Massive black hole binaries in gas-rich galaxy mergers; multiple regimes of orbital decay and interplay with gas inflows

Lucio Mayer

Institute for Theoretical Physics, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
Kavli Institute for Theoretical Physics, UC Santa Barbara, 552 University Road Santa Barbara, CA 93106, USA

E-mail: lmayer@physik.uzh.ch

Received 18 July 2013, in final form 22 October 2013
Published 29 November 2013
Online at stacks.iop.org/CQG/30/244008

Abstract

We revisit the phases of the pairing and sinking of black holes (BHs) in galaxy mergers and circumnuclear discs in light of the results of recent simulations with massive BHs embedded in predominantly gaseous backgrounds. After a general overview we highlight for the first time the existence of a clear transition, for unequal mass BHs, between the regime in which the orbital decay is dominated by the conventional dynamical friction wake and one in which global disc torques associated with density waves launched by the secondary BH as well as co-orbital torques arising from gas gravitationally captured by the BH dominate and lead to faster decay. The new regime intervenes at BH binary separations of a few tens of parsecs and below, following a phase of orbital circularization driven dynamical friction. It bears some resemblance with planet migration in protoplanetary discs. While the orbital timescale is reasonably matched by the migration rate for the Type-I regime, the dominant negative torque arises near the co-rotation resonance, which is qualitatively similar to what is found in the so-called Type-III migration, the fastest migration regime identified so far for planets. This fast decay rate brings the BHs to separations of order $10^{-1}$ pc, the resolution limit of our simulations, in less than $\sim 10^7$ yr in a smooth disc, while the decay timescale can increase to $> 10^8$ yr in clumpy discs due to gravitational scattering with molecular clouds. Eventual gap opening at sub-pc scale separations will slow down the orbital decay subsequently. How fast the binary BH can reach the separation at which gravitational waves take over will be determined by the nature of the interaction with the circumbinary disc and the complex torques exerted the gas flowing through the edge of such disc, the subject of many recent studies. We also present a new intriguing connection between the conditions required for rapid orbital decay of massive BH binaries and those required for prominent gas inflows in gas-rich galaxies undergoing major mergers. We derive a condition for the maximum inflow rate that a circumnuclear disc can host while still maintaining a sufficiently high gas
density at large radii to sustain the decay of a BH binary. We find that gas inflows rates exceeding $10 \, M_\odot \, yr^{-1}$, postulated to form massive BH seeds in some direct collapse models, would stifle the sinking of massive BH binaries in gas-dominated galactic nuclei. Vice-versa, lower inflow rates, below a solar mass per year, as required to feed typical active galactic nuclei (AGNs), are compatible with a fast orbital decay of BH binaries across a wide range of masses.

PACS numbers: 04.25.dg, 98.62.Js, 04.25.dg, 07.05.Tp, 95.75.Pq, 95.30.Lz, 98.80.---k

(Some figures may appear in colour only in the online journal)

1. Introduction

The evolution of massive black hole (BH) binaries in gaseous backgrounds has become an active field of research in the last decade. Galactic nuclei hosting a significant gaseous component are expected to be ubiquitous in the modern scenario of hierarchical galaxy formation since not only they are present in disc-dominated galaxies, which host relatively light central black holes (Greene et al 2010), but are also expected in the disky high redshift progenitors of present-day early-type galaxies (Feldmann et al 2010), the hosts of the most prominent among present-day supermassive black holes (SMBHs) (Kormendy and Richstone 1995, Ferrarese and Merritt 2000). When these massive BHs merge they should produce one of the loudest gravitational waves signal, that can be detected with future gravitational wave experiments even at cosmological distances (Vecchio 2004). Most of the effort of theorists so far has focused on understanding the evolution of the binaries when the BH dynamics can still be treated in the Newtonian regime, although there have been isolated studies of late stages of the decay of BH binaries using a pseudo-Newtonian potential (Bogdanovic et al 2008). Indeed the evolutionary phase preceding the stage at which gravitational wave emission becomes important is still not fully understood. This is a pressing issue now that numerical relativity simulations have progressed far enough to be able to follow the final coalescence of massive BHs (Baker et al 2006, Rezzolla 2009). It is thus fair to say that we still do not know the timescales and efficiency of the BH pairing and merging process due to our incomplete knowledge of the Newtonian phase, despite the fact that past semi-analytical models (Volonteri et al 2003) and cosmological simulations (di Matteo et al 2008) assume that massive BHs merge on short timescales relative to the merging time of the galaxies.

In this paper we will review recent developments concerning BH binary decay in the Newtonian regime, focusing on the orbital evolution in predominantly gaseous backgrounds and, in particular, on the outcome of gas-rich galaxy mergers. We will also present and discuss for the first time novel results on two important aspects of the problem. The first concerns the nature and timescale of the decay process from separations of a few tens of pc to $\sim 0.1$ pc, which is a crucial regime since it marks the transition from a loose pair of BHs to a tightly bound binary in the case of medium-sized black holes of $10^6 \sim 10^7 M_\odot$, the main target of planned gravitational wave experiments such as eLISA (Amaro-Seoane et al 2013). Here we identify a transition to a fast, efficient regime of orbital decay that bears some similarity to planet migration in protoplanetary discs. The second concerns a new notion that we introduce and establish quantitatively, that of the interplay between gas inflows in gas-rich major galaxy mergers and orbital decay of BH binaries, showing how the strongest inflows stifle the decay of BH binaries. The latter analysis bears important implications on direct gas collapse models for seed BH formation and hence on the orbital evolution of high
redshift massive BH binaries. Finally, we will briefly review recent results appeared in the literature on yet another new regime of orbital decay, that of stochastic decay in presence of a clumpy interstellar medium (Fiacconi et al 2013), anticipating further new results obtained in new simulations of galaxy mergers with a multi-phase interstellar medium (ISM) that will be presented in detail elsewhere (Roskar et al 2013). Understanding BH evolution in merger remnants requires modelling hydrodynamics, gravity and radiation physics across several orders of magnitude in density, from galactic scales to sub-nuclear scales when the separation of the holes decreases below parsecs, a daunting task for even the most advanced astrophysical codes available. However, our paper also shows that, inspired by numerical results emerging from such complexity, relatively simple analytical models can be developed and used to make predictions on new aspects of the problem, one example being such interplay between gas inflows and BH binary decay.

2. Overview of orbital decay of massive BH binaries: from kpc to sub-pc scales

Since the seminal paper of Begelman et al (1980) it is assumed that a binary of massive BHs will evolve under a sequence of mainly three phases before gravitational wave emission can take over and bring the binary to coalescence: (I) the pairing phase in which dynamical friction (DF) of the galactic cores embedding the massive black holes, which are not yet mutually bound, drives their orbital decay; (II) a second phase in which the holes dynamically couple to form a binary and continue to sink due to the drag caused by DF against the surrounding background of gas and stars; (III) a final phase in which the binary hardens via three-body scatterings off single stars. As we will discuss in this paper, in gaseous backgrounds there are additional aspects of the environment and nature of the drag that complicate this simple scheme, pointing towards the existence of more than one regime of decay at both large and small scales. We will show, for example, how the second phase is driven not solely by conventional DF since global disc torques are potentially at least as important. Furthermore, in an inhomogeneous background, drag onto BHs might come from clumps, non-axisymmetric features and other sources of gravitational torques that render the orbital decay inherently stochastic in phase II. Additionally, the drag dominated by gas and stars will co-exist in general, with no net separation between the second and last phase. Three-body encounters with stars may take over and drive final coalescence in the remnants of gas-rich galaxies because of favourable orbital structure allowing to refill steadily the loss cone (Khan et al 2012), overcoming the ‘last parsec problem’ noticed in earlier stellar dynamical simulations (Milosavljevic and Merritt 2001, Berczik et al 2006).

In this paper, however, we will focus on the effect of the gas alone. The latest results of numerical simulations suggest the following distinct regimes of BH orbital decay in gas-rich galaxy mergers:

- (1) The large scale pairing of massive BHs as their host galaxy cores are still in the process of merging, for separations from several 10 kpc to a few tens of pc, corresponding to the first phase in the sequence of Begelman et al (1980).
- (2) Orbital decay in a smooth circumnuclear disc driven by DF, for BH binaries with separations of 100 pc to below 1 pc. In this case the drag results from a wake excited by the moving black hole, which exerts a gravitational back-reaction onto it (the wake can be trailing as well as leading in rotating backgrounds). The circumnuclear disc, a few 100 pc in size, forms as a result of the merger between the two galaxies, because gas does not dissipate completely its angular momentum via shocks and gravitational torques during the merger (Barnes 2002, Mayer et al 2008).
• (3) Orbital decay in a smooth circumnuclear disc driven by disc torques, which bears analogies with Type-I and Type-III migration in planetary evolution (Lin and Papaloizou 1979, Papaloizou et al. 2007). In this case the drag at BH separations of a few tens of pc, hence well inside the disc, arises by torques, primarily excited at Lindblad resonances and at the co-rotation resonance. At variance with the DF wake, here it is also the gas at some distance from the secondary BH that extracts angular momentum from it rather than just that along the wake, and torque can be strong even when the relative velocity between the BH and the background is small.

• (4) Orbital decay in a clumpy disc: stochastic migration, at separations of 100 pc to parsecs. This is a whole new regime, which is expected to arise since the interstellar medium in galaxies, including galactic nuclei, is actually clumpy and multi-phase. Clumps of the size of giant molecular clouds (GMCs) can scatter the BHs, eventually ejecting them from the disc.

• (5) Orbital decay in circumbinary disc. It occurs separations of less than 0.1 pc, and can bear some similarity with Type-II migration in planetary evolution if a deep gap is formed between the binary and the disc. It takes place when the interaction between the target body and the rotating gaseous background becomes nonlinear due to a large mass ratio and/or low pressure support in the surrounding gas; disc torques from the binary become strong enough to repel disc material away (Goldreich and Tremaine 1980, Papaloizou et al. 1997), creating a gap and leading to slow migration occurring roughly on the viscous timescale of a circumbinary disc (Armitage and Natarajan 2005). In more realistic self-gravitating or magnetized discs, the emerging scenario is more complex though, as many recent works show that significant gas flow occurs through the gap and that the torque results from the direct tidal interaction between the BH binary and the (tidally deformed) circumbinary disc (Cuadra et al. 2006, Roedig et al. 2012, Roedig and Sesana 2013, Shi et al. 2013, del Valle and Escala 2012). This phase might in principle leads directly to the gravitational wave dominated regime by-passing the third phase driven by three-body encounters present in the Begelman et al. scenario.

Among these three regimes, the first two have been thoroughly studied with 3D hydrodynamical simulations, both smooth particle hydrodynamics (SPH) and adaptive mesh refinement (AMR), (5) has received increased interest in the last few years, with the first global 3D hydro as well as magnetohydrodynamics (MHD) simulations, (4) is beginning only now to be explored, and, finally, the distinction between (2) and (3) has been overlooked so far. A qualitative analysis of the differences between (2) and (3) is one of the main objectives and novel aspects of this paper.

Let us now proceed with a brief overview of our understanding of the orbital decay phase driven by DF. DF has historically been studied primarily for non-dissipative backgrounds, such as to understand the drag exerted by a larger stellar system or dark matter halo on a smaller system or halo, such as a dwarf galaxy satellite or globular cluster orbiting in the Milky Way halo (e.g. White 1978, Colpi et al. 1999). Restricted to a static, infinite, isotropic and homogeneous collisionless medium, the most commonly adopted approach to DF is Chandrasekhar’s formula (1943). The latter assumes that the stars/dark matter particles move along unperturbed trajectories that are straight lines, thus neglecting the self-gravity present in real backgrounds. Despite this limitation, the resulting formula for the drag force has been often used in a local fashion to estimate sinking times also in inhomogeneous self-gravitating backgrounds.

More general approaches to DF have been developed in the last two decades, with the aim of overcoming the various limitations of Chandrasekhar’s formula. Among these we recall
approaches based on global torques amplified at resonances (e.g. Weinberg 1989) and the theory of linear response (TLR) (e.g. Bekenstein and Maoz 1992, Colpi et al 1999), which both attempt to capture the global response in a finite background, taking into account at least partially its self-gravitating nature. We refer to Colpi and Dotti (2009) for a thorough overview on the subject.

For illustrative purpose we can consider the example of a target body orbiting in a stellar (spherical) bulge or dark matter halo, for which a singular isothermal sphere represents a good approximation for the density profile. In a singular isothermal sphere with 1D velocity dispersion \( \sigma \) and density profile \( \rho(r) = \sigma^2/[2\pi G r^2] \). TLR predicts a sinking time, expressed in terms of the circularity \( \epsilon \) (which is the ratio between the angular momentum of the actual orbit relative to that of a circular orbit of equal energy)

\[
\tau_{df} = 1.2 \frac{r_{\text{circ}}^2 V_{\text{circ}}}{\ln(M_{\text{halo}}/M_{\text{BH}})GM_{\text{BH}}\epsilon^{0.4}}
\]

where \( r_{\text{circ}} \) and \( V_{\text{circ}} \) are the initial radius and velocity of the circular orbit with the same energy of the actual orbit, and \( M_{\text{halo}} \) is the mass of the dark matter and stars within \( r_{\text{circ}} \). Applied to the case of a massive black hole, the above equation implies that the latter can sink at the centre of the sphere within a time \( \sim 10^6 \) yr, if released during the merger at a distance of \( r_{\text{circ}} \sim 100 \) pc.

\[
\tau_{df} \sim 5 \times 10^8 \left( \frac{5}{\ln(M_{\text{halo}}/M_{\text{BH}})} \right) \left( \frac{r_{\text{circ}}}{300 \text{ pc}} \right)^2 \left( \frac{V_{\text{circ}}}{\sqrt{2100 \text{ km s}^{-1}}} \right) \left( \frac{10^6 M_{\odot}}{M_{\text{BH}}} \right) e^{0.4} \text{ yr}. \quad (2)
\]

Let us now consider gaseous backgrounds, still infinite and homogenous. In a smooth gaseous background it is useful to consider two limits of DF onto a body moving with velocity \( V \), the supersonic regime (\( V/ V_s > 1 \)) and the subsonic regime (\( V/ V_s < 1 \)). We refer here again to the review by Colpi and Dotti (2009). Following Ostriker (1999) in the steady-state limit and for supersonic motion, the drag on a massive perturber moving with velocity \( V \) across a homogeneous fluid with density \( \rho_{\text{gas}} \) and sound speed \( c_s \) reads

\[
F_{\text{gas}} = -4\pi \left( \frac{b_{\text{max}}}{b_{\text{min}}} (M^2 - 1)^{1/2} \right) \frac{G^2 M^2_{\text{bh}} \rho_{\text{gas}} V}{\sqrt{3}}, \quad \text{for } M > 1
\]

where \( M = V/c_s \) is the Mach number.

The deceleration results, in this regime, by the enhanced density wake that lags behind the perturber and that is confined in the narrow Mach cone. Note that the gaseous drag is enhanced for supersonic motion (with \( M \sim 1 - 2.5 \)) compared to Chandrasekhar’s formula in collisionless backgrounds (for \( \rho_s = \rho_{\text{gas}} \) by a factor \( \sim 2 \) (compare with standard formula in collisionless background on e.g. Binney and Tremaine 1987).

In the subsonic limit, instead, the drag can be much weaker than in the collisionless case. The drag is exactly zero in a homogeneous infinite medium due to the front–back symmetry of the perturbed density distribution present in any stationary solution. However, in a finite medium if the perturber triggers the disturbance at time \( t = 0 \) and moves subsonically along a straight line, the symmetry is broken as long as \( (c_s + V)t \) is smaller than the size of the medium, resulting in a finite drag \( F_{\text{df}} = -(4/3)\pi G^2 M_{\text{BH}}^2 \rho_{\text{gas}} M^5 V/V_s^5 \propto M_{\text{BH}}^2 \rho_{\text{gas}}, V/c_s^4 \) for \( M \ll 1 \).

Consider now the case of a BH moving in a rotating background. Suppose also that the BH is moving on a circular orbit. It can be either corotating or counter-rotating with the background. In the corotating case, which should be a natural configuration (even in galaxy encounters with fairly high relative inclination lead to corotating BH binary configurations in nuclear discs as a result of large scale gravitational torques, see Callegari et al (2011)), the BH would have not net velocity relative to the gas in its immediate vicinity (note that the BH can still have a Mach number larger than unity since that involves the ratio between the speed
of the BH and the thermal sound speed, which is typically smaller than the circular velocity of the disc for astrophysically meaningful situations), gas located further away will have a non-negligible relative velocity, but is gravitational pull will be weaker. This implies that in such conditions equation (3) predicts a small drag, certainly smaller than in the case in which the BH is moving at a significant speed with respect to the background, as on a non-circular orbit in a rotating background (a caveat is, of course, that equation (3) is strictly derived under the assumption of a static background). Hence we would expect the orbital decay of a BH to slow down on circular orbits relative to eccentric orbits in a circumnuclear disc. As we will see in the next section, the orbital decay of a BH in a smooth circumnuclear disc accelerates instead of slowing down on circular orbits because the nature of the drag is more complex than suggested by equation (3).

3. Large scale pairing of massive BHs; the remarkable difference between major and minor mergers

When the two host galaxies with their extended dark matter halos have not completed the merger yet, the two BHs evolve as their host galaxies. The merging timescale of halos and galaxies strongly depends on their mass ratio. Normally in galaxy formation mergers with mass ratios larger than 1:3 are considered minor, while those with lower mass ratio are considered major. In nearly equal mass (major) mergers the outcome is trivial; the two galaxy cores merge rapidly, in a few Gyr, after DF onto the extended dark matter halos has eroded their relative orbital energy and angular momentum.

Multi-scale simulations of galaxy mergers with embedded MBHs, both with SPH and AMR codes, have shown that, after the merging of the two cores, the two BHs decay to parsecs separations, forming a tight binary in a circumnuclear disc on timescales as short as a few Myr for realistic assumptions on the thermodynamics of the ISM (Mayer et al 2007, Chapon et al 2013), as shown in figure 1. Additional effects due to the presence of a clumpy multi-phase ISM, which are studied in more recent simulations (Fiacconi et al 2013, Roskar et al 2013), will be described in section 5.

From a dynamical point view, a mass ratio of about 1:5 constitutes a more physically motivated boundary; indeed for a higher mass ratio the DF timescale becomes longer than the tidal disruption timescale, defined as the characteristic time over which the smaller halo/galaxy loses a significant fraction of mass due to mutual tides (Taffoni et al 2003). As a result, below a mass ratio of about 1:5 the smaller galaxy will be disrupted before the two galaxy cores can merge. This has the consequence that BHs might be left wandering at kiloparsecs from the centre rather than binding into a binary, as first pointed out in Kazantzidis et al (2005).

Indeed, once the more massive galaxy core hosting them has been dissolved by tides, their DF timescale inside the primary galaxy can become longer than the Hubble time due to their relatively tiny mass. Already at the halo level, the disruption timescale depends sensitively on the internal mass distribution of both the lighter and the more massive primary halo. In particular, this translates into a dependence on halo concentration for NFW-like (Navarro–Frenk–White) halo profiles, as thoroughly addressed in Taffoni et al (2003), and could depend on the precise dark halo inner slope if baryonic effects can flatten the cuspy density profiles typical of cold dark matter (CDM) galaxy halos as recently suggested (e.g. Kazantzidis et al 2013).

With all other things being equal, the disruption timescale can be even more severely affected by the physics of the baryonic cores. Indeed, even if the halo is almost entirely disrupted, whether or not a dense core survives in the smaller galaxy depends on the efficiency of gas dissipation, which acts to raise the central density and renders the core more resilient
Figure 1. Evolution of the SMBH binary separation in the simulations of equal mass galaxy mergers with equal mass massive BHs presented in Chapon et al. (2013). The simulations shown have different resolution and polytropic index of the equation of state. The highest resolution simulation, corresponding to the black line, reaches a spatial resolution of 0.1 pc while the other simulations have a spatial resolution ten times lower. They were all carried out with the AMR code RAMSES. The inset shows that the separation of the two massive BHs begins to fluctuate around a fraction of a parsec, suggesting possible stalling. Courtesy of MNRAS.

Callegari et al. (2009, 2011) found that baryonic physics determines the outcome of minor mergers as far as the binding of BHs into a binary is concerned. The efficiency of gas dissipation depends on the relative weight of radiative cooling versus radiative and mechanical heating processes, such as that resulting from feedback due to stellar irradiation and supernovae explosions. These processes also depend indirectly on the mass distribution in galaxies and on the orbit of the galaxy encounter, which affect the strength of torques concentrating gas to the cores of galaxies during the interaction, as well as the efficiency of star formation. This explains the marked dependence on the pairing timescale of BHs on the orbital parameters and gas fraction of the host galaxies found by Callegari et al. (2011), as shown in figure 2. Figure 2 also shows that the conditions in which pairing of binary MBHs is most efficient are also those in which the secondary MBH grows more by gas accretion, which should promote a subsequent efficient hardening of the binary at smaller separations. Note that these results hold strictly for the 1:10 mergers considered by these authors, which, however, are among the most typical merging events in hierarchical galaxy formation (Volonteri et al. 2003).

Simulations are still limited by the small size of the parameter space explored in terms of galaxy orbits, assumptions on the internal structure of galaxies and modelling of various important processes such as star formation, stellar feedback, and black hole accretion and feedback, which all appear to influence the pairing process. Overall, extrapolating the orbital decay rates to smaller separations, hence assuming there is no bottleneck at MBH binary separations below parsecs the simulations suggest that a binary of MBHs with initial mass...
ratio of 1:10 will decay down to a separation at which gravitational waves drive fast coalescence (less than $10^{-2}$ AU) in a time varying from $10^8$ to $10^9$ yr. As explained in the next section, gap opening at small separations (below 0.1 pc) will slow down the orbital decay but the associated sinking timescale at such small separations is still negligible relative to that of pairing at large scales. Note that an overall sinking timescale close to 1 Gyr for minor mergers is significantly longer than the timescale expected for equal mass MBHs sinking in major mergers, which, as stated above, is likely below $10^7$ yr (Mayer et al 2007, Chapon et al 2013).

4. Orbital decay of MBH binaries in circumnuclear gas discs. I: the smooth ISM case

Let us first look now at the decay in a smooth gaseous circumnuclear disc. There are obvious limitations in the analytical approach to the drag force presented in the previous section when the goal is to describe the orbital evolution of MBHs in realistic galaxy mergers, in which the background is inherently highly inhomogeneous and time-dependent, and has stars and dark matter in addition to gas. Nevertheless, Chapon et al (2013), using AMR simulations with the RAMSES code to model galaxy mergers with embedded BHs, have provided convincing evidence that the orbital decay of BH pairs on eccentric orbits is well described by DF once they are embedded in the common circumnuclear disc arising when the two galactic cores merge. They were able to resolve with unprecedented detail the trailing wake produced by the BHs (figure 3), whose strength, as well as that of the associated drag force, was shown to vary with the Mach number (defined as the ratio between the speed of the massive BH and that of the thermal sound speed) as expected in the analytical theory of Ostriker (1999). This is shown in figure 4. Chapon et al (2013) also found that the drag was shutting off at the scale at which DF due to the wake is expected to break down, namely when the separation of the
Figure 3. Colour-coded density maps in the proximity of a massive BH showing the time evolution of the hydrodynamical wake due of DF in one of the AMR galaxy merger runs of Chapon et al (2013). The hydrodynamical wakes and Mach cones are shown for increasing values of $\mathcal{M} = V_{\text{BH}}/c_s$. These wakes are clearly stronger and make the DF much more efficient when the black hole is in a transonic regime ($\mathcal{M} = 1.14, 1.18, 1.23$). Courtesy of MNRAS.

two holes encompasses a gas mass comparable or lower than the sum of their own masses. The orbital decay was thus seen to stall at separations of $\sim 0.1$–$0.5$ pc. Escala et al (2006) had noticed a similar problem in their simulations of circumnuclear discs but had shown that in some cases the decay could be restarted via local torques induced by an ellipsoidal gas deformation around the BH binary. The possibility that orbital decay occurs as a result of the gravitational interaction with gas orbiting around the BH binary is reminiscent of Type-III migration in planetary evolution, a term used to refer to a regime dominated by torques acting in the co-orbital region (see Papaloizou et al 2007), which is actually the fastest migration regime so far identified.

Is this pointing to a potential last parsec problem in gaseous media, in analogy to the difficulties found in purely stellar dynamical backgrounds when the orbital distribution of stars does not allow efficient loss cone refilling (see Milosavljevic and Merritt 2001, Berczik et al 2006)?

The way out likely lies in the existence of the second regime in which torques resulting from the non-axisymmetric distribution of the gas in the surrounding disc, rather than the effect of the wake, dominate the drag. The torques in this case can have a more global character than when the DF wake dominates the drag. Such a regime is analogous to the Type-I and Type-III migration thoroughly studied in planet migration (Papaloizou et al 2007). In a Type-I regime one expects the secondary black hole orbiting around the primary to excite density waves, which in turn cause torques that exchange angular momentum with the hole. If the exchange of angular momentum with the gas in the co-orbital region dominates, namely it is the non-axisymmetric distribution of the gas orbiting the secondary that causes the main torque, then one calls this Type-III migration. The spiral wave triggered by the secondary BH is clearly seen in figure 5, which shows a snapshot of the evolution of a BH binary in a circumnuclear disc modelled using a polytropic equation of state (Fiacconi et al 2013). In linear theory for laminar viscous discs one can see that in the Type-I regime the inward directed torque, which extracts angular momentum, is maximized at the outer Lindblad resonance.
Figure 4. Top: colour-coded density map of the gas disc showing the location of resonances in the disc after orbit circularization (at a time of 5 Myr since the beginning) in a simulation using a disc mass $M_D = 10^8 M_\odot$, a primary black hole with mass of $10^7 M_\odot$ and a secondary black hole with mass of $5 \times 10^5 M_\odot$ (the initial eccentricity is $e \sim 0.2$). The black dots show the location of the primary and secondary BHs. Bottom: the differential torque averaged over a quarter of a Myr around the time shown in the top panel, with the location of resonances overplotted. The actual tangential velocity of the black hole has been used to define the co-rotation resonance, as shown in the panel. Broad peaks corresponding to outer and inner Lindblad resonances and a sharp peak around the co-rotation resonance is evident. The maximum torque is clearly negative, as expected for inward directed migration. The simulation was first presented in Fiacconi et al. (2013) for other purposes.
between the orbital frequency and the spiral pattern frequency. This normally wins over the outward pushing torque, which peaks at the inner Lindblad resonance, as long as there is more disc material outside the orbit of the hole rather than inside. It is thus expected that the negative torque will dominate when the hole is well inside the disc. No saturation of the torque is expected as long as there is gas exterior to the orbit of the holes that can torque them. Therefore there is no natural bottleneck of the orbital decay, contrary to the case of the regime dominated by the local DF wake, hence no ‘last parsec problem’ of the kind found in Chapon et al (2013) is expected in this case. In figure 6, which shows the evolution of the angular momentum of the secondary BH for a set of simulations originally presented in Fiacconi et al (2013), it is evident that there is a first phase of slow decay and then a second phase of faster decay. The first phase is well described by the effect of the local DF wake while

Figure 5. Colour-coded density map in the smooth circumnuclear disc simulations shown in the previous figure at the beginning of the first decay phase dominated by the local DF wake, which is clearly evident (top), and soon after circularization has occurred (bottom), when both the triggered spiral density wave and the accumulation of gas in the co-orbital region of the secondary BH are apparent. The time since the beginning of the simulation is indicated in the panels, and the black dots represent the primary and secondary BHs.
Figure 6. Time evolution of the angular momentum (left) and separation of two massive BHs (right) with two different mass ratio $q$ on two different orbits with eccentricity $e$ as indicated in the labels. The massive BHs evolve in a smooth circumnuclear disc, the mass of which is indicated in the panel. The orbital eccentricity diminishes as a result of DF, as can be seen by the fact that the angular momentum evolves little while the separation decreases in the first part of the evolution. After the orbit circularizes significantly the last part of the decay becomes faster and is driven by spiral wave and co-orbital torques, as shown in the figure 5.

The onset of the second phase coincides with the appearance of a strong triggered spiral density wave as well as an increase of gas density very close to the secondary BH, in the so-called co-orbital region, both regimes being highlighted in figure 5. In order to clarify the nature of the second regime and its actual connection with planet migration, we analysed the torques before and after circularization occurs, and found importance differences. In figure 4 we show the results for one representative simulation. Before circularization the torque varies from negative to positive as the secondary BH goes from pericentre to the apocentre of the orbit, and it is extremely local (most of the torque comes from the region of about 1 pc around the secondary). After circularization, when the triggered spiral density wave becomes evident, the torque is instead always negative and is not strictly local. This can be seen in figure 4, where one can also correlate the location of resonances in the disc with the variation of torque amplitude as a function of frequency. The strongest torque contribution comes from higher order outer Lindblad resonances, which correspond to the region of the spiral wave close to the secondary, and from the gas near co-rotation. Since the peak of the negative torque is located near co-rotation it is tempting to say that the orbital evolution takes place in a Type-III regime, although the existence of a negative torque near the location of the outer Lindblad resonances cannot be neglected. We also caution that in our simulations the gas is self-gravitating, viscous, and the effect of pressure is non-negligible, these being conditions different from those in which standard migration theory has been developed for planets. For example, one expects resonances to be broadened by all these effects, which would explain why there are no sharp peaks in the torque distribution except near co-rotation.

Assuming linear theory for inviscid isothermal discs, the migration rate scales inversely with the mass of the decaying BH as in the case of DF. Under these idealized conditions it is possible to derive a simple equation for the migration rate which applies to the Type-I regime but includes co-orbital torques (Tanaka 2002). While in our case co-orbital torques are stronger than in conventional Type I, other effects such as a finite viscosity due to self-gravity,
and the fact that our discs are polytropic or adiabatic rather than isothermal, generally go in the direction of reducing the total negative torque (Bitsch et al 2013). Therefore we will consider the result of this estimate as purely indicative of the order of magnitude of the decay timescale. We also note that a simple formula describing Type-III migration is not available since results obtained in different conditions in the literature vary a lot, only agreeing on the fact that the migration timescale is significantly shorter than in Type-I migration for the same viscosity and thermodynamical conditions (Peplinski et al 2008). Hence, for relatively light secondary BHs, with masses $< 10^8 M_{\odot}$ that of the disc, following Tanaka (2002), we can write:

$$\tau_{\text{mig}} = (2.7 + 1.1 \beta)^{-1} \frac{M_1^2}{M_2 \Sigma a^3} h^2 \Omega^{-1}$$

(4)

where $M_1$ is the mass of the primary black hole, $M_2$ the mass of the secondary black hole, $a$ is the orbital separation of the two black holes, $\Sigma$ the disc surface density, $h$ its scale height and $\Omega$ its angular velocity. The equation is derived assuming laminar, non-self-gravitating discs, and includes 3D effects and co-rotation torques, for a disc with surface density profile $\Sigma \sim r^{-\beta}$, scale height $h$ and orbital frequency $\Omega$. We can assume $\beta = 1$, as in the convenient Mestel disc model adopted in many simulations (see also section 6). For secondary black holes with masses up to $M_2 \approx 10^8 M_{\odot}$ this equation should be applicable given the large masses of circumnuclear discs adopted in the simulations ($M_2 \gg 5 \times 10^8 M_{\odot}$). For reasonable values of the parameters ($M_1 = 10^7 M_{\odot}$, $\Sigma = M_2 / 2 \pi R_d^2$, with $M_2 = 5 \times 10^8 M_{\odot}$, $R_d = 100$ pc, $a = 10$ pc, $h = 10$ pc, $\Omega = \nu_c / a$, with the circular velocity $\nu_c$ being a constant in a Mestel disc, $\nu_c = 100$ km s$^{-1}$) the result is $\tau_{\text{mig}} \sim 1$ Myr, in broad agreement with the numerical results (see the sharp angular momentum decrease following orbital circularization in figure 6).

Note that this regime of orbital decay was not observed in the published simulations of galaxy mergers (Mayer et al 2007, Chapon et al 2013), which instead seem to be well understood based on the effect of the local trailing wake of DF all the way down to the smallest BH separation reached at the centre of the disc. The fundamental reason behind the difference with respect to the results found in the simulations of circumnuclear discs is not clear yet. However, the fact is that in the circumnuclear disc simulations the Type-I/III migration regime is normally preceded by a phase of orbital circularization induced by DF. Instead, in the merger simulations the orbit of the black hole binary remains eccentric and the transition to the Type-I/III regime never occurs (Mayer et al 2007). The role of circularization might be just coincidental but certainly needs to be clarified since it appears to mark the transition between orbital decay driven by the DF wake and by disc torques. We have verified that an envelope of co-orbital material around the secondary BH only forms after its orbit has circularized, probably because only then reduced relative velocity allows the gas to undergo gravitational capture by the BH. Also, if resonances are crucial in sustaining a spiral density wave, it is clear that such wave can be triggered more effectively when the motion of the BH acquires a well defined frequency after circularization has occurred. The other important ingredient to enter a Type-I/III phase might be the presence of a relatively cold, weakly turbulent disc in which triggered spiral waves can be more efficiently sustained and the BH can capture enough material in the co-orbital region to produce a strong co-rotation torque. The lack of favourable thermodynamical conditions might partially explain why the Type-I/III like regime is not seen the multi-scale merger simulations of Mayer et al (2007) and Chapon et al (2013), which have rather hot and turbulent discs.

The general conditions under which orbit circularization operates effectively are still unclear. Circularization is important not only to understand the transition to the Type-I/III regime but also because the merger rate of BHs once they enter the gravitational wave regime is a strong function of orbital eccentricity. There is evidence that circularization requires a rotating background, be it gaseous or stellar. Indeed, detailed studies of orbital eccentricity
evolution in non-rotating stellar and gaseous backgrounds have never observed circularization even after a large number of orbits; among these the studies of orbital decay of satellite galaxies and/or dark matter halos (e.g. Colpi et al 1999). Circularization has been observed recently in rotating spherical stellar backgrounds (Sesana et al 2011), and had been demonstrated earlier in rotating circumnuclear discs of stars and gas (e.g. Dotti et al 2007). Dotti et al (2007) provide the following qualitative explanation of why circularization occurs. They relate it to the position of the wake at apocentre versus pericentre; they find that at apocentre the wake is in front of the black hole, causing an increase in its tangential velocity because disc material rotates at a speed faster than that of the black hole, causing a positive torque, while at pericentre the wake is behind the black hole and causes a decrease of its velocity. The net result is an increase of the ratio between tangential and total velocity of the black hole over an orbit, which effectively leads to a lower eccentricity. This would not happen without rotation, since in this case the wake is always behind the black hole. The net circularization will be dependent on the rotation curve of the disc since that will determine the magnitude of the effect at apocentre. On the other end, rotation in the background is clearly not a sufficient condition for circularization. Indeed, as we already noticed in the multi-scale simulations of galaxy mergers with embedded BHs circularization is not observed (Mayer et al 2007, Chapon et al 2013). The difference might be that in these mergers the resulting circumnuclear disc is more massive and self-gravitating, with self-sustained spiral waves that could force the eccentricity to remain large, as it has been shown to be possible for migrating planets in asymmetric discs (Papaloizou and Larwood 2000).

Star formation would reduce the density of the gaseous disc, increasing both the DF and the Type-I migration timescales. However, it has been shown that, for comparable disc mass distribution and thickness, BHs sink at similar rates in a mainly gaseous or mainly stellar disc (Dotti et al 2007), suggesting that neglecting the effect of star formation in this entire discussion of the sinking regime is a reasonable approximation.

Even if it enters a fast Type-III phase the decay may still stall if the BHs are able to open a gap, as it is well known in planetary evolution. This happens when the gap-opening gravitational torque exerted by the rotating binary onto the surrounding gaseous medium prevails over the gap-closing viscous torque associated with turbulent/viscous diffusion in the circumbinary gas. A useful form of the gap opening condition was derived by Escala et al (2005), under the assumption that the viscous/turbulent speed characterizing the magnitude of the gap-closing torque is equivalent to the effective sound speed in the nuclear disc, that a clear gap has a size $\Delta r \sim 3h$, where $h$ is the disc scale height set by the balance between pressure and gravitational forces (the criterion was calibrated on simulations’ results), and that the BH obeys the observed $M_{BH} \sim \sigma$ relation between black hole mass $M_{BH}$ and central stellar velocity dispersion $\sigma$ (this being proportional to the total mass within the orbit of the BH binary):

$$M_{BH} \geq (h/pc)^2 \times 7.2 \times 10^5 M_\odot.$$

Assuming a mean disc thickness of 30–40 pc, corresponding the effective sound speed of 50–60 km s$^{-1}$, consistent with the gas velocity dispersions in local circumnuclear discs of interacting or merging galaxies, such as ULIRGs and LIRGs (Downes and Solomon 1998, Medling et al 2013) one finds that the mass at which BHs are expected to open a gap is quite large, of order $10^5 M_\odot$. This neglects the complication of a multi-phase ISM, in which the cold molecular gas phase, distributed in the midplane of the circumnuclear disc, might have a shorter scale height, of order 5–10 pc. Still, even for a massive BH binary confined to the thin cold molecular gas layer a fairly massive pair of BHs will be needed to open a gap, with a total mass $M_{BH} > 10^7 M_\odot$ based on equation (5). However, in the latter case the lower limit is small
enough that BHs in massive early-type spirals as well as in typical elliptical galaxies would be able to open a gap, making the regime relevant to the general evolution of BH binaries.

The possibility of gap opening has motivated studies which model the evolution of a massive BH binary at small separations ($a < 0.1$ pc) within a gap embedded in a common circumbinary disc (Armitage and Natarajan 2005, Cuadra et al 2009). Assuming that a deep gap can be opened, orbital decay in this configuration can still proceed on a timescale proportional to the viscous timescale of the sub-pc scale circumbinary accretion disc with which the binary exchanges angular momentum. Following Armitage and Natarajan (2005), the binary orbit will shrink at a rate:

$$\frac{da}{dt} \propto -\alpha \left(\frac{h}{r}\right)^2 q_d \Omega a$$

where $q_d = \frac{\pi a^2 \Sigma}{M_f^2}$, $a$ being the semi-major axis of the binary orbit, $\Sigma$ the disc surface density, $\Omega$ the angular frequency of the gas, $h$ the disk thickness and $\alpha$ the viscosity, with $\alpha$ in the range $10^{-4} - 0.1$ depending on the mechanisms that can generate viscosity, such as the magneto-rotational instability or gravitational (Toomre) instability. For $\alpha \sim 10^{-3}$ equation (6) leads to a merger timescale of order $10^7$ yr according to simple 2D numerical Newtonian models which simply add the phase of decay via gravitational wave emission to obtain the total merger timescale (Armitage and Natarajan 2005). Note that such timescale is comparable or longer than the orbital decay timescale required for BHs to sink from 100 pc to pc scale separations in galaxy merger simulations or circumnuclear disk simulations (1–10 Myr).

In the last few years there have been several works that have explored the interaction of the binary BH with a circumbinary disk using fully 3D global hydrodynamical simulations with self-gravity of the gas included (Cuadra et al 2009, Roedig et al 2012, Roedig and Sesana 2013, del Valle and Escala 2012, Hayasake et al 2007, Shi et al 2012, Nixon et al 2011, Nixon 2012, Farris et al 2011). They have highlighted several differences relative to the earlier results just described. As a result, the behaviour departs from the predictions of standard Type-II migration. First, they have shown that there is always significant gas flow through the gap, which causes time-dependent torques acting differently on the primary and on the secondary BH (Roedig et al 2012, Roedig and Sesana 2013, del Valle et al 2012). Secondly, the torque that causes the binary to shrink is not associated with the viscosity of the flow as in Type-II migration, rather it is caused by the tidal interaction between the binary and the surrounding disk. The gas flows through the gap even in absence of explicit viscosity in the fluid since self-gravity alone generates an effective viscosity in the flow, transporting angular momentum outward as a result of the nonlinear tidal interaction between the binary and the disc. As a result, the shrinking time of the binary inside the circumbinary disc becomes strongly dependent on the mass of the latter, which would not be the case in Type-II migration; for BH binaries with mass exceeding $10^7 M_\odot$ in low mass discs the decay time in this phase can become longer than the Hubble time (Cuadra et al 2009). Torques from material accreting onto the black hole play an important role too, and in some cases are crucial to ensure that overall a negative torque is felt by the binary, promoting its shrinking (Roedig et al 2012). A much faster shrinking of the binary, below $10^5$ yr, might be promoted by a retrograde circumbinary disc, which could arise in chaotic accretion scenarios because in such a case the tidal torque between the disc and the binary is maximized (Nixon et al 2011, Nixon 2012). While most of these simulations start with a binary already well inside the circumbinary disc, so that some kind of gap is readily opened, del Valle and Escala (2012) have recently revisited the conditions for gap opening in the first place. Using a fairly large set of simulations they have found that, even for massive binaries gaps are opened only in very thin, cold discs, with scale heights of a few pc, namely much smaller than those inferred from observations of circumnuclear disc structure in merger remnants (10–50 pc). A caveat is that observations do not probe the central pc region of merger
remnants yet, hence currently one can only extrapolate from information available at much larger radii (>10 pc). Nevertheless the subject warrants more investigation both with theory and observation since, irrespective of how deep a gap can be formed, the character of the orbital decay changes once it is governed by the interaction of the binary with the circumbinary disc as opposed to that between the secondary BH and the circumnuclear disc at larger scales. Finally, most of these recent simulations do not include effects associated with magnetic fields that might be important in the dense, hot gas sitting close to the potential well of the black holes. Indeed in the circumbinary accretion disc the gas becomes gradually warmer and more ionized towards smaller radii, so that faster angular momentum transport can arise by MRI or, more in general, via magnetic stresses (Menou and Quataert 2001). If this is the case a deep gap would never be cleared even for very massive binaries in cold, thin discs as MHD stresses push more matter inside the gap, allowing the binary to couple more strongly with the surrounding gas and hence be torqued more effectively. This is essentially what has been found in recent general-relativistic MHD simulations of circumbinary discs, where the binary shrinkage occurs nearly three times faster than in hydrodynamical simulations (Shi et al 2012).

5. Orbital decay of a MBH binary in circumnuclear gas discs. II: stochastic decay in a clumpy ISM

In a gravitoturbulent nuclear disc, such as that which should arise in presence of self-gravity, cooling and heating/stirring via feedback mechanisms (Wada and Norman 2001, Agertz et al 2009), the medium is rapidly filled with cold and dense gaseous with sizes comparable with molecular clouds (1–10 pc). A simple way to produce a gravitoturbulent state, without including the various heating and cooling source terms directly, is to introduce a phenomenological dissipative term in the internal energy equation (Fiacconi et al 2013). This dissipation term effectively represents the net cooling rate in the system. If it is large enough massive, gravitationally bound clumps are produced from rapid fragmentation. The resulting gravoturbulent state is an extreme version of a clumpy medium since the resulting clumps are compact and tightly bound while molecular clouds are fluffy objects, often unbound, subject to internal feedback from stellar radiation, stellar winds and supernovae, yet these experiments are useful to understand and quantify the role of a clumpy medium in the context of the orbital decay of BHs (Fiacconi et al 2013). The main consequence is that orbital decay becomes stochastic. BHs can either be slowed down or accelerate their inward migration, and even be ejected temporarily from the disc plane depending on how they exchange energy and angular momentum in gravitational scatterings with clumps. Massive clumps, corresponding to the scale size/mass of GMCs, namely of order $10^6 M_\odot$ or larger, are mostly responsible for scattering holes outside the disc plane.

Once outside the disc plane the hole can sink back down into the gas disc owing to DF from the stars and dark matter. Most of the contribution to the mass, hence to the drag, comes from the stars. The drag in this regime is much smaller than in the disc since the average density of the background is much smaller and DF is also slightly less effective than in gaseous backgrounds (see also Escala et al 2005). For a range of bulge parameters and stellar density profiles (from Plummer to Sersic-like distribution with Sersic index in the range 1–2, to represent from classical bulges to pseudobulges), Fiacconi et al (2013) find that the decay timescale back to the disc varies from 20 Myr to nearly $10^8$ yr. Once back in the disc the holes can shrink to sub-pc scale separations in less than 1 Myr even in clumpy discs, confirming previous results with smooth media in which the BH was never leaving the disc plane (Mayer et al 2007, Dotti et al 2007, Chapon et al 2013). Hence it is the time spent inside the stellar-dominated background that introduces a bottleneck in the decay process.
Note that at $z > 2$ massive disc galaxies are observed to be clumpy at kpc scales, with clumps of masses approaching even $10^7-10^8 M_\odot$, hence much larger than the mass of present-day GMCs. While it is not clear yet whether or not such clumps are long-lived bound condensations produced by gravitational instability in a turbulent ISM as the clumps in the simulations of Fiacconi et al (2013), scattering of massive BHs out of the plane of the nuclear disc should thus occur even more at high redshift. Since at $z > 1$ the detection rate by future gravitational wave experiments should be fairly high, it will be important to clarify the quantitative role of clumpiness in a large set of simulations. Clumps can also aid the orbital decay of MBHs by binding with one of them, thus increasing the effective mass subject to friction and/or disc torques, as found by Fiacconi et al (2013).

We estimated the BH mass threshold $M_{\text{BH}}$ below which these stochastic effects due to a clumpy medium are going to be important. This was obtained by estimating the threshold mass $M_\bullet$ below which a massive BH orbiting in a disc with mass $M_g$ and scale radius $R$ will likely be scattered by clumps with $M_{cl} > M_{\text{BH}}$. We can imagine to describe these encounters as a steady-state diffusion process in which for any massive clump migrating inward and dissolving at the centre there are others forming in the disc at larger radius and maintaining the same level of density fluctuations. Under these assumptions we can adopt a constant number density of clumps $n_{cl}$ in the disc and that the BH moves radially in the disc via a random walk as it encounters various clumps. For a given cross section $\sigma_{cl} = \pi r_g^2$ of the clumps, where $r_g$ can be computed from the gravitational sphere of influence of the clump, we have that the number of clumps that the BH can be scattered by is given by $N_{cl} \sim R^2 n_{cl}^2 \sigma_{cl}^2$. If we consider typical circumnuclear disc parameters and that only the biggest clumps matter, those with mass $M_{cl} \sim 10^6 M_\odot$, comparable to the mass of the secondary BH, we obtain (see Fiacconi et al 2013 for the intermediate steps of the calculation),

$$M_\bullet \sim 7 \times 10^7 \left( \frac{N_{cl}}{4} \right)^{1/4} \left( \frac{\eta}{0.4} \right) \left( \frac{M_g}{10^9 M_\odot} \right) M_\odot,$$

(7)

where $\eta$ is the gas fraction of the circumnuclear disc, $M_g$ its (gas) mass and $N_{cl}$ the number of very massive clumps of mass comparable to that of the sinking BH. From the above equation we see that, at low-$z$, when clump masses are at most as large as those of GMCs ($\sim 10^6 M_\odot$), black holes with up to a few $10^7 M_\odot$ can be affected by gravitational scattering with the clumps, while at high-$z$, when gas clumps possibly produced by gravitational instability can reach masses as large as $10^9 M_\odot$ according to observations of galaxies at high-$z$ (Genzel et al 2011), black holes as large as $\sim 10^9 M_\odot$ can be affected, essentially spanning nearly whole mass spectrum of massive BHs in present-day galaxies.

A further step forward in the modelling of a clumpy medium can be achieved by replacing the single dissipative term with a sub-grid model for a multi-phase ISM, including optically thin radiative cooling, absorption and scattering of radiation at high optical depths via a phenomenological temperature–density relation calibrated with radiative transfer calculations, star formation and supernovae feedback using the blastwave approach successfully employed in galaxy formation simulations with the GASOLINE SPH code (see e.g. Mayer 2010). Multi-scale galaxy merger simulations reaching 0.1 pc resolution with the aid of particle splitting have been carried out with all the latter ingredients and will be presented in a forthcoming paper (Roskar et al 2013). An example of such computationally demanding state-of-the-art simulations, which comprise several millions gas and star particles in the nuclear discs emerging from the merger, timesteps as small as a few hundred years and a spatial resolution of 0.1 pc, is shown in figure 7. In these simulations a gravoturbulent medium arises naturally, with a broad mass spectrum of cloud/clumps, from a few thousand solar masses to $> 10^6 M_\odot$. The effect of clump-BH scattering is confirmed to delay the orbital decay significantly due to
ejections out of the disc plane, bringing it up to nearly $10^8$ yr in some instances, in substantial agreement with the results of Fiacconi et al. (2013).

6. Interplay between orbital decay of massive BHs and inflow-driven massive BH seed formation

Among massive seed BH formation mechanisms many rely on multi-scale gas inflows, either triggered by non-axisymmetric instabilities in marginally unstable protogalaxies, such as bars-within-bars (Begelman et al. 2006, Lodato and Natarajan 2006, Regan and Haehnelt 2009, Choi et al. 2013, Latif et al. 2013) or by merger-driven inflows in collisions between gas-rich massive galaxies at high redshift (Mayer et al. 2010, Bonoli et al. 2013). Mayer et al. (2008) noticed that in circumnuclear discs that are strongly non-axisymmetric, such as those that could feed a massive BH seed in the aforementioned models, the orbital decay of a massive BH binary stalls at a separation of parsecs (the simulations had a resolution as high as 0.1 pc). They argued that the reason was that in such discs the gas inflow is so prominent and fast to overtake the BH binary in its journey to the centre; essentially the BHs undergo only weak friction when they are still at several parsecs from the centre because gas is funnelled to the centre very rapidly, reducing the ambient density considerably and hence the strength of the drag. The implication is that, in order for the binary to shrink to sub-pc separations and coalesce in a gaseous environment, the circumnuclear disc has to be relatively stable so that the inflow does not overtake the holes. If this is true the epochs of massive BH seed formation and efficient massive BH coalescence are likely distinct, since they require different conditions in the galactic nucleus. These conclusions, however, were based on the result of a single very hi-res simulation with a fixed mass of the two BHs (which were of equal mass, $\sim 3.6 \times 10^6 M_\odot$).

We can explore further the overall relevance of this result, as well as if it is physically grounded, by building a simple analytical model. We want to compare the torque exerted by
DF onto a massive BH binary in a circumnuclear disc with the viscous torque causes the gas inflow. Transport of angular momentum via spiral waves, bars or other kinds of non-axisymmetric instabilities in self-gravitating discs can indeed be described based on an effective viscosity to zeroth order (Lin and Pringle 1987, Lodato and Natarajan 2006, Mayer et al 2010). As we mentioned above, Type-I migration torques are also involved in the orbital decay, but they dominate only at small separations between the holes when the orbit of the binary has circularized, hence we can rightly consider the DF dominated phase as the main bottleneck. We will adopt a simple alpha-disc model (Shakura and Sunyaev 1973), with a mass distribution described by a Mestel disc, widely used in simulations (e.g. Dotti et al 2007, Fiacconi et al 2013). The latter is a convenient choice to keep the equations simple. A Mestel disc has indeed the nice feature that it has a constant circular velocity $v_c$ and its surface density is given by $\Sigma = \Sigma_0 \times (R_0/R) = M_d (2\pi R_0^2) R_0/R$, where $M_d$ is the disc mass, $\Sigma_0$ and $R_0$ normalization factors for the surface density and the radius, respectively. This model strictly applies to a razor-thin disc but we will make the assumption that this is a good approximation for a disc in which $h/R \sim 0.05$ (having a finite scale height will be useful in order to be able to include the thermodynamics via consideration of the disc vertical structure). The magnitude of the viscous torque between two adjacent disc annuli at $R$ can be written as:

$$\Gamma_{\text{visc}} = 2\pi v\Sigma R^3 d\Omega/dR$$ (8)

where $\Omega$ is the angular velocity and we set the viscosity $\nu = \alpha c_sh$, where $\alpha \leq 1$ in a standard Shakura–Sunyaev alpha disc, $c_s$ is the sound speed and $h$ the effective disc scale height. This has to be compared with the DF torque acting on a black hole binary of mass $M_{BH}$ and moving at speed $v_{BH}$. We will consider the supersonic case only ($v_{BH} > c_s$) since this is the regime in which orbital decay is more efficient (see section 2) and also the more relevant one for massive BHs delivered in the typical conditions of galaxy mergers (Chapon et al 2013, see also figure 3). In this case we can use equation (3) for the DF force $F_{DF}$ so that the magnitude of the corresponding torque $\Gamma_{DF} = |F_{DF}| R$ is given by:

$$\Gamma_{df} = \frac{4\pi GM_{BH}^2 \Sigma \ln \Lambda R}{\hbar v_{BH}^2}$$ (9)

where we have replaced the usual volume density $\rho_{gas}$ with the surface density $\Sigma$ of the Mestel disc, assuming for simplicity a constant disc scale height $h$ and simply that $\rho_{gas} = \Sigma/h$. The condition that orbital decay is efficient is equivalent to the condition that the inflow does not overtake the two holes as they decay, and can be written as $\Gamma_{df} > \Gamma_{visc}$. Combining (8) and (9) yields the following condition on the magnitude of $\alpha$:

$$\alpha < \frac{2GM_{BH} h^2 \ln \Lambda}{c_sh v_{BH}^2 v_c h^2}$$ (10)

where $\ln \Lambda$ is given by the expression in equation (3), i.e. $\ln \Lambda \sim 3$ for $b_{max}/b_{min} \sim 10$ (which reflects the ratio between disc size and distance of the holes from the centre where stalling was observed in the simulations of Mayer et al (2008)) and Mach numbers in the range 1–2.5 as typical for BHs embedded in circumnuclear discs on orbits with a range of eccentricities (see Chapon et al 2013). Furthermore, we adopt $v_{BH} \sim 200$ km s$^{-1}$, assuming that the circumnuclear disc is hosted in a relatively massive galaxy (note that the velocity will actually increase as the holes sink further in the potential well of the disc). We note that the surface density cancels out in the above equation but the assumption of a Mestel disc is still important since it imposes a simple law for the profile of $\Omega(r)$ given the constancy of $v_c$, so that we could replace $|d\Omega/dR| = |v_c/R|$ in equation (8). For $M_{BH} = 10^8 M_\odot$, $c_s = 60$ km s$^{-1}$, $h = 5$ pc, $R = 10$ pc, $v_c = V_{BH}$, we obtain $\alpha < 0.1$. 

---

Class. Quantum Grav. 30 (2013) 244008

L. Mayer
Let us now examine this result in context. For non-turbulent self-gravitating discs an effective viscosity \(\sim 0.1\) is often interpreted as a threshold between fragmentation and transport of angular momentum via gravitational torques, as long as one assumes that local linear perturbation theory applies to describe gravitational instability, so that one can adopt the local Toomre stability criterion for axisymmetric waves as a global empirical instability criterion for generic modes. In this case the condition \(\alpha < 0.1\) corresponds to \(Q < 1\), which is the standard threshold for fragmentation in a thin disc (Lodato and Rice 2004, 2005). Hence within this simple framework the fact that we find a critical \(\alpha\) close to such threshold has a simple and sensible physical meaning; it implies that DF will be effective as long as there is a stable background, which in this case is the circumnuclear disc.

If \(\alpha\) grows beyond the threshold, however, the inflow will overtake the orbital decay only temporarily since with \(Q < 1\) the disc is expected to fragment, turning gas into stars at a high rate, so that the decay might be restarted later as the disc becomes smooth again one most of the gas has been converted into stars (but this time the drag will be due to friction against a stellar background). However, 3D global disc simulations, performed at various scales and in various contexts, from protogalaxies to protoplanetary discs, show that discs can achieve a gravoturbulent state, with star formation and gas inflows co-existing, and an ‘effective’ \(\alpha \sim 1\) (e.g. Escala 2007, Agertz et al 2009). Such highly gravoturbulent state would seem to be problematic for the orbital decay of binary BHs based on equation (10). If anything, the DF timescale should lengthen considerably. The multi-phase mergers currently in progress (Roskar et al 2013) account such higher physical complexity and will shed light on these issues.

If we start by fixing the viscosity parameter, setting \(\alpha = 0.1\), we can rearrange equation (10) to express a condition on the minimum black hole mass needed for DF to be effective in presence of a strong inflow. This, for the same parameters just adopted, it turns out to be \(M_{\text{BH}} > 10^8M_\odot\), which is larger than the mass of the BH in the most massive spirals. As we said, since a realistic disc can generate a strong inflow in a gravoturbulent state with an even greater effective viscosity in such cases our same argument would imply that the inflow will overtake the binary for black holes even larger than \(10^8M_\odot\).

In order to gain further insight we can examine the case of very strong inflows, such as those postulated to be capable of forming massive BH seeds at high-z during galaxy mergers (Mayer et al 2010), as well as more moderate gas inflows, such as those necessary to supply the gas to nuclear starbursts in local ULIRGs and feed AGNs. We will derive the expected effective \(\alpha\) and compare with the condition expressed by equation (10). We can relate the inflow rate with the effective viscosity in the circumnuclear disc via the following equation, which expresses the maximum inflow rate in a steady-state condition since it essentially neglects any angular momentum or non-radial motion in the flow:

\[
\frac{dM}{dt} = 2\alpha c_s^3/G.
\] (11)

In the simulations of Mayer et al (2010), inflow rates as high as \(10^4M_\odot\) yr\(^{-1}\) have been measured. We can further adopt an effective sound speed in the range 60–100 km s\(^{-1}\), hence a ‘warm’ disc (which favours high inflows at comparatively smaller viscosities relative to the case of a ‘cold’ disc) in order to provide a conservative estimate of the corresponding value of \(\alpha\) for the inflow. With 100 km s\(^{-1}\) and \(\alpha = 0.1\) we obtain \(dM/dt \sim 100M_\odot\) yr\(^{-1}\). While this is a larger number, it is about two orders of magnitude lower than what measured in inner tens of pc of the circumnuclear discs that undergo a central collapse into a supermassive cloud, a possible precursor to a massive BH seed (Mayer et al 2010, Bonoli et al 2013). This implies effective viscosities of order unity at least. Adopting a weaker constraint, namely the minimum
inflow rate necessary to assemble a supermassive star of \( > 10^6 M_\odot \), which is \( > 1 M_\odot \) yr\(^{-1}\) (see Begelman 2010), one would still need \( \alpha > 0.1 \).

In any case, the conclusion is that the conditions required for massive BH seed formation via direct gas collapse are orthogonal to those necessary for efficient BH binary decay, which explains the findings of Mayer et al (2008). Likewise, circumnuclear discs have to enter a sub-critical phase with \( \alpha < 0.1 \), in which central runaway gas collapse cannot take place, in order to allow efficient coalescence of BH binaries. This might well be the natural condition some time after a massive BH seed has formed and the host galaxy merges with another galaxy containing another massive BH (in direct gas collapse models the inflow is a really quick event, operating on timescale less than \( 10^7 \) yr, see e.g. Mayer et al 2010). Of course, even in super-critical nuclei the binary BHs might still merge, albeit on a longer timescale, as the stellar background can still cause drag via DF and three-body encounters with stars. Finally, in presence of more moderate gas inflows such as those needed to feed ‘normal’ QSOs and AGNs (i.e. not the brightest QSOs at hi-z that might require feeding at a rate \( > 10M_\odot \) yr\(^{-1}\) once a massive BH is already in place there is no problem for the orbital decay. Indeed, inflow rates just below those at the critical viscosity threshold are still prominent, being large enough to provide enough gas supply to the accretion disc to grow already existing SMBHs at the typical rates suggested by observations of low redshift AGNs. These are typically occurring at a few per cent of the Eddington limit, with accretion rates in the range \( 10^{-2}–0.1M_\odot \) yr\(^{-1}\) (Raimundo et al 2012), which would correspond to \( \alpha < 10^{-3} \) for the lowest accretion rates, posing thus no problem for the orbital decay of black holes with masses even as small as \( M_{\text{BH}} \sim 4 \times 10^6 M_\odot \) (comparable to the mass of the black hole in the Milky Way), as it can be seen by combining equation (10) and (11).

A caveat in the above discussion is that it stems from results of numerical simulations in which only major mergers (equal mass) were considered. Once may wonder what would happen in the much more common case of unequal mass mergers. There are two important differences at least. First, the circurnuclear gas disc that forms in the minor merger case is going to be 1–2 orders of magnitude less massive, encompassing a mass only \( \sim 10 \) times larger than that of the embedded BHs rather than 100–1000 larger. This is inferred by comparing the gas mass accumulating within the central 50 pc in the 1:10 and 1:4 merger simulations of Callegari et al (2011) as opposed to that in the major merger simulations of Mayer et al (2007) and Chapon et al (2013), with the caveat that to date there is no minor merger simulation with BHs carried out a resolution of pc or better; second, the orbit circularizes at larger distance from the centre in minor mergers because the secondary BH is delivered at larger distance in the nucleus (Callegari et al 2011), hence one can expect the planetary migration-like phase to start earlier, possibly allowing more efficient decay even during inflows relative to the DF case, and therefore invalidating equation (9). Let us look at the first difference more in detail. Figure 5 of Guedes et al (2011) highlights the large difference in the mass of the circumnuclear region. As a result, the circumnuclear region is unlikely to develop very strong self-gravitating gas inflows as in the major mergers of Mayer et al (2008). If we apply equation (11) to the simulations of Callegari et al (2011), which show accretion rates \( < 0.1M_\odot \) yr\(^{-1}\) at all times for both the primary and secondary BH, and since these accretion rates actually reflect of the gas inflow rates at tens of pc scales in the nuclear region (see Callegari et al 2011 on how BH accretion is implemented) we would infer \( \alpha \ll 0.1 \). This clearly would be a drag-dominated regime in the context of the scenario discussed above. Therefore the dominant process will be the orbital decay of the BH binary in this case.

In general, we can also ask how likely is that gas will flow towards the centre of the disc and produce a central gas concentration when two massive BHs that can accrete such gas are also present. We can compare the gravitational pull of the disc potential relative to that of either
BH. For simplicity we can neglect the relative motion of the gas and the BHs and associate a fictitious sphere of influence to the circumnuclear disc, defined as $R_{SD} = GM_D(<r_{in})/v_s^2$. We can compare that with the sphere of influence $R_{BH} = GM_{BH}/(V_{BH} + v_s)^2$ of either BH, where $v_s$ is the thermal sound speed in the medium, $V_{BH}$ is the speed of either BH with respect to the gas disc, and $r_{in}$ is some radius in the disc well within the orbit of the BHs; $M_D$ is thus the mass contained well within the orbit of the black holes, such as in a clump at a few tens of parsecs from the centre within the disc. Clearly accretion onto either BH rather than to the centre of the disc itself will be important if, say, the radius of their spheres of influence is comparable to the effective disc size, hence it can intercept a significant fraction of the gas flowing towards the centre of the disc. This will happen roughly when $R_{BH} >\sim R_{SD}$, namely when $M_{BH} >\sim M_D$ since $V_{BH} \sim v_s$ in most situations (see section 4).

Now, in the minor mergers of Callegari et al (2011) and Guedes et al (2011) this condition is always satisfied for any massive BH since $M_D < 10^6 M_\odot$ already at 50 pc, whereas in the major mergers of Mayer et al (2010) we have $M_D \sim 10^8 M_\odot$ at the same distance even before the large scale inflow has assembled the supermassive cloud, which suggests that only BHs that are already very massive at the time of the merger, $M_{BH} > 10^8 M_\odot$, would be able to accrete at the expense of the inflow. Although this is a simple argument, we conclude that in minor mergers accretion should occur primarily on the BHs while the opposite would happen in major mergers unless there are already very massive BHs present in the system. In this last case, however, Bonoli et al (2013) have argued that the feedback from BHs weighing at least a million solar masses could stifle the gas inflow by creating a warmer, stable circumnuclear disc. Hence we conclude that a massive centralized inflow requires both a major merger and no massive BH to be already present.

7. Summary and conclusions

We have reviewed the various phases of the orbital decay of pairs of massive black holes (BHs), from galactic scale separations to sub-pc scale separation. We have neglected the presence of the stellar background in our analysis, and we have not considered the effect of BH accretion and its energy feedback. The concurrent effect of the stellar background and of black hole accretion has been explored only sporadically in the context of the orbital decay of massive BH binaries, although BH accretion in the Bondi–Hoyle spherical approximation and thermal feedback have been included in recent works (Callegari et al 2011, Van Wassenhoeve et al 2012), while hybrid calculations following the gaseous background at larger scales and three-body stellar encounters at smaller BH separations have been recently pioneered by Khan et al (2012). A detailed study of the effect of BH accretion and feedback in high-resolution simulations of circumnuclear discs, adopting more realistic accretion models rather than the Bondi–Hoyle approximation, is currently under way.

We have presented evidence that the decay of the secondary BH undergoes can undergo transition from a regime dominated by the local DF wake to one governed by more global, near resonant disc torques mediated by density waves as well as co-orbital torques, which resembles the regimes of Type-I and Type-III migration in planetary evolution. In particular, in our simulations co-orbital negative torques appear to be the strongest, and are responsible for the fast decay that brings the secondary BH from a separation of a few tens of pc to 0.1 pc in only a few Myr. This latter regime becomes dominant after the orbit of the secondary BH has circularized by dynamical friction (DF). The spiral density wave torques and the co-orbital torques can in principle continue to act even when the binary hardens to small sub-pc scale separations; in the first case the torquing density is far from the BH rather than being related to the local density while in the second case the asymmetry of the flow in the
vicinity of the secondary is the controlling factor. On the contrary, it has been shown that the local DF wake vanishes at small separations when the gas mass contained with the BH binary is comparable to the mass of the binary (Chapon et al 2013), similarly to what is found in stellar backgrounds (Milosavljevic and Merritt 2001). The migration regime is thus more promising as a channel to bring the secondary BH to the gravitational wave dominated regime in $<10^7-10^8$ yr, with a final phase of decay governed by the exchange of angular momentum with the circumbinary disc. Understanding better the conditions that activate such efficient orbital decay regime driven by disc and co-orbital torques will be of paramount importance in the future. It is likely that disc structure and thermodynamics are both crucial factors, controlling the excitation of density waves as well as how easily the gas can be gravitationally captured by the secondary and create a massive co-orbital envelope of torquing material. If gap opening intervenes below pc separations, in conditions that might be satisfied only by the most massive black holes, $>10^8 M_\odot$, the decay might slow down, with timescales even greater than $10^8$ yr to reach the gravitational wave dominated regime.

The results just recalled hold strictly in an homogeneous circumnuclear disc. In a clumpy, inhomogeneous disc, which is a more faithful model of the multi-phase medium with cold molecular clouds expected in real circumnuclear discs, scattering by gas clumps/clouds and other effects render the decay stochastic, in some cases slowing down the hardening to pc scales and below by more than an order of magnitude. In the latter case the binary will reach $0.1$ pc separations in $\sim 10^8$ yr (Fiacconi et al 2013, Roskar et al 2013).

All these timescales, either in a homogeneous or clumpy disc, are reasonably short—even in the worst case the two black holes should merge on a timescale comfortably below $10^9$ yr. Instead, the real bottleneck might be at larger scales, precisely larger than the size of circumnuclear disc, about a few hundred parsecs. Indeed, while for major mergers there are no particular issues, minor mergers simulations show that, depending on the orbit and combined effect of gas cooling, star formation, feedback processes etc., the two black holes might reach the nucleus of the remnant over a time from as short as a few $10^8$ yr to as long as a Gyrs, or even not reach it at all, becoming wandering black holes (Callegari et al 2009, 2011). Since minor mergers are the preferred mode of galaxy assembly in a CDM universe, future work will have to explore a much larger parameter space with simulations, with the aim at building statistical semi-analytical models that can describe properly the large scale pairing of massive BH as a function of the various relevant parameters. In fact, at the moment assessing the probability distribution of the BH pairing timescales at scales $>100$ pc might be the most pressing issue in order to evaluate how likely is that black holes merge when their host galaxies merge.

Finally, we have begun to unveil an intriguing interplay between efficient orbital decay and efficient formation of massive BH seeds by direct gas collapse, showing that the conditions demanded by one process essentially exclude the other one. This is because prominent gas inflows evacuate the background that is responsible for the drag of the holes. Implicitly, this line of reasoning assumes that the stellar background does not provide an efficient drag. If true, the existence of these two mutually exclusive regimes points to distinct phases of massive BH evolution, an early one in which massive BH seed formation is very likely owing to prominent gas inflows in hi-z mergers among the most massive galaxies at $z \sim 8-10$ or efficient gas accretion in protogalaxies in very high sigma peaks, such as the first halos that can cool via atomic hydrogen, and a later one in which central gas inflows fade away but BH binaries can merge more efficiently at the centre of galactic nuclei. It is tempting to speculate on the effect that this would have on the gravitational wave event rates from massive BH mergers as a function of redshift. For example, at early epochs BH binaries might be held back in their race to the centre due to the inflows overtaking them, at late epochs mergers of massive
BHs become rare because the merger rate of galaxies and halos as well as the gas content of galaxies decrease significantly, while at intermediate epochs BH binaries could find ideal conditions to sink efficiently to the centre as inflows are moderate enough but merger rates and gas content are still high enough. This and other intriguing aspects of the co-evolution of galaxies and massive BH binaries will have to be investigated with semi-analytical models following simultaneously BH seed formation via direct gas collapse and the processes that lead to BH mergers in realistic backgrounds of both gas and stars.

Acknowledgments

We thank all the people that have had a long term collaboration with the author on the topic of co-evolution of galaxies and SMBHs, and who contributed to many of the results discussed in this review; Simone Callegari, Monica Colpi, Stelios Kazantzidis, Massimo Dotti, Davide Fiacconi, Damien Chapon, Andres Escala, Piero Maduakor, Rok Roskar, Romain Teyssier, James Wadsley, Jillian Bellovary and Tom Quinn. Special thanks go to Davide Fiacconi for having produced some of the figures presented in this paper, and to one of the anonymous referees for very helpful comments and questions that improved significantly the original paper. We also thank the Kavli Institute for Theoretical Physics at UC Santa Barbara for hospitality during completion of this work while the program ‘A Universe of Black Holes’ was taking place. The program generated stimulating thoughts and discussions that have been crucial to develop this work.

References

Agertz O, Lake G, Teyssier R, Moore B, Mayer L and Romeo A B 2009 Mon. Not. R. Astron. Soc. 392 294
Amaro-Seoane P et al 2013 European Space Agency on 24th May, 2013 for the L2/L3 selection of ESA’s Cosmic Vision program, submitted (arXiv:1305.5720)
Armitage P and Natarjan P 2005 Astrophys. J. 634 921
Baker J G, Centrella J, Choi D I, Koppitz M and Van Meter J 2006 Phys. Rev. Lett. 96 111102
Barnes J 2002 Mon. Not. R. Astron. Soc. 333 481
Begelman M, Volonteri M and Rees M J 2006 Mon. Not. R. Astron. Soc. 370 289
Begelman M C, Blandford R D and Rees M J 1980 Nature 287 307
Begelman M C 2010 Mon. Not. R. Astron. Soc. 402 673
Bekenstein J D and Maoz E 1992 Astrophys. J. 390 79
Berczik P, Merritt D, Spurzem R and Bischof H P 2006 Astrophys. J. 642 L21
Bitsch A, Boley A and Kley W 2013 Astron. Astrophys. (submitted) (arXiv:1301.1516)
Bogdanovic T, Smith B D, Sigurdsson S and Eracleous M 2008 Astrophys. J. Suppl. 174 455
Bonoli S, Mayer L and Callegari S 2013 Mon. Not. R. Astron. Soc. at press
Callegari S, Mayer L, Kazantzidis S, Colpi M, Governato F, Quinn T and Wadsley J 2009 Astrophys. J. 696 L89
Callegari S et al 2011 Astrophys. J. 729 85
Chandrasekhar S 1943 Astrophys. J. 97 255
Chapon D, Mayer L and Teyssier R 2013 Mon. Not. R. Astron. Soc. 429 3114
Choi J H, Shlosman I and Begelman M 2013 Astrophys. J. 744 149
Colpi M, Mayer L and Governato F 1999 Astrophys. J. 525 720
Colpi M and Dotti M 2009 Advanced Science Letters (Special Issue on Computational Astrophysics) arXiv:0906.4339C
Cuadra J P, Armitage P, Alexander R D and Begelman M C 2009 Mon. Not. R. Astron. Soc. 393 1423
del Valle L and Escala A 2012 Astrophys. J. 761 9
Di Matteo T, Colberg J, Springel V, Hernquist L and Sijacki D 2008 Astrophys. J. 676 33
Dotti M, Colpi M and Haardt F 2006 Mon. Not. R. Astron. Soc. 367 103
Dotti M, Colpi M, Haardt F and Mayer L 2007 Mon. Not. R. Astron. Soc. 379 956
Dotti M, Ruszkowski M, Paredi L, Colpi M, Volonteri M and Haardt F 2009 Mon. Not. R. Astron. Soc. 396 1640
Downes D and Solomon P M J 1998 Astrophys. J. 507 615
Escala A 2007 Astron. J. 167 1264
Escala A, Larson R B, Coppi P S and Maradones D 2004 Astrophys. J. 607 765
Escala A, Larson R B, Coppi P S and Maradones D 2005 Astrophys. J. 630 152
Farris B D, Duffell P, MacFadyen A I and Haiman Z 2013 Astrophys. J. submitted
Farris B D, Liu Y T and Shapiro S L 2011 Phys. Rev. D 84 024024
Feldmann R et al 2010 Astrophys. J. 709 218
Ferrarese L and Merritt D 2000 Astrophys. J. 539 L9
Fiacconi D, Mayer L, Roskar R and Colpi M 2013 Astrophys. J. 777 L14
Genzel R et al 2011 Astrophys. J. 733 101
Goldreich P and Tremaine S 1984 Astrophys. J. 336 263
Greene J E et al 2005 Nature 437 629
Goldreich P and Tremaine S 1980 Astrophys. J. 241 425
Guedes J, Madau P, Mayer L and Callegari S 2013 Astrophys. J. 777 L14
Hayasaki K 2009 Publ. Astron. Soc. Japan 61 66
Khan F M et al 2012 Mon. Not. R. Astron. Soc. 423 1675
Khan F M et al 2012 Astron. J. 756 30
Lin D N C and Papaloizou J C B 1999 Mon. Not. R. Astron. Soc. 315 823
Mayer L, Kazantzidis S, Escala A and Callegari S 2010 Mon. Not. R. Astron. Soc. 415 L35
Mayer L, Kazantzidis S, Madau P, Colpi M, Quinn T and Wadsley J 2007 Science 316 1874
Medling A, Vivian U, Guedes J, Max C E and Mayer L 2013 Astrophys. J. submitted
Menou K and Quataert E 2001 Astrophys. J. 552 204
Milosavljevic M and Merritt D 2001 Astrophys. J. 563 34
Nixon C J 2012 Mon. Not. R. Astron. Soc. 423 529
Papaloizou J C B and Lai Y K M 2004 Mon. Not. R. Astron. Soc. 351 L630
Papaloizou J C B and Lin D N C 1995 Mon. Not. R. Astron. Soc. 276 999
Papaloizou J C B, Nelson R P, Kley W, Masset F S and Artymowicz P 2007 Protostars and Planets V vol 951 (Tucson, AZ: U. Arizona Press) p 655
Peplinski A, Artymowicz P and Mellema G 2008 Mon. Not. R. Astron. Soc. 386 179
Raimundo S, Fabian A C, Vasudevan R V, Gandhi P and Wu J 2012 Mon. Not. R. Astron. Soc. 419 2529
Regan J and Haehnhel M 2009 Mon. Not. R. Astron. Soc. 396 343
Rezzolla L 2009 Class. Quantum Grav. 26 094023
Roedig C and Sesana A 2013 Mon. Not. R. Astron. Soc. submitted
Rodig C, Sesana A, Dotti M, Cuadra J, Amaro-Seoane P and Haardt F 2012 Astron. Astrophys. 545 27
Shi J M, Krolik J H and Hawley J E 2012 Astron. J. 749 26
Taffoni G, Mayer L, Colpi M and Governato F 2003 Mon. Not. R. Astron. Soc. 341 434
Tanaka H, Takeuchi T and Ward W R 2002 Astrophys. J. 565 1257
Van Wassenhove S et al 2012 Astrophys. J. 748 L7
Vecchio A 2004 Phys. Rev. D 70 042001
Volonteri M, Haardt F and Madau P 2003 Astrophys. J. 582 559
Wada K 2011 Astrophys. J. 559 L41
Weinberg M D 1989 Mon. Not. R. Astron. Soc. 239 549
White S D M 1978 Mon. Not. R. Astron. Soc. 184 185