Mon. Not. R. Astron. Soc. **377,** 957–976 (2007) doi:10.1111/j.1365-2966.2007.11694.x

Dynamics of triple black hole systems in hierarchically merging massive galaxies

Loren Hoffman[1⋆] and Abraham Loeb[2⋆]

[1]*Physics Department, Harvard University, 17 Oxford St., Cambridge, MA 02138, USA*
[2]*Astronomy Department, Harvard University, 60 Garden Street, Cambridge, MA 02138, USA*

Accepted 2007 March 1. Received 2007 February 28; in original form 2006 December 18

ABSTRACT

Galaxies with stellar bulges are generically observed to host supermassive black holes (SMBHs). The hierarchical merging of galaxies should therefore lead to the formation of SMBH binaries. Merging of old massive galaxies with little gas promotes the formation of low-density nuclei where SMBH binaries are expected to survive over long times. If the binary lifetime exceeds the typical time between mergers, then triple black hole (BH) systems may form. We study the statistics of close triple SMBH encounters in galactic nuclei by computing a series of three-body orbits with physically motivated initial conditions appropriate for giant elliptical galaxies. Our simulations include a smooth background potential consisting of a stellar bulge plus a dark matter halo, drag forces due to gravitational radiation and dynamical friction on the stars and dark matter, and a simple model of the time evolution of the inner density profile under heating and mass ejection by the SMBHs. We find that the binary pair coalesces as a result of repeated close encounters in ∼85 per cent of our runs, and in ∼15 per cent of cases a new eccentric binary forms from the third SMBH and binary remnant and coalesces during the run time. In about 40 per cent of the runs the lightest BH is left wandering through the galactic halo or escapes the galaxy altogether, but escape of all three SMBHs is exceedingly rare. The triple systems typically scour out cores with mass deficits ∼one–two times their total mass, which can help to account for the large cores observed in some massive elliptical galaxies, such as M87. The high coalescence rate, prevalence of very high-eccentricity orbits, and gravitational radiation 'spikes' during close encounters in our runs, may provide interesting signals for the future Laser Interferometer Space Antenna (LISA).

Key words: black hole physics – methods: numerical – galaxies: elliptical and lenticular, cD – galaxies: interactions – galaxies: nuclei – cosmology: theory.

1 INTRODUCTION

In the favoured cold dark matter (CDM) cosmology, present-day galaxies were assembled hierarchically from smaller building blocks at earlier cosmic times. Since all nearby galaxies with stellar spheroids are observed to host nuclear supermassive black holes (SMBHs) (Kormendy & Gebhardt 2001), hierarchical merging leads inevitably to the formation of SMBH binaries (Begelman et al. 1980). If the binary lifetime exceeds the typical time between mergers, then some galactic nuclei should contain systems of three or more SMBHs. These systems are particularly interesting as they often lead to the ejection of one of the black holes (BHs) at a speed comparable to the galactic escape velocity (Hoffman & Loeb 2006). In massive elliptical galaxies the typical speeds are $\sim 10^3$ km s^{-1}, far greater than attainable through gravitational radiation recoil for non-spinning BHs (Favata, Hughes & Holz 2004; Blanchet, Qusailah & Will 2005; Centrella 2006).

Spatially resolved pairs of nuclei have been observed in a few active galaxies. The most famous example is NGC 6240, an ultraluminous infrared galaxy (ULIRG) in which two distinct active galactic nuclei (AGN) are clearly seen in hard X-rays at a projected separation of ∼1 kpc (Komossa et al. 2003). Maoz et al. (1995, 2005) observed a variable ultraviolet (UV) source, possibly a second active nucleus, at a projected separation of ∼60 pc from the primary nucleus in the spiral galaxy NGC 4736, which shows signs of a recent merger. Rodriguez et al. (2006) have detected what is thought to be an SMBH binary at a projected separation of just 7.3 pc in the radio galaxy 0402+379, through multifrequency radio observations using the Very Long Baseline Array (VLBA). We begin by discussing the theory of how such systems evolve, and the conditions under which they might acquire a third BH.

⋆E-mail: lhoffman@cfa.harvard.edu (LH); aloeb@cfa.harvard.edu (AL)

1.1 Black hole binaries

When two galaxies merge, their dense nuclei sink to the centre of the merger product by dynamical friction. As the nuclei spiral in, tidal forces gradually strip the two SMBHs of their surrounding stars and dark matter. In mergers between galaxies of comparable mass, the BHs are able to come together and form a bound SMBH binary on a time-scale of order 10^9 yr. The binary continues to harden by dynamical friction until it reaches a separation of order

$$a_{\rm hard} \equiv \frac{G\mu}{4\sigma^2} \approx 0.80 \frac{4q}{(1+q)^2} \left(\frac{m_{\rm bin}}{10^8\,{\rm M}_\odot} \right)^{1/2} {\rm pc}, \quad (1)$$

known as the 'hardening radius' (e.g. Quinlan 1996). Here $\mu = m_1 m_2/(m_1 + m_2)$ is the reduced mass of the two BHs with masses m_1 and m_2, σ is the velocity dispersion of the stars beyond the binary's sphere of gravitational influence, q is the binary mass ratio $m_2/m_1 \leqslant 1$ and $m_{\rm bin} = m_1 + m_2$ is the total mass of the binary. For smaller separations the binary looks like a point mass to the distant stars contributing to dynamical friction, but close stellar encounters preferentially harden the binary and so dominate further energy loss. Only stars on nearly radial orbits, with periapsis distances of order the binary separation, can extract energy from ('harden') the binary in this stage. These stars undergo strong three-body interactions with the binary and escape its vicinity with speeds comparable to the BHs' orbital speed. In the low-density nuclei of large elliptical galaxies, the total mass in stars on such 'loss cone' orbits is small compared to the mass of the binary. Furthermore, the two-body stellar relaxation time is long compared to a Hubble time, so once the stars initially on loss cone orbits are cleared out, the loss cone remains empty (Frank & Rees 1976; Lightman & Shapiro 1977; Cohn & Kulsrud 1978). Since the binary must eject of order its own mass per e folding in its semimajor axis, the system stops hardening around $a_{\rm hard}$ unless some other mechanism causes sufficient mass flux through the binary.

If the binary reaches a separation around

$$a_{\rm gw} = 4.5 \times 10^{-2} \left(\frac{m_{\rm bin}}{10^8\,{\rm M}_\odot} \right)^{3/4} \left[\frac{4q}{(1+q)^2} \right]^{1/4} \times \left(\frac{\tau_{\rm gw}}{10^{10}\,{\rm yr}} \right)^{1/4} f^{-1/4}(e)\,{\rm pc}, \quad (2)$$

where e is the orbital eccentricity of the binary and $f(e) = (1 - e^2)^{7/2}/(1 + 73e^2/24 + 37e^4/96)$, then it can coalesce on a time-scale $\tau_{\rm gw}$ through gravitational radiation (Begelman et al. 1980). To get from $a_{\rm hard}$ to $a_{\rm gw}$ it must bridge a gap

$$\frac{a_{\rm hard}}{a_{\rm gw}} \approx 17 \left(\frac{m_{\rm bin}}{10^8\,{\rm M}_\odot} \right)^{-1/4} \left[\frac{4q}{(1+q)^2} \right]^{3/4} \quad (3)$$

by some mechanism other than stellar-dynamical friction or gravitational radiation. The question of whether and how it crosses this gap has become known as the 'final parsec problem' (Merritt & Milosavljević 2005).

In many galaxies there probably are alternative mechanisms for crossing the gap. When gas-rich galaxies merge, tidal torques channel large amounts of gas into the central ~100 pc (Byrd, Sundelius & Valtonen 1987; Hernquist 1989). The gas may lose energy through radiation and angular momentum through viscous torques, and is therefore not subject to a loss cone problem. Using smoothed particle hydrodynamics simulations, Escala et al. (2004, 2005) compute a merger time of order 10^7 yr in an environment typical of the central regions of ULIRGs, which are thought to be gas-rich galaxies caught in the act of merging (Sanders et al. 1988). The nuclei of galaxies are also observed to contain numerous massive perturbers (MPs) such as star clusters, molecular clouds and possibly intermediate-mass black holes (IMBHs). These objects scatter stars into the loss cone much more efficiently than other stellar mass objects, since the relaxation rate scales as the perturber mass for a fixed mass density of perturbers. Perets, Hopman & Alexander (2007) extended the Fokker–Planck loss cone formalism (Frank & Rees 1976; Lightman & Shapiro 1977; Cohn & Kulsrud 1978) to accommodate a spectrum of perturber masses and account for relaxation by rare close encounters with MPs. They show that the population of known MPs in the nucleus of the Milky Way is sufficient to bring a 4 $\times$ 10^6 $\rm M_\odot$ BH binary to $a_{\rm gw}$ in ~6 $\times$ 10^8 yr, and it is reasonable to expect similar perturber populations in other star-forming spiral galaxies.

The final parsec problem is often mentioned as a caveat when predicting the SMBH coalescence signal in low-frequency gravitational wave detectors such as the upcoming Laser Interferometer Space Antenna (LISA). However, the LISA event rate is expected to be dominated by small galaxies at high redshift (Wyithe & Loeb 2003a; Rhook & Wyithe 2005; Sesana et al. 2005), where the gas content and central densities tend to be high and the relaxation times short. For this reason the stalling problem is probably not a significant concern for the LISA SMBH coalescence signal. On the other hand BH ejections by gravitational radiation recoil (Haiman 2004; Merritt et al. 2004) may play an important role in the high-redshift coalescence rate. The long-term survival of SMBH binaries is likewise unlikely in the gas-rich cores of quasars and ULIRGs.

However, none of the gap-crossing mechanisms discussed so far are likely to reduce the coalescence time below a Hubble time in mergers between giant, gas-poor elliptical galaxies. Merritt & Poon (2004) show that a significant fraction of stars on 'centrophilic' orbits in a triaxial potential can greatly increase the mass flux into the loss cone. Some non-axisymmetric potentials can also excite bar instabilities that cause rapid mass flow through the binary and efficient coalescence (Berczik et al. 2006). However, a central SMBH can disrupt box orbits and induce axisymmetry in the inner regions of a triaxial galaxy (Merritt & Quinlan 1998; Holley-Bockelmann et al. 2002), and it is uncertain how often these geometry-specific mechanisms bring the coalescence time below a Hubble time.

One can naively assess the likelihood of coalescence by considering the 'full' and 'empty' loss cone hardening times, $\tau_{\rm full}$ and $\tau_{\rm empty}$, in the nuclei of various galaxies assuming a spherical and isotropic distribution function. $\tau_{\rm full}$ is the hardening time assuming every star kicked out of the loss cone is instantly replaced, while $\tau_{\rm empty}$ is the time assuming stellar two-body relaxation to be the only replenishing mechanism. In small, dense galaxies $\tau_{\rm full} \sim 10^{5-6}$ yr and $\tau_{\rm empty} \sim 10^{9-10}$ yr while in the lowest density cores of giant ellipticals and cD galaxies $\tau_{\rm full} \sim 10^8$ yr and $\tau_{\rm empty} \sim 10^{14}$ yr (Yu 2002). While the empty loss cone rate is difficult to believe in any galaxy given at least some clustering on scales larger than 1 $\rm M_\odot$, it also seems difficult to approach the full loss cone rate if there is no gas around and no strong radial bias in the stellar distribution. From this point of view the stalling of binaries seems unlikely in small galaxies but probable in low-density, gas-poor ellipticals. If some binaries do survive for around a Hubble time, then the hierarchical build-up of galaxies will inevitably place three or more SMBHs in some merging systems.

1.2 Merger-induced binary evolution before three-body interactions: back-of-the-envelope calculations

An inspiralling satellite affects the evolution of a binary SMBH even long before it sinks to the centre, by perturbing the large-scale potential and scattering stars into the loss cone. We may estimate the extent of this effect as a function of satellite mass and distance from the centre of the host galaxy using a rough but simple argument due to Roos (1981). The change in velocity necessary to deflect a star at radius q into the loss cone is $\Delta V \sim h_{\rm lc}/q$, where $h_{\rm lc} \sim \sigma\sqrt{r_{\rm inf} r_{\rm bin}}$ is the characteristic specific angular momentum of stars on loss cone orbits (Frank & Rees 1976), $r_{\rm inf} = Gm_{\rm bin}/\sigma^2$ is the SMBHs' radius of influence and $r_{\rm bin}$ is the binary separation. The dynamical time at this radius is $t_{\rm dyn} \sim q/\sigma$, so the acceleration required to scatter a star into the loss cone is roughly $a_{\rm lc}(q) \sim \Delta V/t_{\rm dyn} \sim \sigma^2\sqrt{r_{\rm inf} r_{\rm bin}}/q^2$. Equating this with the tidal acceleration caused by the satellite, $a_{\rm tid} = 2GM_{\rm sat}(r)q/r^3$, where r is the satellite's radius, yields $q^3 = \sigma^2\sqrt{r_{\rm inf} r_{\rm bin}}r^3/2GM_{\rm sat}(r)$, or with $r_{\rm bin} = a_{\rm hard} = G\mu/4\sigma^2$,

$$q = \left[\frac{\sqrt{m_{\rm bin}\mu}}{4M_{\rm sat}(r)}\right]^{1/3} r. \tag{4}$$

The r dependence of $M_{\rm sat}$ reflects the tidal stripping of the satellite as it spirals inward. Equation (4) defines a critical radius q, outside of which the satellite can deflect stars into (and out of) the loss cone in one dynamical time. The mass flux through the binary induced by the satellite is then approximately

$$\frac{{\rm d}M_{\rm stars}}{{\rm d}t}(q) = 2\pi\rho(q)q^2\sigma\theta_{\rm lc}^2, \tag{5}$$

where $\rho(q)$ is the density of the host galaxy at radius q and $\theta_{\rm lc}^2 \approx r_{\rm inf}\, r_{\rm bin}/r^2$ is the geometrical factor accounting for the fraction of stars on loss cone orbits as a function of radius r, assuming an isotropic distribution function (Frank & Rees 1976). For a fixed satellite mass and distance, we can then define a 'binary feeding' time-scale by

$$\tau_{\rm feed} = \frac{m_{\rm bin}}{{\rm d}M_{\rm stars}/{\rm d}t} = \frac{m_{\rm bin}}{2\pi\rho(q)q^2\sigma\theta_{\rm lc}^2(q)}. \tag{6}$$

To determine whether the scattering of stars into the loss cone by the satellite is sufficient to harden the binary enough to prevent a close three-body encounter before the intruder arrives at the galactic centre, we must compare $\tau_{\rm feed}$ with the time-scale on which the satellite spirals in by dynamical friction. In the approximation of slow inspiral we may write the dynamical friction time-scale as $\tau_{\rm df} \equiv |r/\dot{r}| \approx |v/({\rm d}v/{\rm d}t)_{\rm df}|$. Substituting $({\rm d}v/{\rm d}t)_{\rm df}$ from Chandrasekhar's formula (equation 28 in Section 2.4.1; Chandrasekhar 1943) yields

$$\tau_{\rm df}(r) = \frac{v_{\rm sat}^3(r)}{4\pi G^2\rho(r)M_{\rm sat}(r)\left[{\rm erf}(X) - (2X/\sqrt{\pi}){\rm e}^{-X^2}\right]}, \tag{7}$$

where $X \equiv v_{\rm sat}/\sqrt{2}\sigma$. $v_{\rm sat}(r)$ in equation (7) is computed from $v_{\rm sat}(r) = \sqrt{GM_{\rm host}(r)/r}$, where $M_{\rm host}(r) = M_{\rm stars}(r) + M_{\rm halo}(r) + m_{\rm bin}$ is the mass of the host galaxy enclosed within radius r. $M_{\rm sat}(r)$ is the satellite mass contained within the tidal truncation radius obtained from a simple point mass approximation, $r_{\rm tid} = [M_{\rm sat}/M_{\rm host}]^{1/3}r$ (this slightly underestimates $r_{\rm tid}$ as the satellite approaches the centre of the host). In Fig. 1 we plot $\tau_{\rm df}$ and $\tau_{\rm feed}$ as a function of r for a satellite with one-third the stellar mass of the host, which contains a binary with $(m_1, m_2) = (1.2, 3.7) \times 10^8\,{\rm M}_\odot$. Both host and satellite are modelled as Hernquist profiles (Hernquist 1990), with their masses and effective radii set by observed scaling relations. The details of the galactic model are described further in Sections 2.1 and 2.2.

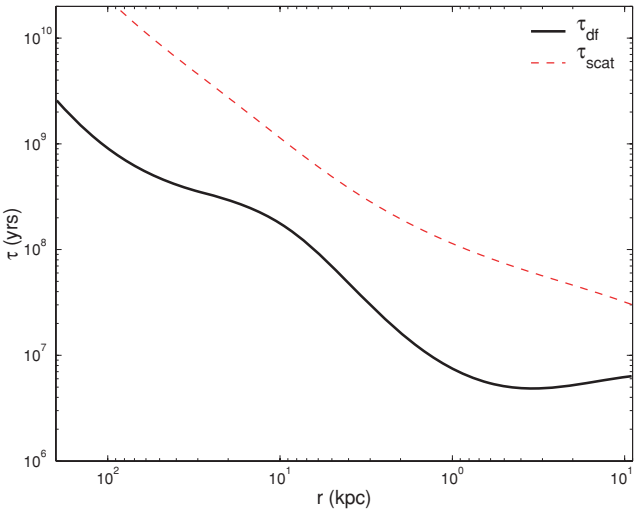

Figure 1. Comparison of the 'feeding time-scale', $\tau_{\rm feed}$, on which an inspiralling satellite scatters mass into the loss cone of an SMBH binary, with the dynamical friction time-scale, $\tau_{\rm df}$, on which the satellite spirals in. Upper (red dashed) line: $\tau_{\rm feed}$ computed from equation (6). Lower curve: $\tau_{\rm df}$ computed from equation (7). Both time-scales are plotted as a function of the satellite's distance from the centre of the host galaxy. For this plot we chose a binary mass of $4.5 \times 10^8\,{\rm M}_\odot$ and merger mass ratio of 3:1 in the stars. The galactic model is discussed in the text.

Since $\tau_{\rm feed}$ remains about an order of magnitude above $\tau_{\rm df}$ throughout the inspiral, this simple calculation makes it plausible that the binary survives the merger process and undergoes close triple interactions with the infalling SMBH. The tidal approximation (as well as our treatment of dynamical friction) breaks down as the satellite approaches $r_{\rm inf}$, so the plot is cut off at a separation of $\sim$100 pc, when the satellite still has $\sim$4 e-foldings to go to reach $a_{\rm hard}$. However, this final stage of the inspiral is found to proceed very rapidly in N-body simulations (Quinlan & Hernquist 1997; Milosavljevic & Merritt 2001; Merritt 2006). The merger's effect on the binary may be dominated by violent relaxation or collective effects such as a bar instability (Berczik et al. 2006), in which case our two-body approach does not capture its essence. The evolution of the core distribution function under the influence of a major merger is an intriguing open problem for simulators.

After the third BH becomes bound to the binary (but still before the onset of close three-body interactions) another hardening mechanism may become important. If the angle of inclination i of the outer binary [formed by the intruder and the inner binary centre-of-mass (COM)] exceeds a critical angle $\theta_{\rm crit} \approx 39^\circ$, then the quadrupolar perturbation from the intruder induces eccentricity oscillations through a maximum (Kozai 1962),

$$e_{\rm max} \approx \sqrt{1 - \frac{5}{3}\cos^2 i}. \tag{8}$$

Since the gravitational radiation rate increases sharply toward high eccentricities, these 'Kozai oscillations' can greatly enhance the radiation, possibly causing the binary to coalesce before it can undergo strong three-body interactions with the intruder (Blaes, Lee & Socrates 2002). General relativistic precession can destroy the Kozai resonance (e.g. Holman, Touma & Tremaine 1997), but Blaes et al. (2002) find that this does not happen for

$$\frac{a_{\rm out}}{a_{\rm in}} \lesssim 43\left[2q_{\rm out}\left(\frac{a_{\rm in}}{1\,{\rm pc}}\right)\left(\frac{10^8\,{\rm M}_\odot}{m_{\rm bin}}\right)\right]^{1/3}\sqrt{\frac{1-e_{\rm in}^2}{1-e_{\rm out}^2}}, \tag{9}$$

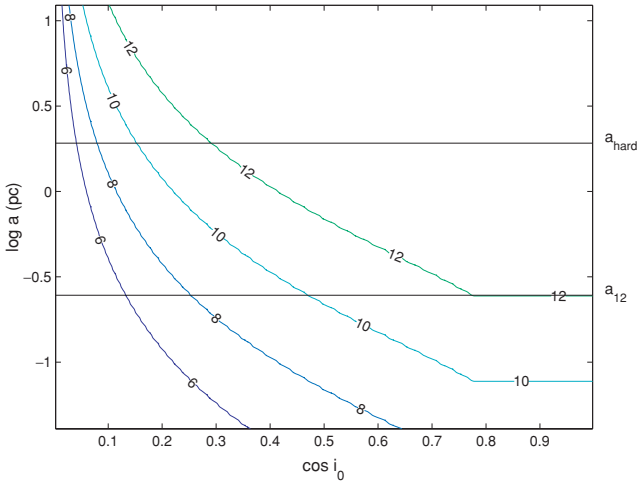

Figure 2. Upper limit on the importance of Kozai oscillations in enhancing gravitational radiation by the inner binary ($m_1 = m_2 = 3 \times 10^8\,\mathrm{M}_\odot$). The cosine of the initial inclination angle is plotted on the horizontal axis, and the inner binary semimajor axis is plotted on the vertical axis. Contours are plotted for gravitational radiation time-scales of $\tau_{\rm gw} = 10^{12}$, 10^{10}, 10^8 and 10^6 yr, if the binary were to stay at maximum eccentricity throughout the whole oscillation cycle. The horizontal lines indicate the hardening radius and the separation such that a circular binary would coalesce on a 10^{12}-yr time-scale.

where $a_{\rm in}$ and $a_{\rm out}$ are the semimajor axes of the inner and outer binaries, $q_{\rm out}$ is the outer binary mass ratio and $e_{\rm in}$ and $e_{\rm out}$ are the inner and outer eccentricities. This leaves a window of about a factor of 10 in $a_{\rm out}/a_{\rm in}$ in which the Kozai mechanism can operate before unstable three-body interactions begin.

The actual enhancement of the gravitational radiation rate of course depends on the amount of time spent at high eccentricity, but one may place an upper limit on the importance of Kozai oscillations by computing the radiation time-scale if the inner binary spends all of its time at $e_{\rm max}$. The orbit-averaged power radiated by gravitational radiation is given by

$$\left|\frac{{\rm d}E}{{\rm d}t}\right|_{\rm gw} = \frac{32G^4m_1^2m_2^2(m_1+m_2)}{5c^5a^5}\frac{1+(73/24)e^2+(37/96)e^4}{(1-e^2)^{7/2}} \tag{10}$$

(Peters 1964), where a is the semimajor axis and e is the eccentricity. In Fig. 2 we plot contours of the gravitational radiation time $\tau_{\rm gw} = |E/({\rm d}E/{\rm d}t)_{\rm gw}|$ in the a–i plane by putting $e_{\rm max}$ into equation (10), for an equal-mass $6 \times 10^8\,\mathrm{M}_\odot$ binary. This may seem like a gross overestimate of the gravitational radiation rate, especially since the shape of the Kozai oscillations is in fact such that the binary spends more time near $e_{\rm min}$ than near $e_{\rm max}$. However, since $\tau_{\rm gw}$ is so strongly dominated by periapsis passages at $e \approx e_{\rm max}$, the shift in the contours for a realistic high-e duty cycle is only modest. See Blaes et al. (2002) for comparison with a detailed study of radiation enhancement by Kozai oscillations in binaries with initial $\tau_{\rm gw} \sim 10^{12}$ yr. For a binary at $a_{\rm hard}$, Kozai oscillations can induce coalescence within 10^{10} yr in $\lesssim$20 per cent of cases assuming $\cos i$ is uniformly distributed. In the remainder of cases the inner binary may survive until the outer binary shrinks to the point of unstable three-body interactions.

1.3 Close three-body encounters

If the intruder comes close enough before it causes sufficient hardening of the (inner) binary, then a strong three-body encounter takes place. Strong encounters are characterized by a significant transfer of energy between the binary's internal degrees of freedom and the COM motion of the binary and third body. When the intruder is slow relative to the binary's orbital speed $v_{\rm bin}$, energy typically flows from the inner binary to the outer components, so that the binary is more strongly bound after the encounter. This is one manifestation of the negative specific heat characteristic of gravitationally bound systems. The encounter ends in the escape of one of the three bodies, usually the lightest, from the system at a speed comparable to $v_{\rm bin}$.

When the lightest body m_3 escapes, momentum conservation requires that the binary COM recoil in the opposite direction with a speed smaller by a factor $m_3/(m_1+m_2)$. It is instructive to compare the expected ejection velocities of the binary and m_3 with the typical galactic escape velocity. For a circular binary with $m_1 = m_2 = m_{\rm bin}/2$, the binding energy at the hardening radius is $E_{\rm B,hard} = Gm_{\rm bin}^2/8a_{\rm hard} \approx 6.8 \times 10^{55}[m_{\rm bin}/(10^8\,\mathrm{M}_\odot)]^{3/2}$ erg. The binding energy at the radius where $\tau_{\rm gw} = 10^8$ yr is $E_{\rm B,gw} = Gm_{\rm bin}^2/8a_{\rm gw} \approx 6.2 \times 10^{57}[m_{\rm bin}/(10^8\,\mathrm{M}_\odot)]^{5/4}$ erg. The mean energy ΔE harvested from the binary in close encounters with slow intruders is about $0.4E_{\rm B}$, though the median ΔE is somewhat lower (Hills & Fullerton 1980). Energy conservation implies that the escaper leaves the system with kinetic energy ${\rm KE}_{\rm sing} = \Delta E/[1 + m_3/(m_1+m_2)]$ while the binary leaves with ${\rm KE}_{\rm bin,cm} = \Delta E/[1 + (m_1+m_2)/m_3]$ in the system COM frame. For an equal-mass binary with $m_{\rm bin} = 5 \times 10^8\,\mathrm{M}_\odot$, this gives ejection velocities of $v_{\rm sing} \sim 290\,{\rm km\,s^{-1}}$ and $v_{\rm bin} \sim 140\,{\rm km\,s^{-1}}$ for the binary at $a_{\rm hard}$, and $v_{\rm sing} \sim 4000\,{\rm km\,s^{-1}}$ and $v_{\rm bin} \sim 2000\,{\rm km\,s^{-1}}$ for the binary at $a_{\rm gw}$.

Any non-zero eccentricity of the binary will increase the semimajor axis corresponding to a fixed $\tau_{\rm gw}$, lowering the ejection velocities for the binary at $a_{\rm gw}$. Also any deviation from equal masses will result in a smaller fraction of the extracted energy being apportioned to the binary and a smaller binary recoil velocity. The typical escape velocity for galaxies hosting $5 \times 10^8\,\mathrm{M}_\odot$ BHs is around $1500\,{\rm km\,s^{-1}}$, accounting for both the stars and the dark matter. From these numbers, it appears that single escapes will be fairly common as repeated encounters harden the binary to $\sim a_{\rm gw}$. However, accounting for realistic mass ratios and eccentricities (the first three-body encounter tends to thermalize the eccentricity even if it starts off circular), binary escapes should be rare. Since the binary must come near the escape velocity to remain outside the nucleus for a significant amount of time, we do not expect triple interactions to empty many nuclei of BHs. We will quantify these statements with our triple BH simulations.

The formation of triple SMBH systems through inspiral of a merging satellite leads to a rather specific initial configuration. The three BHs start off as a bound 'hierarchical triple', consisting of an inner binary with $a_{\rm in} \sim a_{\rm hard}$ and a more widely separated outer binary with semimajor axis $a_{\rm out}$. For very large $a_{\rm out}/a_{\rm in}$ we expect hierarchical triples to exhibit very regular behaviour; in this case the third body sees the inner binary as a point mass and the system essentially consists of two independent (inner and outer) binaries. However, as $a_{\rm out}/a_{\rm in}$ approaches unity, secular evolution gives way to chaotic three-body interactions in which the orbits diverge and the system becomes subject to escape of one its components. Mardling & Aarseth (2001) derive a criterion for the stability of three-body systems based on an analogue with the problem of binary tides. The most distant intruder orbit at which unstable interactions can begin is reliably estimated by

$$\frac{R_{\rm p}^{\rm out}}{a_{\rm in}} \approx 2.8\left[\frac{(1+q_{\rm out})(1+e_{\rm out})}{\sqrt{1-e_{\rm out}}}\right]^{2/5}, \tag{11}$$

where $R_{\rm p}^{\rm out}$ is the periapsis separation of the outer binary, $a_{\rm in}$ is the semimajor axis of the inner binary, $q_{\rm out} = m_3/(m_1 + m_2)$ is the outer binary mass ratio and $e_{\rm out}$ is its eccentricity. This criterion has great practical importance due to the high numerical cost of unnecessarily following weak hierarchical systems. It specifies an optimal starting point for our simulations, which aim to study strong interactions in three-body systems starting off as hierarchical triples.

Naively one might expect a strong three-body encounter following a merger with a galaxy hosting a binary, so long as the intruder does not induce coalescence of the binary before it reaches the centre. However, the stability criterion implies a condition for close interactions much more stringent than this. To undergo a chaotic encounter with the inner binary, the intruder must reach the stability boundary before the outer binary hardens and stalls. A triple system covers somewhat more stellar phase space than a binary of the same size, but not by much for a stable hierarchical system. This means that the merger process cannot cause the binary to harden by more than around an *e*-folding for a nearly circular, equal-mass system before the intruder arrives at the centre. Though the order-of-magnitude estimates in the previous section make this plausible, further study is needed to determine the likelihood of unstable triple interactions in realistic merger situations. An eccentric outer binary relaxes the criterion somewhat, but dynamical friction tends to circularize the orbits of satellites with moderate initial eccentricities before they reach the nucleus (Milosavljevic & Merritt 2001). We therefore assume near-circular initial orbits and begin each simulation from a weakly hierarchical configuration.

1.4 Previous work and goals of this study

Triple SMBH systems in galactic nuclei were first considered by Saslaw, Valtonen & Aarseth (1974), who computed an extensive series of Newtonian three- and four-body orbits, and compared the slingshot ejection statistics to the observed structure of extragalactic radio sources. Valtonen (1976) included a gravitational radiation drag force in the three-body dynamics. He showed that this perturbing force could in some cases yield much higher ejection velocities than would be possible in Newtonian gravity, with associated bursts of gravitational waves.

The more complex problem of three or four SMBHs coming together in the hierarchical merging process and interacting in a galactic potential was first addressed by Mikkola & Valtonen (1990) and Valtonen et al. (1994), who experimented with a variety of initial BH configurations. Heinämäki (2001) studied binary–binary scattering in galactic nuclei using initial conditions (ICs) based on extended Press–Schechter (EPS) theory (Lacey & Cole 1993). Volonteri, Haardt & Madau (2003a) followed the formation of triple BH systems in halo merger trees tracking the hierarchical build-up of SMBHs from $\sim 150\,{\rm M}_\odot$ seeds in high-σ peaks at $z \approx 20$. Using a simple analytic prescription for the ejection velocities, they inferred the presence of a large population of SMBHs and IMBHs wandering through the haloes of galaxies and intergalactic space. Iwasawa, Funato & Makino (2006) performed the first full *N*-body simulations of equal-mass triple BH systems embedded in stellar bulges, an important contribution to our understanding of galactic nuclei. Because of the large computation time required for each run, they could not statistically sample the highly varied outcomes of the three-body encounters as the previous authors did.

In this paper we study the dynamics of repeated triple SMBH interactions in galactic nuclei. Between close encounters we follow the wandering BHs through the galaxy as their orbits decay by dynamical friction. We use physically motivated initial BH configurations and mass distributions, and updated galactic models characteristic of the low-density, massive elliptical galaxies in which SMBH binaries are most likely stall. We include both a stellar and a dark matter component, with the stellar spheroid fixed to lie on the observed $m_{\rm bh}$–σ and $m_{\rm bh}$–$M_{\rm bulge}$ relations (Magorrian et al. 1998; Ferrarese & Merritt 2000; Tremaine et al. 2002; Marconi & Hunt 2003). The close encounters are treated using a KS-regularized Bulirsch–Stoer (BS) integrator provided by Sverre Aarseth (Mikkola & Aarseth 1990, 1993). The inner density profile is updated throughout the simulations to roughly account for core heating by dynamical friction and stellar mass ejection. Gravitational radiation losses are modelled as a drag force determined by the relative coordinates and velocities of each pair. Each simulation takes only a few minutes to run, so we can try a variety of distributions of ICs and statistically sample the outcomes for each. We use this algorithm to study a variety of consequences of the ongoing encounters, such as the merging efficiency of BH pairs, the time spent wandering at various distances from the galactic centre, the distribution of final sizes and eccentricities of the binaries remaining in the galaxy after a steady state has been reached and the extent of the core scouring caused by the triple SMBH systems.

Aside from the motivating order-of-magnitude calculations in previous sections, this paper does not address the question of *whether* close triple SMBH systems form in galactic nuclei. We start our simulations from a state that the system must reach shortly before the onset of unstable three-body interactions *assuming* that they occur, and proceed to derive the subsequent evolution. Our results may be used to argue for or against the occurrence of triple systems in real galaxies, as observations support or disfavour the signatures that we derive.

In Section 2 we describe our model and code methods. In Section 3 we present the results of our study, and in Section 4 we discuss these results and conclude.

2 MODEL AND METHODS

2.1 BH mass distribution and halo model

To get a physically motivated distribution of BH mass ratios, we associate the formation of the inner and outer binaries with the last two major mergers in the history of the galactic halo hosting the triple system. We use EPS theory (Lacey & Cole 1993) to calculate the probability distributions of the halo formation times and progenitor masses, and randomly select the parameters of the previous two mergers from these distributions. We then assign a BH to each progenitor halo using a simple prescription based on the assumption of a flat galactic rotation curve.

Lacey & Cole (1993) derive the instantaneous halo merger rate,

$$r_{\rm LC}(M_1, M_{\rm f}, t) = \frac{{\rm d}^2 p}{{\rm d}M_2 {\rm d}t} = \sqrt{\frac{2}{\pi}} \frac{\delta_{\rm c}}{D(z)} \left| \frac{\dot{\delta}_{\rm c}}{\delta_{\rm c}} - \frac{\dot{D}}{D} \right| \frac{|{\rm d}\sigma/{\rm d}M|_{M_{\rm f}}}{\sigma^2(M_{\rm f})} \times \frac{\exp\left\{-\left(\delta_{\rm c}^2/2D^2(z)\right)\left[\left(1/\sigma^2(M_{\rm f})\right) - \left(1/\sigma^2(M_1)\right)\right]\right\}}{\left[1 - \sigma^2(M_{\rm f})/\sigma^2(M_1)\right]^{3/2}}. \quad (12)$$

This equation gives the probability, per unit time per unit mass of M_2, of a given halo of mass M_1 merging with another halo of mass M_2 to form a product of mass $M_{\rm f} = M_1 + M_2$ at time t. Here $\sigma^2(M)$ is the present-day variance of the linear density field on mass scale M,

$$\sigma^2(M) = \frac{1}{(2\pi)^3} \int_0^\infty P(k) W^2(kr) 4\pi k^2 \, {\rm d}k, \quad (13)$$

where $P(k)$ is the power spectrum of density fluctuations today, W is a tophat window function and r is related to M through $M = (4/3)\pi\, r^3 \rho_{\rm m}$, the volume times the present-day matter density. $P(k)$ is related to the primordial power spectrum through the transfer function $T(k)$, which encapsulates the suppression of perturbations on small scales due to radiation pressure and damping over the history of the Universe. For $T(k)$ we adopt the standard fitting formulae of Eisenstein & Hu (1998). For the linear growth function $D(z)$ we use the approximation

$$D(z) \approx \frac{(5/2)\,\Omega_{\rm m}(z) D_0(z)}{\Omega_{\rm m}^{4/7}(z) - \Omega_\Lambda(z) + [1 + (\Omega_{\rm m}(z)/2)][1 + (\Omega_\Lambda(z)/70)]}, \tag{14}$$

good to within a few per cent for all plausible values of $\Omega_{\rm m}$ and Ω_Λ (Carroll, Press & Turner 1992). $D_0(z) = 1/(1+z)$ is the growth function for an Einstein–de Sitter universe, $\Omega_{\rm m}(z) = \Omega_{\rm m}(1+z)^3/[\Omega_{\rm m}(1+z)^3 + \Omega_\Lambda]$ is the matter density (normalized to the critical density) as a function of redshift, and we take $\Omega_\Lambda(z) = 1 - \Omega_{\rm m}(z)$ assuming the rest of the density is in the form of a cosmological constant. $\delta_{\rm c}$ has the weak redshift dependence (Kitayama & Suto 1996),

$$\delta_{\rm c} \approx \frac{3(12\pi)^{2/3}}{20}[1 + 0.0123 \log_{10} \Omega_{\rm m}(z)]. \tag{15}$$

We adopt the cosmological parameters obtained from 3 yr of data collection by the *Wilkinson Microwave Anisotropy Probe* (*WMAP*), $\Omega_{\rm m}h^2 = 0.127$, $\Omega_{\rm b}h^2 = 0.0223$, $h = 0.73$, $\sigma_8 = 0.74$ and $n_{\rm s} = 0.951$ (Spergel et al. 2006).

Since the merger rate (12) diverges as $M_2/M_1 \to 0$, applications of the formula that track individual merging haloes must employ a cut-off mass ratio $M_2/M_1 \equiv \Delta_{\rm m}$, such that all mergers below $\Delta_{\rm m}$ are treated as smooth *accretion* rather than as discrete mergers (see Manrique & Salavador-Sole 1996, for further discussion). The instantaneous rate of accretion on to a halo of mass M at redshift z is

$$r_{\rm a}(M, t) = \int_M^{M(1+\Delta_{\rm m})} (M' - M) r_{\rm LC}(M, M', t)\, {\rm d}M'. \tag{16}$$

To get the growth history ('accretion track') of a halo of mass M_0 at time t_0 due to accretion since the last merger, one need only solve the differential equation ${\rm d}M/{\rm d}t = r_{\rm a}[M(t), t]$, subject to the IC $M(t_0) = M_0$. We integrate this equation backward in time using a fourth-order Runge–Kutta (RK4) method to get the accretion tracks of the haloes in our simulations. Since we are interested in BH binary formation, we loosely associate $\Delta_{\rm m}$ with the halo mass ratio such that tidal stripping of the satellite would prevent the eventual merging of the two nuclei. N-body simulations of galaxy mergers place this mass ratio in the range $\Delta_{\rm m} \sim 0.1$–0.3, depending on the density and orbital parameters of the satellite (Colpi, Mayer & Governato 1999; Taffoni et al. 2003). Hence our canonical choice is $\Delta_{\rm m} = 0.3$, and we also try values of $\Delta_{\rm m} = 0.1$ (runs D1) and 0.5 (runs D5), the latter being the halo mass that corresponds to a stellar mass ratio of ~3:1 in our prescription.

Following Salvador-Sole, Solanes & Manrique (1998), we write the probability, per unit time, of a halo with mass $M_{\rm f}$ at time t arising from a merger with a smaller halo of mass between M and $M + {\rm d}M$ (the 'capture rate') as

$$r_{\rm c}(M, M_{\rm f}, t)\, {\rm d}M = r_{\rm LC}(M, M_{\rm f}, t)\theta[M_{\rm f} - M(1+\Delta_{\rm m})]\, \frac{N(M,t)}{N(M_{\rm f}, t)}\, {\rm d}M, \tag{17}$$

the EPS merger rate excluding haloes below the threshold $\Delta_{\rm m}$, and weighted by the number of mass M haloes per unit halo of mass $M_{\rm f}$.

The rate at which haloes of mass $M_{\rm f}$ form through all mergers at time t is

$$r_{\rm f}(M_{\rm f}, t) \approx \frac{1}{2} \int_{M_{\rm f}\Delta_{\rm m}/(1+\Delta_{\rm m})}^{M_{\rm f}} r_{\rm c}(M, M_{\rm f}, t)\, {\rm d}M. \tag{18}$$

The probability distribution function (PDF) of formation times of haloes with mass M_0 at time t_0 is

$$\Phi_{\rm f}(M_0, t) = r_{\rm f}[M(t), t]\, {\rm e}^{-\int_t^{t_0} r_{\rm f}[M(t'),t']\,{\rm d}t'}. \tag{19}$$

Given a formation time $t_{\rm f}$ and corresponding mass $M(t_{\rm f})$ along the past accretion track of M_0, the mass of the larger progenitor M_1 is distributed according to

$$\Phi_{\rm p}[M(t_{\rm f}), M_1] = \frac{2G(M_1, M)}{\int_{M\Delta_{\rm m}/(1+\Delta_{\rm m})}^{M/(1+\Delta_{\rm m})} G(M', M)\, {\rm d}M'}, \tag{20}$$

where

$$G(M', M) = \frac{|{\rm d}\sigma(M')/{\rm d}M'|}{M'\sigma^2(M')} \left[1 - \frac{\sigma^2(M)}{\sigma^2(M')}\right]^{-3/2}. \tag{21}$$

By choosing formation times and progenitor masses randomly according to (19) and (20), we capture the stochasticity of the intervals between mergers above $\Delta_{\rm m}$, but treat merging below this threshold only in the mean. See Manrique & Salavador-Sole (1996), Salvador-Sole et al. (1998) and Raig, Gonzalez-Casado & Salvador-Sole (2001) for further details and derivations of (19) and (20). Fig. 3 shows the distribution of formation times, progenitor masses and accretion tracks for a present-day $5 \times 10^{13}\,{\rm M}_\odot$ halo for $\Delta_{\rm m} = 0.1$, 0.3 and 0.5, and the accretion tracks for 1, 2.3 and $5 \times 10^{13}\,{\rm M}_\odot$ haloes with $\Delta_{\rm m}$ fixed at 0.3. All accretion tracks are normalized to the present-day mass M_0. Note the insensitivity of the shape of these tracks to M_0, as expected for masses above the critical mass $M^* \approx 2 \times 10^{12}\,{\rm M}_\odot$.

Our algorithm for generating the BH masses is illustrated schematically in Fig. 4. For each run we begin with a halo of mass $M_0 = 5 \times 10^{13}\,{\rm M}_\odot$ at time $t_0 \equiv t(z=0)$, choose its formation time $t_{\rm f0}$ randomly according to equation (19) and find the mass

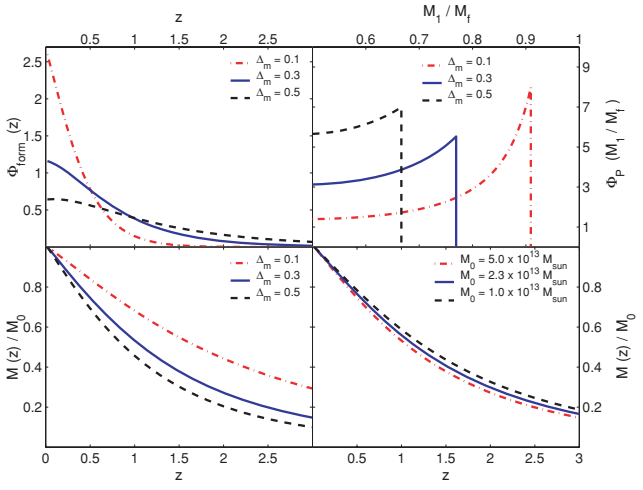

Figure 3. Upper left: PDF of formation redshifts given by equation (19) for a $5 \times 10^{13}\,{\rm M}_\odot$ halo at $z = 0$. Upper right: PDF of masses of the larger progenitor of the same halo given by equation (20), normalized to the mass of the merger product, $M_{\rm f} = 5 \times 10^{13}\,{\rm M}_\odot$. Lower left: past accretion tracks of a present-day $5 \times 10^{13}\,{\rm M}_\odot$ halo back to $z = 3$, normalized to the mass at $z = 0$. Lower right: normalized accretion tracks for three different halo masses.

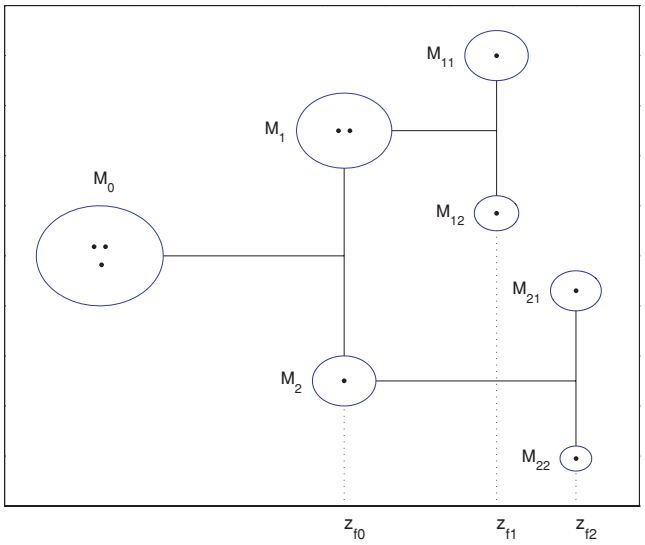

Figure 4. Schematic diagram of our algorithm for generating the BH mass distribution. First, we select a formation time $z_{\rm f0}$ for the halo M_0 hosting the BH triple system randomly from equation (19). Given $z_{\rm f0}$ and $M(z_{\rm f0})$, we select two progenitor masses M_1 and M_2 according to equation (20), assign the binary to the larger one and the third BH to the smaller one. We repeat this process going back one step further in the 'merger tree' to get the masses of the binary constituents.

$M_{\rm f0} = M(t_{\rm f0})$ along its accretion track at that time. The mass $M_{\rm f0}$ is assigned to the dark matter halo hosting the triple BH system, and the physical time for the run to end if other termination conditions are not met first is set to $t_0 - t_{\rm f0}$. To explore the dependence of the results on the absolute mass scale, we also try beginning with a $1 \times 10^{13}\,{\rm M}_\odot$ halo (runs H1).

We model the halo as a Hernquist profile (Hernquist 1990), which is identical to a Navarro, Frenk and White (NFW) profile (Navarro, Frenk & White 1997) in its inner regions if the scale radius $a_{\rm H}$ is related to the NFW scale radius by $a_{\rm H} = a_{\rm NFW}\sqrt{2[\log(1+c) - c/(1+c)]}$, where c is the halo concentration defined by $a_{\rm NFW} = r_{\rm vir}/c$. The Hernquist model falls off as r^{-4} instead of r^{-3} far outside $a_{\rm H}$ (Springel, Di Matteo & Hernquist 2005). The virial radius $r_{\rm vir}(M, z)$ is given by

$$r_{\rm vir}(M,z) = \frac{364}{1+z}\left[\frac{M_{\rm halo}}{10^{13}\,h^{-1}\,{\rm M}_\odot}\frac{\Omega_{\rm m}(z)}{\Omega_{\rm m}}\frac{18\pi^2}{\Delta_{\rm c}}\right]^{1/3} h^{-1}\,{\rm kpc}, \quad (22)$$

where $\Delta_{\rm c} = 18\,\pi^2 + 82[\Omega_{\rm m}(z) - 1] - 39[\Omega_{\rm m}(z) - 1]^2$ (Barkana & Loeb 2001), and c roughly follows the median relation from the ΛCDM simulations of Bullock et al. (2001), $c \approx 9.0\,[(2.1 \times 10^{13}\,{\rm M}_\odot)/M_{\rm halo}]^{0.13}/(1+z)$. The z dependence of $r_{\rm vir}$ and c nearly cancel to make $a_{\rm H}$ depend only weakly on redshift, so we simply use the $z = 0$ relation between $M_{\rm halo}$ and $a_{\rm H}$ in our simulations.

We choose the mass M_1 of the larger progenitor of $M_{\rm f0}$ randomly according to equation (20), and assign a mass $M_2 = M_{\rm f0} - M_1$ to the smaller progenitor. Before the merger the larger progenitor is assumed to have hosted a BH binary, while the smaller one hosted a single BH. Repeating the procedure used for M_0, we assign formation times $t_{\rm f1}$ and $t_{\rm f2}$ to M_1 and M_2 using equation (19), and choose progenitor masses M_{11}, M_{12}, M_{21} and M_{22} according to equation (20).

Having constructed a set of progenitor haloes, we now need a BH–halo relation $m_{\rm bh}(M_{\rm halo}, z)$ to complete our algorithm. We obtain such a relation by equating the halo virial velocity $v_{\rm vir}$ to the circular velocity $v_{\rm c}$ of the stellar spheroid, and using empirical $v_{\rm c}$–σ and σ–$m_{\rm bh}$ correlations to connect $v_{\rm c}$ to $m_{\rm bh}$, similar to the approaches in Wyithe & Loeb (2005) and Erickcek, Kamionkowski & Benson (2006). Combining

$$v_{\rm vir} = 343\,\sqrt{1+z} \times \left(\frac{M_{\rm halo}}{10^{13}\,h^{-1}\,{\rm M}_\odot}\right)^{1/3}\left[\frac{\Omega_{\rm m}}{\Omega_{\rm m}(z)}\frac{\Delta_{\rm c}}{18\pi^2}\right]^{1/6}{\rm km\,s^{-1}} \quad (23)$$

(Barkana & Loeb 2001) with $v_{\rm c} \approx 314\,[\sigma/(208\,{\rm km\,s^{-1}})]^{0.84}\,{\rm km\,s^{-1}}$ (Ferrarese 2002) and $\sigma/(208\,{\rm km\,s^{-1}}) \approx [m_{\rm bh}/(1.56 \times 10^8\,{\rm M}_\odot)]^{1/4.02}$ (Tremaine et al. 2002), we arrive at the relation

$$\left(\frac{M_{\rm halo}}{10^{12}\,{\rm M}_\odot}\right) = 8.28\left(\frac{M_{\rm bh}}{10^8\,{\rm M}_\odot}\right)^{0.626}\gamma(z), \quad (24)$$

where $\gamma(z) \equiv (1+z)^{-3/2}\,[(\Omega_{\rm m}/\Omega_{\rm m}(z))(\Delta_{\rm c}/18\pi^2)]^{-1/2}$.

In our canonical (CN) runs we set the masses of the inner binary members to $m_{\rm bh}(M_{11}, z_{\rm f1})$ and $m_{\rm bh}(M_{12}, z_{\rm f1})$, and that of the intruding BH to $m_{\rm bh}(M_{21} + M_{22}, z_{\rm f2})$. Note that in this prescription the intruder is usually lighter than the heavier binary member, so that most of the three-body interactions result in an exchange. To examine the effect of more interactions without exchange, we try choosing $m_{\rm bh}$ [max $(M_{21}, M_{22}), z_{\rm f2}$] for the intruder mass in runs (MX). As there is neither a direct causal relationship between $m_{\rm bh}$ and $M_{\rm halo}$ predicted by theory (Wyithe & Loeb 2005) nor a tight correlation directly observed between these two variables, and we know that identical haloes may host galaxies of different morphologies and occupation numbers, $m_{\rm bh}(M_{\rm halo}, z)$ should be taken with something of a grain of salt. Nevertheless, it is a useful way to generate simple but physically motivated BH mass distributions when no information other than the halo mass is available.

We make one final modification to the set of BH masses used in our simulations. If the outer binary's hardening radius lies outside the stability boundary given by equation (11) with $a_{\rm in} = a_{\rm hard}$, then the decay of the outer orbit is expected to stall before a strong encounter can begin. To roughly account for this we exclude all ICs where $\mu_{\rm out} > 3\mu_{\rm in}$. The final distribution of BH mass ratios is shown in Fig. 5 for $\Delta_{\rm m} = 0.1$, 0.3 and 0.5. In the upper panel we plot the inner binary mass ratios, while the lower panel shows the distribution of $m_{\rm bin}/m_{\rm esc}$, where $m_{\rm esc}$ is the mass of the lightest BH and $m_{\rm bin}$ is the sum of the masses of the other two BHs. This ratio determines the binary recoil speed when the lightest BH is ejected from the system. The total BH mass is typically $\sim 6 \times 10^8\,{\rm M}_\odot$ in our CN runs.

2.2 Stellar spheroid model

To complete the galactic model we surround the BH system by a smooth stellar potential superimposed on the dark matter halo. The stars are modelled using the 'η models' of Tremaine et al. (1994), with a sharp break to shallower slope $-\gamma$ added at $r_{\rm b} \ll a$:

$$\rho(r) = \begin{cases} \dfrac{\eta}{4\pi}\dfrac{Ma}{r^{3-\eta}(r+a)^{1+\eta}} \equiv \rho_\eta(r) & \text{if } r > r_{\rm b}; \\ \rho_\eta(r_{\rm b})(r/r_{\rm b})^{-\gamma} & \text{if } r < r_{\rm b}. \end{cases} \quad (25)$$

Our canonical model is the $\eta = 2$ (Hernquist) profile, and we also try $\eta = 1.5$ (runs SC) to explore the effect of a steeper inner profile and higher central density ($\rho \sim 800\,{\rm M}_\odot\,{\rm pc}^{-3}$ for $\eta = 1.5$ versus $\rho \sim 180\,{\rm M}_\odot\,{\rm pc}^{-3}$ for the Hernquist profile at the BH radius of influence). $r_{\rm b}$ and γ were initialized to reflect the cusp destruction caused by

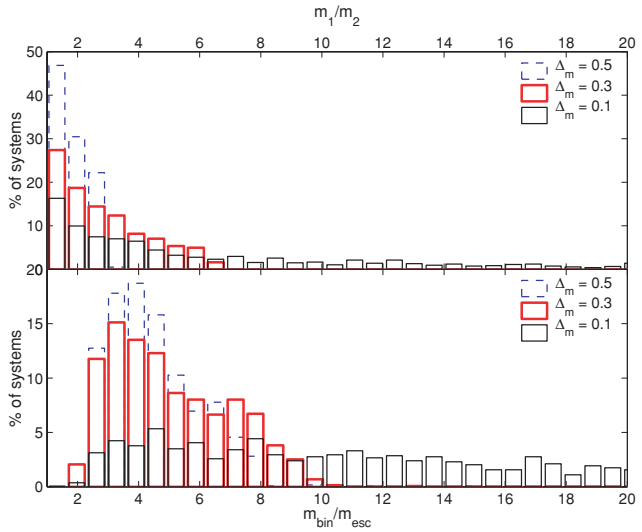

Figure 5. Distribution of BH binary mass ratios. Upper panel: inner binary mass ratio $m_1/m_2 > 1$. Lower panel: $m_{\rm bin}/m_{\rm esc}$, where $m_{\rm esc}$ is the mass of the lightest BH and $m_{\rm bin}$ is the sum of the masses of the other two BHs.

the inspiralling BHs in reaching their initial configuration, and were updated throughout the simulation to account for the continued core heating and mass ejection. Our algorithm for updating the core is described further in Section 3.5.

The parameters M and a in the η models were set based on the tight correlations observed between SMBH mass $m_{\rm bh}$ and stellar bulge mass (Magorrian et al. 1998; Marconi & Hunt 2003; Peng et al. 2005) and velocity dispersion (Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002). Marconi & Hunt (2003) found the relation $M_{\rm MH03} = (4.06 \times 10^{10}\,{\rm M}_\odot)(m_{\rm bh}/10^8\,{\rm M}_\odot)^{1.04}$ between $m_{\rm bh}$ and the virial mass $M_{\rm vir} = kR_{\rm e}\sigma_{\rm e}^2/G$ of the stellar bulge, where $R_{\rm e}$ is the half-light radius and $\sigma_{\rm e}$ is the effective bulge velocity dispersion. They set $k = 3$ (k would be 8/3 for an isothermal sphere) to get an average ratio of unity between $M_{\rm vir}$ and the dynamically measured masses $M_{\rm dyn}$ of galaxies with more direct stellar-dynamical mass determinations (Gebhardt et al. 2003). $\sigma_{\rm e}$ is typically measured over either a circular aperture of radius $R_{\rm e}/8$ (Ferrarese & Merritt 2000) or a linear aperture out to $R_{\rm e}$ (Gebhardt et al. 2000) – Tremaine et al. (2002) discuss the essential agreement between the velocity dispersions measured in these two ways. Thus for each model ($\eta = 2$, 1.5) we compute the projected radius $R_{\rm e} \equiv \kappa_1 a$ containing half the integrated surface brightness (assuming a constant mass-to-light ratio), and velocity dispersion $\sigma_{\rm e}^2 \equiv \kappa_2 GM/a$ at radius $R_{\rm e}/8$. Here κ_1 and κ_2 are constants depending on the density profile (see Tremaine et al. 1994 for relevant formulae). The parameter M is then chosen to satisfy $3R_{\rm e}\sigma_{\rm e}^2/G = 3\kappa_1\kappa_2 M = M_{\rm MH03}(m_{\rm bh}) \Rightarrow M = M_{\rm MH03}/(3\kappa_1\kappa_2)$. For the Hernquist model with $\kappa_1 = 1.815$ and $\kappa_2 = 0.104$, $M_{\rm H} = 1.76 M_{\rm MH03} = (7.15 \times 10^{10}\,{\rm M}_\odot)(m_{\rm bh}/10^8\,{\rm M}_\odot)^{1.04}$. The scale radius is then obtained from $a = GM_{\rm MH03}/3\kappa_1\sigma_{\rm bh}^2(m_{\rm bh})$, where $\sigma_{\rm bh}(m_{\rm bh})$ is the velocity dispersion computed from the $m_{\rm bh}$–σ relation of Tremaine et al. (2002). In each of these relations $m_{\rm bh}$ is set to the total mass of the triple BH system.

In a perfectly smooth, spherically symmetric galactic potential, BHs ejected on distant radial orbits return directly to the centre to interact strongly with any other nuclear BHs. Since real galaxies are clumpy and triaxial, the interaction will more realistically be delayed until the orbit of the ejected BH decays by dynamical friction. To mitigate this problem we flattened the η models by adding two low-order spherical harmonic terms to the spherical potential (de Zeeuw & Carollo 1996):

$$V(r) = u(r) - v(r)Y_2^0(\theta) + w(r)Y_2^2(\theta,\phi), \qquad (26)$$

where $u(r)$ is the potential of the spherical η model, $v(r) = -GMr_1 r^{\eta-1}/(r+r_2)^{\eta+1}$ and $w(r) = -GMr_3 r^{\eta-1}/(r+r_4)^{\eta+1}$. Since near-sphericity in the inner regions is probably a necessary prerequisite for the survival of the inner binary for of order a Hubble time until the next merger (Merritt & Poon 2004; Berczik et al. 2006), the parameters $r_1, \ldots, r_4$ were chosen to give a spherical profile near the galactic centre, and axis ratios approaching 1.3 and 1.5 for $r \gg a$. A similar triaxial modification was applied to the dark matter halo, and the relative orientation of the stellar and halo potentials was chosen randomly. By misaligning their axes we eliminate any artificial stable orbits (e.g. along the long axis of an ellipsoid) near which ejected BHs tend to return on a perfectly radial orbit to the centre. This triaxial modification had the desired effect of preventing frequent strong encounters at periapsis on distant orbits, but had little influence on the global outcome statistics.

2.3 Initial BH configuration

We assume that the three BH system starts off as a hierarchical triple on the verge of unstable three-body interactions. In our CN runs we initialize the inner binary semimajor axis $a_{\rm in}$ to $a_{\rm hard}$. To study the effect of varying $a_{\rm in}$ we also try runs with $a_{\rm in} = 3r_{\rm h}$ (runs BA) and $a_{\rm in} = r_{\rm h}/3$ (runs SA). The outer binary semimajor axis $a_{\rm out}$ is set by the stability criterion of Mardling & Aarseth (2001), equation (11). The initial eccentricity of the inner (outer) binary was chosen uniformly between 0.0 and 0.2 (0.3), in accordance with the low eccentricities found in galaxy merger simulations where dynamical friction tends to circularize the orbits of satellites as they spiral inward (Milosavljevic & Merritt 2001). The three Euler angles of the intruder's orbital plane were chosen randomly relative to the reference plane of the binary orbit, as was the phase of the initial periapsis of the binary. Both orbits were always started at periapsis; since many orbital periods elapse before unstable interactions begin, the relative phase is effectively randomized in any case. Having defined an initial configuration of three BHs embedded in a stellar + dark matter potential, we next describe how we evolve the system forward in time.

2.4 Code method

We treat the close three-body encounters using Sverre Aarseth's CHAIN code, an implementation of the N-body regularization technique of Mikkola and Aarseth (Mikkola & Aarseth 1990, 1993). The masses are first ordered so that neighbours in the chain are the dominant two-body interactions, then the KS transformation (Kustaanheimo & Stiefel 1965) is applied to neighbouring pairs. This transformation eliminates the singularity at $r \to 0$ in Newtonian gravity and transforms the equations of Keplerian motion to the simple harmonic oscillator equation (Stiefel & Scheifele 1971). External perturbing forces of arbitrary strength depending on the coordinates, velocities and/or time are simply incorporated into the formulation (though of course singularities in these perturbing forces need not be eliminated by the change of variables). We use this to add a galactic potential (Sections 2.1 and 2.2), a gravitational radiation back-reaction force and a stellar-dynamical friction force on the intruding BH. The regularized equations of motion are integrated using the BS method (Bulirsch & Stoer 1966) based on Romberg

extrapolation. For unperturbed sinusoidal motion, the BS integrator requires only two or three time-steps per orbital period!

When the binary and third body are far apart we switch to two-body motion (of the single BH and binary COM) using a fourth-order Runge–Kutta (RK4) method. We simultaneously evolve the binary semimajor axis and eccentricity using orbit averaged equations, $\mathrm{d}a = [(\mathrm{d}a/\mathrm{d}t)_{\rm st} + (\mathrm{d}a/\mathrm{d}t)_{\rm gw}]\,\mathrm{d}t$ and $\mathrm{d}e = [(\mathrm{d}e/\mathrm{d}t)_{\rm gw}]\,\mathrm{d}t$, where $(\mathrm{d}/\mathrm{d}t)_{\rm st}$ and $(\mathrm{d}/\mathrm{d}t)_{\rm gw}$ are the contributions from stellar interactions and gravitational radiation. The time-steps are adaptively controlled with a simple step-doubling scheme: at each step the 14 numbers $\{x_1, \ldots, x_6; v_1, \ldots, v_6\,; a, e\}$ are all required to remain the same to within an error $\epsilon = 10^{-n}$ under doubling of the step size. To avoid wasting computation time when any of these values approach zero, we accept agreement to n decimal places as an alternative criterion for convergence. For the calculations reported in this paper we set $n = 12$.

The relative perturbation to the binary from the third body at apoapsis,

$$\delta F = \frac{2 r_{\rm ap}^3 m_3}{\min(m_1, m_2) d^3}, \tag{27}$$

is used to decide which integration method to use at any given time. Here $r_{\rm ap}$ is the apoapsis distance between m_1 and m_2, m_3 is the intruder mass and d is the distance of the intruder from the binary COM. We switch to two-body RK4 integration each time δF falls below 5×10^{-5} and call the CHAIN code again when δF reaches 5×10^{-4}. We choose different δF thresholds for beginning and ending close encounters to prevent overly frequent toggling between the two methods. When <3 BHs remain in the simulation (after coalescence of the inner binary or escape of one or more BHs from the galaxy), we primarily use the RK4 integrator, but call the CHAIN code to treat very close two-body encounters. Since chain regularization is defined only for three or more bodies, we add a light and distant 'dummy' particle when using this method for two-body motion.

During the two-body motion we declare the single BH or binary (remnant) to have escaped if its distance from the galactic centre exceeds 500 kpc and its specific energy $E = \Phi(r, \theta, \phi) + (1/2)v^2$ exceeds $E_{\rm esc}$, the energy needed to escape from $r = 0$ to infinity. We declare the binary to have coalesced during a close encounter when (i) $r_{12} < 3r_{\rm sb}$, where $r_{\rm sb}$ is the Schwarzchild radius of the larger member of the pair or (ii) $|a/\dot{a}|_{\rm gr} < 0.1 t_{\rm dyn}$ and $|a/\dot{a}|_{\rm gr} < 50\,{\rm yr}$ while $\delta F < 10^{-3}$, where $t_{\rm dyn}$ is the current outer binary dynamical time. During the RK4 integration we require that $|a/\dot{a}|_{\rm gr} < 50\,{\rm yr}$ or $r_{12} < 1.1(r_{\rm s1} + r_{\rm s2})$ at periapsis, where $r_{\rm s1,2}$ are the Schwarzchild radii of the two binary members. Upon coalescence we replace the pair with a single body of mass $m_{\rm bin}$ and the COM position and velocity. A run ends when (a) only one SMBH remains in the galaxy and it has settled to the centre of the potential by dynamical friction; (b) two BHs remain and have formed a hard binary at the galactic centre; (c) the physical time exceeds $t_{\rm max} = t_0 - t_{\rm f0}$, the current age of the universe minus the halo formation time; (d) all BHs have escaped the galaxy or (e) the physical time spent in a call to CHAIN exceeds a maximum allowed time $t_{\rm chn}$. The last condition is added to avoid spending too much computation time on very long close encounters.

2.4.1 *Treatment of stellar-dynamical friction*

During the two-body evolution we apply a dynamical friction force given by Chandrasekhar's formula (Chandrasekhar 1943; Binney & Tremaine 1987),

$$\left(\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}\right)_{df} = -\frac{4\pi G^2 \rho m \ln\Lambda[\mathrm{erf}(X) - 2X\mathrm{e}^{-X^2}/\sqrt{\pi}]}{v^2}\hat{\boldsymbol{v}}, \tag{28}$$

where $X \equiv v/(\sqrt{2}\sigma)$, to the single BH and binary COM. The factor in square brackets ≈ 1 for $v \gg \sigma$ and $\approx 0.75X^3$ for $v \ll \sigma$. We take

$$\ln\Lambda = \max\left\{\ln\left[\frac{r(\sigma^2 + v^2)}{Gm}\right], 1\right\} \tag{29}$$

for the Coulomb logarithm, where r is the BH's distance from the galactic centre. For ρ in equation (28) we use $\min\,[\rho(r), \rho(r_{\rm inf})]$, effectively capping the density at its value at the BH radius of influence, $r_{\rm inf} = Gm/\sigma^2$, when the BHs pass through the core.

The semimajor axis a of the binary also evolves under stellar-dynamical friction as it wanders through the galaxy. However, Chandrasekhar's formula applied separately to the binary constituents does not give a good description of this evolution, since the hard binary loses energy through close three-body encounters with stars, while equation (28) relies on the assumption that the energy loss is dominated by weak two-body encounters. We approximate the evolution of a using a formulation for the decay rate of a hard, massive binary in a uniform and isotropic sea of stars developed in Mikkola & Valtonen (1992) and Quinlan (1996). The formulation was calibrated with an extensive series of three-body scattering experiments in Quinlan (1996) and tested against N-body simulations in Mikkola & Valtonen (1992). The binary decay rate is given by

$$\left(\frac{\mathrm{d}a}{\mathrm{d}t}\right)_{\rm st} = -\frac{\rho H a^2}{\sigma}, \tag{30}$$

where the hardening rate H can be approximated by the empirical fitting function (Quinlan 1996)

$$H \approx \frac{16}{\left[1 + (\sigma/w)^4\right]^{1/2}}. \tag{31}$$

Here $w = 0.85\sqrt{G\,\min(m_1, m_2)/a}$ is the characteristic velocity distinguishing the hard binary regime – stars with $v \gtrsim w$ cannot be easily captured into bound orbits and preferentially harden the binary in close encounters. In our simulations the binary COM is often speeding through the stellar medium at $v_{\rm cm} \gtrsim \sigma$ after an energetic ejection, so the stellar medium looks 'hotter' in its frame of reference. To account for this we replace σ in equations (30) and (31) with $\sigma^* \equiv \sqrt{v_{\rm cm}^2 + \sigma^2}$, a good approximation since H is not very sensitive to the shape of the distribution function (e.g. $H \approx 16$ for a Maxwellian versus $H \approx 18$ for a uniform velocity distribution). For ρ in equation (30) we took $\min\,[\rho(r), \rho(r_{\rm inf})]$ as we did for the drag on the COM. We ignored the mild eccentricity evolution $(\mathrm{d}e/\mathrm{d}t)_{\rm st}$, which is shown in Quinlan (1996) to be far weaker than that predicted by Chandrasekhar's formula for hard eccentric binaries.

When the amplitude of oscillation of one of the two masses falls below $Gm/2\sigma^2$, we stop integrating its motion and place it at rest at the galactic centre until the second body returns to within a distance of twice the break radius, $2r_{\rm b}$. If the settled mass is the binary, then we also stop updating its semimajor axis for stellar hardening, assuming that it clears out its loss cone and stalls once it stops moving about the nucleus and encountering new stars. Since the total mass in loss cone stars is small compared to the BH mass in the low-density galaxies that we consider, to good approximation the binary stalls as soon as the replenishing mechanism (motion) shuts off.

During close encounters between the three BHs an orbit-averaged prescription for stellar-dynamical friction is not feasible. However,

the triples are still marginally stable at the boundary given by equation (11), so we apply a drag force given by Chandrasekhar's formula with $\ln \Lambda = 1$ to the intruder at the beginning of each run. At the onset of chaotic interactions in the first encounter (defined loosely by the first time the closest pair is not formed by the original binary members) this perturbation is shut off, and it remains off in all later close encounters. Fortunately the chaotic interactions occur on time-scales very short compared to a dynamical friction time, so it is valid to neglect stellar dissipation during close encounters.

2.4.2 *Treatment of gravitational radiation*

Gravitational radiation is modelled using the $\mathcal{O}[(v/c)^5]$ post-Newtonian (2.5PN) back-reaction acceleration computed by Damour & Deruelle (1981), evaluated in the two-body COM frame (e.g. Gultekin, Miller & Hamilton 2006),

$$\frac{\mathrm{d}\boldsymbol{v_1}}{\mathrm{d}t} = \frac{4G^2}{5c^5}\frac{m_2}{m_1+m_2}\frac{m_1m_2}{r^3}\left\{\hat{\boldsymbol{r}}(\hat{\boldsymbol{r}}\cdot\boldsymbol{v})\left[\frac{34G(m_1+m_2)}{3r} + 6v^2\right] + \boldsymbol{v}\left[-\frac{6G(m_1+m_2)}{r} - 2v^2\right]\right\}. \quad (32)$$

$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ and $\boldsymbol{v} = \boldsymbol{v}_1 - \boldsymbol{v}_2$ are the relative positions and velocities of the two masses. We sum the force linearly over all pairs, a valid approximation provided the perturbations from the third body and other external tidal forces are instantaneously small at periapsis. When averaged over a complete orbit, equation (32) is equivalent to the Peters (1964) equations for the binary semimajor axis and eccentricity:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\frac{64}{5}\frac{G^3m_1m_2(m_1+m_2)}{c^5a^3}\frac{1+(73/24)e^2+(37/96)e^4}{(1-e^2)^{7/2}}, \quad (33)$$

$$\frac{\mathrm{d}e}{\mathrm{d}t} = -\frac{304}{15}\frac{G^3m_1m_2(m_1+m_2)}{c^5a^4}\frac{e+(121/304)e^3}{(1-e^2)^{5/2}}. \quad (34)$$

However, when $|\hat{\boldsymbol{v}}\cdot\hat{\boldsymbol{r}}|$ comes close to one on hyperbolic orbits, so that $(\hat{\boldsymbol{r}}\cdot\boldsymbol{v})^2 \to v^2$, $\dot{E} = \boldsymbol{F}_1\cdot\boldsymbol{v}_1 + \boldsymbol{F}_2\cdot\boldsymbol{v}_2 = m_1\boldsymbol{a}_1\cdot\boldsymbol{v}$ as given by equation (32) becomes *positive*, though we know physically that gravitational waves can only carry energy away from the system. To give the correct answer averaged over an orbit, this positive contribution must be cancelled by extra energy loss near periapsis, making the equation potentially sensitive to numerical error. This effect is much less pronounced in the Damour & Deruelle (1981) form than in other expressions derived for the radiation back-reaction acceleration – they derived the formula specifically for practical use on the problem of two point masses (see the appendix of Lee 1993 and references therein).

For computational ease we neglect the lower order 1–2PN terms (precession of the periapsis) in the post-Newtonian expansion. Though much larger in magnitude than the radiation reaction force, these terms are unimportant in the statistical sense because they conserve the intrinsic properties of the system, such as energy (Iwasawa et al. 2006; Kupi, Amaro-Seoane & Spurzem 2006). We need not concern ourselves with relativistic precession destroying the Kozai resonance since the semimajor axis ratio given by equation (11) is much smaller than that of equation (9).

2.5 Code tests and energy errors

One way to establish the reliability of our integration methods is to test them on problems with known solutions. Fig. 6 shows an example on the two-body problem with gravitational radiation. The upper

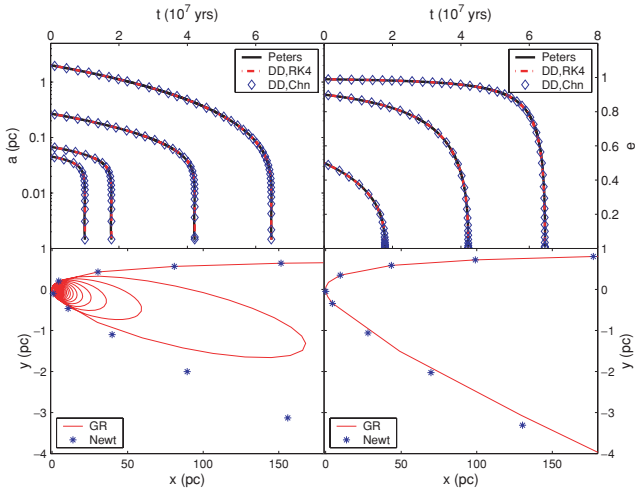

Figure 6. Code tests on the Newtonian two-body problem with gravitational radiation. Upper panels: evolution of elliptical orbits computed using RK4 integration of the Peters equations (black solid), RK4 integration with the Damour & Deruelle (DD) radiation back-reaction acceleration (red dashed) and CHAIN integration with the DD acceleration (diamonds). The ICs were chosen to give $|a/\dot{a}| = 10^7$ yr at the beginning of each integration. Left: semimajor axis evolution for initial eccentricities of $e_0 = 0.0$, 0.5, 0.9 and 0.99 (bottom to top). Right: eccentricity evolution for $e_0 = 0.5$, 0.9 and 0.99. The curves' indistinguishability demonstrates the reliability of all three methods. Lower panels: hyperbolic orbits with impact parameters b set to 80 and 120 per cent of the critical value for gravitational radiation capture, computed using the DD acceleration in the CHAIN code. Left: $0.8b_{\rm crit}$; BH is captured. Right: $1.2b_{\rm crit}$; BH is not captured. The blue asterisks are points along the Newtonian trajectory (without gravitational radiation). The deviation from the Newtonian trajectory after periapsis can be seen in both plots, even though the energy remains positive in the latter.

panels show the evolution of the semimajor axis a and eccentricity e of four decaying elliptical orbits, computed using (a) our RK4 integrator and equation (32), with an error tolerance of $\epsilon = 10^{-9}$, (b) the CHAIN code and equation (32), with $\epsilon = 10^{-14}$ and (c) the Peters (1964) equations (33–34). In each case the initial semimajor axis a_0 was chosen to give a gravitational radiation time-scale of $|a/\dot{a}| \approx 10^7$ yr, and the four curves (from bottom to top) are for eccentricities of 0.0, 0.5, 0.9 and 0.99. The agreement of the three computation methods demonstrates the reliability of both the RK4 integrator and our implementation of the CHAIN code in handling dissipative forces.

The lower panels show two hyperbolic orbits with periapsis distances around 30 times the Schwarzchild radius $r_{\rm sb}$ of the larger BH, computed using equation (32) in CHAIN. The RK4 integrator was found to fail some tests on very close approaches from hyperbolic orbits with gravitational radiation, so we treat all such approaches using the regularized CHAIN code in our runs, even during the unperturbed binary evolution. The blue asterisks are points along the Newtonian orbits while the red solid lines show the trajectories with gravitational radiation. There is a simple analytic expression for the maximum periapsis distance for gravitational radiation capture from a hyperbolic orbit,

$$r_{\rm p,max} = \left[\frac{85\sqrt{2}\pi G^{7/2}m_1m_2(m_1+m_2)^{3/2}}{12c^5v_\infty^2}\right]^{2/7}, \quad (35)$$

where m_1 and m_2 are the masses of the two bodies and v_∞ is their relative velocity at infinity. The orbit on the lower left begins at 80 per cent of the critical impact parameter and the incoming BH

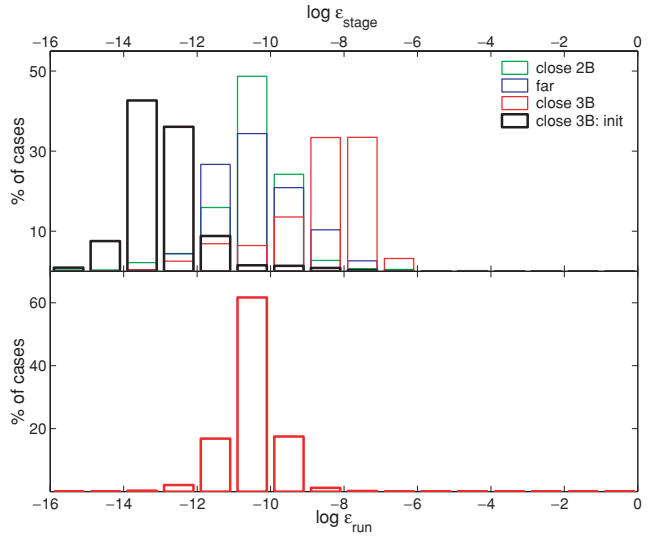

Figure 7. Energy errors. Upper panel: errors given by equation (36), for each code stage: close triple encounters, integrated with CHAIN (red); unperturbed binary evolution with the RK4 integrator (blue) and very close two-body encounters computed with CHAIN after the binary has coalesced (green). The black (heavy) histogram shows the energy errors for close triple encounters normalized to the *initial* energy instead of the energy dissipated. Lower panel: effective energy errors for the entire run, computed from equation (37).

is captured. On the right the intruding BH starts at 120 per cent of the critical impact parameter and is not captured, though the deviation from the Newtonian trajectory due to the energy radiated at periapsis can be seen on the way out. We tried iterating over impact parameters close to the critical value and found that the code reproduces equation (35) to within a part in 10^6 for periapsis distances $r_{\rm peri} \sim 30 r_{\rm s}$, and to within a part in 10^3 for $r_{\rm peri} \sim 3 r_{\rm s}$.

We also evaluated the performance of the code by repeating our canonical set of 1005 runs with a static inner profile to check the precision of energy conservation. In Fig. 7 we histogram the energy errors, computed as

$$\epsilon = \left| \frac{E_0 + \sum_i \left[\int_{t_0}^{t_{\rm f}} (\boldsymbol{F}_{{\rm df},i} \cdot \boldsymbol{v}_i + \boldsymbol{F}_{{\rm gw},i} \cdot \boldsymbol{v}_i) {\rm d}t \right] - E_{\rm f}}{E_0 - E_{\rm f}} \right|, \qquad (36)$$

where E_0 and $E_{\rm f}$ are the initial and final energies, and the two terms in the integral under the sum are the work done by dynamical friction and gravitational radiation during the current stage of the code. In the upper panel we separately plot the errors for close three-body encounters, RK4 integration of the unperturbed binary motion ('far'), and close two-body encounters computed with CHAIN during the unperturbed binary evolution. The plot includes all code stages where the energy dissipated was at least 10^{-3} in code units, or about a part in 10^{5-6} of the initial binding energy of the system. The black (heavy) histogram shows the errors for close encounters normalized to the *initial* energy instead of the dissipated energy in the denominator of equation (36), since the energy dissipated was very small in many close encounters. In the lower panel we combine the energy errors from the various code stages to get an effective energy error for each entire run,

$$\epsilon_{\rm run} = \frac{\sqrt{\epsilon_1^2 \Delta E_1^2 + \epsilon_2^2 \Delta E_2^2 + \cdots + \epsilon_n^2 \Delta E_n^2}}{\Delta E_1 + \Delta E_2 + \cdots + \Delta E_n}. \qquad (37)$$

We had to combine the separate errors to obtain $\epsilon_{\rm run}$ since the galactic potential is handled slightly differently during different stages of the code, e.g. the triaxial modification is applied only during the RK4 integration. In a large majority of cases $\epsilon_{\rm run}$ falls between 10^{-12} and 10^{-9}, and energy is conserved to better than a part in 10^4 in every run. The excellent energy conservation gives us confidence in the robustness of our integration methods.

3 RESULTS

3.1 Outcome statistics

We begin with an overview of the outcomes of our three-body simulations. In subsequent sections we focus on various effects in more detail. Our data consists of eight sets of 1005 runs, each sampling a different distribution of the ICs. A set of 1005 runs took anywhere from ∼4 to ∼30 h to finish on five 2.0-GHz Opteron processors, depending on the ICs.

In our CN runs, we chose $\Delta_{\rm m} = 0.3$ for the threshold merger mass ratio, modelled the stellar bulge as a Hernquist ($\eta = 2$) profile, started off the inner binary at $a_{\rm hard}$, and generated the ICs from a $5 \times 10^{13}\,{\rm M}_\odot$ halo at $z = 0$. In each of the remaining runs we varied one of these assumptions. Runs D1 and D5 used $\Delta_{\rm m} = 0.1$ and 0.5 to explore the effects of widening or narrowing the range of BH mass ratios. In runs MX we assigned a mass $m_{\rm bh}$ [max (M_{21}, M_{22}), $z_{\rm f2}$] instead of $m_{\rm bh}(M_{21} + M_{22}, z_{\rm f2})$ to the intruding BH, as discussed in Section 2.1. In runs BA and SA we started off the inner binary at $3a_{\rm hard}$ and $a_{\rm hard}/3$ instead of at $a_{\rm hard}$. We initialized the stellar bulge to an $\eta = 1.5$ profile in runs SC, to explore the effect of a steeper inner cusp. Finally, in runs H1 we generated the ICs from a $1 \times 10^{13}\,{\rm M}_\odot$ halo at $z = 0$, for total BH masses of $\sim 5 \times 10^7\,{\rm M}_\odot$, about an order of magnitude lower than in our CN runs. Table 1 summarizes the outcomes.

The first two rows give the percentage of cases in which (i) one BH pair coalescenced by the end of the run (i.e. by the time since the last major merger) and (ii) the remaining two BHs also coalesced within the run time. At least one pair coalesced in a large majority of the runs for each set of ICs that we tried. The new system formed from the third BH and binary remnant also coalesced in ∼10–20 per cent of the cases. Since we assume that stellar hardening of the new binary shuts off at $a_{\rm hard}$, it can only coalesce by gravitational radiation from a highly eccentric orbit; we will discuss this topic further in Sections 3.2 and 3.4. The coalescence rate is somewhat lower (∼68 per cent) in set D1, since (a) the hardening effect of the third body is lessened for more extreme mass ratios and (b) mergers with mass ratios as low as $\Delta_{\rm m} = 0.1$ are more frequent, so the run time is typically shorter. Naturally the coalescence rate is somewhat higher (95 per cent) in runs SA, where we begin with a tighter binary ($a_0 = a_{\rm hard}/3$, $\tau_{\rm gw} \sim 10^{11-12}$ yr). Coalescence is also significantly less common in runs H1. This can be understood in light of equation (3) in Section 1.1. The separation between the scale set by the stellar kinematics ($a_{\rm hard}$) and that set by gravitational radiation ($a_{\rm gw}$) is proportional to $m_{\rm bin}^{-1/4}$. Hence in lower mass systems, coalescence is less likely relative to escape. This observation motivates future study of triple BH dynamics in much lighter systems.

The next three rows of the table give the fraction of runs in which (i) the single BH escaped the stellar bulge + halo potential, (ii) all BHs (both the single and the binary) escaped the halo and (iii) the single BH either escaped or remained wandering far out in the halo at the end of the run. The single escaped in ∼15–20 per cent of the runs in all cases. If we also count runs where it remained wandering through the halo for of order a Hubble time, this fraction increases to ∼40 per cent. Double escapes (of both the binary and the single BH)

Table 1. Summary of outcomes for eight different distributions of the ICs. See text for explanation of the entries.

Outcome	CN	D1	D5	MX	BA	SA	SC	H1
Coalescence								
One pair	87 per cent	68 per cent	89 per cent	84 per cent	84 per cent	95 per cent	84 per cent	75 per cent
Two pairs	15 per cent	13 per cent	18 per cent	16 per cent	22 per cent	13 per cent	16 per cent	10 per cent
Escape								
Single	15 per cent	18 per cent	17 per cent	19 per cent	21 per cent	14 per cent	14 per cent	22 per cent
Double	0.0 per cent	0.0 per cent	0.1 per cent	0.0 per cent	0.1 per cent	0.2 per cent	0.1 per cent	0.3 per cent
Wander	37 per cent	51 per cent	38 per cent	42 per cent	42 per cent	26 per cent	45 per cent	50 per cent
Final state								
Binary	60 per cent	67 per cent	54 per cent	57 per cent	51 per cent	65 per cent	56 per cent	63 per cent
Single	38 per cent	31 per cent	45 per cent	41 per cent	47 per cent	33 per cent	43 per cent	35 per cent
No BH	1.3 per cent	2.0 per cent	1.2 per cent	1.5 per cent	1.4 per cent	1.4 per cent	1.2 per cent	1.8 per cent
Core								
$\langle M_{\rm def}\rangle$	1.41	1.20	1.50	1.38	1.31	1.96	1.42	1.61
$\Delta\, M_{\rm def}$	0.48	0.37	0.57	0.44	0.52	0.44	0.57	0.61
Termination								
Sing escape	14 per cent	16 per cent	16 per cent	18 per cent	18 per cent	14 per cent	12 per cent	21 per cent
New binary	61 per cent	44 per cent	61 per cent	55 per cent	53 per cent	73 per cent	54 per cent	49 per cent
T.O. far	22 per cent	30 per cent	22 per cent	23 per cent	20 per cent	12 per cent	31 per cent	28 per cent
T.O. CHAIN	2.1 per cent	10 per cent	1.7 per cent	4.4 per cent	8.4 per cent	0.3 per cent	1.6 per cent	1.5 per cent
Crashed	0.5 per cent	0.1 per cent	0.3 per cent	0.3 per cent	0.7 per cent	0.2 per cent	0.6 per cent	0.1 per cent

were very rare. We get more wandering BHs in set D1, since a larger fraction of the released energy is apportioned to the escaper when it is relatively lighter, the dynamical friction time is longer and the run time is shorter. Runs SA produced less wandering BHs since the binary pair more often coalesced before the intruder had a chance to harvest much of its energy. Wandering was also more common in set H1, due to the $m_{\rm bh}^{-1/4}$ scaling discussed in the previous paragraph. The escape fraction of course depends on the depth of the galactic potential well. Given the uncertainty and scatter in the $m_{\rm bh}$–$M_{\rm halo}$ relation and specificity of the prescription adopted, we must expect these numbers to vary somewhat in studies with different halo or stellar density models.

The entries under 'final state' tell whether, at $z = 0$, the galactic centre hosts (i) a stalled BH binary, (ii) a single BH or (iii) no BHs (neither the single nor the binary has yet returned to the centre by dynamical friction). About 50–70 per cent of the runs ended with a binary at the galactic centre whose gravitational radiation time exceeded the time until $z = 0$. This includes cases where (a) the single was ejected to large distance and the binary settled to the centre before it hardened enough to coalesce by gravitational radiation, (b) when the inner binary coalesced during a close encounter the outer binary coalescence time exceeded the remaining run time or (c) the single and binary remnant both returned to the centre after an ejection and formed a bound pair with a long gravitational radiation time. In most of the remaining cases (30–50 per cent) the run ended with a single BH at the galactic centre, or a binary destined to coalesce before t_0. This occurred when (a) the single was ejected to large distance and the binary (or remnant) settled to the centre after having hardened to the point of coalescence through some combination of repeated interactions with the third BH and stellar dissipation or (b) a new binary with a short gravitational radiation time formed following return from an ejection or coalescence during a close encounter. In only a small fraction (1–2 per cent) of cases the run ended with the centre empty of BHs. Note also that this happened most often in runs where the last merger occurred recently, so the total *time* spent with the centre empty of BHs is still smaller.

The next two entries give the mean and standard deviation of the core 'mass deficit' scoured out by the triple system, in units of the total BH mass $m_{\rm bh}$. For a galaxy modelled as an η model with stellar mass parameter $M_{\rm s}$, bulge scale radius $a_{\rm s}$ and a break to inner slope γ at $r_{\rm b}$, we define the mass deficit $M_{\rm def}$ by

$$M_{\rm def} = 4\pi\left[\int_0^{r_{\rm b}} \rho_\eta(r) r^2 {\rm d}r - \int_0^{r_{\rm b}} \rho(r) r^2 {\rm d}r\right] = M_{\rm s}\left[\left(\frac{r}{r+a_{\rm s}}\right)^{\eta_{\rm s}} - \left(\frac{r_{\rm b}}{r_{\rm b}+a_{\rm s}}\right)^{\eta_{\rm s}}\right] - \frac{4\pi\rho_{\rm bs}}{3-\gamma} r_{\rm b}^3 + {\rm DM}, \tag{38}$$

where $\rho_{\rm bs}$ is the stellar density at $r_{\rm b}$ and DM denotes the corresponding dark matter terms. The mass deficits are highly scattered within each set of runs, with typical $M_{\rm def}/m_{\rm bh} \approx 1.4 \pm 0.5$. More extreme mass ratios (runs D1) tended to produce smaller cores, while a narrower mass range (runs D5) gave somewhat larger ones. The fraction of runs ending with very high mass deficits varied strongly with $\Delta_{\rm m}$; for instance 17 per cent of cases ended with $M_{\rm def}/m_{\rm bh} > 2$ for $\Delta_{\rm m} = 0.5$ versus only (11, 4.4 per cent) for $\Delta_{\rm m} = (0.3, 0.1)$. The large cores in set SA arose mostly from enhanced core scouring during the creation of the initial hard binary, and so are more a consequence of the ICs than of the triple interactions themselves. This sensitivity of $M_{\rm def}$ to the binary stalling radius is an interesting point in its own right. The larger cores in runs H1 probably arise from the higher mean number of ejections and smaller fraction of runs ending in immediate coalescence as the BH mass is decreased. 21 per cent of the runs in this set ended with $M_{\rm def}/m_{\rm bh} > 2$ and 8.7 per cent ended with $M_{\rm def}/m_{\rm bh} > 2.5$. The subject of core scouring will be discussed in further detail in Section 3.5.

Both the core scouring effect and the coalescence rate induced by the encounters are significantly reduced for the extreme mass ratios

in set D1. One must keep in mind, however, that halo mergers with these mass ratios are much more frequent than those above $\Delta_m = 0.3$ (see Fig. 3), so the *cumulative* effect of these events may be as high or higher than that of encounters with near-equal masses. To quantify this statement our simulations would need to be embedded in a merger tree that follows the formation of triple systems.

Finally, the last five lines in Table 1 give the statistics of the condition which formally terminated the run: (i) the single escaped the halo and the binary (or remnant) settled to rest at the galactic centre; (ii) the escaper and binary remnant formed a stalled binary or coalesced (in a few per cent of these cases a bound binary never actually formed; the pair coalesced suddenly upon a very close periapsis passage from an unbound orbit); (iii) the maximum physical time $t_{max} = t_0 - t_{f0}$ was reached (in these cases one or more BHs were left wandering through the halo at the end of the run); (iv) the maximum time for a close encounter ($t_{chn} = 3 \times 10^7$ yr for our CN runs) was exceeded or (v) the time-step went to zero or a limit on the number of time-steps was reached at some stage of the integration, which always occurred in <1 per cent of cases. Runs terminating on condition (iv) or (v) were left out when computing the upper entries in Table 1.

In this slew of runs we have varied only a few of the relevant parameters; one might also try, for instance, varying or adding scatter to the halo mass prescription, further steepening the stellar bulge profiles or adding a disc component, and exploring vastly different BH mass scales, in particular the much lower ($\sim 10^{4-5}$ $M_\odot$) masses that may be relevant at high redshift. One of the advantages of our method is the relative ease of varying the model and ICs. This paper should be viewed as a work in progress, in which we have developed a method that can be applied to three BH systems in whatever context they may arise. Given the qualitative similarity of the outcomes in the runs we have performed so far, we will focus on the CN runs in the more detailed presentation of our results.

3.2 Efficient binary coalescence

The inner binary begins at a_{hard}, where the gravitational radiation time is $\tau_{gw} \sim 10^{13-15}$ yr, in our CN runs. It must shrink by a factor of ∼10 before gravitational radiation can cause coalescence in a Hubble time, or by a factor of ∼40 for τ_{gw} to become comparable to the dynamical friction time. The intruder helps to bridge this gap in several ways: (a) direct hardening of the binary through repeated three-body interactions, (b) enhanced stellar hardening by scattering of stars into the loss cone and motion of the binary about the nucleus and (c) enhanced gravitational radiation losses due to thermalization of the eccentricity during the chaotic encounters and eccentricity growth via the Kozai resonance.

We find that the combination of these mechanisms leads to coalescence of the inner binary within the time $t_0 - t_{f0}$ between the merger that formed the triple system and $z = 0$ in a large majority of the runs. It is instructive to distinguish the systems that coalesce by 'collision' during a close three-body encounter from those that gradually harden enough to coalesce within the time $t_0 - t_{f0}$, through the cumulative effect of repeated encounters and loss-cone refilling while the binary wanders about the nucleus. In runs CN, 23 per cent of the systems undergo collision during the first encounter, and another 19 per cent coalesce during later close encounters, for a net 42 per cent collision rate. Thus about half of the total coalescence efficiency arises from collisions during close encounters, and the other half comes from gradual hardening over the course of the simulation. Kozai oscillations account for most of the collisions during the first encounter, while in later encounters

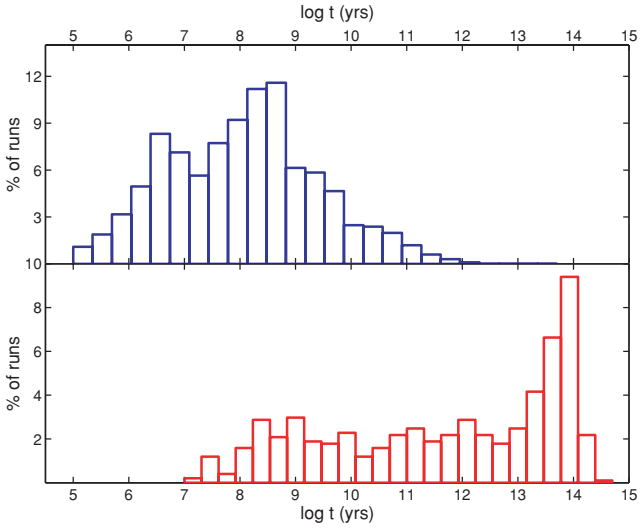

Figure 8. Distribution of coalescence times for the inner binary (upper panel) and new binary (lower panel), formed by the third BH and coalesced binary remnant. The lower plot includes only runs where a new binary formed, and not e.g. cases where the single escaped or was left wandering far from the galactic centre at the end of the run.

random eccentricity variations are more likely to result in coalescence, since the binaries are harder. Our numbers are reasonable based on analytic estimates of the collision rate in chaotic encounters (e.g. Valtonen & Karttunen 2006, chapter 11).

Fig. 8 shows the distribution of binary coalescence times. The upper panel is for the inner binary, while the lower panel is for the new system formed by the third BH and binary remnant. In cases where coalescence occurred during the run, we plot the coalescence time recorded by the code. In other cases we plot $t_{run} + t_{gr,end}$, where t_{run} is the total run time and $t_{gr,end}$ is the time obtained by integrating the Peters (1964) equations from the state at the end of the run to coalescence. The lower plot includes only those runs where the third BH ended up bound to the binary remnant, excluding, for instance, cases where the single escaped the galaxy. In ∼15 per cent of the runs the new binary also coalesced within the time $t_0 - t_{f0}$.

Under circumstances where the gap-crossing mechanisms discussed in Section 1.1 fail, the efficient coalescence in massive triple systems provides a 'last resort' solution to the final parsec problem.

3.3 The three-body interactions

Though the close encounters take up only a small fraction of the physical time in our runs, it is the energy exchanges during these encounters that determine the large-scale BH dynamics. We now take a closer look at the three-body dynamics in a few representative cases.

In ∼20 per cent of the runs the binary swiftly coalesces during the first encounter, usually with the help of the Kozai resonance. Two examples of this are shown in Fig. 9. The time evolution of the inner and outer binary separations is plotted for two different runs in the upper panels. For a circular orbit the separation would be roughly constant over an orbital period, or just a horizontal line in the figure. On the left the inner binary undergoes many Kozai oscillation cycles before coalescing. Observe that at the second-to-last eccentricity maximum, though it does not coalesce, the binary radiates away a large amount of energy and passes through the next eccentricity minimum with a significantly reduced semimajor axis.

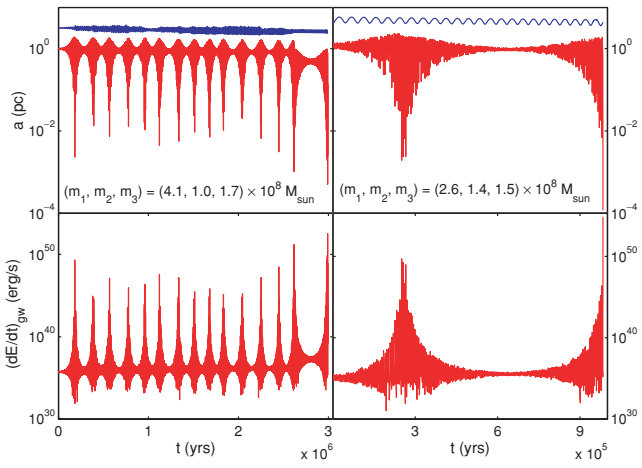

Figure 9. Two examples of rapid coalescence by Kozai oscillations. Upper panels: time evolution of the inner (red, lower) and outer (blue, higher) binary separations. Lower panels: total gravitational radiation power, averaged over BS time-steps. m_1 and m_2 are the masses of the binary members and m_3 is the mass of the intruder.

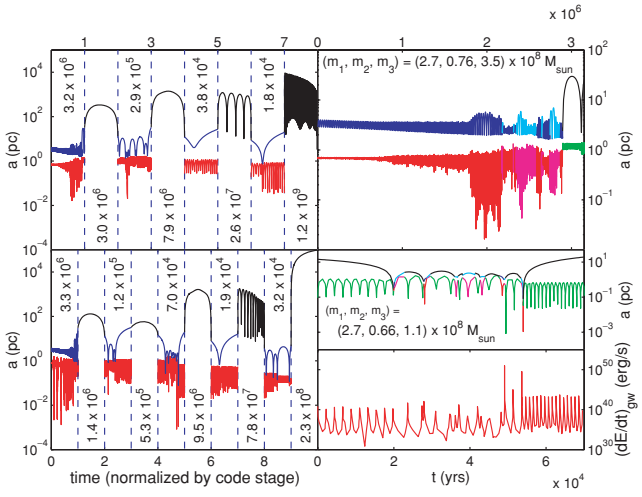

Figure 10. Two examples of longer runs. Left panels: entire run, with time spent in each code stage normalized to unity. Actual times in years are indicated by the numbers on the plot. The red (inner binary) and blue (outer binary) portions show the close encounters, while the black portions show the calls to the RK4 integrator. Upper right: zoom in on the first close encounter in the run at left. Lines are colour coded according to which pair is closest, to highlight the exchanges. Lower right: zoom in on the third close encounter of the lower run, showing also the total gravitational radiation power averaged over BS time-steps. See text for further explanation of this figure.

In the example on the right, the binary coalesces after just one full Kozai cycle. The lower panels show the time evolution of the total gravitational radiation power, averaged over the BS time-steps. Since the system starts on the verge of chaotic interactions where the outer to inner binary semimajor axis ratio is small (so that the quadrupolar approximation breaks down), we get 'messy' Kozai oscillations which can give way to catastrophic eccentricity growth at an unpredictable time.

Fig. 10 shows two examples of more complex runs. The left-hand panels summarize the entire run, including all of the close encounters and ejections in between. Each call to the CHAIN code or the unperturbed binary integrator is separated by dashed vertical lines. The total time in each stage is normalized to unity in order to see the full history of the run at once, and not just distant ejections. The numbers on the plot are the actual times (in yr) spent in each stage.

Each run begins with a short period of secular evolution (illustrating the remarkable stability of hierarchical triples even slightly within the Mardling–Aarseth boundary). Dynamical friction brings the intruder in a bit further to get chaotic interactions underway. This can be seen more clearly in the upper right-hand panel, where we zoom in on the first close encounter at left. In this panel we also colour code the lines according to which two BHs instantaneously form the closest pair, to show the numerous exchanges that occur during close encounters. Large-amplitude Kozai oscillations are present in the first encounter of the lower run, but no oscillations are seen in the upper run, where the initial inclination is below the critical angle.

After the first encounter, in both runs the outer components suffer a few 'near' (∼0.1–1 kpc) ejections before they get shot out to kpc scales and come back by dynamical friction. In the upper run, the single goes out to ∼10 kpc, then comes back and forms a bound pair with the former binary, which has coalesced in the meantime. The new binary is highly eccentric ($e \sim 0.9998$) and quickly coalesces by gravitational radiation. In the lower run the single returns after the first kpc-scale ejection, strongly interacts with the binary one more time and then escapes the galaxy. The binary has a semimajor axis of 0.15 pc at the beginning of the final encounter, and its binding energy increases by 17 per cent in the interaction, imparting a velocity of $\sim$1400 km s^{-1} to the escaper.

In the lower right-hand two panels we focus on the third encounter of this run, which was selected because a significant amount of energy was lost to gravitational radiation over its duration. We can see that the radiation loss occurred during two very close approaches, by two different BH pairs. If close three-body encounters between BHs are sufficiently common in the Universe, such gravitational radiation spikes could be detectable with LISA.

Fig. 11 shows the distribution of post-encounter velocities, for the single and recoiling binary. Included in the plot are all close encounters in which (a) the binary and single are unbound at the end of the encounter; (b) the binding energy of the binary increases by at least 5 per cent (to avoid numerous 'glancing' encounters where δF just barely exceeds the close encounter threshold) and (c)

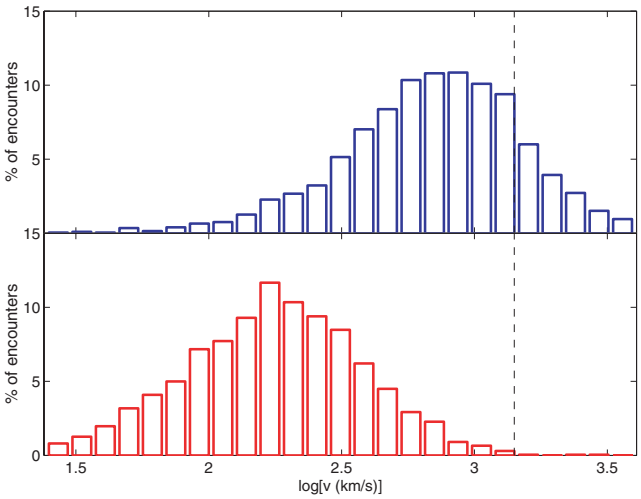

Figure 11. Distribution of ejection velocities. Upper panel: single BH. Lower panel: binary COM. The total BH mass is typically $\sim 6 \times 10^8$ M$_\odot$.

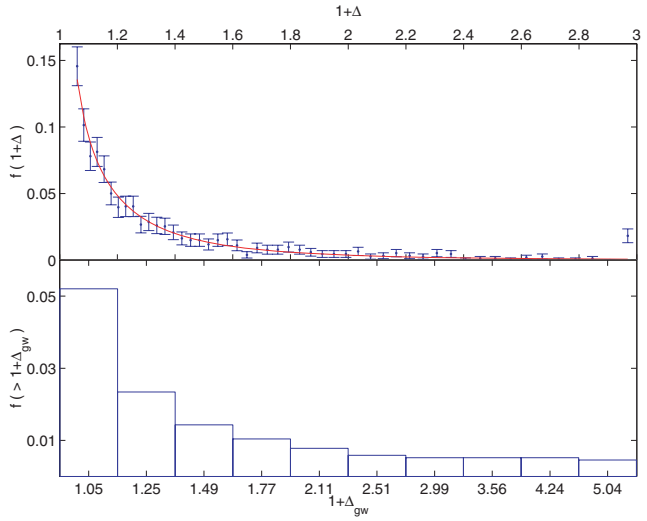

Figure 12. Energy exchanges during close encounters. Upper panel: distribution of the fractional change in the binding energy of the binary, $1+\Delta = 1+(\mathrm{BE_f}-\mathrm{BE_0})/\mathrm{BE_0}$, in close encounters. Blue points with 1.5σ Poisson error bars are the simulation data; the red line shows a comparison with the shape predicted by theory (see text). Lower panel: fraction of encounters with the relative energy radiated as gravitational waves in the encounter greater than $1+\Delta_{\rm gw} = 1+E_{\rm gw}/\mathrm{BE_0}$. Both panels include only encounters that end with the single BH unbound from the binary, and exclude the first encounter of each run, which starts in a special hierarchical triple configuration and also includes some dissipation by the dynamical friction used to bring in the intruder.

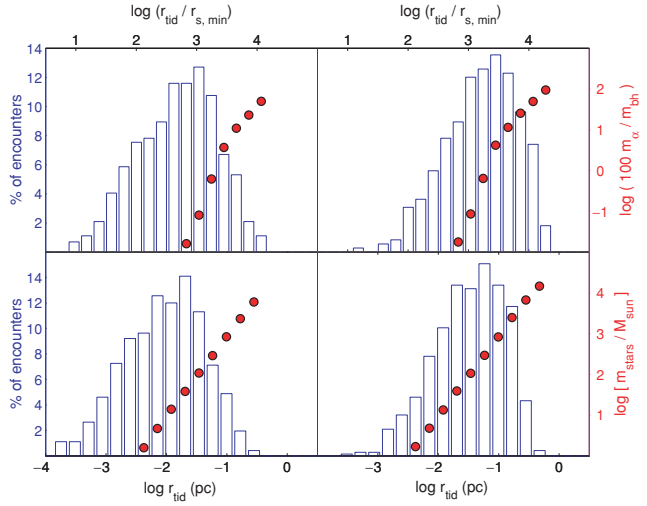

Figure 13. Tidal radius $r_{\rm tid}$ of the lighter pair member at closest approach. In this plot we include only encounters where one BH escaped at a speed above 940 km s^{-1}. Upper panels: $r_{\rm tid}$ in units of the Schwarzchild radius of the lighter pair member, $r_{\rm s,min}$. For reference, the red circles show the per cent of the BH's mass contained within $r_{\rm tid}$ in an α disc if the BH is accreting at the Eddington rate $\dot{m}_{\rm Edd}$, averaged over the encounters in each bin. Lower panels: $r_{\rm tid}$ in pc. Red circles show the mass in stars (in $\mathrm{M}_\odot$) contained within $r_{\rm tid}$ in the Hernquist profile used to model the stellar component of the galaxy. Left panels: closest pair. Right panels: second closest pair.

the encounter ends by δF falling below threshold (and not e.g. by coalescence of the binary or timing out). The dashed vertical line indicates the typical galactic (stellar bulge + halo) escape velocity, $v_{\rm esc} \sim 1400\,\mathrm{km\,s^{-1}}$. We see that the single will sometimes escape the galaxy (or be ejected far out into the halo where the dynamical friction return time exceeds a Hubble time), but the binary will rarely go far.

The upper panel of Fig. 12 shows the distribution of fractional changes in the binding energy of the binary during close encounters, $1+\Delta = 1+(\mathrm{BE_f}-\mathrm{BE}_0)/\mathrm{BE}_0$. The first encounter of each run is excluded from this plot, since it begins in a special hierarchical triple configuration and includes some dissipation by the dynamical friction used to bring in the intruder. The red line shows the best fit to the form $f(1+\Delta) = K\Delta^{-1/2}\,(1+\Delta)^{-9/2}$, with the normalization K depending on the mass ratios and intruder velocity, predicted by theory (Heggie 1975; Valtonen & Karttunen 2006). The lower panel shows the fraction of encounters with the relative energy radiated as gravitational waves in the encounter greater than $1+\Delta_{\rm gw} = 1+E_{\rm gw}/\mathrm{BE}_0$. This shows that gravitational radiation plays a significant role in the dynamics in only a few per cent of the encounters ending in the escape of one component. Another ∼20 per cent of the encounters end in coalescence; gravitational radiation of course plays a significant role in all of these.

Another point of interest is the statistics of the closest approach distances between two-body pairs during the encounters. Besides their intrinsic significance, the distances of closest approach are related to the extent of tidal stripping of the BHs during the encounters. One can imagine that if some stars, or even the inner portion of an accretion disc, remained bound to the individual BHs at the end of an encounter, then some ejected SMBHs might become observable.

Since BS time-steps are not at all infinitesimal (see Section 2.4), we cannot simply take the minimum over the discrete time-steps to be the closest approach distance. When the relative perturbation δF from the third body is small, one can obtain the periapsis distance analytically in the Keplerian two-body approximation. When δF is larger, the minimum over the time-steps should give a better estimate since the time-steps tend to be smaller, but this statement is difficult to quantify. To construct the distance of closest approach in our simulations, we first identify any step where $\mathrm{d}|\boldsymbol{r}|/\mathrm{d}t = \hat{\boldsymbol{r}}\cdot\boldsymbol{v}$ switches sign from negative to positive and $|\boldsymbol{r}| < 30000(r_{\rm s1}+r_{\rm s2})$ for any pair as a 'passing step' containing a close approach. Here $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$, $\boldsymbol{v} = \boldsymbol{v}_1 - \boldsymbol{v}_2$ and $r_{\rm s1,2}$ are the Schwarzchild radii of the two-pair members. We then iteratively bisect the time-step, evaluating $\hat{\boldsymbol{r}}\cdot\boldsymbol{v}$ at each bisection to find the place where it switches sign until the distance between the two bodies converges to within a part in 10^6.

Fig. 13 shows the distribution of the tidal radius $r_{\rm tid}$ at closest approach, for the closest and second closest pair. Since we are interested in observing ejected SMBHs, we only include encounters where the single escaped with a velocity above 940 km s^{-1}, the typical velocity needed to reach the stellar scale radius of ∼3 kpc in our galactic model. $r_{\rm tid}$ is defined by the equation

$$\delta a_{\rm tid} \equiv \frac{Gm_2}{(d-r_{\rm tid})^2} - \frac{Gm_2}{d^2} = \frac{Gm_1}{r_{\rm tid}^2}, \tag{39}$$

where m_1 is the reference mass being stripped (the smaller pair member), m_2 is the other point mass and d is the distance between m_1 and m_2. We solve this polynomial equation for $r_{\rm tid}$ exactly rather than Taylor expanding about $r_{\rm tid}/d = 0$ to get the familiar expression $r = (m_1/2m_2)^{1/3}$ for the tidal radius, since $r_{\rm tid}/d$ is not generally small at closest approach for the near-equal mass problem at hand. The upper panels show $r_{\rm tid}$ in units of the Schwarzchild radius of the smaller BH. The red circles indicate the per cent of this BH's mass contained within $r_{\rm tid}$ in an α disc (Shakura & Sunyaev 1973; Frank, King & Raine 2002; Narayan 2003) accreting at the Eddington rate $\dot{m}_{\rm Edd}$, assuming $\alpha = 0.1$ and a radiative efficiency of $\epsilon = 0.1$. The lower panels give $r_{\rm tid}$ in pc, and here the red circles show the mass

in stars within $r_{\rm tid}$ in the Hernquist model representing the stellar bulge.

The α-disc model assumes that the disc is not self-gravitating and breaks down as $\dot{m} \rightarrow \dot{m}_{\rm Edd}$, so the red circles in the upper panels are merely to give the reader an idea of the bound mass scales associated with the approaches. The tidal approximation is a pessimistic estimate of the extent of the stripping since swift, one-time close passages would be impulsive (Binney & Tremaine 1987, chapter 7). We record only the single closest approach, so we cannot distinguish between such impulsive events and approaches that are part of periodic patterns in the trajectories.

In a significant fraction of cases $r_{\rm tid} \gtrsim 10^4\, r_{\rm s}$ encloses a substantial fraction of the BH's mass in accreting gas, so near-Eddington accretion could continue for a duration of order the Salpeter (1964) time after the slingshot ejection (Hoffman & Loeb 2006). The enclosed stellar mass shown in the lower panels is never nearly comparable to the BH mass, but in most cases the escaper would drag some stars. In principle one can imagine one of these stars entering a giant phase and overflowing its Roche lobe, producing detectable accretion on to the SMBH long after its ejection from the galactic centre (e.g. Hopman, Portegies Zwart & Alexander 2004; Kuranov et al. 2007).

3.4 Distant evolution and binary re-formation

Slingshot ejections in triple encounters produce a population of 'wandering' SMBHs in the haloes of galaxies and intergalactic space (Volonteri et al. 2003a; Volonteri & Perna 2005). Fig. 14 shows the total time spent by the single BH (upper panel) and binary/remnant (lower panel) at various distances from the galactic centre, averaged over all 1005 runs. The time spent in each distance bin is summed over the duration of each run, and if the system reaches a steady state at time $t_{\rm end} < t_0 - t_{\rm f0}$ then the state of the system after the run (until $z = 0$) is included. If a component escapes, then the time $t_0 - t_{\rm end}$ is added to the highest bin; if it settles to the centre then this time is added to the lowest bin. The single is found wandering at large distances nearly half the time, while the binary spends the vast majority of its time at the galactic centre. Over all of our runs, the total fraction of the time spent with no SMBHs within 50 pc of

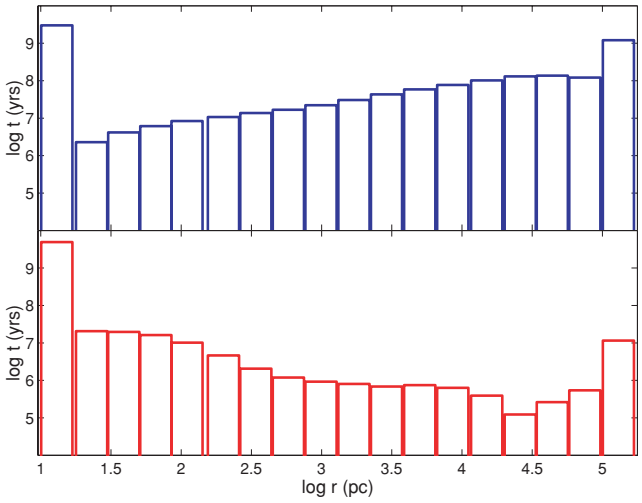

Figure 14. Total time spent by the single BH (upper panel) and binary/remnant (lower panel) at various distances from the galactic centre, averaged over all 1005 runs. The lowest bin includes all distances below 17 pc and the highest bin includes all distances above 10^5 pc.

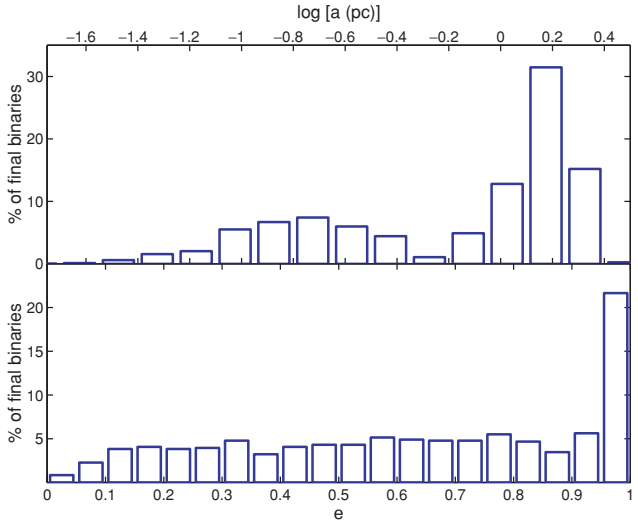

Figure 15. Distribution of parameters of 'final state' binaries. The plot includes binaries that settled to the centre after a single escape, and new binaries that formed from the single BH and inner binary remnant. Upper panel: semimajor axis. Lower panel: eccentricity.

the centre since the formation of the halo hosting the triple system is only ∼1 per cent. Hence we expect the ejections in triple BH encounters at low redshift to produce very few nuclei empty of SMBHs. A cD galaxy cluster, having hosted several dry mergers, might contain up to a few naked SMBHs wandering through the cluster halo as a result of single ejections.

The escaper remains wandering through the halo in only ∼40 per cent of the runs. In the other cases dynamical friction brings it back to the centre, where it becomes bound to the binary remnant and forms a new, hard binary. Fig. 15 shows the semimajor axis (upper panel) and eccentricity (lower panel) distributions of the 'final state' binaries in our simulations. This plot includes binaries formed when a pair coalesces during a close encounter and is replaced by a single BH with its COM coordinates; cases where the original binary never coalesces, but rather settles to the centre and stalls after the single escapes and binaries that form from the third BH and coalesced remnant after ejections.

Whereas in the absence of triple encounters we would expect most SMBHs to sit around $a_{\rm hard}$, the encounters introduce a second population of stalled binaries at smaller separations. The eccentricities of the final binaries span the whole range from 0 to 1. Note the peak at very high eccentricity, arising mostly from cases where the escaper rejoins the binary remnant from a radial orbit following a distant ejection, as in the run shown in the upper panel of Fig. 10. Many of the binaries in this peak are expected to coalesce quickly by gravitational radiation. This result has importance for LISA if three-body ejections are common enough, since the gravitational radiation signature of a highly eccentric binary is quite different from that of a circular binary. An eccentric binary radiates at all integer harmonics of the orbital frequency, so its spectral energy distribution peaks at higher frequencies, possibly enabling the detection of higher mass SMBH binaries (e.g. Pierro et al. 2001; Enoki & Nagashima 2006).

However, we caution the reader that the high-eccentricity coalescence rate appears to be sensitive to the dynamical friction and core updating prescriptions. SMBH binary eccentricity evolution is a delicate question to which simulators have obtained widely discrepant answers (e.g. Milosavljevic & Merritt 2001; Aarseth 2003; Merritt & Milosavljević 2005). While the conclusion that a

high-eccentricity population forms through distant ejections is robust, the question of whether these systems tend to coalesce or remain as an observable population of stalled high-eccentricity binaries may depend on the details.

3.5 Core creation

A number of studies have addressed the mark left on a stellar core by the hardening of one or more BH binaries in independent succession (Milosavljevic & Merritt 2001; Ravindranath, Ho & Filippenko 2002; Volonteri, Madau & Haardt 2003b; Merritt 2006). To estimate the damage, consider a succession of mergers $(M_0, M_0') \rightarrow M_1, (M_1, M_1') \rightarrow M_2, \ldots, (M_{N-1}, M_{N-1}') \rightarrow M_N$, between galaxies containing BHs of mass $(m_0, m_0'), \ldots, (m_{N-1}, m_{N-1}')$, and having insufficient gas for significant stellar cusp regeneration. Suppose that following each merger the BHs spiral in to $a_{\rm hard}$ by dynamical friction on the stars, then cross the gap from $a_{\rm hard}$ to $a_{\rm gw}$ by some non-stellar mechanism, e.g. interaction with a modest amount of gas that ends up in the nucleus through tidal torques associated with the merger. The total energy deposited in the stellar core is roughly

$$E_{\rm dep} \sim \sum_{i=0}^{N-1} \left(\frac{G m_i m_i'}{a_{{\rm hard},i}} - \frac{G m_i m_i'}{a_{{\rm inf},i}} \right), \tag{40}$$

where $a_{{\rm hard},i}$ is the hardening radius of the BH binary formed in each merger, and the radius of influence $a_{{\rm inf},i}$ is the radius containing about twice the mass of the larger binary member in stars (e.g. Merritt 2006). Note that the right-hand side of (40) is dominated by the first term in the parentheses, so the precise definition of $a_{{\rm inf},i}$ is not important. If the inner density profile of a galaxy flattens from $d\ln\rho/d\ln r = 3-\eta$ to a shallower slope γ within some core radius $r_{\rm b}$, then we can define a core 'energy deficit' by

$$U_{\rm def} = \pi \int_0^\infty \left[\rho(r)\Phi(r) r^2 - \rho_\eta(r)\Phi_\eta(r) r^2 \right] {\rm d}r, \tag{41}$$

the difference between the binding energy of the galaxy with the density break and that of the same galaxy, but with the density profile outside the core extrapolated inward to the centre. In this equation $\rho = \rho_{\rm stars} + \rho_{\rm halo}$ and $\Phi = \Phi_{\rm stars} + \Phi_{\rm halo}$ denote the sums of the contributions to the density and gravitational potential from the stellar and dark matter halo components. The cross terms $\rho_{\rm stars}\Phi_{\rm halo}$ and $\rho_{\rm halo}\Phi_{\rm stars}$ contribute about 10–20 per cent of the total binding energy, while the halo–halo term is negligible. We denote the outer slope by $3-\eta$ to match the Tremaine et al. (1994) parametrization used in our galactic models. We can estimate the size of the core created by the cumulative scouring action of the BH binaries formed in the succession of mergers by equating $U_{\rm def}$ of equation (41) to $E_{\rm dep}$ given by equation (40).

The $U_{\rm def} = E_{\rm dep}$ prescription was introduced in order to estimate the extent of cusp destruction *before the binary hardens*. If stalling were prevented by sufficient scattering of stars into the loss cone, an analogous energy argument would grossly overestimate the size of the core scoured out as the binary decayed from $a_{\rm hard}$ to the separation where gravitational radiation could take over. This is because a hard binary loses energy by ejecting stars at high velocities, often exceeding the escape velocity of the entire galaxy. Most of the energy released by the binary goes into excess kinetic energy of these hypervelocity stars rather than heating of the local medium. Equations (40) and (41) capture the essence of the core scouring in the limit of weak encounters (dynamical friction), but for hard binaries we must view the cusp destruction as *mass ejection* rather than *energy injection*, once again following the work of Mikkola & Valtonen (1992) and Quinlan (1996). A hard SMBH binary is defined by the fact that it hardens at a constant rate, ${\rm d}E/{\rm d}t = {\rm const}$. In the limit of very high orbital velocity ($w \gg \sigma$) this implies that a constant mass in stars, comparable to the total BH mass, is ejected from the galactic centre per e-folding of the binary semimajor axis,

$$\frac{1}{m_{\rm bin}} \frac{{\rm d}M_{\rm ej}}{{\rm d}\ln(1/a)} = \frac{1}{m_{\rm bin}} \frac{{\rm d}M_{\rm ej}}{{\rm d}\ln(E_{\rm st})} \equiv J \approx 0.5, \tag{42}$$

where $E_{\rm st}$ is the energy transferred from the BH system to the stars. We can estimate the core damage due to mass ejection by equating the total mass ejected by the binary to $M_{\rm def}$ as defined in equation (38).

Now suppose that instead of coalescing without further damaging the stellar core, the binary formed in the first merger in our sequence stalls at $a_{\rm hard}$ until a third BH sinks in following the second merger. On the one hand, some energy that would have been injected into the stars as the outer binary hardened may now instead be carried off as gravitational radiation or kinetic energy of a fast escaping BH, causing less damage to the stellar core than the decay of two separate binaries. On the other hand, the intruder may continue scattering stars into the loss cone well after the inner binary reaches $a_{\rm hard}$, and ejected BHs heat the core by dynamical friction as their orbits pass repeatedly through the dense nucleus (Boylan-Kolchin, Ma & Quataert 2004).

To quantify these considerations our code evolves the core radius $r_{\rm b}$ and slope γ along with the BH orbits to roughly account for the core heating and mass ejection caused by the triple systems. At the beginning of each run we initialize the core by injecting an energy

$$E_{\rm init} = \frac{G m_1 m_2}{a_{\rm hard,i}} + \frac{G(m_1+m_2)m_3}{a_{\rm init,o}} - \frac{G m_1 m_2}{a_{\rm inf,i}} - \frac{G(m_1+m_2)m_3}{a_{\rm inf,o}} \tag{43}$$

into the parent η model according to equation (41). In runs where the inner binary starts at $a > a_{\rm hard}$ we replace $a_{\rm hard,i}$ with $a_{\rm init,i}$ in equation (43). In runs where it starts at $a < a_{\rm hard}$ we also eject a mass $0.5 m_{\rm bin} \ln(a_{\rm hard,i}/a_{\rm init,i})$ according to equation (38) before the start of the run.

There is an ambiguity in the way we update the profile since energy may be injected (or mass may be ejected) either by increasing $r_{\rm b}$ to make the core larger, or by decreasing γ to make it shallower. We resolved this ambiguity by performing a rough fit to the γ versus $y \equiv M_{\rm def}/m_{\rm bh}$ data in Merritt (2006), to obtain the relation

$$\gamma \approx -0.0281 y^3 + 0.2451 y^2 - 0.7094 y + 1.000 \tag{44}$$

for $\eta = 2$, which gives sensible slopes for all $y \lesssim 5$. This relation at least has the desired property that $\gamma \rightarrow 3-\eta$ as $M_{\rm def} \rightarrow 0$, but the slope becomes quite shallow toward large $M_{\rm def}$. The mass deficits are not sensitive to our prescription for γ.

During the unperturbed binary integration, we increment the energy injected at each time-step $t \rightarrow t + \Delta t$ by

$$\Delta E_{\rm inj} = \sum_{i=1}^{2} \int_t^{t+\Delta t} (\boldsymbol{F}_{{\rm df},i} \cdot \boldsymbol{v}_i) \theta(r_{\rm core} - r_i)\, {\rm d}t, \tag{45}$$

the work done on the single BH and binary COM by dynamical friction while the respective masses are located within a distance $r_{\rm core} \equiv 1.5 r_{\rm b}$ of the galactic centre. When the binary is located within $r_{\rm core}$ and has not yet coalesced or settled to the centre and stalled, we also increment the ejected mass by

$$\Delta M_{\rm ej} = 0.5 m_{\rm bin} \ln\left(\frac{E_0 + \Delta E_{\rm st}}{E_0} \right) \tag{46}$$

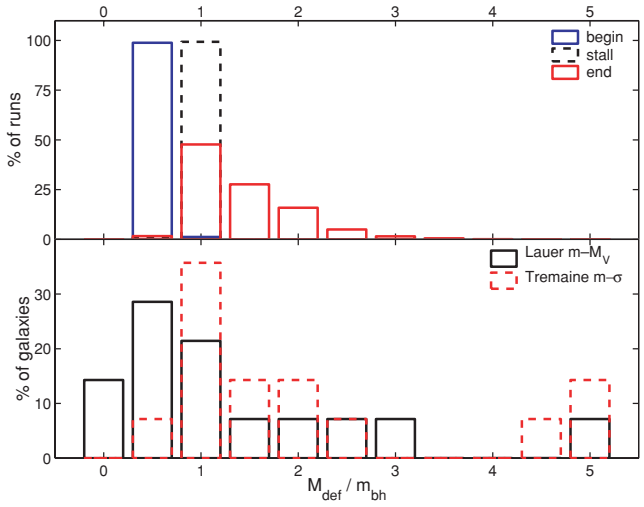

Figure 16. Comparison of calculated and observed mass deficits. Upper panel: calculated mass deficits in units of the total BH mass, $m_1 + m_2 + m_3$. Blue (left): beginning of simulation, based on core heating during inspiral to the initial configuration. Black dashed (middle): net effect of two dry mergers in which core scouring stops at $a_{\rm hard}$. Red (right): end of simulation, reflecting net energy injection and mass ejection caused by close triple encounters. Lower panel: observed mass deficits in units of the SMBH mass (Graham 2004; Ferrarese et al. 2006). In the red dashed histogram we determine the BH masses from the m–σ relation of Tremaine et al. (2002) for all entries. In the black solid histogram we instead use the m–M_V relation of Lauer et al. (2006) for those galaxies with $M_V < -22$, and the dynamically measured BH masses in the four cases where they are available.

at each time-step, where E_0 is the binding energy of the binary at the beginning of the time-step and $\Delta E_{\rm st}$ is the change in binding energy due to stellar hardening during that step. We cannot simply use the semimajor axis increment in equation (46) since the binary may also have hardened by gravitational radiation during the time-step. At the beginning of each run, we also include the energy injected by dynamical friction as the intruder spirals in before the onset of chaotic interactions.

Each time the total energy injected reaches 1 per cent of $E_{\rm hard}$ or the total mass injected reached 1 per cent of the total BH mass we update the core accordingly. At the end of each run we record the final $M_{\rm def}$, $r_{\rm b}$ and γ. Though both mass ejection and energy injection enter into our core growth algorithm, at the end both translate into a single effective $M_{\rm def}$ as given by equation (38), which we record for comparison with observed galaxies.

Fig. 16 compares our calculated mass deficits to 14 cored, luminous elliptical galaxies with BH masses ranging from $\sim 10^8$ to $3 \times 10^9\,{\rm M}_\odot$, with measured mass deficits. The upper panel shows data from our simulations. The blue (left) histogram is the distribution at the beginning of the runs, reflecting the heating of the core by dynamical friction on the BHs as they sink into the initial configuration with the inner binary at $a_{\rm hard}$, equation (43). This distribution also approximately represents the core damage expected for a single merger in which stellar hardening ceases at $a_{\rm hard}$. Note that a significant core ($M_{\rm def}/m_{\rm bh} \sim 0.5$) is scoured out even before the binary hardens and begins ejecting stars. The middle (black dashed) histogram shows the core predicted for a series of two dry mergers, in both of which stellar hardening stops at $a_{\rm hard}$. The red (right) histogram is the distribution of cores at the end of our runs, reflecting the core scouring effect of the three-body interactions.

Table 2. Mass deficits in galaxies with dynamically measured SMBH masses.

Galaxy	$m_{\rm bh}/{\rm M}_\odot$	M_V	$M_{\rm def}/m_{\rm bh}$	Reference
NGC 4291	3.1×10^8	-20.64	1.8	G03
NGC 4374	1.6×10^9	-22.28	1.4	B98
NGC 4486	3.0×10^9	-22.71	2.9	M97, H94
NGC 4649	2.0×10^9	-22.51	1.1	G03

References: G03 – Gebhardt et al. 2003; B98 – Bower et al. 1998; M97 – Macchetto et al. 1997; H94 – Harms et al. 1994.

The mass deficits in the lower panel are obtained in Graham (2004) and Ferrarese et al. (2006) by fitting the outer nuclear density profile to a Sérsic law (Sérsic 1968), and then subtracting a power-law fit to the inner core from the inward extrapolation of the Sérsic profile. The red dashed histogram in the lower panel shows the observed $M_{\rm def}/m_{\rm bh}$ with $m_{\rm bh}$ computed from the $m_{\rm bh}$–σ relation of Tremaine et al. (2002). However, Lauer et al. (2006) argue that luminosity may be a better predictor of BH mass than σ for the most luminous elliptical galaxies ($M_V \lesssim -22$), since their recent merger histories consisted mostly of passive (dissipationless) mergers in which both the BH mass and luminosity are simply additive. The $m_{\rm bh}$–σ relation is thought to arise from self-regulation of accretion on to the SMBH in gaseous mergers (Silk & Rees 1998; Wyithe & Loeb 2003b), which does not apply in this context. Lauer et al. (2006) also show that an extrapolation of the $m_{\rm bh}$–L relation to the highest luminosities is more consistent with the observed $r_{\rm core}$–$m_{\rm bh}$ relation and provides a better match between the $z = 0$ SMBH space density and the quasar population seen at $z \sim 2$ for reasonable quasar duty cycles. In the black histogram, we used the observed BH masses for the four cases with dynamical mass measurements, listed in Table 2 (Harms et al. 1994; Macchetto et al. 1997; Bower et al. 1998; Gebhardt et al. 2003). For the rest of the galaxies we used the Lauer et al. (2006) $m_{\rm bh}$–M_V relation to estimate the BH masses in galaxies with $M_V < -22$, and the Tremaine et al. (2002) $m_{\rm bh}$–σ relation for those with $M_V > -22$. Showing both plots gives an idea of how much the mass deficits vary with $m_{\rm bh}$ estimator.

In set CN ~11 per cent of the runs resulted in cores with $M_{\rm def}/m_{\rm bh} > 2$. In some rare runs where the binary was ejected to a large distance and then brought back by dynamical friction, we obtained even higher mass deficits (up to $M_{\rm def}/m_{\rm bh} = 3$–4). The binary is more efficient at core scouring than the single BH since it is more massive. If each independent binary inspiral adds $\sim 0.5 m_{\rm bh}$ to the mass deficit, then the mean $M_{\rm def}$ enhancement of $\sim 0.5 m_{\rm bh}$ due to the triple encounters is equivalent to one extra merger in the system's history. An $M_{\rm def}$ one standard deviation above the mean is equivalent to two extra mergers.

4 DISCUSSION AND CONCLUSIONS

Triple SMBH systems in galactic nuclei produce a range of phenomena and signatures rather different from those expected if no more than two SMBHs occupy them at a time. We have developed an efficient numerical method for following the evolution of three-body systems in the centres of galaxies, and used it to explore the outcomes of such encounters in massive elliptical galaxies at low redshift.

We find a high efficiency of SMBH coalescence due to the encounters, providing a 'last resort' solution to the final parsec problem. There is, however, a caveat in extending this result immediately to all BH masses. If we define $a_{\rm esc}$ to be the binary semimajor axis

where escape of one BH first becomes likely ($Gm_{\rm bin}/\beta\ a_{\rm esc} = v_{\rm esc}^2$ where β is a factor of order 10 for a Hernquist profile), then since $v_{\rm esc} \propto \sigma$ we have $a_{\rm esc} \propto m_{\rm bin}/\sigma^2 \propto m_{\rm bin}^{1/2}$ if $m_{\rm bin}$ obeys the $m_{\rm bh}$–σ relation $m_{\rm bin} \propto \sigma^4$, so $a_{\rm esc}/a_{\rm gw} \propto m_{\rm bin}^{-1/4}$. In other words, at smaller BH masses, the lightest BH is more likely to escape the galaxy before driving the binary to coalescence by gravitational radiation. By focusing on massive galaxies we have chosen the systems where the binary is *least* likely to coalesce by other means (e.g. gas or MPs), and *most* likely to coalesce in the next merger with the help of three BH interactions. We may address the efficiency of triple-induced coalescence in much smaller mass systems in future studies.

We find that close triple encounters can produce a population of high-eccentricity binaries, whose gravitational radiation signal could potentially be observable by LISA. Such signals originate from Kozai oscillations in hierarchical triples at high initial inclinations and highly eccentric binaries formed following distant ejections. As the eccentricity increases, the radiation spectrum peaks at progressively higher harmonics of the fundamental frequency, approaching a nearly flat spectrum as $e \rightarrow 1$ (Pierro et al. 2001; Enoki & Nagashima 2006). A circular 10^{8-9} $M_\odot$ BH binary remains below the band of frequencies ($\sim 10^{-4}$ to 10^{-1} Hz) detectable by LISA throughout its inspiral, but the occurrence of high-eccentricity coalescences could extend LISA's sensitivity into this mass range, or lengthen the duration of its sensitivity to $\sim 10^{6-7}$ $M_\odot$ events. A highly eccentric binary produces a 'spiky' waveform that looks quite different from that of a circular system (see fig. 7 in Pierro et al. 2001). Gravitational radiation 'spikes' at very close approaches during chaotic three-body interactions could also produce radiation bursts detectable by LISA.

If triple encounters are indeed limited to massive systems at low redshift, then the importance of these considerations is limited by the expected event rate in this mass range, assuming efficient coalescence. This rate is highly uncertain, ranging from ~ 1/yr (Sesana et al. 2005) to $\sim 1/1000$ yr (Rhook & Wyithe 2005) depending on the merger and BH population model adopted. If three BH systems occur in other contexts, e.g. IMBHs in galactic nuclei or star clusters, then the phenomena we have discussed may be observationally relevant even if the high-mass SMBH event rate is low. A detailed look at the gravitational waveforms expected from three-body encounters and their expected detection rates is an interesting topic for a future study.

The slingshot ejections in triple encounters produce a population of 'wandering' SMBHs in and outside the haloes of galaxies. In systems that have undergone several major dry mergers (e.g. cD galaxy systems), one might expect a few such ejected SMBHs to be floating in the vicinity. As of yet, no probable way of observing these wandering BHs has been proposed.[1] In principle one can imagine a star bound to the ejected SMBH entering a giant phase and overflowing its Roche lobe, producing some accretion on to the SMBH and an observable flare. Single ejections could also in principle affect BH–bulge correlations such as the $m_{\rm bh}$–σ relation, but since it is the lightest BH that gets ejected this effect would fall well within the observed scatter in the correlations for just one or two ejection events.

Triple interactions in galactic nuclei can have a large effect on the expected properties of stable SMBH binaries in the local Universe. While many models of binary formation predict mostly circular binaries around $a_{\rm hard}$, three-body encounters produce binaries at all eccentricities. They also create a population of stalled binaries at separations significantly smaller than $a_{\rm hard}$ but still larger than $a_{\rm gw}$, as does any partial gap-crossing mechanism.

Better measurements and statistics on the mass deficits in cored elliptical galaxies may provide clues on the history of the nuclear SMBH activity in these systems. Triple BH encounters produce a highly scattered distribution of core sizes, with mass deficits up to $\sim 2\times$ higher than expected for successive binary coalescences. The apparent peak at mass deficits of $\sim$0.5–1 times the nuclear BH mass in observed cores may very tentatively hint that multiple-BH encounters are not the norm in these systems. This signature of binary or multiple-BH activity is appealing because (a) its duty cycle is the lifetime of the galaxy; (b) it is present whether binary pairs stall or coalesce and (c) it can be observed even in the complete absence of radiative activity, such as disc accretion or jet production. However, the interpretation of galaxy cores is complicated by multiple mergers, the possibility of partial stellar cusp regeneration from traces of cold gas, and observational complications such as projection effects in non-spherical galaxies and optimizing the fitting/extrapolation algorithm to best represent the mass deficit. There is a need for theoretical studies on the cores produced by SMBH mergers in triaxial galaxies, since triaxility seems to be the most likely candidate for a gap-crossing mechanism in dry mergers between gas-poor, giant ellipticals. Inferring the nuclear histories of galaxies from their observed core properties will likely be a topic of much interest in the future.

[1] Gravitational lensing is difficult to search for without knowing in advance the location of the BHs.

ACKNOWLEDGMENTS

We would like to thank Suvendra Dutta for technical help at the parallel computing centre of the Institute for Theory and Computation (ITC), and Sverre Aarseth and Seppo Mikkola for making their *N*-body algorithms and codes available. We are also grateful to Scott Hughes, Michael Eracleous, Marta Volonteri and Fred Rasio for useful discussions. This research was supported in part by an FQXi grant and Harvard University funds.

REFERENCES

Aarseth S. J., 2003, Ap&SS, 285, 367
Barkana R., Loeb A., 2001, Phys. Rep., 349, 125
Begelman M. C., Blandford R. D., Rees M. J., 1980, Nat, 287, 25
Berczik P., Merritt D., Spurzem R., Bischof H., 2006, ApJ, 642, 21
Binney J., Tremaine S., 1987, Galactic Dynamics. Princeton Univ. Press, Princeton, NJ
Blaes O., Lee M. H., Socrates A., 2002, ApJ, 578, 775
Blanchet L., Qusailah M. S. S., Will C. M., 2005, ApJ, 635, 508
Bower G. C. et al., 1998, ApJ, 492, L111
Boylan-Kolchin M., Ma C., Quataert E., 2004, ApJ, 613, 37
Bulirsch R., Stoer J., 1966, Numer. Math., 8, 1
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Byrd G. G., Sundelius B., Valtonen M., 1987, A&A, 171, 16
Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499
Centrella J. M., 2006, in Conf. Proc. Sixth Int. LISA Symp. Am. Inst. Phys., New York, in press (astro-ph/0609172)
Chandrasekhar S., 1943, ApJ, 97, 255
Cohn H., Kulsrud R. M., 1978, ApJ, 226, 1087
Colpi M., Mayer L., Governato F., 1999, ApJ, 525, 720
Damour T., Deruelle N., 1981, Phys. Lett. A, 87, 81
de Zeeuw P. T., Carollo C. M., 1996, MNRAS, 281, 1333
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Enoki N., Nagashima M., 2006, Prog. Theor. Phys., 117, 241
Erickcek A. L., Kamionkowski M., Benson A. J., 2006, MNRAS, 371, 1992

Escala A., Larson R. B., Coppi P. S., Mardones D., 2004, ApJ, 607, 765
Escala A., Larson R. B., Coppi P. S., Mardones D., 2005, ApJ, 630, 152
Favata M., Hughes S. A., Holz D. E., 2004, ApJ, 607, 5
Ferrarese L., 2002, ApJ, 578, 90
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Ferrarese L. et al., 2006, ApJS, 164, 334
Frank J., Rees M. J., 1976, MNRAS, 176, 633
Frank J., King A., Raine D., 2002, Accretion Power in Astrophysics, 3rd edn. Cambridge Univ. Press, Cambridge
Gebhardt K. et al., 2000, ApJ, 539, L13
Gebhardt K. et al., 2003, ApJ, 583, 92
Graham A. W., 2004, ApJ, 613, 33
Gultekin K., Miller M. C., Hamilton D. P., 2006, ApJ, 640, 156
Haiman Z., 2004, ApJ, 613, 36
Harms R. J. et al., 1994, ApJ, 435, L35
Heggie D. C., 1975, MNRAS, 173, 729
Heinämäki P., 2001, A&A, 371, 795
Hernquist L., 1989, Nat, 340, 687
Hernquist L., 1990, ApJ, 356, 359
Hills J. G., Fullerton L. W., 1980, AJ, 85, 1281
Hoffman L., Loeb A., 2006, ApJ, 638, L75
Holley-Bockelmann K., Mihos J. C., Sigurdson S., Hernquist L., Norman C., 2002, ApJ, 567, 817
Holman M., Touma J., Tremaine S., 1997, Nat, 386, 254
Hopman C., Portegies Zwart S. F., Alexander T., 2004, ApJ, 604, L101
Iwasawa M., Funato Y., Makino J., 2006, ApJ, 651, 1059
Kitayama T., Suto Y., 1996, ApJ, 469, 480
Komossa S., Burwitz V., Guenther H., Predehl P., Kaastra J. S., Ikebe Y., 2003, ApJ, 582, 15
Kormendy J., Gebhardt K., 2001, in Wheeler J. C., Martel H., eds, 20th Texas Symposium on Relativistic Astrophysics. Am. Inst. Phys., Melville, p. 363
Kozai Y., 1962, AJ, 67, 591
Kupi G., Amaro-Seoane P., Spurzem R., 2006, MNRAS, 371, 45
Kuranov A. G., Popov S. B., Postnov K. A., Volonteri M., Perna R., 2007, MNRAS, in press (astro-ph/0702525)
Kustaanheimo P., Stiefel E. L., 1965, J. Reine Angew. Math., 218, 204
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lauer T. R. et al., 2006, preprint (astro-ph/0606739)
Lee M. H., 1993, ApJ, 418, 147
Lightman A. P., Shapiro S. L., 1977, ApJ, 211, 244
Macchetto F., Marconi A., Axon D. J., Capetti A., Sparks W., Crane P., 1997, ApJ, 489, 579
Magorrian J. et al., 1998, AJ, 115, 2285
Manrique A., Salvador-Sole E., 1996, ApJ, 467, 504
Maoz D., Filippenko A. V., Ho L. C., Rix H.-W., Bahcall J. N., Schneider D. P., Macchetto F. D., 1995, ApJ, 440, 91
Maoz D., Nagar N. M., Falcke H., Wilson A. S., 2005, ApJ, 625, 699
Marconi A., Hunt L. K., 2003, AJ, 589, L21
Mardling R., Aarseth S., 2001, MNRAS, 321, 398
Merritt D., 2006, ApJ, 648, 976
Merritt D., Milosavljević M., 2005, Living Rev. Relativ., 8, 8
Merritt D., Poon M. Y., 2004, ApJ, 606, 788
Merritt D., Quinlan G. D., 1998, ApJ, 498, 625
Merritt D., Milosavljevic M., Favata M., Hughes S. A., Holz D. E., 2004, ApJ, 607, 9
Mikkola S., Aarseth S., 1990, Celest. Mech. Dyn. Astron., 47, 375
Mikkola S., Aarseth S., 1993, Celest. Mech. Dyn. Astron., 57, 439
Mikkola S., Valtonen M. J., 1990, ApJ, 348, 412
Mikkola S., Valtonen M. J., 1992, MNRAS, 259, 119
Milosavljevic M., Merritt D., 2001, ApJ, 563, 34
Narayan R., 2003, Theory of Thin Accretion Discs, course notes distributed in Astronomy 219, Spring 2003. Harvard University
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Peng C. Y., Impey C. D., Ho L. C., Barton E. J., Rix H., 2006, ApJ, 640, 114
Perets H. B., Hopman C., Alexander T., 2007, ApJ, 656, 709
Peters P. C., 1964, Phys. Rev., 136, 1224
Pierro V., Pinto I. M., Spallicci A. D., Laserra E., Recano F., 2001, MNRAS, 325, 358
Quinlan G. D., 1996, New Astron., 1, 35
Quinlan G. D., Hernquist L., 1997, New Astron., 2, 533
Raig A., Gonzalez-Casado G., Salvador-Sole E., 2001, MNRAS, 327, 939
Ravindranath S., Ho C., Filippenko A. V., 2002, ApJ, 566, 801
Rhook K. J., Wyithe J. S. B., 2005, MNRAS, 361, 1145
Rodriguez C., Taylor G. B., Zavala R. T., Peck A. B., Pollack L. K., Romani R. W., 2006, ApJ, 646, 49
Roos N., 1981, A&A, 104, 218
Salpeter E. E., 1964, ApJ, 140, 796
Salvador-Sole E., Solanes J. M., Manrique A., 1998, ApJ, 499, 542
Sanders D. B., Soifer B. T., Elias J. H., Madore B. F., Matthews K., Neugebauer G., Scoville N. Z., 1988, ApJ, 325, 74
Saslaw W. C., Valtonen M. J., Aarseth S. J., 1974, ApJ, 190, 253
Sérsic J. L., 1968, Cordoba. Observatorio Astronomico, Argentina
Sesana A., Haardt F., Madau P., Volonteri M., 2005, ApJ, 623, 23
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Silk J., Rees M. J., 1998, A&A, 331, L1
Spergel D. et al., 2006, ApJ, in press (astro-ph/0603449)
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Stiefel E. L., Scheifele G., 1971, Linear and Regular Celestial Mechanics. Springer-Verlag, New York
Taffoni G., Mayer L., Colpi M., Governato F., 2003, MNRAS, 341, 434
Tremaine S., Richstone D. O., Byun Y., Dressler A., Faber S. M., Grillmair C., Kormendy J., Lauer T. R., 1994, AJ, 107, 634
Tremaine S. et al., 2002, ApJ, 574, 740
Valtonen M. J., 1976, A&A, 46, 435
Valtonen M., Karttunen H., 2006, The Three-Body Problem. Cambridge Univ. Press, Cambridge
Valtonen M. J., Mikkola S., Heinamaki P., Valtonen H., 1994, ApJS, 95, 69
Volonteri M., Perna R., 2005, MNRAS, 358, 913
Volonteri M., Haardt F., Madau P., 2003a, ApJ, 582, 559
Volonteri M., Madau P., Haardt F., 2003b, ApJ, 593, 661
Wyithe J. S. B., Loeb A., 2003a, ApJ, 590, 691
Wyithe J. S. B., Loeb A., 2003b, ApJ, 595, 614
Wyithe J. S. B., Loeb A., 2005, ApJ, 634, 910
Yu Q., 2002, MNRAS, 331, 935

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