus and of the MBH (if present). Defining the gas pressure as

$$P_d = (\gamma - 1)\rho_d u, \tag{3.3}$$

where γ and u are the gas polytropic index and internal energy respectively, and assuming a single temperature disc, the vertical equilibrium equation can then be written as

$$\frac{1}{\rho_d(R,z)} \frac{\partial P_d(R,z)}{\partial z} = -\frac{\partial \phi(R,z)}{\partial z}, \tag{3.4}$$

where $\phi(R, z)$ is the global gravitational potential of the system. Using eq. 3.3, eq. 3.4 can be solved for the disc surface density:

$$\Sigma_d(R) = \rho_d(R, 0) \int_{-\infty}^{+\infty} \exp\left[-\frac{\phi_z(R, z)}{(\gamma - 1)u}\right] dz, \tag{3.5}$$

where $\phi_z(R, z) \equiv \phi(R, z) - \phi(R, 0)$ is the vertical component of the global potential.

Assuming the thin disc approximation, the Poisson equation can be simplified to

$$\frac{\partial^2 \phi_d}{\partial z^2} + \nabla^2 \phi_b = 4\pi G[\rho_d(R, z) + \rho_b(r)], \tag{3.6}$$

where ϕ_d and ϕ_b are the disc and the nucleus potentials, respectively. Since $\nabla^2 \phi_b = 4\pi G \rho_b$, we can write

$$\frac{\partial^2 \phi_{d,z}}{\partial z^2} = 4\pi G \rho_d(R, z) = 4\pi G \rho_d(R, 0) \exp\left[-\frac{\phi_z(R, z)}{(\gamma - 1)u}\right], \tag{3.7}$$

where we defined $\phi_{d,z} \equiv \phi_d(R, z) - \phi_d(R, 0)$ as the vertical component of the disc potential.

To obtain the density profile we must solve the above equations, which force the vertical hydrostatic equilibrium of the disc, assuming a surface density profile as boundary condition. We start solving eq. 3.7 for $\phi_{d,z}$ by guessing an initial equatorial profile $\rho_d(R, 0)$. Then, we compute the total vertical potential ϕ_z, and by means of eq. 3.5, a new value for $\rho_d(R, 0)$ that satisfies the boundary condition Σ_d is derived. The procedure is iterated until convergence. From eq. 3.7 we obtain $\phi_{d,z}$ and, from eq. 3.4, $\rho_d(R, z)$. As in the iterative procedure $\phi_d(R, 0)$ is a free parameter, we assume a razor thin exponential disc (equation 2-168 in Binney & Tremaine). Finally, the velocity of the disc particles is evaluated by setting the radial component of the velocity equal to 0 (hydrostatic equilibrium assures that the vertical component is null as well), while the tangential velocity is obtained from the Euler equation in the case of a rotationally supported disc.

Concerning stellar-like particles, we evaluate the distribution function f in the 6-dimensional phase-space. We initially consider the Hernquist spherical structure subject only to its own potential and to the MBH potential (i.e., $\phi = \phi_b + \phi_{\rm MBH}$), implying that f depends only on the particle total energy in this

case. From the Eddington's formula (equation 4-140a in Binney & Tremaine) we have

$$f(\epsilon) = \frac{1}{2^{3/2}\pi^2}\frac{d}{d\epsilon}\int_{\epsilon}^{0}\frac{d\rho_b}{d\phi}\frac{d\phi}{(\phi-\epsilon)^{1/2}}, \tag{3.8}$$

where ϵ is the particle energy per unit of mass, and ρ_b can be expressed as a function of ϕ (being ϕ a monotonic function of r). The distribution function is numerically evaluated and used to sample the nucleus particle energy density. We then derive the particle speed $v = \sqrt{2(\epsilon - \phi)}$, where ϕ is computed at the particle position. In order to correct for the neglected contribution of the disc to the global potential, we add to the spherically symmetric component of the potential the approximate contribution of the disc in the form $\phi_d = GM_d(< r)/(3r)$, where $M_d(< r)$ is the mass of gas particles within r.

3.2.1 Code test

In order to test the stability of the initial conditions, we try to initialise and evolve an isolated galactic nucleus (with all the three components) for 10 Myr. We assumed an exponential surface density profile, which reproduces the observed profiles of disc galaxies. We adopted a fiducial value for the initial gas temperature of 2×10^4 K. The evolved disc surface density is shown in fig. 3.2 at different times. The profile changes in the inner $\simeq 20$ pc because of a gas instability developing after $\simeq 2$ Myr from the start. In order to assess the origin of such instability, we numerically estimate the Toomre parameter of the disc Q at initial time $t = 0$ ($Q \equiv kc_s/(\pi\Sigma)$, where k is the epicyclic frequency and c_s is the gas sound speed). Note that, strictly speaking, the initialised disc is not infinitesimally thin, so that Q as defined above represents a lower limit. Fig. 3.3 shows the Toomre parameter at $t = 0$ as a function of the radial distance R. We find $Q > 2$ everywhere, with the notable exception of the region $10 \lesssim R \lesssim 150$ pc, where $1 \lesssim Q \lesssim 2$. The formation of transient spiral arms in this region, clearly seen during the disc evolution, suggests a genuinely physical origin of the disc instability. Such instability results in small changes in the surface density profile in the $10 \lesssim R \lesssim 150$ pc region. The system, now slightly out of equilibrium, undergoes a re-adjustment of the gas distribution down to the very central region of the disc, as observable in fig. 3.2 down to 5–10 pc from the MBH.

Finally, no evidence of any fragmentation instability during the overall evolution is seen, in agreement with Q being always $\gtrsim 1$.

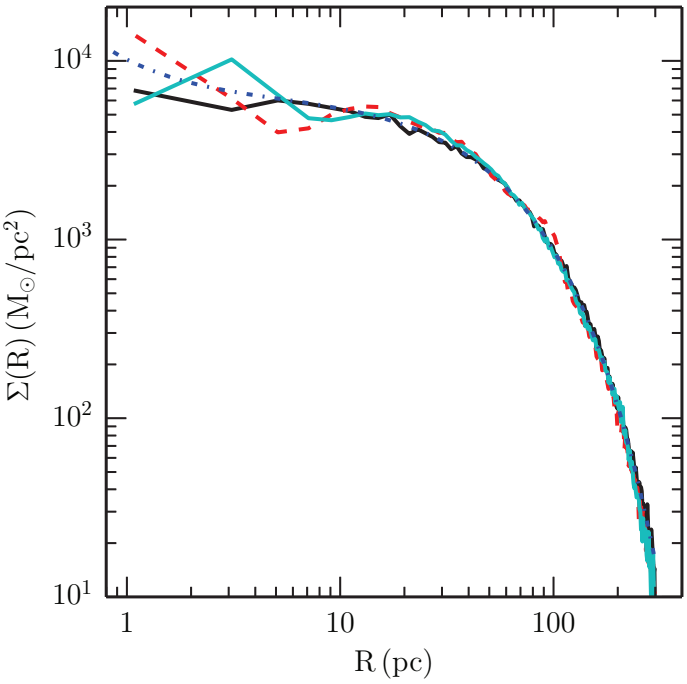

Figure 3.2: *Surface density profile for a disc evolved in isolation using* gadget2 *-2. The solid black, dashed red and solid cyan curves are obtained from the gas particle distribution at $t = 0, 5$ and 10 Myr, respectively. The dash–dotted blue curve is the profile calculated with the algorithm described in the text.*

3.3 Simulation setup

I performed a suite of 6 simulations, using two different codes, the AMR code RAMSES (Teyssier, 2002) ("_R" runs) and the mesh-free code GIZMO (Hopkins, 2015) ("_G" runs). The use of two different powerful numerical techniques is aimed at checking the robustness and reproducibility of our results.

3.3.1 RAMSES Eulerian simulations

I performed two simulations with RAMSES at two different spatial resolutions, namely 0.4 pc ("low" runs) and 0.1 pc ("med" runs). The mass resolution was

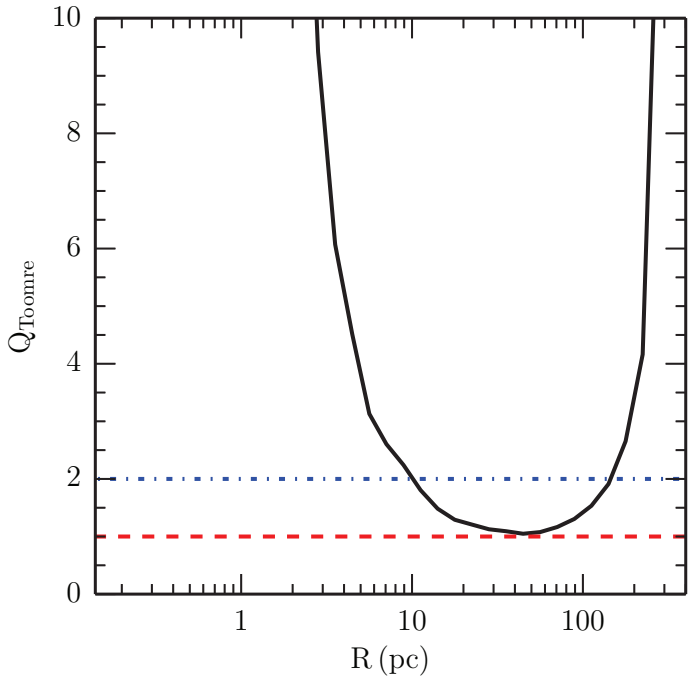

Figure 3.3: *Toomre parameter at $t = 0$ Myr for the isolated disc using the thin disc approximation. The dashed red line corresponds to $Q = 1$, while the dash–dotted blue one to $Q = 2$. The y-axis has been limited in the interval between 0 and 10 to highlight the instability interval, corresponding to $Q < 2$. The region within the inner 1 pc is not visible because of the very large values assumed by Q, outside the axis limits considered.*

$10^3\,\mathrm{M}_\odot$ at the quasi-Lagrangian threshold for refinement. I included the radiative cooling of the gas adopting the standard prescriptions employed in the code (see Teyssier et al., 2013, for details). In order to prevent spurious fragmentation at the highest refinement level I added a polytropic pressure term to the gas component (described as a polytrope with $\gamma = 5/3$ and temperature 10^3 K at $2 \times 10^5\,\mathrm{cm}^{-3}$), ensuring to resolve the Jeans length with at least 4 cells at the highest refinement level.

I set a star formation density and temperature threshold of $\rho_{\rm thr} = 2 \times 10^5$ cm^{-3} and $T_{\rm thr} = 2 \times 10^4$ K, and a typical star formation time-scale of 1 Myr. I also assumed a time delay between star formation and the corresponding SNa explosion event of 1 Myr, with a SNa yield of 0.15 (corresponding to stars with masses above $8\,\mathrm{M}_\odot$ for a Salpeter IMF). In order to model non thermal processes associated with SNa events, I included the blast-wave like

feedback described in Teyssier et al. (2013). In this feedback recipe the SNa energy budget is decoupled from the thermal energy of the gas, preventing, for a typical timescale $\simeq 20$ Myr, the gas to radiatively cool. In our runs I assumed a primordial gas composition and included subsequent metal pollution due to SNe.

3.3.2 GIZMO Lagrangian simulations

I used the same prescriptions of my RAMSES runs for both cooling and SF. Regarding SNa feedback, in GIZMO runs I limited the cooling delay time to 5 Myr only (i.e., 4 times smaller than what assumed in RAMSES runs). I checked that this set up provided consistent results between RAMSES and GIZMO feedback implementations.

I performed three simulations allowing for two different gravitational resolutions, i.e., 0.16 pc ("low" runs) and 0.02 pc ("high" run). I set the same gravitational resolution for gas particles and BH particles. I used 10^5 particles for the "low" runs and 10^7 for the "high" run, corresponding to a mass resolution of $10^3\,\mathrm{M}_\odot$ and $10\,\mathrm{M}_\odot$, respectively. In these runs I used the finite mass mode available in the code, in which mass transfer between particles is forbidden, so that our simulations were purely Lagrangian.

3.3.3 BH accretion

In both RAMSES and GIZMO runs I evaluated the accretion onto the stellar mass BHs using the so-called "flux accretion" prescription. In such a scheme the accretion rate is the mass flux rate within the BH accretion zone (see sections §2.1 and §2.2), i.e.,

$$\dot{M}_{\rm acc} = \int -\nabla \cdot [\rho \Delta \mathbf{v}] d^3\mathbf{x}, \tag{3.9}$$

where the integral is over the volume of the accretion zone and $\Delta\mathbf{v}$ is the gas-BH relative velocity (see Bleuler & Teyssier, 2014, for a detailed description of the implementation).

In order to get a more accurate BH dynamics and to best resolve the accretion rate in RAMSES runs, I forced the region near to each BH to always be at the maximum refinement level, as described in section §4.3. Forcing the resolution close to the BHs at the highest possible level guaranteed that nearby

cells owned a mass $\lesssim 5 \times M_{\rm BH}$ during the whole BH accretion history. This allowed me to set $20\,{\rm M}_\odot$ as the initial mass of the BHs in RAMSES runs.

In GIZMO runs, instead, the large mass of gas particles did not allow me to properly resolve the dynamics of BHs as light as $20\,{\rm M}_\odot$. Therefore, I started from a larger initial BH mass, i.e., $M_{\rm BH} = 100\,{\rm M}_\odot$. With such a choice, BHs in the "high" runs had resolved dynamics since the very beginning of the simulation. In the "low" case, the initial BH dynamics and growth was instead affected by the lack of mass resolution. However, as will be discussed in the next section, some BHs grew above $1000\,{\rm M}_\odot$ in a very short time, making dynamics quickly reliable.

With respect to the standard Bondi-Hoyle model, the "flux accretion" recipe does not make any geometrical assumption for the gas flow, allowing for a more accurate estimation of the accretion rate, where the effect of angular momentum on the resolved scales is taken into account. However, despite the high resolution reached with the "high_G" run, I was unable to properly follow the gas from sub-parsec scales down to the accretion disc scale. This resolution limit could lead to overestimated and more efficient accretion. However, such a convergence study is beyond the scope of this study.

3.3.4 BH feedback

In these simulations I suitably modified the standard recipe for BH feedback to include the effects of accretion in the fashion of slim disc (Sądowski et al., 2014). To this aim, I estimated ϵ using the analytical fit to the numerical results by Sądowski et al. (2014) provided by Madau, Haardt & Dotti (2014):

$$\epsilon = \frac{r}{16} A(a) \left[\frac{0.985}{r + B(a)} + \frac{0.015}{r + C(a)} \right], \tag{3.10}$$

where $r = \dot{M}_{\rm E}/\dot{M}$. Here $\dot{M}_{\rm E} = 16 L_{\rm E}/c^2$ where $L_{\rm E}$ is the Eddington luminosity. A, B, C are fitting functions scaling with the BH spin a as

$$A(a) = (0.9663 - 0.9292a)^{-0.5639}, \tag{3.11}$$
$$B(a) = (4.627 - 4.445a)^{-0.5524}, \tag{3.12}$$
$$C(a) = (827.3 - 718.1a)^{-0.7060}. \tag{3.13}$$

At each accretion event I computed the released energy allowed to feedback on nearby particles/cells (in GIZMO runs) or cells (in RAMSES runs) using this new value for ϵ instead of the fixed value 0.1, while the BH spin has always been fixed at $a = 0.99$ for all BHs. In all simulations I did not include other possible forms of BH feedback, e.g., momentum-driven.

Finally, in order to check whether super-critical accretion is instrumental in leading to very large $M_{\rm BH}$ in a short time, I performed two GIZMO runs setting the radiative efficiency to its custom value, $\epsilon = 0.1$ (low_G_0.1 and high_G_0.1 runs). The details of these six simulations are reported in table 3.1.

Run	Resolution (pc)	BH mass ($\rm M_\odot$)	Accretion radius (pc)	ϵ
low_R	0.40	20	1.6	Slim
med_R	0.10	20	0.4	Slim
low_G	0.16	100	< 1.6	Slim
high_G	0.02	100	< 0.2	Slim
low_G_0.1	0.16	100	< 1.6	0.1
high_G_0.1	0.02	100	< 0.2	0.1

Table 3.1: *Settings of our simulation suite. The second column reports the gravitational resolution (for _G runs) and the highest refinement level resolution (for _R runs). The fourth column is the accretion radius, which is fixed to 4 cells for _R runs and depends on the smoothing length for _G runs. The last column indicates the type of accretion recipe used.*

3.4 Results

Figure 3.4 shows the comparison between the low resolution GIZMO runs with (low_G) and without (low_G_0.1) the slim disc implementation. All the other simulation parameters are the same in the two simulations. It is immediately clear from the comparison that whenever a BH undergoes an intense accretion episode, the large feedback energy available in the radiatively efficient low_G_0.1 case evacuates the BH surroundings, efficiently limiting further BH growth. In the low_G case, on the contrary, even accretion rates significantly higher than $\dot{M}_{\rm E}$ result in moderate luminosities that do not impact on the densest gas clumps, and therefore BHs can grow considerably faster. As

an example, in the low_G run the most mass growing BH (that will be referred to as $\mathrm{BH_{top}}$ in all runs hereafter, red line in the bottom-left panel of figure 3.4) reaches a mass larger by up to 2 order of magnitudes compared to the corresponding $\mathrm{BH_{top}}$ in the radiative efficient case at the end of the simulation (red line in the top-left panel panel). The low radiative efficiency of slim discs has then a double effect: first, for any given accretion rate BHs grow faster simply because less mass is lost as radiation (the "$(1-\epsilon)$-effect"); second, the reduced radiative efficiency results in a reduced feedback on the accreting gas, and larger accretion rates are therefore possible (the "$\dot{M}$-effect").

In order to asses how numerical resolution affects the results, I analysed the two high resolution GIZMO runs (high_G and high_G_0.1), and compared the outputs to the low resolution cases discussed above. Figure 3.5 shows the accretion history of BHs (left panels) and the effect the accretion feedback has on the gas (right panels). Because of the higher resolution the accretion region around each BH which can be resolved is smaller, and this has the net effect of reducing the BH mass growth compared to the corresponding low resolution runs. Nevertheless, it is apparent how, also in these high resolution simulations, BH mass growth is strongly suppressed in the radiatively efficient case (top panels). Indeed, for $\epsilon = 0.1$, $\mathrm{BH_{top}}$ increases its mass by only $\simeq 50\%$ of its initial value. I want to stress again that the different radiative efficiency is only marginally responsible of the different accreted mass in the two cases. As clearly shown in the right panels of Figure 3.5, the largest effect is played by the accretion feedback that, in the standard high-efficiency case, evacuates the region closer to the BHs, hence inhibiting further gas accretion.

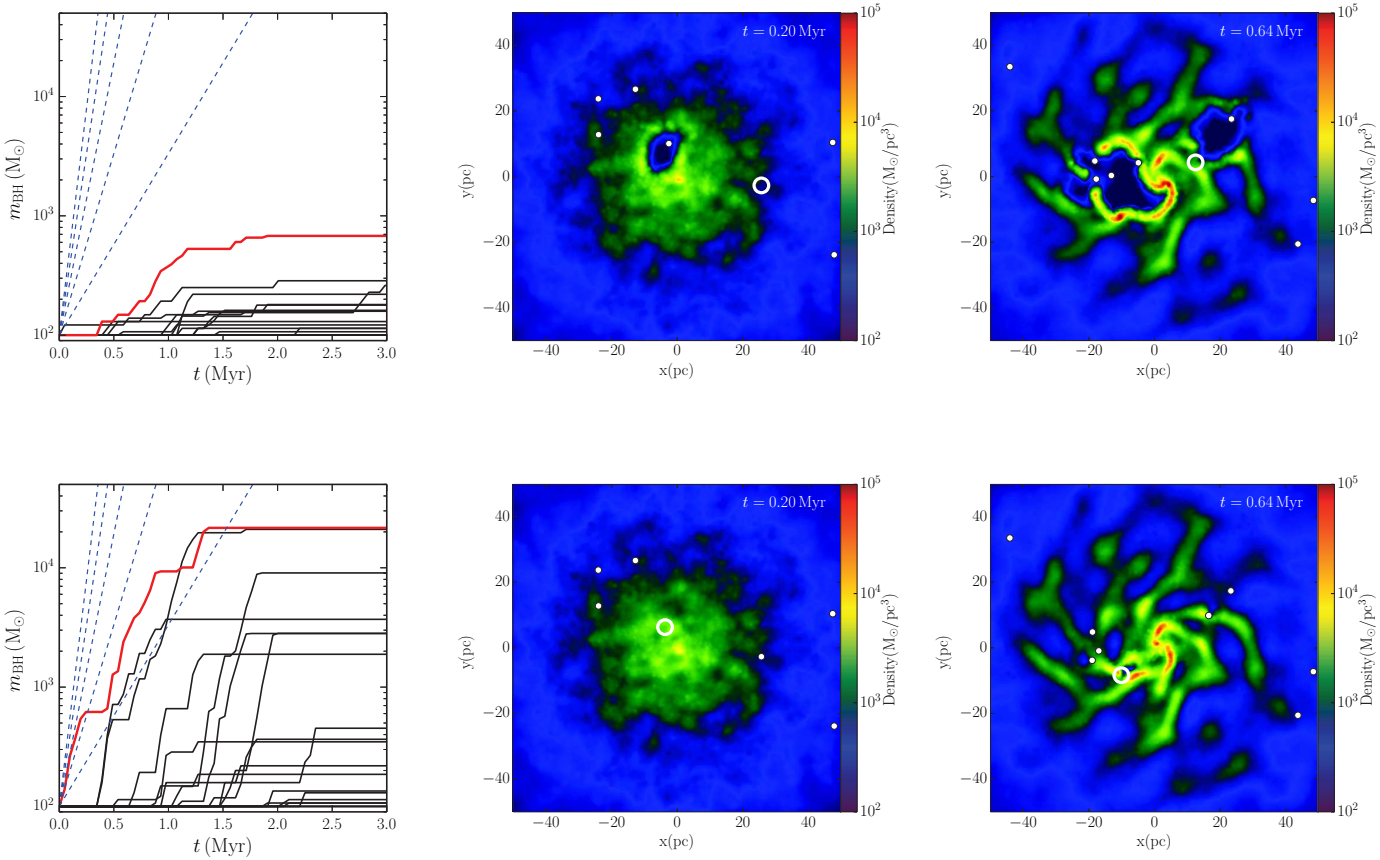

Figure 3.4: *Left panels: BH masses as a function of time for runs low_G_0.1 (top panel) and low_G (bottom panel). The red lines correspond to the most massive BHs ($BH_{\rm top}$) at the end of the runs, while the blue dashed lines trace accretion histories at fixed Eddington ratios of 500, 400, 300, 200 and 100, respectively. Central and right panels: gas density maps for the two runs at $t = 0.2$ and 0.64 Myr, respectively. The white dots mark the positions of the BHs.*

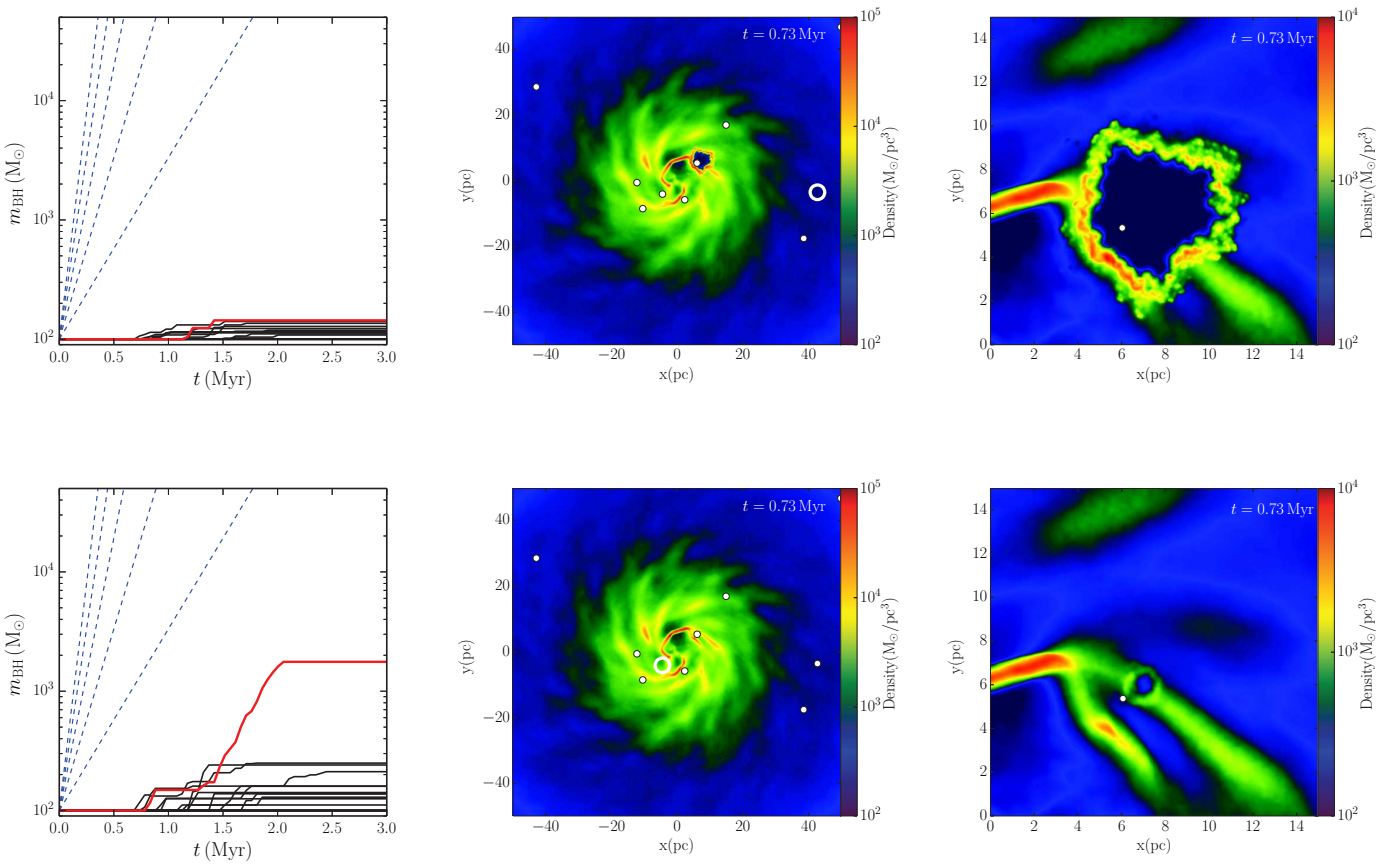

Figure 3.5: *Left panels: BH masses as a function of time for runs high_G_0.1 (top panel) and high_G (bottom panel). The red lines correspond to the most massive BHs ($BH_{\rm top}$) at the end of the runs, while the blue dashed lines trace accretion histories at fixed Eddington ratios of 500, 400, 300, 200 and 100, respectively. Central panels: gas density maps for the two runs at $t = 0.73$ Myr. Right panels: zoom in of a region heated by BH feedback. The white dots mark the positions of the BHs.*

The implementation of a physically motivated radiative inefficient accretion mode is then a necessary condition for a fast, highly super-Eddington growth of BHs in my simulations, but, as I will show next, is not sufficient. In the following I will focus only on runs including the slim disc prescription, in order to link episodes of super-Eddington growth with the physical state of the BHs and of the nuclear disc, with the ultimate aim of understanding the processes that can possibly lead to high accretion rates.

Figure 3.6 shows the results of the highest resolution RAMSES run med_R. The upper left panel reports the mass evolution of the 20 BHs as a function of time. As for the low_G simulation discussed above, the implementation of the slim disc efficiency prescription allows $\mathrm{BH_{top}}$ (shown as a red line) to grow within 3 Myr by up to ~ 3 orders of magnitude in mass. Note that $\mathrm{BH_{top}}$ is not necessary the earliest growing BH of the cluster.

The upper right panel of figure 3.6 focuses on $\mathrm{BH_{top}}$ alone, showing the time evolution of the accretion rate, and the corresponding distance from the gas clump the BH bounds to during the peak of its mass growth. The accreting clump forms out of a spiral stream developing in the cooling disc, and can not be clearly identified as a bound structure before $t \approx 1$ Myr, as shown in the middle left panel. $\mathrm{BH_{top}}$ passes a first time through the overdense stream (middle right panel), and experiences a short $\lesssim 0.1$ Myr super-Eddington accretion episode, but the radial component of its velocity quickly is large enough to displace it from the overdensity (as observed in the $\dot{m}_{\mathrm{BH}}$ plot, upper right panel). As the clump grows in mass (up to a maximum of $\sim 3 \times 10^4\,\mathrm{M_\odot}$ in gas), the $\mathrm{BH_{top}}$ feels its gravitational attraction, and is eventually captured by the clump. At this time $\mathrm{BH_{top}}$ undergoes a longer (~ 0.5 Myr) intense super-Eddington accretion phase. Being the initially small BH surrounded by an overwhelmingly large and cold gas cloud, the BH accretes at the maximum rate allowed by the code (i.e. $500 \times \dot{M}_{\mathrm{E}}$) until almost all gas is turned into stars. At this point $\mathrm{BH_{top}}$ (already grown by 3 order of magnitudes in mass), together with stars exploding as SNe, can evacuate the residual gas condensation (lower right panel). Note that BHs (including $\mathrm{BH_{top}}$) accrete most of their mass from, essentially, a single dense clump they randomly come across during the dynamical evolution of the system.

It is important to realise that the gravitational capture of a BH by a dense gas clump is intrinsically stochastic, as clumps form in the disc via gravitational instabilities of cooling gas independently of the presence of seed holes.

While the BH-capture process is common in all the performed simulations, the number and mass distributions of gas clumps and, consequently, the fraction of BHs that bind to them, in fact depend upon the spatial and mass resolution achieved. Figure 3.7 shows a comparison between runs with different spatial resolution. Among the runs including the slim disc implementation, only run med_R (already shown in Figure 3.6) is left out of the direct comparison.

A first clear difference is observable at early times. The runs with lower resolution show a faster initial growth of each individual BH, and the number of growing BHs right after the beginning of the runs ($t \lesssim 0.5$ Myr) also increases with decreasing resolution. These trends are caused by the larger accretion radius implemented in the lower resolution runs. In these simulations the BHs can start accreting well before the disc develops any significant overdensity. For this reason the feedback of the early BH accretion onto the gas is more efficient, as a larger energy is injected in a lower density medium. As the resolution increases and the accretion radius can be decreased, fewer BHs have an early start, as in the med_R run (upper left panel of Figure 3.6) and, more evidently, in the high_G run (lower left panel of Figure 3.7).

The high_G run, thanks to the exquisite mass and spatial resolutions achieved, shows a richness of structures observable directly in the density map (see the lower right panel of Figure 3.7 in particular), in which the formation of dense clumps as well as the feedback exerted by the ongoing SF are clearly visible. The gas particles tracing the gas evolution allow us to follow the formation of the massive clump from which $\mathrm{BH_{top}}$ gains its mass. Figure 3.8 reports two different projections of $\mathrm{BH_{top}}$ orbit along with the trajectories of 50 gas particles randomly extracted from those forming the massive clump $\mathrm{BH_{top}}$ binds to and accretes from. The clump formation clearly proceeds out of a gas gravitational instability within the dense disc, and starts interacting with $\mathrm{BH_{top}}$ only when their orbits intersect. Strong gravitational perturbations to the BH orbit are clearly seen when the two systems bind gravitationally. The BH growth then exerts a feedback onto the gas particles, that, together with stars exploding as SNe, results in a partial ejection of particles from the BH neighbourhoods and out of the disc plane (as clearly see in Figure 3.8 lower panel).

In summary, I want to note that in these idealised runs the growth of the BHs is finally halted by the star formation-driven gas consumption, and by gas ejection triggered by SNe. However, in a cosmological perspective, the galaxy

nucleus would be replenished of gas coming from large scale filaments and/or galaxy mergers. The very short duration of the super-Eddington accretion bursts allows for the growth of stellar mass BHs up to $\gtrsim 10^4\,\mathrm{M}_\odot$ or more on a time comparable (or even shorter) than the star-formation timescale.

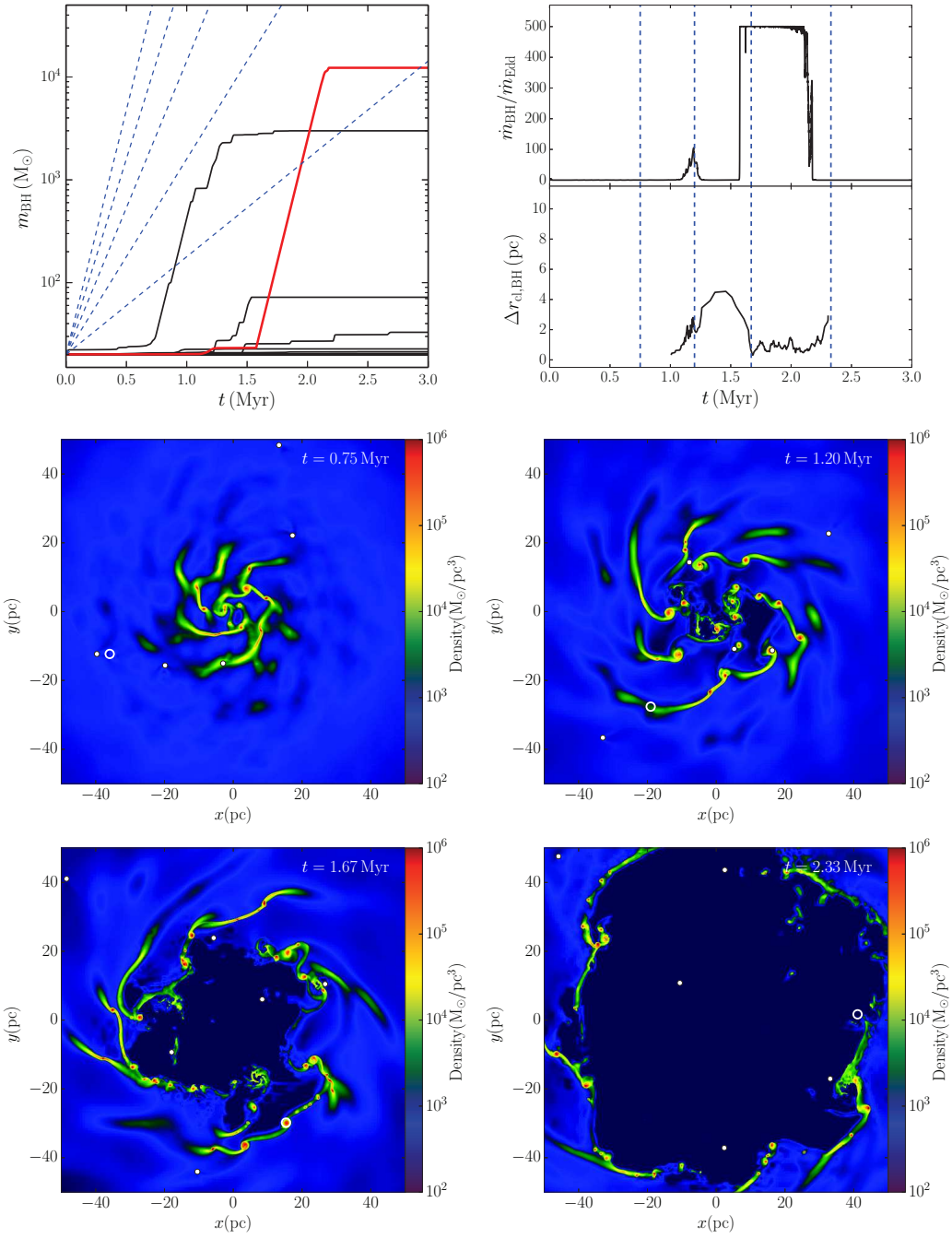

Figure 3.6: *Run med_R. Upper left panel: BH masses vs time for all the 20 BHs. The dashed lines show the slope of accretion episodes at 500, 400, 300, 200 and 100 $\dot{M}_{\rm E}$. Upper right panel: accretion rate for BH$_{\rm top}$, and distance from the clump BH$_{\rm top}$ bounds to during the peak of its mass growth. Middle and lower panels show the density in the equatorial disc plane of the gas at $t = 0.75$, 1.2, 1.67 and 2.33 Myr (corresponding to the times highlighted by the dotted lines in the upper upper right panel). The BH$_{\rm top}$ is reported as large white ring, while the other BHs are shown as smaller white dots.*

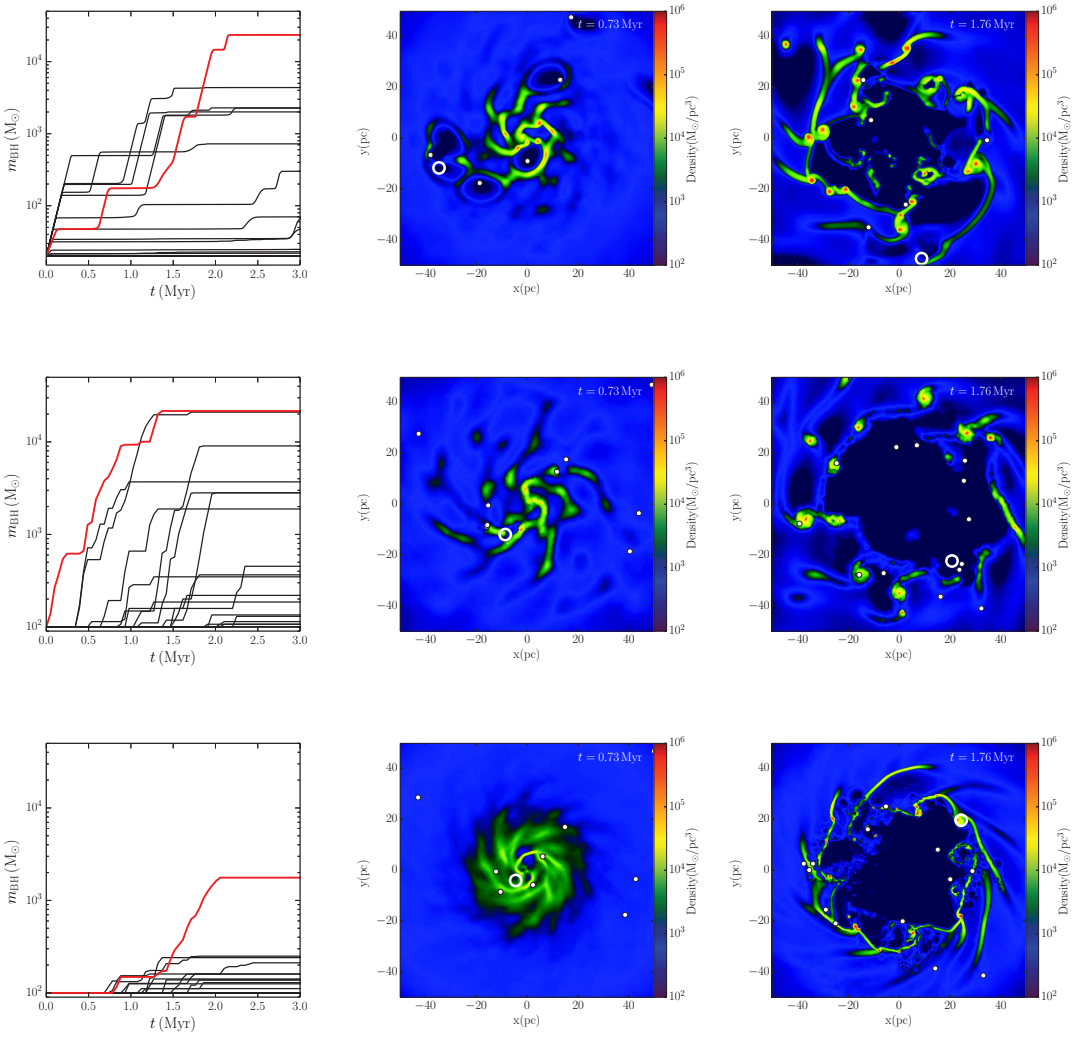

Figure 3.7: *Upper panels: Mass as a function of time of the BHs (left panel), gas density at* $t = 0.73$ *Myr (central panel) and* $t = 1.76$ *Myr for the low_R run. The positions of the BHs are shown as white dots. The growth of* $BH_{\rm top}$ *is highlighted in the left panel with a red line, and its position in the central and right panels is marked with a large white ring. Middle and lower panels, the same as the upper panels for run low_G and high_G, respectively.*

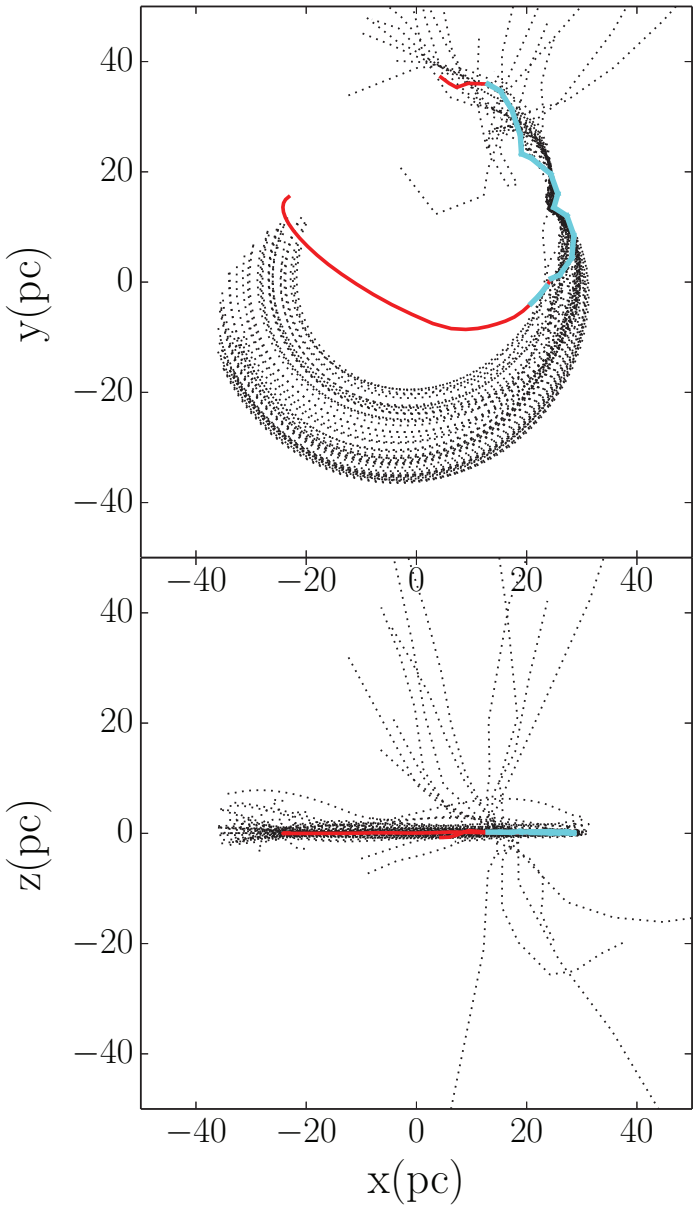

Figure 3.8: *Solid red lines show the face-on (top panel) and edge-on (bottom panel) projections of the trajectory of $BH_{\rm top}$ in run high_G. Black dotted lines trace the orbits of a sample of the gas particles forming the gas clump $BH_{\rm top}$ binds to. The accretion burst due to the $BH_{\rm top}$-clump interaction is highlighted in cyan.*

Chapter 4

Massive black hole binaries in gas-rich galaxy mergers: formation and evolution

Along the cosmic history galaxies assemble their mass via accretion and mergers of smaller units. As described in Chapter 1, MBHs harbouring the centre of these galaxies sink towards the centre of the merger remnant (when the mass ratio between the colliding galaxies is $\gtrsim 1:10$) forming a Keplerian binary. The MBHB shrinks via gas torques (if gas is present) and three-body interactions with stars in the galaxy nucleus and finally undergo coalescence, driven by gravitational wave emission.

To simulate MBHs in galaxy mergers we need a detailed treatment of the dynamics from hundred kpc scales, when galaxies start interacting, down to few 10^{-3} pc, when gravitational wave emission becomes dominant, since it influences the MBH ability to accrete gas (hence, its mass evolution and possible onset of AGN activity), the MBH spin evolution (e.g., Dotti et al., 2010), and the formation and fate of MBHBs. Because of the large dynamic range involved, our ability of describing the whole process is strongly limited. To bypass this problem, different solutions have been proposed, aimed at investigate different phases of the BH route to coalescence: (i) galaxy merger simulations with resolution of $\lesssim 10$ pc, which describe the BH pairing under the action of

gravitational torques by dark matter, gas and stars; (ii) idealised simulation of a single massive circum-nuclear disc in the merger remnant nucleus; (iii) simulations of a pair of circum-nuclear discs, each hosting a MBH in its nucleus, which collide on scales of few hundred pc.

With the smoothed particle hydrodynamics (SPH) code GASOLINE, case (i) was explored by Roškar et al. (2015) who simulated a prograde in-plane 1:1 merger of two late-type galaxies (Milky Way like) where star formation and feedback were turned on at the onset of the simulation to generate a multiphase medium. 5 Gyr after the start of the simulation, particle-splitting of the baryonic particles was performed in an excised zoom-in region to follow the last 100 Myr of evolution when the galaxy's cores touch on scales of ~ 5 kpc. The nature of the multiphase gas which develops clumps affects the MBH dynamics. The MBHs undergo gravitational encounters with massive gas clouds and stochastic torquing by both clouds and spiral modes in the disc relents the pairing process. The MBHs are kicked out of the plane due to their interaction with clumps and this delays the time of binary formation which now is ~ 100 Myr, about two orders of magnitude longer than in the idealised mergers with one-component gas. Thanks to the adaptive mesh refinement technique, instead, a number of grid-based hydro simulations of galaxy mergers have been performed to date (Chapon, Mayer & Teyssier, 2013; Dubois et al., 2014). The two papers assume quite different prescriptions. Chapon, Mayer & Teyssier (2013) assume a smoother IGM, not affected by cooling, star formation (SF) and supernova (SNa) feedback, while these effects are considered in Dubois et al. (2014).

Type (ii) investigations have been performed by del Valle et al. (2015), who recently simulated with an SPH code the sinking of two MBHs in a massive 10^9 $\mathrm{M}_\odot$ circum-nuclear disc with gas forming stars. The orbits of the MBHs are erratically perturbed by the gravitational interaction with the clumps that form as a result of disc's fragmentation, delaying the orbital decay of the MBHs if compared with similar runs with a one-component gas: typical decay times are found close to 10 Myr, when the MBHs are seeded in the disc initially at $\sim 200 - 100$ pc scales. The key result which emerges from these new studies is that the MBH dynamics is sensitive not only to the time varying gravitational background of a merger itself, but also on how fragmentation of gas clouds, star formation and supernova (SNa) feedback shape and change the thermodynamical state of the gas, considered to play a key role

in guiding the orbital decay of the MBH. The mass distribution of the star-forming clumps appears to be a relevant parameter which affects the degree of stochastic forcing of the MBH orbit and the distribution of the sinking times from ~ 100 pc scale down to 0.1 pc.

In this study I used type (iii) simulations, in the aim at describing the intermediate stages of a gas-rich galaxy merger, in which both progenitor galaxies hosted a MBH surrounded by a circum-nuclear disc, which is in the verge of merging within the inner kpc of the merger remnant. The key question of my study was understanding the MBH dynamics in a multiphase gas shaped by cooling, SF and SNa feedback and the role of SNa feedback in shaping the gas mass distribution around the MBHB. The transit from the binary phase II to phase III of gravitational wave inspiral depends on the strength of gas-driven migration in a circum-binary disc surrounding the MBHB (Cuadra et al., 2009; Shi et al., 2012; Roedig et al., 2011, 2012; del Valle & Escala, 2012). My first attempt was to explore under which conditions a circum-binary disc forms around the two MBHs and how this depends on the recipes adopted to model the physics of star-forming regions. The simulations have been performed with the AMR code RAMSES , due to the better treatment of gas shocks with respect to SPH codes (Agertz et al., 2007) which allowed me to accurately describe the gas dynamics when the two gaseous discs collide.

4.1 Initial conditions

I initialised each of the two merging nuclei following the procedure described in section §3.2, by means of the code GD_BASIC. I built two equal mass co-rotating gaseous discs, each described by 10^5 particles, with an initial gas temperature of 2×10^4 K and a polytropic equation of state with index $\gamma = 7/5$. I initially relaxed them adiabatically for about 10 Myr to ensure stability. The discs were initially set at 300 pc on an elliptical orbit with eccentricity $e = 0.3$, and with orbital angular momentum *antiparallel* to the angular momentum of the discs. I stress that each galaxy disc plane is in principle uncorrelated to the orbital plane of the merger, and, to the first order, the same is valid for the CNDs[1]. I arbitrarily chose the geometry that maximises the impact of the

[1]Here I neglect the possible tidal effect exerted by one disc onto the other. This effect would tend to align (or antialign) the two discs, enhancing the chances of having an orientation

two discs along their orbit and that ensures the highest cancellation of angular momentum, enhancing the inflows toward the centremost regions. Such a geometry has not been explored in the literature yet.

Since the system considered was more complex than in the case of single CND, the initial conditions for the AMR runs were obtained by mapping the gas particle distribution produced by GD_BASIC on a grid using the publicly available code TIPGRID.[2] The maximum spatial resolution (at the highest refinement level) for all my simulations was ~ 0.39 pc and the mass resolution for particles forming the stabilising stellar nucleus was $2 \times 10^3\,\mathrm{M}_\odot$. I used the standard Quasi-Lagrangian and Jeans criteria already implemented in RAMSES, as described below. The Quasi-Lagrangian criterion allowed me to resolve a minimum gas mass of $10^3\,\mathrm{M}_\odot$ everywhere. The standard Jeans criterion, on the other hand, ensured that the Jeans length was resolved with at least 4 cells everywhere, so to avoid the formation of spurious clumps due to resolution limits. I also added a pressure support term, modelled as a polytrope with $\gamma = 5/3$ and temperature 2×10^3 K at the star formation threshold, in order to avoid the formation of spurious clumps due to resolution limits.

4.2 Sub-grid physics

In these runs I assumed gas with primordial composition, optically thin and I allowed the gas to cool down under lines and continuum emission. I also included stellar particle creation for gas matching two criteria: (i) the gas temperature dropped below 2×10^4 K, and ii), the gas density in a cell exceeded a pre-defined value. I assumed a typical star formation (SF) timescale of 1.0 Myr and I set the SF density threshold alternatively to $n_\mathrm{H} = 2 \times 10^5$ or $n_\mathrm{H} = 2 \times 10^6$ cm^{-3}, where n_H is the local hydrogen number density. In order to model SNa explosions, I considered each stellar particle as a stellar population following a Salpeter IMF, and a SNa yield of 15%. I further employed two different recipes for SNa feedback. In both SNa feedback recipes the energy budget associated ($10^{50}\,\mathrm{erg}/\,\mathrm{M}_\odot$) is completely released in the parent cell as purely thermal energy. The first recipe (termed "thermal feedback")

between the two CNDs similar to the one I assumed as initial conditions.

[2]The code is available at `http://www.astrosim.net/code/doku.php?id=home:code:analysistools:misctools`.

assumes that the heated gas starts cooling right after the SNa event; the second recipe (termed "blast wave feedback") assumes instead that the energy released by SNe is decoupled from the gas radiative cooling, i.e., it is not radiated away for ~ 20 Myr (Teyssier et al., 2013) and this triggers the formation of a momentum-driven blast wave. This latter scheme is aimed at modelling non-thermal processes energising the blast wave, which are characterised by timescales longer than thermal processes (see e.g, Enßlin et al., 2007). I usually assumed that no star formation occurred within the two discs before the merger, and that SNe exploded after a time $\Delta t_{\rm SN} = 10$ Myr. Stellar mass particles forming the stabilising bulge were not allowed to release energy as SNe. No gas accretion on to the MBHs and AGN feedback has been included in any of the runs.

4.3 Massive particle dynamics in RAMSES

As described above, for this study I used the AMR code RAMSES , which behaves very well when one needs to resolve hydrodynamics, especially when shock occur in the gas. However, the scheme used in grid codes to solve gravitational interactions suffers from numerical noise when single massive particles are considered. These effects have been already observed in galaxy merger simulations (Chapon, Mayer & Teyssier, 2013; Dubois et al., 2014). In Chapon, Mayer & Teyssier (2013) the MBH dynamics depends strongly on the maximal resolution of the simulation. In lower resolution runs ($\Delta x_{\rm min} = 3$ pc) the MBH evolution was significantly slower (because of the underestimated effect of the resolution dependent dynamical friction), and considerably more noisy (well above the resolution level) than in the higher resolution cases ($\Delta x_{\rm min} = 0.1$ pc). A similar noisy evolution of the MBH orbits has been observed by Dubois et al. (2014). In order to prevent spurious oscillations of the MBH due to finite resolution effects, authors introduced an additional drag force onto the MBHs.

Interestingly, a noisy evolution of collisionless particle dynamics (and, in particular, of MBHs) has been also observed in high-resolution AMR simulations of single isolated galaxies, in which the gas is only subject to internal processes such as star formation, SNa feedback, etc. (e.g. Gabor & Bournaud, 2013). In their work, in order to limit numerical MBH wandering, the authors proposed two different approaches. The first one consisted in modelling the

MBH as an extend spherical structure, using few thousands evolving particles. Such BH-forming particles were regenerated over a secondary, coarse time grid. In this case the MBH moved out of the geometrical centre of the galaxy by hundreds to thousands of pc depending on the amount of gas simulated and the noisy effect was only reduced. The second one, instead, consisted in adding an artificial velocity component directed toward the stellar centre of mass, which forced the MBH to orbit close to the galaxy centre.

As noted by Gabor & Bournaud (2013), the noisy evolution of the MBH could be either numerical (due to the limited and time varying spatial resolution), or physical, if caused by interactions with massive and dense gas clouds. This last possibility was particularly interesting when the gas was allowed to cool and actually formed significant compact overdensities, as in the simulations discussed in Gabor & Bournaud (2013) and Dubois et al. (2014). Indeed, a physically motivated noisy orbital evolution of MBHs has been observed in SPH simulations (see e.g. Fiacconi et al., 2013, for a detailed and extensive discussion). It is important to notice that the effects of massive gas clumps on the MBH dynamics are severely altered by the corrections proposed in works discussed above. In this study I considered an alternative solution, which does not alter the MBH dynamics and allows for a more accurate dynamical evolution of both MBHs and gas in their surroundings.

I developed a new refinement criterion aimed at ensuring a fixed accuracy when computing the gravitational force acting on the two MBHs. The new refinement criterion is based on the identity and positions of selected particles, rather than on the global geometry of the system. In my new implementation, refined grids follow the positions of the two MBHs at each time-step. Surrounding cells within two specified, MBH-centred volumes are flagged for further multi-level refinement. Up to $N_{\rm level}$ concentric regions of increasing resolution, where $N_{\rm level}$ is the number of refinement levels used, can be user-defined by setting the corresponding radii. For example, in the runs I will discuss later I enforced the maximum level of refinement, with single cell linear sizes of 0.39 pc, within 10 pc from each MBH, using seven levels of refinement above the coarse resolution level. At larger distances from the MBHs the resolution degrades smoothly unless another refinement criterion is matched.

4.3.1 Simulation suite

I performed a total of five simulations, in order to compare the new refinement criterion with the standard one, and to test the reliability of the dynamical evolution of the MBH binary under different assumptions regarding the sub-grid physics. I also carried out a single simulation with GADGET2 , assuming the same polytropic equation of state with index $\gamma = 7/5$ as in RAMSES runs. The spatial resolution for the SPH simulation was 0.2 pc, while the mass resolution was $10^3\,\mathrm{M}_\odot$ and $2 \times 10^3\,\mathrm{M}_\odot$ for gas and collision-less particles, respectively, equal to the mass resolution in the AMR runs.

Run	Cooling	Star Formation	SNa feedback	New Refinement
Plain	No	No	No	No
Plain+	No	No	No	Yes
Noblast	Yes	Yes	Yes	No
Blast	Yes	Yes	Yes	No
Blast+	Yes	Yes	Yes	Yes

Table 4.1: *The suite of* ramses *runs.*

Table 4.1 shows the suite of RAMSES simulations with the main features highlighted.

In runs "Plain" and "Plain+" no sub-grid physics has been included, while in the following three runs, termed "Noblast", "Blast" and "Blast+", I included both gas cooling and star formation (see section §4.2 for details). I assumed a density threshold for SF of 2×10^6 cm^{-3}. In "Noblast" runs I used the "thermal feedback" recipe, while in the "Blast" and "Blast+" runs I used the "blast wave feedback" one. I want to note that the typical timescale for the onset of SNe is much longer compared to the typical gas inflow timescale in my simulations, i.e., SNe would have little/negligible effects on MBH and gas dynamics. In order to enhance feedback effects, I assumed no time delay for the onset of SNe after star formation. The new dynamic refinement criterion has been implemented in the two "+" runs.

4.3.2 Results

The upper panels of Figure 4.2 show the MBH pair orbit in run Plain compared to the SPH run. While the orbital evolution computed by GADGET2 -2 shows a smooth orbital decay of the pair, run Plain shows an abrupt change in the direction of motion of the two MBHs after $\sim 2-3$ Myr from the beginning of the run. At this time the MBHs suddenly leave the gas (upper panels in Figure 4.3) and stellar overdensities they inhabited. Such an abrupt acceleration could, in principle, have a physical explanation. For example, the sudden swerve could be the outcome of short range encounters between the MBHs and compact massive clumps or stellar clusters. I want note, however, that such an interpretation is unlikely because of two reasons: (i) a strong gravitational perturbation would have affected the gas and stellar nuclei as well as the MBHs, and (ii) as described in Section 3.2.1, the gaseous discs in my simulations were initially stable against fragmentation, and the gas distribution was expected to remain smooth during the entire evolution in run Plain, in which no cooling prescription is implemented. A search for gas and stellar clumps in the snapshots of run Plain confirmed this expectation.

The peculiar and unexpected dynamical evolution of the MBHs in run Plain could be a numerical artefact, due to the rapid variation of the spatial resolution around the two MBHs. Figure 4.1 shows the number of cells at the maximum refinement level within 5 pc from each MBH. The sudden drop of resolution is caused by a density drop during the first stages of the simulation. Such a gas readjustment was expected, since the initial conditions were stable in isolation, and the two circum-nuclear discs were initially set at a finite separation. I stress that, although this initial gas evolution is driven by the procedure used to generate the initial conditions, similar sudden resolution changes are expected also due to the evolution of the gas subject to additional physics, such as SNa explosions, as discussed below.

To check if the unexpected behaviour of the MBHs was a pure numerical effect I ran the same simulation forcing the code to keep a high resolution close to the moving MBHs, through the new refinement implementation. The MBH orbital evolution resulting from this check (run Plain+) is shown in figure 4.2. Run Plain+ shows a dynamical evolution closer to that obtained in the SPH run, that by construction is not affected by any significant fluctuation of the gravitational spatial resolution. Figure 4.3 shows that with the new

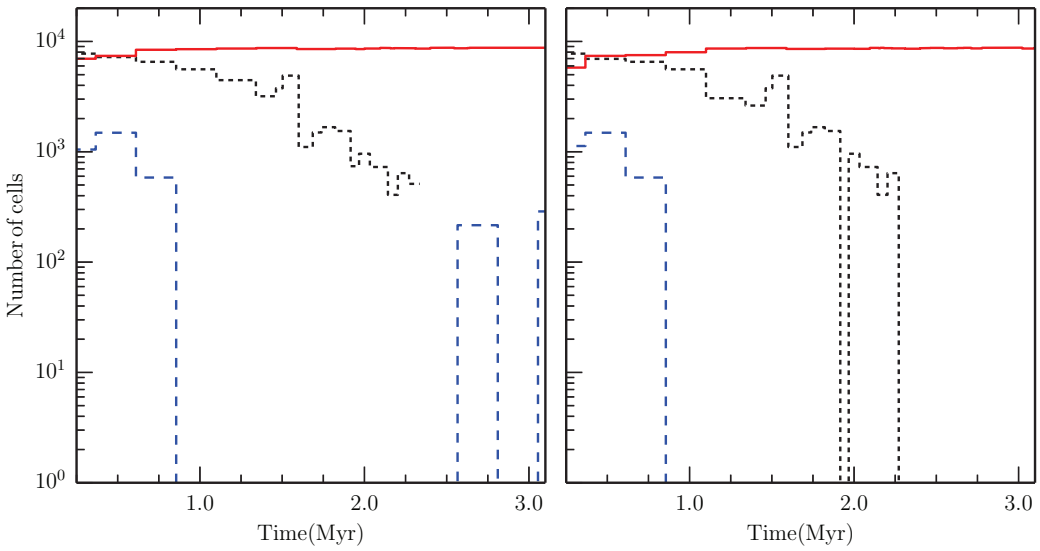

Figure 4.1: *Number of cells for the highest refinement level around each of the two MBHs versus time for simulations Plain (blue dashed line), Noblast (red solid line) and Blast (black dotted line).*

refinement implementation the MBHs do not decouple from the gas structure they are hosted in. I further stress that an enhanced resolution close to the MBHs would facilitate the formation of gas clumps as well as maximise the effect of their gravitational interaction (if clumps would form) with the MBHs. The absence of abrupt kicks in the MBH dynamics in run Plain+ proves that the MBH noisy motion observe in run Plain is numerical and it is caused by poor/rapidly changing resolution in the region surrounding the MBHs. Still some differences in the orbital evolution of the MBHs in run Plain+ and SPH are observable. The initial difference in the vertical motion is probably caused by the resolution increase occurring in the very early stages of the simulation, when the initial conditions (with a maximum resolution of ~ 1.5 pc) are further refined to reach the desired resolution of ~ 0.39 pc. Furthermore, the MBH orbital decay after the first 3 Myr is faster in the the Plain+ run with respect to the SPH run. I checked that this is due to the different magnitude of the gas inflow toward the geometrical centre of the system. Such inflows are caused by the angular momentum removal associated with the shocks developing at the contact surface between the two merging CNDs. The two numerical implementations (SPH and AMR) differ significantly in their treatment of the shocks, resulting in a different MBH dynamics.

To study the effect of the refinement prescriptions onto the MBH dynamics in less idealised simulations, I performed three runs (Noblast, Blast and Blast+) allowing the gas to radiatively cool and form stars. As shown in the following, the orbital evolution strongly depends on the different implementation for the SNa feedback. Figure 4.4 shows the MBH orbital evolution in run Noblast. The MBH dynamics does not show anything similar to the huge kicks that decouple the MBH dynamics from the gas distribution observed in run Plain. On the contrary, figure 4.5 demonstrates that the MBHs are still well within the gas and stellar overdensities close to the centres of the dramatically perturbed nuclear discs.

However, smaller swerves mainly limited to the disc plane are still observable in the MBH orbits (figure 4.4). Figure 4.1 demonstrates that, in run Noblast, the wiggles in the orbits are not related to a decrease of the spatial resolution. Indeed the resolution around each MBH remains almost constant during the entire run, with a high number of cells populating the maximum refinement level. Such a high resolution is ensured by the formation of high density condensation of cooling gas around the MBHs.

The peculiarities in the MBH orbits in run Noblast are due to close interactions with massive clumps, forming in the disc when the gas is allowed to cool. Indeed a large number of massive clumps form during the first stages of the merger, especially along the gas shock surface between the two gaseous discs, as observable in the left panel of figure 4.5. These clumps can lead to very energetic kicks to MBHs, unless they are destroyed before the interaction by SNe. This is not the case for run Noblast, in which the large thermal energy injected by SNe in the gas immediately starts to cool, thus leaving clumps nearly unperturbed. As discussed above, this boosts the probability of having a strong cloud/MBH interaction, and results in a high resolution close to the MBHs (preventing spurious numerical wandering of the MBHs). Figure 4.6 highlights a strong interaction between each MBH and a massive cloud in run Noblast, taking place at $t = 5.8$ Myr. Figure 4.7 shows a later stage ($t = 9$ Myr) of the evolution, when the two MBHs evolve in a smoother environment. In both cases the clouds have been identified extracting the cells with a density exceeding 8×10^5 H/cc and then grouping together the adjacent cells. The detailed analysis of the interactions between MBHs and clouds as well as a broader study of the effect of the gas dynamics onto the MBH pairing is deferred to section §4.4.

If instead the gas is unable to rapidly get rid of the energy injected by SNe I expect a smaller incidence of MBH-cloud interactions, but at the same time the SNe can strongly affect the densest and intensely star-forming regions close to the MBHs. A SNa driven gas depletion may result in a decreasing force resolution when the new refinement discussed here is not implemented. Figure 4.8 shows a comparison between the MBH dynamical evolution observed in runs Blast and Blast+. A peculiar wandering of the two MBHs in the three dimensions is observed in run Blast, similarly to what happens in run Plain. I stopped the run after 2 Myr only, when the MBH motion had already been affected by the numerical effect and MBHs had been scattered very far from the disc plane.

Again, the peculiar motion of MBHs in run Blast could either be a numerical artefact or have a physical origin. I note that in both runs Blast and Blast+ the clumps are disrupted on short timescales by SNe. Hence, gas overdensities are not expected to perturb significantly the dynamical evolution of MBHs. Furthermore, the feedback is energetic enough to deplete the gas from the nuclear regions of both discs, leaving the MBHs in an under-dense region (see figure 4.9). The time evolution of the number of cells at high resolution levels in the MBH vicinities is shown in Figure 4.1. This confirms that the energy injection from SNe drives a significant resolution drop during the first 3 Myr, as also observed in run Plain. In run Blast, however, the loss of resolution does not directly depends on our realisation of the initial conditions, but it is a consequence of the physical evolution of the system.

The different dynamical evolution observed in run Blast+ (lower panels in figure 4.8) finally proves that the jerks in the MBHs paths are numerical artefacts. In facts, in this last case, the MBHs follows a very smooth evolution over multiple orbital timescales, due to the little effect of the transient gas overdensities onto the MBHs. The comparison between the results of run Blast and run Blast+ proves the effectiveness of refinement implementation discussed here in modelling massive particle dynamics in rapidly evolving backgrounds.

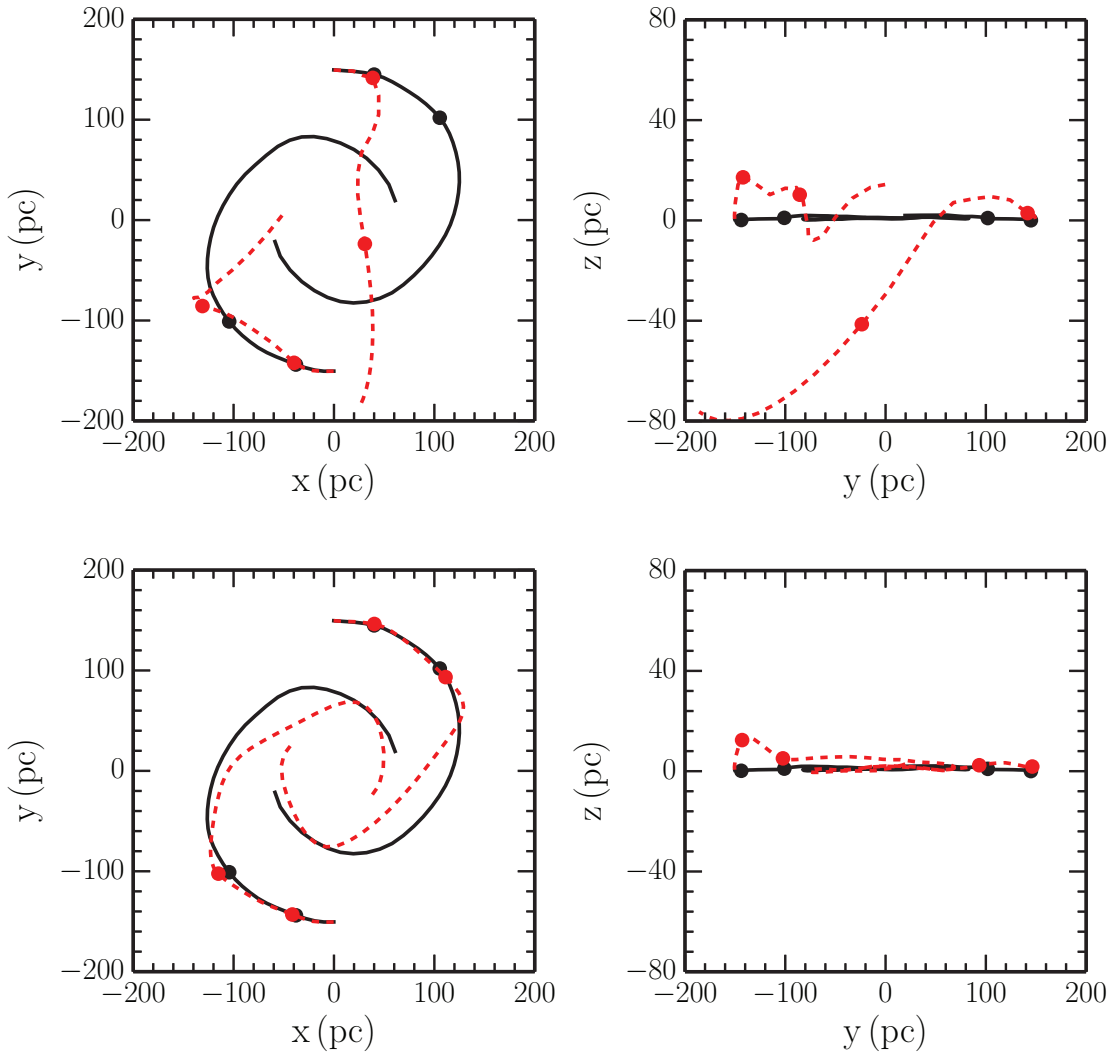

Figure 4.2: *Orbits for the two MBHs from* ramses *runs Plain and Plain+ , compared with the SPH run. The panels on top show the orbits projected in the face-on (on the left) and edge-on (on the right) views for run Plain, plotted as red dashed lines and the SPH run, plotted as black solid lines. The panels at the bottom are the same plots obtained from run Plain+ and the SPH run. The points mark the MBH positions at* $t = 1$ *and 3 Myr for the runs considered.*

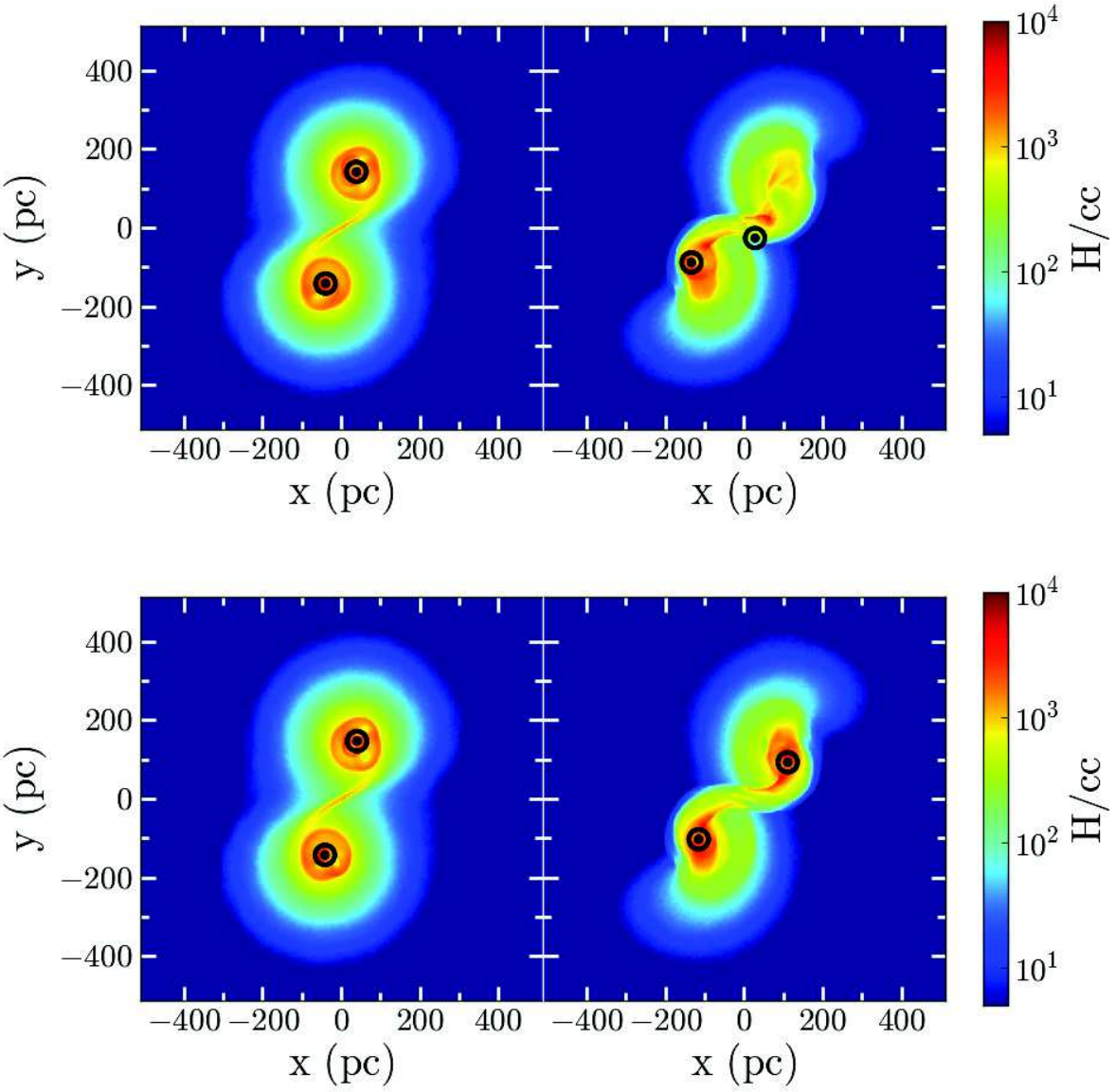

Figure 4.3: *Upper panels: Gas density map at $t = 1$ Myr (left panels) and $t = 3$ Myr (right panels) for run Plain. The MBH positions are identified by the black bullseye symbols. Lower panel: same as upper panel for run Plain+.*

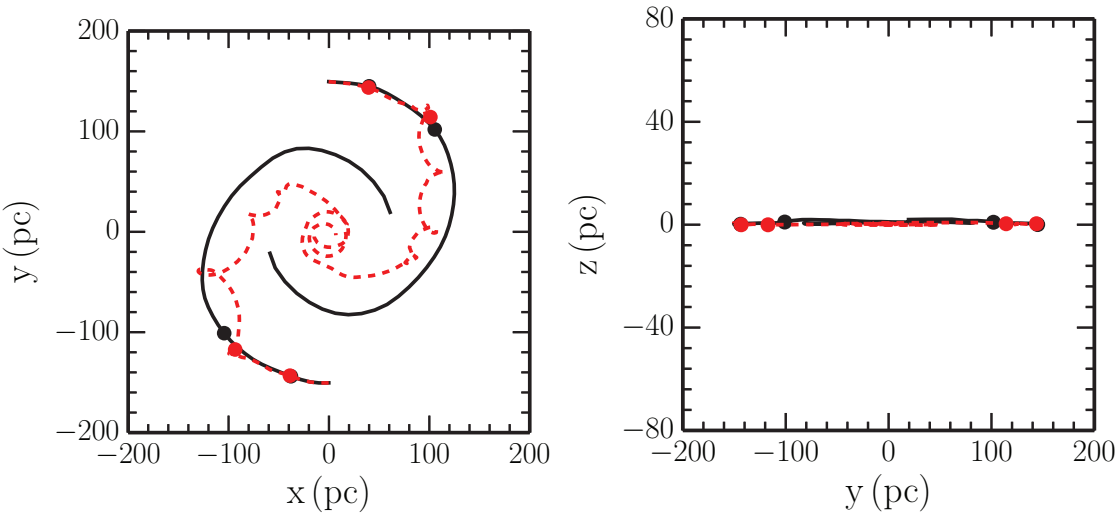

Figure 4.4: *Same as figure 4.2 for the Noblast/SPH runs comparison.*

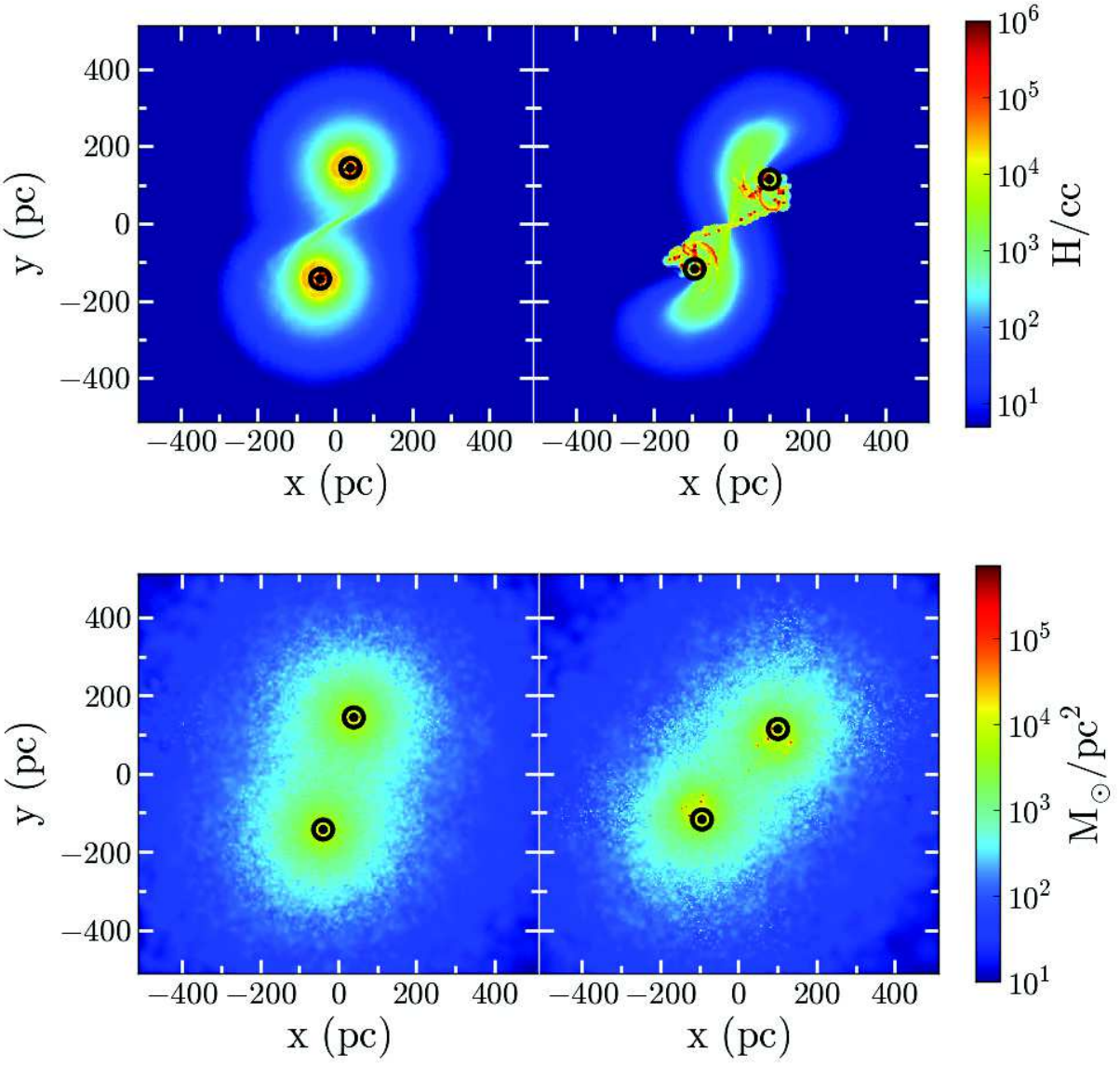

Figure 4.5: *Upper (lower) panel: gas (stellar) density map at $t = 1$ and 3 Myr (left and right panel respectively) in run Noblast. All the notation is the same as in figure 4.3*

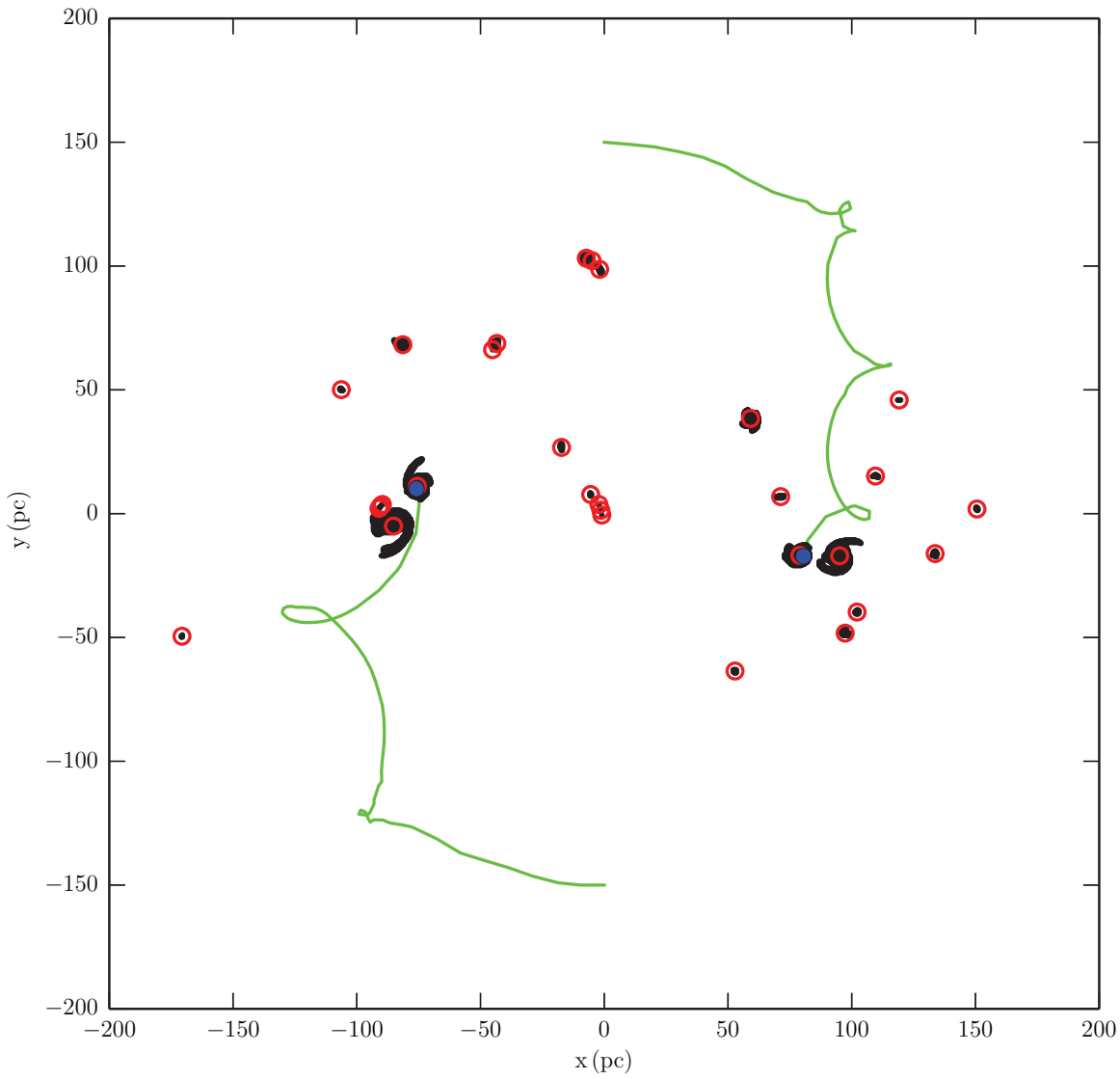

Figure 4.6: Strong interactions between the MBHs and massive gas clouds in run Noblast at time 5.8 Myr. The MBHs orbital path and current positions are marked with green lines and blue dots. The black regions highlight the cells forming the clouds, whose centre of mass is marked by the red empty circles (only for clouds formed by at least 10 cells).

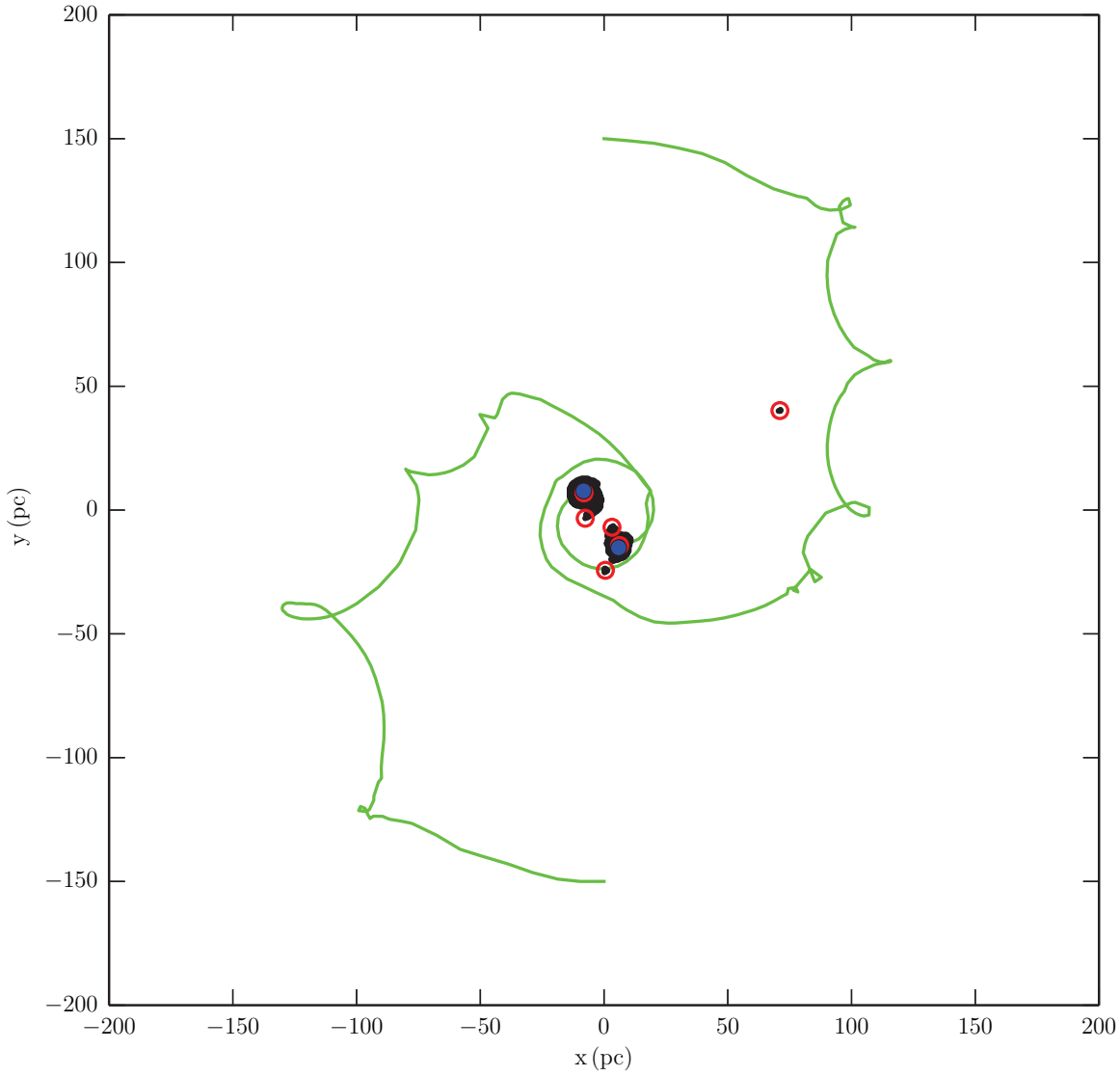

Figure 4.7: *Same as figure 4.6, but at time $t = 9$ Myr. The figure shows a final stage of the orbital evolution, when the MBHs are surrounded by gas overdensities and no significant MBH/cloud interactions are taking place.*

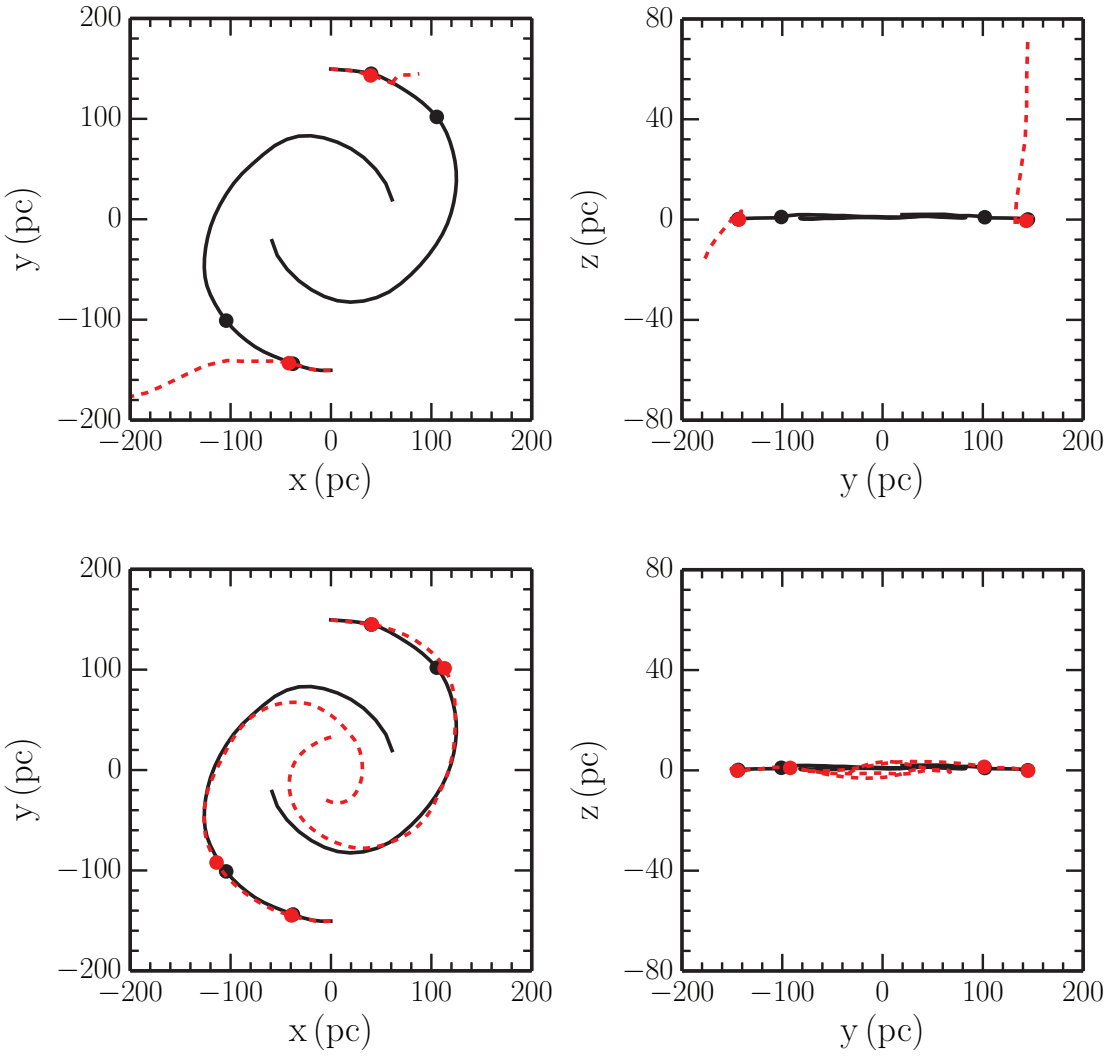

Figure 4.8: *Same as figure 4.2 for the comparison between run SPH and runs Blast and Blast+.*

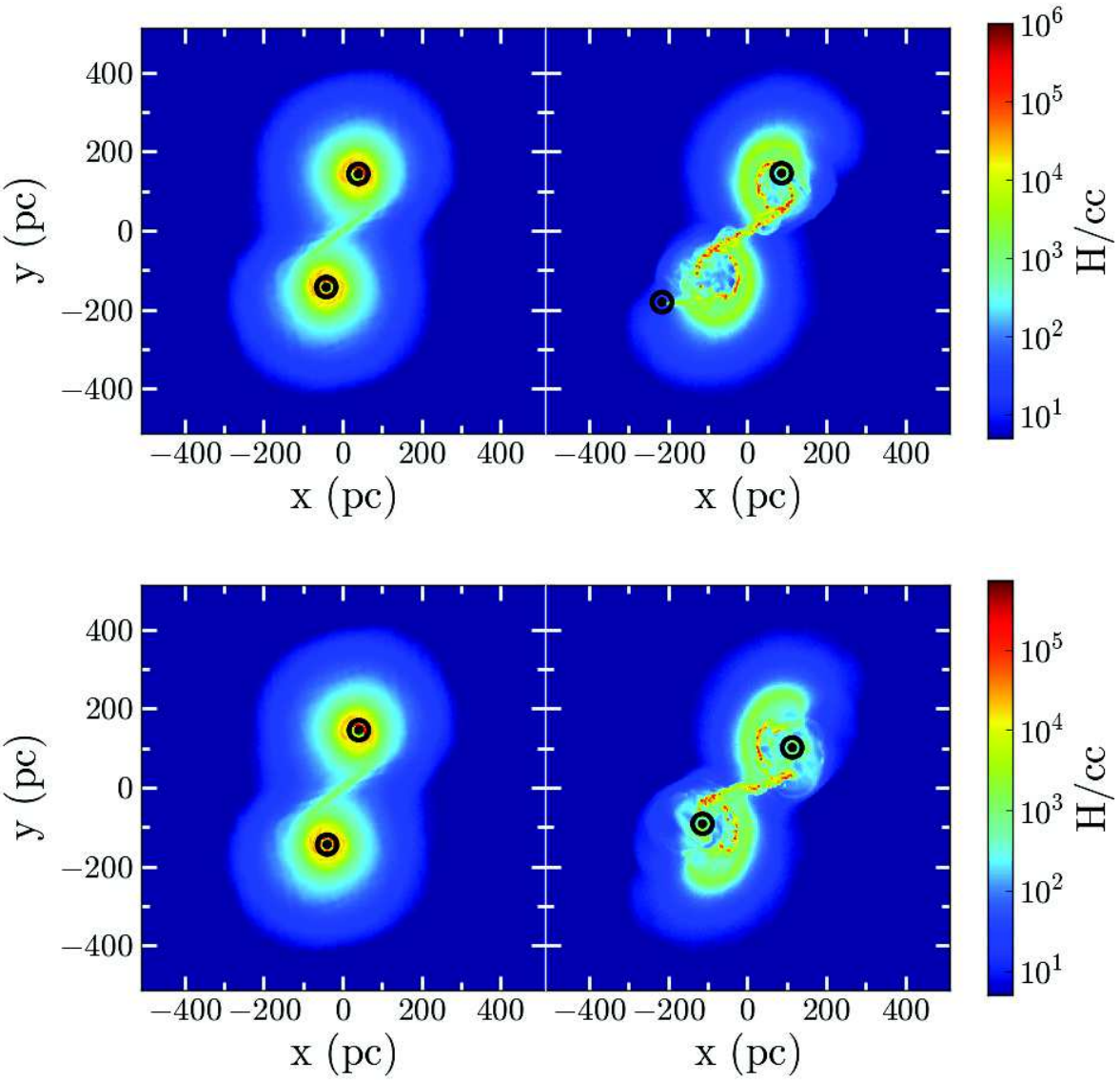

Figure 4.9: *Same as figure 4.3 for runs Blast and Blast+.*

4.4 Gas and MBH evolution: effects of different sub-grid models

In order to survey how the implementation of sub-grid physics could affect the evolution of the system I performed a suite of six simulations, where I changed the prescriptions regarding gas cooling, SF and SNa feedback.

To achieve the best possible treatment of MBH dynamics, I adopted the additional refinement criterion described in section §4.3. In these runs I assumed the two different fiducial values for the SF density threshold already described, i.e., $n_{\rm H} = 2 \times 10^5$ and $n_{\rm H} = 2 \times 10^6$ cm^{-3}. The resulting average mass of stellar particles was ~ 300 M$_\odot$. Such value is significantly more massive than, e.g., what employed in Amaro-Seoane, Brem & Cuadra (2013), who however simulated a lighter and more compact system. I checked that the prescription used resulted in a gas-to-stellar mass conversion rate not lower than the local Kennicutt-Schmidt law.

I ran two further simulations at the highest density threshold for star formation (termed "ThFBh_prompt" and "BWFBh_prompt") assuming no time-lag between star formation and SNa explosion. The aim of these runs was to test the effects on the global (gas and BHs) dynamics of a maximally fast SNa feedback, comparing the results to the standard $\Delta t_{\rm S} = 10$ Myr case. While simulations with standard delay were meant to model star formation as triggered by the merger of the two circum-nuclear discs, the 0-lag case might represent a situation where sustained star formation was already in progress at the time of the merger. Finally, in order to avoid inaccurate integration of the orbits, all runs have been stopped when the MBH separation was approximatively $3 - 4$ times the cell length. Table 4.2 summarises the six simulations with the parameter used.

4.4.1 Black hole dynamics

I start by describing the MBH dynamics for the two simulations characterised by standard thermal SNa feedback and a typical time for SNa explosions of 10 Myr (runs "ThFBl" and "ThFBh"). These two runs are meant to represent a case where star formation is indeed triggered by the merger event, while gas thermodynamics is governed by standard thermal processes. The two different

Run	$n_{\rm H}$ (cm^{-3})	$\Delta t_{\rm SN}$ (Myr)	Feedback
ThFBl	2×10^5	10.0	thermal
ThFBh	2×10^6	10.0	thermal
BWFBl	2×10^5	10.0	blast wave
BWFBh	2×10^6	10.0	blast wave
ThFBh_prompt	2×10^6	0.0	thermal
BWFBh_prompt	2×10^6	0.0	blast wave

Table 4.2: *The complete suite of runs. The second column shows the density threshold for SF, the third column the lifetime of massive stars, the fourth column the type of feedback employed.*

density thresholds for SF are used to asses the effects that the efficiency of gas conversion into stars has on the MBH dynamical evolution.

In Fig. 4.10 I show the MBH projected orbits (left panel), and MBH separation versus time (right panel). The two MBHs exhibit a peculiar orbital motion, which can be explained when considering the gravitational interactions between the MBHs and massive gas/star clumps forming in the merging discs. Such interactions typically accelerate the orbital decay of the MBHs, and a gravitationally bound MBHB forms after ~ 10 Myr (the binary formation time is indicated as a blue dot in the right panel). In the case of ThFBh run (dashed red lines), the MBH orbits appear more perturbed, and the orbital decay is somewhat faster.

Fragmentation of gas, occurring just after the simulation starts, tends to form massive gas clumps, especially in the high density regions surrounding the two MBHs. In high density clumps star formation is very effective, and overall, a large fraction of the initial disc gas is converted into stellar mass within 10 Myr. This is apparent from Fig. 4.11 (left panel), where the stellar mass and the residual gas mass are shown as a function of time. The right panel of Fig. 4.11, instead, shows the star formation rate versus time. A fast increase of star formation occurs initially since gas shocked during the disc collision fragments into small clumps which immediately convert into stellar particles. After ~ 2 Myr, only low density gas survives. Hence, star formation is no longer efficient and almost steadily decreases in time. In Fig. 4.12 we

plot the mass-weighted gas density map at time $t = 2.1$ Myr, defined as the time of the peak in star formation rate (see Fig. 4.11, right panel).

In order to quantify the impact of gas clumps on MBH dynamics, I estimated the total mass in gas/star clumps, along with the clump mass distribution. I identify as "clumps" those gravitationally bound regions that feature a single peak in the 3-D density field. Fig. 4.13 shows the total mass in clumps as a function of time for run "ThFBl". Two distinct phases can be observed, with a peak in the total mass of clumps occurring after a time $t_{\rm peak} \sim 2.5$ Myr. The initial fast growth of the gas locked in clumps is the result of the collision between the two unperturbed gaseous discs. Indeed, gas fragmentation is promoted along the shock surface (resulting also in the peak of star formation rate, see Fig. 4.11, right panel).

Fig. 4.14 shows, for the same "ThFBl" run, the mass distribution of clumps at four different selected times marked as red dots in Fig. 4.13. I selected two times corresponding to a relatively low total clump mass ($\sim 1.8 \times 10^7\ M_\odot$, left panel) and two times corresponding to a larger mass value ($\sim 5 \times 10^7\ M_\odot$, right panel), respectively one before and one after $t_{\rm peak}$. The mass distribution lies in the range $10^{5-7}\,{\rm M}_\odot$, with few clumps as massive as the MBHs. These very massive clumps typically form after $t_{\rm peak}$, most probably due to gas accretion from low density regions and to mergers between less massive clumps, and eventually will merge with the gas overdensity surrounding each MBHs. When one of these more massive clumps manages to approach a MBH at close range, then a transient MBH-clump binary system forms, strong gravitational perturbations develop, and the MBH orbit greatly deviates from its original path. This is the reason behind the "wiggling orbits" seen in Fig. 4.10, right panel. The typical BH-clump distance when the transient binary system forms is $\sim 10 - 20$ pc, which is always resolved with a number of cells $\gtrsim 10$, thanks to our new refinement prescription, allowing us to accurately resolve the BH-clump close interaction.

In the case of run "ThFBh", because of the relatively higher density threshold for SF, a slightly larger number of more massive clumps forms, resulting in the more disturbed orbits (and faster decay) seen in Fig. 4.10.

I then compared the above analysis regarding the MBH dynamics with runs employing the aforementioned blast wave like feedback from SNe (BWFB-like runs). As discussed before, this feedback implementation aims at describing non-thermal processes in the aftermath of SNa explosions. I found that the

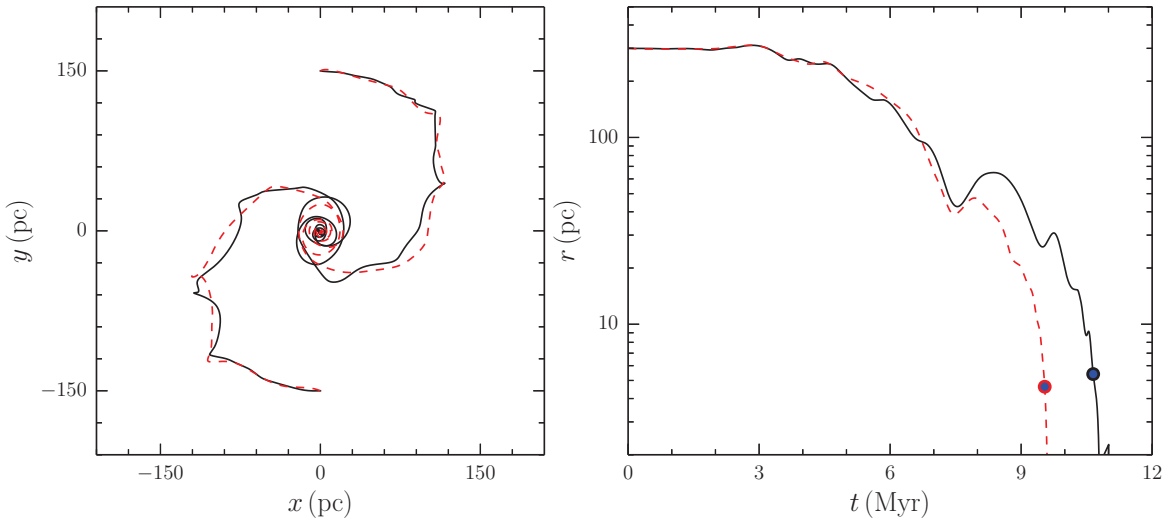

Figure 4.10: *MBH dynamical evolution for runs "ThFBl" (solid black lines) and "ThFBh" (dashed red lines). Left panel: projected orbital evolution. Right panel: MBH separation versus time. The blue dots correspond to the time of binary formation.*

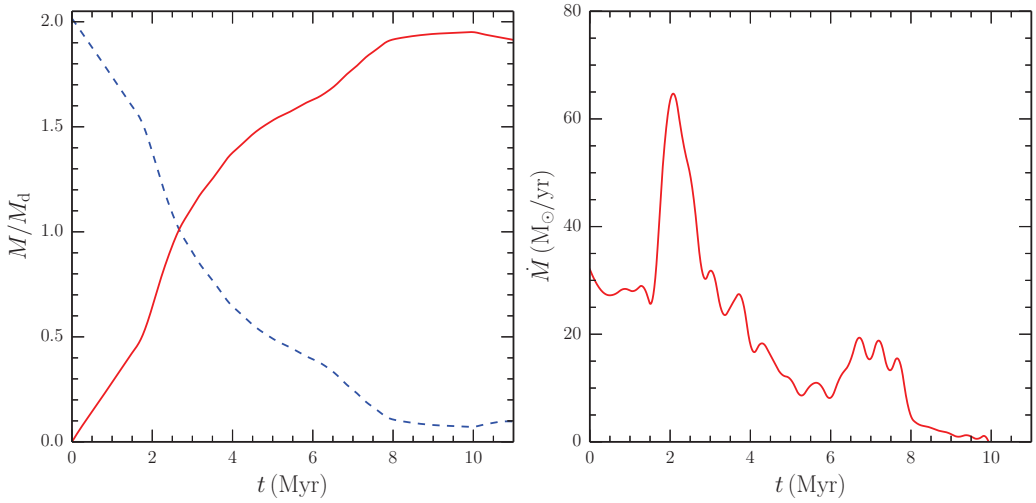

Figure 4.11: *Star formation in run "ThFBl". Left panel: total stellar mass (solid red line) in units of the initial disc mass M_d, and the residual gas mass in units of M_d (dashed blue line) as a function of time. Right panel: star formation rate versus time.*

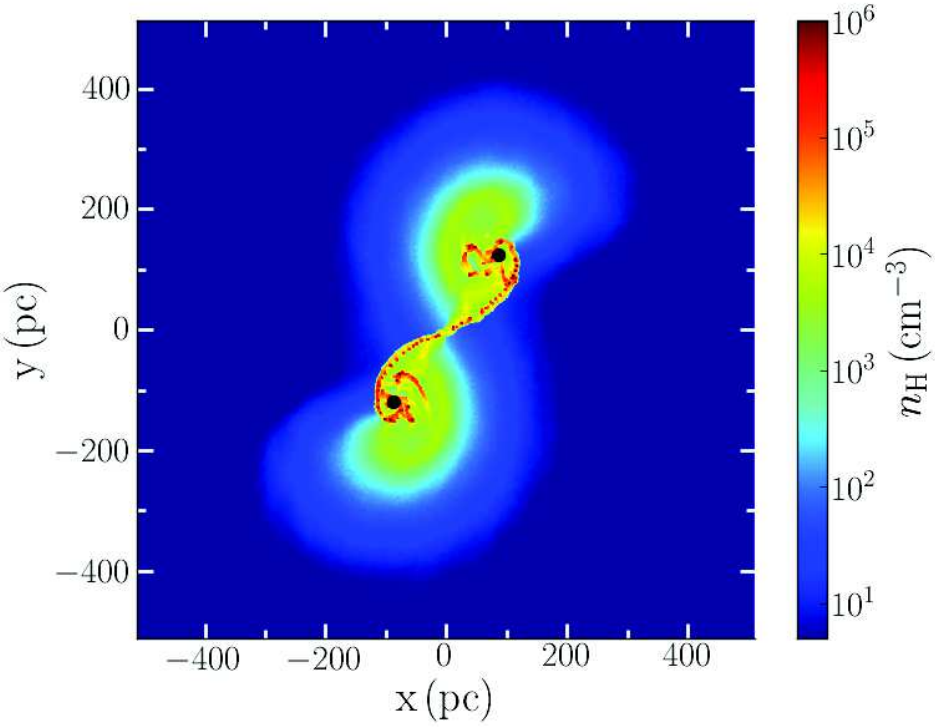

Figure 4.12: *Face-on gas density map for run "ThFBl" at time $t = 2.1$ Myr (i.e., when the star formation rate is maximum, see Fig. 4.11). The gas shocked after the first disc collision fragments into a large number of small clumps which very rapidly convert gas into new stellar mass. The black dots correspond to the positions of the two MBHs.*

dynamical evolution of the MBHs was largely independent upon the details of the SNa feedback employed, making MBH dynamics results fairly robust against the different implementations of sub-grid physics.

4.4.2 Gas dynamics

I discuss here the dynamics of the gas during the merger event. I will focus on the case with the low density threshold for SF (run "ThFBl"), keeping in mind that the higher density case produces a qualitatively and quantitatively similar outcome.

Fig. 4.15 shows the gas distribution around the MBHB after $t = 11$ Myr. On large scale (top panel), the relic disc resulting from the collision of the progenitor discs is almost totally disrupted because of SNa feedback. This residual structure is counter-rotating relative to the MBHB orbit. On scales of order of few pc (bottom panel), the gas which has not been converted into stellar particles settles in a circum-binary disc, with a total mass of few $10^5\,\mathrm{M}_\odot$. The small disc co-rotates with the MBHB thanks to the dragging of gas by the MBHs during their inspiral toward the centre. Note that this implies that the angular momentum of the residual gas changed sign during the evolution of the system.

I reported in Fig. 4.16 the evolution of the modulus of MBH orbital angular momentum and compared it to the modulus of the total angular momentum of the gas which is the closest to the MBHs in the simulation, defined as the gas within a sphere of radius equal to 0.5 times the MBH separation. I observed that at the beginning of the simulation the angular momentum of the gas was larger than that of the MBHs, and I remind that the gas is counter-rotating. After $\gtrsim 4$ Myr, the angular momentum associated to the MBH orbit exceeds that of the gas and in principle there are the conditions for a change in the sign of the gas angular momentum, being dragged by the MBHs. The gas angular momentum actually changes sign after ~ 9 Myr, when the MBH separation is ~ 45 pc. At this evolutionary stage, a large fraction ($\gtrsim 90\%$) of the initial gas mass is already converted in stellar particles. After $\simeq 10$ Myr, when SNe start to explode, the released energy is radiated away by the small amount of residual gas, which is however unable to form further stellar mass at a comparable rate. In other words, star formation is not halted by SNa feedback, rather by gas consumption.

Concerning the impact of blast wave feedback (BWFB-type runs), as expected it does not alter the gas dynamics for a time $\sim \Delta t_{\rm SN}$ (at that point the two MBHs have already reached the centre of the system). After that time, the almost simultaneous SNa events release a fairly large amount of energy which heats the gas up but is not radiated away. The net result is that the remaining gas is pushed at very large distances from the MBHB (up to ~ 500 pc) by the increased pressure. The MBHB lives then in a very low density environment, and no circum-binary disc is formed on any scale.

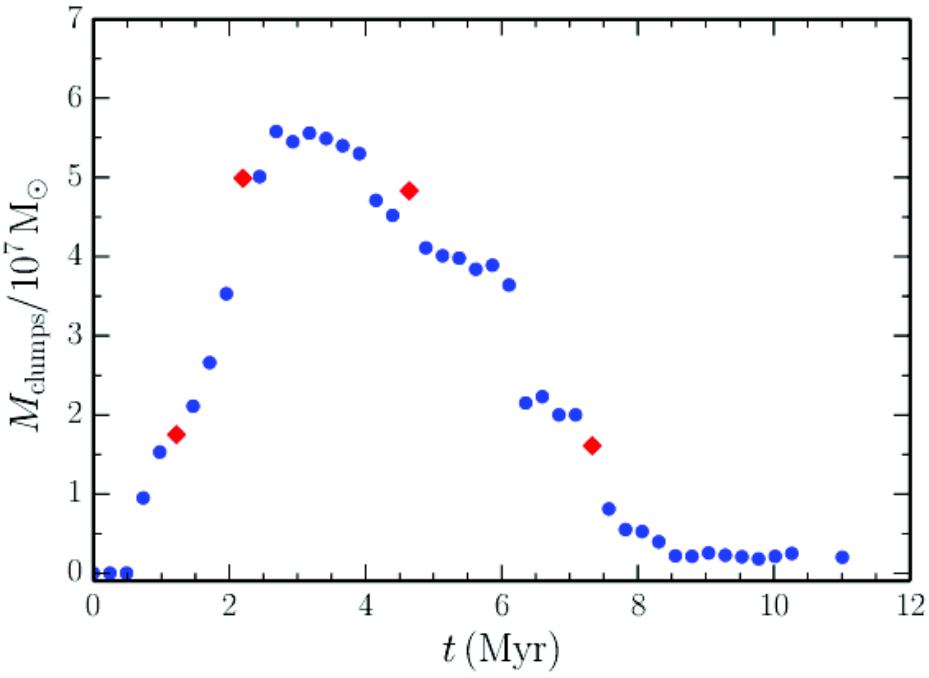

Figure 4.13: *Total mass in clumps for run "ThFBI". The red diamonds correspond to the times at which we computed the clump mass distribution shown in Fig. 4.14.*

4.4.3 Prompt SNa explosions

Both the MBH and gas dynamics are unaffected by feedback for the first 10 Myr as this is the assumed life time of massive stars (and hence for the onset of SNa feedback). To test how the results depend upon such choice, I considered the extreme case of $\Delta t_{\rm SN} = 0$ Myr, i.e, massive stars explode as soon as they form.

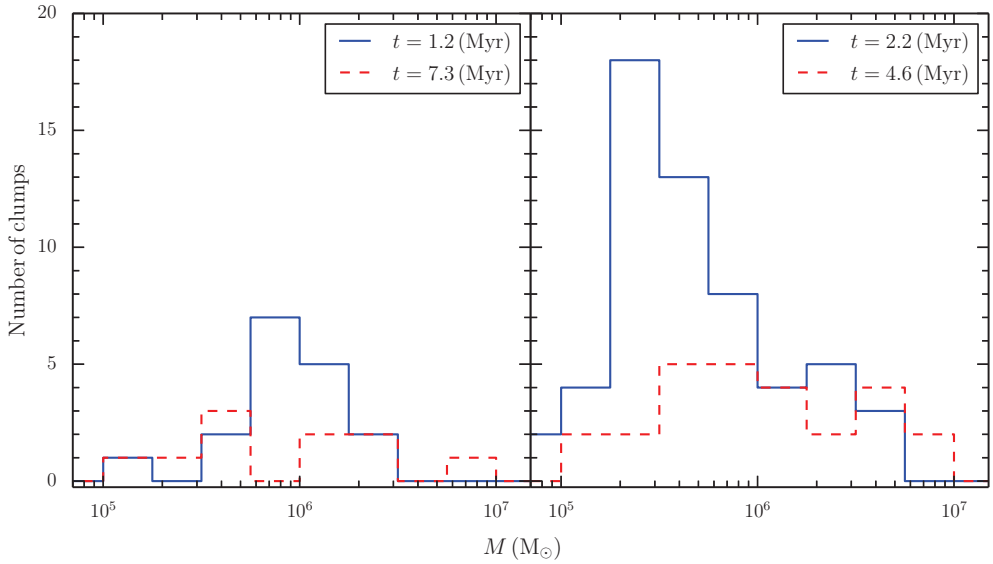

Figure 4.14: *Mass distribution of clumps in run "ThFBI". Left panel: mass distribution at two selected times when the number of clumps is relatively small. Right panel: same as left panel, but at two times when the number of clumps is larger. The four selected times are marked as red dots in Fig. 4.13.*

I found that, as long as the SNa feedback is governed by thermal processes, only small differences in the MBH dynamics exist compared to the standard delay case previously discussed. This similarity occurs because the SNa energy is mostly released in high density clumps, where gas cools down very rapidly, and the clumps can survive the explosion. As a consequence, star formation can proceed until almost all clump gas is consumed.

Large differences occur instead when, along the $\Delta t_{\rm SN} = 0$ assumption, I employed the blast wave recipe for SNa feedback. In Fig. 4.18 I compared the projected MBH orbits (left panel) and the MBH separation versus time (right panel) for runs BWFBh_prompt and ThFBh_prompt. In the case of blast wave like feedback, the orbital decay was slower, with a typical binary formation timescale of $\gtrsim$ 13 Myr. The difference is due to the early SNa explosions that, coupled with the blast wave like feedback, tend to disrupt the gas clumps and to deplete the gas reservoir progressively forming around the MBHs. As a consequence, the two MBHs evolved in a lower density, smoother environment, where low mass clumps were typically unable to induce strong orbital perturbations. The net result is a less disturbed orbital decay (Fig. 4.18, left panel).

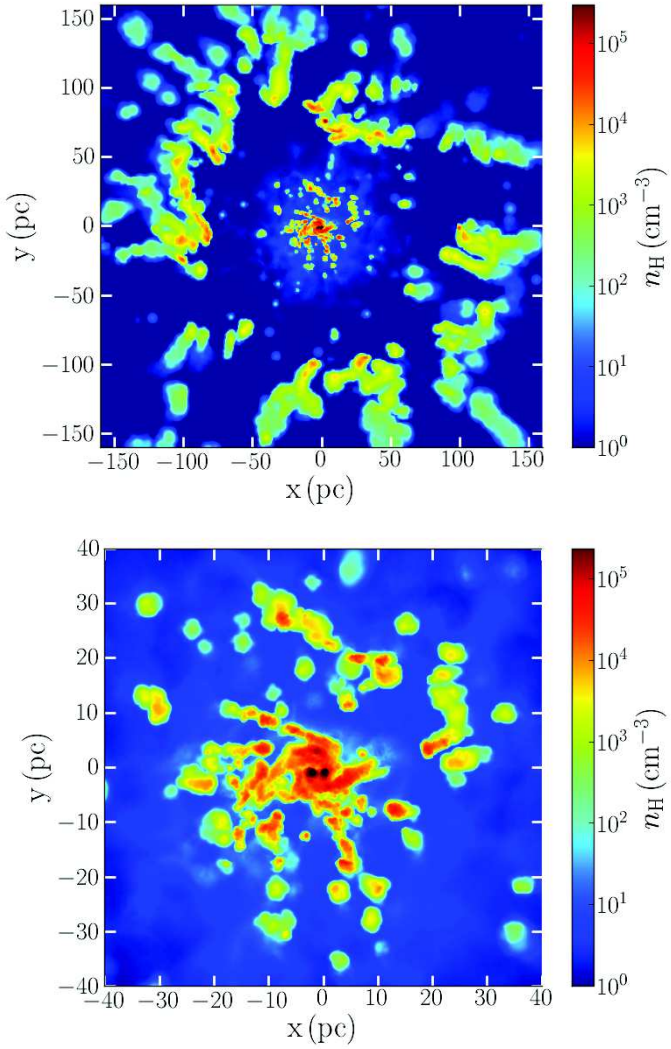

Figure 4.15: *Face-on gas density maps for run "ThFBl" around the MBHB at the end of the simulation ($t \sim 11$ Myr). Top panel: on large scales the disc is almost totally disrupted because of SNa explosions. Bottom panel: zoom in of the nuclear region where an inner co-rotating gas disc forms.*

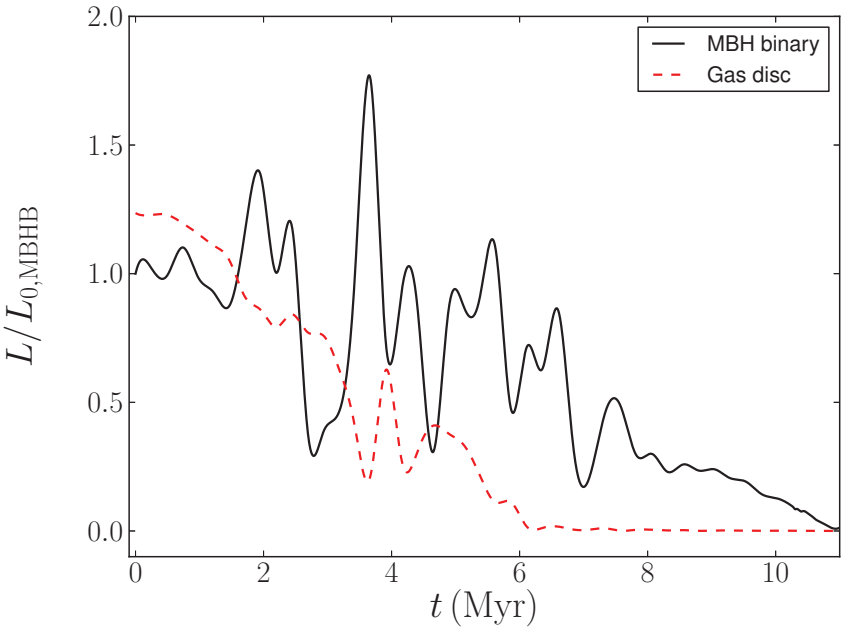

Figure 4.16: *Evolution of the moduli of the angular momenta for run "ThFBl". The MBHB total angular momentum (solid black line) is compared with the gas angular momentum (dashed red line), defined as the total angular momentum of the gas within a sphere centred on one MBH whose radius equals half the MBHB separation. Both curves have been normalised to the initial angular momentum of the MBHB.*

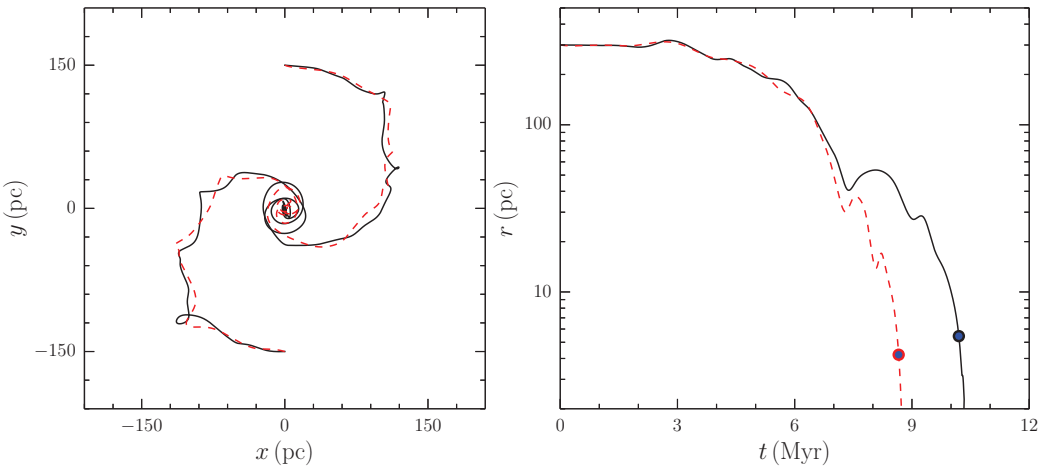

Figure 4.17: *Same as Fig. 4.10 but for runs "BWFBl" (solid black lines) and "BWFBh" (dashed red lines).*

I therefore conclude that in the case of prompt SNa explosions, contrary to the standard delay case, the dynamical evolution of the MBHs is strongly affected by the feedback mechanism employed. The SF density threshold instead does not result in relevant differences anyway.

While MBH dynamics is basically unaffected by the value of $\Delta t_{\rm SN}$ in the case of thermal SNa feedback, substantial differences occur in the dynamics of the gas component. Along with a small scale co-rotating circum-binary disc, I observed a further, much larger disc/ring-like structure on ~ 100 pc scale (see Fig 4.19). Indeed, feedback from SNe did not occur suddenly after 10 Myr but it was instead diluted in time, so that the (rapidly cooling) gas had time to readjust in a disc-like structure. Though several other possible explanations exist (e.g., secular evolution of the Galactic disc), it is tempting to associate such structure to the central molecular zone of the Milky Way (Jones et al., 2011). It is interesting to note that the larger scale disc keeps memory of the initial angular momentum, and it is then counter-rotating with respect to the small inner circum-binary disc which is, as discussed above, dragged by the MBHB.

The case of blast wave like feedback is still different. I did not observe a disc like structure, rather I found a massive triaxial gas distribution surrounding the MBHB with density of few 10^5 cm^{-3} (see Fig. 4.20). This difference is produced by the different nature of the SNa feedback, which is in this case able to heat the gas and provide a pressure support large enough to prevent gas contraction.

Because of the large fraction of gas available (due to the SNa feedback which reduces the net star formation by destroying gas clumps, as discussed above) the gas will continue to cool down, resulting in alternated phases of star formation (due to gas cooling and contraction) and re-heating (due to SNa feedback). I observed a large number of dense gas streams flowing from low density regions toward the centre where the MBHB resides. This large inflow will result in a burst of star formation in the nucleus and in a following phase of SNa explosions. The energy provided by SNe will then reheat the gas, stopping the contraction and eventually expand the entire gas structure into a less dense state. These alternated phases, if occurring for enough time, could convert a large fraction of gas into new stellar mass, which could eventually form a massive nuclear stellar cluster surrounding the MBHB.

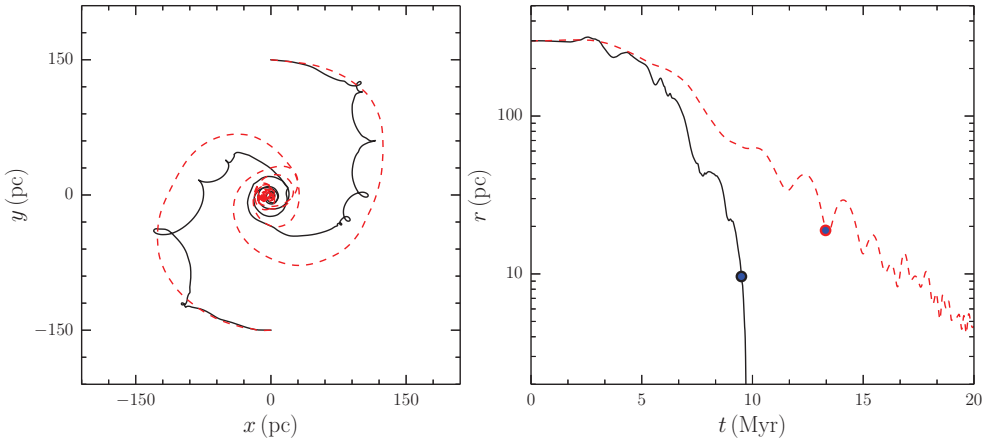

Figure 4.18: *Same as Fig. 4.10 but for runs "ThFBh_prompt" (solid black lines) and "BWFBh_prompt" (dashed red lines). The blue dots correspond to the time of binary formation.*

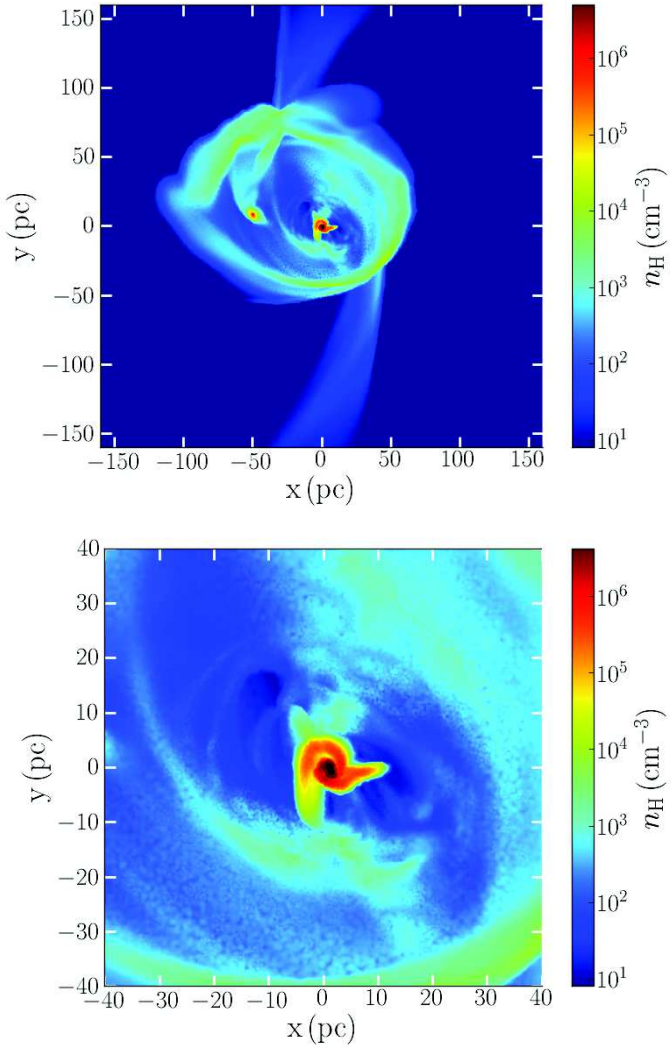

Figure 4.19: *Same as Fig. 4.15 but for run "ThFBh_prompt" at time* $t \sim 10$ *Myr. Top panel: gas settles in a disc/ring like structure which is counter-rotating relative to the MBHs. Bottom panel: zoom in of the region where an inner co-rotating gas disc forms around the MBHB visible on the east side of the left panel.*

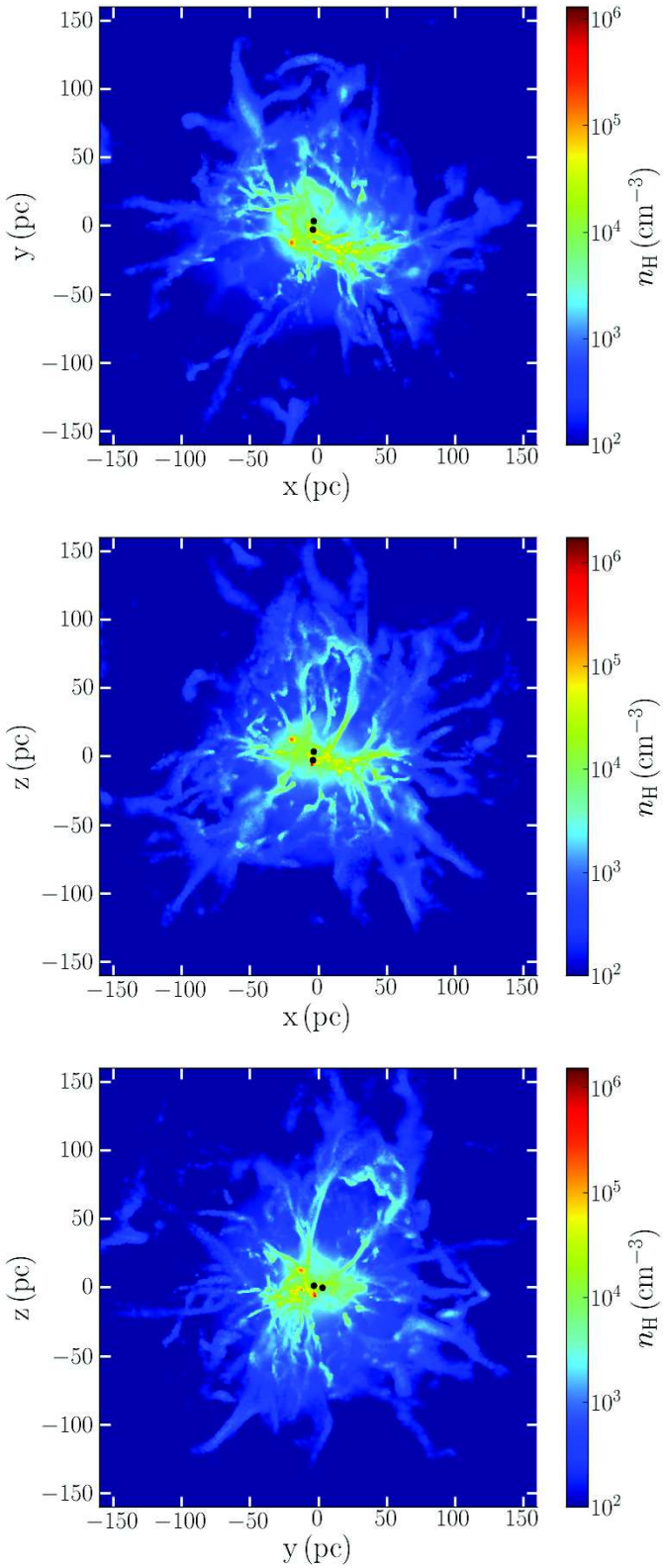

Figure 4.20: *Gas density maps for run "BWFBh_prompt" around the MBHB at the end of the simulation ($t \sim 20$ Myr). The gas settles in a triaxial structure with a denser central core. The core mass is $\sim 10^7\,{\rm M}_\odot$ within a radius ~ 25 pc. The upper panel shows the face-on view, while the central and the bottom ones show the edge-on views of the triaxial gas configuration.*

Chapter 5

Conclusions

In my thesis I investigated the formation and evolution of MBHs in galactic nuclei, focusing on two main aspects: the formation of MBH seeds via super-critical accretion on to SBHs and the formation and evolution of MBHBs during gas-rich galaxy mergers. Both aspects are still very debated and a clear consensus is far from being reached. In my studies I considered highly idealised setups, aimed at better resolving the processes acting on the nuclear scale of a galaxy without being influenced by large scale effects.

5.1 MBH seed formation

By means of high resolution numerical simulations I studied the accretion of stellar mass BHs in nuclear gaseous discs, to investigate whether phases of super-critical accretion on to SBHs could loose the tight constraints on the initial BH seed mass. I implemented a new BH thermal feedback prescription, that takes into account the possible occurrence of radiatively inefficient accretion bursts during which the BHs can actually increase their masses at a significantly super–Eddington pace. I have employed both AMR and Lagrangian mesh-free simulations, achieving comparable results, which strengthen greatly my conclusions and help to disentangle the numerical limits of each code technique and the physical processes involved.

I supposed that a population of SBHs was already in place in a well formed gaseous disc. The latter, at the beginning of the simulation, started cooling and eventually formed stars. Furthermore, these simulations have been evolved in complete isolation, i.e., no gas flew into the nuclear disc from larger scales (e.g., from outer regions of the host galaxy, from cosmological filaments, or through galaxy mergers). As a consequence, every accretion episode halted when star formation and SNa–driven gas consumption evacuated the central disc regions.

The prescriptions adopted in the simulations regarding SF and SNa feedback could, in principle, have affected the growth of BHs. Indeed, a too inefficient SF would have prevented gas consumption, and a too long timescale for SNa explosions would have prevented the clumps from being disrupted, thus leaving large amount of gas available for BH accretion. Nonetheless, my prescriptions are conservative for what concerns BH accretion. First, the SF rate in a sphere of radius $\simeq 1$ pc (corresponding to the average clump radius) around $\mathrm{BH_{top}}$ was $\simeq 0.1\,\mathrm{M_\odot/yr}$, much larger than the average BH accretion rate ($\simeq 10^{-3}\,\mathrm{M_\odot/yr}$). Hence, I can be confident that in my simulations the gas was mostly consumed by SF rather than by accretion. Second, the assumed timescale for SNa explosions (1 Myr) was shorter than the typical lifetime of low metallicity stars in the mass range $8 - 40\,\mathrm{M_\odot}$ ($\gtrsim 4$ Myr; Hurley, Pols & Tout (2000)). The resulting SNa feedback was then highly efficient. In this context, I found that SNe produced a high velocity wind ($v_{\rm ej} \lesssim 3000\,\mathrm{km/s}$), which could expand up to 5 kpc above the disc plane. In principle such gas could form a galactic fountain falling back on to the disc, allowing for a new phase of super-critical accretion. Anyway, the modelling of the large scale galactic potential (essential to asses the fate of the SNa driven wind) was beyond the scope of this study.

Regardless the spatial/mass resolution and the kind of hydrodynamical code used, a coherent picture emerged. If BHs have to grow by 2-3 order of magnitudes in mass, radiative inefficient accretion is a necessary condition, but not a sufficient one. BHs must find themselves embedded in gas structures that need to be: (i) massive enough to provide the gas reservoir, and (ii) dense enough to survive feedback. This may occur when the cooling gas fragments in clumps, and some of the BHs bind to them. Such process allows some of

the BHs to reach masses as high as 10^3 - $10^4\,\mathrm{M}_\odot$ on Myr timescales, making them viable candidates as seeds of the supermassive variety of BHs powering high redshift quasars.

Mass accretion onto the BHs depends upon the number, mass and density of the clumps forming in the disc. I showed that these parameters are affected by the numerical resolution achieved in the different runs and, as discussed in Section 3.4, different resolutions result in different BH accretion histories. I was unable to describe gas dynamics down to the accretion disc scales, even at the highest spatial resolution reached, and this limited my ability to achieve firm estimates of accretion rate and mass growth of the BHs. Yet, the dynamics of gas leading to the formation of dense clouds I observed in all the runs is strongly independent of sub–grid recipes. The gas within the accretion radius of BHs was far from being rotationally supported. Since the relative gas-BH velocity became negligible after the capture process, the gas in fact experienced almost radial inflow toward the BHs. My estimate of the accretion rate is of the same order of the Bondi accretion rate given the temperature and density of the medium surrounding the BHs.

Therefore, despite the accretion histories are not accurate enough from a quantitative point of view, I can be confident about the reliability of the BH-clumps-capture process observed. My study should be considered as a *proof of concept*, robust enough to understand under which conditions and through which processes a cluster of stellar mass BHs can actually experience episodes of super–Eddington growth, and what are the effects on the environment.

The key point of the study is that a radiatively inefficient accretion, together with the aforementioned BH-clumps-capture process, can result in mass growths 10-100 times larger than in the radiatively efficient case, making this mechanism a viable candidate to grow massive BH seeds from stellar mass BHs.

The process I studied can result in a prolonged super–Eddington accretion phase only as long as the masses of the clumps are comparable or larger than the masses of the accreting BHs. While the gravitational capture itself easily binds small clumps to comparatively massive BHs, the available gas reservoir is not sufficient for significant BH growth. Moreover, even feedback from radiative inefficient accretion severely affects such small clumps.

Other feedback processes, e.g., momentum-driven feedback, might be important, and could be explored in the future. If, however, such processes turn

out to be inefficient, this could naturally allow the galaxy to remain highly star–forming despite the fast growth of the MBH, perhaps explaining the new puzzling observation of a high-z star forming galaxy hosting an SMBH well overweight for its stellar mass (Trakhtenbrot et al., 2015).

In addition, as soon as a BH becomes significantly heavier than typical gas clumps, it starts migrating toward the centre of the disc via dynamical friction. This process will naturally bring the most massive BH (the one that by chance had the largest mass growth, i.e., $\mathrm{BH_{top}}$) to the centre of the host galaxy, where MBHs are commonly observed. At this point, however, further clumps forming in the disc no longer interact with the central BH.

In order for the large nuclear gas reservoir assumed in the initial conditions to be present in the galactic nucleus disc angular momentum needs to be removed well before gas turns into stars, so that inflowing material can be, at least partially, accreted by the central BH. This is of course the longstanding fuelling problem of MBH debated in the community (e.g., Hicks et al., 2013, and references therein), and its discussion is beyond the goal of my study.

I finally note that, whenever inflowing gas refills the circum–nuclear disc, the whole simulated process is rejuvenated: a new massive BH seed will be formed, sinking to the centre of the galaxy and eventually forming an intermediate massive black hole binary bound to coalesce owing to gravitational radiation losses. This kind of systems may be a perfect target for the planned eLISA observatory (Amaro-Seoane et al., 2013).

5.2 MBHB formation and evolution

By means of high resolution, adaptive mesh refinement hydrodynamical simulations, I explored the evolution of two massive gas discs hosting at their centre a MBH. The two discs had been set on an elliptic orbit and merged, to mimic the encounter between two very gas–rich disc galaxies. To maximise the strength of the interaction, the orbital angular momentum had been chosen to be antiparallel to the disc's angular momenta. Strong shocks that developed during the merger of the two discs became sites of intense star formation, and stellar feedback altered significantly the thermal and dynamical state of the gas which underwent a major transformation. Most of the gas was turned into new stellar particles through the formation of clumps of mass $\lesssim 10^6\,\mathrm{M_\odot}$. Only few clumps formed as massive as the MBHs, weighing $10^7\,\mathrm{M_\odot}$. In this kind

of environment, where both gaseous and stellar background evolve on short time-scales, a detailed study of the MBH pairing and the subsequent MBHB shrinking require great accuracy in the gravity force computation, to correctly model the orbital decay due to interactions with gas clumps and to the dynamical friction exerted by gas and stars. However, the gravity force accuracy in AMR codes strictly depends on the local resolution and thus on changes in the physical properties of gas and particles during the runs. Since the refinement prescriptions already implemented in RAMSES did not ensure a proper description of the orbit of massive particles, I implemented a new refinement criterion aimed at following in details the orbit of massive particle. In order to trust the simulation results a different refinement criterion, which forces a high and constant resolution near the MBHs, is required. The prescription I designed enforces the region around each massive particle to remain at the maximum resolution allowed. Such region follows the MBHs along their orbits, reducing the computational cost of the runs, and avoiding the spurious effect caused by the resolution changes. In the runs I explored different supernova feedback recipes: the thermal and blast wave feedback, assuming a lifetime of ~ 10 Myr for the massive stars. Furthermore, I considered a case in which prompt SNa explosion was coupled with both thermal and blast wave feedback. I found that the orbits of the two MBHs were perturbed due to their interaction with single clumps during the paring phase I, resulting in impulsive kicks that imprinted sudden changes in the direction and velocity of the orbit. Sinking times of $\sim 10 - 20$ Myr have been found, considering the set of parameters used. The pairing phase terminated with the formation of a Keplerian binary.

The MBH orbit observed was stochastic due to the presence of gas clumps. However, I did not see a sizeable delay or spreading in the sinking time due to gas clumpiness, contrary to what found in Fiacconi et al. (2013), where the level of stochasticity of the orbit was higher. I interpreted this difference as due to the geometry of the collision that mainly confined star formation along the oblique shock forming at the time of impact of the two discs, and to the fact that in the investigated case the mass distribution of the clumps evolved as gas was turned into stellar mass which spreaded due to dynamical relaxation. The simulated MBHs did not leave the orbital plane due to clump–induced kicks, contrary to what seen in Roškar et al. (2015), as my simulations were strictly co–planar. I expect that an inclined encounter would lead to a change in the

orbital plane also in our case, and this could be explored in future. During the pairing phase, the MBH dynamics is mostly affected by the presence of clumps and not by the recipe used to model the feedback processes.

I note that, on the contrary, the gas distribution around the MBHs is significantly affected by feedback. Thermal feedback left no large scale disc around the MBHB. Yet a residual co–rotating circum–binary disc of mass much smaller than the MBH mass formed around the two black holes which I expect will control the further spiral-in via migration like mechanisms.

Blast wave feedback is a way to model the expansion of SNa–driven bubbles. With the code it was then possible to mimic the ballistic phase of the shock triggered by the SNa explosion. As cooling was shut off in this phase, a multi–phase gas formed and the sweeping of the gas induced by the blast wave led to the almost complete evacuation of gas. The MBHB thus inhabited a region completely devoid of gas. Blast wave feedback in the prompt scenario led instead to a configuration in which the MBHB was surrounded by a gas cloud with little angular momentum and triaxial in shape.

The lesson to learn is that star formation in merging gaseous discs is a key process which affects the physical state of the gas in the surroundings of the MBHs. Under these circumstances it is difficult to predict the actual distribution of gas when the most active phase of the merger has subsided, as the outcome depends upon the modelling and on sub-grid physics, and firm conclusions should be taken with caution. Still, the presence of cool gas has deep implications for the evolution and observability of close MBHBs. First, the evolution of a binary on sub-pc scales toward the coalescence is strongly dependent on the gaseous and stellar distribution in its immediate surroundings (Colpi & Dotti, 2011, for a review). The timescale of the MBHs shrinking on sub–pc scales is of fundamental importance as it affects the expected rate of binaries possibly observable as gravitational wave sources. This is particularly true in mergers between gas rich galaxies, a fraction of which can host binaries detectable by future space based gravitational wave detectors such as eLISA (Amaro-Seoane et al., 2013). Second, the presence of gas is a necessary condition for the possible detection of the binary during the hardening phase (Dotti, Sesana & Decarli, 2012) as well as for pinpointing an electromagnetic counterpart of the MBHB coalescence (see, e.g., Schnittman, 2013; Bogdanović, 2015). The lack of a clear consensus on the processes shaping the environment of MBHBs, whose evolution actually depends on the physical modelling, and

the lack of observations available on the small scales I considered in this work witness the need of investigating a wider range of parameters.

5.2.1 Future prospects

Both the studies discussed in this work are worth of future investigations. Different lines can be followed for each of them. In the following paragraphs I briefly discuss some possible prospects.

MBH seed formation

In order to better constrain whether the conditions considered in this study could be plausible in an already evolved massive high redshift galaxy, I will try to study the formation of such a galaxy by means of state-of-the-art zoom-in cosmological simulations (Fiacconi et al. in preparation). I started this investigation together with a group in Zurich and we aim at studying the nuclear properties and gas inflows in a galaxy at $z \sim 6-10$ which by $z = 3$ will have a mass comparable to that of the observed clumpy star forming discs. This will allow us to place the model considered more properly in the context of galaxy formation and test its assumptions and outcomes. Preliminary analysis have shown that the mass enclosed within a hundred pc scale fluctuates between a few times $10^7 M_\odot$ and just above $10^8 M_\odot$. So, as mentioned in section 3.1, my initial conditions seem to be well motivated.

As a possible subsequent study, I would consider a small cosmological box around the massive galaxy just described, to study in better detail the nuclear inflow and gas accretion on to the formed seed MBH.

MBHB formation and evolution

In this study I described a single setup, where the two discs were equal masses, co-rotating and the interaction was coplanar. In a possible future study I would investigate in more detail the binary formation and shrinking for unequal mass systems and inclined initial configuration, with the goal of better understanding whether the orbital decay can be delayed when less idealised conditions are considered.

Appendix A

Numerical modelling

In this chapter I will discuss the most important techniques used to study astrophysical systems through numerical simulations, aimed at describing gravity and hydrodynamics.

A.1 Gravity

The main fundamental force shaping cosmic structures is gravity. According to In order to the ΛCDM cosmology the matter density in our Universe is dominate by dark matter, composed by yet unidentified non-baryonic elementary particles. In order to treat the evolution of this component one needs to model gravitational interaction between particles. In principle, one should follow each particle's orbit, and this could be achieved solving a huge N-body problem. Unfortunately, this is still unfeasible due to the computer technical limitations, so one needs to sample the matter density distribution with fewer particles (obtaining a set of fiducial macro-particles).

In order to describe the particle motion one can compute the gravitational acceleration due to the discrete set of particles, in the form

$$\ddot{\mathbf{r}}_i = -G \sum_{j=1}^{N} \frac{m_j}{[(\mathbf{r}_i - \mathbf{r}_j)^2 + \varepsilon^2]^{3/2}} (\mathbf{r}_i - \mathbf{r}_j), \tag{A.1}$$

where $\mathbf{r}_i$ and $\mathbf{r}_j$ are the particle positions and ε is the so-called gravitational softening. The purpose of the softening length is to avoid short distance interactions, which needs excessively large accuracy in the orbit integration and could result in bound particle pairs with manifestly correlated evolution. [1]

This calculation is exact, but becomes prohibitive for large N, since N partial forces must be computed for each of the N particles. The total computational cost would then be of order $\mathcal{O}(N^2)$. A possible solution is to *approximate* the force calculation through faster, dedicated schemes, which can be summarised in:

- Particle-mesh (PM) algorithm
- Fourier-transform based solvers of Poisson's equations
- Multigrid methods
- Tree algorithms
- TreePM methods

In addition, such methods can also be combined together to increase accuracy.

A.1.1 Particle-mesh technique

This approach uses an auxiliary mesh to reduce the number of elements needed to compute the gravitational forces. It involves four steps:

1. *Construction of a density field ρ on the mesh*
   The N particles with masses m_i and coordinates $\mathbf{r}_i$ (i=1,2,...,N) are put onto a cubical mesh with uniform spacing $h = L/N_{\rm g}$, where L is the mesh size. and a normalised shape function $S(\mathbf{x})$ is associated to each particle. One assigns to each cell a fraction $W_{\mathbf{p}}(\mathbf{x}_i)$ of particle i's mass falling within the cell with index $\mathbf{p}$. $W_{\mathbf{p}}(\mathbf{x}_i)$ corresponds to the overlap of the particle shape function with the cell and it is defined as

$$W_{\mathbf{p}}(\mathbf{x}_i) = \int_{\mathbf{x}_{\rm p}-h/2}^{\mathbf{x}_{\rm p}+h/2} S(\mathbf{x}_i - \mathbf{x}_{\rm p}) d\mathbf{x} \tag{A.2}$$

[1] A system showing this kind of behaviour is called 'collisional', while a system with uncorrelated particle orbits is called 'collisionless'.

The density field on the mesh can then be written as

$$\rho_{\mathbf{p}} = \frac{1}{h^3} \sum_{i=1}^{N} m_i W_{\mathbf{p}}(\mathbf{x}_i), \tag{A.3}$$

which depends on the chosen shape function. The most common functions used are (i) the Nearest grid point assignment, (ii) the Clouds in Cell assignment and (iii) the Triangular Shaped Clouds assignment. For a more detailed description see Springel (2014).

2. *Computation of the potential on the mesh by solving the Poisson's equation*
With the density field just obtained one then computes the potential, discretised on the mesh, by solving the Poisson's equation

$$\nabla^2 \Phi = 4\pi G \rho. \tag{A.4}$$

The most used methods for this are the Fourier-transform methods and the iterative solvers (which will be described in sections §A.1.2 and §A.1.3).

3. *Calculation of the gravitational field from the potential*
After having obtained the gravitational potential, the acceleration is computed from the Newton equation

$$\mathbf{a} = -\nabla \Phi. \tag{A.5}$$

The simplest way to achieve it is by finite differencing the potential to obtain a numerical derivative. Depending on the stencil length assumed the estimations of the acceleration on the mesh can be more or less accurate, at greater or lower computational cost. I report here only the 4th order case along the x direction, which needs a 5×5 cube centred at the cell of interest:

$$a_x^{(i,j,k)} = -\frac{1}{2h} \left\{ \frac{4}{3} \left[\Phi^{(i+1,j,k)} - \Phi^{(i-1,j,k)} \right] - \frac{1}{6} \left[\Phi^{(i+2,j,k)} - \Phi^{(i-2,j,k)} \right] \right\} + \mathcal{O}(h^4), \tag{A.6}$$

where i, j and k are the cell indexes along the x, y and z axes, respectively. We want to stress that a more accurate estimate is also more computationally expensive, but a general solution does not exist for all simulation setups, since other source of errors present in the simulation could be dominant.

4. *Calculation of the forces at the original particle location*
Once the force at the mesh cells are obtained, one needs to de-project them back to the particle locations. To this purpose one must use the same shape

function as used in the density field construction step, since a different choice would result in a non-vanishing self-force and asymmetric forces between particle pairs (Hockney & Eastwood, 1988). Therefore, the acceleration for a given particle with mass m and coordinates $\mathbf{x}$ can be computed as

$$\mathbf{F}(\mathbf{x}) = m \sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} W_{\mathbf{p}}(\mathbf{x}), \tag{A.7}$$

with $W_{\mathbf{p}}(\mathbf{x})$ defined in eq. (A.2).

A.1.2 Fourier techniques

The Fourier transform is a very powerful tool to solve certain kinds of partial differential equations. One of the most important examples which could take advantage of this technique is the Poisson's equation. One starts considering a non-periodic space, where the Newtonian potential generated by a distribution of masses can be written, in the continuum, as

$$\Phi(\mathbf{x}) = -\int G \frac{\rho(\mathbf{x}')d\mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|}. \tag{A.8}$$

The previous equation can be recognised as the convolution integral

$$\Phi(\mathbf{x}) = -\int g(\mathbf{x} - \mathbf{x}')\rho(\mathbf{x}')d\mathbf{x}', \tag{A.9}$$

where $g(\mathbf{x}) = -G/|\mathbf{x}|$ is the *Green's function* of Newtonian gravity. From the convolution theorem, one can compute the potential as

$$\Phi(\mathbf{x}) = \mathcal{F}^{-1}\left[\mathcal{F}(g) \star \mathcal{F}(\rho)\right], \tag{A.10}$$

which in Fourier space is

$$\hat{\Phi}(\mathbf{k}) = \hat{g}(\mathbf{k}) \cdot \hat{\rho}(\mathbf{k}). \tag{A.11}$$

where $\hat{g} = \mathcal{F}(g)$ and $\hat{\rho} = \mathcal{F}(\rho)$. In practice, to solve this equation in Fourier space, one needs to know the Fourier's transforms of both the density field and the Green's function. The problem can be solved both for periodic and non-periodic boundaries, but for sake of simplicity, I only describe the periodic solution, referring to Springel (2005) for a more detailed discussion. In a periodic box of size L, the density field can be expanded as Fourier's series

$$\rho(\mathbf{x}) = \sum_{\mathbf{k}} \rho_{\mathbf{k}} \exp^{i\mathbf{k}\mathbf{x}}, \tag{A.12}$$

where k extends on a discrete spectrum of wave vectors and is defined as

$$\mathbf{k} \in \frac{2\pi}{L} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \tag{A.13}$$

with n_1, n_2 and n_3 positive and negative integer numbers. Since ρ is a real field, the Fourier's series coefficients fulfil $\rho_{\mathbf{k}} = \rho^{\star}_{-\mathbf{k}}$. The Fourier's coefficients can be computed as

$$\rho_{\mathbf{k}} = \frac{1}{L^3} \int_V \rho(\mathbf{x}) \exp^{-i\mathbf{kx}} d\mathbf{x} \tag{A.14}$$

over a single instance of the box. Then, by inserting the Fourier's series expansion of ρ and Φ in the Poisson's equation, one obtains

$$\Phi_{\mathbf{k}} = -\frac{4\pi G}{\mathbf{k}^2} \rho_{\mathbf{k}}, \tag{A.15}$$

since the Poisson's equation should be valid for each of the Fourier modes separately. From eq. (A.15) we can infer that $g_{\mathbf{k}} = -4\pi G/\mathbf{k}^2$. With this definition and the density field Fourier's coefficients (obtained with a discrete Fourier transform algorithm), it is straightforward to compute the gravitational potential on the mesh.

A.1.3 Multigrid techniques

Another possible approach to solve Poisson's equation is the so-called multigrid technique. First, the 3D Poisson's equation is split in 3 separate 1D problems, i.e.

$$\frac{\partial^2 \Phi}{\partial x^2} = 4\pi G \rho(x). \tag{A.16}$$

Assuming that the potential is discretised on a regular mesh of size L with spacing h, one can approximate this expression by finite differences up to the second order, finding

$$\left(\frac{\partial^2 \Phi}{\partial x^2} \right)_i \simeq \frac{\Phi_{i+1} - 2\Phi_i + \Phi_{i-1}}{h^2}, \tag{A.17}$$

where i is the cell index. By inserting this definition in the Poisson's equation, one obtains

$$\frac{\Phi_{i+1} - 2\Phi_i + \Phi_{i-1}}{h^2} = 4\pi G \rho_i. \tag{A.18}$$

Since the considered mesh is regular, this equation is valid for all the N points of the mesh ($L = Nh$) and this results in a system with N equations and N unknowns Φ_i which could in principle be solved algebraically. It can be written as

$$\mathbf{Ax} = \mathbf{b}, \tag{A.19}$$

where $\mathbf{x} = (\Phi_i)$, $\mathbf{b} = \frac{4\pi G}{h^2}\rho$ and $\mathbf{A}$ is the coefficient matrix obtained from eq. (A.18). In order to obtain an exact solution one should use the LU decomposition or the Gauss elimination with pivoting (e.g. Press et al., 1992), but these procedures would become prohibitive for large N, because of their scaling as $\mathcal{O}(N^3)$. An alternative approach is based on 'approximate' iterative solvers, like the Jacobi and the Gauss-Seidel iterations, which are much faster.

Jacobi iteration
First, one considers the following matrix decomposition

$$\mathbf{A} = \mathbf{D} - (\mathbf{L} + \mathbf{U}), \tag{A.20}$$

where $\mathbf{D}$ is the diagonal part of $\mathbf{A}$ and $\mathbf{L}$ and $\mathbf{U}$ are the negative lower and upper diagonal parts, respectively. With few straightforward steps one obtains

$$\mathbf{x} = \mathbf{D}^{-1}\mathbf{b} + \mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})\mathbf{x}, \tag{A.21}$$

which can be used to define an iterative sequence of vectors $\mathbf{x}^{(n)}$:

$$\mathbf{x}^{(n+1)} = \mathbf{D}^{-1}\mathbf{b} + \mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})\mathbf{x}^{(n)}. \tag{A.22}$$

Note that $\mathbf{D}^{-1}$ is trivially obtained, since $\mathbf{D}$ is diagonal, i.e. $\mathbf{D}_{ii}^{-1} = 1/\mathbf{A}_{ii}$. This sequence is called Jacobi iteration (Saad, 2003), and converges only when the eigenvalues of the convergence matrix $\mathbf{M} = \mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})$ are less than 1.

Gauss-Seidel (GS) iteration
The GS iteration relies on the same decomposition used in the Jacobi scheme, but with a different iteration scheme, defined as

$$(\mathbf{D} - \mathbf{L})\mathbf{x}^{(n+1)} = \mathbf{U}\mathbf{x}^{(n)} + \mathbf{b}. \tag{A.23}$$

After few adjustments, the iteration becomes

$$\mathbf{x}^{(n+1)} = \mathbf{D}^{-1}\mathbf{U}\mathbf{x}^{(n)} + \mathbf{D}^{-1}\mathbf{L}\mathbf{x}^{(n+1)} + \mathbf{D}^{-1}\mathbf{b}. \tag{A.24}$$

In principle, a problem could arise, since the n state depends on the $(n+1)$ state. However, it can be noted that if one starts to compute the elements from the first row $i = 1$, the n state does not depend on the $(n+1)$ state anymore, since $\mathbf{L}$ has only elements below the diagonal. Then, when one moves to $i = 2, 3, ...$, only the $(n+1)$ values of the already computed rows are needed. This procedure speeds up convergence, with one GS step being close to two Jacobi steps. A problematic point about GS iteration is that it is strictly sequential and then cannot be parallelised. To overcome this problem one can use the so-called red-black ordering, which is a compromise between Jacobi and GS schemes (see Springel, 2014).

The multigrid technique

Since in every iteration only neighbours cells communicate, the information can only propagate at one cell per step, thus reducing the convergence speed for low frequency errors, which need more than one travels over the whole domain. In order to speed up the convergence, one can use an improved initial guess, obtained by solving the Poisson's equation on a coarser mesh. The main idea behind the multigrid technique is then to **prolong** and **restrict** the mesh in the so-called *V-cycle* to find an optimal guess to solve the Poisson's equation on the finest mesh. A schematic diagram of the multigrid cycle is shown in Fig. A.1.

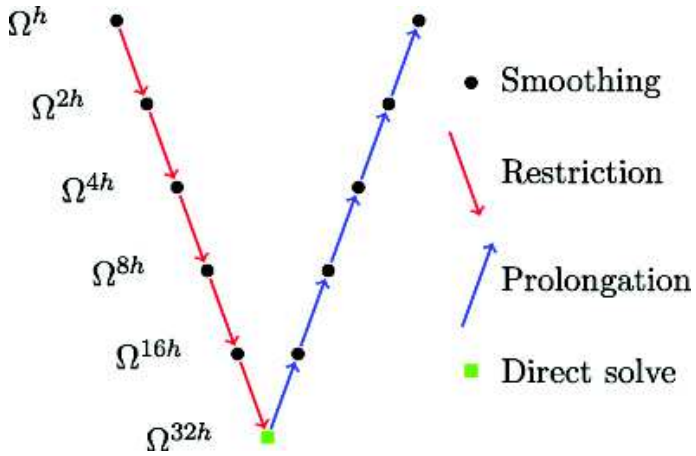

Figure A.1: *The typical V-cycle of a multigrid iteration scheme using 6 levels. The solution on a fine mesh is restricted down to the coarsest level (levels are represented as ω and h is the fine mesh spacing). The corrections found on the coarser levels are prolonged back to finer meshes, interleaving a Jacobi or GS iteration to find a solution for the corresponding level. Figure taken from Sampath, Barai & Nukala (2010).*

The computational cost of a single *V-cycle* is $\mathcal{O}(N_{\rm grid})$, where $N_{\rm grid}$ is the number of grid cells on the fine mesh. Since a convergence to machine precision requires several cycles, the total computational cost becomes $\mathcal{O}(N_{\rm grid} \ln N_{\rm grid})$. At the coarsest grid, however, the problem must be solved without a guess. This leads to the *full multigrid cycle* (see Fig. A.20, which can be summarised as:

1. Initialise the right-hand side on all levels, down to the coarsest grid
2. Solve the problem on the coarsest grid exactly (i)
3. Interpolate the solution to the $(i+1)$ level, obtaining an initial guess for this level
4. Solve the problem at the $(i+1)$ level with a full *V-cycle*

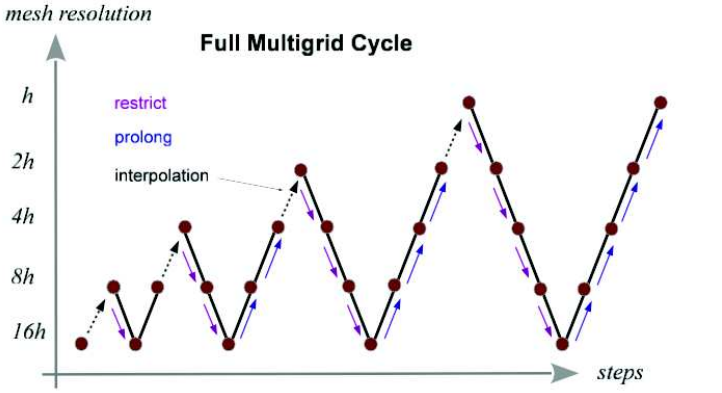

Figure A.2: *The full multigrid cycle aimed at addressing the problem of finding an adequate initial guess. Figure taken from Springel (2014).*

5. Repeat from step 3 until the finest level is reached.

The total computational cost of this scheme is still of the order of the number of mesh cells. This approach is used in the AMR RAMSES (Teyssier, 2002), which has been presented in section §2.1.

A.1.4 Hierarchical multipole methods: "Tree" codes

Another approach to compute the gravitational potential in real space is the so-called "tree algorithm" (Barnes & Hut, 1986).

The central idea is to use multipole expansion of a distant group of particles to describe its gravity, instead of each individual particle like in direct N-body codes. The potential of the group can be written as

$$\Phi(r) = -G \sum_i \frac{m_i}{|\mathbf{r} - \mathbf{x_i}|} = -G \sum_i \frac{m_i}{|\mathbf{r} - \mathbf{s} + \mathbf{s} - \mathbf{x_i}|}, \tag{A.25}$$

where $\mathbf{s}$ is the position of the group's centre of mass and $\mathbf{x}_i$ is the i-th particle position. The denominator can be Taylor expanded assuming $|\mathbf{x}_i - \mathbf{s}| << |\mathbf{r} - \mathbf{s}|$, which means that the angle θ under which the group is seen is small, i.e. $\theta \simeq l/y << 1$, where l is the mean size of the group. The first 2 expansion terms are written as

$$\frac{1}{|\mathbf{y} + \mathbf{s} - \mathbf{x}_i|} = \frac{1}{|\mathbf{y}|} + \frac{1}{2} \frac{\mathbf{y}^T \left[3(\mathbf{s} - \mathbf{x}_i)(\mathbf{s} - \mathbf{x}_i)^T - (\mathbf{s} - \mathbf{x}_i)^2\right] \mathbf{y}}{|\mathbf{y}|^5} + ..., \tag{A.26}$$

where $\mathbf{y} \equiv \mathbf{r} - \mathbf{s}$. The first term is the 'monopole' moment and the second one is the 'quadrupole' moment. The 'dipole' moment is not reported since it vanishes when the

expansion is relative to the group's centre of mass. In order to implement this particle grouping scheme one must rely on a suitable geometrical structure, i.e. a tree-like data structure, which gives the name to this kind of algorithms. After distributing the particles in a tree, the force computation is performed by means of a *tree walk*, where one starts from the root node and checks the node opening angle. If the node opening angle is larger than a tolerance value $\theta_{\rm c}$ the node is opened and all its sub-nodes must be considered, otherwise the multipole expansion is accepted and the walk along this branch is stopped. The computational cost of this algorithm is $\mathcal{O}(N \ln N)$, where N is the number of particles. These schemes are used, for example, in the SPH code GADGET2 (Springel, 2005) and in the new mesh-free Lagrangian code GIZMO (Hopkins, 2015), which descends from GADGET3, itself being a private modified version of GADGET2.

A.1.5 TreePM methods

This approach is at half-way between mesh and tree algorithms and it tries to capture the advantages associated to both of these methods. Indeed, while tree algorithms are the best choice for strongly clustered structures, mesh techniques work particularly well on almost homogeneous coarse grids. These methods have been initially proposed by Xu (1995) and Bagla (2002) and a version similar to this last one has been implemented in GADGET2 (and all codes descending from it). The main idea is to split the gravitational potential in Fourier space in two distinct components, a long-range and a short-range forces, and then use a PM approach to estimate the potential on large scales ($\mathbf{r} >> \mathbf{r}_{\rm s}$, where $\mathbf{r}_{\rm s}$ is the scale of the force split, which should be few times the mesh cell size) and a tree algorithm for the potential on scales smaller than $\mathbf{r}_{\rm s}$.

$$\Phi_{\mathbf{k}} = \Phi_{\mathbf{k}}^{\rm long} + \Phi_{\mathbf{k}}^{\rm short} \tag{A.27}$$

A.2 Hydrodynamics solvers

Galaxies are mainly composed by gas, stars and DM, each of them being intrinsically different from the others. From a purely gravitational point of view (which can be applied when we look at the basic laws describing cosmic structure formation) all of them are equivalent, but when one looks at small scales, the differences cannot be neglected anymore. In particular, we know that the gas component is a fluid and, if

we neglect magnetic fields, it can be described by the Navier-Stokes equations:

$$\begin{cases} \frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0 \\ \frac{\partial(\rho \mathbf{v})}{\partial t} + \nabla(\rho \mathbf{v}\mathbf{v}^T + P) = \nabla \Pi \\ \frac{\partial(\rho e)}{\partial t} + \nabla\left[(\rho e + P)\mathbf{v}\right] = \nabla(\Pi \mathbf{v}) \end{cases} \tag{A.28}$$

where $e = u + \mathbf{v}^2/2$ is the total energy per unit mass, u is the internal energy per units mass and Π is the viscous stress tensor, which is a material property. Each of them corresponds to a continuity equation, for mass, momentum and energy, respectively, and they form a set of hyperbolic conservation laws. However, we should note that this is an open system which cannot be solved, unless we introduce a fundamental closure relation, i.e. the gas equation of state. For typical astrophysical flows, where the gas has extremely low-density, the internal friction could become very small and then could be neglected ($\Pi = 0$). If this is the case, the equations can be simplified to give the ideal gas dynamics described by the Euler equations:

$$\begin{cases} \frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0 \\ \frac{\partial(\rho \mathbf{v})}{\partial t} + \nabla(\rho \mathbf{v}\mathbf{v}^T + P) = 0 \\ \frac{\partial(\rho e)}{\partial t} + \nabla\left[(\rho e + P)\mathbf{v}\right] = 0 \end{cases} \tag{A.29}$$

For an ideal gas, the equation of state is written as $P = (\gamma - 1)\rho u$, where $\gamma = c_p/c_v$ is the ratio of specific heats.

The Euler equations are commonly used to describe the gas component in astrophysics, but one should take into account that in certain regimes, like the hot plasma of rich galaxy clusters, the internal viscosity can become important, thus requiring the more general Navier-Stokes equations, which describe 'real' fluids. In this brief discussion of the main methods used for hydrodynamics I will only consider ideal gas and refer to Springel (2014) for additional details on Navier-Stokes equations.

In order to describe the gas evolution two approaches must be considered: an Eulerian approach, mainly used in mesh-based codes like RAMSES , and a Lagrangian approach, mainly used in SPH codes like GADGET2.

A.2.1 Eulerian hydrodynamics

Euler equations are expressed as partial differential equations (PDEs), a class of differential equations for which a general solution cannot be found. Rather, each PDE has its own class of solvers, suitably built to address it.

In order to solve the Euler equations, which form a system of non-linear equation, one needs a Riemann solver, a scheme devised to solve a Riemann problem, an initial value problem for hyperbolic PDEs where two constant half-spaces meet at a plane at $t = 0$ and for which one wants a solution for $t > 0$. In a Riemann problem, the solution typically contains three self-similar waves, i.e. a contact discontinuity (which marks the original fluid phases), a shock wave and a rarefaction wave. **However, in some cases, also two shock waves or two rarefaction waves can be found, instead of one of both.**

Different solution schemes aimed at solving PDEs have been developed, but I will only consider here the Finite Volume Method, which is probably the most common technique used in astrophysics.

Finite volume discretisation

I describe here how to derive an update scheme which takes advantage of Riemann solvers using a finite volume discretisation of the space. In this scheme the mesh used is fixed in time, that is why this is an example of Eulerian hydrodynamics. One starts defining the initial state $\mathbf{U}$ and the fluxes $\mathbf{F}$ as

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho\mathbf{v} \\ \rho e \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \rho\mathbf{v} \\ \rho\mathbf{v}\mathbf{v}^T + P \\ (\rho e + P)\mathbf{v} \end{pmatrix}, \tag{A.30}$$

where e is the total energy per unit mass and u the internal energy per unit mass. In a finite volume scheme the system is described through an averaged state over a set of finite cells. The state is then defined as

$$\mathbf{U}_i = \frac{1}{V_i} \int_{\text{cell}i} \mathbf{U}(\mathbf{x}) \mathrm{d}V. \tag{A.31}$$

By integrating the Euler equation over a cell and over a finite time interval, one obtains (after few steps)

$$\Delta x \left[\mathbf{U}_i^{(n+1)} - \mathbf{U}_i^{(n)}\right] + \int_{t_n}^{t_{n+1}} \mathrm{d}t \left[\mathbf{F}(x_{i+1/2}, t) - \mathbf{F}(x_{i-1/2}, t)\right] = 0. \tag{A.32}$$

Now, $\mathbf{F}(x_{i+1/2}, t)$ for $t > t_n$ is given by the solution of the Riemann problem on the cell's "right" side. At the interface, the solution is independent of time, so one can write

$$\mathbf{F}(x_{i+1/2}, t) = \mathbf{F}^\star_{i+1/2}. \tag{A.33}$$

The update scheme can then be written as

$$\mathbf{U}_i^{(n+1)} = \mathbf{U}_i^{(n)} + \frac{\Delta t}{\Delta x} \left[\mathbf{F}^\star_{i+1/2} - \mathbf{F}^\star_{i-1/2}\right]. \tag{A.34}$$

The idea behind the use of a Riemann solution in the update step is due to Godunov (Godunov, 1959) (*Godunov schemes*). A caveat of the procedure is that the obtained solution is no longer valid when the emanated waves cross the cell boundaries. This constraint leads to the so-called **Courant-Friedrichs-Levy** time-step criterion ($\Delta t \leq \Delta x/c_{\max}$, where $c_{\max}$ is the maximum wave-speed), which needs to be satisfied to get a correct solution to the problem.

The Godunov methods can be defined with a 3-step procedure called REA (Reconstruct, Evolve, Average): (i) *Reconstruct step*, when the cell-averaged quantities are used to define the quantities everywhere in the cell; (ii) *Evolve step*, when each cell's state is evolved forward in time by Δt; (iii) *Average step*, when the wave structure resulting from the Riemann problem solution is spatially averaged in a conservative fashion to compute the new states. This is implicitly obtained from the update scheme described above. Then the cycle is repeated again.

This scheme can be extended to multiple spatial dimensions and also to higher order integration accuracy both in space and time. The extension to multiple spatial dimensions can be done via:

- Dimensional splitting, where we split the 3D problem in three 1D problems. An example of a second order accurate operator in 3D is:

$$\mathbf{U}^{(n+1)} = \mathcal{X}(\Delta t/2)\mathcal{Y}(\Delta t/2)\mathcal{Z}(\Delta t)\mathcal{Y}(\Delta t/2)\mathcal{X}(\Delta t/2)\mathbf{U}^{(n)}, \tag{A.35}$$

  where each operator leads to an update of $\mathbf{U}$, thus requiring that fluxes have to be recomputed before the next time-evolution operator is applied. Each of these 1D operators is called *sweep*.

- Unsplit schemes, where all the flux updates are computed simultaneously to a cell. The update step can be used also for irregular shaped cells and is defined as

$$\mathbf{U}^{(n+1)} = \mathbf{U}^{(n)} - \frac{\Delta t}{V}\int \mathbf{F}\cdot \mathrm{d}\mathbf{S}, \tag{A.36}$$

  where the integration is over the whole cell surface, defined by the surface vectors $\mathrm{d}S$.

In order to extend the scheme to higher order of integration one needs to improve the reconstruction of the cell quantities. Different schemes have been developed for different orders, like the MUSCL-Hancock solver (piece-wise linear reconstruction, used in RAMSES), the PPM solver (with parabolic shapes instead of linear reconstruction), and so on.

A.2.2 Lagrangian hydrodynamics

In Lagrangian hydrodynamics the fluid is approximated through a set of "tracer" particles. Different techniques have been developed to describe the fluid in such a mesh-

free fashion. The fundamental technique is the so-called SPH. Recently new techniques have been developed to solve some of the known problems associated with the SPH formalism, like the new Godunov mesh-free algorithms implemented in the code GIZMO by Hopkins(2015; see also Springel, 2010, for the moving-mesh technique).

SPH formalism

Hydrodynamical equation of motions are derived for particles, leading to a simple and intuitive formulation of gas dynamics. This technique also shows excellent conservation properties. Moreover, the scheme does not show advection errors and is also Galilean invariant, unlike mesh-based Eulerian techniques.

The core of the SPH scheme is the kernel, an interpolation function used to estimate the fluid properties necessary to determine the rest of SPH equations through variational formalism. For any field one can define a smoothed interpolated version obtained through convolution with a kernel $W(\mathbf{r}, h)$, where h is the kernel size. The kernel is normalised to unity and approximates a δ-function for $h \to 0$. Another requirement is that the kernel is symmetric and sufficiently smooth to make it differentiable twice. The usually adopted kernel function by SPH codes is the cubic spline, which has finite support, defined in 3D as

$$W(r;h) = \frac{8}{\pi h^3} \begin{cases} 1 + 6\left(\frac{r}{h}\right)^2 \left(\frac{r}{h} - 1\right), & 0 < \frac{r}{h} \leq \frac{1}{2} \\ 2\left(1 - \frac{r}{h}\right)^3, & \frac{1}{2} < \frac{r}{h} \leq 1 \\ 0, & \frac{r}{h} > 1 \end{cases} \tag{A.37}$$

Alternative kernels have also been considered (Read, Hayfield & Agertz, 2010; Dehnen & Aly, 2012). In order to get an accurate property estimation, one needs at least a minimum number of neighbours ~ 33, corresponding to the number of points on a Cartesian mesh with spacing d included in a sphere of radius $2h$. The fluid properties in SPH schemes can then be computed as

$$F(\mathbf{r}) \simeq \sum_j \frac{m_j}{\rho_j} F_j W(\mathbf{r} - \mathbf{r}_j, h), \tag{A.38}$$

where the sum is performed over neighbour particles. As an example, the i-th particle density is

$$\rho_i = \sum_{j=1}^{N} m_j W(\mathbf{r}_i - \mathbf{r}_j, h_i). \tag{A.39}$$

Since the chosen kernel is differentiable, one can define a derivative operator on the fluid quantities, taking into account that the interpolation error would be higher than

that on the smoothed quantity. In general, the smoothing kernel can vary in space, and this is one of the key advantages of SPH, as it allows for intrinsically adaptive resolution.

The variability of the kernel size can be included in two different ways: a 'scatter' approach where the kernel function is $W(\mathbf{r}-\mathbf{r}_j, h(\mathbf{r}))$ with h evaluated at the neighbour's location (Hernquist & Katz, 1989) or a 'gather' approach, where $W(\mathbf{r}-\mathbf{r}_j, h(\mathbf{r}_i))$ where h is evaluated at the reference particle location. In the last approach the equations of motion implicitly include the variation of h in a self-consistent fashion. Another issue related to the kernel is how one determines h. Some codes like GASOLINE (Wadsley, Stadel & Quinn, 2004) use a discrete number of neighbours, while others like GADGET2 use a constant mass in the kernel volume, i.e. $\frac{4\pi}{3}\rho_i h_i^3 = N_{\rm NGB}\bar{m}$, where $\bar{m}$ is the average particle mass and $N_{\rm NGB}$ is an 'effective' number of neighbours.

SPH equations of motion

The SPH equations of motion can be then derived starting from the Euler equations written in a Lagrangian formulation:

$$\begin{cases} \frac{\mathrm{d}\rho}{\mathrm{d}t} + \rho\nabla\cdot\mathbf{v} = 0 \\ \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} + \frac{\nabla P}{\rho} = 0 \\ \frac{\mathrm{d}u}{\mathrm{d}t} + \frac{P}{\rho}\nabla\cdot\mathbf{v} = 0 \end{cases} \tag{A.40}$$

where $\mathrm{d}/\mathrm{d}t = \partial/\partial t + \mathbf{v}\cdot\nabla$ is the convective derivative. This formulation manifestly expresses conservation of mass, momentum and energy. Following Eckart (1960) the equations of motion can be derived discretising the fluid Lagrangian, written as

$$L_{\rm SPH} = \sum_i \left(\frac{1}{2}m_i\mathbf{v}_i^2 - m_i u_i\right), \tag{A.41}$$

where u_i is the internal energy per unit mass, which is expressed as an entropic function which implicitly relates it to the particle pressure P. The equations of motion can be obtained using the Euler-Lagrange equations, which give

$$m_i\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = -\sum_{j=1}^{N} m_j\frac{P_j}{\rho_j^2}\frac{\partial\rho_j}{\partial\mathbf{r}_i}. \tag{A.42}$$

Depending on the formulation used for the smoothing kernel (constant mass in the volume, discrete number of particles, etc...) the formulation is different. However, it

can be generally written as

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\sum_{j=1}^{N} m_j \left[f_i \frac{P_i}{\rho_i^2} \nabla_i W(\mathbf{r}_i - \mathbf{r}_j, h_i) + f_j \frac{P_j}{\rho_j^2} \nabla_j W(\mathbf{r}_i - \mathbf{r}_j, h_j) \right], \quad \text{(A.43)}$$

where f is a coefficient related to the formalism (Wadsley, Stadel & Quinn, 2004; Springel, Yoshida & White, 2001; Springel, 2005). Due to the intrinsic relation between the P, u and ρ, this equation can also be formulated in different ways, like pressure-entropy, density-energy, and so on. Note that the above formulation exhibits perfect energy, momentum and angular momentum conservation. From the entropic function equation one can also derive the time evolution for u, which gives

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = f_i \frac{P_i}{\rho_i^2} \sum_{j=1}^{N} m_j (\mathbf{v}_i \mathbf{v}_j) \cdot \nabla W(\mathbf{r}_i - \mathbf{r}_j, h_i). \quad \text{(A.44)}$$

Artificial viscosity

A problem with the Lagrangian formalism is that it is based on the differential form of Euler equations, which break down when gas experiences shocks, i.e. real discontinuities. When a shock occurs, entropy is not conserved anymore and the conservation laws must be described through the integral form of Eulerian equations. Therefore, the SPH formalism needs to be corrected to account for these effects, which typically occur in gas dynamics and lead to energy dissipation. The solution is provided by an *artificial viscosity* term, added to the SPH equations of motion to convert kinetic energy into heat when a shock occurs. The actual effect of artificial viscosity is to broaden the shock surface, so that the shock is resolved with a finite number of particles and thus described with the differential form of Euler equations. The additional terms can then be written as

$$\left.\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t}\right|_{\mathrm{visc}} = -\textstyle\sum_{j=1}^{N} m_j \Pi_{ij} \nabla \overline{W}_{ij} \quad \text{(A.45)}$$

$$\left.\frac{\mathrm{d}u_i}{\mathrm{d}t}\right|_{\mathrm{visc}} = \tfrac{1}{2} \textstyle\sum_{j=1}^{N} m_j \Pi_{ij} (\mathbf{v}_i - \mathbf{v}_j) \cdot \nabla_i \overline{W}_{ij}, \quad \text{(A.46)}$$

where $\overline{W}_{ij}$ is the average kernel between i and j particles (or, in some cases, the kernel computed with the average smoothing length $h_{ij} = 1/2(h_i + j_j)$). In order to avoid spurious dissipation outside a shock, the viscosity term Π is usually defined as a piece-wise function given by

$$\Pi_{ij} = \begin{cases} \left[-\alpha c_{ij} \mu_{ij} + \beta \mu_{ij}^2 \right] / \rho_{ij} & \mathbf{v}_i \cdot \mathbf{v}_j < 0 \\ 0 & \text{otherwise,} \end{cases} \quad \text{(A.47)}$$

where $\mu_{ih} = \frac{h_{ij}(\mathbf{v}_i - \mathbf{v}_j)\cdot(\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^2 + \epsilon h_{ij}^2}$, $\rho_{ij} = 1/2(\rho_i + \rho_j)$, $c_{ij} = 1/2(c_i + c_j)$ and $\epsilon \simeq 0.01$ to avoid singularities. The viscosity parameters α and β are typically chosen in the range $\alpha \simeq 0.5 - 1.0$ and $\beta = 2\alpha$.

Godunov mesh-free methods

In the last years a great effort has been done in the scientific community to improve Lagrangian methods, trying to solve the different problems affecting both mesh-based and SPH methods, with the purpose of capturing the advantages of both techniques. One of these new approaches is the "moving-mesh method" implemented by Springel (2010) in the new code AREPO, while another one has been proposed by Hopkins (2015) in the code GIZMO . While AREPO is not publicly available, GIZMO has been publicly released in a basic version. I here present the idea behind this second one only, since we employed it in my studies.

The scheme is derived considering the integral form of Euler equations (this is the first fundamental difference with SPH) in a frame moving with velocity $\mathbf{v}_{\rm frame}$, i.e.

$$\frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t} + \nabla \cdot (\mathbf{F} - \mathbf{v}_{\rm frame} \otimes \mathbf{U}) = 0, \tag{A.48}$$

where $\frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t} = \frac{\partial \mathbf{U}}{\partial t} + \mathbf{v}_{\rm frame}\nabla \otimes \mathbf{U}$ with $\otimes$ the tensor product. To solve the equations ensuring that shocks are correctly taken into account, it relies on a Galerkin-type method, where a weak solution is found by multiplying the equations by a test function ϕ (assumed to be Lagrangian, i.e. $\mathrm{d}\phi/\mathrm{d}t = 0$) and then integrating the result over a volume ω, obtaining

$$\int_\Omega \phi \left(\frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t} + \nabla \cdot \tilde{\mathbf{F}} \right) \mathrm{d}\Omega = 0, \tag{A.49}$$

where $\tilde{\mathbf{F}} = \mathbf{F} - \mathbf{v}_{\rm frame} \otimes \mathbf{U}$. Integrating by parts and assuming that the fluxes or ϕ vanish at infinity, one gets

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \mathbf{U}(\mathbf{x}, t)\phi \, \mathrm{d}^\nu \mathbf{x} - \int_\Omega \tilde{\mathbf{F}}(\mathbf{U}, \mathbf{x}, t) \cdot \nabla\phi \, \mathrm{d}^\nu \mathbf{x} = 0, \tag{A.50}$$

where the boundary term has been eliminated and the time derivative has been pulled out of the integral. If one wants to discretise this integral, it is necessary to choose a discretisation scheme for the domain. The idea is to consider a volume partition scheme in which a differential volume $\mathrm{d}^\nu \mathbf{x}$ at coordinates $\mathbf{x}$ is fractionally associated to the nearest cells (defined by their central tracer particle) through a weighting function W. In order to get a second-order accurate method conserving energy, linear and angular momentum W should be continuous, symmetric and with compact support. Based on these constraints, a good choice for such a function is the cubic spline kernel

usually adopted in SPH codes, even if the purpose of W is totally different from that of SPH schemes. According to the weighting function, the fraction associated to a particle i is

$$\psi_i(\mathbf{x}) = \frac{1}{\omega(\mathbf{x})} W(\mathbf{x} - \mathbf{x}_i, h(\mathbf{x})) \tag{A.51}$$

of the volume $\mathrm{d}^\nu \mathbf{x}$, where $\omega(\mathbf{x})$ is the sum of W over all the particles and $h(\mathbf{x})$ is the 'kernel' size, a continuous function defined all over the domain aimed at describing how many particles the volume is assigned to (if one defines W as a δ-function the Voronoi tessellation used in AREPO is recovered). Including the partition scheme into the volume integral and Taylor expanding all terms up to second-order, one obtains

$$\sum_i \phi_i \left[\frac{\mathrm{d}}{\mathrm{d}t}(V_i \mathbf{U}_i) + \sum_j (\tilde{\mathbf{F}}_{ij} \cdot \mathbf{A}_{ij}) \right] = 0, \tag{A.52}$$

where $V_i = \int \psi_i(\mathbf{x}) \mathrm{d}^\nu \mathbf{x} \simeq \omega(\mathbf{x}_i)^{-1}$ is the particle volume,$\tilde{\mathbf{F}}_{ij}$ is the flux solution of a time-centred Riemann problem and $\mathbf{A}_{ij}$ is an 'effective face area' associated with the particle pair(see Hopkins, 2015, for details). The equation above must be valid for an arbitrary test function ϕ, so the term in bracket must vanish itself, giving

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_i \mathbf{U}_i) + \sum_j (\tilde{\mathbf{F}}_{ij} \cdot \mathbf{A}_{ij}) = 0, \tag{A.53}$$

which is the form of traditional Godunov-type methods. The kernel size is obtained as in the SPH code GADGET2, as the length encompassing an effective number of neighbours, but with a formulation depending on the number density instead of the mass density, so that it only depends on the particle coordinates and not on the fluid properties.

Bibliography

Abramowicz M. A., Czerny B., Lasota J. P., Szuszkiewicz E., 1988, ApJ, 332, 646

Abramowicz M. A., Lasota J. P., 1980, Acta Astron., 30, 35

Agertz O. et al., 2007, MNRAS, 380, 963

Alvarez M. A., Wise J. H., Abel T., 2009, ApJ, 701, L133

Amaro-Seoane P. et al., 2013, GW Notes, Vol. 6, p. 4-110, 6, 4

Amaro-Seoane P., Brem P., Cuadra J., 2013, ApJ, 764, 14

Armitage P. J., 2013, Astrophysics of Planet Formation

Armitage P. J., Natarajan P., 2005, ApJ, 634, 921

Arun K. G. et al., 2009, Classical and Quantum Gravity, 26, 094027

Backer D. C., Jaffe A. H., Lommen A. N., 2004, Coevolution of Black Holes and Galaxies, 438

Bagla J. S., 2002, Journal of Astrophysics and Astronomy, 23, 185

Balick B., Brown R. L., 1974, ApJ, 194, 265

Barnes J., Hut P., 1986, Nature, 324, 446

Barth A. J., Martini P., Nelson C. H., Ho L. C., 2003, ApJ, 594, L95

Baumgarte T. W., Shapiro S. L., 1999, ApJ, 526, 941

Begelman M. C., 2010, MNRAS, 402, 673

Begelman M. C., Blandford R. D., Rees M. J., 1980, Nature, 287, 307

Begelman M. C., Rees M. J., 1978, MNRAS, 185, 847

Begelman M. C., Shlosman I., 2009, ApJ, 702, L5

Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289

Bellovary J. M., Governato F., Quinn T. R., Wadsley J., Shen S., Volonteri M., 2010, ApJ, 721, L148

Berczik P., Merritt D., Spurzem R., Bischof H.-P., 2006, ApJ, 642, L21

Berger M. J., Colella P., 1989, Journal of Computational Physics, 82, 64

Berger M. J., Oliger J., 1984, Journal of Computational Physics, 53, 484

Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press

Bleuler A., Teyssier R., 2014, MNRAS, 445, 4015

Bogdanović T., 2015, Astrophysics and Space Science Proceedings, 40, 103

Booth C. M., Schaye J., 2009, MNRAS, 398, 53

Bowyer S., Byram E. T., Chubb T. A., Friedman H., 1965, Science, 147, 394

Bromm V., Larson R. B., 2004, ARA&A, 42, 79

Bromm V., Loeb A., 2003, ApJ, 596, 34

Callegari S., Kazantzidis S., Mayer L., Colpi M., Bellovary J. M., Quinn T., Wadsley J., 2011, ApJ, 729, 85

Callegari S., Mayer L., Kazantzidis S., Colpi M., Governato F., Quinn T., Wadsley J., 2009, ApJ, 696, L89

Capelo P. R., Volonteri M., Dotti M., Bellovary J. M., Mayer L., Governato F., 2015, MNRAS, 447, 2123

Chapon D., Mayer L., Teyssier R., 2013, MNRAS, 429, 3114

Chevalier R. A., 1974, ApJ, 188, 501

Ciotti L., D'Ercole A., Pellegrini S., Renzini A., 1991, ApJ, 376, 380

Ciotti L., Ostriker J. P., 1997, ApJ, 487, L105

Ciotti L., Ostriker J. P., 2007, ApJ, 665, 1038

Clark P. C., Glover S. C. O., Klessen R. S., Bromm V., 2011, ApJ, 727, 110

Colpi M., 2014, Space Sci. Rev., 183, 189

Colpi M., Callegari S., Dotti M., Mayer L., 2009, Classical and Quantum Gravity, 26, 094029

Colpi M., Dotti M., 2011, Advanced Science Letters, 4, 181

Comastri A., Gilli R., Marconi A., Risaliti G., Salvati M., 2015, A&A, 574, L10

Couchman H. M. P., Rees M. J., 1986, MNRAS, 221, 53

Cuadra J., Armitage P. J., Alexander R. D., Begelman M. C., 2009, MNRAS, 393, 1423

Davies M. B., Miller M. C., Bellovary J. M., 2011, ApJ, 740, L42

Davies R. I., Müller Sánchez F., Genzel R., Tacconi L. J., Hicks E. K. S., Friedrich S., Sternberg A., 2007, ApJ, 671, 1388

Debuhr J., Quataert E., Ma C.-P., Hopkins P., 2010, MNRAS, 406, L55

Dehnen W., Aly H., 2012, MNRAS, 425, 1068

del Valle L., Escala A., 2012, ApJ, 761, 31

del Valle L., Escala A., Maureira-Fredes C., Molina J., Cuadra J., Amaro-Seoane P., 2015, ApJ, 811, 59

Devecchi B., Volonteri M., 2009, ApJ, 694, 302

Devecchi B., Volonteri M., Colpi M., Haardt F., 2010, MNRAS, 409, 1057

Devecchi B., Volonteri M., Rossi E. M., Colpi M., Portegies Zwart S., 2012, MNRAS, 421, 1465

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Dijkstra M., Haiman Z., Mesinger A., Wyithe J. S. B., 2008, MNRAS, 391, 1961

Dotti M., Colpi M., Haardt F., 2006, MNRAS, 367, 103

Dotti M., Colpi M., Haardt F., Mayer L., 2007, MNRAS, 379, 956

Dotti M., Ruszkowski M., Paredi L., Colpi M., Volonteri M., Haardt F., 2009, MNRAS, 396, 1640

Dotti M., Sesana A., Decarli R., 2012, Advances in Astronomy, 2012

Dotti M., Volonteri M., Perego A., Colpi M., Ruszkowski M., Haardt F., 2010, MNRAS, 402, 682

Dubois Y., Volonteri M., Silk J., 2014, MNRAS, 440, 1590

Dubois Y., Volonteri M., Silk J., Devriendt J., Slyz A., 2014, MNRAS, 440, 2333

Eckart C., 1960, Physics of Fluids, 3, 421

Eisenstein D. J., Loeb A., 1995, ApJ, 443, 11

Enßlin T. A., Pfrommer C., Springel V., Jubelgas M., 2007, A&A, 473, 41

Escala A., Larson R. B., Coppi P. S., Mardones D., 2005, ApJ, 630, 152

Fan X. et al., 2006, AJ, 131, 1203

Farris B. D., Duffell P., MacFadyen A. I., Haiman Z., 2014, ApJ, 783, 134

Farris B. D., Duffell P., MacFadyen A. I., Haiman Z., 2015, MNRAS, 447, L80

Ferland G. J. et al., 2013, Rev. Mexicana Astron. Astrofis., 49, 137

Ferrara A., Haardt F., Salvaterra R., 2013, MNRAS, 434, 2600

Ferrarese L., Merritt D., 2000, ApJ, 539, L9

Fiacconi D., Mayer L., Roškar R., Colpi M., 2013, ApJ, 777, L14

Fiore F. et al., 2012, A&A, 537, A16

Freese K., Spolyar D., Aguirre A., 2008, J. Cosmology Astropart. Phys., 11, 14

Fukugita M., Peebles P. J. E., 2004, Astrophys. J., 616, 643

Gabor J. M., Bournaud F., 2013, MNRAS, 434, 606

Ghisellini G., Celotti A., Tavecchio F., Haardt F., Sbarrato T., 2014, MNRAS, 438, 2694

Ghisellini G., Foschini L., Volonteri M., Ghirlanda G., Haardt F., Burlon D., Tavecchio F., 2009, MNRAS, 399, L24

Ghisellini G., Tagliaferri G., Sbarrato T., Gehrels N., 2015, MNRAS, 450, L34

Godunov S. K., 1959, Matematicheskii Sbornik, 89, 271

Gould A., Rix H.-W., 2000, ApJ, 532, L29

Greif T. H., Springel V., White S. D. M., Glover S. C. O., Clark P. C., Smith R. J., Klessen R. S., Bromm V., 2011, ApJ, 737, 75

Gültekin K. et al., 2009, ApJ, 698, 198

Gürkan M. A., Fregeau J. M., Rasio F. A., 2006, ApJ, 640, L39

Haehnelt M. G., Rees M. J., 1993, MNRAS, 263, 168

Haiman Z., 2004, ApJ, 613, 36

Haiman Z., Loeb A., 2001, ApJ, 552, 459

Häring N., Rix H.-W., 2004, ApJ, 604, L89

Hawking S., 1971, MNRAS, 152, 75

Heger A., Fryer C. L., Woosley S. E., Langer N., Hartmann D. H., 2003, ApJ, 591, 288

Hernquist L., 1990, ApJ, 356, 359

Hernquist L., Katz N., 1989, ApJS, 70, 419

Hicks E. K. S., Davies R. I., Maciejewski W., Emsellem E., Malkan M. A., Dumas G., Müller-Sánchez F., Rivers A., 2013, ApJ, 768, 107

Hockney R. W., Eastwood J. W., 1988, Computer simulation using particles. CRC Press

Hopkins P. F., 2015, MNRAS, 450, 53

Hurley J. R., Pols O. R., Tout C. A., 2000, MNRAS, 315, 543

Iocco F., 2008, ApJ, 677, L1

Johnson J. L., Bromm V., 2007, MNRAS, 374, 1557

Jones D. I., Burton M., Jones P., Walsh A., Rowell G., Aharonian F., 2011, ArXiv e-prints

Kazantzidis S. et al., 2005, ApJ, 623, L67

Kennicutt, Jr. R. C., 1998, ApJ, 498, 541

Khan F. M., Berentzen I., Berczik P., Just A., Mayer L., Nitadori K., Callegari S., 2012, ApJ, 756, 30

Khan F. M., Holley-Bockelmann K., Berczik P., Just A., 2013, ApJ, 773, 100

Khan F. M., Just A., Merritt D., 2011, ApJ, 732, 89

Khlopov M. Y., Rubin S. G., Sakharov A. S., 2005, Astroparticle Physics, 23, 265

Kim J.-h. et al., 2014, ApJS, 210, 14

Kim J.-h., Wise J. H., Alvarez M. A., Abel T., 2011, ApJ, 738, 54

Kormendy J., 2013, Secular Evolution in Disk Galaxies, Falcón-Barroso J., Knapen J. H., eds., p. 1

Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Koushiappas S. M., Bullock J. S., Dekel A., 2004, MNRAS, 354, 292

Latif M. A., Bovino S., Grassi T., Schleicher D. R. G., Spaans M., 2015, MNRAS, 446, 3163

Lodato G., Natarajan P., 2006, MNRAS, 371, 1813

Lupi A., Colpi M., Devecchi B., Galanti G., Volonteri M., 2014, Monthly Notices of the Royal Astronomical Society, 442, 3616

Lupi A., Haardt F., Dotti M., 2015, MNRAS, 446, 1765

Madau P., Haardt F., Dotti M., 2014, ApJ, 784, L38

Madau P., Rees M. J., 2001, ApJ, 551, L27

Madau P., Rees M. J., Volonteri M., Haardt F., Oh S. P., 2004, ApJ, 604, 484

Magorrian J. et al., 1998, AJ, 115, 2285

Mayer L., 2013, Classical and Quantum Gravity, 30, 244008

Mayer L., Fiacconi D., Bonoli S., Quinn T., Roškar R., Shen S., Wadsley J., 2015, ApJ, 810, 51

Mayer L., Kazantzidis S., Escala A., Callegari S., 2010, Nature, 466, 1082

Mayer L., Kazantzidis S., Madau P., Colpi M., Quinn T., Wadsley J., 2007, Science, 316, 1874

McConnell N. J., Ma C.-P., Gebhardt K., Wright S. A., Murphy J. D., Lauer T. R., Graham J. R., Richstone D. O., 2011, Nature, 480, 215

McKee C. F., Tan J. C., 2008, ApJ, 681, 771

Medling A. M. et al., 2014, ApJ, 784, 70

Merloni A., 2008, Mem. Soc. Ast. It., 79, 1310

Merritt D., 2013, Dynamics and Evolution of Galactic Nuclei

Merritt D., Mikkola S., Szell A., 2007, ApJ, 671, 53

Merritt D., Milosavljević M., 2005, Living Reviews in Relativity, 8, 8

Merritt D., Poon M. Y., 2004, ApJ, 606, 788

Mestel L., 1963, MNRAS, 126, 553

Milosavljević M., Bromm V., Couch S. M., Oh S. P., 2009, ApJ, 698, 766

Milosavljević M., Merritt D., 2001, ApJ, 563, 34

Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319

Mortlock D. J. et al., 2011, Nature, 474, 616

Novak G. S., 2013, ArXiv e-prints

Oh S. P., Haiman Z., 2002, ApJ, 569, 558

Omukai K., Nishi R., 1998, ApJ, 508, 141

Omukai K., Schneider R., Haiman Z., 2008, ApJ, 686, 801

Plummer H. C., 1911, MNRAS, 71, 460

Portegies Zwart S. F., Makino J., McMillan S. L. W., Hut P., 1999, A&A, 348, 117

Portegies Zwart S. F., McMillan S. L. W., 2002, ApJ, 576, 899

Power C., Nayakshin S., King A., 2011, MNRAS, 412, 269

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in C. The art of scientific computing. Cambridge: University Press

Preto M., Berentzen I., Berczik P., Spurzem R., 2011, ApJ, 732, L26

Pringle J. E., 1991, MNRAS, 248, 754

Quinlan G. D., Shapiro S. L., 1987, ApJ, 321, 199

Rasera Y., Teyssier R., 2006, A&A, 445, 1

Read J. I., Hayfield T., Agertz O., 2010, MNRAS, 405, 1513

Rees M. J., 1978, in IAU Symposium, Vol. 77, Structure and Properties of Nearby Galaxies, Berkhuijsen E. M., Wielebinski R., eds., pp. 237–242

Ripamonti E., Mapelli M., Ferrara A., 2007, MNRAS, 375, 1399

Roedig C., Dotti M., Sesana A., Cuadra J., Colpi M., 2011, MNRAS, 415, 3033

Roedig C., Sesana A., Dotti M., Cuadra J., Amaro-Seoane P., Haardt F., 2012, A&A, 545, A127

Roškar R., Fiacconi D., Mayer L., Kazantzidis S., Quinn T. R., Wadsley J., 2015, MNRAS, 449, 494

Russell D. M., Gallo E., Fender R. P., 2013, MNRAS, 431, 405

Saad Y., 2003, Iterative Methods for Sparse Linear Systems, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA

Safronov V. S., 1960, Annales d'Astrophysique, 23, 979

Saijo M., Baumgarte T. W., Shapiro S. L., Shibata M., 2002, ApJ, 569, 349

Sampath R. S., Barai P., Nukala P. K. V. V., 2010, Journal of Statistical Mechanics: Theory and Experiment, 2010, P03029

Sądowski A., Narayan R., McKinney J. C., Tchekhovskoy A., 2014, MNRAS, 439, 503

Schawinski K., Thomas D., Sarzi M., Maraston C., Kaviraj S., Joo S.-J., Yi S. K., Silk J., 2007, MNRAS, 382, 1415

Schawinski K., Virani S., Simmons B., Urry C. M., Treister E., Kaviraj S., Kushkuley B., 2009, ApJ, 692, L19

Schnittman J. D., 2013, Classical and Quantum Gravity, 30, 244007

Sesana A., Barausse E., Dotti M., Rossi E. M., 2014, ApJ, 794, 104

Sesana A., Khan F. M., 2015, MNRAS, 454, L66

Sesana A., Volonteri M., Haardt F., 2007, MNRAS, 377, 1711

Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337

Shi J.-M., Krolik J. H., Lubow S. H., Hawley J. F., 2012, ApJ, 749, 118

Shibata M., Shapiro S. L., 2002, ApJ, 572, L39

Shlosman I., Frank J., Begelman M. C., 1989, Nature, 338, 45

Silk J., Rees M. J., 1998, A&A, 331, L1

Soltan A., 1982, MNRAS, 200, 115

Spaans M., Silk J., 2006, ApJ, 652, 902

Spitzer L., 1987, Dynamical evolution of globular clusters

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2010, MNRAS, 401, 791

Springel V., 2014, ArXiv e-prints

Springel V., Yoshida N., White S. D. M., 2001, New A, 6, 79

Stacy A., Greif T. H., Bromm V., 2012, MNRAS, 422, 290

Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074

Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253

Taffoni G., Mayer L., Colpi M., Governato F., 2003, MNRAS, 341, 434

Tanaka T., Haiman Z., 2009, ApJ, 696, 1798

Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, ApJ, 474, 1

Teyssier R., 2002, A&A, 385, 337

Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, MNRAS, 429, 3068

The Enzo Collaboration, Bryan G. L. et al., 2014, ApJS, 211, 19

Toomre A., 1964, ApJ, 139, 1217

Trakhtenbrot B. et al., 2015, Science, 349, 168

Treister E., Schawinski K., Volonteri M., Natarajan P., Gawiser E., 2011, Nature, 474, 356

Trenti M., Stiavelli M., Michael Shull J., 2009, ApJ, 700, 1672

Van Wassenhove S., Volonteri M., Mayer L., Dotti M., Bellovary J., Callegari S., 2012, ApJ, 748, L7

van Wassenhove S., Volonteri M., Walker M. G., Gair J. R., 2010, MNRAS, 408, 1139

Vasiliev E., Antonini F., Merritt D., 2015, ApJ, 810, 49

Volonteri M., 2010, A&A Rev., 18, 279

Volonteri M., Bellovary J., 2012, Reports on Progress in Physics, 75, 124901

Volonteri M., Dotti M., Campbell D., Mateo M., 2011, ApJ, 730, 145

Volonteri M., Haardt F., Madau P., 2003, ApJ, 582, 559

Volonteri M., Natarajan P., 2009, MNRAS, 400, 1911

Volonteri M., Rees M. J., 2006, ApJ, 650, 669

Volonteri M., Silk J., Dubus G., 2015, ApJ, 804, 148

Wadsley J. W., Stadel J., Quinn T., 2004, New A, 9, 137

Wang L., Berczik P., Spurzem R., Kouwenhoven M. B. N., 2014, ApJ, 780, 164

Wild V., Heckman T., Charlot S., 2010, MNRAS, 405, 933

Willott C. J., Percival W. J., McLure R. J., Crampton D., Hutchings J. B., Jarvis M. J., Sawicki M., Simard L., 2005, ApJ, 626, 657

Wise J. H., Turk M. J., Abel T., 2008, ApJ, 682, 745

Xu G., 1995, ApJS, 98, 355

Yu Q., 2002, MNRAS, 331, 935

Zel'dovich Y. B., Novikov I. D., 1967, Soviet Ast., 10, 602

Zeldovich Y. B., Novikov I. D., 1971, Relativistic astrophysics. Vol.1: Stars and relativity